Università
Ca'Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di ricerca
in Scienze Ambientali, Informatica e Statistica
Ciclo XXVI
Anno di discussione 2014

The mapping and quantification of CO_2 leakage and its potential impact on groundwater quality

SETTORE SCIENTIFICO DISCIPLINARE DI AFFERENZA: GEO/08
Tesi di Dottorato di Stanley Eugene Beaubien, matricola 955906

Coordinatore del Dottorato

Prof. Gabriele Capodaglio

Tutore del Dottorando

Prof. Salvatore Lombardi

Co-tutore del Dottorando

Prof. Giancarlo Rampazzo

TABLE OF CONTENTS

DEDICATION iii

ACKNOWLEDGMENTS iv

ABSTRACT v

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. THEORETICAL BACKGROUND 5

2.1. CO_2 characteristics 5

2.2. Soil gas CO_2 origins 5
2.2.1. Biogenic CO_2 7
2.2.2. Shallow inorganic CO_2 10
2.2.3. Geogenic CO_2 10
2.2.4. Atmospheric and Anthropogenic CO_2 14

2.3. Gas migration 14
2.3.1. Migration mechanisms 14
2.3.1.1. Advection 15
2.3.1.2. Diffusion 16
2.3.1.3. Attenuation or migration via dissolution 18
2.3.2. Deep Pathways 19
2.3.2.1. Faults 19
2.3.2.2. Boreholes 23
2.3.3. Shallow pathways 25
2.3.3.1. Free-phase gas in sediments below the water table 26
2.3.3.2. Gas movement and distribution in the unsaturated zone 27
2.3.4. Release to the atmosphere 28

2.4. Surface leakage characterisation 29
2.4.1. Finding Leaks 30
2.4.2. Quantifying Leaks 33
2.4.2.1. Single survey of flux measurements 33
2.4.2.2. Multiple flux surveys 36
2.4.2.3. Monitoring of soil gas tracers 36
2.4.2.4. Combined flux and soil gas measurements 36

CHAPTER 3. LATERA CALDERA DESCRIPTION 38

3.1. Geology 39

3.2. Hydrogeology 41
3.2.1. Geothermal fluids 41
3.2.2. Springs and Groundwaters 41

3.3. Gas leakage at the ground surface 43
3.3.1. General overview 43

CHAPTER 4. THE LEAKAGE OF CO_2 AT LATERA 46

4.1. Introduction 46

4.2. Study site 47

4.3. Material and Methods **50**
4.3.1. Flux measurements 50
4.4. Latera flux results and discussion **50**
4.5. Summary **54**
CHAPTER 5. MODELLING OF SAMPLING DENSITY **56**
5.1. Introduction **56**
5.2. Program architecture **56**
5.2.1. Data Input 56
5.2.2. Algorithm description 58
5.2.2.1. Synthetic versus Real Data 58
5.2.2.2. Sub-sampling approaches 64
5.2.2.3. Data contouring and flux calculations using Surfer 9 66
5.2.2.4. Boundary effects 67
5.3. Synthetic data modelling **68**
5.3.1. Leak detection 68
5.3.2. Leak quantification 77
5.3.2.1. Scenario 1 – no leakage 78
5.3.2.2. Scenario 2 – single vent 81
5.3.2.3. Scenario 3 – three vents, simulations at three different distances 86
5.4. Latera data modelling **89**
5.5. Summary **93**
CHAPTER 6. IMPACT OF CO_2 ON GROUNDWATER QUALITY **96**
6.1. Introduction **96**
6.2. Study site **98**
6.3. Material and Methods **101**
6.3.1. Sampling 101
6.3.2. Analytical Methods 103
6.4. Latera groundwater results and discussion **104**
6.4.1. Borehole log and piezometric surface results 104
6.4.2. Data Statistics 105
6.4.3. Groundwater chemistry compared to regional data from the literature 107
6.4.4. Spatial distribution of groundwater chemistry 108
6.4.5. Spatial distribution of surface water chemistry 116
6.5. Summary **117**
CHAPTER 7. CONCLUSIONS **119**
CHAPTER 8. REFERENCES **122**
Appendix 1 **136**

Dedication

To my wife, Elena, who understood better than anyone my problems and difficulties, who cheered me up when I was stressed, and who gave me the space, the time, and the understanding which allowed me to make this happen.

To my mom, Joyce, whose unwavering enthusiasm and support, from the day I announced I would undertake the PhD to the day I submitted, has been a constant source of inspiration and motivation.

To my dad, Joffre, who I wish could be here to share this with me.

Acknowledgments

The untimely passing of Prof. Giovanni Maria Zuppi, my original thesis advisor, was a shock to all. His energy and enthusiasm for his research was an inspiration, and I will be forever indebted to him for accepting me as his student and giving me this opportunity.

I thank Prof. Giancarlo Rampazzo for kindly accepting me as his student within the department, and Prof. Bruno Pavoni for his very sincere support during his tenure as coordinator of the doctoral program.

I cannot express enough my gratitude to Prof. Salvatore Lombardi of the Department of Earth Sciences at the Università di Roma "La Sapienza", with whom I have worked for so many years. His constant drive and energy to carry his group forward and to open new avenues of research, ever more challenging in these difficult times, has never failed to amaze me. His very human approach to research has always focussed on treating all who work with him with respect and trying to use the different capabilities that each person has. It was he who convinced me, after many years, to attempt the PhD - something for which I have cursed him many times in the past, but for which I thank him sincerely now!

I would like to thank the many friends from Professor Lombardi's team who, over the years, have helped me in so many ways to make this journey, including Giancarlo Ciotoli, Anna Baccani, Aldo Annunziatellis, Stefano Graziani, Sabina Bigi, Samuela Vercelli, Chiara Tartarello, Livio Ruggiero, Pietro Sacco, Davide de Angelis, and Mariagrazia Finoia. Thanks especially to Giancarlo and Mariagrazia for their reading of, and comments on, parts of the thesis, and to Pietro and Davide for their help in the field.

I would also like to acknowledge the kind access given by Signor Iaccarelli to the field where the Latera soil flux grid measurements were performed.

Finally I would like to thank my family, for always being there for me.

Parts of Section 2.4 ("Surface Leakage Characterization") were originally written by the author for the report "Quantification Techniques for CO_2 Leakage" (IEAGHG, Report Number 2012/02, 2012); the kind permission granted by the IEAGHG to reproduce and/or summarize segments of this text is greatly appreciated. Chapter 6, on the impact of CO_2 leakage on groundwater quality at the Latera site, was written by the author as part of our research group's contribution to the European Community (EC) funded RISCS project.

Abstract

The present study relates to CO_2 leakage from the deep subsurface towards the atmosphere. Work involved writing a computer code to study the impact of different sampling strategies and densities on the capability of a gas flux survey to find and quantify the leakage of CO_2, with simulations being performed using both synthetic and real data. In addition, the potential impact of a CO_2 leak on drinking water quality was studied by measuring groundwater chemistry along a transect through a major natural CO_2 leakage area. The research detailed here relates to both natural leaking systems as well as to man-made systems where leakage may hypothetically occur (such as carbon capture and storage, CCS, projects).

A dense grid (10 m spacing) of 550 CO_2 flux measurements were made in a field within the Latera Caldera where numerous gas leakage points are known to occur. The obtained results are discussed looking at issues related to finding and quantifying leakage (e.g., sample density, background subtraction) and migration processes and pathways (e.g., faults, sediment control). In addition, another critical goal of this survey was to obtain a real-world dataset of spatially distributed CO_2 flux measurements on which sub-sampling simulations could be conducted using the software developed during this research, to compare these real-world results with those obtained from simulations using purely synthetic data.

A major focus of the present research was to develop a new computer code capable of examining the effects of different sampling strategies and sampling densities on the end results of gas flux surveys aimed at CCS monitoring or quantification of natural CO_2 emissions. The developed program creates a highly detailed (1 m node spacing) original dataset of synthetic data (or imports real CO_2 flux data from a grid), which is in turn randomly sub-sampled N times for each of M different sample densities using one of four different sample strategies. By performing multiple, but unique, simulations on a known input, the resulting output can be interpreted in terms of probabilities and statistics. Various calculations and monitoring of different parameters permits one to determine, under the simulation conditions and sampling settings, the number of samples needed to find a leakage point at a given confidence level and the precision and accuracy of estimated leakage flux values compared to the "true" original value. Various representative simulations are presented and the observed trends are discussed in terms of how they may influence and impact the results of actual surveys.

Finally, a groundwater study was performed to examine the potential impact of CO_2 leakage on water chemistry within the Latera Caldera. This site consists of deep carbonate units that are overlain by volcanic lithologies and surface alluvial and fluvial sediments formed from the same volcanic rocks. At this location gas (>95% CO_2) is leaking over a wide area, however there is no clear indication of co-migration of deep water, meaning that any observed changes in chemistry may be due to gas-induced in situ reactions only. A series of boreholes were augered by hand along a transect parallel to the groundwater flow direction, with sampling points located up gradient, within, and down gradient of the leakage area. Samples were analysed in the laboratory for major and trace elements, whilst some physical-chemical parameters, such as pH and temperature, were measured in the field. Results are discussed in terms of the spatial trends along the flow gradient, the potential chemical mechanisms that may be controlling the release or immobilization of various elements and compounds, and how local site conditions can control and influence potential impact.

Chapter 1. Introduction

The Earth is a dynamic system, with a constant exchange of mass between solid, liquid, and gaseous phases and a constant cycling of the fluid phases to and from the solid earth, the oceans, and the atmosphere. Research into these transformation and migration processes can lead to a greater understanding of how best to manage the complex geological environment in which we live and on which we depend.

Carbon dioxide (CO_2) gas is a critical component of the global carbon cycle in terms of both the abiotic and biotic systems (and the interaction between them). A very simplified overview of the abiotic system involves CO_2 in the atmosphere as a minor component (but having a critical role in regulating global climate), dissolution in the oceans, transformation and precipitation as carbonate mineral phases, sedimentation, subduction, transformation back to CO_2 via metamorphic and metasomatic processes, and eventual re-release to the atmosphere and hydrosphere via volcanoes and fault systems. A very simplified overview of the biotic system involves the fixation of CO_2 in plant material via photosynthesis, its incorporation in the shells, tests, or exoskeletons of aquatic organisms, and the production and release of gaseous CO_2 as a by-product of the respiration process. Clearly these two cycles have many points of overlap and interaction, and it is the balance and equilibrium amongst the numerous processes which control CO_2 concentrations within, and fluxes between, the various compartments. For the goals of the present work, research has focused on the leakage of deep, geologically produced CO_2 from the ground surface to the atmosphere and the potential impact that this leaking gas may have on groundwater quality.

Italy is geologically young and highly dynamic, formed by a series of tectonic events that caused both crustal thinning and orogenesis (Di Stefano et al., 2009; Devoti et al., 2011; Giacomuzzi et al., 2011). These in turn have formed areas of high heat flow, volcanism, subduction, and deep faulting, all processes that contribute to the production and migration of abiotic, geogenic CO_2. Central Italy is particularly interesting from this point of view, with extensive CO_2 release being associated with high geothermal gradients and predominantly extinct Quaternary volcanism along the western Tyrrhenian coast (Minissale 2004; Chiodini et al., 2004). Originating from mantle / magma degassing or metasomatism of subducted carbonate-rich crust, this gas migrates upwards along permeable pathways within faults and fracture networks and is released to the atmosphere from individual points known as gas vents or moffates (e.g. Annunziatellis et al., 2008) or over more diffuse areas via aquifer degassing.

An extensive literature exists on the origins, migration, and release of CO_2 within central Italy (Hooker et al., 1985; Minissale 2004; Chiodini et al., 2004; Frondini et al., 2008; Frezzotti et al., 2009) and other geologically similar areas throughout the world, with studies trying to map and quantify the total amounts of leaking CO_2 for various purposes. For example, there are health and safety concerns due to the fact that the denser than air CO_2 can accumulate in confined or depressed areas (like basements or cellars) and reach life-threatening concentrations, and thus mapping can provide a critical tool for local governments to define, and limit construction within, high risk areas (e.g. Beaubien et al., 2003). Other studies contribute to the need to quantify all natural and anthropogenic sources and sinks of atmospheric greenhouse gases like CO_2 to improve climate change predictive modelling efforts (e.g. Chiodini et al., 2004), or use CO_2 leakage rates (combined with chemical geothermometers) to estimate the potential in situ energy level that could be exploited from geothermal reservoirs (e.g. Chiodini et al., 2007).

Natural leaking sites have also been studied in light of human activities related to underground gas storage. In particular, Carbon Capture and Storage (CCS), is one issue that has greatly benefitted through the study of these natural laboratories. CCS involves the capture of anthropogenic CO_2 from point sources (such as coal-fired power plants) and its injection into deep saline aquifers or depleted hydrocarbon reservoirs for permanent storage as a mitigating technology to reduce greenhouse gas loading to the atmosphere (White 2003; IPCC 2005, Leung et al., 2014). Although a well-chosen and well-engineered CCS site is not expected to leak, all risks must be assessed to ensure safety, it must be shown that technologies exist that will give early warning and aid in remediation / mitigation if necessary, and accurate leakage quantification will be critical for carbon credit auditing and liability issues. In this regard, research at such leaking sites has been conducted to better understand gas migration pathways and processes (Pearce et al., 2004; Shipton et al., 2005; Stevens, 2005; Lewicki et al., 2007; Annunziatellis et al., 2008; Dockrill et al., 2010; Bigi et al., 2013), test new and existing technologies for site assessment, safety monitoring, and leakage quantification (Pettinelli et al., 2008; Pettinelli et al., 2010; Bateson et al., 2008; Jones et al., 2009; Sauer et al., 2014; Schutze et al., 2013; Schutze et al., 2012), and to assess potential health and ecosystem risks and impacts should such a site leak (Beaubien et al., 2008; Oppenheimer et al., 2010; Kruger et al., 2011; Frerichs et al., 2012).

The present research focusses on CO_2 leakage at surface and its potential impacts, examining a natural leakage site, the Latera caldera in central Italy, with the goal of addressing questions relative to both "pure" research of such natural sites as well as "applied" research related to CCS. The work on CO_2 leakage examines how sampling strategy and density can influence the possibility of finding and accurately quantifying leakage flux by inputting both real and synthetic data into a simulation program developed by the author, while the work on CO_2 impact looks at how the leakage of this gas could potentially alter the chemistry of near-surface potable groundwater resources via water-rock-gas interaction. This research was conducted in collaboration with the Fluid Geochemistry Laboratory, run by Professor Salvatore Lombardi, at the Department of Earth Sciences, University of Rome "La Sapienza".

The most common approach for leakage mapping and quantification is to measure the flux of this gas at individual points over a given study area and then estimate the total flux based on the interpolation of the obtained dataset or by applying a leakage population average to a given surface area (Bergfeld et al., 2006; Bergfeld et al., 2012; Cardellini et al., 2003; Chiodini et al., 2007; Chiodini et al., 2008; Evans et al., 2009; Hernandez et al., 2012; Lewicki et al., 2005; Lewicki et al., 2010). The precision and accuracy of the obtained results will depend on various factors, including whether the sampling density is sufficient to find most leakage points, whether the related anomalies can be spatially constrained, and whether it is possible to accurately separate the leakage flux rate from the omnipresent and diffuse biogenic CO_2 flux produced by soil respiration processes. Despite extensive research in this field, questions still remain regarding the magnitude of potential errors in the total flux estimates and how survey design could be improved to minimise them and increase the potential of finding leaks. To address some of these issues a simulation program has been written by the author which creates a high density (1 m spacing) large (1 x 1 km) grid populated with synthetic data that consists of background, coherently distributed, log-normal biogenic CO_2 flux values onto which individual areas of geogenic gas leakage are superimposed. This matrix is subsampled using different spatial approaches and at different densities, with each sampling density being repeated 100 – 200 times to define the statistical probability of finding a leak or the statistical distribution of total leakage estimates compared to the "true" value of the entire synthetic data set. In addition, high density real data can also be imported into the program and similar simulations can be conducted to

verify that the trends observed in the simplified synthetic system are applicable in the real world; this was done using a dataset of 547 flux measurements collected on a 10 m spacing grid over a field in the Latera caldera that has various leakage points. The main objectives of this flux and modelling work were to better understand what sampling strategies and densities yield the best chances of finding CO_2 leakage areas of different sizes and shapes, and to quantify the potential errors that may result during leakage estimates due to the chosen sampling density.

Although leakage is not expected from well-constructed CCS sites (based also on the fact that numerous natural, non-leaking CO_2 geological reservoirs have been discovered during deep hydrocarbon exploration drilling programs), concern has been raised regarding the potential risk that such industrial activities may pose to potable groundwater resources overlying storage reservoirs. As detailed in a number of recent review articles (Harvey et al., 2013; Lemieux et al., 2011; Lions et al., 2014; Little et al., 2010), potential impact could occur via two primary mechanisms: i) the leakage of CO_2 gas into a potable aquifer, which would cause in situ water-rock reactions due to acidification and ii) the displacement of deep brines from the storage reservoir into an overlying aquifer, which could make the aquifer saline and increase the concentration of associated toxic elements. To date the majority of studies related to the former have used laboratory measurements (Humez et al., 2013; Kirsch et al., 2014; Wunsch et al., 2014), geochemical modelling (Zheng et al., 2009; Darby et al., 2009; Navarre-Sitchler et al., 2013), small scale field experiments (Cahill et al., 2013; Humez et al., 2014; Kharaka et al., 2010; Trautz et al., 2013), or a combination of the three to estimate potential impacts, whereas those addressing the latter are typically conducted using physical flow models (Kirk et al., 2009; Birkholzer et al., 2011; Cihan et al., 2012; Walter et al., 2013). Although these approaches provide critical data needed to better understand these issues, laboratory measurements are limited by the fact that they cannot address geological complexity, scale, or slow kinetics, while modelling, by its very nature, requires simplifying assumptions that may not always be valid. For these reasons numerous researchers have recently started to study natural sites where these processes have been occurring (Choi et al., 2014; Keating et al., 2013a; Keating et al., 2013b; Keating et al., 2014; Lewicki et al., 2013, Lions et al., 2013), thus allowing for geochemical trends and processes to be defined for very long time periods (thousands to millions of years), over very large scales (hundreds to thousands of metres), and within complex and variable lithologies, structures, and hydrogeological systems. Whereas most of these recent studies have used natural springs and existing wells to examine these processes, the present work involved the unique approach of drilling six shallow boreholes along a transect which parallels the groundwater flow direction and which crosses a large CO_2 gas leakage area. One well was drilled up-gradient in background conditions, two were drilled within the leakage area itself, and the remaining three were drilled progressively further down-gradient away from the leak. The main objectives of this work were to determine how groundwater chemistry changes as it flows through a CO_2 leak, to understand the reactions and mineral phases which control how the groundwater chemistry is impacted, and to learn how the aqueous geochemistry continues to evolve as the impacted water migrates further down gradient out of the leakage zone.

The thesis is organised as follows. Chapter 2 gives the theoretical background related to CO_2 gas in terms of its multiple origins and the processes which control its migration through the lithosphere, giving an overview of issues that control leakage distribution and complicate its interpretation. This chapter also discusses the various difficulties associated with finding and quantifying geogenic (or potentially anthropogenic CCS) CO_2 leakage, as well as the various approaches used in the literature to address these issues. Chapter 3 describes the Latera Caldera, the site where both CO_2 flux and groundwater geochemistry studies were conducted for this research. The nature, origin, migration pathways, and surface leakage manifestations of the

geogenic CO_2 is described, along with the general geology, structural geology, hydrogeology, and gas geochemistry of the site. This chapter gives an overview of the extensive work already conducted in this area, and helps define the context and the rationale behind the work reported here. Chapters 4, 5, and 6 present the new research results and interpretations resulting from this dissertation work. Chapter 4 describes detailed CO_2 flux sampling conducted at the Latera site, data that is used to describe gas migration pathways and styles at the site and which is also used in the subsequent chapter as input for modelling purposes. Chapter 5 gives a detailed description of the simulation software developed by the author, followed by its use to simulate the influence of sampling strategy and density on the obtained results. These simulations look to quantify these effects in the hopes of improving survey design and success rates. Chapter 6 describes the groundwater work conducted at the Latera site, the first such detailed study to look at along-gradient aqueous geochemical modifications as groundwater flows through a major CO_2 leaking area. This chapter aims to determine impacts, as well as trying to understand if there are subsequent processes or reactions which can mitigate the impact down-gradient (e.g. natural attenuation). Each of Chapters 4-6 has a final sub-section which gives a summary of the results and some specific conclusions related to that work, while Chapter 7 discusses high-level conclusions for the entire research, and how the results may be used to move forward our understanding of the issues discussed.

Chapter 2. Theoretical Background

2.1. CO_2 characteristics

CO_2 is denser than air (Figure 1a) (thus it tends to accumulate in depressions above ground and migrate laterally via density-assisted flow in unsaturated soil), has a lower diffusion coefficient than smaller gas molecules like CH_4 or He (and thus slower to diffuse), is less viscous than air (Figure 1b) (thus less resistant to advective flow), and is highly soluble in water (thus gas-phase CO_2 can be attenuated in groundwater via dissolution, and its concentration in unsaturated soil can be reduced by dissolution in percolating meteoric water).

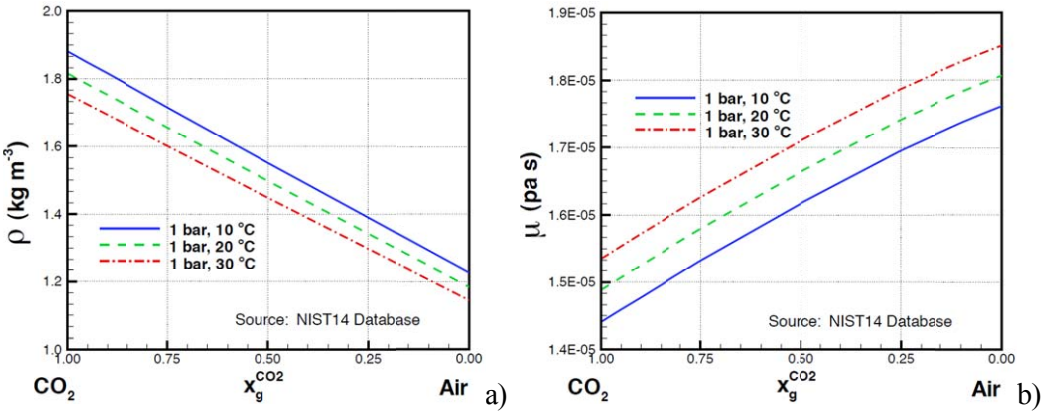


Figure 1. Variations of density (a) and viscosity (b) of a gas mixture having from 0 to 100 % CO_2 in air (Oldenburg et al., 2003).

2.2. Soil gas CO_2 origins

Soil gas CO_2 and CO_2 flux values can originate from two main sources: in situ, primarily biological production or the leakage of deep-origin, external CO_2 into the shallow system. The range of typical stable carbon isotopic values, reported as $\delta^{13}C$ relative to the Pee Dee Belemnite (PDB) standard, of various different organic and inorganic sources are reported in Figure 2.

In situ, "biogenic" production primarily involves soil respiration by autotrophs (plant roots and rhizospheric organisms) and heterotrophs (bacteria, fungi, etc.) (Kuzyakov, 2006; Arevalo et al., 2010), resulting in the concurrent consumption of O_2 and production of CO_2. This process is ubiquitous in soil, with soil biogenic CO_2 flux to the atmosphere being an important driver of the terrestrial carbon cycle (Section 2.2.1). Rates are controlled by a wide variety of soil and meteorological parameters, creating a background spatial distribution of values that is typically log-normal. Within this biogenic background distribution, anomalous biogenic values can occur as a result of auspicious production or accumulation conditions. Under normal conditions, maximum biogenic CO_2 soil concentrations and flux rates are typically less than 1% and 20 g m^{-2} d^{-1}, although values up to 20% and up to (and higher than) 100 g m^{-2} d^{-1} are possible. Intertwined and difficult to separate from biogenic production is in situ inorganic CO_2 production via dissolution / precipitation reactions involving pedogenic carbonate minerals (Section 2.2.2).

Leakage, on the other hand, occurs from natural geological sources ("geogenic") in geothermal and volcanic areas (Section 2.2.3), and there is also concern that it may occur in the future from man-made CO_2 geological storage reservoirs ("anthropogenic" CCS) (Section 2.2.4). A migration pathway, such as a fault or leaky well, is needed for leakage to occur (see Section 2.3.2). The addition of this external gas into the soil environment will result in dilution of the pre-existing gases (e.g. Annunziatellis et al., 2008), which are similar to atmospheric air concentrations but with greater water vapour and CO_2. The leaking gas tends to be localised in proximity above the vertical migration pathway, although specific conditions (e.g., a thick unsaturated zone or a CO_2 – charged aquifer) may result in spatial dispersion. The magnitude of the resultant values will depend on leakage rates, level of lateral dispersion, and soil conditions, however in extreme cases shallow soil gas concentrations can reach 100% CO_2 with leakage rates > 10,000 g m^{-2} d^{-1}.

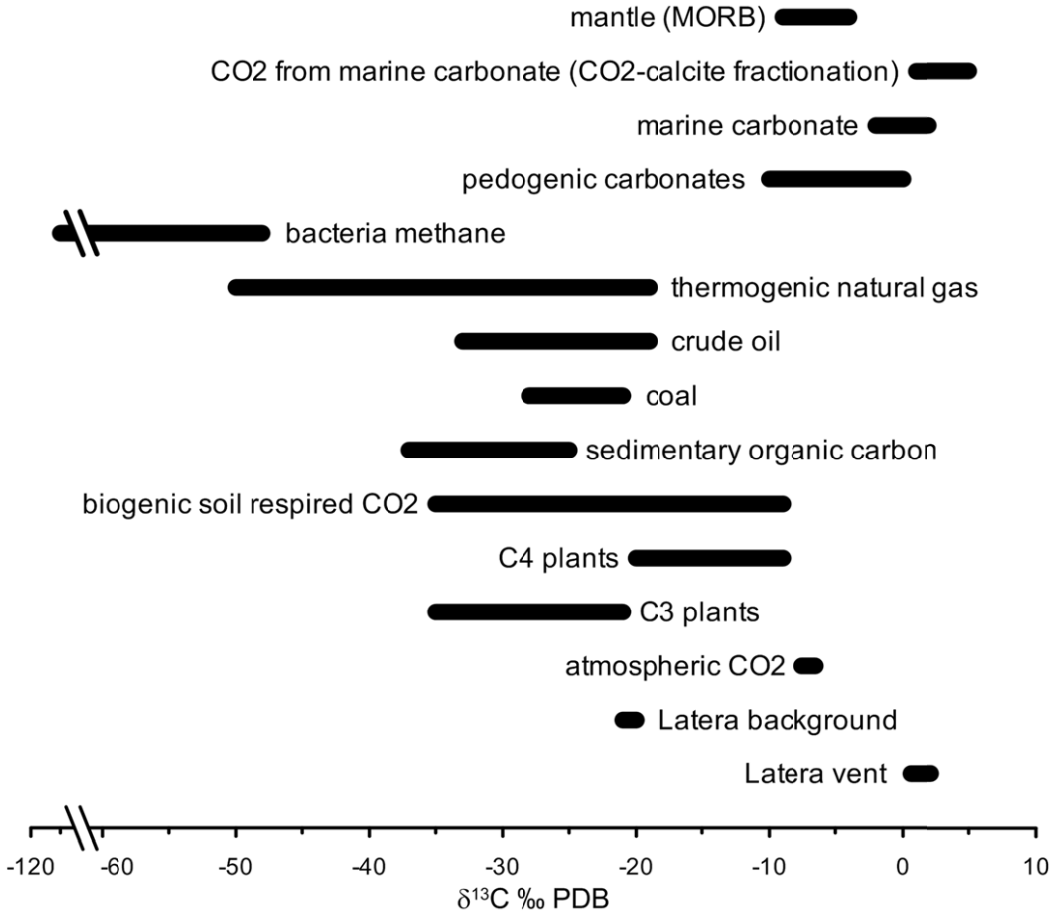


Figure 2. Range of $\delta^{13}C$ values for various inorganic and organic carbon sources (after Bergfeld et al., 2001a; Sano and Marti, 1995; Whiticar, 1996; Ramnarine et al., 2012). Note that the value for "Latera vent" (i.e. the deep-origin, geogenic CO_2 leaking at the surface) lies within the range of marine carbonates and "Latera background" (i.e. soil gas CO_2 collected in a non-leaking area) has values at the boundary between C3 and C4 plants (Latera data from Beaubien et al., 2008b).

Although part of the range of values caused by biogenic and leakage processes can overlap, different lines of evidence can be used to separate them and determine the true origin.

2.2.1. Biogenic CO_2

A solid understanding of soil biogenic CO_2 is critical for deep CO_2 leakage detection and quantification because it forms the baseline "noise" in which a deep-origin anomaly would occur. Separation of the deep anomaly from this background is challenging but critical to avoid false positives (biogenic anomalies interpreted as leakage) and false negatives (leakage anomalies interpreted as being part of the background biogenic distribution), and for an accurate baseline subtraction and/or threshold determination when calculating total flux from natural or human-induced leaks (section 2.4.2). The production, accumulation, and migration in the soil, and the exchange rate between the soil and the atmosphere, of this biogenic gas depends on many factors, like plant type, soil parameters, rainfall, temperature, and land use.

The biogenic CO_2 produced in situ in the soil originates from either living plants or from the breakdown of soil organic matter (SOM) formed from dead plant material (litter, roots, etc.), and as such its $\delta^{13}C$ is a direct function of these original sources (Amundson et al., 1998). Based on the photosynthetic pathway that they use for energy production, the bulk of most plants can be subdivided into either C3 or the more recently evolved (since the mid-Miocene) C4 plants. Due to the fractionation processes inherent in the different photosynthetic pathways, the C3 plants have a more depleted $\delta^{13}C$ signature (-34 to -21‰) compared to the C4 plants (-19 to -8 ‰) (Smith and Epstein, 1971); these ranges are illustrated in Figure 2. Autotrophic root respiration will produce CO_2 having a $\delta^{13}C$ signature that is directly related to the types of plants (C3 or C4) growing at that specific time, whereas heterotrophic breakdown of SOM will produce CO_2 that has a $\delta^{13}C$ signature that is an average of the different types of plants that have grown in that soil over time (although enrichment and depletion can take place during the plant-to-SOM and SOM-to-CO_2 conversion steps (Nadelhoffer and Fry, 1988; Schweizer et al., 1999)). Rotating of C3 and C4 crops over time can thus change and potentially widen the biogenic $\delta^{13}C$ signature in CO_2 soil gas and flux (Fassbinder et al., 2012). Other processes that can modify the values of $\delta^{13}C$ of biogenic soil CO_2 include mixing with atmospheric CO_2, diffusion fractionation, dynamic fractionation, and temporal variability in autotrophic respiration (Amundson et al., 1998; Nickerson and Risk, 2009; Risk et al., 2012), with diurnal changes in the $\delta^{13}C$-CO_2 value of up to 3‰ (Fassbinder et al., 2012).

Microbial and root respiration produces biogenic CO_2 via the consumption of O_2 in an almost 1:1 stoichiometric ratio. This is represented by the simple formula (Stumm and Morgan, 1995):

Eqn 1 $$\mathbf{CH_2O + O_2 \leftrightarrow CO_2 + H_2O}$$

where CH_2O represents organic carbon as a simple carbohydrate; in essence, this formula is simply the reverse of photosynthesis. The comparison of CO_2 with the other major soil gases can be useful to illustrate this process and to differentiate respiration from leakage, as the relative behaviour of these gases will vary depending on the controlling chemical / physical processes (Beaubien et al., 2004; Romanak et al., 2012a; Beaubien et al., 2013; Romanak et al., 2014). Figure 3 shows a scatter plot of CO_2 versus N_2 and O_2+Ar, comparing the values obtained above the Weyburn CO_2-EOR site and those from a shallow horizontal profile conducted across a CO_2 vent in the Latera caldera, together with the expected trends for both respiration and leakage (solid and dashed lines, respectively) (data from Beaubien et al., 2013). The Weyburn data clearly follows the shallow biogenic origin trend (implying no leakage), while the Latera results track along the leakage trend. This figure shows how the biogenic trend has a slope of about -1 for O_2+Ar versus CO_2 while the slope of N_2 versus CO_2 is essentially zero (as N_2 is not involved

in respiration). In contrast, the observed leakage trend is due to the dilution of both O_2+Ar and N_2 as CO_2 is added to the system. The limited number of Weyburn results which diverge from the biogenic trend actually lie in the opposite direction from that expected for a leakage signal, with off-trend N_2 values occurring above the normal soil gas concentration of 76-78% and the O_2+Ar values below the biogenic trend. Processes that might explain this behaviour include selective dissolution of CO_2 and associated enrichment of N_2 (Romanak et al., 2012a), denitrification, or methane oxidation.

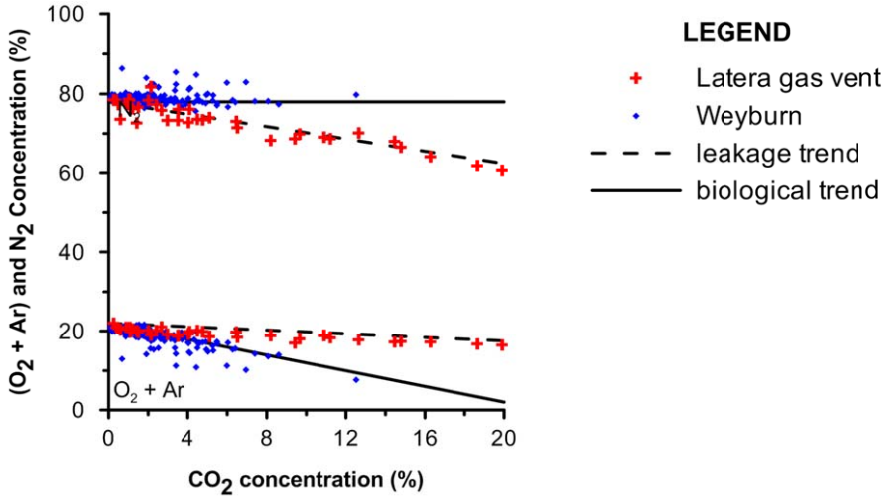


Figure 3. Scatter plot showing the relationships between CO_2 and O_2+Ar and CO_2 and N_2 from a gas leakage point at Latera compared to a normal, non-leaking site at Weyburn, Canada (data from Beaubien et al., 2013). See text for discussion.

The primary drivers of biogenic CO_2 production and subsequent transport in the shallow soil are temperature and water content, via their joint impact on metabolic rates and gas permeability of the soil. It is generally accepted that, of these two parameters, temperature has the most significant effect on soil gas CO_2 production (Moncrieff and Fang, 1999), with most studies indicating an exponential relationship (Hamada and Tanaka, 2001; Gorczyca et al., 2003; Hirano et al., 2003; Hashimoto and Komatsu, 2006). It has been shown that this relationship is stable and robust in deeper samples (Risk et al., 2002b). CO_2 flux also increases with increasing temperature via an exponential (Mielnick and Dugas, 2000) or Arrhenius (Lloyd and Taylor, 1994) relationship. However, diurnal hysteretic behaviour has also been observed, such that higher concentrations are found in the evening compared to the morning for the same given temperature (Arevalo et al., 2010; Barron-Gafford et al., 2011). This effect, which may be due to the time lag between photosynthesis and the delivery of labile carbon to the roots (Barron-Gafford et al., 2011), could potentially influence the inter-comparability of spatial datasets if a large number of flux measurements are made over the arc of a day.

Precipitation can be highly variable year on year, and thus it can have both a seasonal and annual influence (Beaubien et al., 2013; Flechard, 2007; Hamada and Tanaka, 2001; Mielnick and Dugas; 2000; Flanagan and Johnson, 2005). Rainfall can influence both near-surface CO_2 production and its subsequent migration in the soil and flux to the atmosphere. Numerous authors have shown that an optimal water content level exists for CO_2 production, as microbial activity is limited by a lack of water below the threshold and limited above by reduced gas diffusion rates to replenish the consumed O_2 (Bouma and Bryla, 2000; Risk et al., 2002b; Hashimoto and Komatsu, 2006). This effect on CO_2 production can also be very transitory, as soil respiration rates can increase significantly for a short period of time after a dry soil has been

re-wetted. While increased CO_2 concentrations in association with significant rainfall events have been observed to a soil depth of at least 85 cm (Maier et al., 2010), this mechanism is more important for very shallow soil (DeSutter et al., 2008; Bowling et al., 2011) and CO_2 flux (Barron-Gafford et al., 2011; Fierer and Schimel, 2003). The size and duration of an induced CO_2 flux peak appears to be greater for larger rainfall events (Borken et al., 2003) and when the rain comes after a dry period (Fierer and Schimel, 2003; Bowling et al., 2011), and it has been shown that this effect is much more pronounced in the summer than the fall (Borken et al., 2003). Regarding CO_2 migration, rainfall will decrease the effective gas porosity and permeability of the surface soil (Maier et al., 2010), thus potentially leading to an increase in soil concentrations (i.e., accumulation) and a decrease in flux to the atmosphere. High rainfall may also raise the water table, which could in turn cause an upward shift of the CO_2 depth-distribution curve (see Section 2.3.3). Such a shift was observed by Maier et al. (2010) at a site in SW Germany, where soil CO_2 concentrations increased strongly and converged towards a similar value at 27, 55, and 85 cm depths after rainfall caused the water table to rise.

The combined effects of temperature and precipitation are manifest in seasonal trends of biogenic CO_2 production and flux. Flux has been found to be particularly influenced, with maximum values during the wet, hot summers of continental temperate climates that are typically 1 (sometimes 2) orders of magnitude greater than those measured in the late fall or winter (Beaubien et al., 2013; Mielnick and Dugas, 2000; Frank et al., 2002; Risk et al., 2002b; Gorczyca et al., 2003; Jassal et al., 2005). This seasonal effect is also observed in soil gas data (Hamada and Tanaka, 2001; Risk et al., 2002b; Drewitt et al., 2005; Bekele et al., 2007) and continuous data from soil gas CO_2 monitoring probes (Beaubien et al., 2013; Hirano et al., 2003; Jassal et al., 2005; Arevalo et al., 2010; Maier et al., 2010; Barron-Gafford et al., 2011). In these studies minimum CO_2 values were encountered during the winter while the timing of maximum values depended on the local climate and vegetative cover, in agreement with soil respiration modelling results (e.g., Moncrieff and Fang, 1999). It must be noted, however, that the majority of these studies were conducted in continental North America or northern Europe. In contrast, in a Mediterranean climate like that of Italy, the period of lowest biogenic CO_2 values is often mid-summer, as the very hot, dry conditions decrease soil moisture content, which limits biogenic production and increases gas permeability and exchange with the atmosphere; in contrast, winter conditions in this climate are wet, which can promote soil CO_2 accumulation.

Biogenic CO_2 production and flux can also be highly spatially variable due to numerous factors. For example, different land-use practices (e.g. crop types, fertilizer use, tilling, etc.) have been shown to influence the spatial distribution of biogenic CO_2 concentrations and flux (Arevalo et al., 2010, Brüggemann et al., 2011, Carbonell-Bojollo et al., 2011). Beaubien et al. (2014) showed how soil CO_2 concentrations were significantly lower in forest soils than in the soils of cultivated or brush fields in central Denmark due to lower temperatures induced by the tree canopy. Lateral changes in soil type can influence in situ production based on the amount and reactivity of organic carbon, and can influence gas migration/accumulation based on soil properties like porosity and permeability. Finally some researchers have also found greater CO_2 respiration in low-lying, moist riparian areas compared to the adjacent hill slopes (Riveros-Iregui and McGlynn, 2009; Pacific et al., 2010). This may be due to the fact that topography can control various parameters that are important for biogenic CO_2 production in normal soils; for example depressed areas often have greater soil thickness and specific soil types (King et al., 1983), greater organic carbon (Landi et al., 2004), and higher soil moisture content (Brocca et al., 2007) than the higher grounds. This correlation has been used to distribute synthetic biogenic CO2 flux values in a spatially coherent manner (Section 5.2.2.1) for the modelling experiments described in Chapter 5.

2.2.2. Shallow inorganic CO_2

In addition to the primary organic source described above, a portion of CO_2 flux can also come from inorganic sources in the soil. This can potentially impact both the amount of surface CO_2 flux and its carbon isotopic signature.

Carbonates in the soil can be either primary "lithogenic" minerals that have a $\delta^{13}C$ equal to that of marine limestone (-2 to +2‰) or secondary "pedogenic" minerals that have $\delta^{13}C$ values between -10 to 0‰ (Ramnarine et al., 2012) (Figure 2). While lithogenic minerals are the result of protolith erosion, the pedogenic minerals are formed by in situ precipitation of mainly calcite due to vadose zone pore water over saturation. This occurs primarily by removal of water (i.e., resulting in a more concentrated solution) via soil drying or evapotranspiration (Serrano-Ortiz et al., 2010), meaning that this process is more important in arid to semi-arid climates. The bicarbonate ion necessary for calcite precipitation comes from both carbonate mineral dissolution (i.e. inorganic) and soil respired CO_2 (i.e., organic), while the calcite ion comes primarily from carbonate minerals but with a secondary source being the slower dissolution of Ca/Mg silicates (Emmerich, 2003).

According to the following two equations (Jin et al., 2009; Serrano-Ortiz et al., 2010), the dissolution of calcite (either primary or secondary) can be both a source and a sink for gaseous CO_2:

Eqn 2 $$\mathbf{2HNO_3 + CaCO_3 \rightarrow Ca^{2+} + 2NO_3^- + H_2O + CO_2}$$

Eqn 3 $$\mathbf{CO_2 + H_2O + CaCO_3 \rightarrow Ca^{2+} + 2HCO_3^-}$$

While the first formula represents the presence of other acids (such as via denitrificaiton or pyrite oxidation), the latter addresses only the carbonate system and the weaker carbonic acid. Considering only the second formula, many authors have pointed out that calcite precipitation (i.e. driving the reaction to the left) should be a source of CO_2 (Serrano-Ortiz et al., 2010). This has been used to explain increased CO_2 flux rates measured after rainfall events, with short term dissolution and subsequent re-precipitation of calcite causing a short-term release of inorganic origin CO_2. The cycling of both inorganic and organic carbon via this reaction (in both directions) explains the range of pedolithic carbon isotope values mentioned above (i.e. -10 to 0‰), as illustrated in the following formula for the dissolved organic carbon in the pore waters (Jin et al., 2009):

Eqn 4 $$\mathbf{\partial^{13}C_{DIC} = (\partial^{13}C_{H_2CO_3} + \partial^{13}C_{CaCO_3})/2}$$

This process has been found to be most important in soils with higher carbonate mineral contents (Emmerich, 2003), with higher inorganic flux rates being observed during dry seasons due to soil drying (Emmerich, 2003) and after rainfall events where complex dissolution/precipitation processes (thus degassing) occur over relatively short time periods (Serrano-Ortiz et al., 2010).

2.2.3. Geogenic CO_2

There are numerous inorganic, geological processes that can produce CO_2 at depth, which can, given appropriate conditions, migrate towards the ground surface and leak to the atmosphere. Although each process has a relatively narrow distribution of $\delta^{13}C$-CO_2 values (Figure 2) which

should allow interpretation of its origin, it is more common that multiple processes are involved and/or that the rising gas mixes with other CO_2 of a different origin, thus mixing the stable isotopic signature. Because this yields non-unique values, many authors have noted that $\delta^{13}C$ data by itself is insufficient to determine origin (Sano and Marty, 1995) and that it should be combined with other trace gases like He or Ne. In addition, the ^{14}C of deep origin CO_2, be it natural or anthropogenic from burning hydrocarbons, is zero due to the combination of the ^{14}C half-life and the great age of the CO_2 sources; for this reason ^{14}C can be useful for separating shallow biogenic from deep geogenic CO_2.

The main processes / origins of deep CO_2 include:

- Mantle degassing. The $\delta^{13}C$ of this primordial CO_2 is between -9 and -4 ‰ based on bubbles in quenched lava from Mid-Ocean Ridge Basalt (MORB) samples. Considering that the mantle is typically isolated deep below thick continental crust, CO_2 from this source typically occurs only at locations of active crustal rifting or subduction, or mantle arching or upwelling. In areas of subduction there is a high probability of metasomatism or assimilation which will mix the mantle and crustal signatures.

- Magma degassing. This source can be much more diverse because the composition of a magma will be highly site specific, being formed by the local tectonic, structural, lithological, fluid, and heat flow conditions, with the potential for mineral/volatile fractionation and host rock assimilation along the ascent pathway (Hoefs, 2009). The principle, original origin of the magma, such as mantle or crustal melting, plus the various mixing and contamination processes involved in its evolution, will determine the specific $\delta^{13}C$ value for each magma; this value may, however, change in time if conditions evolve at depth. Experimental petrology studies have also shown that the process of limestone assimilation causes precipitation of pyroxene and olivine (desilication of the melt, which renders it more carbon rich and buoyant) and the release of large quantities of CO_2 (Iacono-Marziano et al., 2007).

- Thermal decarbonation, metamorphism, and metasomatism. These three processes often overlap and can be difficult to separate. Thermal decarbonation involves the release of carbon (as CO_2) from carbonate mineral phases due to elevated temperatures, metamorphism can involve high temperature deformation, while metasomatism (often associated with metamorphism) involves alteration via liberated fluids. The resultant stable isotopic signature of the CO_2 produced from these processes will depend on that of the source, temperature, and involvement of fluids. In addition, isotopic fractionation between calcite and CO_2 at temperatures above 900°C during thermal decarbonation has been shown to enrich the CO_2 by up to c. 3‰ (Rosenbaum et al, 1994). Possible source material involved in these processes include:

    - Marine carbonates. These have a relatively narrow $\delta^{13}C$ range of -2 to +2 ‰. In localised geographical areas this range can be even narrower, such as typical values between 0 to +2 ‰ in the Apennine mountains; with fractionation this can increase to +1 to +5 ‰.

    - Hydrothermal calcite veins. These values will be more site specific, but are expected to be more depleted than the marine carbonates.

    - Organic-rich marine sediments. These include clays and siltstones, and because their carbon content is biogenic the $\delta^{13}C$ of this carbon is depleted, with a typical

range of -25 to -37 ‰. The conversion of this organic carbon to CO_2 can take place via either oxidation and/or hydrolysis (Bergfeld et al., 2001b).

Clearly the mixing of some of these processes and origins will have a significant impact on the resultant isotopic signature of the CO_2 that may leak at the surface. There are many examples of inferred multiple, complex origins described in the literature, such as degradation of organic matter and dissolution of calcite veins at the Geyser geothermal area, USA (Bergfeld et al., 2001a), magma degassing and limestone metamorphism at Santorini Island, Greece (Parks et al., 2013), and magma degassing and limestone assimilation in the Alban Hills, Italy (Iacono-Marziano et al., 2007).

In addition to the obvious release of CO_2 from active volcanic edifices, large volumes are also released to the atmosphere via more wide-spread, diffuse areas linked to faulting and fracturing. While a portion of this is still linked with volcanics via release along their flanks, the last 20 years has shown the volumetric importance of what is referred to generically as "non-volcanic diffuse degassing". These latter are typically found in active tectonic areas where thinned crust, deep faults, and/or high heat flow can produce (and allow to migrate) large volumes of CO_2 via a number of the mechanisms listed above.

The western part of central Italy represents a very large area over which diffuse CO_2 degassing occurs locally at surface, given appropriate gas migration pathways (faults and fracture networks). This area extends from north of Florence to south of Naples and from the Tyrrhenian Sea to some unknown point within the Apennine mountains. The eastern limit in some way appears to be linked with the east-west geological/tectonic divide which defines central Italy, in that the western part has a high heat flow, thin crust, and a depth to Moho of about 25 km, whereas the eastern part (from the Apennine mountains to the Adriatic coast) has normal heat flow, thick crust, and a depth to Moho of about 35 km. The western gas emissions manifest as, or are associated with, groundwaters with elevated pCO_2 values, thermal springs, bubbling pools, and/or dry gas emissions ("gas vents" or moffetes). Although this degassing area overlaps and is often associated with Quaternary volcanics (Roman Magmatic Province) and actual volcanism (e.g. Vesuvius), other non-volcanic manifestations are observed, for example, above geothermal fields (e.g. Lardarello; Minissale, 2000) and within "cold" intra-montane basins (e.g. the San Vittorino valley; Giustini et al., 2013).

The origin of this widespread CO_2 degassing in western Italy has been debated over the years, with theories including: metamorphic hydrothermal reactions of Mesozoic carbonates, based on travertine and gas isotopes (Panichi and Torgiori, 1975); mantle degassing, based on few $\delta^{13}C$-CO_2 samples and coincidence with the shallow Moho (Minissale, 1991); magmatic and metamorphic sources, based on a re-interpretation of the same isotopic data (Marini and Chiodini, 1993); metamorphism and decarbonation of carbonate formations, based on $\delta^{13}C$-CO_2 data (Chiodini et al., 1995); metamorphic alteration of carbonates and isotopically depleted organic carbon from Neogene basins (Minissale, 1997); crustally contaminated mantle, based on $\delta^{13}C$ of dissolved inorganic carbon (DIC) in regional groundwater and the assumption that fractionation during carbonate metamorphism would produce higher $\delta^{13}C$ results (Chiodini et al., 2000; Chiodini et al., 2004; Frondini et al., 2008); and limestone metamorphism triggered by mantle melting and intrusion, based on stable isotopes of carbon and helium (Minissale, 2004). This range of possibilities highlights the complexity of gas generation and migration in a tectonically active area like central Italy.

On a regional scale Chiodini et al. (2000) used a mass balance approach on DIC from groundwater springs to illustrate that deep-origin CO_2 is not only leaking in the western peri-Tyrrhenian area, but that it may also be leaking in the western part of the Apennine mountains. Unlike the area west of the mountains where there are evident emissions, leakage is not typically visible in most of the mountains themselves. Instead these authors contend that excess "external" CO_2 is leaking into the large, high transmissivity, regional carbonate aquifers in the mountains, which have the capacity to dissolve this gas and thus prevent it from leaking at the surface. Minissale (1997; 2004) also highlighted the importance of the regional Mesozoic carbonate aquifers in controlling fluid flow and gas leakage anomaly magnitude and location, while Minissale (2000) illustrated the tectonic control that separates the CO_2-rich domain to the west and CH_4-rich domain to the east and north-east of the Apennine mountains.

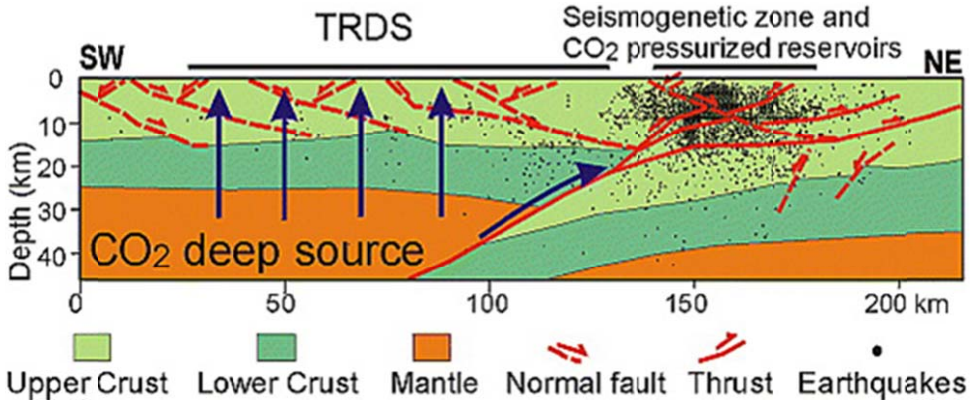


Figure 4. Schematic geological section across the central Italian peninsula showing the proposed link between a regional CO_2 source, normal and thrust faulting, and earthquake occurrence in the western Apennine mountains (Chiodini et al., 2004).

Chiodini et al. (2004) extended their earlier work regarding CO_2 degassing in western central Italy by proposing a large-scale conceptual model that addresses gas origin and migration, and their potential link with Apennine seismicity. In this model (Figure 4), the shallow mantle upwells and infiltrates into the ductile deeper crust, where crustal contamination occurs and CO_2 is liberated and migrates along faults towards the surface. On the eastern side of this domain the gas follows in-active thrust planes and then eventually intersects active normal faults in the western portion of the Apennines, where structural traps allow for accumulation, over-pressurisation, fault weakening, and eventual earthquakes.

The potential for regional mantle melting and interaction with, and contamination by, carbonate-rich crustal material has also been recently proposed by Frezzotti et al. (2009) using non-geochemical data. These authors used seismic results to define a low velocity layer (60-130 km deep beneath the Tyrrhenian, 30-50 km deep beneath the Roman magmatic complex) that they interpret as a geochemically anomalous mantle caused by low fractions of volatile-rich melts or fluids. These authors contend that carbonate metasomatism and melting induced by a hot mantle causes higher buoyancy and lower viscosity in the resultant magma, inducing upward migration into shallower levels where large volumes of CO_2 can degas.

In all these papers the importance of the Mesozoic carbonates as shallower traps which accumulate CO_2 and define surface leakage routes is noted. The origin of the gas leaking at the Latera site, which fits within this regional context, is addressed in Chapter 3.

2.2.4. Atmospheric and Anthropogenic CO_2

There is a constant exchange between gases in the atmosphere and the soil, such that the composition of shallow soil gas is typically very similar to that of air (aside from higher water vapour and CO_2). The concentration of CO_2 in the atmosphere is about 380 ppm, while its $\delta^{13}C$ is about -7.5 ‰ (Amundson et al., 1998). In a natural system these parameters are at pseudo steady state via the balance between biogenic and geogenic inputs and consumption via rock weathering. As is well known, however, anthropogenic emissions linked to fossil fuel burning for energy production have upset this balance by introducing a new input term. As such concentrations have increased by about 100 ppm since the start of the industrial revolution, increasing the heat trapping capacity of the atmosphere, and contributing to human-induced climate change.

Anthropogenic CO_2 is in fact organic carbon, in the form of the fossil fuels that are burnt for energy production. Due to their organic origin they have a strongly depleted $\delta^{13}C$ signal (Figure 2) which is imparted to the CO_2 that is produced upon their burning. As described above, CCS involves the capture of fossil fuel emissions, primarily CO_2, from point sources and their injection into the deep subsurface for permanent trapping within porous geological reservoirs. If CCS eventually obtains approval and moves forward from the present research stage to full industrial deployment world-wide, coal will likely be one of the first target applications due to its high CO_2 emissions per unit energy produced and its significant global reserves. In theory the isotopic signature of coal, and more importantly of the CO_2 produced upon its burning, could potentially be a useful tracer to define if a portion of the stored CO_2 is leaking at surface. Unfortunately this is not the case, as the stable C isotopic signature of coal is very similar to that of CO_2 produced by modern C3 plants, as recently illustrated by Beaubien et al. (2013); these authors had to use other techniques to prove that an alleged CCS leak was due instead to near-surface biological processes.

2.3. Gas migration

2.3.1. Migration mechanisms

A complete review of all mechanisms involved in the migration of CO_2 from hundreds of metres depth towards the surface is beyond the scope of this work. The interested reader is referred to the literature for discussions on such topics as: migration of super critical CO_2 in CCS storage reservoirs laterally via pressure differentials (Bickle et al., 2007) and vertically via buoyancy forces (Hesse et al., 2010); density-driven convective flow induced by the dissolution of CO_2 in brines in fractures (Chen and Zhang, 2010) and saline aquifers (Neufeld et al., 2010); and dissolved phase CO_2 migration in a CCS reservoir (Emberly et al., 2005) and in stacked aquifers above a natural CO_2 reservoir (Kampman et al., 2013).

Instead the following addresses the processes of advection and diffusion of CO_2 in the unsaturated horizon, which, depending on the depth to the water table, generally encompasses the upper 1 to 100 m of the sub-surface. Due to conditions at the Latera study site this discussion will be limited to soil and unconsolidated sediments, and will not address unsaturated fractured rock. Typically sediments will consist of a three phase system made up of solid grains, water and gas. The porosity (ϕ) of a sediment is defined as the volume of the pores (Vp) divided by the total

volume (V_T), with the pore volume consisting of the sum of the water filled (V_W) and gas-filled (V_G) volumes:

Eqn 5

$$\varphi = \left(\frac{V_p}{V_T}\right) = \frac{(V_G + V_W)}{V_T}$$

Gas saturation is defined by $S_G = V_G/V_p$ and water saturation by $S_W = V_W/V_p$, such that $S_G + S_W = 1$. In this system water will be the wetting fluid, meaning that it will first occupy the smaller pores at low S_W while gas will fill the larger pores. As S_W increases (and S_G decreases) the water will progressively fill the larger pores. This distribution impacts on the effective permeability of each of these two phases (and thus flux) at different water/gas saturation levels.

Gas in the system described above partitions into each of the phases, such that:

Eqn 6

$$C_T = \theta_G C_G + \theta_l C_l + \rho_b C_{ad}$$

where C_G, C_l, and C_{ad} are the mass concentrations in the gas, liquid, and adsorbed phases, θ_G is the volumetric gas content ($\theta_G = V_G/V_T$), θ_l is the volumetric water content ($\theta_l = V_W/V_T$), and ρ_b is the bulk density of the sediment.

Concentrations in the gaseous (C_G) and liquid (C_l) phases are related via the Henry's constant ($C_G = K_H C_l$) and adsorption can be related to liquid concentration via an adsorption isotherm (e.g. $C_{ad} = K_d C_l$). Adsorption is not important for CO_2 because this gas is non-polar, however its reactivity related to precipitation and dissolution of carbonate minerals means that there is the potential for the solid phase to act as a sink or a source of CO_2 in sediments and soil (see Section 2.2.2). Clearly the phase in which the CO_2 occurs will have a defining effect on its mobility.

Recently a number of excellent reviews have addressed gas movement in the unsaturated zone (Scanlon et al, 2002; Webb, 2006a; Webb, 2006b; Kuang et al., 2013). The following is a brief, simplified summary based on these reviews, focussing primarily on Scanlon et al. (2002) and following the mathematical conventions used in that article (unless otherwise cited). Note that although here (and elsewhere) advection and diffusion are described separately as if they are completely independent, it must be noted that in a complex multicomponent gas system it is not possible to fully separate the effects of these processes due to the influence of one on the other (Scanlon et al., 2002); under certain conditions these interactions can be significant (Webb 2006a).

2.3.1.1. Advection

Advection is movement induced by a pressure gradient. This gradient can be formed by near-surface processes in the soil via strong winds, and by rapid changes in barometric pressure ("barometric pumping") or the height of the water table (Kuang et al., 2013). Advection caused by these processes tend to occur, however, only when the changes are abrupt, the permeability of the soil is high, and, in the last case, the water table is shallow (Scanlon et al., 2002). While these processes will be intermittent, continual natural advection is observed at sites where large volumes of deep-origin gas are leaking to the surface (e.g. Annunziatellis et al., 2008). In addition, advection can also occur via human-induced forcing, such as soil vapour extraction methods used for the remediation of subsurface gasoline spills (e.g. Nguyen et al., 2013). Advective flux is termed "non-separative" because it does not separate the gas components or the component isotopes (e.g. Camarda et al., 2007), as does most forms of diffusion.

Because air is about 50 times less viscous than water (and CO_2 is less viscous than air; Figure 1b), significant gas flow can occur via advection at relatively small pressure gradients. Viscosity is included in Darcy's Law, which is the most commonly used formulation to describe advective flow:

Eqn 7

$$\mathbf{J_G = -\frac{k_G}{\mu_G} \nabla P}$$

where J_G is volumetric flux, k_G is the permeability, μ_G is the dynamic gas viscosity, and ∇P is the pressure gradient. In soil pores containing both gas and water, gas permeability (k_G) must be expressed as the effective permeability at a given gas saturation level:

Eqn 8

$$\mathbf{J_G = -\frac{k_{rG}(S_G)k_G}{\mu_G} (\nabla P) = -\frac{k_G\,(S_G)}{\mu_G} \nabla P}$$

where k_{rG} is the relative gas permeability at a given gas saturation (S_G). The ideal gas law can be used to convert the volumetric flux J_G into a molar flux N^V, such that:

Eqn 9

$$\mathbf{N^V = -\frac{P}{RT}\frac{k_G\,(S_G)}{\mu_G} \nabla P}$$

Note that Darcy's Law is an empirical formula developed for water flow in sediments. However, water is essentially incompressible while gas is compressible. If the pressure gradient is sufficiently small the gas compressibility can be ignored in the formulation of the governing equations, however if it is large this factor will have to be taken into account. Another important deviation from Darcy's Law for gas flow is that while the water velocity is zero at the pore walls, the gas velocity is not zero (called the "slip velocity"). This additional "viscous slip flux" (also known as the "Klinkenberg effect") means that Darcy's Law tends to underestimate gas flux, with the deviation being most pronounced at low pressures and absent at high pressures where the gas behaves as a liquid. Finally, at high velocities the relationship between flux and the pressure gradient becomes non-linear, flow becomes turbulent, and inertial effects become important. Under these conditions Darcy's Law overestimates gas flux, thus necessitating the use of an additional non-linear flow resistance term called the Forchheimer Extension (Webb, 2006a).

In addition to standard, pressure-driven advection, density-driven advection is also possible under certain conditions. Based on numerical simulations this mechanism should be most important in highly permeable sediments and depend on the density contrast between the gas of interest and the surrounding gas. Although often studied in terms of the movement of dense volatile organic contaminants, this process was also found to be important in explaining the movement of dense CO_2 in the unsaturated zone at a natural geogenic gas leak at Mammoth Mountain, California (Altevogt and Celia, 2004).

2.3.1.2. Diffusion

Diffusion is movement induced by a concentration gradient. In normal soils diffusion is the main process that controls the primarily vertical movement of biogenic CO_2 (produced in the soil by plants and microbes), and is important for lateral and vertical movement of geogenic CO_2 once outside of the advective leak regime. In general diffusion is a "separative" process, in that differences in diffusion coefficients between heavier and lighter gases, and between heavier and lighter isotopes, result in a separation of these components along the path length. As an example, it has been shown that the $\delta^{13}C$-CO_2 of diffusing biogenic-origin CO_2 can be depleted by up to

4.4‰ relative to its source organic matter due to fractionation between the faster ^{12}C and the slower ^{13}C (Amundson, 1998), while significant separation would be expected between a light gas like He (MW = 1 g) and a heavy one like CO_2 (MW = 44 g). As discussed below, however, one diffusion mechanism (non-equimolar) is non-separative. Because gas molecular diffusion coefficients are about 4 orders of magnitude greater than those in water, gas diffusive fluxes are generally much greater than in water (Scanlon et al., 2002).

Diffusion in the gas phase is more complicated than in the liquid phase, as it can include various types like molecular and non-equimolar diffusion (grouped together as "bulk diffusive flux"), and Knudsen diffusion. The distinction and mode of movement of these different forms of diffusion are often a function of the relationship between the pore size (λ_p), which is the average distance between sediment particles, and the mean free path (λ), which is the average distance that a molecule travels before colliding with another molecule.

Molecular diffusion occurs under isobaric, isothermal conditions in which equimolar gas pairs (e.g. CO and N_2) counter-diffuse in a porous medium such that $\lambda_p >> \lambda$. In this manner molecule-molecule collisions dominate over molecule-wall collisions, thus minimising the effect of the solid phase and approaching, conceptually, free air diffusion. The effective gas diffusion coefficient in sediments is derived from that in free air by introducing variables that take into account the volume available for movement and the pathway length:

Eqn 10 $$\mathbf{D_{ij}^e = \tau\, \theta_G\, D_{ij}}$$

where D_{ij}^e is the effective diffusion coefficient of gas *i* in gas *j*, τ is the tortuosity (i.e. the length of the "tortuous" path in a porous media divided by the corresponding straight line length), θ_G is the volumetric gas content, and D_{ij} is the free-air diffusion coefficient.

Non-equimolar diffusion is due to gas components having different molecular weights. Given that all gas molecules in an isothermal and isobaric setting must have the same kinetic energy, it follows that lighter molecules move more rapidly than heavier molecules. In a binary system the more rapid diffusion of the lighter molecules will result in an increase in the pressure, which in turn will result in a pressure induced flux known as non-equimolar flux or "diffusive slip flux" (some authors refer to this as a type of induced advection, rather than a type of diffusion; e.g. Webb 2006a). Because of the influence of pressure in this process, non-equimolar flux is non-separative, unlike the other diffusion types.

Knudsen diffusive flux (or "free-molecule flux") occurs when $\lambda >> \lambda_p$, such that molecule-wall collisions dominate over molecule-molecule collisions. This type of diffusion is important in fine-grained materials with small pore sizes, and depends on the molecular weight of the molecules but is not influenced by the presence of other gas species.

Modelling of diffusive flux is typically conducted using Fick's Law, however this is truly only valid for molecular equimolar diffusion under isobaric and isothermic conditions, predicting the movement of only one component. Originally it was developed empirically for molecular diffusion of solutes in a liquid, and has been applied in the gas phase to describe molecular diffusion of gas *i* in gas *j*:

Eqn 11 $$\mathbf{J_{iM}^m = -D_{ij} C \nabla x_i}$$

where J_{iM}^m is the molar flux of gas i, D_{ij} is the diffusion coefficient of gas i in gas j, C is the total molar concentration, and ∇x_i is the mole fraction gradient; as stated by Webb (2006a), in this

form the mole flux is relative to the molar-average velocity, not to stationary coordinates. Other limitations of Fick's Law include the necessity that the mole fraction of the diffusing gas should be low, the diffusion coefficients of the two gases should not differ by more than a factor of 2, and the flux of one component must not depend on the flux of the other components (Scanlon et al., 2002).

A more rigorous approach to modelling gas diffusion in a porous medium is given by the Dusty Gas Model (DGM), which is not empirical but rather was developed based on the kinetic theory of gases. The name originates from the fact that the grains of the sediment are treated as giant molecules ("dust") that constitute an immobile component of the gas phase. In contrast to Fick's Law, the DGM considers all types of diffusion and can predict the flux of all components in a gas mixture; the DGM system of equations have been solved both analytically and numerically. The DGM is the only model that can couple advection and diffusion (Webb, 2006a) because it is the only one that considers Knudsen diffusion (Scanlon et al., 2000). Finally the Stefan-Maxwell equation (SME) can also be applied; this formulation can be obtained from the DGM in cases where Knudsen diffusion is negligible (i.e. no molecule-wall collisions).

Because the DGM can predict the flux of all components its use in multicomponent systems is critical, as the improper use of Fick's Law can result in attributing incorrect processes to secondary components (such as microbial degradation of oxygen) when the true process may be another (like physical displacement). Model application can be summarized, according to Scanlon et al. (2002), as follows: Fick's law can be used for high permeability material under isobaric conditions and low diffusing gas concentration (although the DGM and SME can also be applied), while the SME should be used if the concentrations are high; the DGM should be used for low or high diffusing gas concentrations in low permeability material under isobaric conditions; the DGM coupled with Darcy's Law for advection should be used for low or high diffusing gas concentrations in low permeability material under non-isobaric conditions (i.e., presence of a pressure gradient); and the DGM should be used for high diffusing gas concentrations in high permeability material under non-isobaric conditions while Fick's law coupled with Darcy's law can be used for the same situation for low diffusing gas concentrations.

2.3.1.3. Attenuation or migration via dissolution

As a result of the high solubility of CO_2 in water, the interaction between these two phases can influence the migration and spatial distribution of this gas. If the water content distribution is relatively stable then gaseous and dissolved CO_2 will equilibrate as a function of the Henry's constant and the movement of the dissolved CO_2 will be attenuated, because it is partitioned into an essentially immobile phase. As discussed earlier (see Section 2.2.2), based on mineral saturation levels, this CO_2 may be re-released via calcite precipitation. If, instead, water content changes as a result of the infiltration of meteoric water, the dissolution of CO_2 into this downward migrating water can result in the transfer of shallow CO_2 (e.g. biogenic CO_2 produced in the root zone) deeper into the soil column. If this water reaches the water table, the dissolved CO_2 can then be transported laterally with groundwater flow.

In addition to this chemical dissolution effect, the presence of two fluid phases (water and gas) influences physical flow through the mutually dependant effective permeabilities for each phase (as a function of relative volumetric water and gas contents, and the fact that water is the wetting phase so gas will tend to occupy the larger pores). For example, although not reported here, Darcy's law can be rewritten to address two phase, unsaturated flow, taking into account the

pressure difference (capillary pressure) between the two phases and their relative permeabilities (Webb, 2006b). In the case of diffusion, changes in relative water and gas contents can impact on the porous media tortuosity value.

2.3.2. Deep Pathways

Migration and leakage of gas is a natural geological process, as shown, for example, by hydrocarbon signatures observed in the shallow subsurface above oil and gas reservoirs or tracer gas signals measured above geothermal fields. In general, potential leakage pathways can be either natural (such as fracture systems and faults) or man-made (such as deep wells). Other potential natural leakage pathways, such as diffusion through caprocks or up-dip movement within a sloped reservoir, are not discussed here due to their extremely slow migration rate.

2.3.2.1. Faults

The principle natural pathways that can transport fluids from deep underground towards the surface are faults and fracture zones, which are structural discontinuities that cross-cut and potentially compromise the vertical barrier properties of otherwise intact rock units. The hydraulic characteristics of faults are, however, highly complex, and the fact that they can represent fluid conduits, barriers, or a combination of the two means that they have a critical influence on numerous subsurface processes. Extensive research has thus been conducted to better understand, for example, the impact of fault zone permeability on hydrocarbon reservoir integrity and exploitation, groundwater flow, earthquake development, and isolation of man-made waste (such as industrial fluids, nuclear waste, or anthropogenic CO_2).

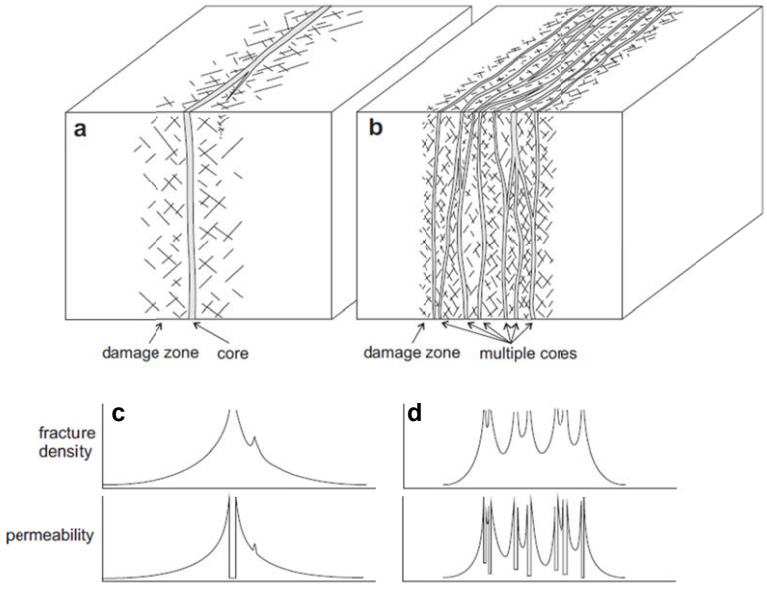


Figure 5. Spatial distribution of the fault core and damage zone with single (a) and complex multiple (b) fault planes, as well as associated fracture density and permeability (c, d) (Faulkner et al., 2010).

In its simplest form a fault can be subdivided into two zones, a fault core (FC) and a damage zone (DZ) (Figure 5a). The FC is the interval across which the majority of strain and displacement is accommodated, while the DZ represents a surrounding interval where secondary

slip and dilational features are formed as a result of deformation and strain dissipation into the protolith (Caine et al., 1996). More complex geometries commonly occur, however, whereby multiple FCs branch and anastomose, associated DZs overlap, and protolith blocks are entrained into the fault structure (Figure 5b); in addition, at a larger scale, the linking of multiple, sub-parallel faults along strike can form additional dilational and folding features (Faulkner et al., 2010; Bense et al., 2013). Fault form, style, size, and permeability are all variable both spatially and temporally, with spatial changes occurring both along dip with depth and along strike; this variability and the factors that control it are discussed below.

Vertical variations are controlled primarily by the combined effects of pressure and lithology, with lithology being sub-divided into the broad categories of competent, brittle rocks with low intrinsic porosity (e.g., crystalline, volcanic, and carbonates rocks) and unlithified, porous rocks (e.g., sands, clays, shales).

The FC of brittle rocks vary markedly with depth, such that a uniform unit may produce fault gouge and/or breccia near surface that evolves into cataclasite, phyllite, and mylonite with increasing lithostatic pressure (Sibson, 2000). Gouge is fine grained material formed by either equant particles or phyllosilicates / clays and has a low intrinsic permeability, while breccia consists of angular fragments in a matrix of <30% fine grained material that can have low permeability if compacted or cemented or a moderate permeability if not (Bense et al., 2013; Uehara and Shimamoto, 2004). With increasing pressure fault movement breaks and crushes grains (comminution) to form fine grained, impermeable cataclasites. Finally the combined effects of increasing pressure and temperature causes ever more ductile deformation which results in phyllites and mylonites. With the exception of non-cemented breccias, these various FC styles are isotropically impermeable and if continuous tend to form effective barriers to both vertical and horizontal flow.

In sharp contrast to the pore-scale-controlled low-permeability of FC rocks, permeability in the DZ of brittle rocks tends to be higher and fracture controlled (Caine et al., 1996) (Figure 5 c,d). The DZ, if present, is thought to first form during the initial stages of a new fault, with the creation of a band of dilational fractures and slip surfaces in the protolith during the period prior to failure along the proto-FC (e.g. Agosta and Kirschner, 2003). With subsequent movement the width of this interval can remain the same if strain is concentrated along the principle slip surface ("strain softening") or can continue to enlarge if strain dissipation is more diffuse ("strain hardening") (Gray et al., 2005). Conjugate fracture systems tend to form in these conditions, although examples have been found where orientations can be conditioned by pre-faulting discontinuities, such as cooling cracks in volcanic rocks (Riley et al., 2010). The size and number of fractures tend to decrease exponentially moving away from the FC and into the protolith (Bense et al., 2013). The level of lithostatic pressure can have a significant impact on the hydraulic characteristics of these features, with lessening pressure allowing for more open fractures. Field and modelling studies have shown the importance not only of the permeability of individual fractures, but also their three dimensional geometry, orientation, and how well they are inter-connected (Gudmundson et al., 2001; Bigi et al., 2013; Faulkner et al., 2010). Gas migration along fractures beneath the water table is a function of relative fluid input pressure, with low gas pressures resulting in individual bubble migration and high gas pressures causing the gas to move as a single continuous phase in the centre of the fracture while water flows along the fracture surface (annular flow) (Ranjith et al., 2006). In general, if not sealed by the precipitation of secondary mineral phases or completely closed by lithostatic pressure, brittle rock DZs tend to be highly permeable.

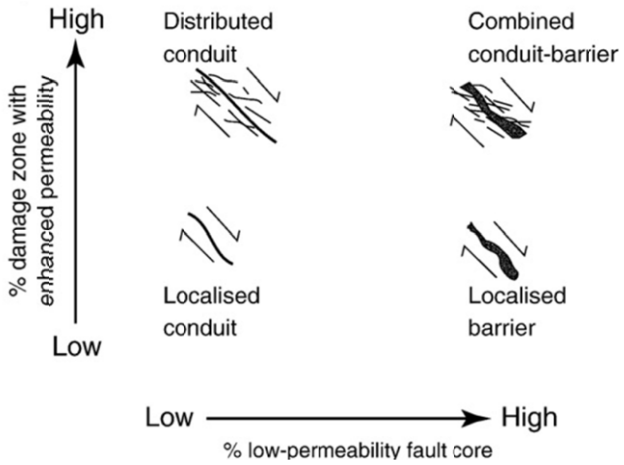


Figure 6. Schematic classification of the hydraulic conduit / barrier characteristics of faults based on the relative importance of core versus damage zone development (Bense et al., 2013).

If both the FC and the DZ are well developed this type of fault zone is often anisotropic, with permeability being high parallel to the structure within the DZ and low perpendicular as a function of the FC and protolith. However, considering that these two zones can be present or absent, and when present can take different forms, a qualitative descriptive classification of the hydraulic character of a fault in coherent rocks has been defined based on four end-members (Caine et al., 1996): low DZ and FC is classified as a localized conduit, high DZ but low FC as a distributed conduit, high FC but low DZ as a localized barrier, and high FC and DZ as a combined conduit-barrier (Figure 6). Examples of such end-members are described from an underground research laboratory in the volcanic tuffs of Yucca Mountain, where various deformation styles can be observed to grade into each other along strike and along dip as the stress regime and lithology changes (Gray et al., 2005). Besides anisotropy, various faults also show permeability asymmetry, with the permeability being greater in either the footwall (e.g. Agosta and Kirschner, 2003) or the hangingwall (e.g. Géraud et al., 2006); as noted by Riley et al. (2010), a wider hanging wall DZ may indicate near-surface faulting.

The deformation products in non-coherent, plastic, or porous rocks are different than those observed in brittle rocks due to how strain is dissipated; this difference is primarily in the shallower subsurface, as greater lithostatic pressures at depth causes compaction, de-watering, and lithification that results in rocks with more brittle characteristics (Takahashi, 2003; Cuisiat and Skurtveit, 2010). Initial movement in these shallow, typically sedimentary rocks takes place via particulate flow, which consists of the movement of grains past or away from each other. Because of this, in very shallow rocks there may not be any significant change in the original permeability. With further movement and greater depth these dilation bands (which can increase permeability) evolve into shear bands (which can decrease permeability); these features occur primarily within the fault DZ (Bense et al., 2013).

Instead, various processes occur in the FC which can result in either increased or decreased permeability, including grain orientation, grain mixing, and clay/shale smearing, all of which involve entrainment of particles along the slip surface (Takahashi, 2003). Grain orientation involves the alignment of elongated grains parallel to the fault plane, resulting in a permeability anisotropy that can reach two orders of magnitude. Grain mixing is the processes of dragging grains along the fault plain from one bed to different point in the stratigraphy, causing mixing of grain sizes; if sand is incorporated along the fault surface there can be an increase in permeability (Bense et al., 2013). Clay or shale smear (Figure 7) is an important process that has been studied

extensively in the hydrocarbon industry due to its impact on reservoir integrity in sedimentary basins (and is also observed in carbonate rocks). Here the dragging of a plastic layer like clay forms a barrier to cross-fault flow. The sealing capacity of a clay smear is a function of the competency contrast between adjacent layers (e.g. shale and sandstone) and the ratio between bed thickness and fault throw (Cuisiat and Skurtveit, 2010). In particular, if throw is less than bed thickness the smear will be incomplete, but if it is much more the smear will thin and lose part of its sealing capacity (Takahashi, 2003). Smearing of alternating sand-clay sequences may result in the incorporation of sand lenses within the fault core, with the potential for increased fault-parallel permeability (Cuisiat and Skurtveit, 2010). Finally, clay content in a sand can have a significant impact on deformation style (Caine et al., 1996), as experiments have shown that as little as 10% clay can reduce the overall fault width by up to 75% compared to a pure sand (Antonelli and Aydin, 1995).

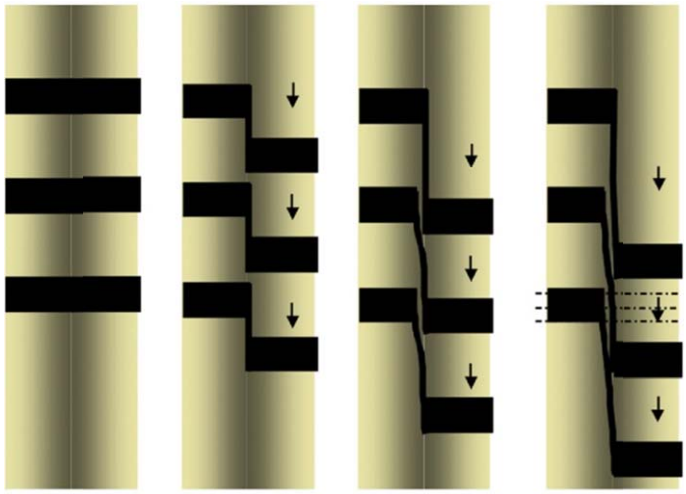


Figure 7. Schematic diagram showing the formation and potential evolution of a clay smear (Cuisiat and Skurtveit, 2010). Increased throw could result in thinning and eventual breaching of the clay membrane, thus changing it from a local barrier to a local conduit.

The combined effect of coherent and non-coherent rocks in a single stratigraphy being impacted by the same fault has been shown, for example, in the study of welded and non-welded tuffs. In this research deformation style was strongly conditioned by lithologic rheology, with the welded tuffs showing extensive fracturing in a well-defined DZ while the non-welded tuffs had deformation bands (Riley et al., 2010).

Horizontal, along-strike variations are controlled primarily by changes in the stress field, although lateral lithological changes can also have an effect. These changes are manifest by the formation of dilational features that have the potential for long-lived, high-volume fluid flow because they are more likely to remain open and stress conditions have likely limited the formation of impermeable gouge or smear (Gartrell et al., 2004). Numerous studies have shown that flow tends to be concentrated in such "complex" zones (e.g. Dockrill and Shipton, 2010). These dilational features can be the result of the linking of in-line smaller faults (e.g. relay ramps, folds), lithology changes (e.g. dilational jogs), or primary fault architecture (e.g. fault tips, horsetail fractures in shallow hanging walls, fracture intersections) (Bigi et al., 2013; Gartrell et al., 2004; Faulkner et al., 2010). In particular the intersection of two faults has been found to be particularly effective, as low stress regimes are needed to cause dilation and fault locking processes can help maintain the pathways open (Gartrell et al., 2004).

The hydraulic characteristics of a fault can also vary in time, as a function of the many processes that can open or seal a structural feature. For example there is a clear, even if scattered, relationship between throw and fault width, gouge development, smear thinning, cataclasis, and the relative proportion between FC and DZ (Bense et al., 2013). This increased fault maturity tends, on average, to focus strain within the fault core, often creating a progressively more effective barrier. If, instead of this strain softening, fault evolution results in the formation of progressively more rigid FC rocks, subsequent strain hardening could widen the DZ and increase fault-parallel permeability (Gray et al., 2005). Mineral precipitation is also a time-dependent process that can affect both fault permeability and fault strength in complex feed-back processes, particularly in carbonate rocks due to the higher mineral solubility (Polak et al., 2004). Increased DZ permeability may increase fluid flow, but secondary mineral precipitation of calcite or quartz can eventually seal those pathways and cause fluid pressures to increase. If minerals like calcite or quartz are precipitated, the resultant veins can contribute to a partial recovery of rock strength and make it more prone to further brittle deformation, whereas the formation of phyllosilicates (typically involving the breakdown of feldspars via hydrothermal reactions) can weaken a slip surface (Bense et al., 2013). In this sense one also has to consider the seismic cycle, whereby stress build-up, the formation of dilational microcracks, subsequent failure, and inter-seismic sealing all affect (and are all affected by) mineral precipitation, fluid flow, and permeability changes (Uehara and Shimamoto, 2004; Gartrell et al., 2004; Tenthorey and Fitz Gerald, 2006; Sibson, 2000).

Taken together it is clear that in a natural 3D system, consisting of different lithologies at different depths exposed to different stress fields, the migration of fluids along faults can be highly complex and spatially variable, with the potential for channelled flow, accumulation in capped porous units (as in stacked oil reservoirs), fault cross-flow, lateral migration along strike (where it may exploit a barrier weakness) or complete sealing (Faulkner et al., 2010). Permeability along a fault will be highly heterogeneous and anisotropic, with the potential that fluid migration, especially for gas, will occur along a limited number of fractures (Bense et al., 2013; Faulkner et al., 2010). This complexity makes modelling efforts of real-world situations highly challenging (Dockrill and Shipton, 2010).

2.3.2.2. Boreholes

Both active and abandoned wells could represent migration pathways for CO_2 from CCS sites because they form a direct connection between the engineered CO_2 storage reservoir and the surface, and because they are composed of man-made materials that may corrode over long periods of time (Figure 8). Not all boreholes are created using the same techniques, however, and thus newer wells are generally more secure than older ones while horizontal wells may be safer than vertical ones. Leakage along these pathways can be subdivided into 2 general categories based on the mechanism and potential rate of flow: i) blowouts during injection; and ii) long-term, slower leakage along wells. Blowouts (i.e., uncontrolled flow of well and/or formation fluids from the borehole to the surface) are extreme and obvious emergencies that will require immediate remediation, and thus this type of leakage would not be studied using the techniques described in this research. As such, only the process of slow leakage will be discussed below.

The various potential well leakage pathways / mechanisms can again be divided into two main groups: i) CO_2-induced chemical reactions that compromise the integrity of the cement/casing barrier; and ii) direct migration of fluids along existing fractures in the cement or gaps at the cement-casing or cement-rock interfaces. Clearly these two processes, the first diffusion controlled and the second advection controlled, are closely linked, as cement corrosion could put

CO_2 fluids in contact with an open pathway or flow along an open pathway will result in chemical alteration of the fracture walls. Research into well-bore leakage processes have been recently reviewed by both Choi et al. (2013) and Rochelle and Milodowski (2013).

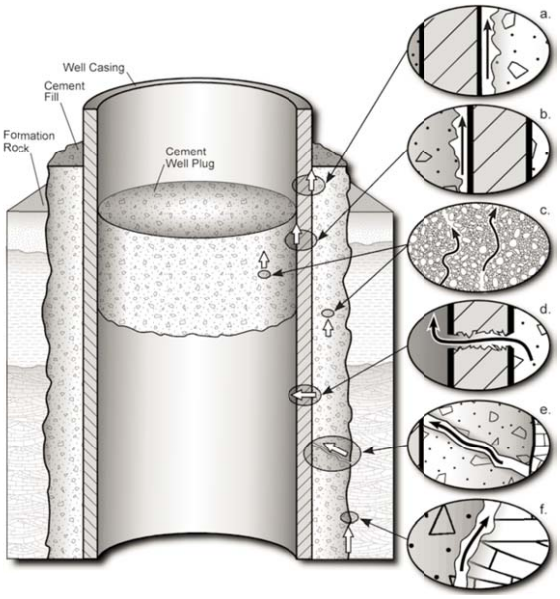


Figure 8. Schematic drawing of possible leakage pathways through an abandoned well (Gasda et al., 2004): a) between casing and cement; b) between cement plug and casing; c) through the cement pore space as a result of cement degradation; d) through casing as a result of corrosion; e) through fractures in cement; and f) between cement and rock

Kutchko and co-workers have conducted a number of studies into diffusion-controlled static experimental systems. These researchers defined three different alteration zones within cement samples exposed to CO_2-saturated water under different pressure and temperature conditions (Kutchko et al., 2007): 1) an inner calcium leaching zone (decreased Ca and $Ca(OH)_2$, and slightly increased porosity); 2) a carbonation zone (increased calcite content, and decreased porosity); and 3) an outer amorphous zone (highly depleted in Ca, increase in amorphous silica, significant increase in porosity). The migration of these zones into the cement appears to be controlled by the pH gradient (>12 in cement porewater, and potentially <3 in a CO_2-saturated brine), with reservoir-representative conditions yielding more well-defined zonation and less penetration compared to ground-surface conditions. Kutchko et al. (2008) defined non-linear penetration rates of the alteration fronts (initially rapid but which slowed with time) and an estimated penetration depth of 1.00 ± 0.07 mm for a 30 year period for cement in contact with CO_2 saturated brine, as well as a linear rate of 1.68 ± 0.24 mm over the same time for cement in contact with supercritical CO_2. Finally, Kutchko et al. (2009) examined the effect of cement composition on alteration, showing that a flyash - Portland cement mix resulted in a more rapid, non-zoned alteration which had, however, less impact on physical properties like permeability. Such results indicate that direct alteration of cement in contact with static supercritical CO_2 or CO_2-saturated brine is a slow process that will likely not have a significant detrimental effect on the barrier capabilities of a well that has been cemented properly and is of good quality.

Other experiments have looked instead at flowing fluids, showing that fluid flow along gaps that may exist at the cement-casing or cement-rock interfaces has the potential to have a much greater impact than a strictly diffusion controlled system. For example, Duguid et al. (2005) conducted flow-through experiments for 31 days at different temperatures and pH values, at a flow rate of

about 9 ml/minute. These authors observed very rapid, linear alteration which, for the carbonation layer, ranged from 300 mm over 30 years in the 20°C / pH 3.7 experiment up to almost 1000 mm in the 50°C / pH 2.4 experiment. The enormous difference compared to the diffusion controlled systems is due to constant elevated ionic and pH gradients causing the system to always be far from equilibrium. Flow-through experiments conducted by Carey et al. (2010) mimicked the narrow aperture gaps that may occur at the casing-cement interface by embedding standard-grade carbon steel with grooves in a 6 cm long cement plug. A 50:50 mix of supercritical CO_2 and NaCl brine was passed through limestone (to simulate equilibrium in a carbonate reservoir) prior to passing through the steel-cement plug at 40°C and 14MPa for about 16 days at a flow rate of 20 ml/h for the first 294 h and 10 ml/h for the remaining 120 h. At the end of the experiment the steel was corroded to a depth of about 25-30 μm; in some locations scale formation along the steel surface appears to protect it from further corrosion, however at other locations it was still exposed. Instead, the cement was carbonated to a depth of 50 to 250 μm, with little evidence of erosion. Iron and calcium carbonate precipitates formed in all of the gaps, showing the potential importance of self-sealing processes along gaps and fractures (Carey et al., 2007; Viswanathan et al., 2008) and the chemical interaction of products from cement and steel alteration. Bachu and Bennion (2009) stress, however, that because well-related leakage is likely to be 1D vertical, the occurrence of a very well-sealed interval will impede fluid flow and reduce the effective permeability along the entire length of the well, even if some intervals experience increased corrosion and dissolution.

The analysis of core from wells that have been exposed to CO_2 for up to 30 years support the laboratory findings described above. For example Carey et al. (2007) show reaction front penetration rates in samples collected about 3 m above the reservoir-caprock contact at a CO_2-EOR field to be similar to those predicted by Kutchko et al. (2008), and evidence of interface flow with the precipitation of 0.1 to 0.3 cm thick layer of self-sealing carbonate minerals at the casing-cement contact (but no sign of casing corrosion itself). A more recent study (Crow et al., 2010) sampled casing, cement and formation rock from a 30 year old well in a natural CO_2 production reservoir in the Dakota Sandstone in central USA. In this study there was evidence of diffuse carbonation but cement interfaces with the casing and formation rocks were tight with no significant calcite deposition and all casing samples were almost in pristine conditions.

While the slow leakage mechanism will result in much smaller flow rates than that which could occur during a blowout, it would be more difficult to detect and thus could potentially release a significant volume of CO_2 to the atmosphere (or an aquifer) if the leak is not recognised, located, and remediated within a reasonable length of time. For this reason near-surface gas geochemistry methods can give an important contribution to the monitoring of such sites (e.g., Beaubien et al., 2013; Jones et al., 2011; Loizzo et al., 2011).

2.3.3. Shallow pathways

For simplicity, the term "shallow" is used here to describe that depth interval which includes un-consolidated sediments and the soil, typically consisting of the upper 100 m. The discussion below has been divided into two main settings: i) free-phase CO_2 gas beneath the water table caused by upwardly migrating deep-origin gas; and ii) CO_2 gas movement in the unsaturated soil horizon. This subdivision has been chosen to highlight some processes that can affect CO_2 leakage at the ground surface (both in terms of location and rate), as well as any impact it may have on groundwater quality or ecosystem health. In addition, these processes can also influence the capabilities of various methods to locate and quantify leakage.

2.3.3.1. Free-phase gas in sediments below the water table

One issue linked to the leakage of deep-origin, geogenic or anthropogenic CO_2 that has not received much attention in the literature is that of the entry of free-phase gas into shallow sediments from faults or leaking boreholes and the processes that control its subsequent lateral and vertical migration.

For example, if faulted bedrock is overlain by water-borne sediments deposited after the last fault movement (i.e., the sediments themselves are not faulted), what is the behaviour and migration pathway of the gas after it exits the bedrock fault? The answer to this question is likely linked to the relative timing and duration of leakage and sedimentation. If leakage was continuous during the entire sedimentation process it is possible that bubble-induced turbulence would have maintained a constantly open vertical pathway through the newly forming sediments. If, instead, the leak is intermittent, or has newly formed due to such mechanisms as bleeding of over-pressurised reservoir fluids (Sibson, 2000) or the sealing of one leak point and opening of another (Dockrill and Shipton, 2010), the leaking gas will encounter stratified sediments (below the water table) with a classic anisotropic permeability field (i.e., Kh>>Kv). If a permeable unit is first encountered, like a lacustrine sand or a fluvial gravel, the gas could accumulate within it and migrate laterally as a bubble at the base of an overlying impermeable strata. The groundwater would begin to dissolve a portion of this CO_2, transporting it laterally in the dissolved phase along the groundwater flow direction. Over time this aquifer may become locally saturated with respect to CO_2. If leakage rates are high relative to groundwater flow rates and recharge, dissolution would not be sufficient to balance CO_2 input and thus the bubble would grow. If the boundary conditions of the permeable unit are open, groundwater will be displaced with no significant pressure build-up and the gas bubble will expand until a upward pathway is encountered. For example, if the contact is quasi horizontal the bubble may expand until a permeable fracture is encountered, or if the contact is dipping the bubble could migrate upwards and leak towards the atmosphere if the unit outcrops on surface or sub-crops in the unsaturated zone. Instead if the boundary conditions are closed, the continual addition of gas could increase over-pressuring within the aquifer; this may lead to fracturing if the unit is already subjected to a tectonic stress regime or to localised liquefaction. If one of these processes occurs, the leak may be able to propagate itself.

For the case of leakage along a borehole, much of the processes are the same except for the issue of leakage during sedimentation. If a borehole is permeable only to a certain level in the shallow stratigraphy, the leaking CO_2 may migrate vertically to that point and then laterally if it encounters a sufficiently permeable unit. Once in the permeable strata, the migration, accumulation, and leakage processes described above would be similar.

Implicit in the above discussion of hypothetical leakage into and through recent, non-faulted overlying sediments is that the eventual release point at ground surface may not coincide vertically with the source (fault or leaking borehole), but rather may be offset due to the movement towards, or creation of, different leakage pathways in the shallow sediments. Another implication relates to the footprint of a potential CO_2 source gas, as the gas bubble within an aquifer (and the associated CO_2-saturated groundwater) could cover a much larger area than the original, spatially limited "point" source. This gas-charged aquifer could then feed multiple leakage areas if different pathways are intersected, or could feed wider-scale, lower-level diffusive flux given sufficient time and suitable geological conditions.

2.3.3.2. Gas movement and distribution in the unsaturated zone

Various processes and mechanisms can influence the spatial distribution of biogenic and geogenic gas in the vadose zone, both in terms of total concentrations and isotopic values. The level of impact that any given process will have on eventual distribution will depend on the balance and inter-play between various site-specific characteristics, such as concentration and pressure gradients, soil permeability and organic matter content, source isotopic composition, etc.

Biogenic CO_2 concentrations typically increase with depth in normal agricultural or forest soils (Hamada and Tanaka, 2001; Risk et al., 2002b; DeSutter et al., 2008; Maier et al., 2010), due to the balance between the gradients of in situ biogenic CO_2 production and gas diffusivity (Amundson et al., 1998; Risk et al., 2002b; Hashimoto and Komatsu, 2006). Although the vast majority of biogenic CO_2 production (60 to 90%) occurs in the upper 5 to 20 cm of the soil column (Gaudinski et al., 2000; Risk et al., 2002b; Elberling, 2003; Hirano et al., 2003; Drewitt et al., 2005; Jassal et al., 2005) due to more labile carbon and the presence of roots (Gaudinski et al., 2000; Kuzyakov, 2006; Phillips et al., 2012), concentrations increase with depth due to the balance between this production curve and a decrease in diffusivity with depth caused by compaction and increased water content. In addition, infiltrating meteoric water can dissolve and transport shallow biogenic CO_2 deeper in the soil column.

The $\delta^{13}C$ of biogenic soil CO_2 is a complex function of: i) the $\delta^{13}C$ of the living plants or decomposing organic matter; ii) the difference in the diffusion coefficient of ^{12}C and ^{13}C; and iii) the rate at which CO_2 is produced by biological activity in the soil (Amundson and Davidson, 1990). In most cases, the $\delta^{13}C$ decreases from atmospheric value (about -7.5‰) at the soil surface to more negative values, closer to that of decomposing organic matter with increasing soil depth. However, because of diffusive fractionation, the soil gas CO_2 will be enriched by up to 4.4‰ while the CO_2 fluxes at the soil surface, assuming steady state, should have an isotopic signature representative of the source biological production (Cerling et al., 1991; Davidson, 1995; Admundson et al., 1998).

In addition to the effects that water content (i.e., precipitation) can have on gas permeability and thus on CO_2 flux to the atmosphere, as already described in Section 2.2.1, other atmospheric related parameters that can affect the rate and transfer of CO_2 to the atmosphere include wind via the Bernoulli effect (Lewicki et al., 2007) and atmospheric pressure via "barometric pumping"; these processes can affect the pressure gradient and induce transient advection.

For environments where deep CO_2 is leaking, advection and concentration gradients have a significant influence on distribution. In a series of detailed soil gas and gas flux surveys across various gas vents at Latera, Annunziatellis et al. (2008) and Beaubien et al. (2008) highlighted the complex interplay between processes within the core of a gas leak, where advection dominates, and those within the surrounding halo, where diffusion dominates. Both studies show CO_2 flux values in the centre of the gas vent of up to 3000 g m^{-2} d^{-1} and soil gas concentrations at a depth of about 70 cm that approach 100%; this high CO_2 concentration results in essentially anoxic soil conditions in which reduced gas species like CH_4 and H_2S are stable and where sulphides are precipitated. Outside of this core, over a distance of less than a metre, flux values drop rapidly to below 500 g m^{-2} d^{-1} (even though soil gas CO_2 concentrations are still high), CH_4 and H_2S values decrease to background, He concentrations also decrease to background values, and O_2 levels slowly start to increase. Moving further away, soil gas CO_2 values gently decrease while O_2 and N_2 values increase in a constant ratio that indicates mixing between leaking CO_2 and atmospheric gases.

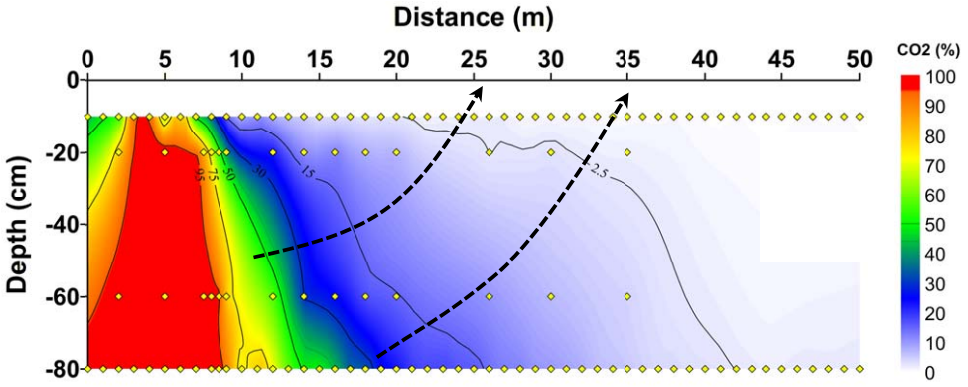


Figure 9. Cross-section of soil gas CO_2 concentration values through a gas vent at Latera (data re-processed from Beaubien et al., 2008). The red values, exceeding 95%, are in the core of the gas vent where advective flux dominates. Laterally (beyond 8 m) the lower values in the halo around the vent contribute to diffusive flux. Dashed lines indicate diffusion path perpendicular to concentration gradient.

Processes in the vent core are almost certainly advection controlled; other studies have shown in similar conditions that a slight vertical pressure gradient occurs within such emission points (Natale et al., 2000) and that isotopic fractionation does not occur in a fully advective flow regime (Camarda et al., 2007). Below a threshold of about 500 g m^{-2} d^{-1} at the Latera site the CO_2 flux likely becomes diffusion controlled and the back-diffusion of atmospheric O_2 is able to oxidise, via biochemical reactions, trace gases like CH_4 and H_2S. This is illustrated by re-elaborating the multiple depth soil gas samples taken by Beaubien et al. (2008) as a contoured cross-section, rather than the original scatter plots (Figure 9). This plot shows that high flux values within the gas vent core correspond with constant CO_2 concentrations with depth (similar to that observed by Camarda et al., 2007). In contrast, outside of this zone the soil gas CO_2 concentration contours occur at progressively deeper levels moving further away from the source, resulting in an ever decreasing concentration gradient and hence lower, diffusion-controlled, deep-origin flux rates on surface. This distribution implies the lateral migration of CO_2 in the soil via density effects, an idea supported by the observation that He, a much lighter, more mobile molecule, is restricted to the high advective flux interval (Beaubien et al., 2008). Because He initially migrates from depth with CO_2, and because He is non-reactive chemically and biologically, it is assumed that the observed difference in lateral distribution is due to differences in diffusion coefficients, with the heavier CO_2 "sinking" and the lighter He "rising" towards the surface, resulting in a form of mass fractionation. A similar relationship has also been observed at other sites (e.g. Gal et al., 2011). Finally, Figure 9 also illustrates why 1D models are not appropriate in such settings, as the flow pathways defined based on movement perpendicular to the iso-concentration lines are not vertical.

2.3.4. Release to the atmosphere

Once the migrating gas is released from the soil surface it will be mixed and diluted in the atmospheric air, primarily by turbulent wind motion. This mixing will be transient (as a function of the changing wind conditions) and spatially variable (due to the influence of topography and vegetative cover density / height). The existence of a thin boundary layer in the first few centimetres above the soil, where wind speed decreases to almost zero due to friction (Figure 10), and the greater density of pure CO_2 relative to atmospheric air, can result in accumulation of this gas and the creation of anomalous concentrations immediately above gas leakage points.

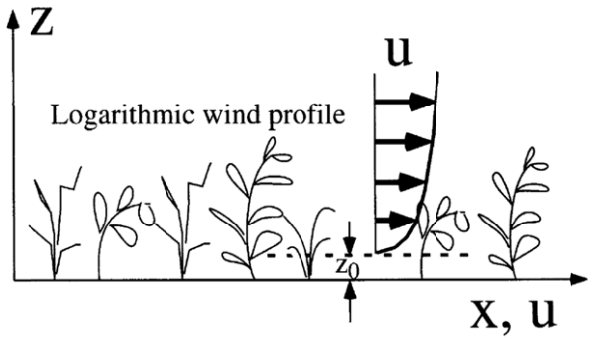


Figure 10. Schematic diagram showing the wind strength distribution near the ground surface, illustrating how the boundary layer (Z_0) has essentially zero wind mixing and thus can represent a zone of accumulation for leaking CO_2 (from Oldenburg and Unger, 2004).

This potential for accumulation was illustrated by Annunziatellis et al. (2008b) in experiments conducted above gas vents and background sites within the Latera caldera. In this work the inlet tube of a pumped, infrared CO_2 detector was placed at different heights above the ground surface and allowed to collect data at a 1 Hz frequency for 2-3 minutes. Whereas the values at all measured heights were very similar above the background site (Figure 11c), a clear enrichment was observed in the near ground surface measurements above the two studied gas vents (Figure 11a, b), despite the very short grass and light breeze at the time of the observations. Whereas anomalous values were clearly defined to a height of 20 cm above the ground surface at the high-flux vent (Figure 11a), only the ground surface sampling showed clearly anomalous results above the lower flux vent (Figure 11b).

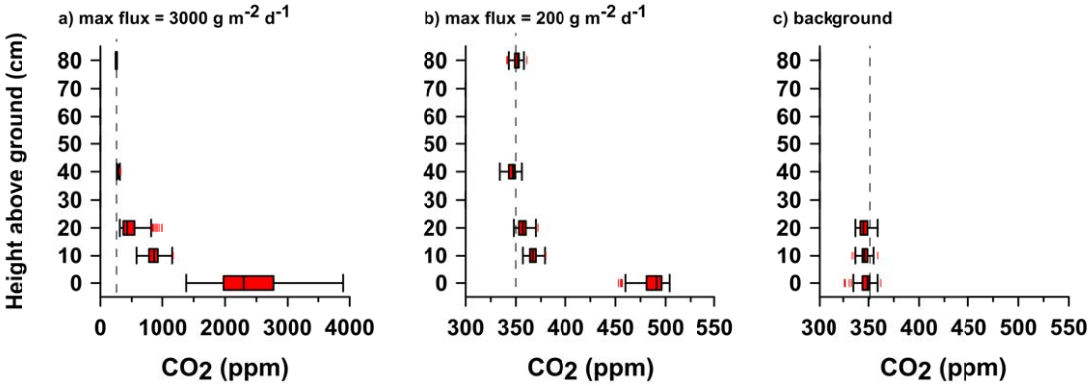


Figure 11. Box and whisker plot showing the statistical distribution and range of CO_2 concentration values observed at discrete heights above large (a) and small (b) gas vents and at a site where no leakage was observed (c) using an infrared CO_2 analyser deployed at each height for about 2 minutes at a sampling frequency of 1 Hz (Annunziatellis et al., 2008b).

2.4. Surface leakage characterisation

Two main factors will influence the success of soil gas and gas flux surveys for studying the leakage to the atmosphere of geogenic or anthropogenic CO_2 originating from deep natural or man-made geological reservoirs. First, these methods must locate the leak and define its physical extent. This will be a function of the size, style (i.e. point vs diffuse) and magnitude (subtle or obvious anomalies) of the leak combined with the chosen sampling density. Second these

methods must be able to quantify this leakage and accurately separate it from baseline flux originating from biogenic processes. As natural, near-surface, biologically-produced CO_2 flux can be both significant and highly variable (both spatially, as a function of soil type and underlying geology, topography, land-use, etc., and temporally, based on diurnal and seasonal effects of rainfall, temperature, etc.), great care must be taken to correctly subtract baseline fluxes so as to ensure that calculated leakage rates are not grossly over- or under-estimated. Although there is clearly overlap between these two arguments, they are separated below for convenience.

The discussion that follows focuses on near surface gas geochemistry techniques because of the scope of the thesis and the fact that these are a direct, precise measurement of the parameter.

2.4.1. Finding Leaks

The problem of locating a specific leak is more closely related to CCS, because the goal of CCS monitoring is to ensure site safety and integrity. In the unlikely event of a leak at surface from a geological CO_2 storage reservoir, emission will be new and probably localised, and surveying will be conducted to ensure rapid discovery to minimise any impacts and plan remediation. Natural, geological leakage from geothermal or volcanic areas will, on the other hand, tend to occur from long-lived, multiple gas release points, with research focused more on quantifying total leakage rates over larger areas. That said, these latter studies may benefit from the leakage location methods developed for CCS because leakage quantification precision will depend on discovering (and delineating) as many leakage areas as possible.

Because both soil gas and gas flux surveys are point measurements, the issue has been raised regarding the number of samples required to locate a leak (or leaks) of a given dimension in an area that could potentially be on the order of 1 to 100 km^2. To illustrate this Oldenburg et al. (2003) present a statistical analysis that relates the probability (P) of finding a leak based on the size of the leak (x), the size of the total survey area (A), and the number of randomly selected sample points (n), assuming that the method used has 100% capability to distinguish a gas leak anomaly:

Eqn 12

$$\mathbf{P = 1 - \left[1 - \frac{x}{A}\right]^n}$$

Using this formula, the number of purely randomly collected samples that will be needed to locate at least one gas leak point at the 95% confidence level will be 30 at x/A = 0.1, 300 at x/A = 0.01, and 3000 at x/A = 0.001. As discussed in section 5.3.1, different results can be obtained using different sampling strategies other than purely random.

Another approach described by Oldenburg et al. (2003) calculates the number of samples (n) required to attain a required confidence level. For example, a low confidence level (e.g. 0.1) would be set in the case where there is a low possibility of a leak, resulting in n = 7, 70, 700, and 7000 for x/A = 0.1, 0.01, 0.001, and 0.0001. In contrast a high confidence level (e.g. 0.9) would be required in the case where other information indicated that a leak likely exists, resulting in n = 50, 500, 5000, and 50,000 for the same the same range of x/A values.

These calculations assume that the method is 100% capable of detecting a CO_2 leak anomaly. In reality, however, both false positives and negatives will likely occur due to the statistical distribution of both the background and anomaly populations (i.e., overlap between the two populations). Parameters can be included in the calculations that estimate probability

distributions (based on baseline studies and leakage modelling), and a threshold can be set above which a value is considered anomalous. Monte Carlo simulations that take into account these parameters, as well as x/A, can be used to estimate the number of measurements that will be required to determine with a desired confidence level whether a gas anomaly exists in the sampling area. Oldenburg et al. (2003) state that these examples highlight the importance of collecting site specific geological information to delineate the most probable gas leakage locations to minimise A, which will maximise the potential for successful location and quantification of a leak while at the same time minimising the number of samples and specialised analyses (i.e. costs).

To address the issue of capability to recognise a leak during a survey, various additional geochemical methods can be used to help distinguish a biogenic anomaly from a leak anomaly. As discussed above the carbon stable isotopic signature of CO_2 can give an indication of origin, however only if there is a significant difference between the deep leaking gas and the shallow biogenic gas. For instance, $\delta^{13}C$ can be very useful to distinguish biogenic CO_2 (typically between -8 to -35‰) from geogenic CO_2 (often between -10 to +2‰) in natural diffuse degassing studies (e.g. Chiodini et al., 2008), however the similar values of biogenic (particularly of the C3 photosynthetic pathway) and anthropogenic CO_2 (originating from burning coal or petroleum products) often precludes its use in CCS monitoring (e.g. Beaubien et al., 2013). Examples where stable isotopes were successfully used to map leakage from controlled release experiments, in which injected and background biogenic CO_2 had significantly different $\delta^{13}C$ signatures, include Krevor et al. (2010), Moni and Rasse (2014), and Jones et al. (2014). Another option is the radiogenic carbon isotope, ^{14}C, of CO_2. This isotope is a measure of age, with higher values indicating a more modern carbon source; for example $\Delta^{14}C_{CO2}$ of atmospheric air is around 70‰ and CO_2 of a typical forest soil is about 128‰, whereas geogenic or anthropogenic CO_2 is basically free of ^{14}C (Oldenburg et al., 2003). Other tracer gases that can be associated with deep-origin gas (and absent or very low in shallow gases) have also been proposed, such as noble gas ratios like Ar/Kr and He/Kr (Rouchon et al., 2010) and noble gas isotopes of species like Xe (Nimz and Hudson, 2005) for geogenic gas, or man-made gases added to anthropogenic CO_2 prior to injection such as perfluorocarbons and sulphur hexafluoride (Wells et al., 2007; Strazisar et al., 2009).

A number of researchers have proposed different sampling soil gas / flux strategies to improve leak detection success rates. For example, Cortis et al. (2008) describe a dynamic sampling campaign of soil gas or gas flux measurements directed via an artificial neural network (ANN) model coupled with particle swarm optimisation (PSO). The ANN defines a regression correlation between the baseline CO_2 point measurements conducted prior to injection and various easily measured system properties (e.g. topography or vegetation), with points lying outside this regression being defined as anomalous. The PSO is then used to dynamically manage a sampling campaign, with subsequent sampling points (direction and distance) for multiple technicians in the field being directed by their current direction and by minimising the regression coefficient between their previous measured values and the ANN model at both the local (individual) and global (group or "swarm") level. Computer testing of this approach was conducted on simulated data for a hypothetical site having variable topography, vegetation, and an average background CO_2 flux of about 10 g m^{-2} d^{-1}. On this baseline distribution a 20 times CO_2 flux anomaly was superimposed in one area within the grid. Multiple PSO simulations were conducted with different random initialisations, with the majority being capable of finding the anomaly. Unfortunately, however, the authors do not state the average number of iterations (i.e. sampling points) that were required for the various simulations.

Another approach involves multiple measurement campaigns combined with spatial and temporal filters to both locate and subsequently quantify a leakage (Lewicki et al., 2005a; Lewicki et al., 2006). This approach is based on the assumption that: i) shallow background processes like soil respiration are generally spatially heterogeneous and controlled by meteorological and biological processes that operate on diurnal or seasonal time scales; ii) leakage anomalies are relatively coherent in space and constant in time; and iii) both baseline and leakage signals are modified by near-surface processes on predictable time scales. To test the idea that these differences can be exploited to locate a leak, the authors conducted multiple simulations on synthetic data consisting of a log-normal background CO_2 flux distribution and local leakage values created as a scaled Gaussian distribution. Results indicate that the approach has potential to find a leak given favourable conditions.

Finally, ground surface CO_2 concentration or isotope mapping, based on the concept of near surface gas accumulation described above in Section 2.3.4, has been proposed as a potential leakage mapping technique. In this approach measurements are made continuously at a fixed height above the ground surface while constantly moving over the area of interest. By keeping track of analysis location via coupled GPS measurements, results can be mapped and leakage-associated anomalies can be delineated. The advantage of this approach is that it is rapid and continuous along the surveyed lines, however it may be limited by a lower sensitivity compared to point flux and / or soil gas measurements. Published examples of this approach include the use of cavity ring-down spectroscopy (WS-CRDS) for the analysis of $\delta^{13}C$-CO_2 at a height of about 9 cm above the ground surface (Krevor et al., 2010), CO_2 concentration mapping at a height of between 30 to 60 cm using a mobile open path infrared laser (Jones et al., 2009), and CO_2 concentration just at ground surface using an infrared sensor (Jones et al., 2009; Annunziatellis et al., 2008). Jones et al. (2009) show how a gas vent having a flux rate of 200 g m^{-2} d^{-1} was clearly visible with the measurements made at ground level (Figure 12) but that no anomaly was observed for this vent at the subsequent height of 10 cm (for the given sensitivity and response time of the sensor system used). These results indicate that continuous ground surface CO_2 measurements may be a viable method for reconnaissance leakage mapping given sufficiently high enough flux rates above background (probably >100 g m^{-2} d^{-1}) and good ground conditions.

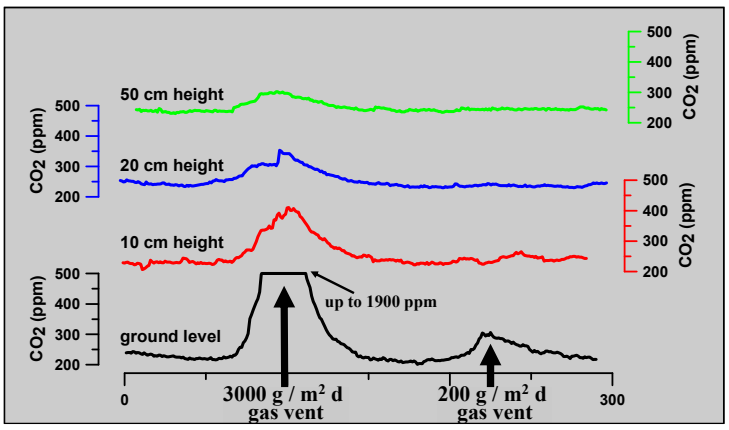


Figure 12. Plot showing the CO_2 concentrations at four different heights along a transect crossing two gas vents (Jones et al., 2009); the surveys were performed at the pace of a slow walk. Note that only the survey performed at ground surface was able to define the second, lower-flux gas vent.

Other, more large-scale methods that have been suggested for leakage detection, like Eddy Covariance (Leuning et al., 2008; Lewicki et al., 2012) and remote sensing (Bateson et al., 2008; Verkerke et al., 2014) are beyond the scope of the present study.

2.4.2. Quantifying Leaks

Leakage quantification has been applied to both CCS and natural leakage studies, with primary challenges being the separation of deep-origin flux from shallow biogenic flux, and the size of the leakage area with respect to the sample spacing and survey area (e.g. interpolation between measurement points). Four main approaches have been discussed in the literature: i) single survey of flux measurements; ii) multiple flux surveys; iii) monitoring of soil gas tracers; iv) combined flux and soil gas measurements. Approach (i), which can be further subdivided based on data processing methodologies (as discussed below), is the most commonly used approach and is the method which is applied in the present research.

2.4.2.1. Single survey of flux measurements

This method involves a single CO_2 flux survey over the area of interest, with sample spacing being chosen based on target size, total area size (i.e. time and money available to conduct the survey), logistics, experience, etc. Based on how the resultant data is treated (i.e., for calculating total flux and for estimating the biogenic and deep flux contributions) this approach can be further sub-divided into "dataset averaging", "population averaging", and "area calculating" (with some overlap between the techniques). Note that because flux data is typically log-normally distributed, appropriate statistical methods must be used.

For *dataset averaging*, the entire dataset is processed to calculate the simple arithmetic mean (which is valid for normal distributions) (Bergfeld et al., 2001; Lewicki et al., 2005; Bergfeld et al., 2006; Evans et al., 2009), the mean flux using a minimum variance unbiased estimator (MVUE) on log-transformed data (Bergfeld et al., 2006; Lewicki et al., 2005), or the declustered mean (Lewicki et al., 2007). These average values (in g m^{-2} d^{-1}) are then multiplied by the total survey surface area (m2) to yield a total flux (i.e. undifferentiated in terms of origin) for the study site (g d^{-1}). To separate biogenic from leaking flux, an estimate of the average or maximum biological flux rate is multiplied by the total surface area, and this value is subtracted from the total flux to estimate the portion of leaking flux. Note that this approach implicitly assumes that biological production continues at essentially the same rate whether deep leakage is present or not; this assumption has not been proven, as soil biological processes could be affected by both low (plant fertilization) and high (soil anoxia) CO_2 flux. The biogenic flux can be estimated based on a visual assessment of the lowest population in a log-normal cumulative distribution plot of all data (Bergfeld et al., 2001), on a sub-set of samples collected at the same time in a non-leaking area (Lewicki et al., 2007; Chiodini et al., 2007), or ignored completely if most sampling is conducted in an area of elevated flux and little vegetation (Evans et al., 2009). Note that all the referenced studies were conducted on natural sites except for Lewicki et al. (2007), which instead was performed at an experimental site where CO_2 was injected into the shallow subsurface via a horizontal borehole equipped with packers. Because the actual injection rate of this study was known, which is not possible on natural sites, these authors were able to show that the estimated steady-state leakage flux rates were very close to the actual injection rates of 0.1 t CO_2 d^{-1} (Lewicki et al., 2007) and 0.3 t CO_2 d^{-1} (Lewicki et al., 2010). This gives confidence in the method, provided that sample density is sufficient to capture spatial variability, although the small size and limited variables in this controlled experiment would be very difficult to duplicate in a more regional, natural-leakage study.

The *population averaging* approach, better known in the literature as Graphical Statistical Analysis (GSA) (Chiodini et al., 1998; Lewicki et al., 2005) involves the definition of more than

one population in a log-normal distribution using the graphical methods defined in Sinclair (1991). This approach is based on the fact that a log-normal distribution (i.e., a single population) will plot as a straight line on such a graph, whereas multiple populations will result in a graph that has straight intervals (one for each population) connected by inflection points (due to overlap between the populations). The mean and proportion are calculated for each population, the proportion is multiplied by the total survey area to define the surface area of that population, and then that number is multiplied by the mean flux for that population to calculate the total flux due to that population. In the simpler case where two populations are defined, one can assign the lower population to background biogenic flux and the upper population to deep leakage flux. In contrast to the previous, this approach separates the two mechanisms and thus all flux in a leakage area is assumed to be of deep origin. Note that GSA has also been used to define background populations to assist in the sequential Gaussian simulation approach described below (e.g. Fridriksson et al., 2006). It should be pointed out, however, that multiple background populations (e.g. Mazot et al., 2013), spatial/temporal variations in background distributions (Leon et al., 2014), and the broad overlap of background and leakage populations can complicate data interpretation.

Area calculating involves the use of contouring algorithms to calculate total flux rates, with the most common being kriging and the newly applied stochastic method called sequential Gaussian simulation (sGs).

Kriging uses a weighted linear combination of neighbouring observations to estimate a value at an unsampled location. Based on a minimised least squares approach it provides the "best" fit for the available dataset, however it does not take into consideration the statistical (histogram) distribution of the original data, nor does the interpolated dataset match the spatial (variogram) distribution of the original dataset. This approach tends to produce maps that are "smoothed", such that high and low values are under- and over-estimated, respectively. Two approaches have be used to quantify leakage using the kriging method. One is to calculate the arithmetic mean of the interpolated dataset and multiply it by the total area (Lewicki et al., 2005), similar to the procedure described above under dataset averaging. The second is to use the volume and area integration algorithms on the interpolated grids to calculate total flux, with area (m^2) of the survey being multiplied by the Z values of the grid (i.e., flux in g m^{-2} d^{-1} for each grid cell. (Gerlach et al., 2001; Annunziatellis et al., 2007); this is the approach applied in this study (see section 5.2.2.3). Subtraction of the biogenic background, as described above, can be applied to these approaches.

The newest approach is sequential Gaussian simulation (sGs), a stochastic method (Deutsch and Journel, 1998) first applied to flux quantification by Cardellini et al. (2003) and recently applied by numerous researchers (Lewicki et al., 2005; Fridriksson et al., 2006; Chiodini et al., 2007; Evans et al., 2009; Padrón et al., 2009; Bergfeld et al., 2012; Hernandez et al., 2012; Rissman et al., 2012). This stochastic method has kriging at its core, with the additional advantage that the resultant interpolated data respects the histogram and variogram of the original dataset. The method works as follows. First the typically skewed flux data is transformed into a normal distribution. A grid of a user defined cell size is created over the survey area and a random pathway is created which passes through each node on the grid. The first point is calculated using kriging, according to the semivariogram model for the original dataset, to estimate a mean and variance for that node based on the surrounding points, and then randomly selecting a value from that "conditional cumulative distribution" and applying it to that node. The new data point becomes part of the total dataset, such that each successive node calculation takes into account not only the original measured points but also all points calculated sequentially to that stage in

the random path. In this way the spatial and statistical distribution of the original dataset are preserved. After all nodes have been calculated the data is back transformed into the original data units, with the user setting the extrapolation model(s) for the upper and lower tails and maximum and minimum values allowed during this back transformation. Once complete, the algorithm starts again from step one, creating a new interpolated grid based on a new random starting point and pathway, guaranteeing a different but equi-probable grid based on the same input data. The software performs a user-defined number of such realisations, typically between 100 and 300. The final product is created by performing a linear arithmetic average of all simulations for each point, producing a final interpolated grid which is the average of all simulations. The total flux is calculated by multiplying the average value for each cell by the cell size and summing all values over the entire survey grid.

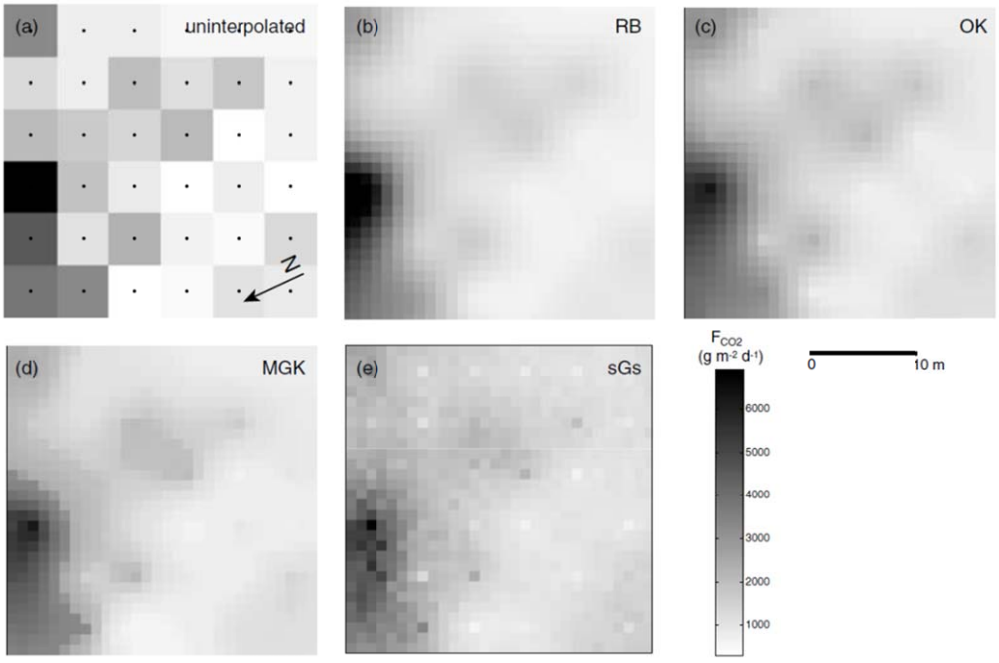


Figure 13. Raw, un-interpolated data (a) compared to results using different interpolation algorithms, including radial basis function (b), ordinary kriging (c), multi-gaussian kriging (d), and sequential Gaussian simulation (e) (Lewicki et al., 2005).

Although numerous studies have shown that the sGs approach yields total flux rates that are quite similar to that estimated using other approaches (Cardellini et al., 2003; Lewicki et al., 2005), the main advantages of sGs are that it gives a less smoothed spatial distribution map compared to other contouring algorithms (e.g. Figure 13), which may be more representative of the complex leakage that occurs in nature, and it gives an estimate of the error based on the range of values calculated for all realisations (although this is the error of the interpolation based on the available dataset, and not the error of the actual flux which is determined by how well the leakage points are located and delineated - see section 5.3.2). Biogenic baseline flux is typically calculated using GSA (Fridriksson et al., 2006) or a subset of samples that are clearly not affected by leakage phenomena (Chiodini et al., 2007). One recent approach for separating baseline from leakage flux, although total flux was not calculated, was given by Chiodini et al. (2008), where $\delta^{13}C$ of CO_2 in the flux chamber was used to define origin and to show how an intermediate population of mixed biogenic and leakage CO_2 lies between the two end members.

2.4.2.2. *Multiple flux surveys*

The multiple-campaign approach described above (Lewicki et al., 2005a; Lewicki et al., 2006), which automatically filters out the spatial and temporal variability of the baseline CO_2 flux field for a direct quantification of the leakage flux, has been proposed based on modelling results. In these articles, strategy success was evaluated based on the fraction mis-estimation of the calculated leakage rate versus the model-imposed synthetic leakage rate. These simulations showed that for a fixed A_S/A_T (leak versus total surface area) and F_S/F_B (leak versus baseline CO_2 flux), the error of the leakage rate estimate decreases non-linearly with increasing number of sampling campaigns, becoming relatively insensitive after a certain number depending on the used parameters. It was also found that leakage estimates improve with a large leakage area relative to the total area or a high flux rate compared to background, but not necessarily both together. This implies that judicious selection of sampling areas for CCS monitoring based on known geology and structure will decrease the total measurement area (i.e. A_T), thus increasing the potential for leak detection and quantification.

2.4.2.3. *Monitoring of soil gas tracers*

Perfluorocarbon tracers (PFTs) have many advantages for CO_2 leakage monitoring, including being soluble in CO_2, non-reactive and mobile, having very low natural concentrations, and being detectable at very low concentrations. Disadvantages include the fact that sampling points can be easily contaminated because of their high mobility and low concentrations, and the uncertainty regarding whether they migrate at the same velocity (and along the same pathways) as the host CO_2. PFT tracer experiments at the West Pearl Queen depleted oil formation in south-eastern New Mexico (Wells et al., 2007) has yielded the most successful results thus far. This study involved the injection of about 2,090 tonnes of CO_2 into the reservoir over a 53 day period, together with slugs of three different PFTs. The PFTs were monitored using capillary adsorption tubes (CATs) that were deployed for extended time periods in a number of shallow boreholes (max 1.5 m deep) arranged in a radial pattern up to 300 m away from the injection well. Results showed NW and SW aligned "hot spot" anomalies within 100 m of the injection well. The PFT data was used to quantify leakage from the site as follows. The monitoring matrix was divided into 9 equal sized areas. Within each section, background concentrations were subtracted from points with anomalous values, an average calculated, and this value was projected for the entire section. The values for all sections with observed anomalies were then summed and this PFT leakage rate was converted to a CO_2 leakage rate by multiplying by the injected CO_2/PFT ratio. With this procedure a total CO_2 leakage rate of 2.82×10^3 g CO_2 yr^{-1} was calculated for one of the three tracers, which corresponds to 0.014% of the total injected CO_2. Possible errors in the quantification calculations, as discussed by the authors, include the use of the average tracer concentration rather than a spatially distributed one, enhanced dilution of the tracer in the CO_2 over time during migration in the reservoir compared to the original injected ratio, and the possibility that "chromatographic" separation could have occurred as the tracer is far less water-soluble and reactive than the CO_2.

2.4.2.4. *Combined flux and soil gas measurements*

Klusman (2003) used both stable and unstable carbon isotopic data, combined with CO_2 flux and other measurements, to estimate the amount of CO_2 leaking at surface from the Rangely CO_2-EOR project in Colorado, USA. Approximately 23 million metric tonnes of inorganic geological

CO_2 were injected into the Weber Unit from 1986 to 2003 over the 78 km^2 area of the field. The geogenic CO_2 used has an isotopic signature that is significantly different from that of biogenic CO_2, with $\delta^{13}C$ values near –4‰ and ^{14}C (expressed as Percent Modern Carbon, PMC) near zero, thus rendering it a clearly recognisable tracer compared to the background CO_2. Work consisted of a limited number of initial measurement points (41 above the oil field and 16 in the control area) being sampled for soil gas along shallow vertical profiles at 30, 60, and 100 cm depths and triplicate CO_2 and CH_4 flux measurements. The results were then used to choose 5 locations to install deep multilevel boreholes (possible because of the 60m deep water table), with sampling intervals at 1, 2, 3, 5 m and end-of-hole depths (typically 7-8 m). Soil gas samples were analysed for CO_2, CH_4 and other alkanes, $\delta^{13}C$ in CO_2 and CH_4, and ^{14}C in CO_2. Through the combination of CO_2 flux measurements, $\delta^{13}C_{CO2}$ results from the flux chambers, modelling of CH_4 distributions in the deep boreholes (to define methanotrophic rate constants of CH_4 consumption and its eventual control on CH_4 and CO_2 concentrations), and the ratio of ancient vs modern carbon in the deep boreholes, Klusman (2003) estimated that less than 170 t CO_2 yr^{-1} is leaking from the reservoir. This value may over-estimate the actual amount of CO_2 leakage, as the high CH_4 values in the deep wells (from 1-10%) and the modelled oxidation rates in the soil may indicate that much of the observed CO_2 is the product of CH_4 oxidation rather than leakage of the injected CO_2.

Chapter 3. Latera Caldera description

The Latera caldera, located about 110 km to the NW of Rome, Italy (Figure 14, inset), is a NE-SW trending elliptical depression consisting of a relatively flat central agricultural plain surrounded by tree-covered, 100m high hills. The caldera hosts a water-dominated geothermal field that was studied intermittently from 1970 to 2000 to see if this energy source could be exploited. Despite construction of a 4.5 MW pilot power plant in the early 1980's, and the start of construction of a full-scale 30 MW plant in the early 1990's, exploration and exploitation plans were stopped in 2000 due to technical problems and environmental concerns.

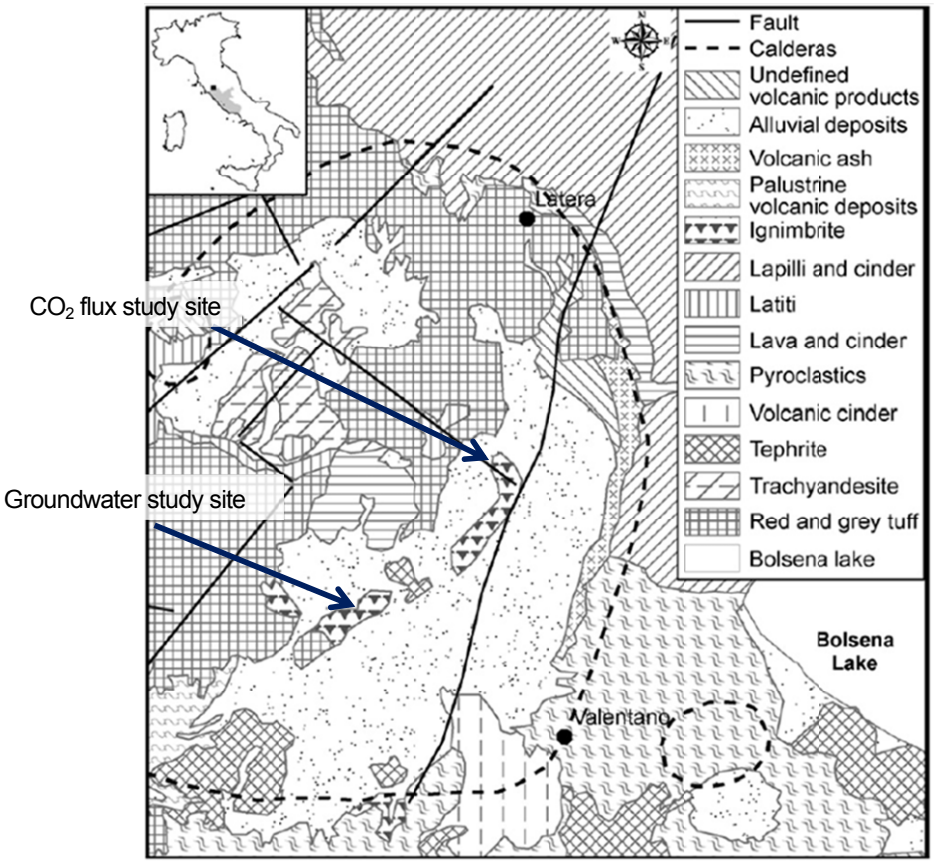


Figure 14. Geology map of the Latera caldera, showing the two locations studied in the present research.

This area was chosen for study due to the occurrence of numerous points where deep-origin CO_2 (produced via thermo-metamorphic reactions linked with the geothermal field) leaks to the surface via various fault and fracture systems (see Section 3.3). In addition, because there are different theories regarding the near surface leakage mechanisms and distribution, and thus the total amount of geogenic CO_2 leaking to the atmosphere, it was decided that this research could contribute to a more detailed understanding of these problems. Finally, the fact that the area has been studied extensively in the past permitted the selection of appropriate sites for both the gas flux and groundwater impact field studies.

3.1. Geology

The geology of the west-central Italian peninsula is the product of two main tectonic events, a compressive phase during the Eocene to Late Miocene and a subsequent extensional phase from the Late Miocene through the Quaternary (Carmignani and Kligfield, 1990).

The regional substratum of the study area consists of metamorphic phyllites and micaschists, onto which a series of "tectonic units" were placed during the initial compressive period via thrusts that trend primarily N-S. In the area of the Latera caldera, and throughout central Italy, these units are mainly formed by the Mesozoic "Tuscan nappe" and the Cretaceous to Eocene "Ligurian flysch". The Tuscan nappe consists primarily of carbonates and some siliciclastic successions; because of high secondary permeability in the carbonates, this unit forms a regional aquifer and locally hosts the geothermal reservoir beneath the Latera caldera. The overlying Ligurian flysch is made up of various units, including impermeable shales and siltstones which act as a regional aquitard and which locally form the caprock for the geothermal reservoir and associated CO_2.

The subsequent extensional regime, which was related to the opening of the Tyrrhenian Sea, resulted in the formation of a series of NW-SE to N-S trending grabens and normal fault systems which dissected the previous compressional fold and thrust belt (Di Filippo et al., 1999; Funiciello and Parotto, 1978). The extensional phase also caused crustal thinning, resulting in significant volcanic activity during the Quaternary. The Vulsini volcanic district, consisting of the three main eruptive centres of Latera, Bolsena, and Montefiascone, was formed during this period, starting about 0.8 Ma ago (Evemden and Curtis, 1965; Nicoletti et al., 1979).

The Latera Volcanic complex (Figure 14) developed in the western part of the Vulsini district between 0.3-0.1 Ma (Palladino and Simei, 2006; Simei et al., 2006; Vezzoli et al., 1987). Although the Latera caldera has a relatively simple topography, its eruptive history is complex, with multiple depocentres and a close link to Quaternary extensional tectonics (Annunziatellis et al., 2008). The explosive nature of this volcano resulted in a polygenic collapse caldera which has been dated at 0.3 Ma (Palladino and Simei, 2006; Vezzoli et al., 1987), followed by recurrent Plinian activity 0.28-0.23 Ma which formed a central volcanic edifice. The subsequent Sovana eruption resulted in pyroclastic flow activity (around 0.19 Ma) and is related to the main collapse of the Latera caldera. The largest volume of volcanic products was then erupted during the Onano eruption (0.17 Ma) (Freda et al., 1990; Nappi, 1969; Vezzoli et al., 1987), whereas the final phases of activity were characterised by the migration of emission epicentre from east to the northwest, where the Pitigliano eruption took place at 0.16 Ma and generated the small Vepe caldera (Nappi et al., 1991).

Two magmatic series have been defined in this area, a potassium (K) and a high-potassium (HK) series. The HK rocks are strongly to mildly silica under-saturated, ranging from leucitites to phonolitic trachytes (Holm, 1982); the more mafic products are slightly porphyritic lavas, with phenocrysts of clinopyroxene, olivine and leucite, while the intermediates have no olivine and scarce plagioclase. The K series are nearly silica-saturated trachybasalts, latites and trachytes; porphyritic mafic lavas of this series contain olivine, clinopyroxene and minor plagiclase phenocrysts, while the pyroclastic products are dominantly trachytes with sanidine, plagioclase, biotite and titano-magnetite phenocrysts (Nicoletti et al., 1981).

Due to anomalously high heat flow beneath the caldera floor a number of geothermal exploration wells (up to 3000 m deep) were drilled during the 1970's. These boreholes provide valuable information regarding the geology, structure, and geochemistry of the caldera (Barberi et al., 1984). The geothermal reservoir is located within a structural high of Tuscan limestones (Barberi et al., 1984; Palladino and Simei, 2006), a feature which may represent a recumbent fold (Bertrami et al., 1984), an ancient caldera rim (Barberi et al., 1984), or a series of superimposed thrust sheets (Annunziatellis et al., 2008). This structural high is bounded laterally by sealed faults and thermo-metamorphosed carbonates, below by the metamorphic basement, and above by low permeability Ligurian flysch and volcanic pyroclastics. The reservoir limestones are heavily fractured and folded, and locally re-sealed with a composite hydrothermal assemblage (Cavarretta et al., 1985). Maximum permeability within the reservoir occurs along the major, NNE-SSW-trending axis of the structure (Sabatelli and Mannari, 1995).

The CO_2 that leaks at the surface (see Section 3.3) is likely the product of metamorphic alteration of the reservoir limestones, driven by the water-dominated system and the high heat flow of the geothermal reservoir (presently ranging from 170 to 230°C) linked to the original magma intrusion (Chiodini et al., 1995; Duchi et al., 1992; Minissale, 2004), although a portion may also come from a deeper mantle component (Chiodini et al., 2007). Because the geothermal system is water dominated the majority of gas samples collected from the deep wells are dissolved gases that exsolve at the lower pressures at surface. In general, most reservoir waters contain 3 to 6% "Non Condensable Gases" (NCG) (Sabatelli and Mannari, 1995), with well L2 having a representative composition of 98.35% CO_2, 0.05% CH_4, 1.22% H_2S, 0.4% N_2 and trace levels of H_2 (Bertrami et al., 1984). One well (L11) did intersect a gas-only reservoir having 98% CO_2 (Lombardi, 1993).

Normal faults related to edifice collapse, combined with thrusts, regional normal faults, and extensive fracturing associated with the volcanic eruptions themselves, became conduits for upwardly migrating hydrothermal fluids and CO_2. Secondary minerals observed in the deep geothermal cores from Latera show that the area has likely undergone three alteration phases, including a self-sealing episode. This self-sealing process has made much of the overlying flysch and volcanic units impervious, except where gas is locally escaping (see Section 3.3), and helped to partially isolate the underlying geothermal reservoir.

At the surface the central part of the Latera caldera is characterised by a NE-SW to N-S trending hydrothermally-altered area, with silica, alunite, and kaolinite formed in the vicinity of several gas vents and thermal / cold springs (Gianelli and Scandiffio, 1989; Lombardi and Mattias, 1987; Lombardi, 1993); at least 3 kaolinite quarries were once active in this area. These two directions have also been observed by Buonasorte et al. (1991) in a geomorphological study of the dominant morphological elements located throughout the Vulsini mountains, while a recent structural study in various caldera quarries yielded primary fracture directions of N10E and N45E (Annunziatellis et al., 2008).

Surface lithology throughout the caldera floor is dominated by alluvial sediments and fertile soils derived from erosion of the alkali-potassic units of the Latera volcanic complex (Figure 14). Soil sampling to a depth of 70 cm (Beaubien et al., 2008) indicates that the shallow soils are predominantly made up of clay sized particles. The major minerals found in this study were quartz, K-feldspar, and pyroxene (augite), with minor amounts of hematite and albite, and trace amounts of cristobalite, mica, and pyrite. The clay minerals were primarily halloysite with typically less than 20% illite, while a significant proportion of amorphous, non-crystalline material was also observed which is most likely volcanic glass. Low-permeability, unconsolidated volcanic sediments also represent the main lithologies encountered in 20m deep

boreholes drilled in the same field for a shallow gas injection experiment (Annunziatellis et al., 2006). This work did observe, however, a highly fractured, incompetent interval from about 6 to 9 m at one location in a borehole drilled near a gas leakage area.

3.2. Hydrogeology

The various hydrogeological units in this area basically follow the geological units described above, as the permeability and self-sealing capacities of the dominant lithologies in each genetic/structural unit controls groundwater flow through that horizon. Therefore in a simplified sense the hydrogeological units can be considered as follows: metamorphic basement – aquitard; Tuscan limestones – regional aquifer and geothermal reservoir; Ligurian flysch – regional aquitard; volcanic flows – local aquifer; volcanic pyroclastics – local aquitard; and post orogenic sediments – local aquitard or aquifer. As described earlier the local hydrogeology is greatly complicated by folds and faults (which allow for cross-unit flow, mixing, and complex flow patterns) as well as by changes in primary porosity and permeability. Geochemistry has defined three main water types in the above aquifers: a hot NaCl geothermal fluid in the confined reservoir within the Tuscan limestone; a Ca-SO_4-HCO_3 water in the regional Tuscan limestone aquifer; and a Ca-HCO_3 water in the very shallow leucite and K-feldspar bearing volcanics.

3.2.1. Geothermal fluids

The Latera geothermal reservoir is water dominated, with temperatures in the range of 210 to 230°C. In general the reservoir waters are Na-Cl type with high concentrations of HCO_3 and SO_4 (with the latter coming from either late stage magmatic volatiles like H_2S or SO_2, or from sedimentary sources). Most wells have high Li, B, Cs and As values and very high Cl/Br ratios (750-2000). The $\delta^{18}O$ values are quite positive with respect to the groundwaters in the area (on the order of $\delta^{18}O$ = -0.2‰). Most geothermal fluids are at equilibrium with alteration minerals that indicate interaction with igneous rocks, such as muscovite, microcline and Mg/Ca montmorillonite, while being under-saturated with such mineral phases as anhydrite and calcite (Gianelli and Scandiffio, 1989). Results show that almost all hydrothermal waters could have resulted from large scale mixing of two end-members, well L3D water and the Ca-SO_4-HCO_3 water of the regional aquifer. The origin of the end-member brine found in L3D is a matter of debate, however modelling indicates that inflow of hydrothermal or magmatic brine into the sedimentary rocks and cooling may be most likely (as the fluid becomes acidic and progressively depleted in SiO_2, Fe, SO_4, H_2S, Na, and K).

3.2.2. Springs and Groundwaters

A number of studies have sampled springs throughout the western Mount Vulsini area to better understand regional flow and water origin (e.g. Barberi et al. 1984; D'Amore et al., 1979; Checcucci et al., 1988). As stated, it is believed that there are two primary aquifers in the Latera region – a shallow one in the volcanics and a more-regional, deeper one in the fractured limestones of the Mesozoic Tuscan nappe. Geochemical data generally supports this interpretation, however results indicate the existence of 4 sub-types within this general classification. Types A and B are thought to be shallow waters while types C and D likely originate from the deep aquifer system.

Type A water is the most common type in the area, as it represents approximately 75% of the measured springs. These cold, HCO_3-Ca-Mg waters have low total dissolved solid (TDS) values of <700 ppm and are believed to circulate within, and be involved in the alteration of, the shallow volcanic units; this latter is shown by their low Na/K ratios, which indicate reaction with K-rich vitreous volcanics or dissolution of leucite (Gianelli and Scandiffio, 1989). They are under-saturated with respect to calcite, dolomite and anhydrite, and instead fall in the stability field for kaolinite. In addition Checcucci et al. (1988) state that their elemental composition is quite variable (due to different flow paths in the various volcanic lithologies), their dissolved CO_2 content shows a bimodal distribution (with the higher mode interpreted as deep CO_2 input) and higher SO_4 values are correlated with anomalous pCO_2 (possibly due to addition of H_2S with deep CO_2). Type B chemistry is similar to that of Type A, but is distinguished by higher SO_4 content (approaching gypsum solubility), a lower pH, generally higher temperatures and higher TDS values (500-2000 ppm); this water is thought to circulate in deeper volcanics. The origin of the sulphate in these waters is a matter of debate, as Bertrami et al. (1984) state that it arises from dissolution of sulphates in the volcanics whereas Checcucci et al. (1988) believe it to be the result of the addition of H_2S gas.

Types C and D have not been encountered in wells in the Latera caldera. Type C springs occur in the Viterbo, Canino and Pitigliano areas and likely originate from the deep regional limestone aquifer. Type C comprises thermal, Ca-SO_4-HCO_3 waters with high TDS values (2000-3000 ppm) that are saturated with respect to calcite, dolomite and albite and weakly oversaturated with respect to anhydrite. Regionally the SO_4 in these waters is believed to come from interaction with anhydrites in the Tuscan nappe (e.g. Minissale et al., 2000). However, as evaporites were not encountered in the deep Latera drill holes, Gianelli and Scandiffio (1989) suggest that the SO_4 originates from remobilisation of secondary anhydrite that was precipitated during a previous hydrothermal episode. Type D are deep, weakly-thermal Na-Cl waters with high TDS values (2000-7000 ppm) and Na - K values which are very close to equilibrium with albite and microcline at high temperatures. Occurring in the Tuscania and Canino areas, this is the only type which is at all similar to the geothermal reservoir waters described above, although they are located a great distance away. Although leakage from these reservoirs is one explanation, another is that these waters are fed by descending fossil seawater from the Neogene sediments (Duchi et al., 1992).

Isotope analyses of these water types show that they all lie on the meteoric water line for the area, indicating that recharge is precipitation falling in the surrounding hills. Trace element analyses were also conducted on all types, however these results do not show any strong trends, other than, perhaps, the enrichment of Li, Cs and H_3BO_4 in the weakly thermal waters of the caldera (Type B) and those of Tuscania (Type D). Both thermal and cold springs seem to have SiO_2 values in equilibrium with chalcedony (Minissale et al., 1997).

Considering the maximum sampling depth of about 3 m attained in this study using a hand-auger, the installed observation wells likely do not intersect even the most shallow of these aquifers. That said, because the intersected alluvial/fluvial surface sediments were formed by the erosion of the surrounding volcanic rocks, it is expected that background samples collected in this study would be most similar to those of the shallow volcanic aquifers (i.e. Type A), although those samples impacted in the CO_2 leaking area could, in theory, show signs of mixing with one of the other, deeper waters, if water and gas co-migration is occurring at this particular study site (see section 6.4.3).

3.3. Gas leakage at the ground surface

3.3.1. General overview

Near surface studies began in the Latera area during a period of extensive geothermal exploration, first with regional soil gas studies of the Vulsini Mountains (Baldi et al., 1984; Bertrami et al., 1984; Lombardi et al., 1985; Lombardi et al., 1986; Lombardi et al., 1989) and then more detailed research within the Latera caldera itself (Lombardi and Pinti, 1992; Lombardi et al., 1993; Lombardi et al., 1994). These early studies focussed primarily on the measurement of radon and helium, and showed how these tracer gases were found in elevated concentrations above the geothermal reservoir and in correspondence with the main faulting structures. Recently the area has undergone a new phase of research, both in terms of improving our knowledge of the natural geothermal system via gas flux surveys (Chiodini et al., 2007) and as a natural test site in the context of CO_2 geological storage (CCS) to study gas migration pathways, monitoring technologies, and potential environmental impact of the leaking CO_2 (Bigi et al., 2013; Donders et al., 2013; Oppermann et al., 2010; Pettinelli et al., 2010; Annunziatellis et al., 2008; Bateson et al., 2008; Beaubien et al., 2008; Pettinelli et al., 2008; Lombardi et al., 2008; Lombardi et al., 2006). The following is a short synopsis of some of the regional results from these more recent studies, while more detailed studies conducted near the areas sampled in this thesis are described in Sections 4.2 and 6.2.

The gas leaking at surface is composed of approximately 98.3% CO_2, 1% N_2, 1200 ppm CH_4, 2000 ppm H_2S, 10 ppm He, and 10 ppm H_2, with $\delta^{13}C$-CO_2 of around +2‰ and $^3He/^4He$ ratio of 0.38. (Chiodini et al., 2007). CO_2 flux rates as high as 20,000 g m^{-2} d^{-1} and soil gas CO_2 concentrations at 70 cm depth as high as 100% have been measured within the caldera (Annunziatellis et al., 2008).

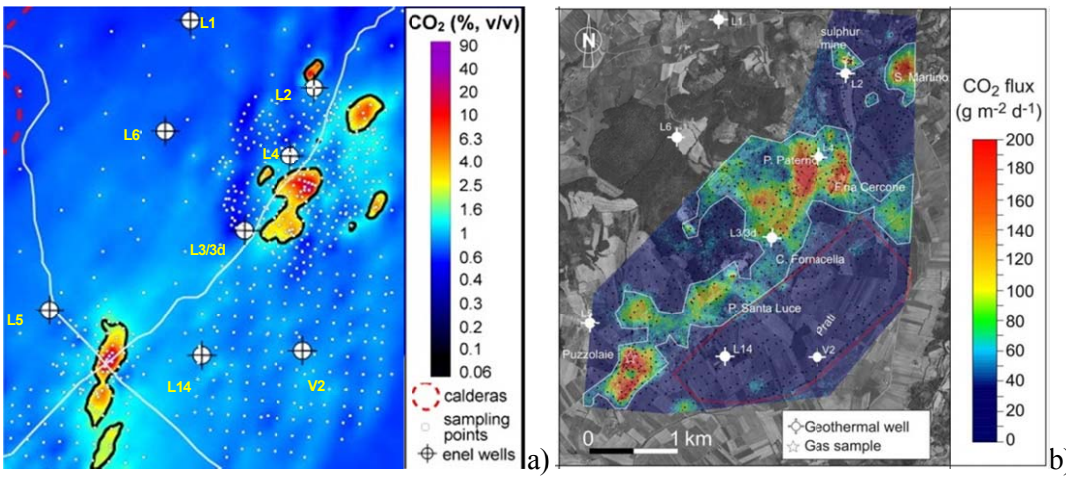


Figure 15. Regional soil gas CO_2 concentrations (a) as reported in Annunziatellis et al. (2008) and regional CO_2 flux rates (b) as reported by Chiodini et al. (2007).

Leakage of deep-origin geogenic gas does not occur throughout the entire Latera caldera, but rather is concentrated primarily along a central belt that runs more or less NE-SW. This can be seen in the two figures presented in Figure 15, which give (a) soil gas CO_2 concentrations at a depth of about 70 cm reported by Annunziatellis et al. (2008) and (b) CO_2 flux rates reported by

Chiodini et al. (2007). Although the major leakage areas are observed in both surveys, such as at Puzzolaie, on the SE side of the road between wells L3 and L4, at S. Martino, and at the sulphur mine, the gas flux map defines much wider anomalies as well as some additional anomalous areas that are not observed in the soil gas data. These differences could be due to:

i) different sampling locations and sample spacing (relative to target size), as these parameters will determine what leakage points are sampled and how the final results are contoured. One example is located to the immediate NW of P. Santa Luce, where a number of flux anomalies were observed but where very few soil gas samples were collected;

ii) the regional flux data was collected in October while all the soil gas samples were collected during the hot summer months. It is possible that cooler, moister conditions in October may have influenced baseline biological CO_2 production compared to the hot, dry summer;

iii) thresholds to separate background CO_2 from leaking CO_2 were chosen in both studies using the cumulative distribution plot method of Sinclair (1991), however different thresholds could have influenced the extent of the defined anomalous areas;

iv) the flux data were all collected during a single campaign whereas the soil gas data represent a composite of various campaigns. Although unlikely, there is a small possibility that leakage changed between the sampling periods;

v) the different data processing approaches (sGs for flux and ordinary kriging for soil gas) may have influenced the final plot appearance. Numerous authors have shown, however, that while sGs can give more heterogenous distributions compared to the more smoothed output from kriging, the overall anomaly trends and extents produced with the two methods are typically quite similar (e.g. Lewicki et al, 2005).

While many differences are subtle and could be explained by the points listed above, a number of strong differences between the two maps are more difficult to explain. For example, the large flux anomaly located between well L4 and L3/3d (with extension to the NW) is marked by quite low soil gas CO_2 values.

Gas leakage within the caldera can take many forms on the ground surface, depending on the flux rate, the surface vegetation, and if the flow is concentrated or diffused. Some examples of visible impacts are given in Figure 16, showing photos of the strongest leakage area at Puzzolaie (a), a single isolated gas vent with a bare-soil core surrounded by a halo of impacted vegetation located 100 m south of well L4 (known as gas vent A, "GVA", in the literature; Beaubien et al., 2008), and a series of smaller, elongate leakage areas defined by taller yellow grass, located in the same field as GVA. Based primarily on the assumption that leaking CO_2 will have an impact on plant health, like that observed in Figure 16c, remote sensing (RS) surveys were conducted by plane and helicopter across the Latera caldera to try and map leakage on a large scale. Using such techniques as plant stress indicators in multi-spectral data, scores were assigned for areas showing anomalous values for one or more techniques and then checked by conducting soil gas and CO_2 flux measurements (Bateson et al., 2008). A significant number of gas leakage areas were defined within this study, however many false positives were identified (i.e., remote sensing anomalies that were not due to CO_2 leakage) and a certain number of false negatives also occurred (i.e., leakage areas not recognised by the remote sensing methods). Clearly in its present state remote sensing is not capable of delineating, with enough sensitivity and enough reliability,

all leakage areas. In particular, areas with low flux rates (e.g., < 100 g m^{-2} d^{-1}) likely do not have an impact on surface vegetation that is measureable with present-day airborne techniques.

Estimates have been made of the total amount of CO_2 leaking from the Latera geothermal field using different approaches. Gambardella et al. (2004) used the total surface heat flux (estimated based on deep borehole data), the enthalpy of liquid water at reservoir temperature, and the CO_2 molality of the single liquid phase circulating in the geothermal reservoir to calculate the total CO_2 flux released upwards from the Latera geothermal reservoir, yielding an estimate of 2.58 x 10^8 mol yr^{-1} (about 11 kT CO_2 yr^{-1}). Annunziatellis et al. (2007) conducted detailed surface flux surveys (10 to 25 m sample spacing) over the largest eight gas leakage areas and used kriging and volume algorithms in the program Surfer (see section 2.4.2) to estimate a total CO_2 flux of 1.75 x 10^8 mol yr^{-1} (about 8 kT yr^{-1}), which is only about 25% lower than the estimate of Gambardella et al. (2004). These authors caution, however, that their total flux value is an underestimate of the true value because some known leakage areas had not been sampled for the calculation. Finally Chiodini et al. (2007) conducted a regional CO_2 flux survey over much of the caldera (see Figure 15b) and used GSA and sGs (see section 2.4.2) to estimate a total CO_2 flux of 350 t d^{-1} (about 130 kT yr^{-1}), which is an order of magnitude larger than the previous two.

Figure 16. Photographs of different gas leakage points within the Latera caldera: a) Puzzolaie, a large leakage area where the groundwater study was performed (Chapter 6); b) medium and c) small gas vents, visible as longer yellow grass, in the field where the gas flux survey was conducted (Chapter 4). Note that (b) is referred to in the text as gas vent A (GVA). Note also that these photos were taken during different seasons and were not taken during the present field work.

As described by Bigi et al. (2013), leakage from a bedrock fault exposed in a Latera quarry (sulphur mine in Figure 15b), where there is no sedimentary overburden, shows gas flow within the damage zone (up to 5 m wide) but very low flow in the impermeable fault core. By modelling the flow and permeability characteristics of five different types of damage zone fracture networks exposed in the quarry wall, these authors clearly illustrate that leakage is highly spatially variable. Flow within and along the damage zone is channelled and localized, with the gas flow pathway being controlled not only by the permeability of the fracture but also by its orientation and, most importantly, its interconnectivity with other permeable fractures; maximum flow was observed to be parallel to the intersection of conjugate fractures. As discussed previously in Section 2.3.2, areas of dilational deformation also result in the formation of open fractures that can act as conduits, such as relay ramps, folding along transfer zones between normal faults, or the intersection of two faults; these areas could be highly permeable and could cover a wider area compared to the damage zone of a single fault. However, for the most part, leakage areas in the Latera caldera are not from exposed bedrock faults, but rather emanate from the alluvial sediments that cover the bedrock. As outlined in Section 2.3.3, the sediments may or may not be fractured in relation to the fault, and if they are not their highly anisotropic permeability field has the potential to control, focus, and/or offset gas flow in the shallow sediments beneath the water table as well as in the unsaturated zone.

Chapter 4. The leakage of CO_2 at Latera

4.1. Introduction

The reasons for studying the leakage of non-biogenic gas from the soil, both in terms of spatial distribution and total mass flux, can be subdivided into three main groups.

i) Resource exploration. This includes the search for hydrocarbons, mineral deposits, and geothermal energy sources, and is based on the fact that certain tracer gases associated with these resources can migrate towards the surface where their anomalous concentrations or fluxes can help (together with other geological and geophysical methods) focus further exploration.

ii) Geochemical cycling and mass balance research. The exchange between the solid (rock), liquid (groundwater), and gaseous phases regulate (and are regulated by) many geological, chemical and biological processes, and their study can give important information regarding large-scale processes. One of the most important of these is global climate, as controlled by the sources and sinks of various greenhouse gases. In addition to man-made pollution and well-known natural sources, like biological production and volcanic eruptions, the leakage of CO_2 from geothermal fields or the flanks of quiescent volcanos may be significant (Chiodini et al., 2004) and thus it requires accurate quantification. Another area where mass balance is important regards the monitoring of CO_2 capture and storage (CCS) sites, as any leakage to the atmosphere of originally stored CO_2 has to be accurately quantified to offset any carbon credits originally awarded for the storage of that gas.

iii) Health and safety. The leakage from the soil and accumulation in enclosed areas on the surface of certain gases can represent both a health and environmental risk, both from natural and man-made sources. An example of a natural source is once again the flux of large volumes of CO_2 in geothermal / volcanic areas, while man-made systems include the potential for leakage from CCS sites, hydro-fracturing petroleum recovery activities, natural gas storage reservoirs, or buried pipelines. For natural sites where leakage is relatively constant, gas flux and soil gas measurements can be used to define high risk areas for zoning bylaws and building permits (e.g. Beaubien et al., 2003). For man-made sites, where leakage may never occur, monitoring is needed to ensure that the public is safeguarded if it does. In this regard the development and testing of monitoring methods that are both sensitive and capable of separating leakage from near-surface biogenic anomalies is fundamental.

The work presented in this chapter primarily addresses the latter two points, looking at issues related to finding and quantifying leakage (e.g., sample density, background subtraction) and migration processes and pathways (e.g., faults, sediment control) via the measurement of CO_2 flux over a detailed grid in a field in the centre of the Latera caldera. In addition, another critical goal of this work was to obtain a real-world dataset of spatially distributed CO_2 flux measurements on which sub-sampling simulations could be conducted using the software developed during this research (Chapter 5), to compare with the results obtained applying the same software to purely synthetic data.

A total of 550 flux measurements were performed July 16 and 17, 2014 in one field (Figure 17) following an offset grid pattern (see section 5.2.2.2 for an explanation of sampling approaches) and using three identical flux measurement systems. Although mid-July is typically a very hot, dry month during which the soil dries out, thus enhancing exchange with the atmosphere and decreasing soil respiration, the summer of 2014 experienced anomalously high rainfall. In fact, a heavy rain occurred in the Latera area only four days before the measurement campaign, and the study site was found to have moist soil and green grass (as opposed to the dry, yellow grass which occurs during this period in a regular year). As discussed below, these conditions influenced the obtained results.

4.2. Study site

The study site itself is found in the centre of the caldera (Figure 14) within an a small, irregularly shaped field used for animal pasture (Figure 17). Detailed study of one isolated gas vent in this field ("GVA"; Figure 17) has shed light on the spatial distribution of leaking gases, their impact, and has been used to study gas migration in the near surface. In particular, a horizontal profile conducted across (Annunziatellis et al., 2008) and a grid performed over (Pettinelli et al., 2010) this vent shows different behaviours for the different measured parameters (Figure 18).

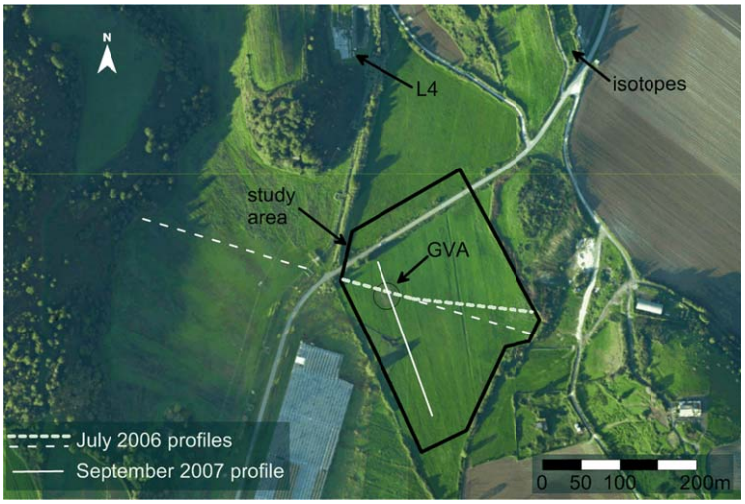


Figure 17. Air-photo of the present study area, together with locations of previous measurements discussed in the text. "GVA" refers to 'gas vent A', "isotopes" refers to a 6m deep vertical profile on which $\delta^{13}C$-CO_2 analyses were conducted, and L4 is a deep geothermal well. Air-photo courtesy of Michela Vellico, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Trieste, Italy.

Soil gas CO_2 has the widest area of anomalous values around the gas vent centre (Figure 18a) due, likely, to lateral diffusion and effects induced by the higher density of CO_2, and shows a NNW-SSE alignment of low-level anomalous values (Figure 18d) that may trace the underlying gas-permeable fault. CO_2 flux distribution is less smooth, with extremely high values in the core of the vent relative to the surrounding area (Figure 18a) and a spatial distribution of anomalous values that appears to be less than that of soil gas CO_2 (Figure 18e). Re-plotting of the flux data on a logarithmic scale (Figure 18b) clearly shows that flux returns to background values (around 10 g m^{-2} d^{-1}) essentially at the same location as the CO_2 concentration values, approximately 25 m away from the vent core. This implies that leakage occurs in an area around a central area (be it a point or a line) where the barrier characteristics of the sediments have been locally

compromised. The link between flux and soil CO_2 is further illustrated in Figure 9, where curved contour lines define the concentration gradients that control CO2 diffusion from the soil to the atmosphere around a central leakage point.

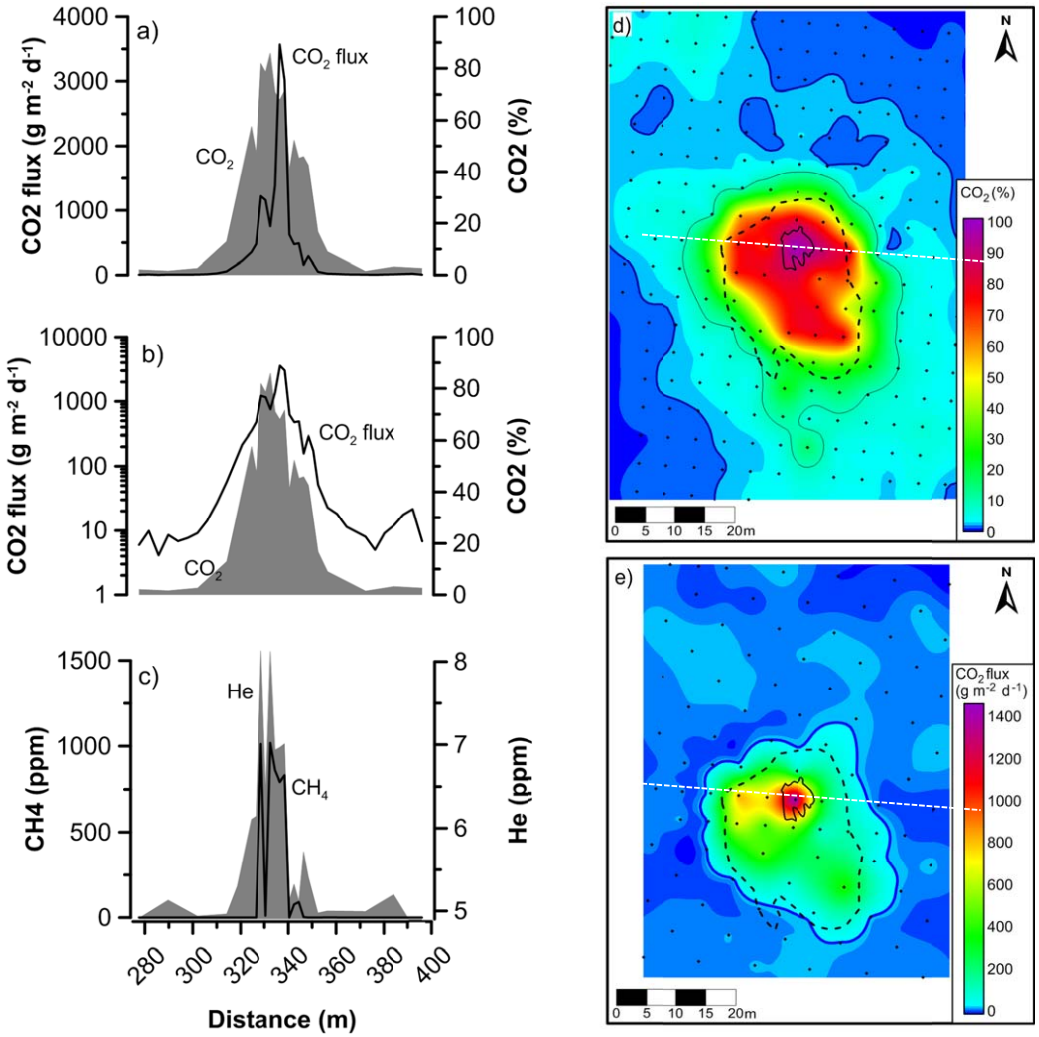


Figure 18. Plots showing part of the results from a long horizontal profile (Figure 17) that crosses GVA (a - c) (modified from Annunziatellis et al., 2008) and from a grid of samples around the same gas vent (d, e) (Pettinelli et al., 2010). Plots a) and b) present the same data, but with the latter having a logarithmic CO_2 flux scale. The dashed white line in both contour plots shows the approximate trace of the profile presented in a) – c).

For these reasons, and for the high level of detailed information available for this location, GVA was used as the "template" for the creation of artificial gas vents in the modelling work conducted in Chapter 5 below. In addition to adopting the circular shape based on GVA's form (Figure 18e), the change in flux values moving away from the core of GVA towards its outer edge (Figure 18a) was used to develop the formula that creates the synthetic vent flux rates within the specified vent radius (see section 5.2.2.1). Finally, Figure 18c shows that both He and CH_4 have narrower spatial distributions, He because of its higher mobility and CH_4 (and other redox-reactive gases like H_2S) because of their rapid oxidation outside the vent core where CO_2 flux and CO_2 concentrations decrease and O_2 values increase (Annunziatellis et al., 2008). The relationship between CO_2 and O_2 in the soil across GVA can be seen in Figure 3.

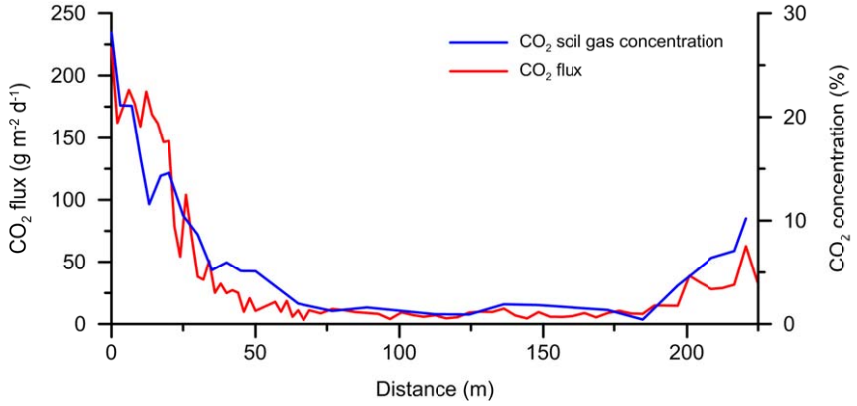


Figure 19. Soil gas CO_2 concentrations at 80 cm depth and CO_2 surface flux values for the western-most segment of the long horizontal profile shown in Figure 17 (data re-drawn from Annunziatellis et al., 2008).

Other research conducted on GVA includes electrical resistivity tomography profiles that show low resistivity at surface in correspondence with the gas vent core and laterally at a depth below about 4 m (Pettinelli et al., 2010), while unpublished data shows a low seismic wave velocity unit at 5-7m depth in a multichannel analysis of surface wave (MASW) survey and a hard layer at 2-3 m depth just outside of the vent measured via a penetrometer test (pers. comm. F. Cecchini and S. Margottini). Put together this data implies that gas has accumulated beneath a shallow impermeable layer, possibly formed by precipitation of carbonate minerals induced by the CO_2 itself, which focusses leakage within the gas vent and impedes vertical migration from depth over a larger area, with lateral extent of surface flux controlled more by movement in the upper, shallow unsaturated zone (see Figure 9).

Data from the western half of the same long horizontal profile described above are presented in Figure 19 (thick dashed white line in Figure 17). Here gas leakage is observed from 0 to 75 m and from 180 to 225 m, with low background values in between that are equal to about 0.9-1.9% and 3-12 g m^{-2} d^{-1}, respectively. This result, similar in form to that observed for GVA in Figure 18a, clearly implies that leakage is spatially restricted, with diffuse degassing being controlled by a central leakage point and by lateral migration within the unsaturated zone. Note that the trend of the first vent (Figure 19), from the start of decreasing values around 20m to background values at around 50m, is matched well using the simulation formula developed in section 5.2.2.1 (see Figure 29b).

Finally, CO_2 isotope and concentration data (Figure 20a and b, respectively) have been collected from a 6m deep vertical profile through a gas vent located on a small near-by hill ("isotopes" in Figure 17) and from a background non-leaking site located on the edge of the caldera (location not shown). At the leaking site the $\delta^{13}C$-CO_2 values are relatively constant at 1‰, but increase towards the surface above 2m depth, likely due to diffusive fractionation as described in Camarda et al. (2007). The constant value of 1‰ is consistent with a thermogenic origin due to breakdown of marine carbonates (see section 2.2.3). In contrast, the isotopic value at the background site, at a depth of almost 2m, is equal to -20.9‰, a value which is consistent with the biologically respired CO_2 produced via the near-surface breakdown of organic matter originating from C3 or C4 plants (see section 2.2.1).

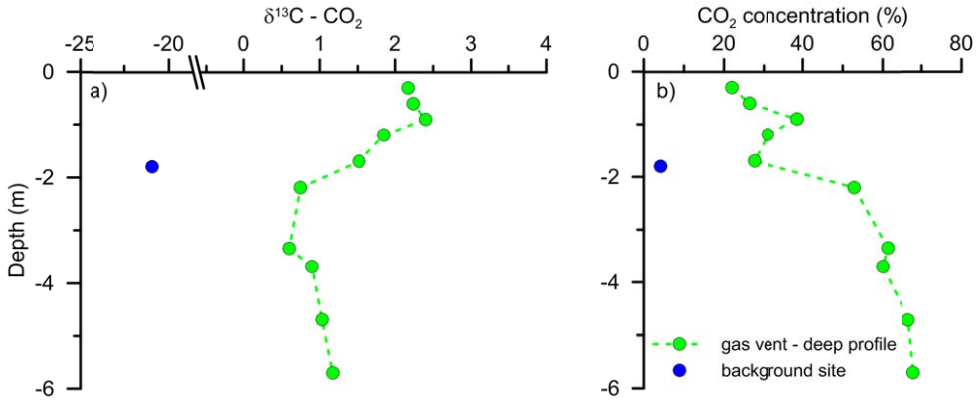


Figure 20. Stable carbon isotope (a) and concentration (b) data for CO_2, collected from a deep unsaturated zone profile performed in a gas vent located on a small hill near the study area ("isotopes" in Figure 17) and from a background site located about 2km NNE of the site near the town of Latera (Beaubien et al., 2008b). Note the broken X axis in (a).

4.3. Material and Methods

4.3.1. Flux measurements

Carbon dioxide flux measurements were made using a closed-circuit accumulation system (e.g. Hutchinson and Livingston, 1993) that was constructed "in-house" by researchers at the Università di Roma "La Sapienza". The control unit is equipped with a 0-3000 ppm infrared CO_2 detector while the accumulation chamber itself is a small aluminium container with an inlet and outlet tube for the IR detector and a small equilibration tube to prevent artefacts above high flux points and on windy days (Hutchinson and Livingstone, 2001). The accumulation chamber was pressed firmly to the ground to insure a proper seal after the removal of any surface vegetation, and CO_2 concentrations were measured and stored every second for about 90 seconds. Flux values were calculated using the following equation (Lewicki et al., 2005b):

Eqn 13 $$\phi CO_2 = (d[CO_2]/dt)*(V/A)*(T_0/T)*(P/P_0)*k$$

where $(d[CO_2]/dt)$ is the first linear slope of the CO_2 concentration increase (ppm/s), V is the chamber volume (m^3), A is the chamber surface area (m^2), T and T_0 are measured and standard temperature (°K), P and P_0 are measured and standard pressure (kPa), and k is a unit conversion factor (169.71). Quantification of CO_2 leakage rates using gas flux measurements usually consists of sampling on a grid, interpolating between points, conversion to total CO_2 flux for the measurement area, and subtracting near-surface contributions based on baseline studies or soil gas data, as discussed above in section 2.4.2.1.

4.4. Latera flux results and discussion

The spatial distribution of CO_2 flux values within the measured grid are shown in Figure 21, plotted using the entire flux range up to 1000 g m^{-2} d^{-1} (a) as well as a narrower range up to 200 g m^{-2} d^{-1} to highlight lower level anomalies (b). A total of six strong and two weaker anomalies are visible (Figure 21a), occurring primarily as individual "hot spots" distributed throughout the

study grid that range in size from about 3 to 40 m in diameter. Although the minimum diameter anomaly that can be found at the 95% confidence level at this sample spacing using the offset grid approach is about 12 m (see Figure 40 below), one smaller anomaly was found due to random chance (vent 8) while another was due to an additional sample being added to the grid based on the visual observation of impacted vegetation (vent 6). Thus it is possible that other leakage points smaller than 12 m were not intersected with the chosen grid spacing.

The isolated nature of these leaks, as discussed by Annunziatellis et al. (2008) for the Latera site, is likely due to channelling along the highest gas permeable pathways within the buried faults followed by vertical, buoyancy-driven ascent within the saturated sediments. Final movement within the unsaturated vadose zone tends to result in some lateral movement, due to diffusion and density driven migration of the denser CO_2 gas, that widens and "smears" the anomaly as a function of the maximum advective flow in the centre of each gas vent. Similar hot spot anomalies have also been observed in controlled experiments, such as the injection of CO_2 into a fault within the Siena Graben in central Italy (Ciotoli et al., 2005), into a shallow buried horizontal pipe at the ZERT test site in Montana USA (Lewicki et al., 2010; Oldenburg et al., 2010), and at the CO_2 Field Lab in southern Norway (Jones et al., 2014).

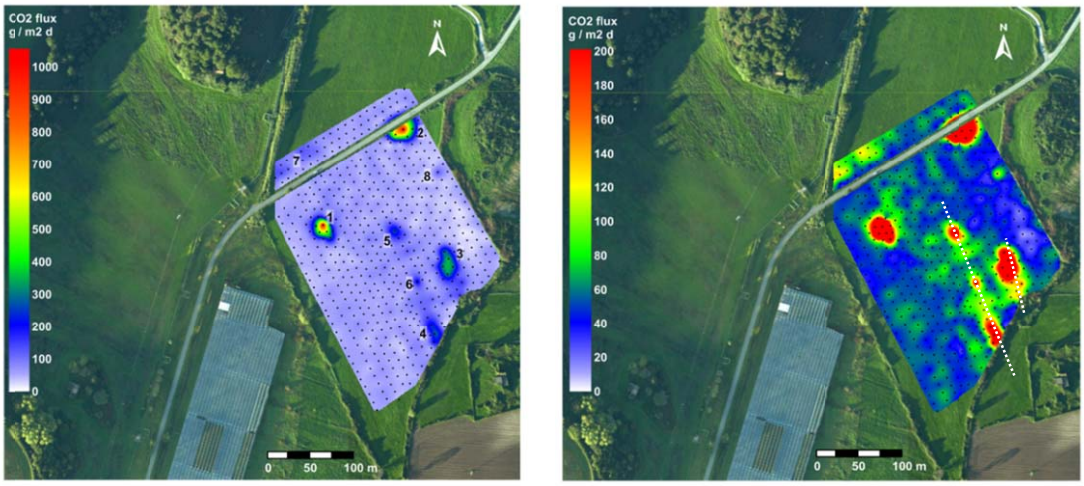


Figure 21. Contour maps of CO_2 flux measurements conducted in July 2014, plotted using the entire range of values (a) and a narrower range to highlight lower level anomalies (b). Base map courtesy of Michela Vellico, OGS, Trieste Italy.

A number of these anomalies appear to be elongate and/or aligned along a N20W or N10W direction, as shown by the dashed white lines in Figure 21b. Although this differs from the more regional lineament and structural trends of about N10E observed by Annunziatellis et al. (2008), it is essentially the same as that observed by Pettinelli et al. (2010) for lineaments occurring in the same field around GVA. This illustrates that the local controlling structure that provides the migration pathway from the deep geothermal source likely consists of pervasive fracturing along this direction.

As log CO_2 flux data collected from a single population (i.e., single origin and controlling migration process) tends to lie on a straight line when plotted on a probability plot (Sinclair, 1974), the inflection points observed in the 2014 grid dataset (Figure 22) show the possible existence of three different populations. The dominant feature in this figure is the change in trend occurring around 70 g m^{-2} d^{-1}, which separates highly anomalous values up to 1000 g m^{-2} d^{-1} from the rest of the dataset. Whereas the upper population clearly belongs to the leakage hot

spots defined in Figure 21, the origin of the values between 9 and 70 g m^{-2} d^{-1} is less clear. One possible explanation is that the smaller convex inflection point at about 40 g m^{-2} d^{-1} separates two different populations, with the lower one (grey line) being of shallow biogenic origin and the upper one (green line) being low level diffusive leakage. An alternative is that this is essentially all one population linked with biogenic production and that the minor inflection at 40 g m^{-2} d^{-1} separates populations due to other processes (such as, for example, diurnal variability during the sampling) or artefacts of the sub-sampling statistics (see section 5.3.2.1 and Figure 42); such multiple background populations have also been observed by Mazot et al. (2013), which they attributed to differences in soil permeability and productivity. As discussed below, determining which interpretation is correct for the present data set and, by association, what the average biological CO_2 flux actually is, has a significant influence on the eventual estimate of the overall leakage flux rate (see also section 2.4.2).

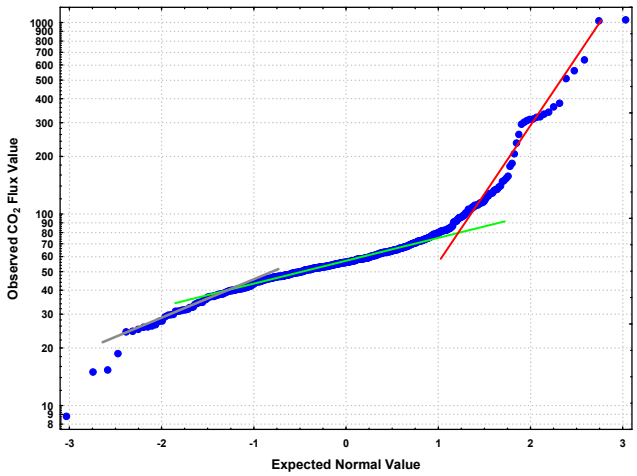
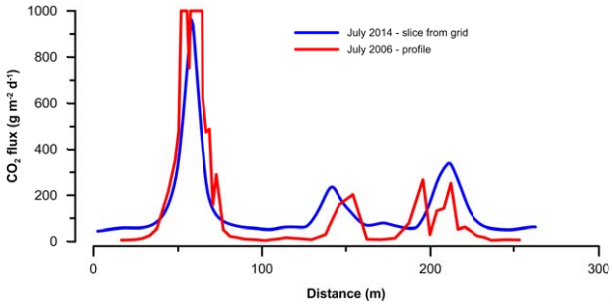


Figure 22. Normal Probability Plots (NPP) of the entire July 2014 grid CO_2 flux dataset plotted on a log scale. The three coloured lines mark possible different populations based on similar trends.

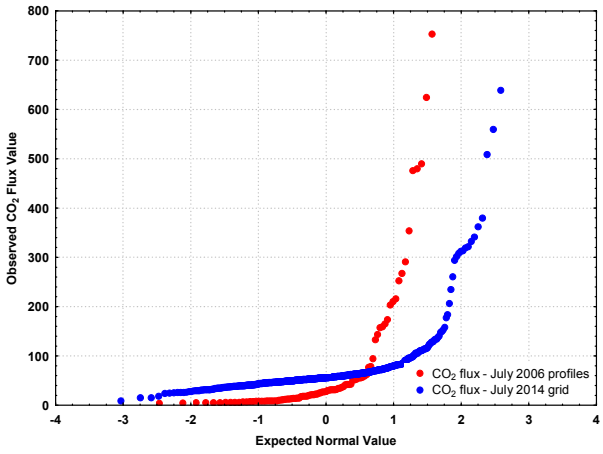
Whereas values of 70 g m^{-2} d^{-1} are on the high side for biogenic soil respiration, such values and higher have been observed in the literature (e.g. Beaubien et al., 2013). In this regard, the positive correlation between CO_2 soil respiration and water content and temperature is well known (e.g. Lloyd and Taylor, 1995; Barron-Gafford et al., 2011), and thus the anomalously wet summer experienced in central Italy during 2014 may have increased biogenic flux. To examine this possibility the contour grid in Figure 21 was sliced to produce a plot of CO_2 flux data along the same line as a transect performed in July of 2006 (grey dashed line in Figure 17), a year with more normal, low rainfall. The sliced data (this work) and the transect data (re-processed from Annunziatellis et al., 2008) both show three main peaks associated with leakage points (Figure 23a), with slight horizontal offsets due to positional accuracy. The transect is more irregular because it is made from individual measurement points spaced 4 to 6 m apart, thus highlighting spatial variability, whereas the data sliced from the contoured grid is more smoothed because of the nature of the contouring algorithm.

Despite these differences the maximum values of the two smaller peaks are quite similar in amplitude for the two campaigns. In sharp contrast, however, the baseline values between the peaks are very different, with the July 2006 transect showing values in the range of 5 – 15 g m^{-2} d^{-1} and the July 2014 grid data showing values of around 50 g m^{-2} d^{-1} (Figure 23a). These results imply an overall rise in the background CO_2 flux values throughout the entire field during the 2014 campaign. A comparison of the statistical distribution of all the 2006 profile data collected

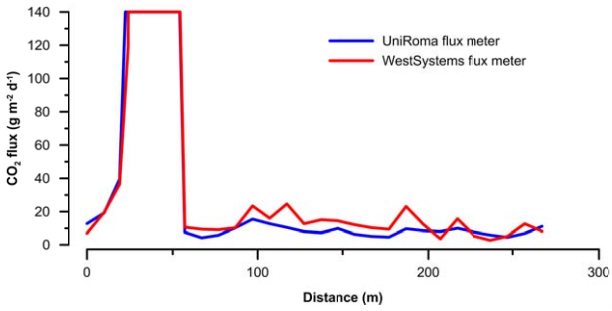
in this field (100 points, location of profiles shown in Figure 17) against all the 2014 grid data (548 points, Figure 21) in a probability plot highlights this difference (Figure 23b). This plot, using a normal scale to a maximum of 800 g m^{-2} d^{-1}, clearly shows how the bulk of the samples were much lower and had a less variable statistical distribution during the dry 2006 conditions compared to the wetter 2014 conditions.



a)



b



c)

Figure 23. a) comparison of CO_2 flux values from a transect performed across the study field in July 2006 during very dry conditions and results extracted from the grid data collected during the wetter July 2014 conditions. b) probability plot comparing the statistical distribution of the complete 2006 and 2014 datasets. c) comparison of results from two different flux instruments along a profile in the study field conducted in September 2007.

Similar low results were also obtained in September 2007, when a profile of 30 points measured using both a commercial unit (West System Srl) and the University of Rome unit showed background values on the order of 5 to 20 g m^{-2} d^{-1} (Figure 23c). These results also indicate that the in-house developed flux meter used in the present study gives reliable results that are very similar to that of a commercial unit.

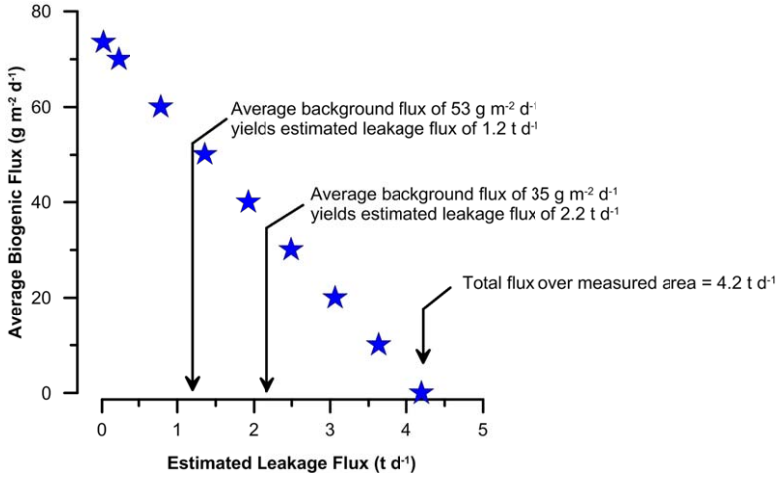


Figure 24. Estimated leakage flux as a function of estimated average biogenic flux.

As mentioned, the importance of an accurate estimate of the baseline biological CO_2 flux is needed for a similarly accurate estimate of the leakage flux, as this latter value is calculated by multiplying the average baseline flux per square metre by the total surface area and subtracting it from the measured total flux (see section 2.4.2.1). In the present study the uncertainty arising from the possible two populations below the flux value of 70 g m^{-2} d^{-1} (Figure 22) complicates the calculation of the average background value, although comparison with previous campaigns implies that even some of the values on the higher end of this range may have a biological origin. The influence of the eventual background value chosen is shown in Figure 24, where various average background values are plotted against the resultant estimated leakage flux using that value. Values range from a no biogenic flux resulting in the leakage flux equalling the total measured flux of 4.2 t d^{-1}, to a value of no leakage flux if the average background flux is assumed to be 73 g m^{-2} d^{-1}. Instead of these two completely unrealistic extremes, the calculated leakage flux rate using the average of the lower population (i.e. grey line in Figure 22) yields a leakage rate estimate of 2.2 t d^{-1} whereas the average of all values below 70 g m^{-2} d^{-1} yields a leakage rate estimate of 1.2 t d^{-1}. Although the present data does not allow one to choose between these two possible values, the comparison with previous data (Figure 23a) implies that the higher background average, and thus lower leakage flux rate of 1.2 t d^{-1}, may be closer to the actual value. Isotopic analyses plus more seasonal work would be needed to better verify this hypothesis.

4.5. Summary

A total of 550 CO_2 flux measurements were made over an area of 57,000 m^2 within a single field in the centre of the Latera caldera. Eight hot spot anomalies were defined that range in size from 3 to 40 m in diameter and that attain maximum flux values of between 200 and 1000 g CO_2 m^{-2} d^{-1}. The minimum vent diameter that can be found at the 95% confidence interval using the applied offset grid spacing of 10 m is about 12 m (see Chapter 5), however smaller vents can (and were) found by chance and by adding additional samples where leakage was inferred. A number of the leakage hot spots are aligned along a N20W to N10W direction, in agreement with other studies conducted in the same field. Deep gas-permeable faults underlying the alluvial sediments in the field are likely aligned along this direction and control the leakage on the surface.

These hotspots are located within a background field that ranged from 9 to 70 g CO_2 m^{-2} d^{-1}. A statistical analysis of the entire dataset shows that although there is a clear distinction between high flux hotspot values and the rest of the measured values, it is difficult to determine the cut-off level between low level flux resulting from purely near-surface biogenic origins compared to that which is biogenic plus varying levels of low-level geogenic leakage. Compared to the newly collected grid data during the wet summer of 2014, historical data collected during much dryer periods in 2006 and 2007 show much lower CO_2 flux levels away from the main hot spots, implying a strong biogenic control in the background areas linked to different environmental conditions. Soil gas sampling, involving isotopic analyses and vertical profiles, would greatly aid in better determining the balance of these two sources in the areas outside the leakage hotspots.

Average background values must be subtracted from the total measured flux to estimate the leakage flux rates. Based on the statistical distribution of the data, two different average "biogenic" flux rates were calculated and used to show how they can influence the final estimated leakage flux rate. An average background (applied to the entire surveyed area, including those areas with known geogenic CO_2 leakage) of 35 g m^{-2} d^{-1} yielded a leakage estimate of 2.2 t d^{-1} while 53 g m^{-2} d^{-1} yielded 1.2 t d^{-1}. This difference, where one estimate is almost double the other, illustrates the importance of an accurate estimate of the biological near surface flux when calculating leakage flux rates. Clearly this issue is critical for both the estimation of natural leakage for geological or atmospheric mass balance studies as well as for carbon credit auditing above CCS sites.

The obtained results can be used to address various issues related to these types of studies:

- The chosen sample approach (offset grid) and sample spacing (10 m) may potentially have missed leakage points that have a diameter less than about 12 m, however considering that smaller vents have smaller leakage rates (e.g., vent 6 in Figure 21) any such missed features would probably not have a significant effect on the overall leakage estimate. For example a 100 m^2 vent (i.e., 10x10 m) with an average flux of 200 g m^{-2} d^{-1} would contribute a total of 20,000 g d^{-1}, which only represents about 1.5% of the lower leakage rate of 1.2 t d^{-1}.
- A more significant error can come from the chosen level of the background biogenic flux, with the present example showing a 100% difference between two estimates due strictly to the chosen average background value.
- The present results appear to support the theory that the bulk of leakage within this field occurs from individual leakage points (or hotspots), where a preferential pathway has been created via pressure-driven advective flow in the sediments above a gas permeable fault. Although diffusion of geogenic gas is important outside the vent cores, it appears to be more related to lateral movement in the unsaturated zone rather than a constant diffuse flow from below. That said, more detailed work involving isotopes and vertical soil gas profiles would be needed to better quantify the ratio between geogenic and biogenic flux, and over what surface area it is significant.
- Due to the natural variability of different fields in terms of land use, soil type, water content, etc., the use of a different field distant from the studied site as an estimate of baseline values risks to under- or over-estimate the biogenic background, thus impacting on the eventual leakage flux calculations.

Chapter 5. Modelling of sampling density

5.1. Introduction

A major focus of the present research was to develop a new computer code capable of examining the effects of different sampling strategies and sampling densities on the end results of gas flux surveys aimed at CCS monitoring or quantification of natural geological CO_2 terrestrial diffuse emissions. The program creates a highly detailed (1 m node spacing) original dataset of synthetic data (or imports real CO_2 flux data from a grid), which is in turn randomly sub-sampled N times for each of M different sample densities using one of four different sample strategies. Various calculations and monitoring of different parameters permits one to determine, under the simulation conditions and sampling settings, the number of samples needed to find a leakage point and the range of leakage flux values calculated compared to the "true" original value. In the following, section 5.2 outlines the functioning of the program, section 5.3 describes the results of simulations conducted on synthetic datasets, while section 5.4 discusses the simulation results obtained using real data from the Latera test site. The program was written by the author in Visual Basic 6 and the complete code is provided in APPENDIX I.

5.2. Program architecture

5.2.1. Data Input

Each simulation is run by choosing various options and input files via the Graphical User Interface (GUI) shown in Figure 25. A general introduction and overview of the various options available to the user are presented below, making reference to this figure and the highlighted numbers for each frame, whereas the actual workings of the program are discussed in Section 5.2.2.

The program is started in 1 – Initial Settings, where clicking on the Activate and Confirm path buttons loads the default program settings and file paths. Here the maximum number of X and Y grid points can be chosen based on a fixed 1 metre spacing, with the total number of points in the grid matrix being calculated automatically. Note that "X max" and "Y max" can be user defined when synthetic data is used, but they are set externally if real data are read into the program. A description of the simulation can be written in 8 – Description of the simulation, and this information, together with all settings chosen on this window, is saved in a log file for future reference. The type of data is chosen in 2 – Data Type, which can be either simulated data or real data that is read into the program.

If simulated data is chosen a background dataset of biologically produced CO_2 flux is generated and distributed spatially on the X-Y grid and artificial gas leakage points ("vents") are added to this background distribution at defined points. For this option, two types of background data are available under 4 – Background Flux Rate, either a random normal distribution calculated by the program itself ("calculate") and randomly distributed over the grid, or the importation of an external file ("import") having a chosen statistical distribution (e.g. log-normal) and a non-

random spatial distribution. The location, size, and strength of the gas vents are set under 5 – Vent Settings. Here either a single point can be inserted, or multiple vents can be defined by importing an external list file.

If, instead, real data is chosen an external file is read into the program. As this is real data there is no need to generate a background distribution or artificial gas vents. Instead only the fixed average option is available under 4 – Background Flux Rate, which is a number selected prior to simulations based on expert knowledge of the real site and associated dataset, while 5 – Vent Settings is disabled.

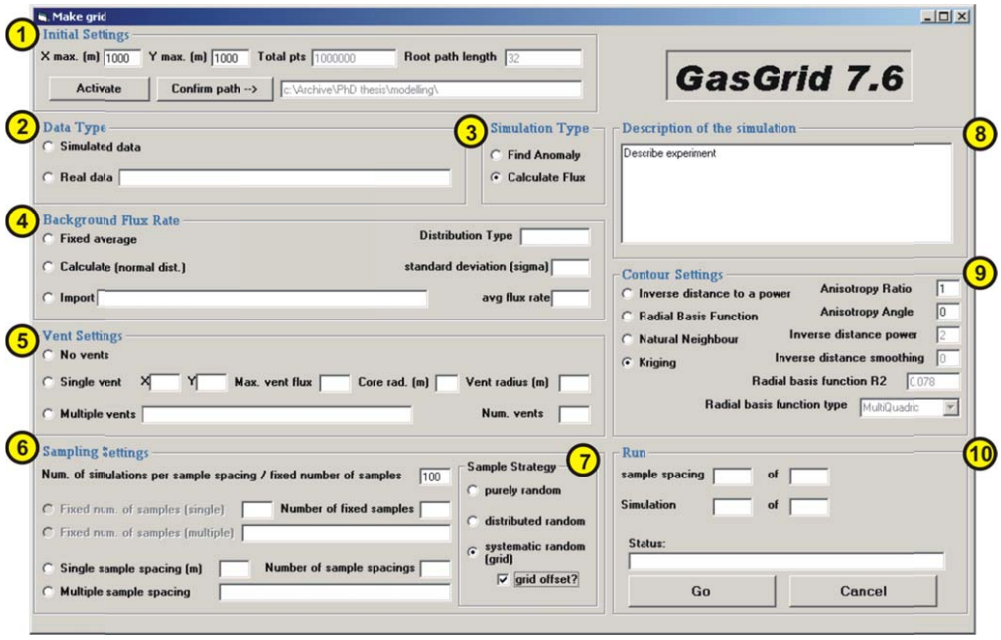


Figure 25. Screen capture of the GasGrid data input window.

These synthetic or real datasets can then be processed, choosing in 3 – Simulation Type either a complete treatment (option "Calculate Flux") or a simplified, faster version which bypasses Surfer9 contouring and calculations for each simulation (option "Find Anomaly"). Whereas the first option examines how sampling density can influence interpolated spatial distribution and estimates of total leakage flux rates, the latter looks only at the probability of finding that leak at a given sampling density and a given ratio between leak and survey area, assuming that the leak can be distinguished with absolute certainty from the background values.

The sub-sampling settings are chosen in 6 – Sampling Settings and 7 – Sample Strategy; here it is decided how many simulations will be performed for each given sample spacing (or fixed number of samples) and what sampling strategy is applied, including purely random, uniform random, grid, and offset grid. Random sampling is based on a fixed number of samples collected for each simulation, while the other three are based on a nominal minimum (uniform random) or exact (grid and offset grid) sample distance which results in a slightly different number of samples collected for each simulation. As above, either a single value or a multiple value file can be chosen for the sample spacing or fixed number of sample options. If the "calculate flux" option is chosen in 3 – Simulation Type, one of four different interpolation algorithms can be chosen in 9 – Contour Settings (along with respective options) for contouring using objects called from Surfer9. Finally 10 – Run contains the Go button to start the simulations, as well as various fields that show the progress of the simulations.

5.2.2. Algorithm description

A flowchart showing the basic structure of the program is given in Figure 26, which can be briefly summarised as follows (with specific details given in subsequent sub-sections). Simulations can be conducted on either synthetic or real data. Synthetic data consists of a random background flux field that is either generated by the program or imported from an external file, onto which are superimposed gas leakage areas; the total flux of both the background and background plus leakage grids are calculated using Surfer9. If, instead, real data is used, it is read in from an external file and again the total flux for this dataset is calculated using Surfer9.

Sub-sampling is conducted on these detailed, original grids, using one of four options, purely random, uniform random, regular grid, and offset grid. The program loops until the grid has been completely sub-sampled using the "sampling density" criteria of the chosen method (sample spacing or fixed number of samples) and then the sub-sampled dataset is processed using Surfer9. These two steps (i.e. subsampling the complete matrix and Surfer9 processing) are repeated N times based on the number of simulations chosen by the user. Once completed the program checks if other "sampling density" values have been specified, and repeats the above loop with the new value if there are. Once all simulations for all sample densities are complete the program exports a series of summary files and ends.

5.2.2.1. Synthetic versus Real Data

As stated, there are two different ways to generate synthetic data. The first is to have the program itself create an array of random, normally-distributed flux values based on user-defined average and standard deviation values; these flux values are then distributed randomly on the grid, meaning that there is no spatial relationship between points (Figure 27a). This option was developed in an early version of the program because Visual Basic 6 only provides the normal random number generator, however it is not considered a realistic representation of the background flux field and thus was not used in any of the simulations reported here.

Instead the option of importing externally produced synthetic background data was added to permit the use of other, third-party programs that have a wider choice of random number generators. Of particular importance was to be able to create log-normal datasets, considering that most authors believe that CO_2 flux datasets typically follow this type of distribution. For the present work the free EXCEL add-on "PopTools" (Hood, 2010) was used. Although PopTools has a wide range of functions focused on the analysis of matrix population models and simulation of stochastic processes, it was used here exclusively to create random, log-normal distributions based on average and standard deviation values chosen by the user (within PopTools) and considered realistic for background CO_2 flux datasets.

The use of an external background file also allowed for additional data processing prior to importing into the program, specifically to order the dataset such that some form of spatial continuity could be created (i.e., a non-random spatial distribution). To this end topographic data was used to spatially distribute the synthetic CO_2 data, based on the fact that soil biogenic CO_2 production can be enhanced in local, low-lying areas due to greater moisture content and potentially higher organic matter content with respect to more well drained high ground (see section 2.2.1). Although a slight oversimplification in the case of CO_2 flux, which does not always correlate well with soil gas CO_2 concentrations due to surface soil effects (e.g. low permeable surface soil), it does allow for the creation of spatially coherent data.

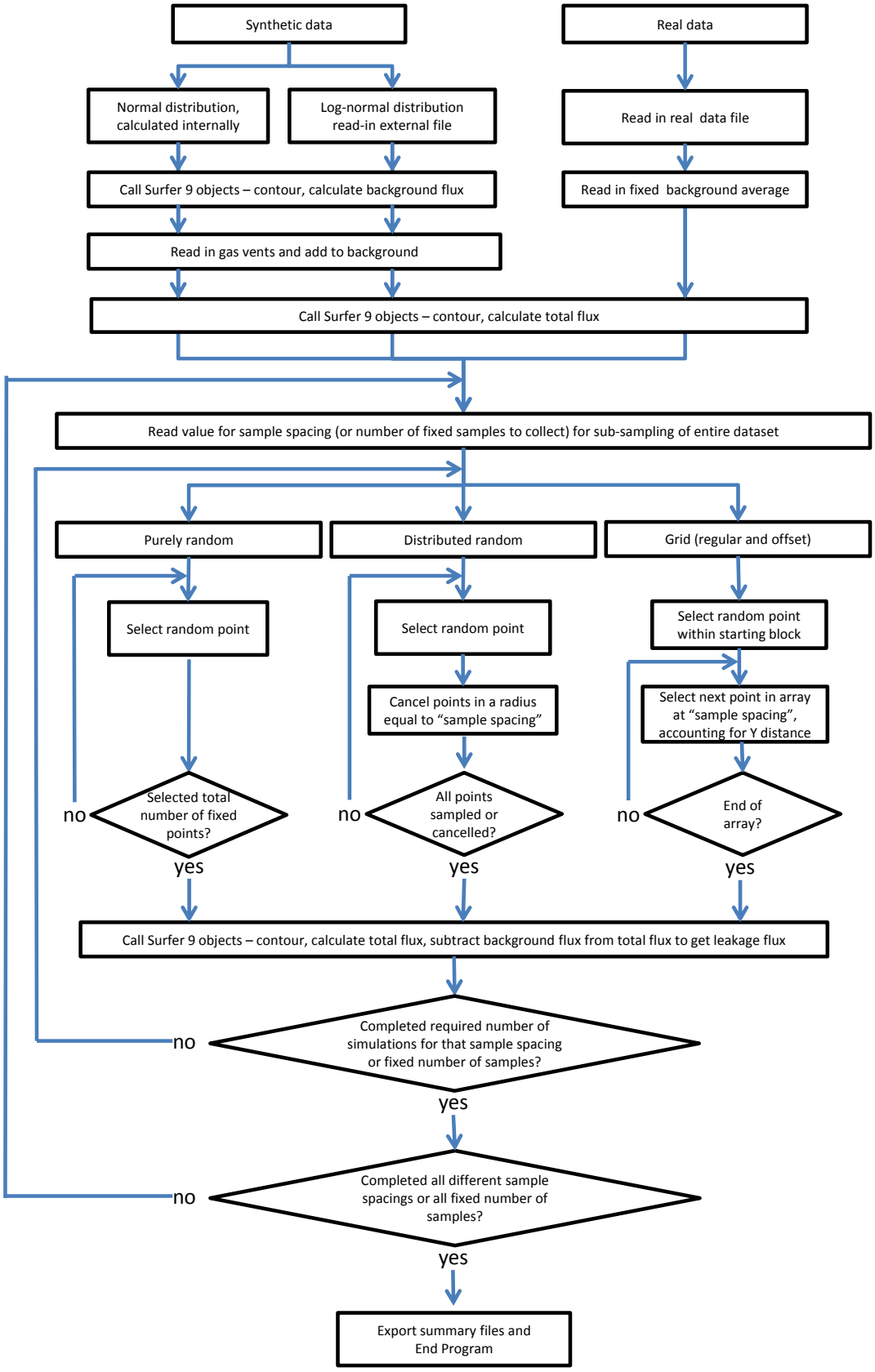


Figure 26. Flowchart showing the general algorithm and control logic of the GasGrid program.

To accomplish this a digital elevation model (DEM) of an arbitrary area in northern Canada was downloaded from the Global Data Explorer (GDEX) site of the United States Geological Survey (USGS) (ASTER GDEM, 2014), transformed from the true 25 m node spacing to a 1 m node spacing for a 1000 x 1000 m grid with corresponding 1 to 1000 Cartesian coordinates, and the entire 3 column dataset (i.e. X, Y, elevation) was ordered from lowest to highest elevation. The PopTools-generated flux data was then ordered from highest to lowest and the two datasets merged, resulting in a gridded array with high flux corresponding with low elevation. At this point the elevation data was cancelled, the remaining three columns were re-ordered based first on increasing Y and then on increasing X coordinate values, and then the X and Y data columns were cancelled, leaving a spatially coherent flux array that was saved as a comma delimited, single column text file for eventual import into the GasGrid program (Figure 27b). A rigid naming convention was adopted for this file to allow the program to read in dataset characteristics at the time of import, such as: DEM1a020s005x1000y1000tLN.txt, where DEM1 refers to the downloaded DEM, a020 is an average flux of 20 g m^{-2} d^{-1}, s005 is a standard deviation of 5 g m^{-2} d^{-1}, x1000y1000 gives the size of the grid, and tLN defines this as a log-normal type of distribution. Because these values are read into the program at the time of importation, these parameters are fixed and cannot be subsequently changed in the data input window (Figure 25).

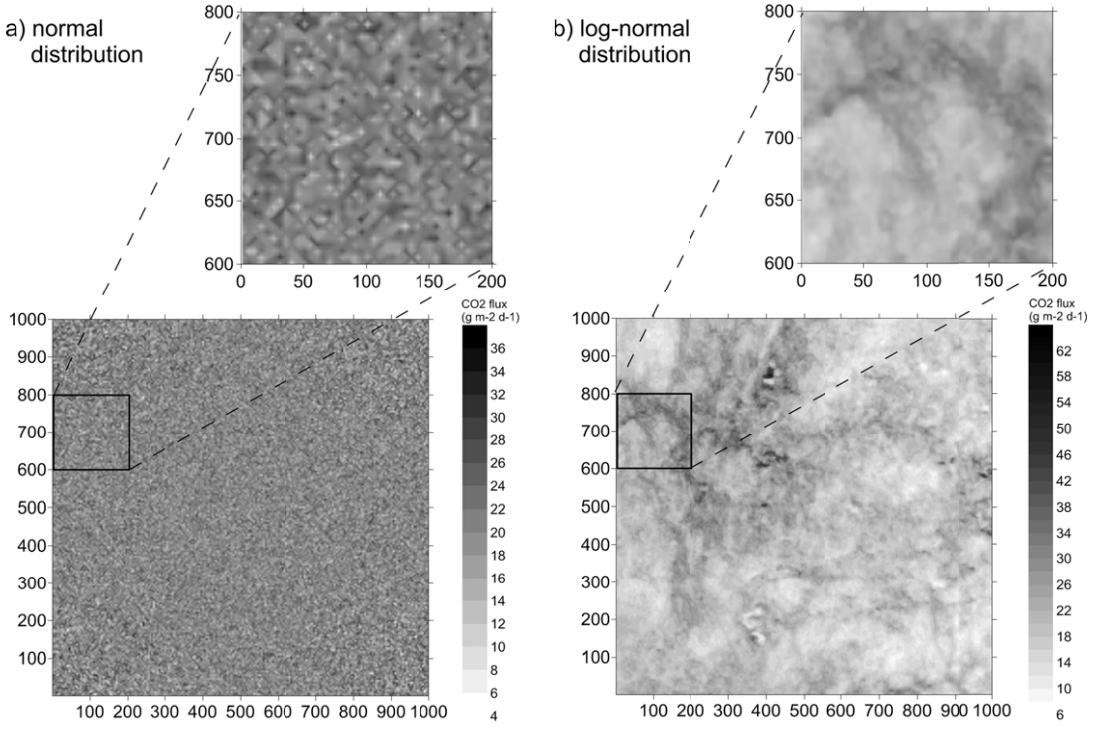


Figure 27. Contoured CO_2 flux examples of a random generated and random dispersed, normal-distribution synthetic background field (a) and a random generated, spatially distributed, log-normal synthetic background field (b).

One or multiple gas leakage areas ("gas vents") can be superimposed on either type of simulated background flux fields, with the method used depending on whether the goal of the simulation is to examine how different sampling strategies can influence our ability to i) find a leak or ii) accurately estimate its total flux.

In the case of the former, the selection of the "find anomaly" option bypasses code relating to Surfer9 and associated contouring (thus making it faster), and assigns the maximum vent flux value to all points within the vent radius. During simulations each sub-sampled point value is then compared against the maximum background value for the entire array; if it is higher, a vent anomaly counter is incremented. For these simulations the selection of a high vent flux (e.g., 1000 g m^{-2} d^{-1}) versus a moderately high background maximum (e.g., 60 g m^{-2} d^{-1}) allows for a clear distinction (Figure 28) and an accurate counting of vent-intersecting sample points, to look at the probability of finding a leak of a given size in a given survey area. In this case, for simplicity, only a single vent is created for each simulation.

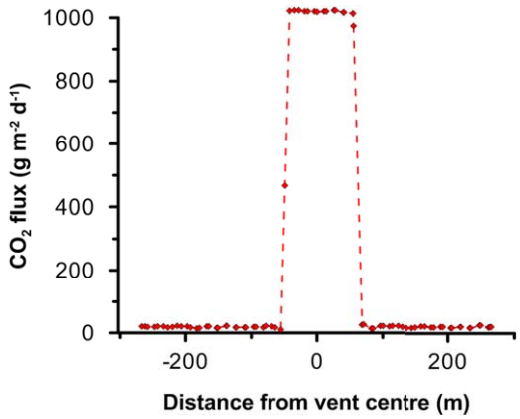


Figure 28. A simulated gas vent for the "find anomaly" option, in which the maximum flux is assigned to all points within the vent radius to create an clear distinction between leakage and background values.

If, instead, the goal is the latter, i.e. to study the effect of sampling strategy on total leakage estimates, then the "calculate flux" option is chosen. Here the entire code is used, and the creation of vents are treated differently. The number, location, maximum value, and diameter of the gas vent(s) are chosen by the user at the time of the simulation, whereas the flux value for each point as one moves away from the vent core towards the outer vent edge is calculated using the following empirically derived formula:

Eqn 14 $$\textbf{flux} = \exp(-(0.065 * \ln(F_{vm}) - 0.1525)*((20 / R_v)*D))*V_{mf}$$

where F_{vm} is the vent maximum flux, R_v is the vent radius, and D is the distance between the point being calculated and the vent centre. This formula was derived using the results of CO_2 flux profiles performed across different gas vents in the Latera caldera. Figure 29a and Figure 29b compare the real measured data against the distribution calculated using Eqn 14. The general trends of the real and calculated values show a good level of similarity. Figure 29c instead shows a range of possible distributions for vents having maximum flux rates of 1000 and 200 g m^{-2} d^{-1} and radii of 20 and 56 m. These four trends are presented here to give an idea of possible distributions, and also because some of these settings are used in the simulations discussed below.

Addition of the synthetic gas vents involves locating the centre of the gas vent within the background array and assigning the user-defined maximum flux rate to that point. Subsequently a grid of points surrounding the vent centre is queried sequentially to see if they lie within the user-defined vent radius. If affirmative, Eqn 14 is used to calculate a leakage flux rate for that point, and the resultant value is added to the already existing background value. Once complete the circular vent is defined by a lateral gradient that is the same in all directions around the

central point (Figure 30a, b). If two vents overlap, the first vent is created and then during the creation of the second the array is queried as to whether the point has already been previously changed; if yes, then the higher of the two values (i.e. the values calculated for that point for the two vents) is used.

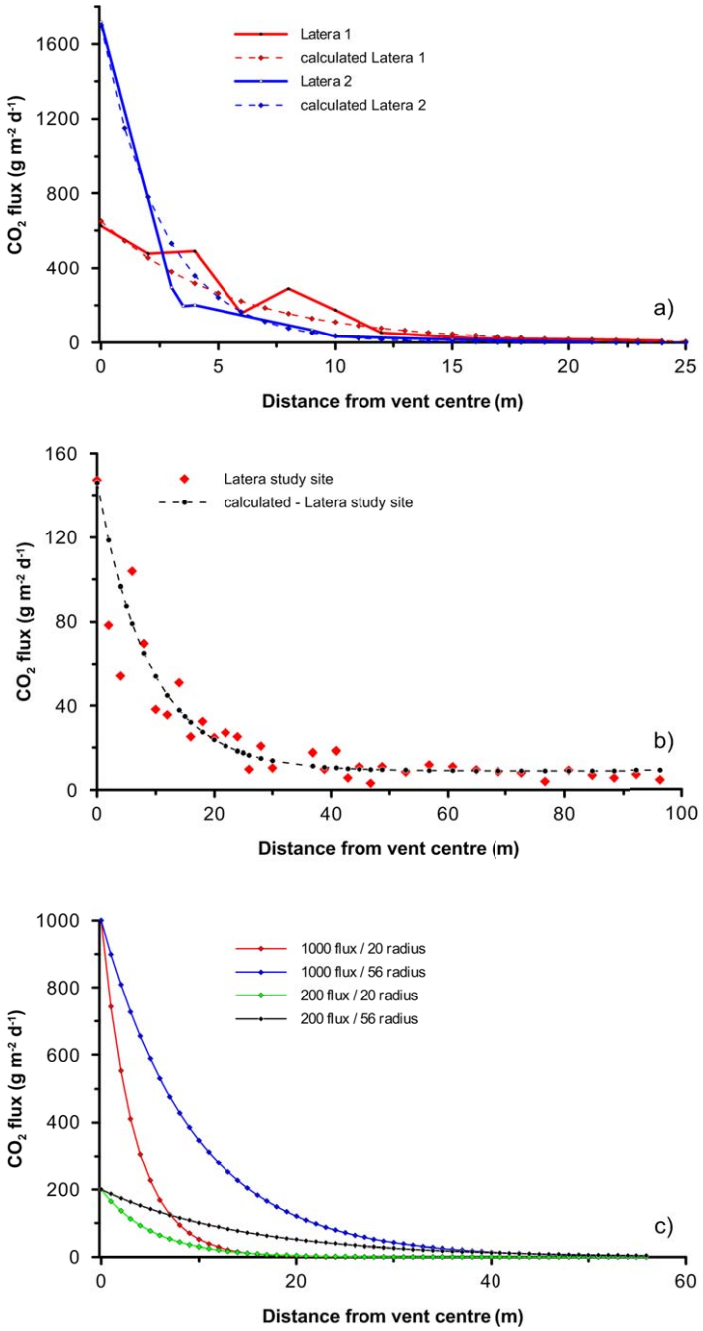


Figure 29. Rate of change in CO_2 flux values moving away from a vent centre at distance 0. a) comparison between real data from two vents within the study field and the estimated values calculated using the developed empirical formula; Latera 1 and 2 were calculated using maximum flux rates of 1700 and 650 g m^{-2} d^{-1} and radii of 17 and 30 m, respectively. b) same as (a) but for a gas vent with a lower flux rate located in a different field (western extreme of dashed line in Figure 17); based on a maximum flux at the edge of the vent of 145 g m^{-2} d^{-1}, a radius of 30 m, and the addition of a constant background value of 10 g m^{-2} d^{-1}. c) examples of different calculated distributions for two maximum flux rates (1000 and 200 g m^{-2} d^{-1}) and two vent radii (20 and 56 m).

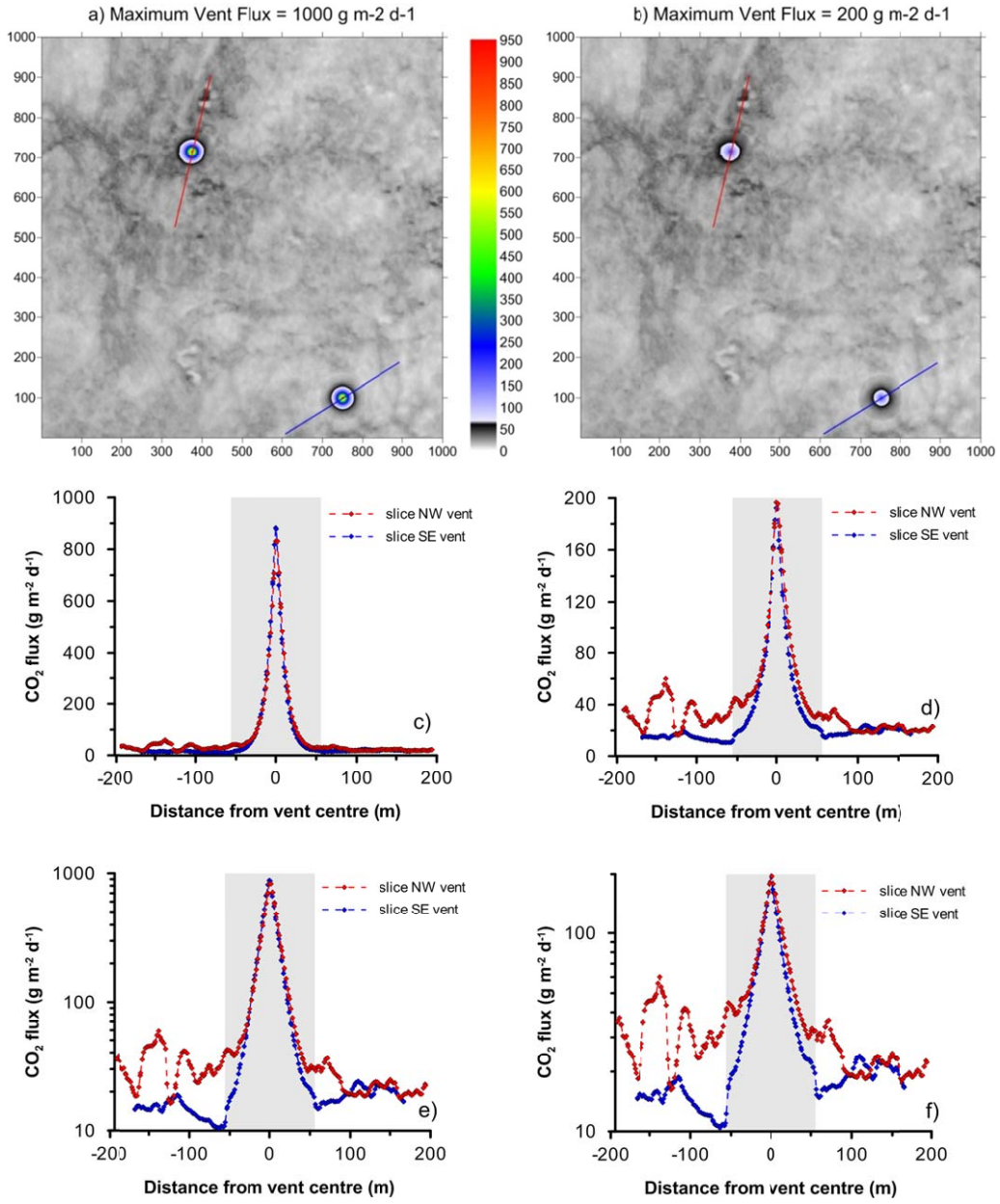


Figure 30. Simulated vents and flux background, as contour maps (a, b) and scatter plots sliced from the grids (c to f), with the vents having a maximum flux rate of 1000 g m^{-2} d^{-1} (left column) and 200 g m^{-2} d^{-1} (right column). Vent locations were chosen to illustrate the same vent surrounded by a high (NW vent) and low (SE vent) background field. Note in the scatter plots that the vent interval is highlighted in grey, and that plots (c) and (d) are normal Y scale while (e) and (f) are the same data on a logarithmic scale.

To illustrate how the created gas vents fit in the surrounding background flux field, two vents have been placed in a high background (NW corner) and a low background (SE corner) part of the grid (Figure 30a, b), with both vents being assigned a maximum flux of 1000 g m^{-2} d^{-1} (Figure 30a) or 200 g m^{-2} d^{-1}(Figure 30b). These vent areas have been "sliced" along the traces shown using a function in Surfer9, and the CO_2 flux values along both lines have been plotted on both normal (Figure 30c, d) and logarithmic scales (Figure 30e, f). Note that whereas the vent flux values in the low-background area are consistently higher than the surrounding values, the same vents in the high background area may have values that are actually lower than some

surrounding background values. This is important, as "anomalous" background and leakage values may be combined during the contouring of low-resolution subsampling thus falsely uniting anomalies of different origins and potentially over-estimating the size and total flux of a leakage area.

If real data, instead of synthetic data, is chosen, an external file is read into the program. This file must be a single column text file containing only the flux data, ordered sequentially for each X point along progressive Y lines. As this is real data there is no need to generate a background distribution. Instead only the fixed average option is available, which is a number selected prior to simulations based on population definition using normal probability plots. Within the program, this value is multiplied by the entire surface area of the grid to give an estimate of the total background flux, following the approach taken by Chiodini et al. (2007) and others. Clearly for real data there is also no need to generate artificial gas vents, and as such the Vent Setting form in the input window is disabled. As it is not feasible to collect real data at a 1m sample spacing for a large area, in the present study samples were collected at a 10 m spacing (see section 4.4), a square grid was created by assigning background values outside of the actual sampled polygon, and then a 1 m grid of values was created using simple Kriging interpolation within the program Surfer9 prior to import in the program.

5.2.2.2. Sub-sampling approaches

As stated above, four different sampling approaches have been included within GasGrid, purely random, uniform random, regular grid, and off-set grid. These different approaches and how they are implemented are described below.

The purely random case uses the random number generator within Visual Basis 6 to sequentially select a fixed number of samples from the total array. As each point is selected it is marked to ensure that it is not sampled a second time. Being purely random, each point has equal possibility of being selected, and as such there is no control over spatial distribution or distance between samples. Figure 31a shows how this can result in significant clustering of points and can leave other areas with no samples at all.

The uniform random approach again sequentially selects a random point from the entire array, however once chosen all the points in the "sample spacing" radius defined for that simulation are cancelled from the array, eliminating the possibility that they can be chosen in subsequent iterations of the code. The program loops until all points in the grid are either selected or cancelled. In this way the resultant points chosen are completely random, but they are forced to be relatively uniformly distributed over the entire sample grid (Figure 31b), with no clustering and no sample gaps. Although the "sample spacing" parameter is used in this process, it may be more correct to refer to it as a minimum sample spacing. This is explained with reference to Figure 32, where three points have been chosen by the program based on a "sample spacing" of 10 m. In this example, the first point at X=50, Y=50 is chosen and the surrounding points are cancelled. The next point at X=50, Y=60 is located at the edge of the first cancelled area, and thus the distance between points 1 and 2 is equal to the 10 m sample spacing. However the third point, at X=70, Y=50, results in very little overlap between the cancelled areas around points 1 and 3, thus ensuring that no points can be selected in this interval and that the distance between these points will be 20 m. As such the average sample distance is likely on the order of $(r + 2r)/2$. For a given "sample spacing" value, each simulation will select a similar, but not necessarily equal, number of samples.

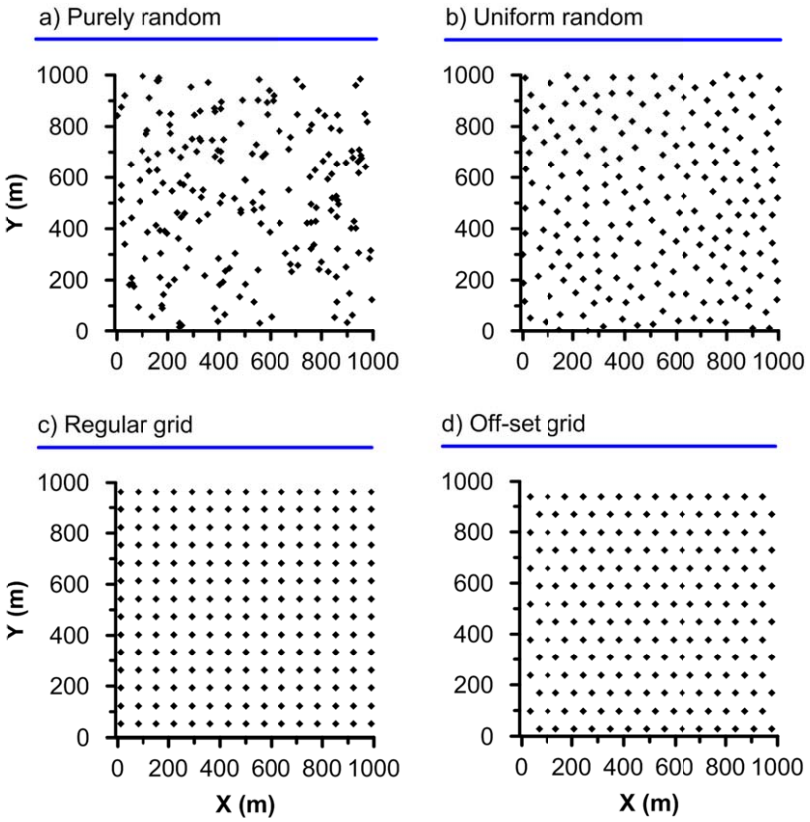


Figure 31. An example from each of the four sampling strategies available within the GasGrid program. Each plot shows about 200 points for each subsampling, out of a total matrix of 1,000,000 points.

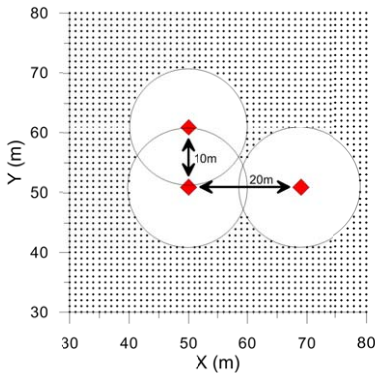


Figure 32. Detail of a grid sampled using the uniform random technique, showing how distances between samples can range from the "sample spacing" parameter value to twice that value.

The regular grid approach chooses one point at random within a subset area defined by (X = 1 to sample spacing) by (Y = 1 to sample spacing). Once this first point is chosen each subsequent point is selected from the array based on a regular grid, with sample distance being fixed in the X and Y directions by the sample spacing variable defined by the user (Figure 31c). Note that the initial subset area is used to choose the first point because the use of a rigid grid means that shifting the grid along the X or Y axis beyond the sample spacing distance will result in repetition of the same grid points. Because of this limitation, the number of unique simulations

using this approach will be limited by (sample spacing)^2, such that a maximum of 100 unique simulations can be conducted with a sample spacing of 10 m, 400 for 20 m, 900 for 30 m, etc. Compared to the other two approaches, the regular grid method gives a fixed sample spacing and the number of samples for each simulation at a given sample spacing will typically yield 2 or three values (depending on how close to the boundary the sampling grid is). Finally the off-set grid option uses the same algorithm as the regular grid, with the only difference being the shifting laterally by (sample spacing / 2) on every second row.

5.2.2.3. Data contouring and flux calculations using Surfer 9

Four contouring / interpolation methods were included as options within GasGrid, each being applied by calling Surfer 9 objects:

1) The Inverse Distance to a Power gridding method is a weighted average interpolator that can be either exact or smoothing. With this interpolation method, data are weighted such that the influence of one point relative to another declines with distance from the grid node. Weighting is assigned to data through the use of a user-defined weighting power that controls how this factor drops off further from a grid node. One of the characteristics of this method is that it tends to produce "bull's-eyes" surrounding the grid observation points, which can be reduced by adjusting the smoothing parameter.

2) Natural Neighbour (NN) interpolation finds the closest subset of input samples to a query point and weights them according to proportionate areas to interpolate a value. This method is local as it only uses the subset of samples that surround the grid point, such that interpolated values will be within the range of the samples used. The natural neighbours of any point are those associated with what are known as the neighbouring Thiessen polygons. The resultant surface passes through the input samples and is smooth everywhere except at locations of the input samples. It adapts locally to the structure of the input data, works well with regularly and irregularly distributed data, and requires no user input for search radius, sample count, or shape.

3) Radial Basis Function interpolation is a diverse group of data interpolation methods, all of which are exact interpolators (i.e., they attempt to honour the original data). This method is conceptually like a rubber membrane that is fitted to each of the measured data points while minimizing the total curvature of the surface. There are five different basis functions that can be chosen, including multilog, inverse multiquadric, multiquadric, natural cubic spline, and thin plate spline.

4) Kriging, a common geostatistical gridding method, is an optimal interpolation based on regression against observed z values of surrounding data points, weighted according to spatial covariance values. These covariance values are represented by the variogram, which is a function that describes the spatial variation of the gridded parameter. The default linear variogram often provides acceptable results, although data-specific variograms can be constructed. The kriging equation is essentially the same as that used for inverse distance weighted interpolation, except that rather than using weights based on an arbitrary function of distance, the weights used in Kriging are based on the variogram. Kriging tends to result in more smoothed contour results, is very flexible, and allows a variety of map outputs including predictions, prediction standard errors, probability, etc.

The GasGrid program uses Surfer9 to interpolate three different types of datasets, including the complete synthetic background dataset, the complete synthetic background plus gas vents dataset (or real data), and the various sub-sampled datasets of the synthetic background plus vent data (or real data). At each stage the total flux for the survey area is calculated using a function in Surfer9 called GridVolume. This method is typically applied to calculate the volume of geological unit, multiplying each grid cell (in units of m^2) by the height difference (m) between upper and lower bounding surfaces. Instead, by setting the lower boundary as a constant flat surface equal to zero and the upper boundary as the contoured grid of flux values in g m^{-2} d^{-1}, the calculating formula becomes $\sum_1^n (A_c * \varphi_c)$, where n is the total number of grid cells, A_c is the surface area of each cell (m^2), and φ_c is the CO_2 flux calculated for cell n (g m^{-2} d^{-1}), thus yielding the total flux for the entire grid area (g d^{-1}). For these calculations the trapezoidal rule option was used.

Table 1. Cross-validation parameters monitored during the gridding of the sub-sampled datasets; "array index" refers to the array of statistical results generated by the Surfer9 Cross.Validate command.

Category	Array index	Explanation
Estimation Error Statistics	srfCV_R_RootMeanSquare	root mean square
Spatial Regression Statistics for Estimation Error	srfCV_R2	coefficient of multiple determination error
Inter-Variable Correlations	srfCV_ZE_Correlation	correlation between Z and estimated statistics
	srfCV_ZR_Correlation	correlation between Z and estimation error
Elevation / Estimation Statistics / Estimation Error Rank Correlations	srfCV_ZE_RankCorrelation	rank correlation between Z and E
	srfCV_ZR_RankCorrelation	rank correlation between Z and R

Finally, each time that a sub-sampled dataset is generated and contoured, the quality of the gridding of that dataset is assessed by performing a cross-validation. Cross-validation, which in Surfer9 can be used with all gridding methods, involves the sequential removal of each individual point from the original dataset and interpolation of a value for that point using all remaining data and the specified interpolation algorithm. The interpolation error is then calculated for that point as "interpolated value – original value". By repeating this procedure N times for the N points in the grid array (i.e., removing only one point, performing the interpolation, replacing the point, repeat for next point), a total of N interpolation errors are generated. Various statistics computed for the errors can be used as a quantitative, objective measure of quality for the gridding method. These statistical values are saved in a Surfer9 array called srfCVResults which contains 62 parameters. For the present work six of these parameters (Table 1) were chosen for output by the GasGrid program.

5.2.2.4. Boundary effects

Initial testing of the program using the contouring and flux calculation option highlighted a potential source of error in the automatic approach needed for the creation of large numbers of simulations. Under the conditions of a vent close to a boundary, a sub-sample within the vent, and no samples between the vent sample and the grid boundary (most common in low sample density simulations), the contouring algorithm extended the area of the high flux values over a large, completely un-realistic area along the edge of the boundary (e.g. Figure 33a). Although

most prevalent in the grid-types of sampling, this effect was also observed in the other sampling strategies. Although a common effect in contouring spatial data that is easily overcome by blanking the grid area beyond the edge of the sample points, the large number of automated simulations, each of which is slightly different, did not allow for the use of this method. In addition, by blanking a part of the grid a certain area would be excluded in the total flux calculations.

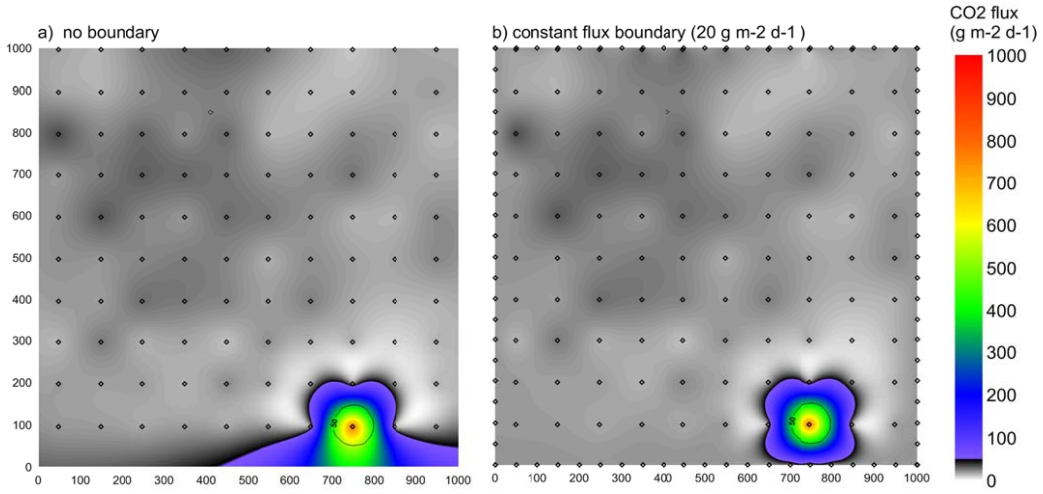


Figure 33. Contouring results of one simulation consisting of 100 m sample spacing. a) raw data. b) the same data, but with a value equal to the background average (20 g m^{-2} d^{-1}) assigned to points every 50 m around the edge of the grid to eliminate contour boundary effects. The outline of the "true" leakage area is marked, for reference, with the black circular isoline labelled as 50 g m^{-2} d^{-1} in both images.

To address this problem a different approach was used, which involved adding grid nodes to the exported sub-sampled data file at 50 m intervals around the entire edge of the grid and assigning them a fixed flux value equal to the average background flux, in this case 20 g m^{-2} d^{-1} (Figure 33b). This figure shows that the imposed boundary condition has little or no impact on the background values calculated throughout the grid, while at the same time greatly constrains the spatial extent of the contoured leakage area. This approach thus allows one to examine the influence of sampling density on contouring and total flux calculations without the periodic occurrence of artefacts in some simulations. For example, the total flux (i.e. background plus leakage) calculated for the plots in Figure 33 was equal to 32.2 tonnes CO_2 d^{-1} for the original data and 27.6 tonnes CO_2 d^{-1} for the data with the imposed background boundary, equivalent to a difference of about 15%. Note that the added boundary nodes were not used in the calculation of the number of samples collected per simulation, nor were they used for any statistical analysis of the sub-sampled datasets.

5.3. Synthetic data modelling

5.3.1. Leak detection

As outlined in section 2.4.1, the point nature of soil gas and gas flux measurements means that the potential for successfully locating a leak will depend on the number of samples collected, the size of the leak relative to the total survey area, and the ability of the measurement method to

define with 100% certainty if an anomaly is due to a leak or not. In addition to these factors, the sampling strategy can also influence survey success. For example, Eqn 12 predicts the probability of finding a leak for a given number of collected samples, however this is assuming that the sample locations are chosen in a purely random manner. To better understand how sampling strategy can affect survey success, a series of GasGrid simulations were performed using the four different sampling strategies available (i.e., purely random, uniform random, regular grid, and offset grid) on the five leakage shapes and orientations shown in Figure 34. Note that all five shapes have the same surface area (10,000 m^2) and thus the same x/A ratio of 0.01 (see Eqn 12).

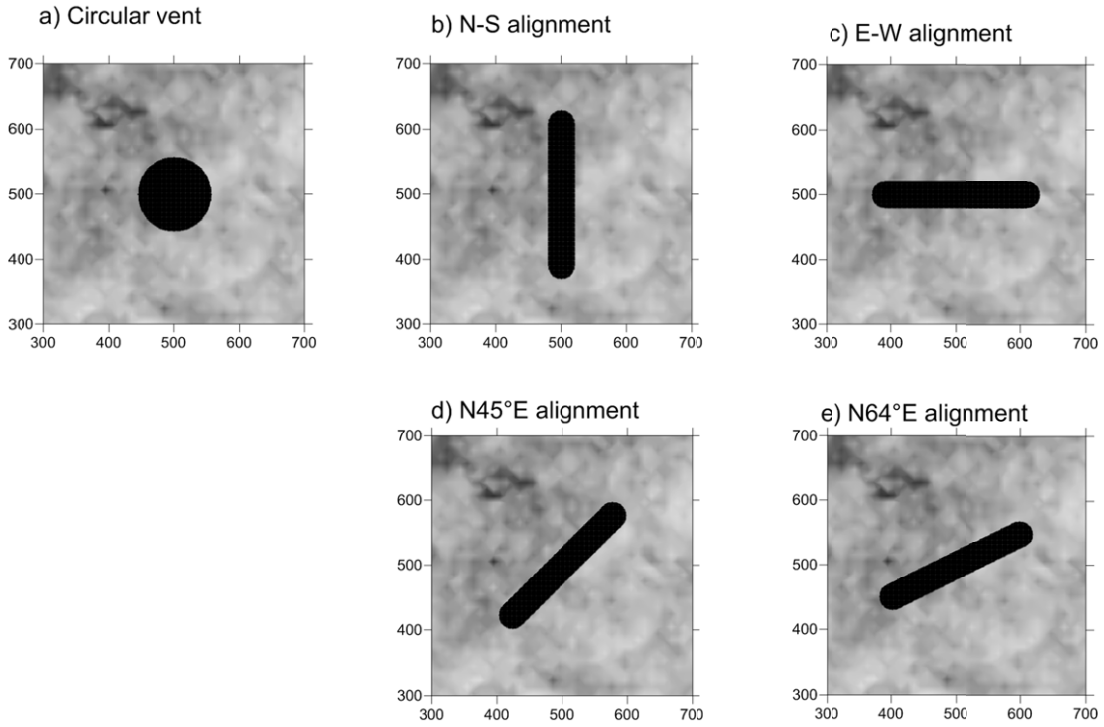


Figure 34. Plots showing the different leakage shapes tested to understand the influence of the chosen sampling strategy on the success rate of finding an anomaly. a) represents a circular gas vent while b) to e) represent different orientations of an aligned anomaly (such as along a fault trace). Note that all 5 shapes have the same surface area of 10,000 m^2.

Figure 35 shows the results of a series of simulations conducted on the single, circular gas vent. The following procedure was used for this and all subsequent tests presented in this section. First simulations were conducted choosing the "find anomaly" option, which bypasses Surfer9 calculations, assigns the maximum flux rate to each point in the defined gas vent, and keeps track of whether a sampled point is within the leakage area or not. For each given "sample density" (i.e., sample spacing for grid and offset grid, minimum sample spacing for uniform random, or fixed number of samples for purely random), a total of 500 simulations were performed. Once complete the number of samples that were within the leak anomaly for each simulation were converted into a binary value (1 if one or more points were in the vent and 0 if none were) and the percentage of success was calculated for the first 100, 200, 300, 400, and 500 simulations.

Figure 35, which presents the results from all 500 simulations, shows an excellent agreement between the theoretical probability of success using purely random sampling based on Eqn 12 and the GasGrid simulation results for the same sampling strategy. As stated by Oldenburg et al.

(2003), approximately 300 samples are needed to have a 95% confidence of sampling, at least once, a gas leak with the modelled characteristics. Interestingly, however, the other three methods require much fewer samples for the same confidence level, with about half as many needed for uniform random sampling (147) and even fewer for regular grid (118) and offset grid (100) sampling. The reason for this difference can be seen examining Figure 31, which shows how purely random sampling can leave large areas un-sampled where a leak may occur. This effect is much reduced in the uniform random approach, although the slight variability of sample distances can result in smaller un-sampled areas. Instead the grid approaches yielded the best results because the sampling distances are rigidly uniform.

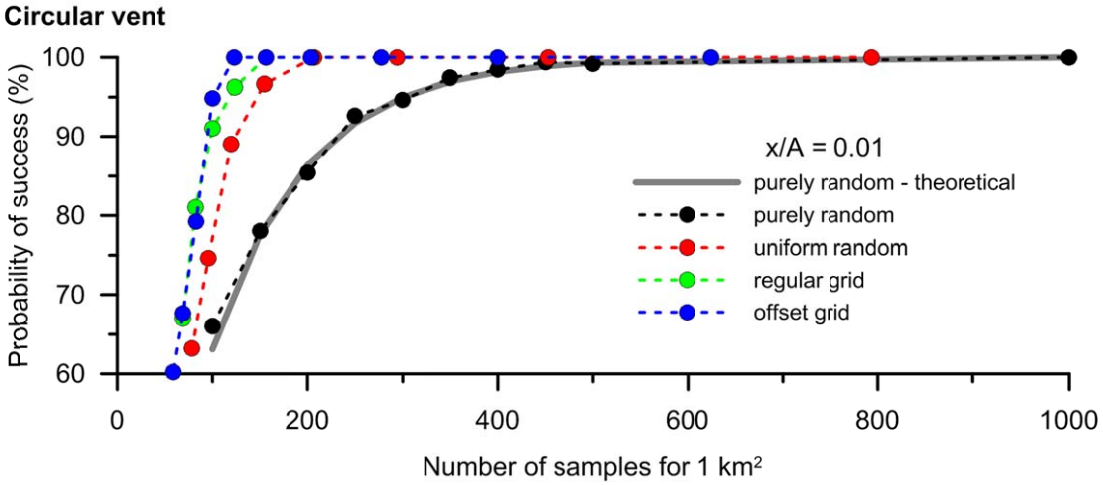


Figure 35. Comparison of the number of samples required to locate a gas leak using different sampling strategies, based on 500 simulations and a circular gas leakage area "x" of 10,000 m^2 (i.e., a radius of 56 m) within a total survey area "A" of 1,000,000 m^2 (i.e. 1000 x 1000 m). Note the good agreement between the theoretical and program-generated distribution for purely random sampling.

To assess the impact of the number of simulations on the obtained results, the data presented in Figure 35 have been re-plotted in Figure 36, but this time with the results of each of the four different sampling strategies divided in blocks of the first 100 simulations, the first 200, etc. up to all 500 simulations. The purely random simulations are given in Figure 36a, together with the theoretical distribution. This plot shows how the simulation response stabilises towards the theoretical one only after about 300 simulations, in contrast to the 200 needed for consistent results in the uniform random (Figure 36b) and offset grid simulations (Figure 36d), and the stability observed right from 100 simulations for the regular grid simulations (Figure 36c). While all tests reported in this section were conducted with 500 simulations each, because of the much longer processing times needed for simulations involving contouring with Surfer9, a compromise of 200 simulations was chosen (based on these results) for the simulations presented below in section 5.3.2. The only exception to this is the regular and offset grids having a 10 m sample spacing, for which only a maximum of 100 simulations can be performed (as explained above in section 5.2.2.2).

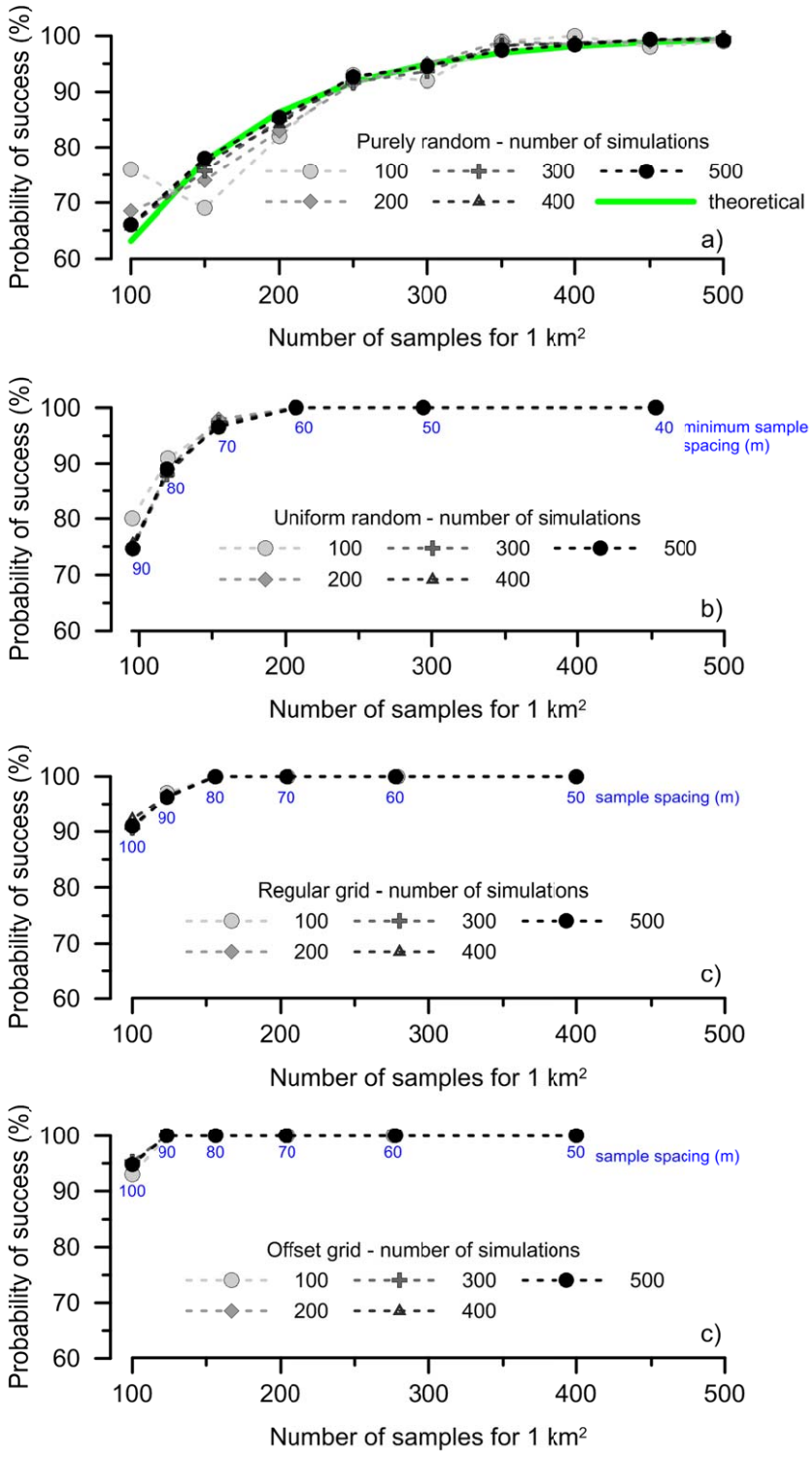


Figure 36. For each sampling strategy the probability of success versus the number of samples collected over 1 km^2 is compared for the number of simulations conducted. For random sampling over 300 simulations are needed to approach a constant (i.e. statistically stable) result, as compared to 200 for uniform random and offset grid, and only 100 for regular grid. Approximate sample spacing is given for the latter three plots for reference.

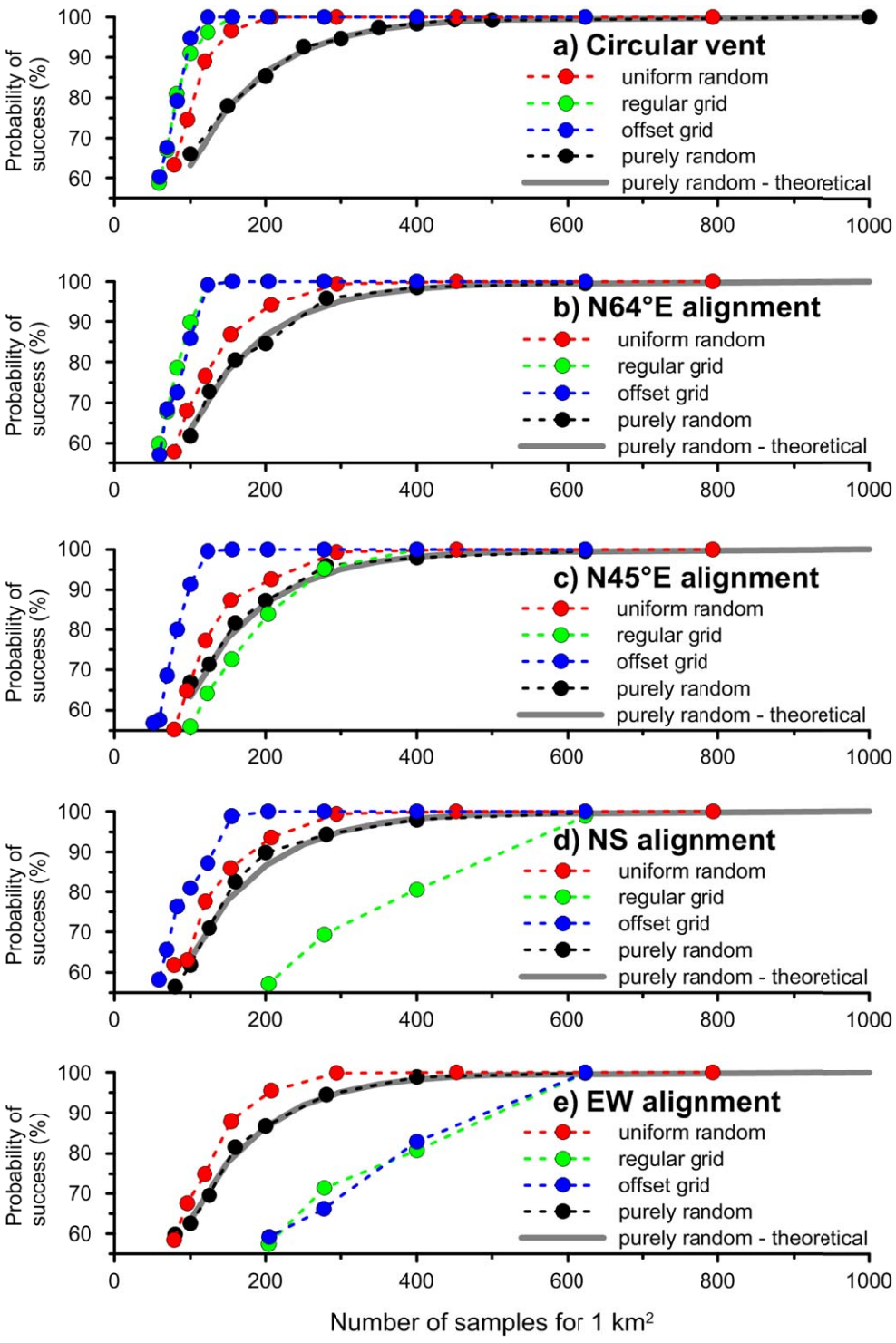


Figure 37. Comparison of probability of sampling at least once a gas leakage anomaly of different shapes and orientations using the four different sampling strategies available in the GasGrid program. All leakage areas are equal to 10,000 m^2, for an x/A ratio of 0.01, and each data point represents 500 simulations.

Table 2. Number of samples needed to intersect at least once, at the 95% confidence level, an anomaly having a surface area of 10,000 m^2 in a total 1 km^2 survey area, comparing different anomaly shapes and different sampling strategies.

	Offset grid	Regular grid	Uniform random	Purely random
Circular vent	99	117	149	300
N64°E alignment	115	113	221	300
N45°E alignment	109	276	236	300
NS alignment	140	575	231	300
EW alignment	560	563	206	300

Results for all simulations using the four sampling strategies on the five different leakage areas are summarised in Figure 37, with the results for the circular gas vent given in Figure 35 reported again to facilitate comparison with the other tests. In addition the number of samples needed to intersect the studied anomalies, at the 95% confidence level, using the various sampling strategies given in Table 2.

As expected, the purely random sampling approach yielded exactly the same results in all five tests because the leakage surface area was the same for all 5 anomalies (Table 2), despite the different leakage shapes and orientations. In contrast, the uniform random method gave much better results than the purely random approach for the circular vent (149 vs 300 samples, respectively, at the 95% probability level) (Figure 37a), whereas for all the aligned anomaly tests it gave slightly poorer success rates (Figure 37b - e) that were still consistently higher than the purely random approach (typically requiring 65 to 95 samples fewer at the 95% level; Table 2). The fact that the uniform random method consistently requires fewer samples to obtain the same results is because this method does not allow for data clustering, which can cause gaps in the areas sampled.

The two grid sampling approaches often yielded the highest success rates, although these approaches are much more affected by the orientation of elongated anomalies. For the circular vent the offset grid method gives very slightly better results than the regular grid, due to the geometry of the anomaly relative to the gaps in the two different grids. This is illustrated in Figure 38, where the 56 m radius vent used in the simulations above is plotted together with hypothetical sample point distributions for the regular (a) and offset (b) grids, both having the same horizontal sample spacing. In this example, the regular grid would miss the anomaly whereas the offset grid would have one point within the anomaly. Based on this sample spacing, the smallest vent that would be intersected with 100% certainty is a function of the diagonal distance "c". For the regular grid method, assuming that x=y, this value corresponds to $c = \sqrt{2(x)^2}$, whereas for the offset grid approach it corresponds to $c = \sqrt{(x/2)^2 + (y)^2}$. Substituting an 80 m sample spacing in the regular grid equation yields c = 113 m, which corresponds to the diameter of the smallest vent that could be observed with 100% certainty at this sample spacing. Instead, substituting the same value in the offset grid equation yields c = 90 m, which corresponds to a vent radius of about 97 m. This means that with the same sample spacing (and thus number of samples), the offset grid method is capable of locating circular gas leakage areas that are about 14% smaller in diameter. For a 10 m offset grid, like that used at Latera, this would correspond to a vent diameter of about 12 m.

Re-writing the respective equations in terms of the sample spacing ("x" or "y") and using a vent diameter equivalent to that in the above simulations (i.e., radius of 56 m, diameter of 112 m), we

can estimate that the smallest sample spacing needed to locate the simulated vent with 100% certainty would be 79 m for the regular grid and 90 m for the offset grid, in good agreement with that observed in the simulation results given in Figure 36c and Figure 36d, respectively (see numbers in blue for sample spacing, and the X axis for the corresponding number of samples).

In contrast to the search for a symmetrical anomaly like that discussed above, Figure 37b – e shows that the grid sampling approaches can yield significantly different results depending on the orientation of the same sized and shaped elongate anomaly. For the case of a small oblique angle between the grid and the anomaly, such as that at N64°E (Figure 37b) both grid approaches yield similarly high success rates. Instead, at an alignment of N45°E (i.e. NE-SW), the offset grid approach maintains a similar level of success however the regular grid decreases to the level of the purely random sampling approach (Figure 37c). This can be understood by examining Figure 39 a and b, where it is shown schematically how the anomaly is perfectly aligned within a gap in the regular grid which does not exist in the offset grid. Although having a slightly lower success rate, the offset grid results for the NS alignment case is still quite high, whereas the success rate of the regular grid has decreased to a quite low level (e.g., 575 samples needed at the 95% confidence level; Table 2). Figure 39c illustrates how the anomaly aligns along the y-axis-parallel gap in the grid, and how success is controlled by the width of the anomaly (here only 20 m) with respect to the grid sample spacing. Instead Figure 39d illustrates how the offset of alternating rows halves this vertical gap, thereby increasing the potential for success. Finally, both grid approaches give equally poor results for the EW alignment (Figure 37e) because both approaches have the same x-axis-parallel gap (Figure 39 e and f).

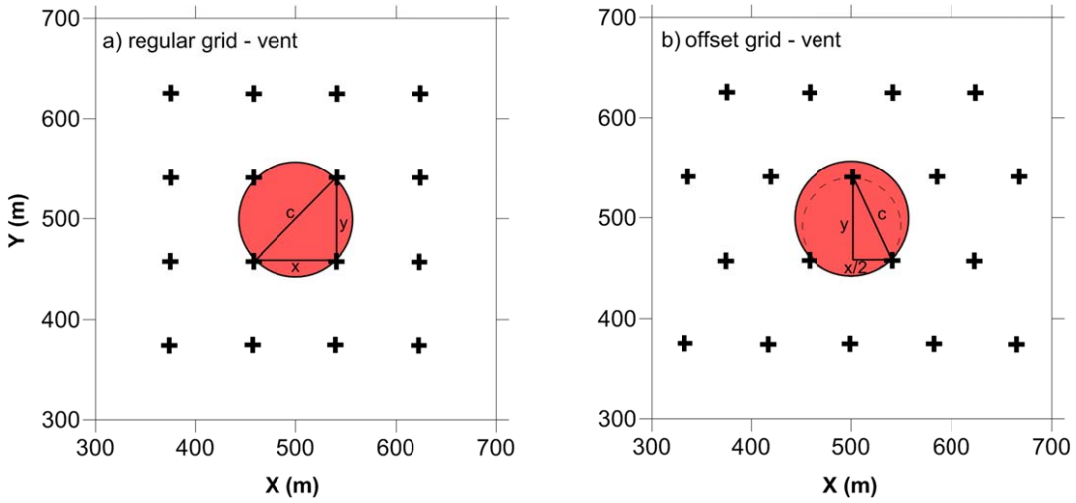


Figure 38. Relationship between the diameter of a circular gas vent and the same sample spacing for a regular (a) and an offset (b) grid sampling strategy.

In summary, these results indicate that the offset grid approach gives the highest probability of success compared to the other sampling strategies. It must be highlighted, however, that if there is the potential for long, narrow anomalies, such as may occur along the trace of a fault, it is important that the shift in the alternating rows takes place in the direction perpendicular to the strike of the anisotropy, to obtain results similar to the NS alignment simulations and not those of the EW alignment simulations.

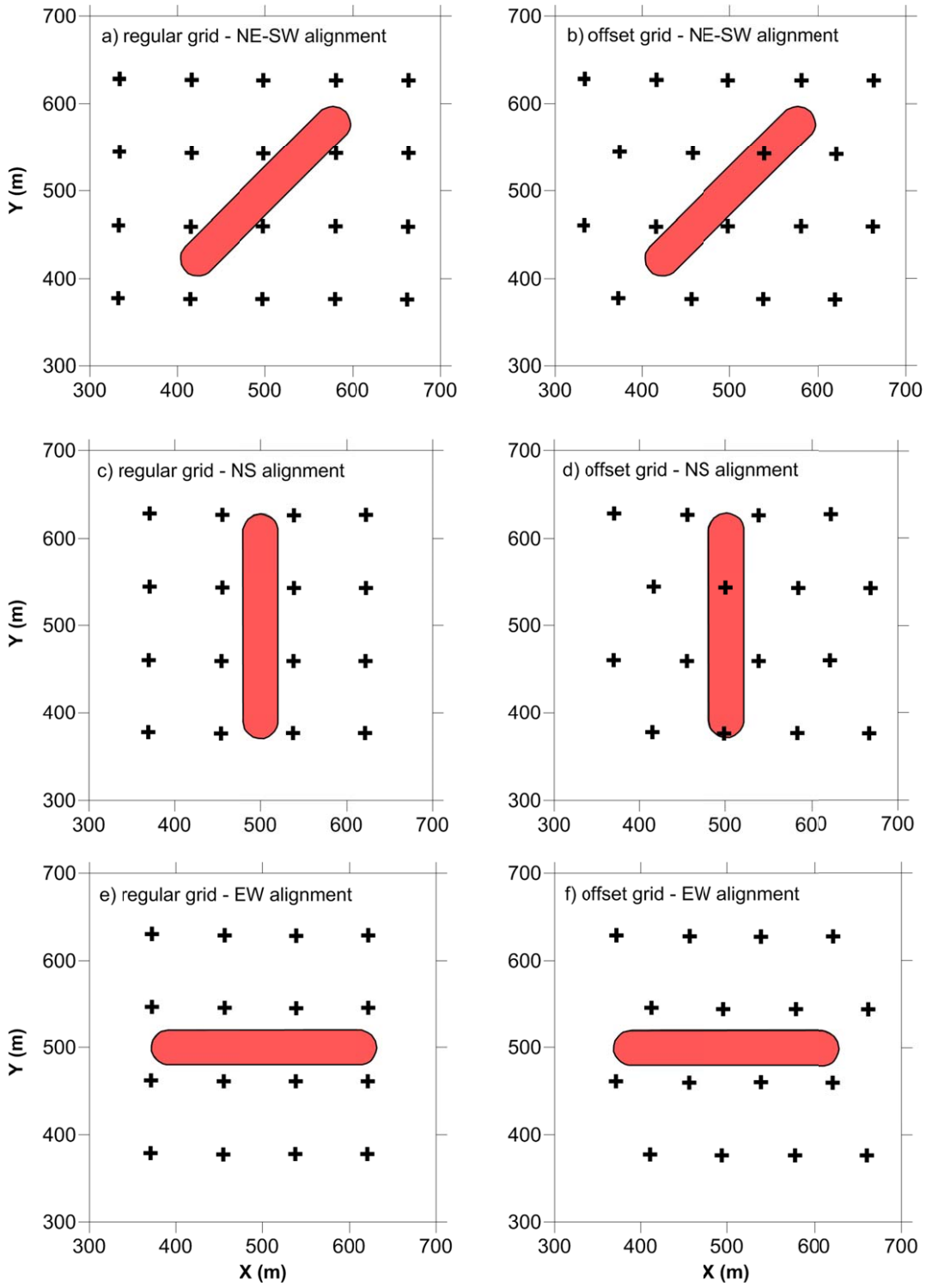


Figure 39. Series of schematic diagrams showing how the orientation of an elongate anomaly influences the potential success of a regular grid survey (left column) and an offset grid survey (right column). Note that while the offset grid has a higher potential of success in the top two examples (NE-SW and NS orientations), there is no such advantage in the last example (EW orientation).

Based on these results a series of simulations were conducted using the offset grid sampling strategy on circular gas vents having progressively smaller radii, to define the relationship between the number of samples needed to find a given vent at the 95% confidence level. These results are given in Figure 40, which show the rate of increasing number of samples with decreasing vent radius for the total simulation survey area (A) of 1 km^2 (1,000,000 m^2); equivalent sample spacing on the offset grid are given as blue labels.

A power law equation was fit to the distribution (with a very high r^2 value almost equal to 1.0) which can be used to predict the number of samples necessary to find a gas vent of a given radius in a 1 km^2 area. The smallest radius modelled, 17.8 m, has a target surface area (x) equal to 1000 m^2 and thus represents an x/A ratio of 1,000/1,000,000 = 0.001. To find such a vent a total of 1000 samples would be needed at a sample spacing of 32 m using an offset grid. While this is a very large number of samples, it should be pointed out that to find this vent using a random sampling approach, according to Eqn 12, a total of about 3000 samples would be required (Oldenburg et al., 2003).

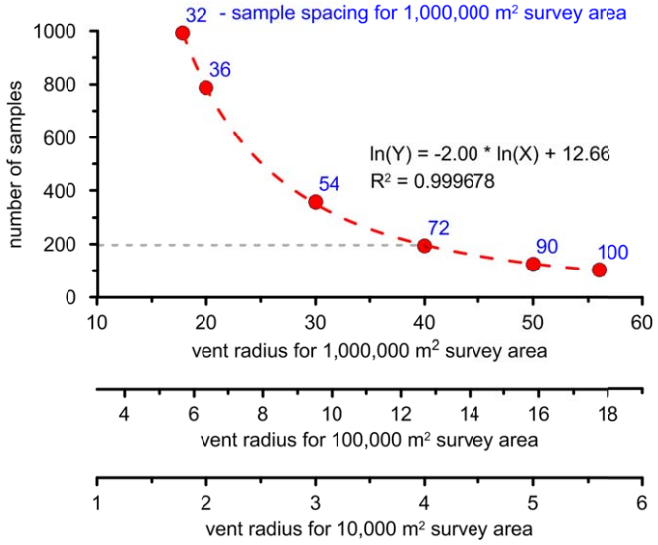


Figure 40. Plot reporting the estimated number of samples needed to find (at the 95% confidence level) circular gas vents having different radii using the offset grid strategy. The dashed line is a power law equation fit to the data, with the equation itself and the r^2 reported on the plot. Sample spacing (in metres) are reported in blue for the 1 km^2 survey area for reference. The multiple X axes are explained in the text.

These results can be extended, considering that the value of x/A can be calculated for each vent, given that the target size is calculated from each vent radius and the survey surface area is fixed at 1 km^2 for the simulations. By taking each x/A value and back-calculating target radii based on different values of A, we can estimate the number of samples needed to find vents in smaller areas. For example, by collecting a fixed number of 200 samples (horizontal dashed line), one would be able to find a vent with a 40 m radius in a 1 km^2 area, a vent with a 13 m radius in a 100,000 m^2 area (i.e. 316m x 316m), or a vent with a 4 m radius in a 10,000 m^2 area (i.e., 100m x 100m). For monitoring of CCS sites this highlights the importance of focussing sampling on smaller, high-risk areas like well heads or inferred fault traces, or to merge monitoring technologies to increase the potential for success. For quantification of natural CO_2 leakage areas it gives an idea of the level of sample resolution that is needed to produce accurate results.

5.3.2. Leak quantification

As discussed in section 2.4.2, the quantification of deep gas leakage from the ground surface has many important applications in which the resultant estimates can influence decisions made. For example, it has been proposed that "diffuse degassing" of natural CO_2 in volcanic or geothermal areas may contribute almost as much CO_2 to the atmosphere as emissions from the more obvious plumes released from volcanic calderas (Chiodini et al., 2004). Diffuse degassing values thus may represent a small but not insignificant contribution to the atmospheric carbon cycle and thus should be considered in climate modelling. Another example is for carbon credit auditing at an active or decommissioned CCS site, because if a portion of the sequestered CO_2 leaks to the atmosphere the credits originally assigned to the operator for that volume of leaked CO_2 must be re-paid.

As outlined in section 2.4.2.1 the most common approach is to conduct a single survey (based on a regular or random sampling strategy) and to use either statistical or interpolation approaches to calculate the total flux, and to calculate and subtract the estimated background flux to estimate the leakage flux. The flux grid and data interpretation given in Chapter 4 addressed some of the difficulties that may be encountered with this approach, particularly related to an accurate estimate of biogenic flux rates over the measured area. One of the few attempts to estimate how many samples are needed to obtain an accurate estimate of this leakage flux was given by Cardellini et al. (2003), who obtained an empirical relationship between the standard deviation of sequential Gaussian simulations (sGs) on sub-sampling of real datasets and the number of samples lying within a circle range area (CRA) having a radius equal to the range of the CO_2 flux variogram for the entire original dataset. These authors showed a good correlation that defined a minimum of 90 samples needed in the CRA to obtain an estimated sGs error of <10%, based on sampling densities (samples km^{-2}) at 5 different study sites equal to: 296 (Solfatara), 3235 (Vesuvius), 239 (Poggio dell'Olivo), 1442 (Nisyros caldera), and 2408 (Horseshoe lake).

Instead, the present research is the first to take a purely statistical modelling approach to address the question of how sampling density influences estimated leakage flux rates. This approach sub-samples known synthetic data (both background and leakage flux rates) which allows one to: i) remove uncertainties that can arise in studies of incompletely known real data; ii) test different scenarios related to size, shape, and spatial distribution of the leakage areas; iii) perform a statistically valid number of simulations at any desired sample density over any desired total surface area; and iv) test the influence of different sampling strategies (purely random, uniform random, grid, and offset grid), different statistically and spatially distributed background datasets, and different contouring algorithms. The methods used by the developed program for calculating total, background, and leakage flux rates, and for subsampling the total synthetic dataset are described in detail in section 5.2.

Three different modelling scenarios using synthetic data are presented below to examine the influence of different settings or parameters on potential errors and uncertainties in total, background, and leakage flux rates:

i) no leakage, to understand possible errors in background flux estimation;
ii) a single circular leakage area having a surface area of 10,000 m^2, to examine the impact of different sampling strategies on leakage estimates;
iii) three circular vents aligned along a N45E direction, at progressively greater distances from each other, to study if isolated anomalies may be merged by the contouring procedure.

In the following scenarios, 200 simulations were performed for each sampling strategy. For the first two scenarios, all four sampling strategies were modelled for 15 different sampling densities; only results from the random and the offset grid sampling strategies are presented as they represent the two extremes. For the second two scenarios, to reduce computational times and the amount of data produced, and based on the earlier results, only the random and offset grid simulations were performed for a total of 9 different sampling densities. To simplify and limit the number of different scenarios examined, all simulations were performed using the same synthetic background dataset, consisting of a randomly generated log-normal CO_2 flux distribution having an average of 20 g m^{-2} d^{-1} and a standard deviation of 5 g m^{-2} d^{-1}, distributed spatially on the grid in relation to a digital elevation model such that high flux rates correspond with low elevations (see section 5.2.2.1).

It is acknowledged that the decisions made in the automation of the process do not wholly mimic what would be done by a researcher with a single dataset (variograms, experience, different kriging models, etc.), however by performing large numbers of simulations on synthetic data it is possible to highlight general problems and issues related to spatial geochemical sampling and subsequent interpolation.

5.3.2.1. Scenario 1 – no leakage

As described in Sections 2.4.2.1 and 4.4 the average background value is a critical parameter for calculating total leakage flux, as it must be subtracted from the total flux measured over the entire survey area. As even small flux rates distributed over large areas yield large total CO_2 flux (e.g. 20 g m^{-2} d^{-1} * 1 km^{-2} = 20 tons CO_2 d^{-1}), an error in this estimate could contribute to an over- or under-estimation of the calculated leakage flux rate (see section 4.4).

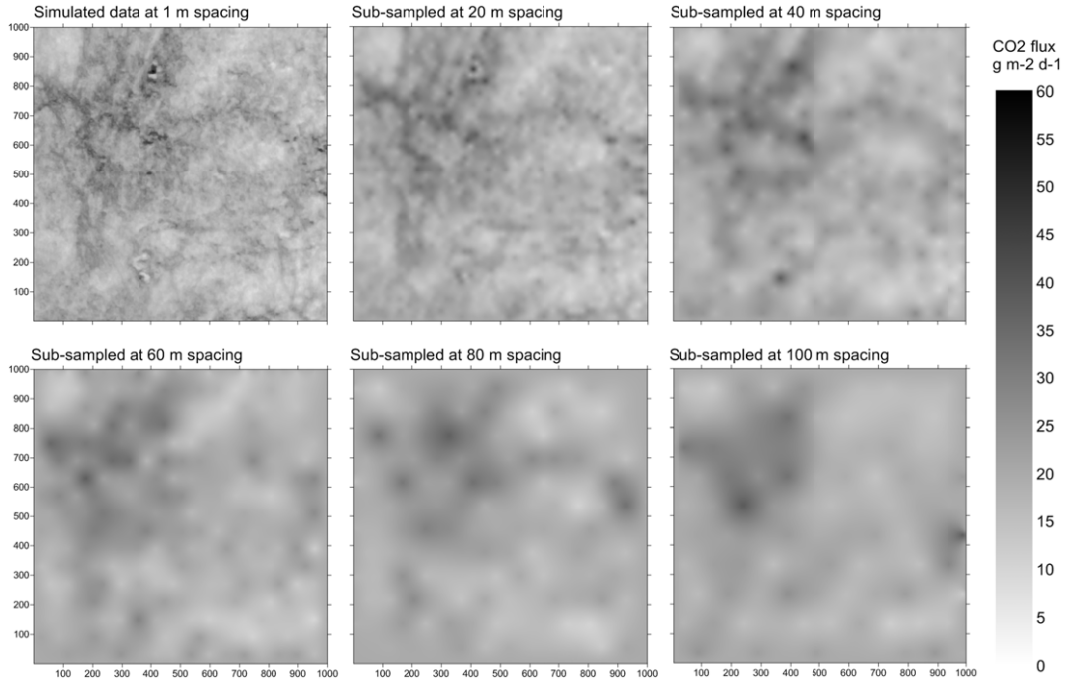


Figure 41. Contour maps showing progressively wider offset grid sample spacing of the synthetic background CO_2 flux dataset. All maps use the same colour scale shown to the right.

To give a conceptual idea of how different sampling densities can influence the sub-sampled results of the background flux field, six CO_2 flux contour maps are given in Figure 41 that show randomly chosen examples of ever decreasing offset grid sampling densities. Although general trends remain visible, such as the higher values in the NW sector, as expected detail is rapidly lost and spatial errors increase with greater sample spacing; in general results are more "smeared" and averaged over much larger areas.

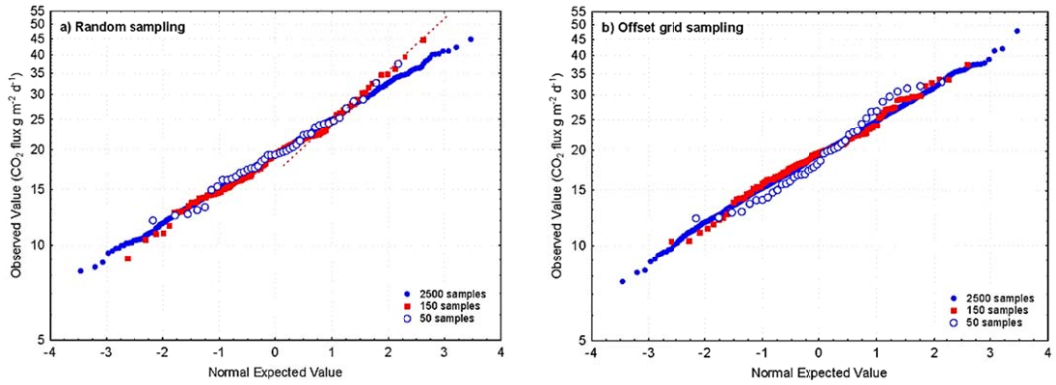


Figure 42. Comparison of the statistical distribution of one simulation for three different sample densities for the random (a) and offset grid (b) sampling strategies. In general trends are similar, although some deviations (e.g. dashed line in (a)) are observed.

The subsampled flux values from the first of the 200 simulations conducted for three sampling densities (2500, 150, and 50 samples km^{-2}) are presented in log-normal probability plots for the random (Figure 42a) and offset grid (Figure 42b) sampling approaches, showing that in general the widely different number of samples yield a similar statistical distribution. Subtle differences do occur, however, in the low density sampling that deviate from the straight line defined by the high density (blue) sampling points, such as a different trend seen in the 150 samples km^{-2} simulation (red) and marked by a dashed line in Figure 42a. In contrast, plotting of two simulations that represent the extremes of estimated total flux rates for the offset grid sampling approach at the 50 samples km^{-2} sample density (Figure 43) shows how subsampling of the same population can yield different statistical distributions. In this example the calculated average values range from 19 to 21 g m^{-2} d^{-1}.

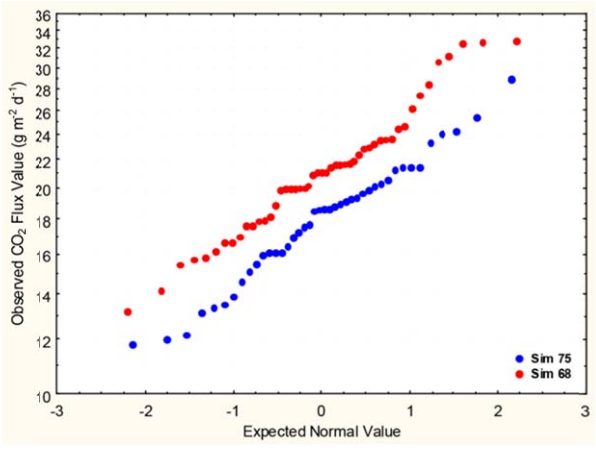


Figure 43. Probability plot of the two 50 samples km^{-2} simulations with the minimum and maximum estimated total CO_2 flux value, showing that different subsampling of the same population can yield different sub-populations.

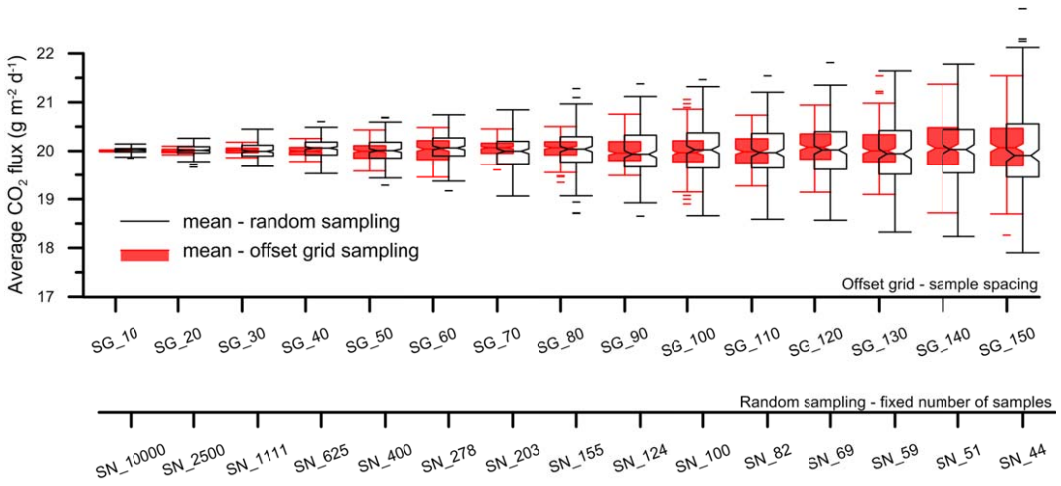


Figure 44. Box and whisker plot showing the statistical distribution of the arithmetic means calculated for each of 200 simulations for each sample density, for both the random and offset grid sampling approaches.

To better understand the potential range of average background values that could be obtained, the arithmetic mean was calculated for each simulation of each sample spacing (offset grid sampling) or fixed number of samples (random sampling) and the resultant dataset (200 calculated averages for each sample density) was plotted as a box and whisker plot (Figure 44). For both sampling methods there is a clear increase in the range of average values possible as sampling density decreases, with the offset grid method producing a slightly narrower distribution. At its most extreme this could produce an error, in this example, of ± 2 tonnes CO_2 d^{-1} (c. 10%), although in general the upper and lower quartiles for all sampling densities are < ± 0.3 tonnes CO_2 d^{-1}. The magnitude of such variability within the results of a real survey would depend on the true range of values, their statistical distribution, and how they were spatially distributed (e.g., high values clustered in a single area or spread more evenly throughout).

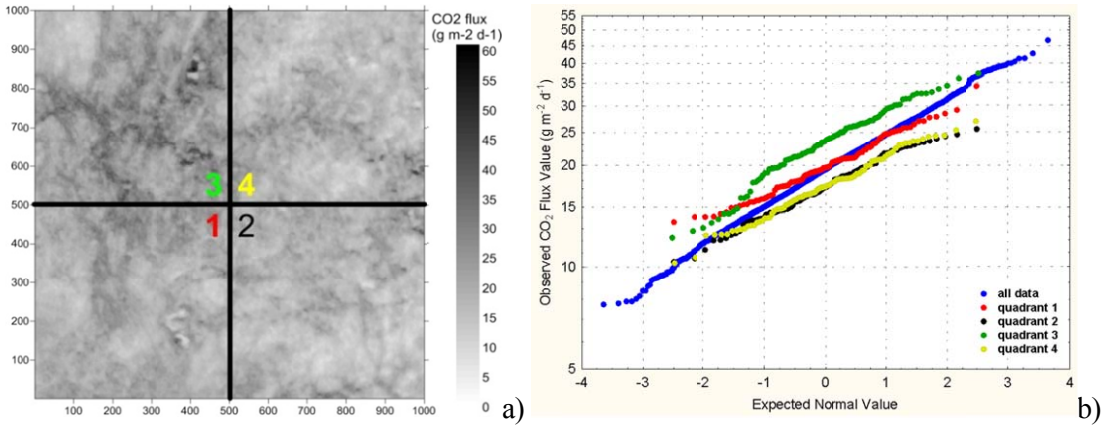


Figure 45. Map of the synthetic background flux grid divided into four quadrants (a) and a log-normal probability plot of 100 randomly selected points from each quadrant compared to the total data set (b). The strong difference between quadrants 2 and 3 in (b) can clearly be seen in (a).

While one approach is to calculate the average background flux using the lower population defined in a log-normal probability plot of all the data collected over the entire sampling grid, essentially that described above, another option is to select a separate area where it is known that no CO_2 is leaking and assume that its average biogenic flux rate is also representative of the leakage area. To examine this the total synthetic background dataset was divided arbitrarily into

four quadrants (Figure 45a) and the statistical distribution of 100 randomly selected points in each quadrant was plotted together with the entire dataset in a log-normal probability plot (Figure 45b); the corresponding statistics are given in Table 3.

Table 3. Statistical parameters describing the entire dataset as well as randomly sub-sampled datasets from each of the four quadrants shown in Figure 45.

	Samples	Mean	Median	Min.	Max.	Up. quart.	Low. quart.	Variance	St. Dev.
all data	1000000	20.0	19.4	5.9	62.7	16.4	22.9	25.0	5.0
quadrant 1	106	20.2	19.7	13.6	34.1	17.3	22.2	15.6	3.9
quadrant 2	107	17.5	17.2	10.4	25.5	14.8	19.8	11.8	3.4
quadrant 3	115	23.7	23.7	12.2	37.2	20.1	26.8	28.4	5.3
quadrant 4	103	17.6	17.3	10.3	26.9	15.0	19.8	12.3	3.5

The four quadrants yield different statistical populations compared to the original total dataset, such that the calculated arithmetic mean values range from the "true" value of 20 g m^{-2} d^{-1} to a minimum of 17.5 for quadrant 2 and a maximum of 23.7 for quadrant 3. For the example used, this could result in an over- or under-estimation of c. 3 tonnes CO_2 km^{-2} d^{-1} for a related leakage flux calculation, however it is conceivable that other situations could yield more extreme differences; the importance of such a potential error will depend on its magnitude relative to that of the actual true leak. In addition, it should be remembered that this error will combine with the error induced by the contouring of any observed leakage anomalies (see scenario 2 below), which will either add or cancel depending on the signs of the two values. In addition to this purely spatial effect, which could result, for example, from different areas having different soil types, water contents, or land uses (e.g. Beaubien et al., 2014), there could also be a temporal effect caused by general or localised events (occurring before or after specific areas are sampled) that influence CO_2 flux rates. These could include, amongst other things, fertilization, tillage, or rainfall, all of which could cause a relatively short-lived increase or decrease in CO_2 soil flux rates (see discussion in Beaubien et al., 2013).

5.3.2.2. Scenario 2 – single vent

A single gas vent with a maximum flux rate of 1000 g m^{-2} d^{-1}, in an area surrounded by a log-normal background flux distribution having an average of 20 g m^{-2} d^{-1} (Figure 46), was created to examine the range of flux rates calculated at different sampling densities for a simple system. Note that the algorithm adds the leakage to the existing background value; while this effect is small when the leakage flux is high, the background value can make a significant contribution to the total flux when the leakage rate is low (i.e., on the outer edge of the vent).

The modelled vent has a radius of 56 m and an approximate surface area of 10,000 m^2. See Figure 29c for a comparison of the trend of flux values for the vent, and Figure 30 c,d to see how the vent relates to the surrounding background along a profile ("slice") through the gridded data. Only the offset grid and random sampling strategy results are presented here.

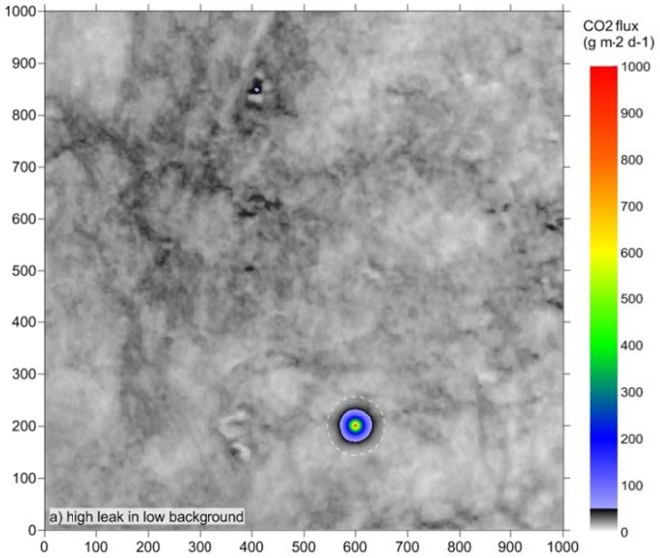


Figure 46. Map showing the location of the modelled gas vent (maximum 1000 g m^{-2} d^{-1}) in an area having a generally low background flux distribution. The dashed white line delimits the calculated vent limits; note that between the dashed line and the start of the blue in the vent, values are within the range of the regional background.

The complete data sets used in this simulation, i.e. 1,000,000 points, are given in the normal probability plot (NPP) of Figure 47. Here the complete background dataset (blue) is given as well as the background plus the single gas vent (red). Note that although there is a clear division between the two, as marked by the point where the slope of the trend changes dramatically, the inflection point does not occur at the upper bound of the background population. In other words, although the total background plus leakage dataset appears to show the boundary between the two populations is around 43 g m^{-2} d^{-1}, the background data plotted alone clearly shows values up to about 60 g m^{-2} d^{-1} and the occurrence of about 600 samples (< 0.1% of the total) greater than 43 g m^{-2} d^{-1}. This observation implies that one should be careful when using such plots to define populations and their respective average values, at least in terms of a small percentage of outlier background values.

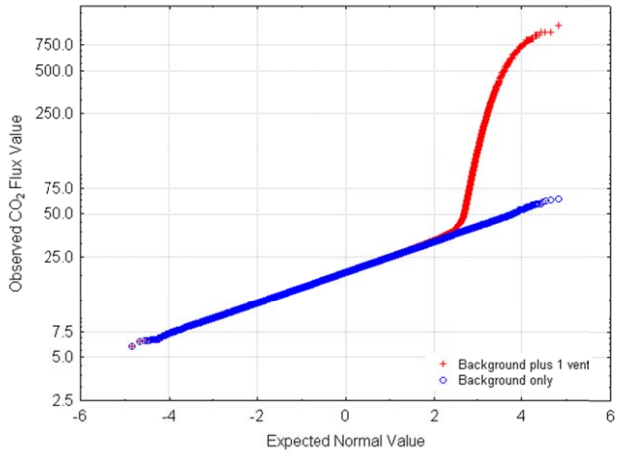


Figure 47. Normal Probability Plot (NPP) showing the statistical distribution of only the background data (blue) and the background plus the single modelled gas vent (red).

The statistical distribution of the estimated leakage flux rate for the 200 simulations for each studied number of sub-samples collected for the two sampling strategies are shown in the box and whisker plots of Figure 48; plot (a) gives all the data while plot (b) is a detail for the first 8 sub-sample simulations without outliers. The blue line in both plots represents the "true" leakage rate of about 0.56 t CO_2 d^{-1}. As expected, 10,000 samples yielded a very narrow and accurate estimate, followed by an ever wider range of estimated leakage flux rates with decreasing number of samples collected. Whereas quartile and 10/90 percentiles attain fairly constant values when less than 200 samples are collected, outlier values continue to increase with each decreasing number of samples collected, up to a maximum of about 15 t CO_2 d^{-1} (i.e., about 30 times the true value) for 44 samples collected.

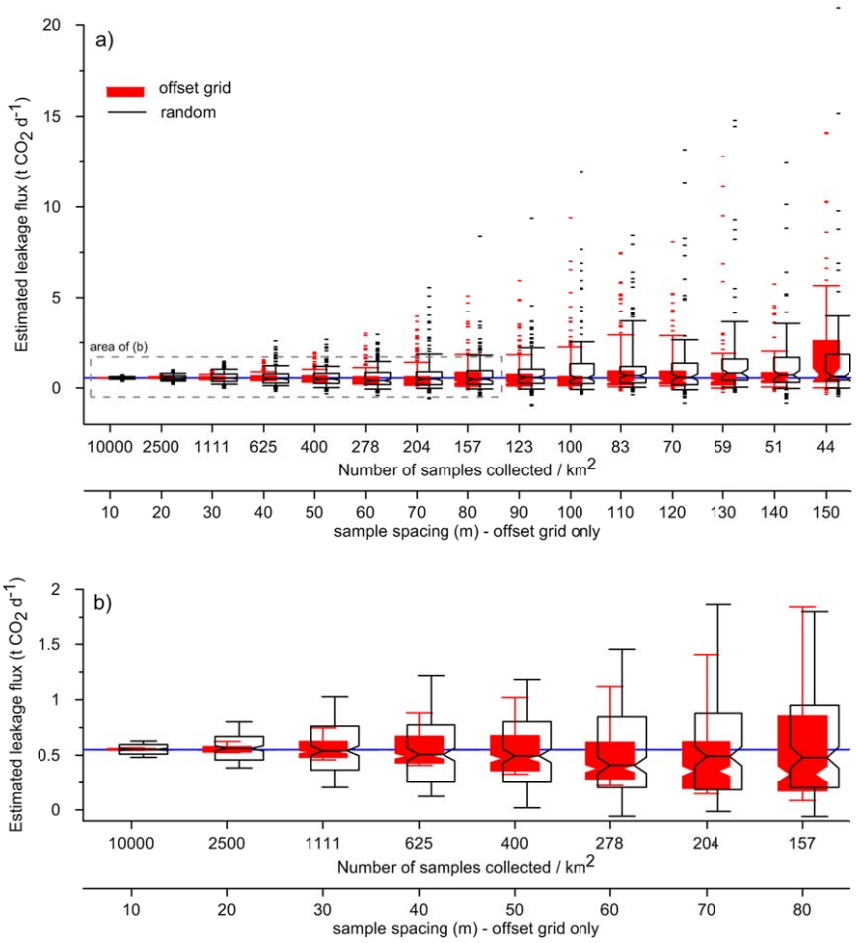


Figure 48. Box and whisker plots of data for the case given in Figure 46 using both the offset grid and random sampling approaches; plots show all data (a) and the first 8 sub-sample numbers without outliers (b). Each result represents the statistical distribution of 200 simulations: the symbols are outliers, the whiskers are the 10/90 percentiles, the box delimitates the upper and lower quartiles, and the indent marks the median. The horizontal blue line marks the true leakage flux.

The tests with high numbers of samples show almost normal distributions centred on the true value, whereas with progressively lower sample numbers the distribution becomes more skewed to higher values and the median drops below the true value. The higher values are the result of the vent being intersected by a single point and interpolation to more distant background values that "smear" the anomaly over an area that is much larger than the true vent size. The lower median values are due to the fact that there is a high potential that at low sample densities the

single gas vent will not be intersected at all and that the vast majority of samples will instead be measured in the background. In theory this should result in a leakage flux rate of zero, however due to the decision to subtract the "true" background value (calculated via gridding of the total background dataset prior to addition of the leakage points, chosen for automated computational simplicity) from the simulation-calculated total flux rate and the fact that subsampling of the background can yield slightly different average values (see Scenario 1 and Figure 43 above) there is the possibility that the simulation background may be slightly higher or lower than the true background used. This can result in a few outlier cases where the calculated leakage flux rate is negative. This artefact of the chosen algorithm is not expected to have a significant impact on the overall interpretation of the results.

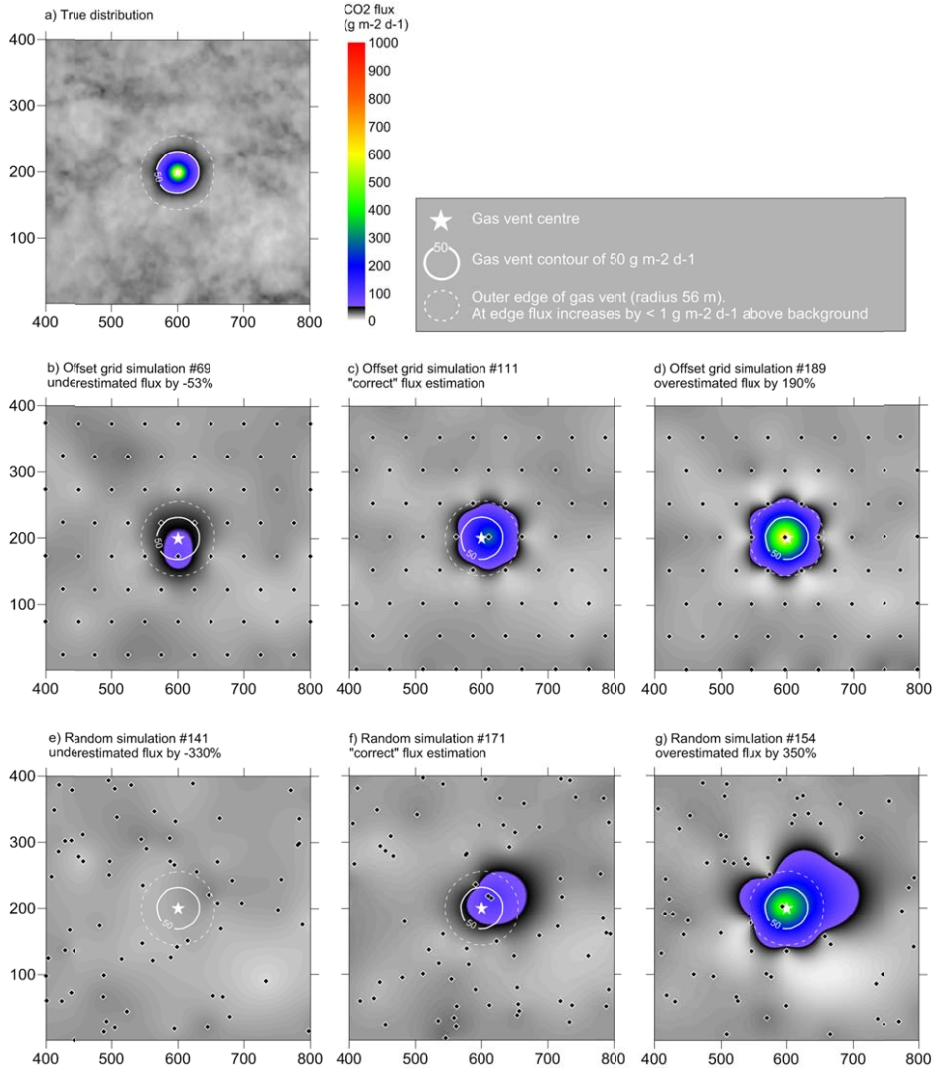


Figure 49. Selected simulations for the 400 sub-sampling of the case given in Figure 46a. The true location of the gas vent is marked by the solid white line outlining the 50 g m^{-2} d^{-1} contour and the dashed white line delimiting its outer boundary. Black diamonds mark sampling points.

The greater detail given in Figure 48b shows how the offset grid distribution is significantly narrower than that of the random sampling method above 200 samples km^{-2}, whereas below this threshold the two different methods yield more similar results. Examples to explain this wide difference between the two sampling approaches at the higher sampling densities are given in

Figure 49, which shows results from the 400 samples km^{-2} density. In addition to the "true" vent distribution (Figure 49a), this figure shows simulations with the lowest, a "correct", and the highest leakage flux estimate for the offset grid (Figure 49 b, c, and d, respectively) and the random (Figure 49 e, f, and g, respectively) sampling methods. The regular nature and spacing of the offset grid guarantees that the vent is intersected at least once at this sampling density (Figure 49a), while the fact that sample spacing is essentially equal to the vent radius ensures that its lateral extent is bounded thus limiting the size of any flux over-estimation (Figure 49c). In contrast, even at this relatively high number of 400 samples, the random sampling approach risks to completely miss the vent (Figure 49c), or to leave holes in the sample distribution into which falsely elevated values can be interpolated (Figure 49g). Finally, in both examples where the leakage rate was accurately estimated (Figure 49c, f), neither sampling method defined the true leakage distribution within the vent (which is to be expected at this sampling density) and the correct calculation was a fortuitous balance between the measured anomaly and the interpolated surface area.

The seven datasets in Figure 49 are presented on log-normal probability plots in Figure 50, illustrating the separation of the two populations. Similar to that discussed above (see Figure 47), however, this separation is not always complete, as shown by a close comparison of these two figures. For example, Figure 49b shows three points within the vent area of simulation-69, however Figure 50a shows only one simulation-69 sample in the anomalous population; the other two points (having values < 50 g m^{-2} d^{-1}) plot instead within the tail end of the background population.

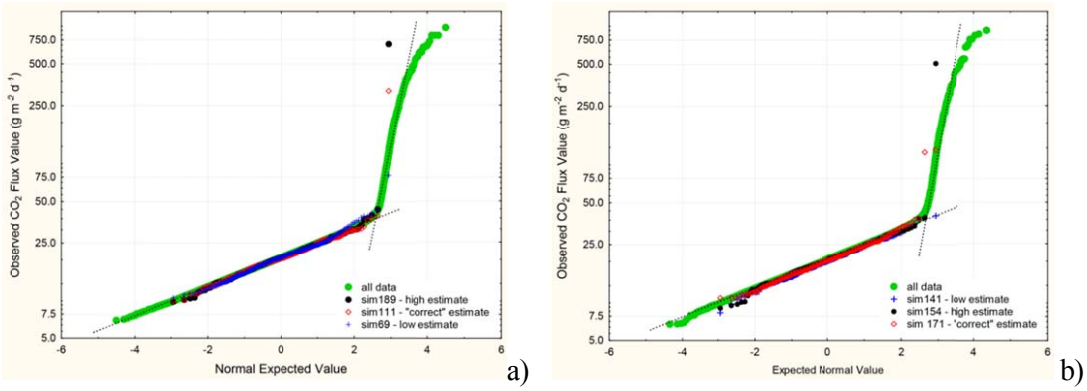


Figure 50. Log-normal probability plots for the data presented in Figure 49: a) offset grid sampling simulations; b) random sampling simulations.

The issue raised above regarding a wide offset-grid sample spacing yielding high leakage estimates due to the lack of bounding sample points is illustrated in Figure 51b. This plot shows the contoured results of the simulation having the highest leakage estimate for this sampling approach on the 100 m spaced grid (8.4 t CO_2 d^{-1}, compared to the "true" value of only 0.56 t CO_2 d^{-1}). In this example the centre of the vent with the highest flux was intersected with one point and then values were extrapolated over a radius of 100 m to the next point. Such an occurrence in a real campaign could be avoided relatively easily, however, with limited infill sampling around significant anomalies to constrain their spatial extent. For example, the addition of only 4 points would reduce the over-estimate from 1400% of the true value to 360% (Figure 51c), whereas the addition of 8 points would yield a correct estimate of the total leakage flux (Figure 51d). Such infilling could be conducted during sampling, or could be performed at the end of a campaign once all data had been collected and anomaly thresholds defined.

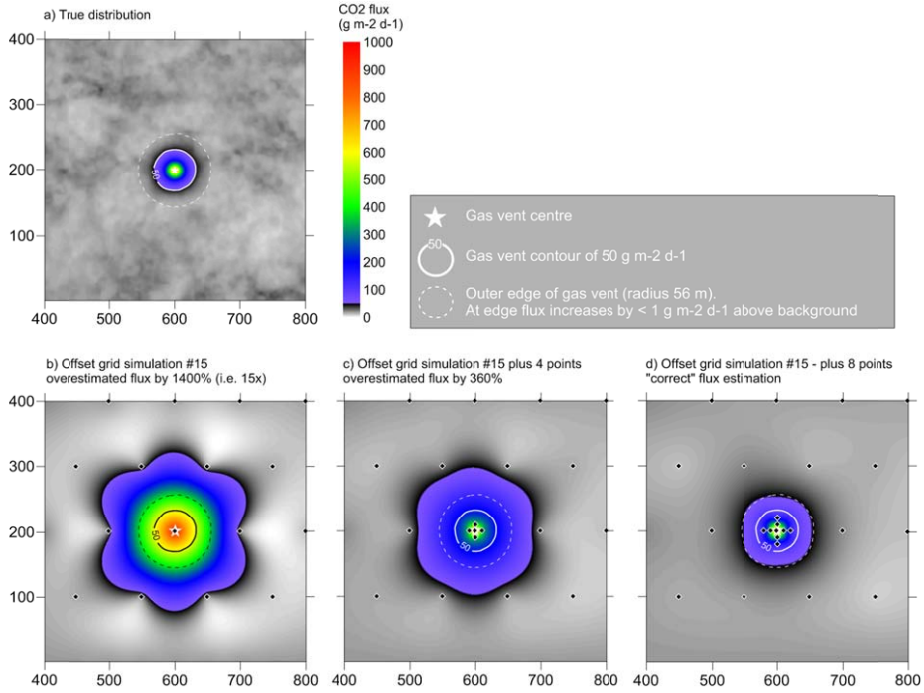


Figure 51. Contour maps of the true gas vent distribution (a), the most anomalous 100m-spacing simulation (b), and the re-calculation of flux distribution and total flux amount for the same dataset with the addition of 4 (c) and 8 (d) sample points. Note that the additional 8 sampling points result in an accurate estimate of the total leakage flux and its spatial distribution.

5.3.2.3. Scenario 3 – three vents, simulations at three different distances

Three gas vents, identical in size and flux rate to the single vent simulated in the previous section, were aligned along a NE-SW direction to look at the potential for merging of anomalies in the contouring process. As shown in Figure 52, three different positions were tested, with the centre of the vents located about 1 (position 1), 1.5 (position 2), and 2 (position 3) vent diameters away from each other. As before the vents were superimposed on a background log-normal baseline grid (average CO_2 flux of 20 g m^{-2} d^{-1} and a one sigma of 5 g m^{-2} d^{-1}) that was distributed based on topography. The vents themselves were positioned to the south of the higher background area in the NW. Only the offset grid and random sampling strategies were used for these simulations.

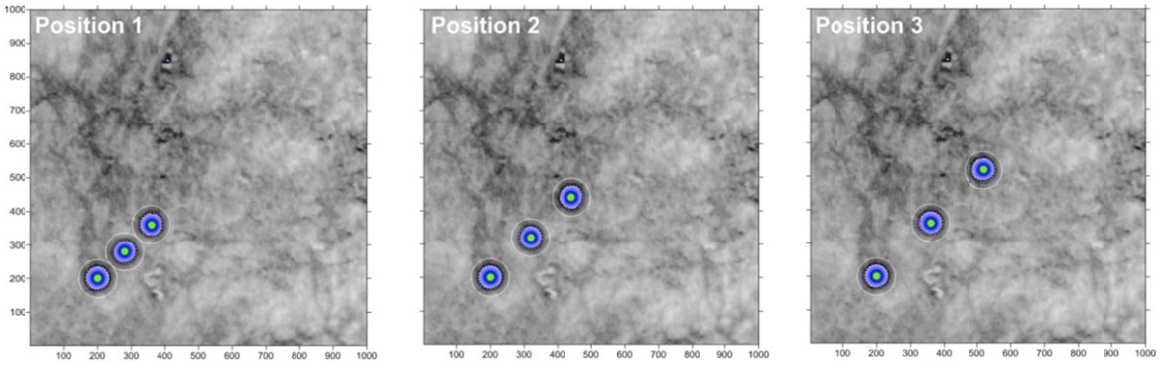


Figure 52. Maps showing the location of the three modelled gas vents, positioned at progressively increasing distances from each other for different tests.

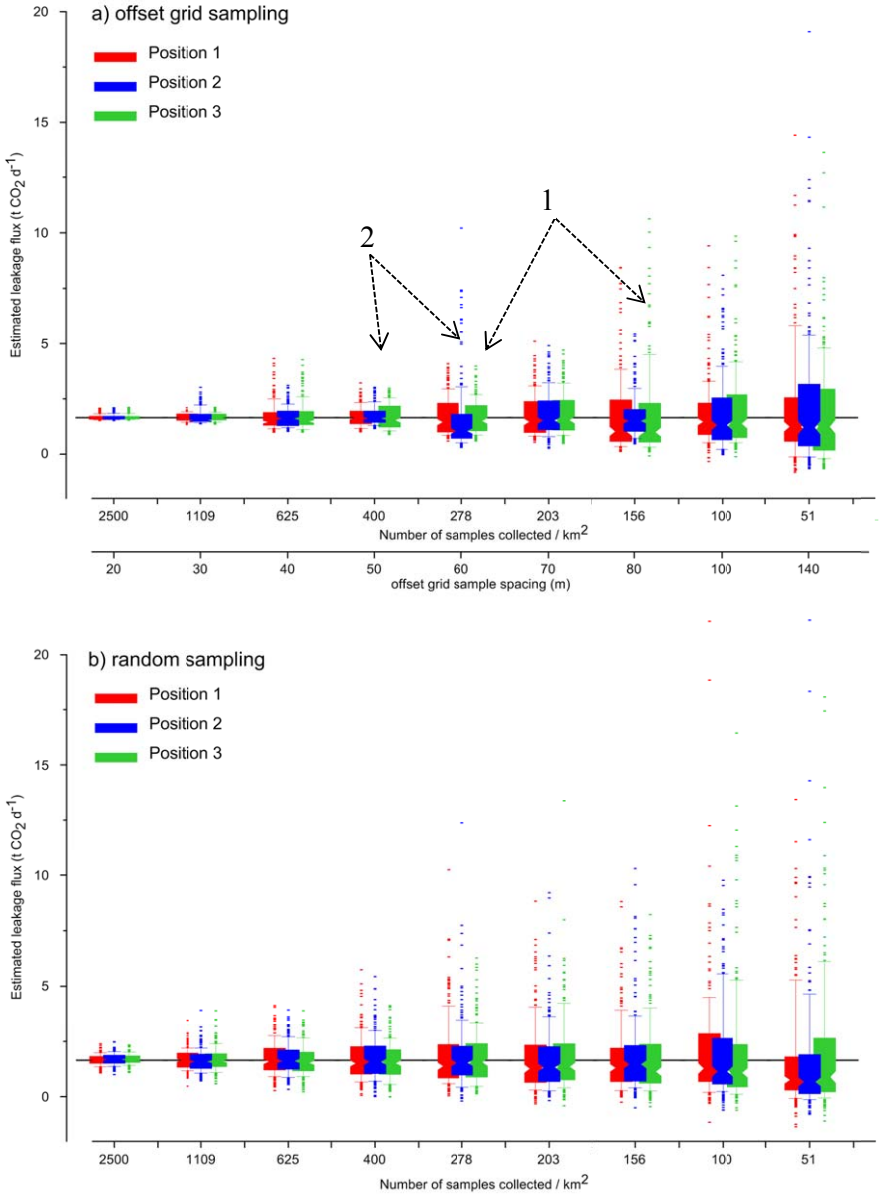


Figure 53. Box and whisker plots of the estimated leakage rates for the three different vent positions using both the offset grid (a) and random (b) sampling strategies. Numbered arrows discussed in text.

The statistical distribution of the estimated leakage flux for each of the three simulated positions are shown in the box and whisker plots of Figure 53, for both the offset grid (a) and random (b) sampling strategies. The horizontal black line in both plots represents the true leakage rate of about 1.65 t CO_2 d^{-1}.

In general, the offset grid sampling approach (Figure 53a) provides a relatively narrow range of calculated total leakage rates at the high sampling densities, with a marked widening of the ranges and greater differences between the three vent position results above a sample spacing of about 50 m. In contrast the random approach (Figure 53b) yields a wider statistical range, especially at the high sample numbers, shows a more regular increase in ranges and outlier values, and more similar results amongst the three vent positions for any given sampling density. Some of these observations are discussed in more detail below.

Although the offset grid sampling results often show similar statistical distributions for the three vent positions, for a given sampling density, some exceptions do occur ("1" in Figure 53a). For example at 60 m grid spacing the second position of the vents yields generally lower values but much higher outlier estimates, whereas at 80 m the first and third positions give much higher outlier estimates. An explanation for this behaviour can be found by examining the spatial distribution of some of the simulations, specifically those that gave the highest estimations (Figure 54). This figure shows how the geometrical relationship between the sample spacing and the distance between vent centres controls how many vents are intercepted and where. For example, the spacing of the vents in position 2 results in the potential that all three vent cores are sampled (Figure 54b), something that is not possible for the other two positions (Figure 54a, c). Similarly, however, this same relationship between sample spacing, vent radius, and vent distance also means that there is a greater likelihood that positions 1 and 3 will intersect one relatively high flux value in at least one vent, while position 2 may miss all three. These observations explain how the overall distribution of positions 1 and 3 are higher but that the outliers are much higher for position 2 at this sample spacing.

Instead, the centres of all three vents can be intersected at the 80 m sample spacing when the vents are spaced as in position 3 (Figure 54f), two can be intersected with vent position 1 (Figure 54d), whereas only one vent centre can be intersected with vent position 2 (Figure 54e). In addition, a comparison between the highest estimate at 60 m spacing (Figure 54b) and at 80 m spacing (Figure 54f) shows how the wider sample spacing can yield much higher overall estimates, due to over-extrapolation during the contouring process, similar to that described in the discussion of Figure 51.

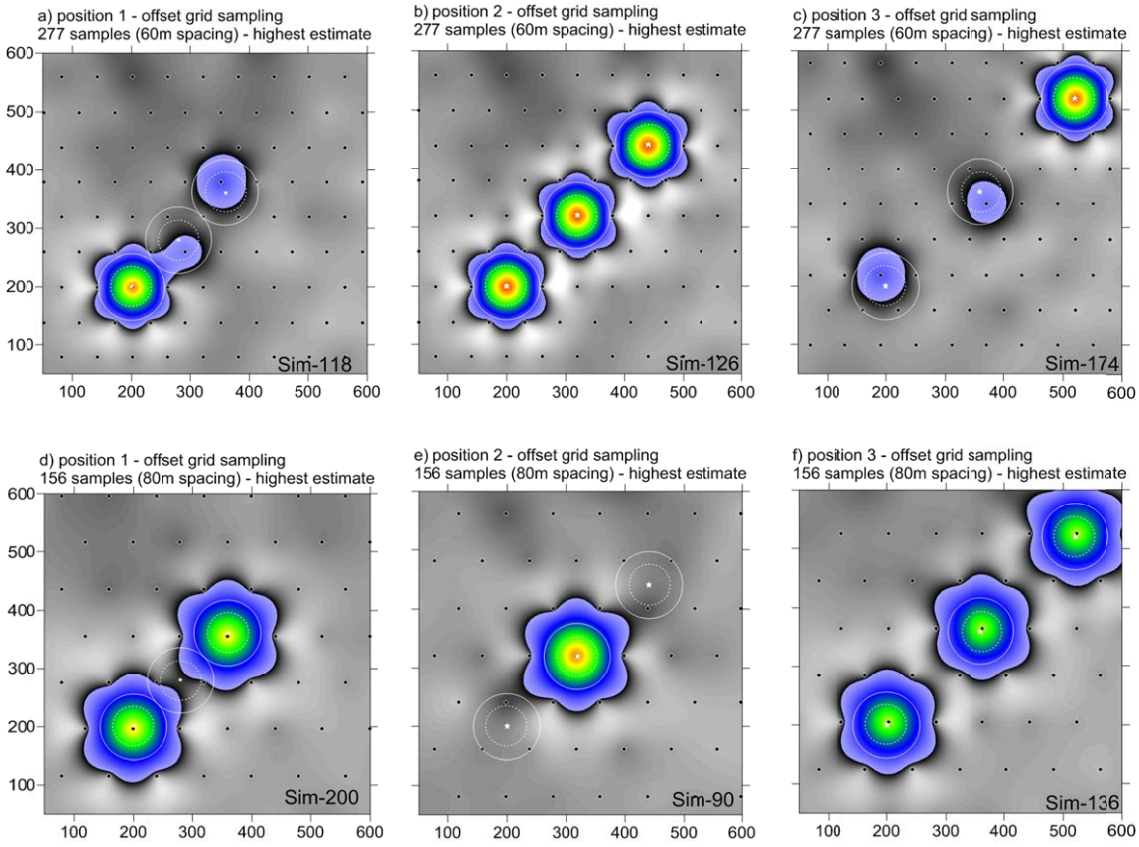


Figure 54. Comparison of the highest leakage flux estimate using the offset grid sampling strategy for the three different vent positions, for 60 m spacing (a – c) and 80 m spacing (d – f).

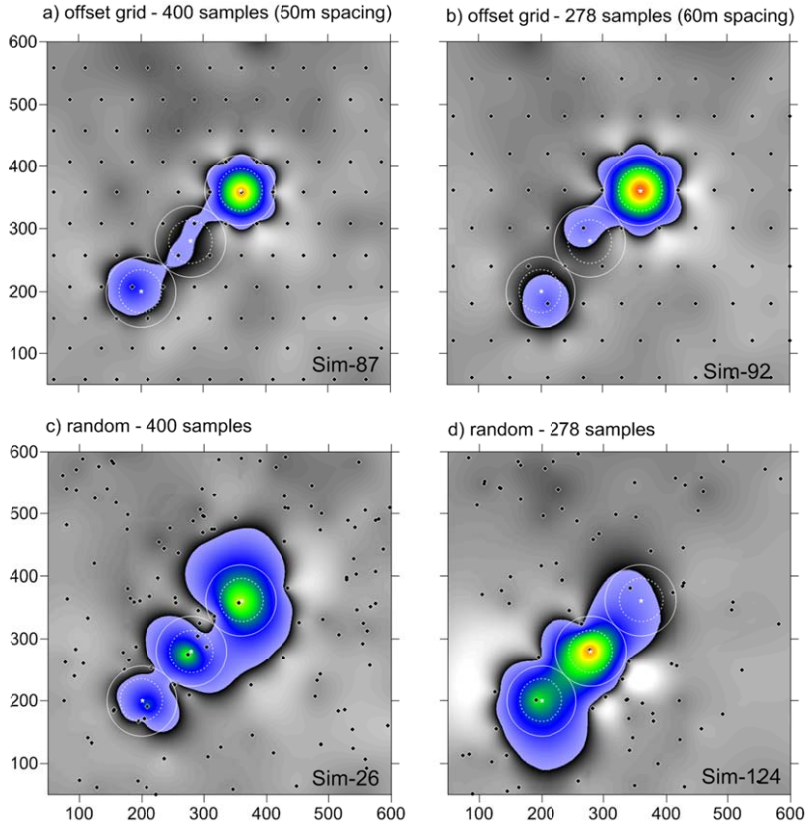


Figure 55. Representative results from the 400 samples km^{-2} simulations (left column) as well as the 278 samples km^{-2} simulations (right column) for the offset grid (a, b) and random (c, d) sampling strategies.

The change observed in the offset grid results between the sample spacing of 50 and 60 m ("2" in Figure 53a) is likely, as observed above, due to the geometry and relative dimensions of the grid spacing versus the vent diameter. This is illustrated in Figure 55, where results from 50 m spacing (a) and 60 m spacing (b) show how the wider spacing beyond that of the diameter of the vent increases the size of the interpolated anomaly and, thus, its total flux. It is interesting to note that whereas the offset grid approach seemed often capable of separating the multiple anomalies due to the occurrence of intermediate samples (Figure 55a,b), the random approach was much more prone to merging the vents and smearing them over a much larger area (Figure 55c,d).

5.4. Latera data modelling

These simulations were performed by inputting real data collected from a single field at the Latera caldera site (Chapter 4). This CO_2 flux data was collected using 10 m sample spacing on an offset grid pattern. As the measured field was not regular, importation into the program was facilitated by expanding the grid area to create an orthogonal rectangle (Figure 56). Points outside of the measured grid but inside the rectangle were assigned random values taken from the lower population of the measured dataset (from 20 to 50 g m^{-2} d^{-1}; see section 4.4). Modelling was performed again using the offset grid and random approaches; the former included sample spacing of 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, and 100 m while the number of samples for the latter was chosen to match the average value for each of the listed sample spacings.

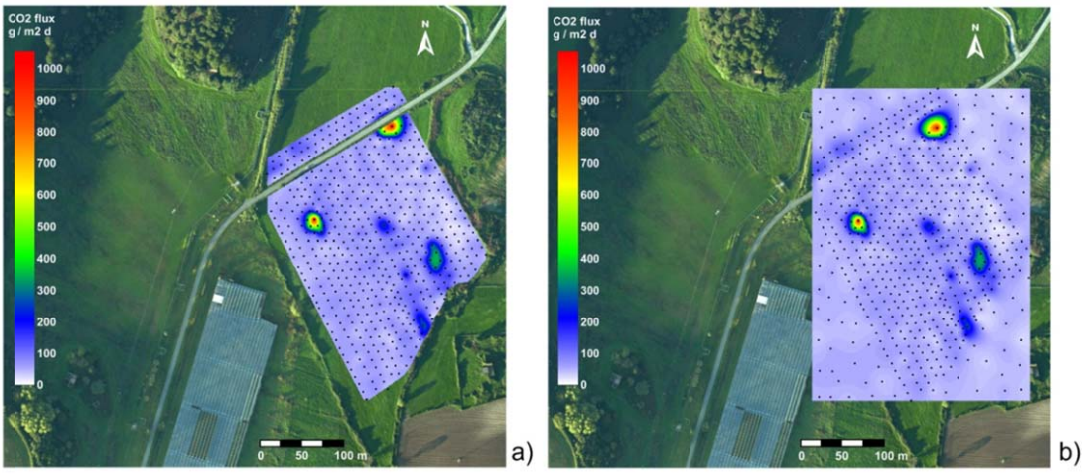


Figure 56. Original Latera CO_2 flux data (a) and expanded area to facilitate import into model (b).

In contrast to the synthetic data, the true average background value of the real data is unknown. As discussed in section 4.4 estimates range from 35 to 53 g m^{-2} d^{-1}, with the comparison with historical data implying a real value closer to the higher of these two options. As a compromise, a value of 45 g m^{-2} d^{-1} was used; as the total background (i.e., average multiplied by surface area) is subtracted from each simulation the selection of this value will affect the absolute flux value calculated but will not affect trends.

A summary of all simulations is given in Figure 57 for the offset grid (a) and random (b) sampling approaches. The two methods give a similar trend, with decreasing sample numbers resulting in a shift of the average and quartile range (i.e. the red box) to progressively lower values (as the possibility increases than few or no gas vents are sampled, or only the low flux edge) but with outliers becoming increasingly larger (as the greater distance between samples means that if a high-flux portion of a vent is sampled its elevated values will be extrapolated well beyond its actual physical extent) (Figure 58).

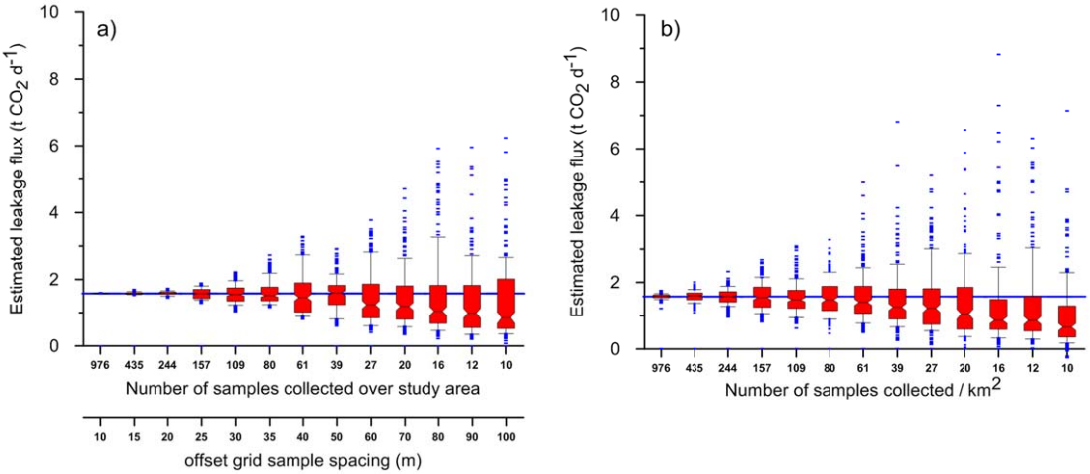


Figure 57. Box and whisker plots of the estimated CO_2 leakage rate at 13 different sample densities, using the offset grid (a) and random (b) sampling approaches. Statistical distributions for each sample density are based on 200 simulations each. The horizontal blue lines mark the "true" leakage rate.

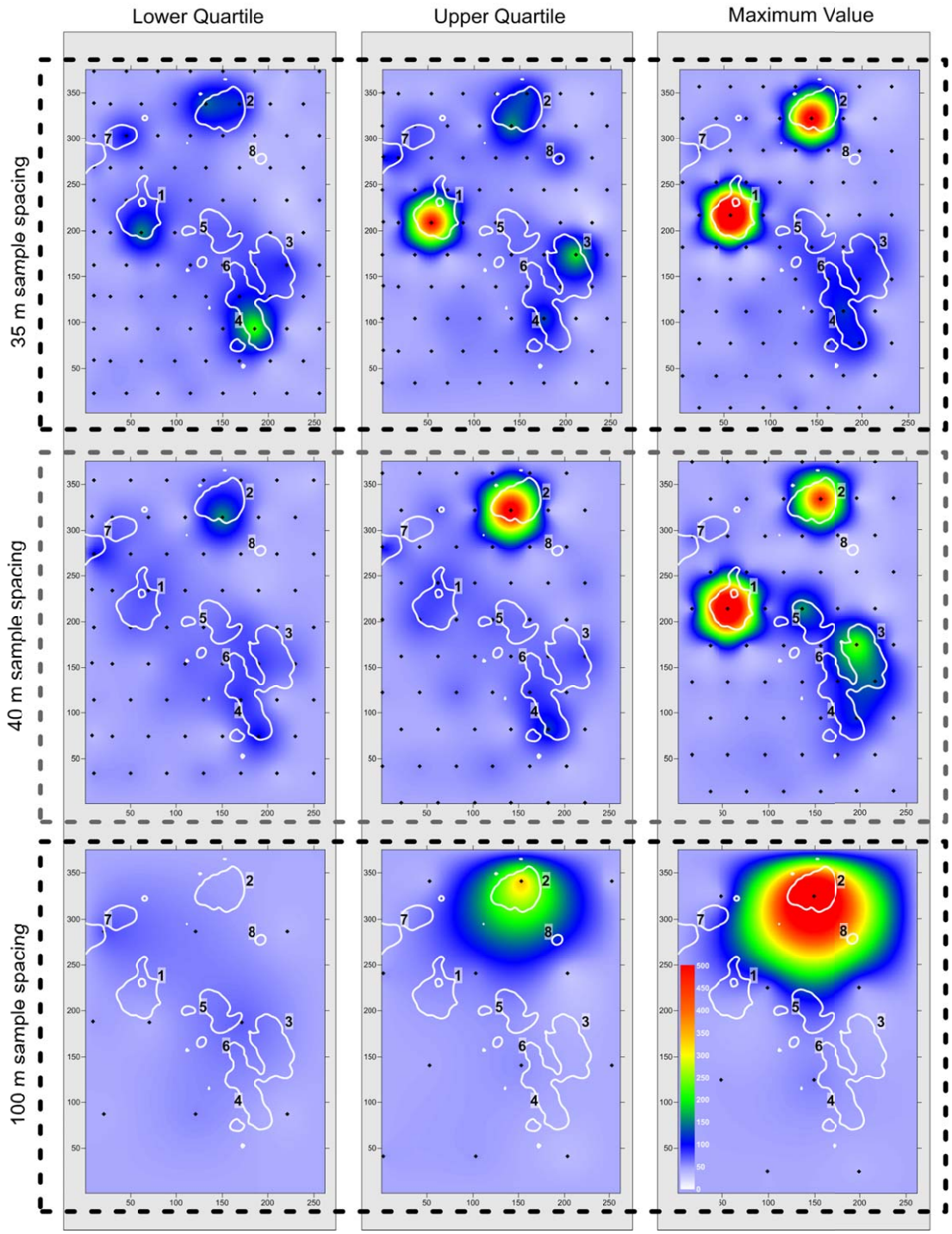


Figure 58. Contoured sub-sampling simulations of the real field data collected at Latera, at different points in the statistical distribution (i.e., lower and upper quartile, maximum value) and for three different sampling distances (35, 40, and 100 m) for the offset grid approach. The white lines represent the 75 g m^{-2} d^{-1} contour of the original field data (10 m sample spacing) and the black dots the sub-sampled points collected during the simulations. The colour scale, reduced to the interval of 0-500 g m^{-2} d^{-1}, is the same for all plots.

Examining the offset grid simulations (Figure 57a), results for the first three sample spacings are all very similar to the true value (1.6 t d^{-1}) and to each other, and very narrow in their distribution. For the subsequent three spacings, up to 35 m, the average value is equal to the true value and the upper and lower quartiles are only about +/- 10%, although outliers start to increase. Beyond this level the quartile range starts to increase, the average value decreases, and outliers become very large. This boundary between the 35 and 40 m sample spacing is illustrated in Figure 58, where simulations representative of the lower quartile, upper quartile, and maximum outlier values are plotted for three different spacings. This figure shows that the step between 35 and 40 m is likely due to 40 m exceeding the average radius of the vents, thus extending an anomaly past the actual size of the vent when encountered or, in contrast, completely missing the vents. Note also that almost all of the vents are intersected at the 35 m spacing.

The largest modelled sample distance of 100 m (i.e., 100 samples km^{-2}) gave the widest range of simulated values, although not very different from that obtained for the 80 m spacing. Relative to the "true" value of 1.6 t d^{-1}, the 100 m spacing yielded 0.08 t d^{-1} as the lowest estimate, a lower quartile of 0.5 t d^{-1} (-69%), an average of 0.8 t d^{-1} (-53%), an upper quartile of 2 t d^{-1} (+25%) and a maximum outlier of 6.2 t d^{-1} (+287%). Representative examples are shown in Figure 58 (bottom row). These results clearly show that great care must be taken not only in planning a sampling campaign beforehand, but also having the flexibility during the work to add additional sample points to constrain found anomalies (e.g., see Figure 51) or to test anomalous looking points that may be sites of leakage (e.g., vent 6 in Figure 21).

Finally the statistical distribution of all simulations for the offset grid and random sampling approaches are compared in Figure 59, with the outliers removed to better highlight general trends. This figure shows how the offset grid results are much more accurate and precise at the higher sample densities, but that this difference decreases at the lower densities. At the extreme end of 100 m sample spacing the random method distribution is narrower, but it is significantly lower than that of the offset grid approach (e.g. the upper quartile of the random data is significantly below the true value.

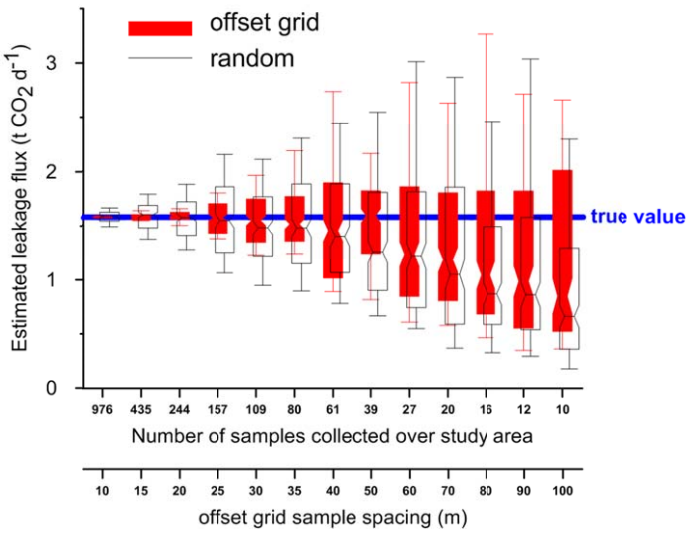


Figure 59. Detailed comparison of the offset grid and random sampling results presented in Figure 57 a and b, respectively, with outlier values removed.

5.5. Summary

This chapter documents the development of a computer program in Visual Basic, "GasGrid", and its application to goals related to monitoring of Carbon Capture and Storage (CCS) sites and carbon mass balance studies. The program subsamples highly detailed synthetic or real data at different densities, multiple times, and uses different spatial distribution approaches to obtain a measure of the probability of finding a leak or the possible range of CO_2 leakage rates that could be obtained by modifying these two variables. In theory any spatially variable parameter can be modelled, however for the scope of this research the flux of CO_2 from the soil (originating from both shallow biological soil sources and from CO_2 leaking from the deep subsurface) was chosen for study.

When synthetic data is used, a user defined grid of points (in this case 1km x 1 km) with a 1 m spacing is created and populated with a log-normal background biological CO_2 flux that is spatially distributed based on topography (i.e. low lying areas associated with higher flux due to potential for higher moisture and nutrient content in the soil). Onto this background, leakage point(s) of user defined maximum flux and anomaly width (with decreasing trends defined based on real leakage areas) are superimposed. When real data is used, a grid as detailed as logistically feasible should be collected over an area with both background and leakage areas, results of which are then imported into the program. The program is then used to sub-sample the created / imported dataset using either a grid, offset grid, uniform random, or random approach. The user defines a list of multiple sample densities and the program performs a user defined number of unique simulations for each of those densities, thus producing a statistical dataset of potential results for a given sampling density and approach. Two main issues can be addressed with the program. First, how many samples must be collected to find a leakage point of a given size and shape at a given confidence level ("leak detection"), and second, how does sampling density affect the overall estimate of CO_2 leakage flux from an area ("leak quantification"). For the former, the program keeps track every time the leakage area is intersected and various plots and statistics are used to interpret the results. For the latter, the sub-sampled data for each simulation is contoured and total flux calculated using algorithms within the program Surfer 9, then background estimates are subtracted to obtain the estimated total leakage flux.

Leak detection was first studied by performing simulations of each sampling approach on five different anomalies having the same surface area and ratio of anomaly to total surface area, but different shapes or orientations (one circular vent and four elongate belts oriented NS, EW, N45°E, and N64°E). The interpretation of initial tests performed with 500 simulations for each experiment showed that most sampling approaches produced statistically stable results after about 200 simulations, and thus this number was used for all subsequent work to reduce computational times. Whereas the random sampling approach yielded the same results (300 samples to find the anomaly at the 95% confidence level) for each anomaly shape because the ratio between anomaly and total surface area was constant, the efficiency of the other methods was affected by the shape and orientation of the anomaly. Of all the methods, the offset grid sampling approach performed best, for example requiring anywhere from a third to a half as many samples (100 – 140 samples at the 95% confidence level) as that needed for the random approach for most shapes. The exception to this was the EW alignment, where 80% more samples were needed (560 samples) due to the narrow geometry of the anomaly often laying between grid node rows. The uniform random approach yielded intermediate results between the offset grid and random approaches and gave a relatively stable number of samples for each

anomaly shape, however the difficulty of applying this approach in the field means that it may not be a feasible alternative.

Based on these tests the offset grid approach was found to yield the best overall results, however if there is the potential for long narrow anomalies it is best to align the grid offset direction perpendicular to the strike of the anisotropy. A power law was developed for the offset grid approach which estimates the number of samples required to find a circular anomaly (at the 95% confidence level) having a specific anomaly area to total survey area ratio (for example, 72 samples would be needed to locate a vent with a 40 m radius in an area of 1 km^2 or to locate a vent with a radius of 4 m in an area of 10,000 m^2).

Three scenarios using synthetic data were studied for the leakage quantification applications of this model, including: i) no leakage to understand possible errors in background flux estimates; ii) a single circular gas vent to examine the effect of sampling approach and sample density; and iii) three circular vents aligned along a NE direction to look at the potential for the merging of individual vents via the contouring process. For this work only the random and offset grid results were presented as they represent the two extremes; in contrast to the leak detection simulations, only circular vents were modelled to maintain the number of tests at a reasonable level. Subsequent to this work on synthetic data, sub-sampling simulations were also conducted on real data from the Latera Caldera. Although not reproducing all data processing aspects that may occur in a real, single survey (such as variogram analysis, survey flexibility based on researcher experience, etc.) the automated, "brute-force" approach used here can give useful general indications on possible problems and inaccuracies, and can help improve survey planning.

The no leakage scenario was examined because of the importance of accurately estimating the biological background CO_2 production, as this value must be subtracted from the total measured flux to calculate the leakage flux. In the first test the 1 million point synthetic grid was subsampled at different densities, and the statistical distributions compared. Although the trends and average values are generally similar, the range of potential arithmetic mean values produced for a given sample density become wider with increased sample spacing. At its most extreme this can result in a range of ± 10% around the true value despite the fact that it is the same large dataset that is being subsampled, although the quartile range is typically closer to the much lower value of ± 1.5%. In the second test different quadrants were subsampled and their statistical distributions compared, as some researchers use non-leaking background areas to estimate biological flux. In the scenario modelled here a variability of about ± 15% was observed relative to the true total background flux. Conceivably even higher differences could occur depending on the actual site conditions (land use, water content, topography, or changing conditions during sampling like rainfall events or fertilization), and thus this approach could potentially introduce significant errors, especially if the estimated leakage rate is relatively small compared to the biogenic flux estimate error.

The single vent scenario was studied to look at sampling density and sampling approach and the effect that these two parameters have on the estimated leakage flux. As observed previously for leakage detection, the offset grid approach yielded much better leakage quantification results compared to the random approach, especially at the high to medium sample densities. Detailed examination of the statistical distribution of the various datasets using probability plots shows how members of one population can plot within the field of the other population when they occur in the transition zone. This fact is important when defining the different populations, especially as related to background flux estimates.

Results show how at high sample densities the statistical distribution of the simulation results are almost normal with average values close to the true value and quartile values within a small percentage. At lower sampling densities the interquartile range widens, the average value decreases below that of the true value, and the outlier values become progressively larger and more skewed. The lower average values are due to the vent being less likely to be sampled (thus sampling is almost exclusively of the background), while the high outliers are due to the extrapolation of the intersected vent well beyond its actual size. Regarding the offset grid approach, quartile ranges are about ±15% of the true value for the 30 m spacing, ± 25% for 50 m, and + 20% and – 70% for 100 m, while maximum estimated values (i.e., greatest outliers) are +180%, +390%, and +1200% for the same sample spacings. It is shown that limited additional "in-fill" sampling around found anomalies, to constrain their spatial extent during the contouring procedure, can greatly improve final leakage estimates.

The three vent scenario, which actually consisted of tests on three different distances between those vents, looked at the potential merging of anomalies during the contouring process and how it may affect final leakage estimates. While many of the behaviours observed in the previous single vent simulations were again noted, an additional aspect that influenced the obtained results was the interaction of sample distribution geometry, vent size, and distance. For example, some sample spacings using the offset grid approach were able to intersect all three vent centres or no vents at all, whereas other spacings tended to hit at least one vent, meaning that the former resulted in a lower average but with higher outliers compared to a higher average with lower outliers for the latter. The random sampling approach tended to merge the three vents more than the offset grid approach, as there was more potential for sample gaps between the vents.

Finally the detailed grid of CO_2 flux measurements conducted in a single field within the Latera caldera were input into the program and sub-sampling simulations were conducted to see the range of leakage flux estimates that would be possible at different sample densities. Similar to that observed above, the lower the sampling densities resulted in the average leakage estimates decreasing but with outlier values becoming progressively larger. Based on the estimated average background flux rate used for these simulations, equal to 45 g m^{-2} d^{-1}, the true leakage rate was estimated to be about 1.6 t d^{-1}. The average leakage estimate was around this value up to a sampling spacing of about 50 m, with extreme estimates equal to about double the actual value. At wider spacing the average value drops towards half that of the true value, while maximum estimates were about 4 times the true value. Finally, as observed above in the three vent synthetic data simulations, a relationship was found between the average diameter of the leakage areas and the sample spacing, which influenced how often the vents were intersected, where, and if the multiple vents were defined as individual entities or merged during the contouring process.

Chapter 6. Impact of CO_2 on groundwater quality

6.1. Introduction

Although a well-chosen and well-engineered geological CO_2 storage site is not expected to leak, all potential risks must be addressed to ensure human health and safety and ecosystem integrity and to comply with regulations, such as the EU Directive on CO_2 storage (Directive 2009/31/EC of 23rd April 2009). One issue that has received much attention, and raised concern amongst regulators, some NGOs, and the public at large, is the impact that geological carbon capture and storage (CCS) may have on potable groundwater resources. This potential impact could come about through two main mechanisms: i) by leakage of a portion of the injected CO_2 gas into an overlying aquifer, resulting in acidification and in situ chemical reactions that change water quality; or ii) by the migration of deep reservoir brines into an overlying aquifer that changes its water quality, both by the addition of high concentrations of the elements dissolved in the brine itself and by in situ reactions induced by the different pH or Eh conditions of the leaking brine relative to the aquifer. Clearly these two mechanisms could also be combined, with the leakage of both injected CO_2 and brine together. In addition to the many research papers that have been published on various aspects of potential groundwater impacts due to CCS, some review articles have recently summarized the processes involved, the present state of the art, and potential knowledge gaps (Harvey et al., 2012; IEAGHG, 2011; Lemieux, 2011; Lions et al. 2013).

Of the two potential mechanisms, in situ reactions caused by the addition of CO_2 has received the most research attention. This work can be grouped based on the different approaches taken, including geochemical and transport modelling (Apps et al., 2010; Birkholzer et al., 2008; Romanak et al., 2012; Zheng et al., 2009; Zheng et al., 2013), laboratory batch experiments (Carroll and Knauss, 2005; Humez et al., 2013; Little and Jackson, 2010; Mickler et al., 2013; Montes-Hernandez et al., 2013; Wunsch et al., 2013b), field experiments involving the injection of CO_2 gas or CO_2 charged water (Cahill and Jakobsen, 2013; Kharaka et al., 2010; Mickler et al., 2013; Trautz et al., 2013; Yang et al., 2013; Zheng et al., 2012), and the study of natural analogues (Keating et al., 2013a); many of these studies also combine approaches, such as the geochemical modelling of field experimental results. In addition to these studies on potential impacts, the chemistry and isotope chemistry of potable aquifers have been monitored above sites where CO_2 is actively injected (or produced from deep reservoirs) to look for signs of leakage and impact (Caritat et al., 2013; Lions et al., 2013; Romanak et al., 2012); to date none of these sites have shown any signs of leakage or impact.

All of these studies have highlighted the site specific nature of any potential impact, and how a detailed characterisation and baseline hydrogeological and hydrogeochemical survey of any potable aquifer above the injection reservoir is paramount for understanding any potential risk. Site specific parameters will include, amongst others:

- Aquifer mineralogy. The types of minerals will control the pH and Eh buffering capacity of the aquifer, while the chemical composition of the minerals (and their solubility under the given pH-Eh conditions) will control the availability for potential major and trace element release into the dissolved phase.

- Confined versus unconfined aquifer. In an unconfined aquifer system the leaking gas is able to migrate into the unsaturated horizon and eventually into the atmosphere, meaning that it will not accumulate (other than directly along the flow path) and that the amount of CO_2 dissolved in the water, and its eventual impact, will be more limited. In contrast, leakage into a confined aquifer will result in "pooling" of a separate gas phase at the top of the aquifer and the creation of a long-term source over a much wider area.
- Leakage rate versus groundwater flow rate. The balance between these two parameters (together with the previous point) will greatly influence the spatial extent of impact, and will also influence the resultant chemical changes (e.g. dilution, etc.)
- Chemical composition of leaking gas. The impact of only CO_2 on an aquifer will principally be in the area of a lowering of the pH, however if a significant percentage of the leaking gas consists of reduced gases like CH_4 or H_2S there is also the potential to change the Eh of the system. This will impact on the stability of various mineral phases, making them more or less soluble and changing the potential for trace element release.

The principal concern regarding this type of impact is the potential increase of various trace element concentrations above accepted drinking water standards, often quoted as the Maximum Contaminant Level (MCL) for each regulated element. The concentration of any given trace element in the dissolved phase (and its mobility) will be controlled by two main mechanisms: i) mineral dissolution or precipitation; and ii) adsorption or desorption on mineral surfaces like clays, oxides, or oxyhydroxides. Both of these processes are controlled by pH and Eh, and by the overall chemistry of the water (e.g. presence of competing ions, complexing agents, etc.). While both are important, it should be noted that the latter tends to have more rapid kinetics.

As conditions tend to change down-gradient, due to mixing with the background water and various reactions, a plume of impacted water within an aquifer will ususally be zoned, with some conservative elements (like Cl) migrating essentially at the same velocity as the groundwater flow rate, while those elements affected by solubility and/or sorption processes will tend to migrate much more slowly (often being released in one zone but then newly immobilized in another). As such the issue of "impact" must be considered not only in terms of chemical changes, but also on the spatial extent and distribution of such changes for each individual chemical species of interest.

The second impact mechanism, that is the leakage of brine associated with the storage reservoir, is of concern because the pressure gradient induced by the CO_2 injection process (i.e. the driving force that could potentially cause the brine to migrate upwards) will occur much faster and cover a much wider area than that occupied by the injected CO_2 itself. This implies, especially in a multi-well injection field, that initial site characterisation will have to ensure storage integrity also in the far field and that monitoring of groundwater resources may have to cover a wider area. Some researchers contend that this mechanism has a greater potential to impact the water quality of an overlying aquifer (e.g., Keating et al., 2013a). This topic has received less focus than the gas-only leakage scenario, possibly due to the greater difficulty of realistically simulating this process in laboratory or field-based experiments. In fact most of the literature on this item tends to focus on groundwater modelling (Birkholzer et al., 2011; Bricker et al., 2012; Cihan et al., 2012; Noy et al., 2012) and natural analogues (Keating et al., 2013; Kirk et al., 2009; Wigley et al., 2013), while known brine chemistries (Wunsch et al., 2013) or reacted reservoir rocks (Karamalidis et al., 2013) have been used to examine possible source input characteristics.

While most of the processes described above regarding the mechanisms that control element mobility will also play an important role in this second type of impact, there are some additional

features that should be considered. The most important of these differences is the fact that any leaking brine will have a chemistry that is potentially very different from that which could be developed via in situ reactions in the aquifer. In other words, whereas the potential for eventual maximum dissolved concentrations of a given trace metal or other inorganic pollutant for the CO_2-only leakage case will be a function of what is available in the aquifer minerals, brine leakage can transport elevated concentrations of major (e.g. nuisance chemicals like Cl) or trace (toxic elements like As) components that may be present in the aquifer in only very low concentrations. Other important issues for this type of scenario include the potential for the brine to have a more significant impact (chemically and spatially) on the redox conditions of an aquifer compared to the gas-only scenario, the importance of brine leakage rate versus groundwater flow rate, and the potential impact that the brine may have not only on water quality for potable resources but also on water used for irrigation (Wunsch et al., 2013).

The present research addresses a number of the issues raised above through the study of the Latera natural laboratory site in central Italy. This site consists of deep carbonate units that are overlain by volcanic lithologies and surface alluvial and fluvial sediments formed from the same volcanic rocks. At this location gas (>95% CO_2) is leaking over a wide area, however there is no clear indication of co-migration of deep water, meaning that any observed changes in chemistry may be due to gas-induced in situ reactions only. A series of boreholes were augered by hand along a transect parallel to the groundwater flow direction, with sampling points located up gradient, within, and down gradient of the leakage area. Samples were analysed in the laboratory for major and trace elements, whilst some physical-chemical parameters, such as pH and temperature, were measured in the field.

The goal of this work was to determine the chemical and spatial impact of CO_2 leakage on the groundwater quality of this volcanic terrain, both in terms of better understanding such natural hydrogeological systems as well as using this site to predict the potential of a hypothetical CO_2 leak from a CCS reservoir on potable groundwater resources.

6.2. Study site

The Latera groundwater study site was chosen based on a groundwater flow direction that passes through a CO_2 leakage area (to observe potential chemical impacts caused by the leak) and on the occurrence of a relatively shallow water table (to facilitate field logistics, considering that all holes were hand augered). As shown in Figure 14, this site is located in the centre-south of the caldera near a large CO_2 leakage area. The site is at the intersection of two roads, thus easily accessible, and is crossed by a 2-3m deep creek gulley which shows water table depth, water flow direction, and which exposes outcrops of the surface geology of the area.

Some examples of the lithologies exposed in the gulley walls are shown in Figure 60. Units are principally recent alluvial deposits formed by erosion of the surrounding volcanic rocks. Fine grained clayey sediments are most common, sometimes in multi-coloured bands of 10-20 cm thickness (Figure 60a) but most often as finely laminated, grey beds (Figure 60b), although poorly sorted units with gravel clasts supported in a silt matrix (Figure 60c) and higher energy river bottom deposits (Figure 60d) can also be observed. In terms of ground water flow and hydrogeology, at the local scale the dominant fine-grained material will represent aquitards while the more gravel-rich layers can be considered “low-permeability” aquifers that are likely discontinuous and irregular in space.

Figure 60. Exposures of the surface lithologies at the Latera study site. (a) multicoloured alluvial clays; (b) clay exposure in area of main gas leak; (c) angular gravel clasts in clay rich matrix; and (d) bottom clays cut by fluvial deposits with gravel and cobbles.

In addition to the regional soil gas survey that was conducted throughout the caldera (Figure 15), detailed CO_2 flux measurements have been performed around the chosen site, both by the University of Rome "La Sapienza" during the EC-funded CO_2GeoNet project (Figure 61a; S. Lombardi, unpublished data) and by the research group led by the Istituto Nazionale di Geofisica e Vulcanologia in Naples (Figure 61b; Chiodini et al., 2007). Based on the latter data, which covers a larger area, gas leakage occurs along a general NE-SW alignment (that parallels the local creek) at a rate of between 20 to 40 t d^{-1}.

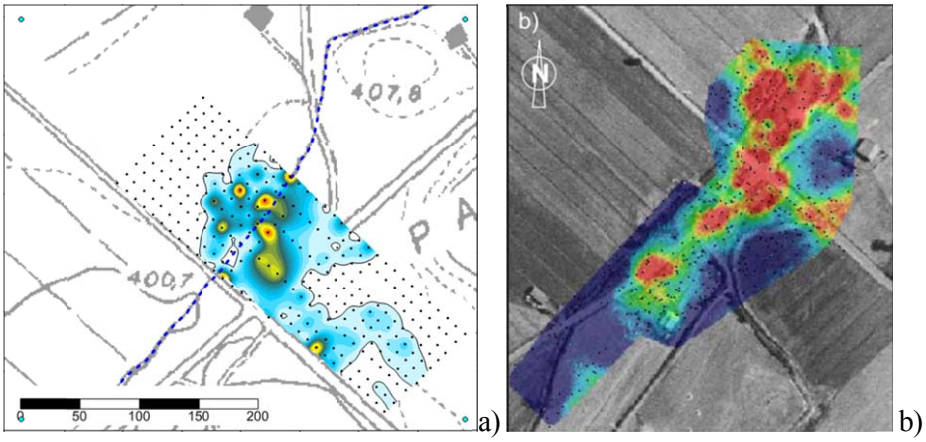


Figure 61. Gas flux surveys conducted at the site in the past by Lombardi et al. (University of Rome, unpublished data) and by Chiodini et al. (2007).

The impact of this gas leakage can clearly be seen on the ground surface. For example, the area with the strongest flux shows numerous areas where bare grey clay is exposed and no vegetation is growing (Figure 62a; located just south of the main intersection in Figure 61b). It should be noted, however, that this impact may be due to the combined effects of both CO_2 and H_2S which leak in this area. In contrast the surrounding areas, within the lower flux regions or outside of the leakage area, are cropped with wheat and other grains or used for sheep pasture (Figure 62b; located in the SE corner of Figure 61b).


a) b)

Figure 62. Photographs showing the difference in surface appearance, within the area of highest gas emission (a) and about 500 m to the SW in a pasture field with no leakage (b). For locations of piezometers P1 (a) and P5, see Figure 64a.

Based on topography and surface water flow direction in the creek, it is inferred that groundwater flow converges towards the leak and then has a general NE to SW direction (Figure 63). This figure reports the outline of the combined CO_2 flux maps given in Figure 61, and shows how it is expected that non-impacted, shallow groundwater will migrate through the leakage area and then exit on the other side to the south west.

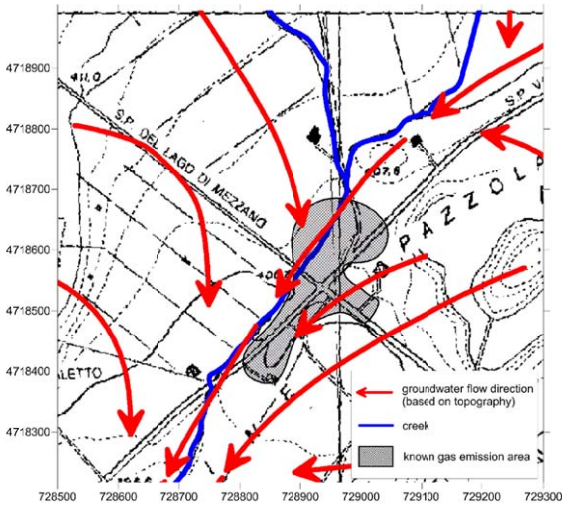


Figure 63. Groundwater flow direction estimated based on topography and surface water flow direction. Note the area in grey, which delineates the area of known CO_2 gas leakage based on the combination of the two maps presented in Figure 61.

6.3. Material and Methods

6.3.1. Sampling

The combined gas flux and topography-inferred groundwater flow directions described above were used to choose sites for both groundwater sampling (Figure 64a) and surface water measurements (Figure 64b). Groundwater sampling was conducted at 6 different points, with two directly within the gas leakage area (P1, P2), three down gradient towards the south west (P4, P3, P5), and one up gradient to the north (P6). Surface water measurements were conducted approximately every 50 m across and either side of the leakage zone. This work was performed July 2-4, 2013.

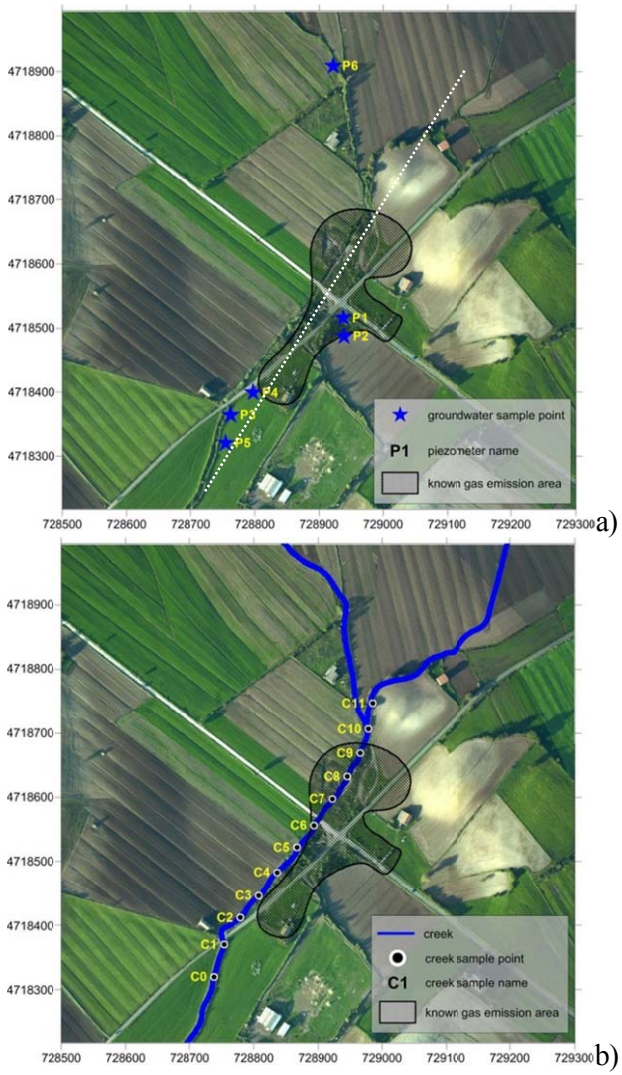


Figure 64. Maps showing groundwater (a) and surface water (b) sampling locations. Note the hatched polygon, which delineates the area of known CO_2 gas leakage based on the combination of the two maps presented in Figure 61. The dashed white line in (a) marks the central line onto which the sampling points are projected for cross-section plots, with 0 in the NE. Base map photo courtesy of M. Vellico (OGS, Trieste, Italy).

The boreholes were drilled using an extendable auger; in this work the maximum depth necessary was about 3.5m. An example is given in Figure 65a, which shows the entire length of the auger used to drill the hole at P3 in front of the creek gulley with the exposed lithology. After augering, each hole was first purged using a hand bailer to remove the water with the most suspended material (Figure 65b), followed by subsequent purging using a peristaltic pump. The first five boreholes were drilled over 2 days, then on the third day all wells were purged again and sampled; only P6 was drilled and sampled in the same day, one week later. Because of the lack of permission to leave material on site, a piezometer was lowered to the bottom for each hole for final purging and sampling and then removed when completed. Note that a sand pack was not used at this site to minimise potential changes to groundwater chemistry that may be caused by introducing material that could react with the highly acidic waters encountered in some holes.


a)


b)

Figure 65. Photographs showing the full length of the auger after drilling of P3 in the pasture field south of the emanation area (a), and use of a bailer to clean muddy water from a hole prior to pumping(b).

Water was pumped to surface into a plastic container and measured for pH, dissolved oxygen, temperature, and conductivity using a PCD650 field probe (Eutech). An inertia ball-valve pump was used to bring the water level near surface in a wide bore tube, into which a narrow bore tube attached to a 60 ml plastic syringe was inserted. The syringe was then filled and a 0.45 um pore diameter, regenerated cellulose syringe filter was added. About 10ml of sample was passed through the filter to rinse it, and then two 10ml samples were filtered into separate acid-washed HDPE bottles for anion and cation analyses; the small sample volume was necessitated by the highly turbid nature of the groundwater and the associated difficulty in filtering. The cation bottle was acidified with ultrapure nitric acid to 0.4% v/v to ensure stability during storage. Samples were also collected for silica and for dissolved hydrogen sulphide. For silica, a 10 ml sample was filtered with the same filter into a plastic vial; these samples were acidified to 1% v/v HCl and left open for 1 day to remove any H_2S present, because high concentrations of this gas interferes with the silica analytical method used. Samples for H_2S analysis instead consisted of 8ml of unfiltered well water, stabilized with 0.4ml of zinc acetate gelatine solution (50 mg of gelatine dissolved in 25 cm^3 O_2-free water, plus 261 mg ZnAc * 2 H_2O). All sample bottles were stored in a refrigerator at 4°C prior to analysis.

Finally dissolved gas samples were collected by filling a 12 mm OD PE tube using the inertial pump; as the groundwater sample is pushed from below with this technique there is minimal degassing (which tends to occur with systems that rely on a vacuum, like peristaltic pumps). The sample was transferred from the tube to a 45ml amber VOA vial by attaching a 6mm OD transfer tube and gently inclining it so as to minimise turbulence that may cause degassing. Sample vials were capped with Teflon-coated, silicon septa and stored at 4° C prior to analysis.

To examine the potential impact of CO_2 leakage (and possible co-migrating deeper waters) on the chemistry of surface water, a preliminary survey of field-measured parameters was undertaken of the creek that passes through the gas leakage zone. Measurements were performed along the creek (Figure 66a) by immersing the Eutech PCD650 probe in situ and leaving it to equilibrate for about 2 minutes prior to noting the values for the various parameters, as shown at the site of highest gas bubbling (Figure 66b) which is located at site C8 (Figure 64).

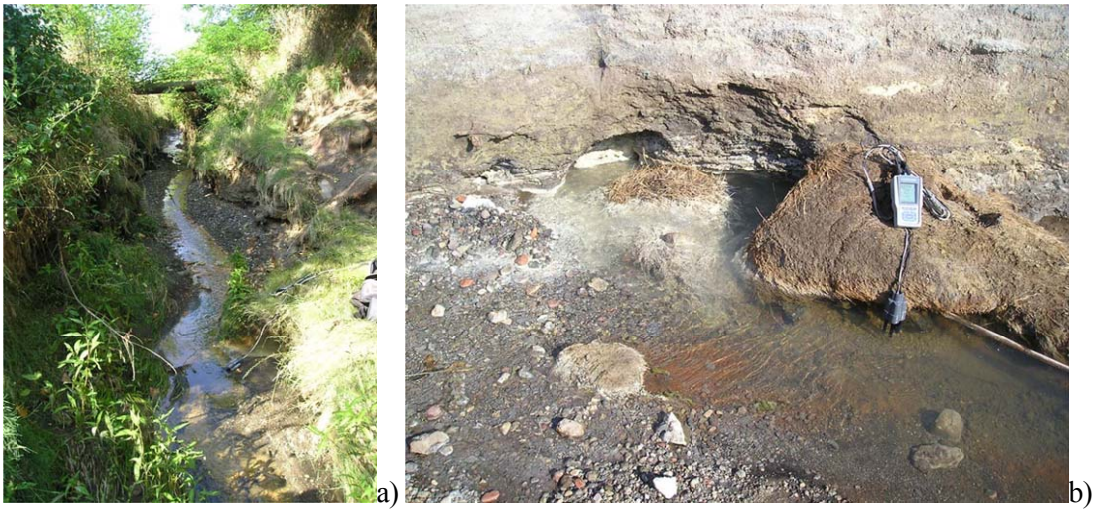

a) b)

Figure 66. Photographs of the creek along which field analyses were conducted. (a) intermediate point along the transect; (b) probe in water adjacent to a strong gas leak bubbling on the edge of the creek.

6.3.2. Analytical Methods

The headspace technique was used for the quantification of dissolved gases (Capasso & Inguaggiato 1998). Analysis involved the injection of 10 mL of pure helium gas into the VOA sample vial through the septum (displacing an equal volume of water through a second needle), an equilibration period of at least 12 hours, and analysis of aliquots of the head-space gas using a gas chromatograph. Samples were analysed for CO_2 (with additional analyses for O_2, N_2, and CH_4) on a 8000 Series Carlo Erba gas chromatograph equipped with TCD and FID detectors.

Both major anion (Cl, F, SO_4, NO_3) and cation (Li, NH_4, Ca, Mg, K, Na and Sr) concentrations were measured separately with a CX-500 ion chromatograph (Dionex Ltd.). Each element was calibrated using at least three different concentrations which bracketed the expected values of the dissolved components. Calibrations were re-done periodically if the accuracy of quality control samples (which were analysed each day along with the samples) deviated by +/- 10%. Due to the high concentrations encountered, all samples were diluted 1:200. Trace element analyses were performed on a Thermal, X series II ICP-MS (inductively coupled plasma – mass spectrometer). The instrument was used in sweep mode for the analysis of a total of 29 elements, from lithium up to uranium.

Both hydrogen sulphide and silicate were analysed using colorimetric methods. For hydrogen sulphide analyses, 10 ml of colorant (400 mg N,N-Dimethyl-1,4-phenylenediamine-dihydrochloride dissolved in 100 ml 6N HCl) and 10 ml of catalyst (1.6 g $FeCl_3$ * 6 H_2O dissolved in 100 ml 6N HCl) were added to the stabilised 1 ml of sample, shaken, and left for one hour to develop the colour. Samples were transferred into micro-volume (1 ml) disposable PS cuvettes and analysed at a wavelength of 670 nm. For silicate, 0.1 ml of heptamolybdate solution (6.3 g ammoniumheptamolybdate-tetrahydrate dissolved in 50 ml pure water added to 50 cm^3 of 4.5M H_2SO_4) was added to 5ml of sample. This was shaken and left for 30 minutes. 0.1 ml oxalic acid (10 g oxalic acid in 100 ml pure water) and 0.1 ml ascorbic acid (2.8 g ascorbic acid in 100 ml pure water) were then added, the vials shaken again and left for another 30 minutes. Absorbance was measured at a wavelength of 810 nm.

6.4. Latera groundwater results and discussion

6.4.1. Borehole log and piezometric surface results

Boreholes were typically augered to about 2.5 to 3.5 m to intersect the water table. Impermeable clayey units were encountered in all holes above the water bearing interval, thus the sampled unit can be considered confined. A short description of each hole follows (see Figure 64a for locations).

Hole P1. Located within the gas leakage area (see photo Figure 62a) on a site with no vegetation, exposed grey soil, and the smell of H_2S in the air. Lithology: grey clay from 0 to 55 cm; very dark grey to black, compact clay from 55 to 202 cm; end of hole at 202 cm. First sign of water at 165 cm. Initial water was black. At the end of sampling the hole was re-filled with unaltered, brown- to grey-coloured sandy sediments from the creek bed. One week later the hole was re-excavated and this material was found to be completely black, indicating the highly reducing nature of the system, the impact of H_2S and, possibly, the precipitation of iron sulphides (such as amorphous iron (II) monosulfide (FeS_{am}), mackinawite (FeS) or greigite (Fe_3S_4)).

Hole P2. Located about 30 m south of the non-vegetated leakage area, on the boundary between an agricultural field and scrub brush. The vegetation appears healthy and there is no visual indication of gas leakage or impact on surface. Lithology: grey soil from 0 to 50 cm; grey clay with red stringers (that increase in amount with depth) from 50 to 140 cm; red clay from 140 to 200 cm; red sand- to gravel-sized fragments in clay matrix from 200 to 278 cm; end of hole at 278 cm. End of hole very resistant to augering, possibly top of volcanic bedrock. First sign of water at 183 cm. Initial water was red.

Hole P3. Located 5 m from the creek, with creek gulley giving a good exposure of the lithology of the first c. 2.5 m (see photos in Figure 60a and Figure 65a), and normal vegetation in this pasture field (see photo in Figure 62b). Lithology: dark brown clay from 0 to 100 cm; compact black clay from 100 to 138 cm; red clay from 138 to 180 cm; grey clay from 180 to 230 cm; grey clay with lithoid fragments from 230 to 310 cm; end of hole at 310 cm. First sign of water at 220 cm. Initial water was brown.

Hole P4. Located at the northern end of the same field as P3, chosen as an intermediate point between the highly anomalous P2 and the more normal P3. Lithology: brown soil from 0 to 143 cm; brown clay with red inclusions and stringers from 143 to 209 cm; grey to brown, wet clay

from 209 to 230 cm; dry, compact black clay from 230 to 233 cm; end of hole at 233 cm. First sign of water at 209 cm. Initial water was black to dark brown.

Hole P5. Located 40 m south of P3 in the same pasture field, 10 m from the creek gulley. Lithology: brown soil from 0 to 90 cm; plastic grey clay with intervals with red stringers from 90 to 282 cm; grey sandy clay with lithoid fragments from 282 to 294 cm; grey clay from 294 to 300 cm; end of hole at 300 cm. First sign of water at 270 cm. Initial water was grey to brown.

Hole P6. Located about 250 m north of the known gas leakage area, chosen as a background, up-gradient site. 5 m west of a shallow, dry creek bed. Lithology: dark brown clayey soil from 0 to 120 cm (note that a tube was lowered to 120 cm and CO_2 was measured directly using a field instrument – a value of 3.6 % was observed, which is higher than the 1-2% that might be expected in a normal soil at that depth); light grey clay from 120 to 198 cm; compact grey clay with small intervals with lithoid fragments from 198 to 225 cm; grey sandy clay from 225 to 280 cm; gravel in sandy-clay matrix from 280 to 290 cm; end of hole at 290 cm. First sign of water at 242 cm. Initial water was grey, after being left for the rest of the day it was clear.

All wells were purged 2 to 3 times then left overnight prior to measuring the piezometric surface, followed by an additional single purge prior to sampling. Note that all wells had a very slow recharge rate indicating the low permeability of the measured hydrological units.

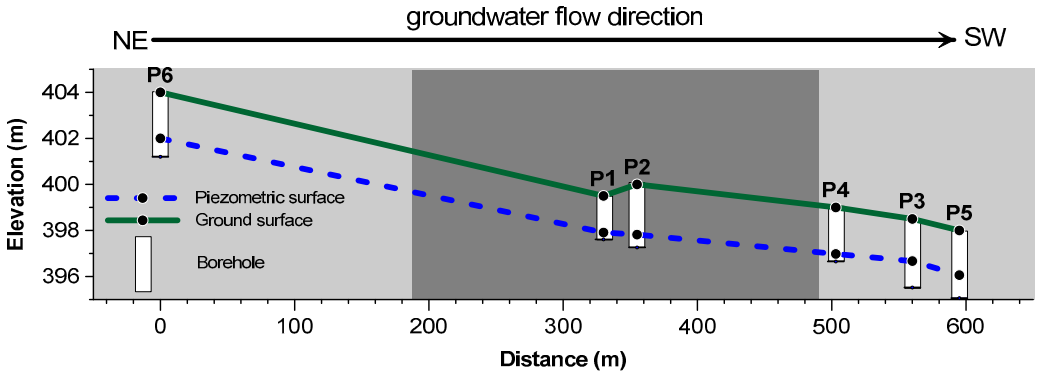


Figure 67. Ground surface and piezometric surface along a central line through the study area onto which the various sample points are projected. See Figure 64a for location.

Figure 67 shows the piezometric surface along a central projection line crossing the study area from NE to SW (see Figure 64a for location). In agreement with the initial interpretation of groundwater flow direction based on topography, the borehole data indicates a general NE to SW hydraulic gradient that flows through the main gas leakage area. The gradient from P1 to P5 is on the order of 0.007. Note that because of the highly variable nature of the surface lithologies it was not possible to create a geological cross-section.

6.4.2. Data Statistics

A correlation matrix for all collected data is given in Table 4. The colour coding clearly shows the strong negative correlation between many parameters and pH that is typically on the order of about -0.7. Strong (0.7 to 0.9) and very strong (>0.9) positive correlations are observed for numerous other parameters, including SiO_2, SO_4, K, Hg, and Fe. It must be noted that part of the reason for such a good correlation is the very wide range of values along the transect, especially due to the very high values occurring in the gas leakage area, as discussed below.

Table 4. Correlation matrix of r^2 values for all chemical - physical parameters measured on the Latera samples. Note legend for explanation of colour coding.

	pH	El. Cond.	Alk.	Ionic Str.	Ag	Al	As	B	Ba	Ca	Cd	CH4	Cl	CO2	Cr	Cs	Cu	Fe	H2S	HCO3	Hg	K	Li	Mg	Mn	Na	Ni	NO3	Pb	Rb	SiO2	SO4	Sr	Zn	U	Be	V	Co	Sb	Ga	Mo
pH	1.00																																								
El. Cond.	-0.74	1.00																																							
Alk.	0.63	-0.17	1.00										Legend																												
Ionic Str.	-0.71	1.00	-0.23	1.00									0.9 to 1.0																												
Ag	-0.50	0.84	0.05	0.80	1.00								0.7 to 0.9																												
Al	-0.72	0.96	-0.30	0.98	0.92	1.00							0.7 to 0.9																												
As	0.11	0.41	0.82	0.33	0.48	0.23	1.00						-0.9 to -1.0																												
B	-0.72	0.38	-0.45	0.39	0.35	0.41	-0.16	1.00																																	
Ba	0.55	-0.19	0.98	-0.26	-0.03	-0.35	0.87	-0.35	1.00																																
Ca	-0.66	0.94	0.04	0.90	0.75	0.89	0.59	0.24	0.05	1.00																															
Cd	-0.62	0.85	-0.09	0.80	0.97	0.91	0.35	0.58	-0.13	0.68	1.00																														
CH4	-0.74	0.82	-0.42	0.81	0.47	0.73	0.11	0.16	-0.41	0.83	0.42	1.00																													
Cl	-0.58	0.45	0.12	0.40	0.42	0.34	0.44	0.78	0.25	0.49	0.55	0.15	1.00																												
CO2	-0.79	0.83	-0.40	0.82	0.48	0.74	0.14	0.17	-0.38	0.85	0.43	1.00	0.19	1.00																											
Cr	-0.71	0.96	-0.33	0.98	0.85	0.97	0.20	0.29	-0.39	0.83	0.80	0.83	0.21	0.83	1.00																										
Cs	-0.71	0.98	-0.19	0.99	0.95	0.99	0.35	0.44	-0.23	0.85	0.93	0.70	0.44	0.72	0.95	1.00																									
Cu	0.00	-0.39	-0.39	-0.33	-0.14	-0.17	-0.62	0.52	-0.39	-0.64	0.06	-0.57	0.07	-0.58	-0.32	-0.21	1.00																								
Fe	-0.71	0.95	-0.36	0.96	0.73	0.92	0.21	0.26	-0.38	0.88	0.69	0.94	0.24	0.94	0.97	0.90	-0.45	1.00																							
H2S	-0.42	0.26	-0.35	0.23	-0.22	0.09	-0.08	-0.12	-0.27	0.42	-0.27	0.75	-0.06	0.74	0.25	0.06	-0.64	0.48	1.00																						
HCO3	0.61	-0.16	1.00	-0.23	0.03	-0.30	0.83	-0.44	0.99	0.05	-0.10	-0.41	0.13	-0.38	-0.34	-0.20	-0.41	-0.35	-0.32	1.00																					
Hg	-0.71	0.99	-0.21	1.00	0.91	0.99	0.34	0.34	-0.26	0.88	0.87	0.79	0.34	0.80	0.98	0.99	-0.33	0.95	0.18	-0.21	1.00																				
K	-0.80	0.95	-0.32	0.94	0.70	0.85	0.27	0.31	-0.31	0.92	0.66	0.95	0.35	0.96	0.93	0.88	-0.49	0.99	0.53	-0.31	0.92	1.00																			
Li	-0.83	0.80	-0.43	0.83	0.80	0.87	0.05	0.79	-0.42	0.60	0.92	0.50	0.61	0.51	0.77	0.87	0.23	0.70	-0.09	-0.43	0.81	0.70	1.00																		
Mg	-0.52	0.83	0.17	0.80	0.63	0.67	0.68	0.06	0.19	0.98	0.52	0.80	0.41	0.82	0.72	0.72	0.72	0.84	0.50	0.19	0.72	0.80	0.41	1.00																	
Mn	-0.34	0.30	0.35	0.24	0.34	0.18	0.57	0.64	0.48	0.38	0.45	-0.04	0.96	0.00	0.03	0.29	0.05	0.06	-0.19	0.37	0.19	0.17	0.44	0.34	1.00																
Na	-0.34	0.73	0.47	0.66	0.64	0.54	0.89	0.12	0.51	0.87	0.55	0.51	0.62	0.54	0.52	0.64	-0.66	0.58	0.21	0.48	0.64	0.65	0.38	0.90	0.63	1.00															
Ni	-0.50	0.04	-0.35	0.05	0.10	0.10	-0.26	0.93	-0.24	-0.10	0.35	-0.19	-0.70	-0.18	-0.06	0.13	-0.72	-0.09	-0.32	-0.35	0.00	-0.05	0.58	-0.26	0.62	-0.10	1.00														
NO3	0.47	-0.77	-0.20	-0.71	-0.57	-0.60	-0.69	-0.15	-0.26	-0.87	-0.50	-0.69	-0.55	-0.72	-0.60	-0.67	0.65	-0.74	-0.42	-0.22	-0.68	-0.78	-0.44	-0.89	-0.51	-0.92	0.10	1.00													
Pb	-0.32	0.05	0.10	0.02	0.14	0.01	0.19	0.79	0.24	0.04	0.34	-0.27	0.89	-0.24	-0.16	0.10	0.46	-0.17	-0.35	0.11	-0.03	-0.08	0.43	-0.05	0.90	0.24	0.88	-0.16	1.00												
Rb	-0.74	0.98	-0.17	0.99	0.93	0.98	0.38	0.44	-0.20	0.87	0.92	0.73	0.47	0.74	0.95	1.00	-0.25	0.91	0.11	-0.18	0.99	0.90	0.87	0.76	0.33	0.68	0.13	-0.72	0.12	1.00											
SiO2	-0.89	0.90	-0.56	0.92	0.66	0.87	0.01	0.47	-0.55	0.79	0.69	0.91	0.33	0.91	0.92	0.87	-0.23	0.96	0.47	-0.55	0.89	0.96	0.81	0.67	0.11	0.45	0.14	-0.63	-0.02	0.89	1.00										
SO4	-0.72	0.97	-0.38	0.98	0.86	0.99	0.17	0.42	-0.42	0.82	0.85	0.81	0.33	0.82	0.99	0.97	-0.22	0.96	0.22	-0.38	0.98	0.93	0.86	0.70	0.14	0.52	0.10	-0.62	-0.02	0.97	0.95	1.00									
Sr	0.51	-0.25	0.61	-0.26	0.16	-0.17	0.37	0.02	0.58	-0.32	0.18	-0.72	0.22	-0.70	-0.35	-0.11	0.40	-0.49	-0.88	0.59	-0.23	-0.50	-0.02	-0.34	0.41	0.06	0.24	0.10	0.46	-0.12	-0.51	-0.30	1.00								
Zn	-0.54	0.12	-0.42	0.15	0.20	0.22	-0.30	0.94	-0.34	-0.07	0.45	-0.12	0.66	-0.11	0.05	0.23	0.76	0.00	-0.34	-0.43	0.11	0.03	0.67	-0.25	0.56	-0.11	0.98	0.06	0.82	0.23	0.24	0.21	0.25	1.00							
U	0.34	0.04	0.93	-0.04	0.20	-0.12	0.91	-0.15	0.96	0.25	0.11	-0.27	0.45	-0.23	-0.19	0.01	-0.38	-0.19	-0.28	0.94	-0.04	-0.11	-0.17	0.35	0.65	0.68	-0.11	-0.44	0.37	0.04	-0.34	-0.20	0.59	-0.18	1.00						
Be	-0.76	0.46	-0.47	0.49	0.50	0.55	-0.16	0.97	-0.41	0.25	0.78	0.17	0.74	0.18	0.40	0.56	0.57	0.34	-0.24	-0.47	0.45	0.36	0.85	0.05	0.56	0.11	0.85	-0.16	0.71	0.55	0.53	0.54	0.11	0.93	-0.20	1.00				
V	-0.61	0.85	-0.35	0.86	0.61	0.81	0.15	0.01	-0.40	0.81	0.51	0.94	-0.03	0.94	0.92	0.78	-0.57	0.96	0.56	-0.34	0.86	0.92	0.50	0.77	-0.20	0.47	-0.35	-0.60	-0.45	0.78	0.87	0.86	-0.61	-0.25	-0.27	0.09	1.00			
Co	-0.46	-0.06	-0.31	-0.05	-0.03	-0.02	-0.26	0.89	-0.18	-0.15	0.22	-0.22	0.71	-0.21	-0.17	0.01	0.69	-0.17	-0.24	-0.31	-0.11	-0.11	0.47	-0.28	0.64	-0.11	0.99	0.10	0.89	0.02	0.07	-0.02	0.20	0.95	-0.07	0.81	-0.42	1.00		
Sb	-0.34	0.77	0.26	0.71	0.97	0.83	0.61	0.31	0.19	0.67	0.93	0.27	0.47	0.30	0.70	0.85	-0.12	0.56	-0.38	0.24	0.79	0.54	0.73	0.57	0.45	0.67	0.12	-0.53	0.24	0.84	0.47	0.73	0.37	0.19	0.40	0.45	0.43	-0.01	1.00	
Ga	-0.88	0.77	-0.33	0.76	0.78	0.89	0.13	0.68	-0.30	0.59	0.91	0.43	0.72	0.45	0.70	0.89	0.24	0.64	-0.13	-0.33	0.76	0.65	0.99	0.41	0.58	0.44	0.65	-0.47	0.56	0.84	0.74	0.80	0.07	0.72	-0.04	0.91	0.40	0.55	0.72	1.00	
Mo	0.11	0.40	0.30	0.43	0.76	0.54	0.38	0.00	0.16	0.21	0.70	-0.10	0.04	-0.09	0.43	0.55	0.13	0.23	-0.69	0.27	0.49	0.14	0.43	0.12	0.09	0.27	-0.05	-0.14	0.01	0.52	0.13	0.42	0.62	0.06	0.27	0.22	0.17	-0.19	0.83	0.41	1.00

6.4.3. Groundwater chemistry compared to regional data from the literature

Extensive geothermal research in this area during the 1970's to 1980's created various regional groundwater datasets (from different aquifers and lithologies) that can be used for comparative purposes. In addition, because of the nature of that research, chemical analyses of the brines within the >2000m deep reservoir are also available.

Bertrami et al. (1984) give a review of the geothermal exploration conducted at the Latera site by the power company ENEL, detailing results from geological, geophysical, and geochemical studies, including sampling of groundwater as well as the deep reservoir brine. These authors define three water types in the area of Latera. i) Shallow flow paths are characterised by a Ca-HCO_3 chemistry, low temperatures and low salinity, resulting from alteration of the volcanic rocks. ii) Deeper circulation in the volcanics results in waters with a similar chemistry, but with higher salinities, slightly higher temperatures, and progressive saturation with respect to gypsum; the supply of deeper geothermal fluids is not evident in these waters. iii) deep hydrothermal flow paths yield Na-Cl waters, with anomalous concentrations (as illustrated by deep well Latera 2) of Li (14 mg/l), H_3BO_3 (2000 mg/l), SiO_2 (400 mg/l) and As (70 mg/l). Figure 68 plots the shallow and deep waters from the Latera caldera reported by Bertrami et al. (1984) together with the samples collected in the present study. As can clearly be seen, the shallow samples collected in this study plot along the same trend defined by the spring waters (i.e. within the shallow volcanic units), far from the chemistry of the Na-Cl brines of the geothermal reservoir. This implies that deep water is not leaking at this study site, and that any impacts observed should be due only to the leaking gas and its influence on in situ reactions and chemistry.

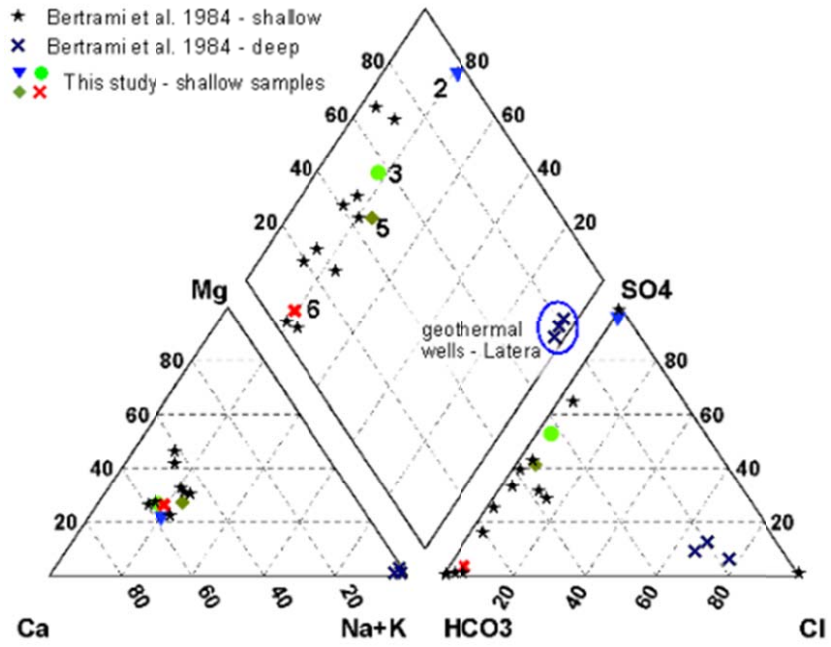


Figure 68. Piper diagram (modified after Bertrami et al., 1984) comparing the chemistry of shallow and deep groundwater samples from the literature with the samples collected in this study. Note that the collected samples are significantly different from the deep brines of the geothermal reservoir.

Similarly, comparison of collected data with the K and Na results of Duchi et al. (1987) indicate circulation in volcanic rocks (Figure 69a). In addition, results plot within the fields of the shallow volcanic units ("A" and "B") defined by Checcucci et al. (1988) using a Ludwig-Langelier plot (Figure 69b), with group B being similar to group A but with the addition of CO_2 and H_2S which increases acidity and shifts the carbonate system equilibrium away from bicarbonate towards dissolved CO_2.

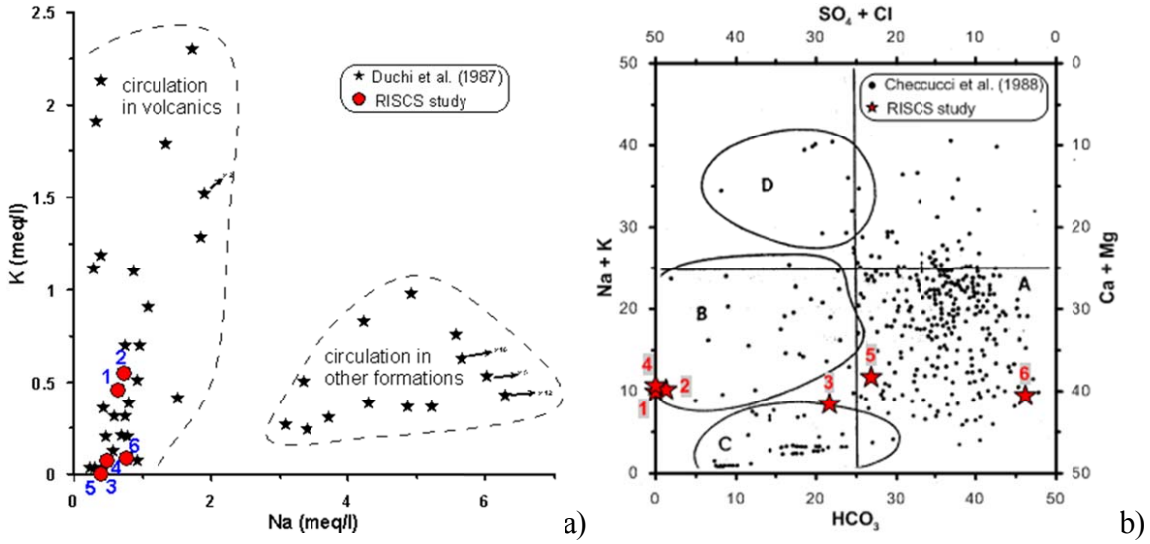


Figure 69. Results from the present study plotted against literature data. a) K versus Na (Duchi et al., 1987) shows that all collected samples lie in the field of volcanic groundwaters. b) the Ludwig-Langelier plot (Checcucci et al., 1988) shows that all samples lie in the shallow groundwaters of fields A or B.

6.4.4. Spatial distribution of groundwater chemistry

The data from the six boreholes augered at the Latera site have been projected onto a central transect that parallels the estimated groundwater flow direction, with one point (P6) up-gradient, two points within the leakage area (P1, P2), and three points progressively further down-gradient (P4, P3, P5) (see Figure 64a for location). All boreholes drilled at Latera yielded sufficient water for sampling. Figure 70 shows pH and major ion chemistry along the 600 m transect, plotted with the groundwater flow direction from left to right through the zone of gas leakage marked in dark grey. Almost all parameters show a clear trend of background values up-gradient, impacted values within the leakage area, and progressively less anomalous values moving further down-gradient of the leak.

In the up-gradient region the pH at P6 is close to neutral, after which it drops significantly to values less than 4 in the leakage area, followed by a smooth increase back up to neutral values as one moves down-gradient from the leak (Figure 70a). The very low pH values within the leakage area are due primarily to the lack of carbonate mineral buffering in the volcanic terrain at Latera. In addition, H_2S gas leaks at Latera and its oxidation contributes to acid production (see discussion below).

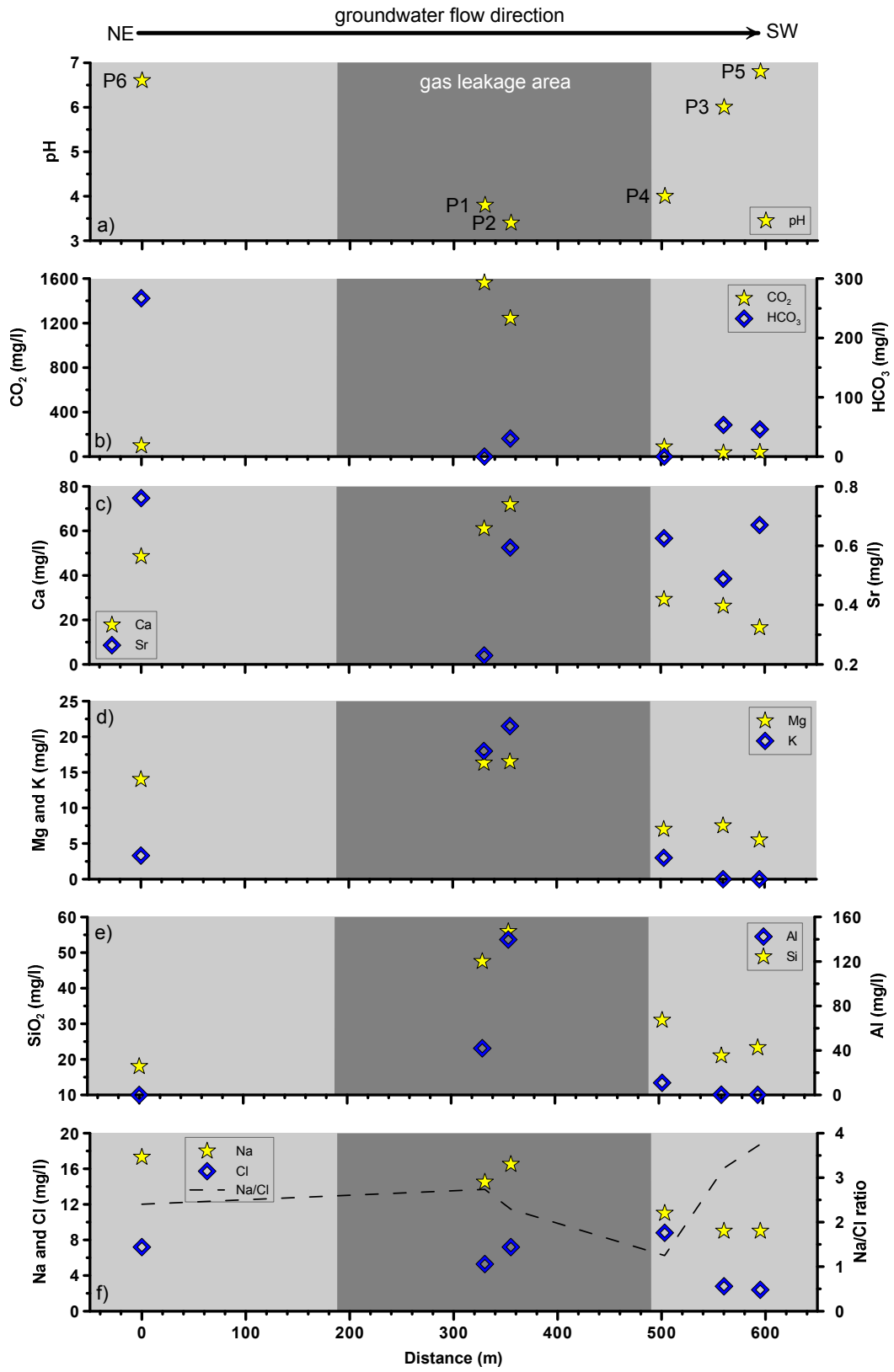


Figure 70. pH and major ion trends along the measured transect. Note the gas leakage area in dark grey and the presumed groundwater flow direction from left to right.

Figure 70b shows the values for the measured inorganic carbon species of CO_2 and HCO_3, with these two parameters having opposite trends due to the controlling influence of pH on species distribution; for example, at pH values less than 4 essentially all of the inorganic carbon is in the form of CO_{2aq} (Figure 71). Along the transect the bicarbonate ion is highest up gradient, essentially zero in the leakage area (including P4), and low in the down-gradient area, whereas CO_2 is near saturation within the leakage area with significantly lower values both up and down gradient. Note that the values for these two parameters observed in the up gradient sample at P6 are slightly higher than those expected for volcanic rocks, indicating that the chemistry at this point may be partially impacted by gas leaks that are known to occur further up slope to the north. Down gradient both parameters decrease away from the leakage area.

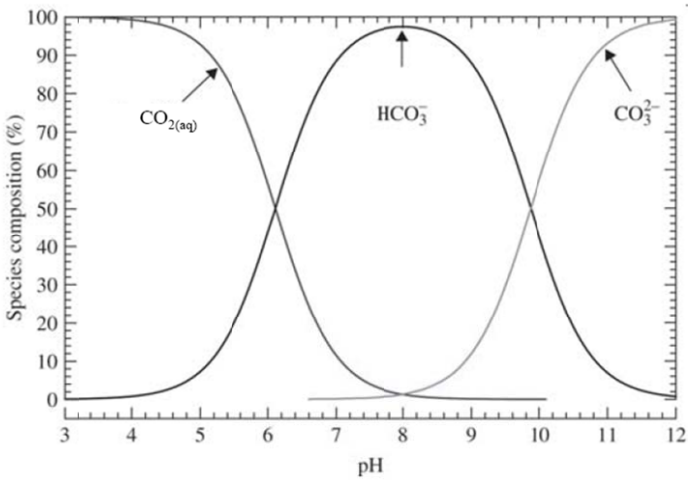


Figure 71. Bjerrum plot showing the relative amounts of the different inorganic carbon species as a function of pH. As this plot varies with salinity it is only provided as a qualitative illustration.

Of the various other ions plotted in Figure 70, basically all show highest concentrations in the gas leakage area followed by moderate to sharp decreases in the three down gradient samples. What distinguishes the different parameters is the relationship between up gradient (P6) and leakage-area (P1, P2) samples, as some show concentrations at P6 which are relatively similar to those in the leakage area (e.g. Na, Cl), those which are slightly lower (e.g. Ca, Mg), and those which are much lower (e.g. K, Al, and SiO_2). This implies that minerals containing the last group of elements (and to a lesser extent group 2) are being dissolved within the gas leakage interval. As stated, potassium (K) and a high-potassium (HK) magmatic series have been identified in this area (see section 3.1), with the pyroclastics of the former containing sanidine (a K-feldspar), plagioclase, and biotite, while the latter contains clinopyroxene, olivine, and leucite. In addition, mineralogical analyses conducted on surface sediments located about 3 km from the present study site (Beaubien et al., 2008) yielded, in decreasing order of abundance, halloysite (clay mineral), K-feldspar, augite (a clinopyroxene), albite (a plagioclase), and quartz, with minor hematite and pyrite (high background values in XRD analyses also imply significant amounts of volcanic glass).

An examination of the chemical formula of these minerals (see Table 5) shows that the dissolution of halloysite, sanidine and/or leucite could explain the increased Al, K, and Si concentrations in the leakage area relative to the up gradient sample. The slight increase in Ca and Mg in the leakage area may be due to augite dissolution, or Ca alone could be due to dissolution of secondary calcite that

may occur as a fracture filling material. Instead the slight decrease from background to leakage area concentrations for Na is curious, as one would expect this element to increase if albite (Table 5) was an important constituent.

The saturation indices (SI) of various potential "primary" mineral phases and "secondary" alteration products are plotted along the transect in Figure 72. Saturation indices for the Latera dataset were calculated using the program PHREEQC (Version 3; Parkhurst and Appelo, 2013) with the associated PHREEQC.dat thermodynamic database.

The rapid decrease of all major elements down gradient from the leakage area implies a removal mechanism from the dissolved phase, either via mineral precipitation and/or adsorption. Table 6 gives the chemical formula of various mineral phases whose solubility may control equilibrium concentrations within the leakage area and decrease dissolved concentrations down gradient.

Table 5. Types and formulas of mineral phases that likely form the volcanic-origin alluvial sediments in the study area, prior to leaking gas - induced impact.

mineral name	type	chemical formula
halloysite	clay mineral	$Al_2Si_2O_5(OH)_4 \bullet 4H_2O$
sanidine	K-feldspar	$KAlSi_3O_8$
augite	clinopyroxene	$(Ca, Na)(Mg, Fe^{2+}, Al)(Si, Al)_2O_6$
albite	plagioclase	$NaAlSi_3O_8$
leucite	feldspathoid	$KAlSi_2O_6$
hematite	oxide	Fe_2O_3
pyrite	sulfide	FeS_2

Table 6. Types and formulas of mineral phases that may represent alteration products formed within or down-gradient from the leaking gas area.

mineral name	type	chemical formula
alunite	hydroxide	$KAl_3(SO_4)_2(OH)_6$
gibbsite	hydroxide	$Al(OH)_3$
kaolinite	clay	$Al_2Si_2O_5(OH)_4$
smectite	clay	$R_{0.33}Al_2Si_4O_{10} \bullet 2H_2O$, where R = Na^+, K^+, Mg^{2+} and/or Ca^{2+}
illite	clay	$K_y(Al_4Fe_4Mg_{10})(Si_{8-y}Al_y)O_2(OH)_4$
ferrihydrite	oxide	$Fe_{10}O_{15} \bullet 9H_2O$
hematite	oxide	Fe_2O_3
magnetite	oxide	Fe_3O_4
goethite	hydroxide	FeO(OH)
iron hydroxide	hydroxide	$Fe(OH)_{3\ amorphous}$
greigite	sulfide	Fe_3S_4

Regarding the alumino-silicate mineral phases, one can see that albite, anorthite, and K-feldspar are all in equilibrium with the surrounding groundwater, all are under-saturated in the leakage area, and then all show increasing values that return to near equilibrium values at the farthest sampled point at P5. This indicates, at least from a thermodynamic point of view, that these minerals can be dissolved in the leakage area but not at the two extremes of the transect.

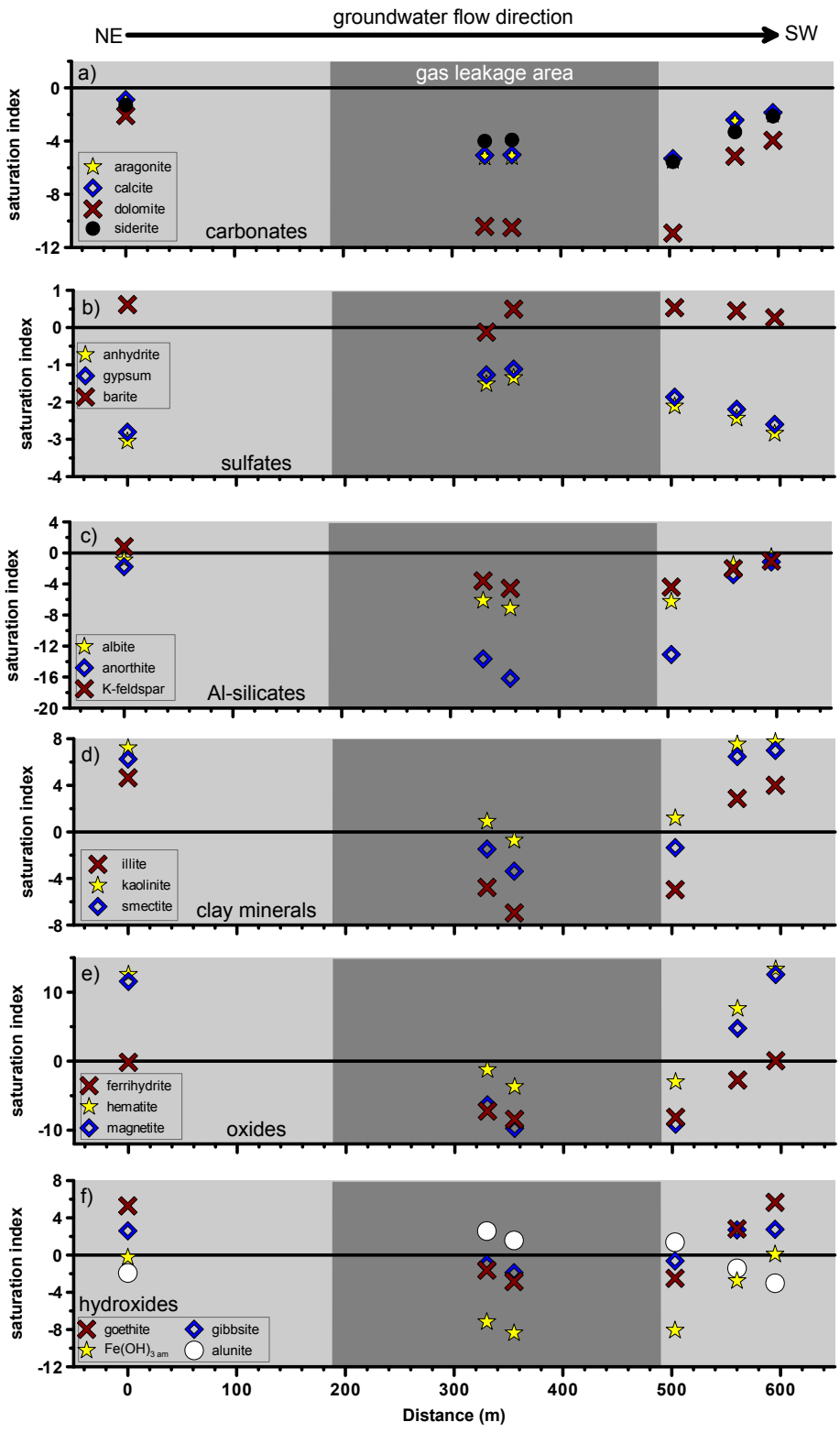


Figure 72. Saturation indices of various mineral phases along the measured transect. Note the gas leakage area in dark grey and the presumed groundwater flow direction from left to right.

For Al the most likely minerals are alunite (considering the high SO_4 concentrations), gibbsite, kaolinite, smectite, and illite, for K it would be alunite, smectite, and illite, for Si it could be any of the clay minerals, while for other cations it would be smectite. In addition, cation adsorption on the surfaces of newly formed clays and/or hydroxides is also possible. A plot of the saturation indices of these minerals (Figure 72) shows that kaolinite and gibbsite are over-saturated outside of the leakage area but near saturation within it. In contrast alunite is under-saturated outside the leakage area but over-saturated within it (note that unlike the other mineral SI values, that for alunite was calculated with the LLNL database, as the PHREEQC database shows relatively constant, positive values across the entire transect, which seemed un-realistic considering the changes in chemistry. Down gradient from the leakage area the solubility indices of all modelled clay and hydroxide minerals increase towards over saturation conditions, showing that all or some of these phases could contribute to the observed decreases in the other cation species.

Figure 70 also compares the trends of Na and Cl, as well as the ratio between them (plot "f"). The higher Na concentrations relative to Cl are an indication that the Na results from reaction with Na-bearing silicate mineral phases. The low Cl concentrations combined with the high Na/Cl ratio within the centre of the gas leakage area indicates that geothermal reservoir brine is not co-migrating with the leaking gas at this site, as discussed in Section 6.4.3. The only point outside of this trend is P4 at ~ 500 m, where the Cl concentration, while still low, is closer to that of Na, resulting in a ratio value closer to 1. The cause of this slight increase is unknown.

A number of redox sensitive species are plotted in Figure 73, including SO_4 and H_2S (b), Fe (c), Mn (d), and CH_4 (e). The overall trends of SO_4, Fe, and CH_4 are relatively similar to each other (and similar to those of Al, K, and Si; Figure 70), with low values up gradient, very high concentrations within the leakage area, and low, decreasing concentrations down gradient from the leak. The source of the Fe is likely the dissolution of pyrite and hematite (Table 5), with the latter being near or slightly below saturation within the leakage area (Figure 72e). Decreasing Fe concentrations down gradient may be controlled by precipitation of hematite, magnetite, or goethite (Table 6, Figure 72e,f). The high concentrations of SO_4 in the leakage area are probably caused by oxidation of the dissolved H_2S that is migrating with the CO_2 (for example, note that only P1 had high H_2S concentrations, whereas all other points were below detection; Figure 73b) or by dissolution of pyrite.

The rapid decrease in SO_4 concentrations down gradient cannot be explained by precipitation by common sulphate phases like gypsum and anhydrite, as these are under saturated along the entire transect with progressively more negative values moving away from the leakage area (Figure 72b). Instead, barite is at saturation levels throughout, however the low Ba concentrations preclude significant SO_4 removal. Alunite (Table 6) may be a possible removal mechanism for both Al and SO_4, given its over-saturation state within the leakage area (see above), however as stated, different SI values were observed using different thermodynamic databases. Simple dilution / spread of the plume may also explain this rapid decrease, although the slow flow rate does not, perhaps, support this hypothesis. Finally CH_4, like H_2S, originates in the leaking gas; it is only observed in the leakage area and apparently is rapidly oxidised beyond its borders.

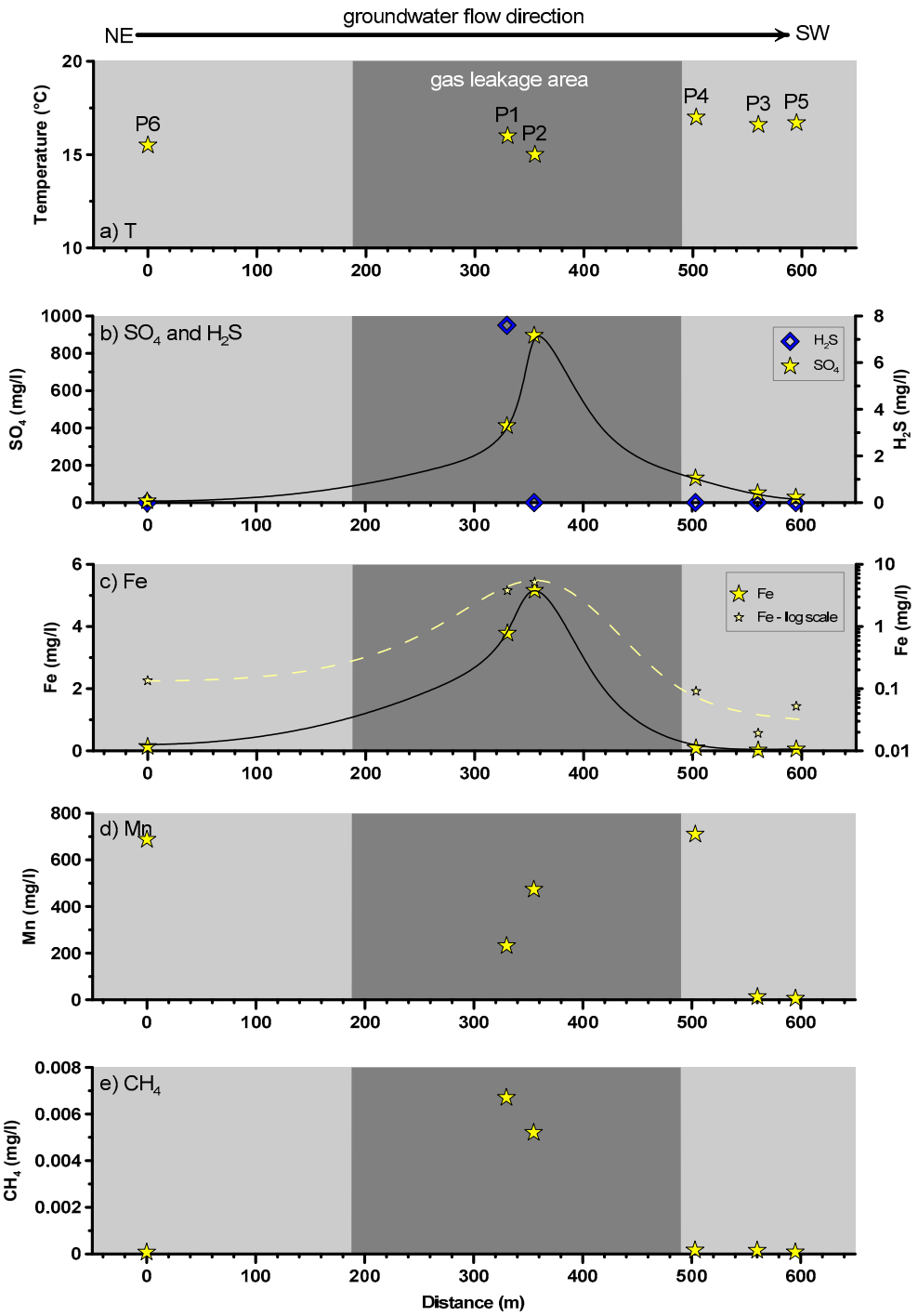


Figure 73. The trends of redox sensitive parameters along the measured transect. Note the gas leak area in dark grey and presumed groundwater flow direction from left to right.

The results of many of the trace elements analysed at this site are reported in Figure 74, grouped into three separate plots based on the general trend of the data. Note that despite the different behaviours of the three groups, all trace elements reported in this figure have down gradient concentrations that decrease towards, or below, the concentrations observed in the up gradient background sample, despite elevated concentrations in the leakage area.

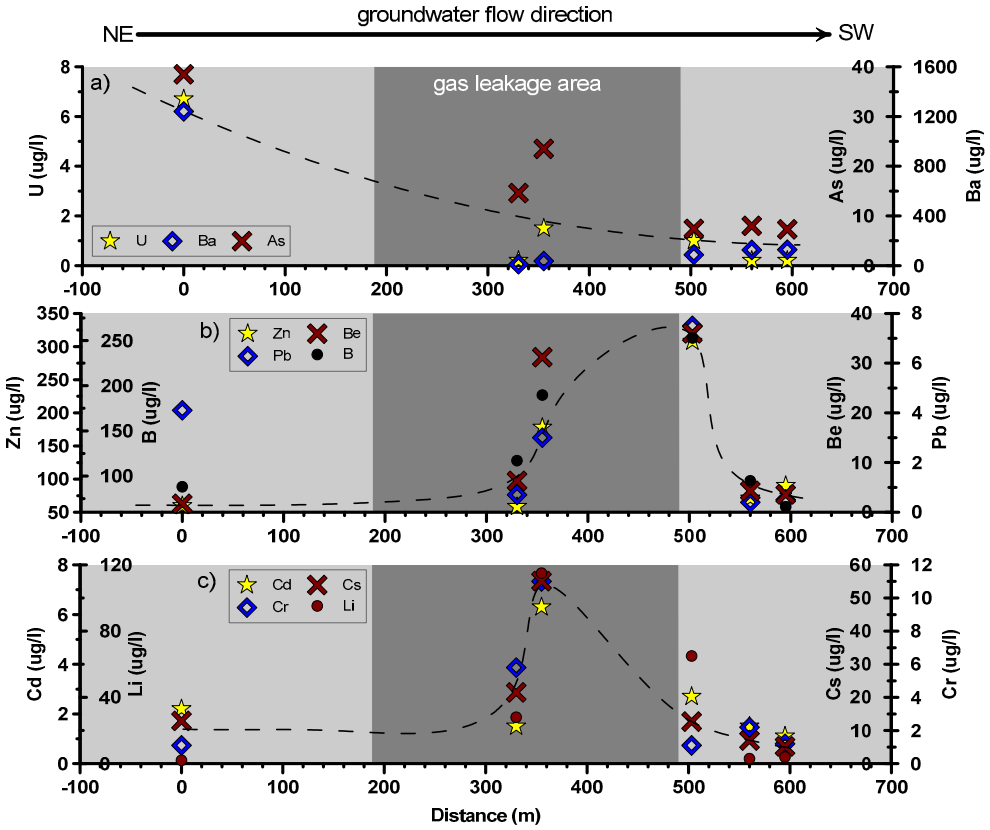


Figure 74. Trace element trends along the measured transect. Note the gas leakage area in dark grey and the presumed groundwater flow direction from left to right.

The trace elements in Figure 74a show a general decreasing trend along the entire transect, with background values higher than leakage area values. Uranium will occur as U(VI) under reducing conditions and will likely precipitate out as the oxide uraninite. The uraninite stability field occurs over a wide pH range, helping to explain why concentrations continue to decrease down gradient to 0.2 ppb in the furthest sample (P5), where pH values approach neutral. In addition, the low concentrations of complexing agents in this interval (such as bicarbonate) that could potentially increase U mobility, may contribute to the lower concentrations. Decreasing As concentrations could also be linked to redox conditions, with the possible precipitation of orpiment. Ba is the only element which goes from a high background value (>1200 ppb), to almost 0 in the leakage area (<30 ppb), and then a slight increase down gradient (c. 130 ppb). This behaviour could be explained by barite saturation conditions along the entire transect combined with the fact that of all divalent cations, Ba has the highest affinity for adsorption.

The trace elements given in Figure 74b show relatively low values in the first leakage point (P1), anomalously high values in the second leakage point (P2), and the highest values in the point sampled directly down gradient from the leakage area (P4), before rapidly decreasing in the last two points down gradient (P3, P5). P4 is noteworthy for the fact that it has quite a low pH (c. 4) but relatively low dissolved CO_2 (98 ppm) and essentially no bicarbonate, as well as being the only point where the Na/Cl ratio approaches 1. The reason for this point having the highest values for these trace elements is not clear, but it could be linked to the combined effects of pH, limited carbonate

complexing agents, Eh, and mineral phases present. pH appears to be particularly important, as the last two points along the transect have pH values that are approaching neutral.

Finally the trace elements plotted in Figure 74c show values that increase slightly from the background site to the first leakage sample point, increase sharply to the highest concentrations observed at the second leakage sample point, and then decrease smoothly (Cs, Cd) or irregularly (Cr, Li) in the down gradient part of the transect. This trend is very similar to that observed for Fe, SO_4, Al, Si, and K, and thus it is assumed that these four trace elements are linked with the mineral phases described above to explain the behaviour and spatial distribution of these major elements.

6.4.5. Spatial distribution of surface water chemistry

As described above in Section 6.3.1, field-based water-quality measurements were made at 50 m intervals along the creek that passes through the gas leakage area to ascertain the spatial impact of the leaking CO_2 on surface water quality and to look for indications of possible co-migration of deep-origin water with the gas. Because only in situ measurements were conducted these results can only be considered as a preliminary assessment of the potential impact.

The creek flow rate was small during the sampling period (see, for example, the photo in Figure 66a) and, although not measured, it is estimated to have been < 1 l/s during the very dry summer sampling period (i.e. mid-July). The observed results could thus represent an extreme, as higher surface water flow rates (for example, during the rainy winter season) would result in greater dilution.

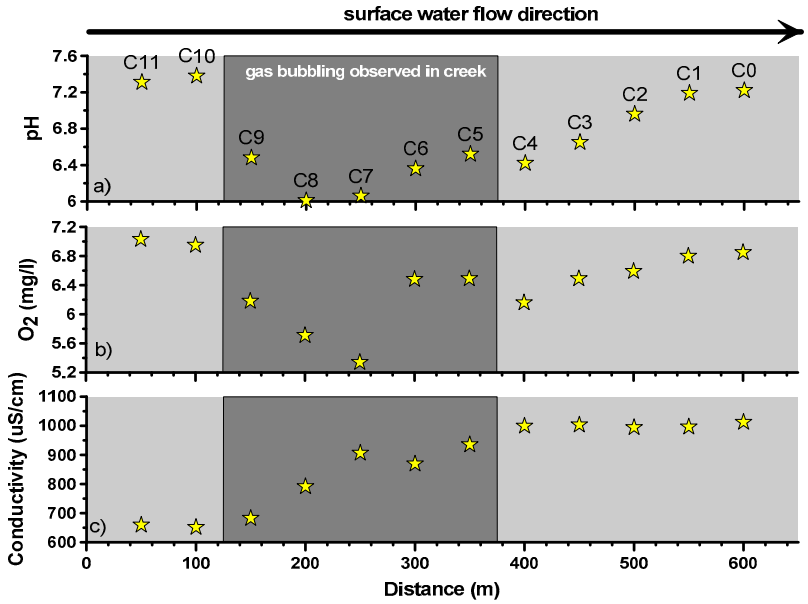


Figure 75. N-S transect along the creek where it crosses the gas leakage area (see Figure 64b for location). Points are spaced c. 50 m apart and water flow direction is from NE to SW.

The impact of the gas leakage area is clear on all three field-measured parameters presented in Figure 75; this figure shows surface water flowing from left (NE) to right (SW) through the interval where gas bubbling was directly observed within the creek water (marked in dark grey). Trends are different, however, in regards to how each parameter behaves beyond the leakage area. For example,

whereas pH and dissolved O_2 decrease in the leakage area but then recover relatively quickly to near background values down-stream, conductivity values rise in the leakage area and then remain stable and elevated even 200m beyond the southern boundary of the leakage area.

The pH trend can be explained by the dissolution and subsequent exsolution of CO_2 along the flow path. O_2 likely decreases due to the leakage of small amounts of reduced gases like H_2S and CH_4 together with the CO_2, which would consume oxygen, followed by subsequent re-equilibration with atmospheric O_2 down-stream. In contrast, the conductivity trend indicates that the increased dissolved ions acquired within the leakage area do not decrease rapidly despite the return to near-neutral pH values down-stream. This is likely due to the appropriate mineral phases always being under-saturated and thus not precipitating out of solution. This increase in conductivity could be due to either water-rock interaction in the stream bed caused by CO_2-induced acidity or to a small amount of co-migrating deeper water, although there is no evidence of the latter in the groundwater samples collected nearby. Understanding the actual source, and understanding if this impact on water quality is significant, would require subsequent chemical (and possibly) isotopic analyses.

6.5. Summary

Data at the Latera site were collected along a 600 m long transect that parallels groundwater flow and crosses a large and significant gas leakage area to observe changes in groundwater chemistry, looking for potential impacts on, and evolution of, the near-surface groundwater as it flows through the zone of high CO_2. Boreholes were hand augered to a depth of about 2.5 to 3.2 m in highly variable alluvial volcaniclastic sediments dominated by a aluminum-silicate mineralogy. Water was sampled for various physical-chemical parameters in the field (e.g., temperature, pH, conductivity) and for laboratory analysis of dissolved gases and major and trace elements. The water bearing horizon was encountered at a depth of about 2.8 to 3.0 m, with the piezometric surface typically occurring about 50 cm above this unit. The water bearing horizon in some locations appeared to be associated with a sediment - "bedrock" contact, while in others with a sandier fluvial unit. The hydraulic gradient along the transect was about 0.007 m/m.

The present study, like all field studies in which natural complexity can make interpretation challenging, has some limitations. These include: i) the sampled units are not "aquifers", as defined as a geological unit that can provide useable volumes of water, but rather are thin permeable horizons through which water flows relatively slowly; ii) this natural leaking site is long-lived and therefore is more likely to be at equilibrium, something that will likely not be the case should a leak occur and be discovered shortly afterwards; and iii) although a hydraulic gradient was observed along the transect, pumping tests were not conducted to determine if the various wells are hydraulically connected. Despite these issues, the obtained results are unique in their scope.

Data from the profile was compared with regional data from the literature, showing that all values plot within the field of the volcanic aquifers of the area. In addition, comparison of the collected samples with chemistry of deep brines associated with the geothermal reservoir at >2000 m shows no similarity, implying that no deep waters are migrating upwards with the leaking gas. This means that the observed results are likely due to only in situ reactions caused by the introduction of CO_2 (and associated H_2S).

Due to the predominately silicate mineralogy occurring at this site, carbonate minerals have little influence on the resulting impact. The buffering capacity of the carbonate system is exceeded and minimum pH values are on the order of 3.5. Whereas dissolved CO_2 reaches almost saturation values within the leakage area, HCO_3^- values decrease to almost 0 due to the extremely low pH (i.e. almost all DIC is in the form $H_2CO_3°$). Concentrations of Ca and Mg do not increase much moving from background to vent samples, however K, Al, and Si all increase significantly. All of these major elements decrease (and pH increases) rapidly down gradient, typically attaining values at <100 m that are less than or equal to the concentrations measured in the up gradient sample.

Examination of the saturation indices of numerous mineral phases shows how carbonates are always under saturated, and sulphates are always under saturated but with highest values in the venting area. The silicate minerals are at saturation up gradient, most under saturated within the venting area, and rise towards saturation levels in the down gradient samples. Finally clay minerals, oxides, and hydroxides are generally highly over saturated up gradient, saturated or slightly under saturated in the vent, and then increasingly over saturated moving down gradient. These results imply that the dominant reactions in the vent area involve dissolution of the main potassium aluminium silicates known to occur in the surface sediments (e.g. K-feldspar, leucite), with possible precipitation of clay minerals like kaolinite and smectite, oxides like hematite, and hydroxides like goethite and gibbsite controlling the decreasing concentrations down gradient (together with adsorption onto these newly formed mineral phases).

The overall trend of redox sensitive species is similar to that of, for example, silica, with sharp increases in SO_4, CH_4, and Fe in the vent area followed by rapidly decreasing values down gradient; H_2S, in contrast, is only high in one vent sample and essentially absent in all other samples. While SO_4 is likely formed by the oxidation of co-migrating H_2S, and CH_4 is also a co-migrating gas, elevated Fe concentrations are likely due to the dissolution of pyrite and hematite. The rapid decrease of these species is linked to the over saturation conditions of various mineral phases beyond the leak area.

Trace metal behaviour can be sub-divided into three groups, those that decrease along the flow path (U, Ba, As), those that increase in the leak area, have the highest concentration just outside the leak area, and then decrease to background levels further down gradient (Zn, Pb, Be, B), and those that have a general distribution that is very similar to that of Si and Al, with the highest values within the leakage area and rapidly decreasing values down gradient (Cd, Cr, Cs, Li). Again the individual behaviour of each trace element would need to be studied with detailed geochemical modelling and mineralogical analyses, however the similar behaviour within each of the three groups indicates that the principle mechanisms are probably controlled by a limited number of precipitation and adsorption processes. In any case, regardless of the mechanisms, there is a clear decrease of all measured trace elements to at or below background concentrations within 50 to 100 m of the gas leakage area. These values are typically below the drinking water limits.

Chapter 7. Conclusions

The present study has focussed on the phenomenon of CO_2 leakage from the deep subsurface towards the atmosphere. Work involved modelling surface flux measurement strategies and sampling densities (using synthetic and real data) to understand the influence that these parameters have on finding and quantifying leaks, as well as sampling groundwater along a transect through a major CO_2 leakage area to examine the potential impact of a leak on groundwater quality. This work relates to both natural leaking systems as well as to man-made systems where leakage may hypothetically occur (such as carbon capture and storage, CCS, projects). As a detailed summary of the results is given at the end of each research chapter (i.e., Chapters 4 – 6), to which the interested reader can refer, only high-level implications are discussed below.

A new model was developed and applied to determine what sampling strategies yield the greatest potential for finding a leak and how leak shape and orientation can influence success rate. Of the four sampling strategies tested, the offset grid approach was the most efficient at finding a circular anomaly, followed closely by the regular grid and uniform random approaches; instead the purely random approach required a much larger number of samples to attain the same probability of success. It was found that the diameter of the smallest circular anomaly guaranteed to be found at the 95% confidence interval using the offset grid approach is about 20% greater than the grid spacing itself, although the work conducted at the Latera field site showed that smaller anomalies can be located by chance (i.e., at a lower probability) or by selective additional sampling based on visual clues and other information. A power law was developed to estimate the number of samples needed to find a circular anomaly of a given size using the offset grid approach, providing a valuable tool for the planning of surveys and site monitoring. Tests examining the influence of anomaly shape and direction showed that while the random method always required the same, relatively high, number of samples for different anomalies of the same size, the responses of the other methods were influenced by anomaly orientation. The offset grid consistently gave the best results for most tests, however it gave the worst results (together with the standard grid approach) when the anomaly was aligned parallel to the offset direction. Based on these results the offset grid method is recommended for leak detection, however if elongate features are expected (e.g., along fault traces) it is best to align the offset direction so that it is not parallel to the anomaly anisotropy. Preliminary studies combining other sources of information, such as structural geology, remote sensing, and reconnaissance surface mapping, can greatly improve sample strategy and spacing decisions.

The same model was used to examine the effect of sample strategy and spacing on the estimate of total leakage flux. In a comparison of the offset grid and random sampling approaches, the former yields much better results at high sampling densities but only marginally better or equivalent results at lower densities. This boundary occurs where the grid spacing exceeds the average diameter of the anomalies present. This work also highlighted various potential sources of error during the leakage flux estimation process, as outlined below.

The estimation of the average background biological flux is critical, as this value must be subtracted from total flux measurements to determine the leakage flux. Several approaches can be taken, each having the potential to introduce errors. It was found that different sub-sampling of the same background population can yield a range of average values, although this was found to be small

(error of only a couple of percentage points). The use of a separate, distant area where leakage is known to not occur as a "background surrogate" has, instead, the potential to introduce a much larger error. If, for example, the surrogate area has different characteristics (such as water content, organic matter content, or vegetation type) or was sampled under different conditions than the leakage area (such as after a heavy rainfall or tillage), the two areas may have significantly different background flux rates. This was illustrated with the developed program by sub-sampling different quadrants of a single synthetic background population distributed over the entire grid. In this example the range of calculated difference was about 15%, however much more significant differences are quite feasible. Even when the statistical distribution of all samples collected only over the survey area are used to separate leakage from background populations, the spatial complexity of background distributions as well as the overlap between these two populations can also lead to potential errors. This was shown in the interpretation of the highly detailed grid performed at the Latera site, where a clear boundary in the log probability plot defined high leakage flux but complexity in the lower level distribution made it difficult to classify these samples as being purely biogenic or a mix of biogenic and geogenic flux (although the comparison with historical samples collected under different environmental conditions helped in the interpretation). In this example the use of different average background values, both based on the statistical distribution, yielded estimates of the leakage flux rate that differed by 100%.

A theoretical example can be used to highlight the need for an accurate background estimate. If the background flux is estimated to be 10 g m^{-2} d^{-1} when in reality it is 20 g m^{-2} d^{-1}, the difference will be incorrectly attributed to leakage flux. This would result in an error of 10 t d^{-1} for a 1 km^2 area and 100 t d^{-1} for a 10 km^2 area. While difficult to address, various possible approaches could be used to improve this estimate: subdivide large areas and treat them separately, as each may have different background populations; characterise the different soils over the study area in terms of their content (e.g. organic matter), water retention capacity, and their ability to respire CO_2 (i.e., incubation experiments); use vertical profiles of soil gas concentrations, isotopes, and seasonal / temporal monitoring to separate biogenic from geogenic gas fluxes.

Another potential source of error relates to the sampling density and how interpolation of the collected data into data gaps can extend leakage anomalies far beyond their actual physical extents, or merge anomalies that are in reality separate entities. While at higher sampling densities this is less of an issue, as sample spacing increases so does the potential for this effect to result in total leakage flux estimates that are much larger than the real flux, as seen in simulations using both synthetic and real data (for example, > 10x more in some outlier simulations). While this effect has the potential to introduce the largest error, it is actually the one that is simplest to address. Modelling examples have shown how limited infill sampling around found anomalies, such as 5 to 10 samples around an anomaly to delineate a leakage area or a profile between anomalies to separate them, can greatly improve leakage estimates for a small investment of time. This requires, however, flexibility in the sampling strategy, real-time data analysis during the sampling campaign, and allocation of additional time (perhaps a day or two) at the end of the campaign to allow this work to be done. The use of new technologies that rapidly measure CO_2 concentrations at the ground surface may also be used as a rapid reconnaissance tool prior to the flux survey, or as a spot verification tool after it is complete.

Although developed specifically to examine CO_2 flux, the GasGrid program now has the potential, with minor modifications, to be applied to other geological parameters that have a spatial distribution with localized anomalies.

Groundwater sampling along a transect parallel to the inferred flow direction and crossing a large CO_2 gas leak was used to better understand possible changes in water quality due to water-rock-gas reactions induced by CO_2 leakage. Within the leakage area itself pH values decreased to very low values due to the lack of significant carbonate buffering capacity in these silicate volcanic rocks. Within this interval very elevated values of Al, K, and Si were observed due to dissolution of mineral phases such as leucite and potassium feldspar, as well as redox sensitive compounds like CH_4 (co-leaking gas), SO_4 (due to oxidation of co-leaking H_2), Fe (due to dissolution of pyrite) and various trace elements likely occurring as trace components in the dissolving mineral phases. Down gradient over a distance of 50 – 100 m all anomalous parameters return to essentially up-gradient, non-impacted levels, likely due to secondary precipitation and adsorption processes. Additional work is needed, however, to verify these results given the uncertain hydraulics of the studied system.

Based on the present study, some observations can be made regarding the potential impact of CO_2 leakage on groundwater chemistry. First, redox conditions were seen to be as important as pH in controlling the evolution of the groundwater and the associated mobility of various major and trace elements. These are controlled not only by the Eh and pH buffering capacity of the aquifer itself (i.e., its mineralogy), but also the leakage rate, the leakage type (gas only versus gas plus deep water), and the composition of the leaking gas (CO_2 +/- H_2S, CH_4, etc.). Samples collected immediately down gradient, and further along the flow path, show how trace element concentrations decrease to background levels. This shows that the impact at this site is relatively limited, however other sites (with, for example, greater leakage rates, greater aquifer permeability and flow, higher concentrations in co-migrating water, etc.) could result in a wider impact area. Similar studies at different leakage sites are recommended to understand the potential range of impacts that may occur.

Some observations can also be made regarding groundwater monitoring at industrial sites (such as CCS) where there is concern about the potential for leakage from deep reservoirs and impact on overlying potable groundwater resources. As shown here, there will potentially be a significant difference in the types of parameters that should be monitored based on: i) whether the leak is gas only, water / brine only, or both phases together; ii) leakage rate; and iii) ph-Eh buffering capacity of the impacted aquifer. For example if brine is leaking one would expect high Cl concentrations; considering that this species is conservative and flows essentially at the groundwater flow rate, it could be a potential early warning parameter compared to more attenuated (and more harmful) elements like As or Pb. In addition, as Cl contributes to free ion concentrations, conductivity could be used as a simple, robust monitoring tool. Depending on background aquifer conditions, Eh measurements would also be useful. If instead only CO_2 gas is leaking, Cl and Eh may not change significantly, while increases in other major elements will depend on aquifer mineralogy and related dissolution kinetics; on the other hand trace metal concentrations may increase due to more rapid desorption processes. Considering that trace element measurements are prone to sample contamination and are expensive, in this case pH and dissolved CO_2 would be good parameters to be monitored, as would DIC; as pointed out by some authors conductivity may not be a good indicator in this case as DIC at lower pH values will be primarily in the neutral form $H_2CO_3°$, which does not contribute to conductivity. The advantage of using pH and pCO_2, as opposed to DIC, is that sensors can be installed for continuous in situ monitoring, and together with temperature these parameters can be used to calculate carbonate alkalinity. Implicit in this is the need for an accurate characterisation of the baseline conditions, both spatially to integrate aquifer heterogeneity and temporally to address potential seasonal variability, for the correct identification of a leakage signal.

Chapter 8. References

Agosta, F., and Kirschner, D.L., 2003, Fluid conduits in carbonate-hosted seismogenic normal faults of central Italy: Journal of Geophysical Research: Solid Earth, v. 108, p. 2221.

Altevogt, A.S., and Celia, M.A., 2004, Numerical modeling of carbon dioxide in unsaturated soils due to deep subsurface leakage: Water Resour. Res., v. 40, p. W03509.

Amundson, R., Stern, L., Baisden, T., and Wang, Y., 1998, The isotopic composition of soil and soil-respired CO_2: Geoderma, v. 82, p. 83-114.

Amundson, R.G., and Davidson, E.A., 1990, Carbon dioxide and nitrogenous gases in the soil atmosphere: J. Geochem. Explor. , v. 38, p. 13-41.

Annunziatellis, A., Beaubien, S.E., Bigi, S., Ciotoli, G., Coltella, M., and Lombardi, S., 2008a, Gas migration along fault systems and through the vadose zone in the Latera caldera (central Italy): Implications for CO_2 geological storage: Int. J. Greenhouse Gas Control, v. 2/3, p. 353-372.

Annunziatellis, A., Beaubien, S.E., Ciotoli, G., Coltella, M., and Lombardi, S., 2007, Total CO_2 flux from the Latera caldera and how flux rates affect the transfer of other reactive gas species to the atmosphere: the results of highly detailed surveys on and across individual gas vents., European Geosciences Union 2007; Geophysical Research Abstracts, Vol. 9, 07469, 2007; SRef-ID: 1607-7962/gra/EGU2007-A-07469: Vienna, Austria.

—, 2008b, Development of a rapid, low-cost technique for sensitive CO_2 leakage mapping, Vol. 10, EGU2008-A-10641, EGU General Assembly 2008: Vienna, Austria.

Annunziatellis, A., Beaubien, S.E., Ciotoli, G., and Lombardi, S., 2006, Studies into the migration of gas injected at shallow depths: implications for CO_2 geological sequestration, 8th International Conference on Greenhouse Gas Control Technologies (GHGT8): Trondheim, Norway.

Antonellini, M., and Aydin, A., 1995, Effect of faulting on fluid flow in porous sandstones; geometry and spatial distribution: AAPG Bulletin v. 79, p. 642-671.

Apps, J.A., Zheng, L., Zhang, Y., Xu, T., and Birkholzer, J.T., 2010, Evaluation of Potential Changes in Groundwater Quality in Response to CO_2 Leakage from Deep Geologic Storage: Transport in Porous Media, v. 82, p. 215-246.

Arevalo, C.B.M., Bhatti, J.S., Chang, S.X., Jassal, R.S., and Sidders, D., 2010, Soil respiration in four different land use systems in north central Alberta, Canada: J. Geophys. Res., v. 115, p. G01003.

ASTER GDEM, 2014, The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) is a product of the U.S. National Aeronautics and Space Administration (NASA) and Japan's Ministry of Economy, Trade, and Industry (METI). Accessed at http://gdex.cr.usgs.gov/gdex/ on 04/02/2014.

Bachu, S., and Bennion, D.B., 2009, Experimental assessment of brine and/or CO_2 leakage through well cements at reservoir conditions: Int. J. Greenhouse Gas Control, v. 3, p. 494-501.

Baldi, P., Ceccarelli, A., Bertrami, R., Friedman, I., and Lombardi, S., 1984, Helium in soil gases in geothermal areas, Seminar on Utilization of Geothermal Energy for Electric Power Production and Space Heating: Florence, Italy, United Nations, Economic Commission for Europe, p. 1-19.

Barberi, F., Innocenti, F., Landi, P., Rossi, U., Saitta, M., Santacroce, R., and Villa, I.M., 1984, The evolution of Latera Caldera (central Italy) in the light of subsurface data: Bull. Volcanol., v. 47, p. 125–141.

Barron-Gafford, G.A., Scott, R.L., Jenerette, G.D., and Huxman, T.E., 2011, The relative controls of temperature, soil moisture, and plant functional group on soil CO_2 efflux at diel, seasonal, and annual scales: J. Geophys. Res., v. 116, p. G01023.

Bateson, L., Vellico, M., Beaubien, S.E., Pearce, J.M., Ciotoli, G., Annunziatellis, A., Coren, F., Lombardi, S., and Marsh, S., 2008, Preliminary results of the application of remote sensing techniques to detecting and monitoring leaks from CO_2 storage sites: Int. J. Greenhouse Gas Control, v. 2/3, p. 388-400.

Beaubien, S.E., Bigi, S., Lombardi, S., Sacco, P., and Tartarello, M.C., 2014a, Groundwater changes caused by flow through naturally occurring gas (±water) leakage points, Fourth EAGE CO_2 Geological Storage Workshop: Stavanger, Norway.

Beaubien, S.E., Ciotoli, G., Coombs, P., Dictor, M.C., Krüger, M., Lombardi, S., Pearce, J.M., and West, J.M., 2008a, The impact of a naturally-occurring CO_2 gas vent on the shallow ecosystem and soil chemistry of a Mediterranean pasture (Latera, Italy): Int. J. Greenhouse Gas Control, v. 2/3, p. 373-387.

Beaubien, S.E., Ciotoli, G., and Lombardi, S., 2003, Carbon dioxide and radon gas hazard in the Alban Hills area (central Italy): Journal of Volcanology and Geothermal Research, v. 123, p. 63-80.

Beaubien, S.E., Jones, D.G., Gal, F., Barkwith, A.K.A.P., Braibant, G., Baubron, J.C., Ciotoli, G., Graziani, S., Lister, T.R., Lombardi, S., Michel, K., Quattrocchi, F., and Strutt, M.H., 2013, Monitoring of near-surface gas geochemistry at the Weyburn, Canada, CO_2-EOR site, 2001-2011: International Journal of Greenhouse Gas Control, v. 16, Supplement 1, p. S236-S262.

Beaubien, S.E., Jones, D.G., Le Pierres, K., van Bergen, F., Schulz, H.M., S., L., and the JRAP5 team, 2008b, Deliverable JRAP-5/5: Final report – data integration and interpretation, Internal CO_2GeoNet Report (EC funded project SES6-CT-2004-502816).

Beaubien, S.E., Ruggiero, L., Annunziatellis, A., Bigi, S., Ciotoli, G., Dieiana, P., Graziani, S., Lombardi, S., and Tartarello, M.C., 2014b, The importance of baseline surveys of near-surface gas geochemistry for CCS monitoring, as shown from onshore case studies in northern and southern Europe: Oil and Gas Science Technology Journal.

Beaubien, S.E., Strutt, M.H., Jones, D.G., Baubron, J.-C., Cardellini, C., Lombardi, S., Quattrochi, F., Penner, I., 2004. D20 Report: Soil Gas Surveys in the Weyburn Oil Field (2001–2003). British Geological Survey, Keyworth, Nottingham; Commissioned Report CR/04/030.

Bekele, A., Kellman, L., and Beltrami, H., 2007, Soil Profile CO_2 concentrations in forested and clear cut sites in Nova Scotia, Canada: Forest Ecology and Management, v. 242, p. 587-597.

Bense, V.F., Gleeson, T., Loveless, S.E., Bour, O., and Scibek, J., 2013, Fault zone hydrogeology: Earth-Science Reviews, v. 127, p. 171-192.

Bergfeld, D., Evans, W.C., Howle, J.F., and Farrar, C.D., 2006, Carbon dioxide emissions from vegetation-kill zones around the resurgent dome of Long Valley caldera, eastern California, USA: Journal of Volcanology and Geothermal Research, v. 152, p. 140-156.

Bergfeld, D., Evans, W.C., Lowenstern, J.B., and Hurwitz, S., 2012, Carbon dioxide and hydrogen sulfide degassing and cryptic thermal input to Brimstone Basin, Yellowstone National Park, Wyoming: Chemical Geology, v. 330-331, p. 233-243.

Bergfeld, D., Goff, F., and Janik, C.J., 2001a, Carbon isotope systematics and CO_2 sources in The Geysers-Clear Lake region, northern California, USA: Geothermics, v. 30, p. 303-331.

—, 2001b, Elevated carbon dioxide flux at the Dixie Valley geothermal field, Nevada; relations between surface phenomena and the geothermal reservoir: Chemical Geology, v. 177, p. 43-66.

Bertrami, R., Cameli, G.H., Lovari, F., and Rossi, U., 1984, Discovery of Latera Geothermal Field: Problems of the Exploration and Research, Seminar on Utilization of Geothermal Energy for Electric Power Production and Space Heating: United Nations, Economic Commission for Europe. EP/SEM.9/R.20.: Florence.

Bickle, M., Chadwick, A., Huppert, H.E., Hallworth, M., and Lyle, S., 2007, Modelling carbon dioxide accumulation at Sleipner: Implications for underground carbon storage: Earth and Planetary Science Letters, v. 255, p. 164-176.

Bigi, S., Battaglia, M., Alemanni, A., Lombardi, S., Campana, A., Borisova, E., and Loizzo, M., 2013, CO_2 flow through a fractured rock volume: Insights from field data, 3D fractures representation and fluid flow modeling: International Journal of Greenhouse Gas Control, v. 18, p. 183-199.

Birkholzer, J.T., Apps, J.A., Zheng, L., Zhang, Y., Xu, T., and Tsang, C.-F., 2008, Research Project on CO_2 Geological Storage and Groundwater Resources: Water Quality Effects Caused by CO_2 Intrusion into Shallow Groundwater. Technical Report LBNL-1251E, Lawrence Berkeley National Laboratory, Berkeley, CA.

Birkholzer, J.T., Nicot, J.P., Oldenburg, C.M., Zhou, Q., Kraemer, S., and Bandilla, K., 2011, Brine flow up a well caused by pressure perturbation from geologic carbon sequestration: Static and dynamic evaluations: International Journal of Greenhouse Gas Control, v. 5, p. 850-861.

Borken, W., Davidson, E.A., Savage, K., Gaudinski, J., and Trumbore, S.E., 2003, Drying And Wetting Effects On Carbon Dioxide Release From Organic Horizons: Soil Sci. Soc. Am. J., v. 67, p. 1888-1896.

Bouma, T.J., and Bryla, D.R., 2000, On the assessment of root respiration for soils of different textures: interactions with soil moisture contents and soil CO_2 concentrations.: Plant and Soil, v. 227, p. 215-221.

Bowling, D.R., Grote, E.E., and Belnap, J., 2011, Rain pulse response of soil CO_2 exchange by biological soil crusts and grasslands of the semiarid Colorado Plateau, United States: J. Geophys. Res., v. 116, p. G03028.

Bricker, S.H., Barkwith, A., MacDonald, A.M., Hughes, A.G., and Smith, M., 2012, Effects of CO_2 injection on shallow groundwater resources: A hypothetical case study in the Sherwood Sandstone aquifer, UK: International Journal of Greenhouse Gas Control, v. 11, p. 337-348.

Brocca, L., Morbidelli, R., Melone, F., and Moramarco, T., 2007, Soil moisture spatial variability in experimental areas of central Italy: Journal of Hydrology, v. 333, p. 356-373.

Bruggemann, N., Gessler, A., Kayler, Z., Keel, S.G., Badeck, F., Barthel, M., Boeckx, P., Buchmann, N., Brugnoli, E., Esperschutz, J., Gavrichkova, O., Ghashghaie, J., Gomez-Casanovas, N., Keitel, C., Knohl, A., Kuptz, D., Palacio, S., Salmon, Y., Uchida, Y., and Bahn, M., 2011, Carbon allocation and carbon isotope fluxes in the plant-soil-atmosphere continuum: a review: Biogeosciences Discuss., v. 8, p. 3619-3695.

Buonasorte, G., Ciccacci, S., De Rita, D., Fredi, P., and Lupia Palmieri, E., 1991, Some relations between morphological characteristics and geological structure in the Vulsini Volcanic Complex (Northern Latium, Italy): Zeitschrift fuer Geomorphologie Supplement BD, v. 82, p. 59–71.

Cahill, A.G., and Jakobsen, R., 2013, Hydro-geochemical impact of CO_2 leakage from geological storage on shallow potable aquifers: A field scale pilot experiment: International Journal of Greenhouse Gas Control, v. in press.

Caine, J.S., Evans, J.P., and Forster, C.B., 1996, Fault zone architecture and permeability structure: Geology, v. 24, p. 1025-1028.

Camarda, M., De Gregorio, S., Favara, R., and Gurrieri, S., 2007, Evaluation of carbon isotope fractionation of soil CO_2 under an advective-diffusive regimen: A tool for computing the isotopic composition of unfractionated deep source: Geochimica et Cosmochimica Acta, v. 71, p. 3016-3027.

Capasso, G., and Inguaggiato, S., 1998, A simple method for the determination of dissolved gases in natural waters. An application to thermal waters from Vulcano Island: Applied Geochemistry, v. 13, p. 631-642.

Carbonell-Bojollo, R., Gonzalez-Sanchez, E.J., Veroz-Gonzalez, O., and Ordonez-Fernandez, R., 2011, Soil management systems and short term CO_2 emissions in a clayey soil in southern Spain: Science of The Total Environment, v. 409, p. 2929-2935.

Cardellini, C., Chiodini, G., and Frondini, F., 2003, Application of stochastic simulation to CO_2 flux from soil: Mapping and quantification of gas release: Journal of Geophysical Research: Solid Earth, v. 108, p. 2425.

Carey, J.W., Svec, R., Grigg, R., Zhang, J., and Crow, W., 2010, Experimental investigation of wellbore integrity and CO_2-brine flow along the casing-cement microannulus: International Journal of Greenhouse Gas Control, v. 4, p. 272-282.

Carey, J.W., Wigand, M., Chipera, S.J., WoldeGabriel, G., Pawar, R., Lichtner, P.C., Wehner, S.C., Raines, M.A., and Guthrie, J.G.D., 2007, Analysis and performance of oil well cement with 30 years of CO_2 exposure from the SACROC Unit, West Texas, USA: International Journal of Greenhouse Gas Control, v. 1, p. 75-85.

Caritat, P.d., Hortle, A., Raistrick, M., Stalvies, C., and Jenkins, C., 2013, Monitoring groundwater flow and chemical and isotopic composition at a demonstration site for carbon dioxide storage in a depleted natural gas reservoir: Applied Geochemistry, v. 30, p. 16-32.

Carmignani, L., and Kligfield, R., 1990, Crustal extension in the Northern Apennines: the transition from compression to extension in the Alpi Apuane core complex.: Tectonics, v. 9, p. 340-349.

Carroll, S.A., and Knauss, K.G., 2005, Dependence of labradorite dissolution kinetics on CO_2(aq), Al(aq), and temperature: Chemical Geology, v. 217, p. 213-225.

Cavarretta, G., Giannelli, G., Scandiffio, G., and Tecce, F., 1985, Evolution of the Latera Geothermal System II: Metamorphic, Hydrothermal Mineral Assemblages and Fluid Chemistry: Journal of Volcanology and Geochemical Research, v. 26, p. 337-364.

Cerling, T.E., Solomon, D.K., Quade, J., and Bowman, J.R., 1991, On the isotopic composition of carbon in soil carbon dioxide: Geochimica et Cosmochimica Acta, v. 55, p. 3403-3405.

Checcucci, R., Chiodini, G., and Giaquinto, S., 1988, Idrogeochimica dei M.ti Vulsini: sintesi e rielaborazione dei dati esistenti. p. 269-290.

Chen, C., and Zhang, D., 2010, Pore-scale simulation of density-driven convection in fractured porous media during geological CO_2 sequestration: Water Resources Research, v. 46, p. W11527.

Chiodini, G., Baldini, A., Barberi, F., Carapezza, M.L., Cardellini, C., Frondini, F., Granieri, D., and Ranaldi, M., 2007, Carbon dioxide degassing at Latera caldera (Italy): Evidence of geothermal reservoir and evaluation of its potential energy: Journal of Geophysical Research, v. 112, p. doi:10.1029/2006JB004896.

Chiodini, G., Caliro, S., Cardellini, C., Avino, R., Granieri, D., and Schmidt, A., 2008, Carbon isotopic composition of soil CO_2 efflux, a powerful method to discriminate different sources feeding soil CO_2 degassing in volcanic-hydrothermal areas: Earth and Planetary Science Letters, v. 274, p. 372-379.

Chiodini, G., Cardellini, C., Amato, A., Boschi, E., Carliro, S., and Frondini, F., 2004, Carbon dioxide Earth degassing and seismogenesis in central and southern Italy: Geophysical Research Letters, v. L07615, p. doi:10.1029/2004GL019480.

Chiodini, G., Cioni, R., Guidi, M., Marini, L., and Raco, B., 1998, Soil CO_2 flux measurements in volcanic and geothermal areas.: Applied Geochemistry, v. 13, p. 543-552.

Chiodini, G., Frondini, F., Cardellini, C., Parello, F., and Peruzzi, L., 2000, Rate of diffuse carbon dioxide Earth degassing estimated from carbon balance of regional aquifers: The case of central Apennine, Italy: Journal of Geophysical Research: Solid Earth, v. 105, p. 8423-8434.

Chiodini, G., Frondini, F., and Ponziani, F., 1995, Deep structures and carbon dioxide degassing in central Italy: Geothermics, v. 24, p. 81-94.

Choi, B.-Y., Yun, S.-T., Kim, K.-H., Choi, H.-S., Chae, G.-T., and Lee, P.-K., 2014, Geochemical modeling of CO_2-water-rock interactions for two different hydrochemical types of CO_2-rich springs in Kangwon District, Korea: Journal of Geochemical Exploration, v. 144, Part A, p. 49-62.

Choi, Y.-S., Young, D., Nesic, S., and Gray, L.G.S., 2013, Wellbore integrity and corrosion of carbon steel in CO_2 geologic storage environments: A literature review: International Journal of Greenhouse Gas Control, v. 16, Supplement 1, p. S70-S77.

Cihan, A., Birkholzer, J.T., and Zhou, Q., 2012, Pressure Buildup and Brine Migration During CO_2 Storage in Multilayered Aquifers: Groundwater, v. 51, p. 252-267.

Ciotoli, G., Etiope, G., Guerra, M., Lombardi, S., Duddridge, G., and Grainger, P., 2005, Migration of gas injected into a fault in low-permeability ground: Quarterly Journal of Engineering Geology and Hydrogeology, v. 38, p. 305-320.

Cortis, A., Oldenburg, C.M., and Benson, S.M., 2008, The role of optimality in characterizing CO_2 seepage from geologic carbon sequestration sites: International Journal of Greenhouse Gas Control, v. 2, p. 640-652.

Crow, W., Carey, J.W., Gasda, S., Brian Williams, D., and Celia, M., 2010, Wellbore integrity analysis of a natural CO_2 producer: International Journal of Greenhouse Gas Control, v. 4, p. 186-197.

Cuisiat, F., and Skurtveit, E., 2010, An experimental investigation of the development and permeability of clay smears along faults in uncemented sediments: Journal of Structural Geology, v. 32, p. 1850-1863.

D'Amore, F., Panichi, C., Squarci, P., Bertrami, R., and Ceccarelli, A., 1979, Studio idrogeologico e geochimico dei sistemi termali della zona Latera-Canino.

Darby, E.B., Bumgarner, J., and Hovorka, S.D., 2009, Geochemical modeling of near-surface CO_2 interactions: The critical element in cost-effective long-term monitoring: Energy Procedia, v. 1, p. 2389-2395.

Davidson, G.R., 1995, The stable isotopic composition and measurement of carbon in soil CO_2: Geochimica et Cosmochimica Acta, v. 59, p. 2485-2489.

DeSutter, T.M., Sauer, T.J., Parkin, T.B., and Heitman, J.L., 2008, A Subsurface, Closed-loop System For Soil Carbon Dioxide And Its Application To The Gradient Efflux Approach: Soil Sci. Soc. Am. J., v. 72, p. 126-134.

Deutsch, C.V., and Journel, A.G., 1998, GSLIB: Geostatistical Software Library and Users Guide: New York, USA, Oxford Univ. Press.

Devoti, R., Esposito, A., Pietrantonio, G., Pisani, A.R., and Riguzzi, F., 2011, Evidence of large scale deformation patterns from GPS data in the Italian subduction boundary: Earth and Planetary Science Letters, v. 311, p. 230-241.

Di Filippo, M., Lombardi, S., Nappi, G., Reimer, G.M., Renzulli, A., and Toro, B., 1999, Volcano-tectonic structures, gravity and helium in geothermal areas of Tuscany and Latium (Vulsini volcanic district), Italy: Geothermics, v. 28, p. 377-393.

Di Stefano, R., Kissling, E., Chiarabba, C., Amato, A., and Giardini, D., 2009, Shallow subduction beneath Italy: Three-dimensional images of the Adriatic-European-Tyrrhenian lithosphere system based on high-quality P wave arrival times: Journal of Geophysical Research: Solid Earth, v. 114, p. B05305.

Dockrill, B., and Shipton, Z.K., 2010, Structural controls on leakage from a natural CO_2 geologic storage site: Central Utah, U.S.A: Journal of Structural Geology, v. 32, p. 1768-1782.

Donders, T.H., Decuyper, M., Beaubien, S.E., van Hoof, T.B., Cherubini, P., and Sass-Klaassen, U., 2013, Tree rings as biosensor to detect leakage of subsurface fossil CO_2: International Journal of Greenhouse Gas Control, v. 19, p. 387-395.

Drewitt, G.B., Black, T.A., and Jassal, R.S., 2005, Using measurements of soil CO_2 efflux and concentrations to infer the depth distribution of CO_2 production in a forest soil: Canadian Journal of Soil Science, v. 85, p. 213-221.

Duchi, V., Minissale, A., Ortino, S., and Romani, L., 1987, Geothermal prospecting by geochemical methods on natural gas and water discharges in the Vulsini Mts Volcanic District (central Italy): Geothermics, v. 16, p. 147-157.

Duchi, V., Minissale, A., Paolieri, M., Prati, F., and Valori, A., 1992, Chemical relationship between discharging fluids in the Siena-Radiocfani Graben and the deep fluids produced by the geothermal fields of Mt Amiata, Torre Alfina and Latera: Geothermics, v. 21, p. 401-413.

Duguid, A., Radonjic, M., and Scherer, G., 2005, Degradation of Well Cements Exposed to Carbonated Brine, Fourth Annual Conference On Carbon Capture And Sequestration DOE/NETL: Alexandria Virginia, USA.

Elberling, B., 2003, Seasonal trends of soil CO_2 dynamics in a soil subject to freezing: Journal of Hydrology, v. 276, p. 159-175.

Emberley, S., Hutcheon, I., Shevalier, M., Durocher, K., Mayer, B., Gunter, W.D., and Perkins, E.H., 2005, Monitoring of fluid-rock interaction and CO_2 storage through produced fluid sampling at the Weyburn CO_2-injection enhanced oil recovery site, Saskatchewan, Canada: Applied Geochemistry, v. 20, p. 1131-1157.

Emmerich, W.E., 2003, Carbon dioxide fluxes in a semiarid environment with high carbonate soils: Agricultural and Forest Meteorology, v. 116, p. 91-102.

Evans, W.C., Bergfeld, D., McGimsey, R.G., and Hunt, A.G., 2009, Diffuse gas emissions at the Ukinrek Maars, Alaska: Implications for magmatic degassing and volcanic monitoring: Applied Geochemistry, v. 24, p. 527-535.

Evemden, J.F., and Curtis, G.H., 1965, K/Ar of late Cenozoic rocks in Eastern Africa and Italy: Current Anthropology, v. 6, p. 343–385.

Fassbinder, J.J., Griffis, T.J., and Baker, J.M., 2012, Interannual, seasonal, and diel variability in the carbon isotope composition of respiration in a C3/C4 agricultural ecosystem: Agricultural and Forest Meteorology, v. 153, p. 144-153.

Faulkner, D.R., Jackson, C.A.L., Lunn, R.J., Schlische, R.W., Shipton, Z.K., Wibberley, C.A.J., and Withjack, M.O., 2010, A review of recent developments concerning the structure, mechanics and fluid flow properties of fault zones: Journal of Structural Geology, v. 32, p. 1557-1575.

Fierer, N., and Schimel, J.P., 2003, A Proposed Mechanism For The Pulse In Carbon Dioxide Production Commonly Observed Following The Rapid Rewetting Of A Dry Soil: Soil Sci. Soc. Am. J., v. 67, p. 798-805.

Flanagan, L.B., and Johnson, B.G., 2005, Interacting effects of temperature, soil moisture and plant biomass production on ecosystem respiration in a northern temperate grassland: Agricultural and Forest Meteorology, v. 130, p. 237-253.

Flechard, C.R., Neftel, A., Jocher, M., Ammann, C., Leifeld, J., and Fuhrer, J., 2007, Temporal changes in soil pore space CO_2 concentration and storage under permanent grassland: Agricultural and Forest Meteorology, v. 142, p. 66-84.

Frank, A.B., Liebig, M.A., and Hanson, J.D., 2002, Soil carbon dioxide fluxes in northern semiarid grasslands: Soil Biology and Biochemistry, v. 34, p. 1235-1241.

Frerichs, J., Oppermann, B.I., Gwosdz, S., Möller, I., Herrmann, M., and Krüger, M., 2013, Microbial community changes at a terrestrial volcanic CO_2 vent induced by soil acidification and anaerobic microhabitats within the soil column: FEMS Microbiology Ecology, v. 84, p. 60-74.

Frezzotti, M.L., Peccerillo, A., and Panza, G., 2009, Carbonate metasomatism and CO_2 lithosphere-asthenosphere degassing beneath the Western Mediterranean: An integrated model arising from petrological and geophysical data: Chemical Geology, v. 262, p. 108-120.

Fridriksson, T., Kristjansson, B.R., Armannsson, H., Margretardottir, E., Olafsdottir, S., and Chiodini, G., 2006, CO_2 emissions and heat flow through soil, fumaroles, and steam heated mud pools at the Reykjanes geothermal area, SW Iceland: Applied Geochemistry, v. 21, p. 1551-1569.

Frondini, F., Caliro, S., Cardellini, C., Chiodini, G., Morgantini, N., and Parello, F., 2008, Carbon dioxide degassing from Tuscany and Northern Latium (Italy): Global and Planetary Change, v. 61, p. 89-102.

Funiciello, R., and Lombardi, S., 1984, Composizione isotopica del carbonio e dell'ossigeno nelle mineralizzazioni a calcite e nei carbonati di alcuni pozzi geotermici italiani, , University of Rome "La Sapienza", Roma.

Gal, F., Joublin, F., Haas, H., Jean-prost, V., and Ruffier, V., 2011, Soil gas (^{222}Rn, CO_2, ^{4}He) behaviour over a natural CO_2 accumulation, Montmiral area (Drome, France): geographical, geological and temporal relationships: Journal of Environmental Radioactivity, v. 102, p. 107-118.

Gambardella, B., Cardellini, C., Chiodini, G., Frondini, F., Marini, L., Ottonello, G., and Vetushi Zuccolini, M., 2004, Fluxes of deep CO_2 in the volcanic areas of central-southern Italy: Journal of Volcanology and Geothermal Research, v. 136, p. 31-52.

Gartrell, A., Zhang, Y., Lisk, M., and Dewhurst, D., 2004, Fault intersections as critical hydrocarbon leakage zones: integrated field study and numerical modelling of an example from the Timor Sea, Australia: Marine and Petroleum Geology, v. 21, p. 1165-1179.

Gasda, S., Bachu, S., and Celia, M., 2004, Spatial characterization of the location of potentially leaky wells penetrating a deep saline aquifer in a mature sedimentary basin: Environmental Geology, v. 46, p. 707-720.

Gaudinski, J.B., Trumbore, S.E., Davidson, E.A., and Zheng, S., 2000, Soil carbon cycling in a temperate forest: radiocarbon-based estimates of residence times, sequestration rates and partitioning of fluxes: Biogeochemistry, v. 51, p. 33-69.

Geraud, Y., Diraison, M., and Orellana, N., 2006, Fault zone geometry of a mature active normal fault: A potential high permeability channel (Pirgaki fault, Corinth rift, Greece): Tectonophysics, v. 426, p. 61-76.

Giacomuzzi, G., Chiarabba, C., and De Gori, P., 2011, Linking the Alps and Apennines subduction systems: New constraints revealed by high-resolution teleseismic tomography: Earth and Planetary Science Letters, v. 301, p. 531-543.

Gianelli, G., and Scandiffio, G., 1989, The Latera geothermal system (Italy): chemical composition of the geothermal fluid and hypotheses on its origin: Geothermics v. 18, p. 447-463.

Giustini, F., Blessing, M., Brilli, M., Lombardi, S., Voltattorni, N., and Widory, D., 2013, Determining the origin of carbon dioxide and methane in the gaseous emissions of the San Vittorino plain (Central Italy) by means of stable isotopes and noble gas analysis: Applied Geochemistry, v. 34, p. 90-101.

Gorczyca, Z., Rozanski, K., Kuc, T., and Michalec, B., 2003, Seasonal variability of the soil CO_2 flux and its isotopic composition in southern Poland: Nukleonika, v. 48, p. 187-196.

Gray, M.B., Stamatakos, J.A., Ferrill, D.A., and Evans, M.A., 2005, Fault-zone deformation in welded tuffs at Yucca Mountain, Nevada, USA: Journal of Structural Geology, v. 27, p. 1873-1891.

Gudmundsson, A., Berg, S.S., Lyslo, K.B., and Skurtveit, E., 2001, Fracture networks and fluid transport in active fault zones: Journal of Structural Geology, v. 23, p. 343-353.

Hamada, Y., and Tanaka, T., 2001, Dynamics of carbon dioxide in soil profiles based on long-term field observation: Hydrological Processes, v. 15, p. 1829-1845.

Harvey, O.R., Qafoku, N.P., Cantrell, K.J., Lee, G., Amonette, J.E., and Brown, C.F., 2012, Geochemical Implications of Gas Leakage associated with Geologic CO_2 Storage: A Qualitative Review: Environmental Science & Technology, v. 47, p. 23-36.

Hashimoto, S., and Komatsu, H., 2006, Relationships between soil CO_2 concentration and CO_2 production, temperature, water content, and gas diffusivity: implications for field studies through sensitivity analyses: Journal of Forest Research, v. 11, p. 41-50.

Hernandez, P.A., Padilla, G.n., Padron, E., Perez, N.M., Calvo, D., Nolasco, D., Melian, G., Barrancos, J., Dionis, S., Rodriguez, F., and Sumino, H., 2012, Analysis of long- and short-term temporal variations of the diffuse CO_2 emission from Timanfaya volcano, Lanzarote, Canary Islands: Applied Geochemistry, v. 27, p. 2486-2499.

Hesse, M.A., and Woods, A.W., 2010, Buoyant dispersal of CO_2 during geological storage: Geophysical Research Letters, v. 37, p. L01403.

Hirano, T., Kim, H., and Tanaka, Y., 2003, Long-term half-hourly measurement of soil CO_2 concentration and soil respiration in a temperate deciduous forest: J. Geophys. Res., v. 108, p. 4631.

Hoefs, J., 2009, Stable Isotope Geochemistry, Springer, 286 p.

Holm, P.M., 1982, Mineral Chemistry of Perpotassic Lavas of the Vulsinian District, Roman Province, Italy: Miner. Mag. , v. 46.

Hood, G.M., 2010, PopTools version 3.2.5. Available on the internet. URL http://www.poptools.org.

Hooker, P.J., Bertrami, R., Lombardi, S., O'Nions, R.K., and Oxburgh, E.R., 1985, Helium-3 anomalies and crust-mantle interaction in Italy: Geochimica et Cosmochimica Acta, v. 49, p. 2505-2513.

Humez, P., Lagneau, V., Lions, J., and Negrel, P., 2013, Assessing the potential consequences of CO_2 leakage to freshwater resources: A batch-reaction experiment towards an isotopic tracing tool: Applied Geochemistry, v. 30, p. 178-190.

Humez, P., Negrel, P., Lagneau, V., Lions, J., Kloppmann, W., Gal, F., Millot, R., Guerrot, C., Flehoc, C., Widory, D., and Girard, J.-F., 2014, CO_2-water-mineral reactions during CO_2 leakage: Geochemical and isotopic monitoring of a CO_2 injection field test: Chemical Geology, v. 368, p. 11-30.

Hutchinson, G.L., and Livingston, G.P., 1993, Use of chamber systems to measure trace gas fluxes, Agricultural ecosystem effects on trace gases and global climate change, Volume ASA Special Publication, n.55.

—, 2001, Vents and seals in non-steady-state chambers used for measuring gas exchange between soil and the atmosphere: Eur.J. Soil Sci. , v. 52, p. 675-682.

Iacono Marziano, G., Gaillard, F., and Pichavant, M., 2007, Limestone assimilation and the origin of CO_2 emissions at the Alban Hills (Central Italy): Constraints from experimental petrology: Journal of Volcanology and Geothermal Research, v. 166, p. 91-105.

IEAGHG, 2011, Potential Impacts on Groundwater Resources of CO_2 Storage, 2011/11, October, 2011.

IPCC, 2005, IPCC Special Report on Carbon Dioxide Capture and Storage: Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press, 442 p.

Jassal, R., Black, A., Novak, M., Morgenstern, K., Nesic, Z., and Gaumont-Guay, D., 2005, Relationship between soil CO_2 concentrations and forest-floor CO_2 effluxes: Agricultural and Forest Meteorology, v. 130, p. 176-192.

Jin, L., Ogrinc, N., Hamilton, S.K., Szramek, K., Kanduc, T., and Walter, L.M., 2009, Inorganic carbon isotope systematics in soil profiles undergoing silicate and carbonate weathering (Southern Michigan, USA): Chemical Geology, v. 264, p. 139-153.

Jones, D.G., Barkwith, A.K.A.P., Hannis, S., Lister, T.R., Gal, F., Graziani, S., Beaubien, S.E., and Widory, D., 2014, Monitoring of near surface gas seepage from a shallow injection experiment at the CO_2 Field Lab, Norway: International Journal of Greenhouse Gas Control, v. 28, p. 300-317.

Jones, D.G., Barlow, T., Beaubien, S.E., Ciotoli, G., Lister, T.R., Lombardi, S., May, F., Möller, I., Pearce, J.M., and Shaw, R.A., 2009, New and established techniques for surface gas monitoring at onshore CO_2 storage sites: Energy Procedia, v. 1, p. 2127-2134.

Jones, D.G., Lister, T.R., Smith, D.J., West, J.M., Coombs, P., Gadalia, A., Brach, M., Annunziatellis, A., and Lombardi, S., 2011, In Salah gas CO_2 storage JIP: Surface gas and biological monitoring: Energy Procedia, v. 4, p. 3566-3573.

Kampman, N., Bickle, M.J., Maskell, A., Chapman, H.J., Evans, J.P., Purser, G., Zhou, Z., Schaller, M.F., Gattacceca, J.C., Bertier, P., Chen, F., Turchyn, A.V., Assayag, N., Rochelle, C., Ballentine, C.J., and Busch,

A., 2013, Drilling and sampling a natural CO_2 reservoir: Implications for fluid flow and CO_2-fluidâ€“rock reactions during CO_2 migration through the overburden: Chemical Geology, v. 369, p. 51-82.

Karamalidis, A.K., Torres, S.G., Hakala, J.A., Shao, H., Cantrell, K.J., and Carroll, S., 2013, Trace Metal Source Terms in Carbon Sequestration Environments: Environmental Science & Technology, v. 47, p. 322-329.

Keating, E., Newell, D., Dempsey, D., and Pawar, R., 2014, Insights into interconnections between the shallow and deep systems from a natural CO_2 reservoir near Springerville, Arizona: International Journal of Greenhouse Gas Control, v. 25, p. 162-172.

Keating, E.H., Hakala, J.A., Viswanathan, H., Carey, J.W., Pawar, R., Guthrie, G.D., and Fessenden-Rahn, J., 2013a, CO_2 leakage impacts on shallow groundwater: Field-scale reactive-transport simulations informed by observations at a natural analog site: Applied Geochemistry, v. 30, p. 136-147.

Keating, E.H., Newell, D.L., Viswanathan, H., Carey, J.W., Zyvoloski, G., and Pawar, R., 2013b, CO_2/Brine Transport into Shallow Aquifers along Fault Zones: Environmental Science & Technology, v. 47, p. 290-297.

Kharaka, Y., Thordsen, J., Kakouros, E., Ambats, G., Herkelrath, W., Beers, S., Birkholzer, J., Apps, J., Spycher, N., Zheng, L., Trautz, R., Rauch, H., and Gullickson, K., 2010, Changes in the chemistry of shallow groundwater related to the 2008 injection of CO_2 at the ZERT field site, Bozeman, Montana: Environmental Earth Sciences, v. 60, p. 273-284.

King, G.J., Acton, D.F., and St. Arnaud, R.J., 1983, Soil-landscape analysis in relation to soil distribution and mapping at a site within the Weyburn Association: Canadian Journal of Soil Science, v. 63, p. 657-670.

Kirk, M.F., Crossey, L.J., Takacs-Vesbach, C., Newell, D.L., and Bowman, R.S., 2009, Influence of upwelling saline groundwater on iron and manganese cycling in the Rio Grande floodplain aquifer: Applied Geochemistry, v. 24, p. 426-437.

Kirsch, K., Navarre-Sitchler, A.K., Wunsch, A., and McCray, J.E., 2014, Metal Release from Sandstones under Experimentally and Numerically Simulated CO_2 Leakage Conditions: Environmental Science & Technology, v. 48, p. 1436-1442.

Klusman, R., 2003, A geochemical perspective and assessment of leakage potential for a mature carbon dioxide-enhanced oil recovery project and as a prototype for carbon dioxide sequestration; Rangely field, Colorado: AAPG Bulletin, v. 87, p. 1485-1507.

Krevor, S., Perrin, J.-C., Esposito, A., Rella, C., and Benson, S., 2010, Rapid detection and characterization of surface CO_2 leakage through the real-time measurement of C signatures in CO_2 flux from the ground: International Journal of Greenhouse Gas Control, v. 4, p. 811-815.

Kruger, M., Jones, D., Frerichs, J., Oppermann, B.I., West, J., Coombs, P., Green, K., Barlow, T., Lister, R., Shaw, R., Strutt, M., and Moller, I., 2011, Effects of elevated CO_2 concentrations on the vegetation and microbial populations at a terrestrial CO_2 vent at Laacher See, Germany: International Journal of Greenhouse Gas Control, v. 5, p. 1093-1098.

Kuang, X., Jiao, J.J., and Li, H., 2013, Review on airflow in unsaturated zones induced by natural forcings: Water Resources Research, v. 49, p. 6137-6165.

Kutchko, B.G., Strazisar, B.R., Dzombak, D.A., Lowry, G.V., and Thaulow, N., 2007, Degradation of Well Cement by CO_2 under Geologic Sequestration Conditions: Environmental Science & Technology, v. 41, p. 4787-4792.

Kutchko, B.G., Strazisar, B.R., Huerta, N., Lowry, G.V., Dzombak, D.A., and Thaulow, N., 2009, CO_2 Reaction with Hydrated Class H Well Cement under Geologic Sequestration Conditions: Effects of Flyash Admixtures: Environmental Science & Technology, v. 43, p. 3947-3952.

Kutchko, B.G., Strazisar, B.R., Lowry, G.V., Dzombak, D.A., and Thaulow, N., 2008, Rate of CO_2 Attack on Hydrated Class H Well Cement under Geologic Sequestration Conditions: Environmental Science & Technology, v. 42, p. 6237-6242.

Kuzyakov, Y., 2006, Sources of CO_2 efflux from soil and review of partitioning methods: Soil Biology and Biochemistry, v. 38, p. 425-448.

Landi, A., Mermut, A.R., and Anderson, D.W., 2004, Carbon Distribution in a Hummocky Landscape from Saskatchewan, Canada: Soil Sci. Soc. Am. J., v. 68, p. 175-184.

Lemieux, J.-M., 2011, Review: The potential impact of underground geological storage of carbon dioxide in deep saline aquifers on shallow groundwater resources: Hydrogeology Journal, v. 19, p. 757-778.

Leon, E., Vargas, R., Bullock, S., Lopez, E., Panosso, A.R., and La Scala Jr, N., 2014, Hot spots, hot moments, and spatio-temporal controls on soil CO_2 efflux in a water-limited ecosystem: Soil Biology and Biochemistry, v. 77, p. 12-21.

Leung, D.Y.C., Caramanna, G., and Maroto-Valer, M.M., 2014, An overview of current status of carbon dioxide capture and storage technologies: Renewable and Sustainable Energy Reviews, v. 39, p. 426-443.

Leuning, R., Etheridge, D., Luhar, A., and Dunse, B., 2008, Atmospheric monitoring and verification technologies for CO_2 geosequestration: International Journal of Greenhouse Gas Control, v. 2, p. 401-414.

Lewicki, J.L., Bergfeld, D., Cardellini, C., Chiodini, G., Granieri, D., Varley, N., and Werner, C., 2005a, Comparative soil CO_2 flux measurements and geostatistical estimation methods on Masaya volcano, Nicaragua: Bull. Volcanol., v. 68, p. 76-90.

Lewicki, J.L., and Hilley, G.E., 2012, Eddy covariance network design for mapping and quantification of surface CO_2 leakage fluxes: International Journal of Greenhouse Gas Control, v. 7, p. 137-144.

Lewicki, J.L., Hilley, G.E., Dobeck, L., McLing, T.L., Kennedy, B.M., Bill, M., and Marino, B.D.V., 2013, Geologic CO_2 input into groundwater and the atmosphere, Soda Springs, ID, USA: Chemical Geology, v. 339, p. 61-70.

Lewicki, J.L., Hilley, G.E., Dobeck, L., and Spangler, L., 2010, Dynamics of CO_2 fluxes and concentrations during a shallow subsurface CO_2 release: Environmental Earth Sciences, v. 60, p. 285-297.

Lewicki, J.L., Hilley, G.E., and Oldenburg, C.M., 2005b, An improved strategy to detect CO_2 leakage for verification of geologic carbon sequestration: Geophysical Research Letters, v. 32, p. doi:10.1029/2005GL024281.

—, 2006, An improved strategy to detect CO_2 leakage for verification of geologic carbon sequestration, 8th International Conference on Greenhouse Gas Control Technologies: Trondheim, Norway.

Lewicki, J.L., Hilley, G.E., Tosha, T., Aoyagi, R., Yamamoto, K., and Benson, S.M., 2007a, Dynamic coupling of volcanic CO_2 flow and wind at the Horseshoe Lake tree kill, Mammoth Mountain, California: Geophysical Research Letters, v. 34, p. doi:10.1029/2006GL028848.

Lewicki, J.L., Oldenburg, C.M., Dobeck, L., and Spangler, L., 2007b, Surface CO_2 leakage during two shallow subsurface CO_2 releases: Geophysical Research Letters, v. 34, p. L24402, doi:10.1029/2007GL032047.

Lions, J., Devau, N., de Lary, L., Dupraz, S., Parmentier, M., Gombert, P., and Dictor, M.-C., 2014a, Potential impacts of leakage from CO_2 geological storage on geochemical processes controlling fresh groundwater quality: A review: International Journal of Greenhouse Gas Control, v. 22, p. 165-175.

Lions, J., Humez, P., Pauwels, H., Kloppmann, W., and Czernichowski-Lauriol, I., 2014b, Tracking leakage from a natural CO_2 reservoir (Montmiral, France) through the chemistry and isotope signatures of shallow groundwater: Greenhouse Gases: Science and Technology, v. 4, p. 225-243.

Little, M.G., and Jackson, R.B., 2010, Potential Impacts of Leakage from Deep CO_2 Geosequestration on Overlying Freshwater Aquifers: Environmental Science & Technology, v. 44, p. 9225-9232.

Lloyd, J., and Taylor, J.A., 1994, On the Temperature Dependence of Soil Respiration: Functional Ecology, v. 8, p. 315-323.

Loizzo, M., Lombardi, S., Deremble, L., Lecampion, B., Quesada, D., Huet, B., Khalfallah, I., Annunziatellis, A., and Picard, G., 2011, Monitoring CO_2 migration in an injection well: Evidence from MovECBM: Energy Procedia, v. 4, p. 5203-5210.

Lombardi, G., and Mattias, P., 1987, The kaolin deposits of Italy: L'Industria Mineraria, v. 6, p. 1-34.

Lombardi, S., Altunina, L.K., and Beaubien, S.E., 2006, Advances in the geological storage of carbon dioxide, NATO Science Series. IV. Earth and Environmental Sciences Volume 65, Springer, p. 362.

Lombardi, S., Annunziatellis, A., Beaubien, S.E., Ciotoli, G., and Coltella, M., 2008, Natural analogues and test sites for CO_2 geological sequestration: experience at Latera, Italy: First Break, v. 26, p. 39-43.

Lombardi, S., Di Filippo, M., and Manfra, L., 1985, Prospezione dell'elio nei gas del suolo nell'area dei Vulsini Orientali. Primi risultati del rilevamento del settembre 1984, Seminario Informativo sulle Attività di Ricerca del Sottoprogetto Energia Geotermica - Progetto Finalizzato Energetica (PFE) 1983-1987: Roma, Italia, Consiglio Nazionale delle Ricerche.

Lombardi, S., Etiope, G., and Pinti, D.L., 1994, Variations of ^{222}Rn in soil gas as seismic precursor signals at Latera Caldera (central Italy), in Furlan, G., and Tommasino, L., eds., Radon Monitoring in Radioprotection, Environmental and/or Earth Sciences, World Scientific Pub Co Inc, p. 418-434.

Lombardi, S., Manfra, L., and Baccani, A., 1989, Prospezione di elio nei gas del suolo nell'area Vulsina, Seminario Generale delle Unità Operative del PFE. Progetto Finalizzato Energetica 2, Sottoprogetto Energia Geotermica: Montecatini, Italia, Consiglio Nazionale delle Ricerche, p. 147-151.

Lombardi, S., Manfra, L., Brunamonti, F., and Romagnoli, C., 1986, Studio delle emanazioni dell'elio nell'area Vulsina, Seminario Generale delle Unità Operative del PFE. Progetto Finalizzato Energetica 2, Sottoprogetto Energia Geotermica: Pisa, Italia, Consiglio Nazionale delle Ricerche.

Lombardi, S., and Pinti, D.L., 1992, ^{222}Rn behavior at the Latera geothermal field (northern Latium, Italy): Journal of Radioanalytical and Nuclear Chemistry, v. 161, p. 365-375.

Lombardi, S., Pinti, D. L., Rossi, U., and Fiordalisi, A., 1993, ^{222}Rn in soil gases at Latera geothermal field:a preliminary case history: Geologica Romana, v. 29, p. 391-399.

Maier, M., Schack-Kirchner, H., Hildebrand, E.E., and Holst, J., 2010, Pore-space CO_2 dynamics in a deep, well-aerated soil: European Journal of Soil Science, v. 61, p. 877-887.

Marini, L., and Chiodini, G., 1994, The role of carbon dioxide in the carbonate-evaporite geothermal systems of Tuscany and Latium (Italy): Acta Vulcanol., v. 5, p. 95-104.

Mazot, A., Smid, E., Schwendenmann, L., Delgado-Granados, H., and Lindsay, J., 2013, Soil CO_2 flux baseline in an urban monogenetic volcanic field: the Auckland Volcanic Field, New Zealand: Bulletin of Volcanology C7 - 757, v. 75, p. 1-9.

Mickler, P.J., Yang, C., Scanlon, B.R., Reedy, R., and Lu, J., 2013, Potential Impacts of CO_2 Leakage on Groundwater Chemistry from Laboratory Batch Experiments and Field Push-Pull Tests: Environmental Science & Technology, v. 47, p. 10694-10702.

Mielnick, P.C., and Dugas, W.A., 2000, Soil CO_2 flux in a tallgrass prairie: Soil Biology and Biochemistry, v. 32, p. 221-228.

Minissale, A., 1991, Thermal springs in Italy: their relation to recent tectonics: Applied Geochemistry, v. 6, p. 201-212.

—, 2004, Origin, transport and discharge of CO_2 in central Italy: Earth-Science Reviews, v. 66, p. 89-141.

Minissale, A., Evans, W.C., Magro, G., and Vaselli, O., 1997, Multiple source components in gas manifestations from north-central Italy: Chemical Geology, v. 142, p. 175-192.

Minissale, A., Magro, G., Martinelli, G., Vaselli, O., and Tassi, G.F., 2000, Fluid geochemical transect in the Northern Apennines (central-northern Italy): fluid genesis and migration and tectonic implications: Tectonophysics, v. 319, p. 199-222.

Moncrieff, J.B., and Fang, C., 1999, A model for soil CO_2 production and transport 2: Application to a florida Pinus elliotte plantation: Agricultural and Forest Meteorology, v. 95, p. 237-256.

Moni, C., and Rasse, D.P., 2014, Detection of simulated leaks from geologically stored CO_2 with 13C monitoring: International Journal of Greenhouse Gas Control, v. 26, p. 61-68.

Montes-Hernandez, G., Renard, F.o., and Lafay, R., 2013, Experimental Assessment of CO_2-Mineral-Toxic Ion Interactions in a Simplified Freshwater Aquifer: Implications for CO_2 Leakage from Deep Geological Storage: Environmental Science & Technology, v. 47, p. 6247-6253.

Nappi, G., 1969, Genesi ed evoluzione della Caldera di Latera: Boll. Serv. Geol. It., v. 90, p. 61–8.

Nappi, G., Renzulli, A., and Santi, P., 1991, Evidence of incremental growth in the Vulsinian calderas (central Italy): J Volcanol. Geotherm. Res., v. 47, p. 13–31.

Natale, G., Hernández, P., Mori, T., and Notsu, K., 2000, Pressure gradient measurements in volcanic diffuse gas emanations: Geophysical Research Letters, v. 27, p. 3985-3987.

Natelhoffer, K.J., and Fry, B., 1988, Controls on Natural Nitrogen-15 and Carbon-13 Abundances in Forest Soil Organic Matter: Soil Sci. Soc. Am. J., v. 52, p. 1633-1640.

Navarre-Sitchler, A.K., Maxwell, R.M., Siirila, E.R., Hammond, G.E., and Lichtner, P.C., 2013, Elucidating geochemical response of shallow heterogeneous aquifers to CO_2 leakage using high-performance computing: Implications for monitoring of CO_2 sequestration: Advances in Water Resources, v. 53, p. 45-55.

Neufeld, J.A., Hesse, M.A., Riaz, A., Hallworth, M.A., Tchelepi, H.A., and Huppert, H.E., 2010, Convective dissolution of carbon dioxide in saline aquifers: Geophysical Research Letters, v. 37, p. L22404.

Nguyen, V.T., Zhao, L., and Zytner, R.G., 2013, Three-dimensional numerical model for soil vapor extraction: Journal of Contaminant Hydrology, v. 147, p. 82-95.

Nickerson, N., and Risk, D., 2009, Physical controls on the isotopic composition of soil-respired CO_2: J. Geophys. Res., v. 114, p. G01013.

Nicoletti, M., Petrucciani, C., Piro, M., and Trigila, R., 1979, Nuove datazioni vulsine per uno schema di evoluzione dell'attività vulcanica: il quadrante nord-occidentale: Periodico di Mineralogia, v. 47/48, p. 153–165.

—, 1981, Nuove datazioni vulsine per uno schema di evoluzione dell'attività vulcanica: nota II: il quadrante sud-occidentale.: Per. Miner. , v. 50, p. 141-169.

Nimz, G.J., and Bryant Hudson, G., 2005, The Use of Noble Gas Isotopes for Monitoring Leakage of Geologically Stored CO_2, Carbon Dioxide Capture for Storage in Deep Geologic Formations: Amsterdam, Elsevier Science, p. 1113-1128.

Noy, D.J., Holloway, S., Chadwick, R.A., Williams, J.D.O., Hannis, S.A., and Lahann, R.W., 2012, Modelling large-scale carbon dioxide injection into the Bunter Sandstone in the UK Southern North Sea: International Journal of Greenhouse Gas Control, v. 9, p. 220-233.

Oldenburg, C., Lewicki, J., Pan, L., Dobeck, L., and Spangler, L., 2010, Origin of the patchy emission pattern at the ZERT CO_2 release test: Environmental Earth Sciences, v. 60, p. 241-250.

Oldenburg, C.M., Lewicki, J.L., and Hepple, R.P., 2003, Near-surface monitoring strategies for geologic carbon dioxide storage verification.

Oldenburg, C.M., and Unger, A.J.A., 2004, Coupled vadose zone and atmospheric surface layer transport of carbon dioxide from geologic carbon sequestration sites: Vadose Zone Journal, v. 3, p. 848-857.

Oppermann, B.I., Michaelis, W., Blumenberg, M., Frerichs, J., Schulz, H.M., Schippers, A., Beaubien, S.E., and Kruger, M., 2010, Soil microbial community changes as a result of long-term exposure to a natural CO_2 vent: Geochimica et Cosmochimica Acta, v. 74, p. 2697-2716.

Pacific, V.J., McGlynn, B.L., Riveros-Iregui, D.A., Welsch, D.L., and Epstein, H.E., 2010, Landscape structure, groundwater dynamics, and soil water content influence soil respiration across riparian–hillslope transitions in the Tenderfoot Creek Experimental Forest, Montana: Hydrological Processes, v. 25, p. 811-827.

Padron, E., Hernandez, P.A., Toulkeridis, T., Perez, N.M., Marrero, R., Melian, G., Virgili, G., and Notsu, K., 2008, Diffuse CO_2 emission rate from Pululahua and the lake-filled Cuicocha calderas, Ecuador: Journal of Volcanology and Geothermal Research, v. 176, p. 163-169.

Panichi, C., and Tongiorgi, E., 1976, Carbon isotopic composition of CO_2 from springs, fumaroles, mofettes, and travertines of central and southern Italy: a preliminary prospection method of geothermal areas., Proc. 2nd U. N. Symp. Development and Use of Geothermal Energy., p. 815-825.

Parkhurst, D.L., and Appelo, C.A.J., 2013, Description of input and examples for PHREEQC version 3--A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Book 6, Chap. A43, U.S. Geological Survey Techniques and Methods. pp.497. Available only at http://pubs.usgs.gov/tm/06/a43/.

Parks, M.M., Caliro, S., Chiodini, G., Pyle, D.M., Mather, T.A., Berlo, K., Edmonds, M., Biggs, J., Nomikou, P., and Raptakis, C., 2013, Distinguishing contributions to diffuse CO_2 emissions in volcanic areas from magmatic degassing and thermal decarbonation using soil gas ^{222}Rn-^{13}C systematics: Application to Santorini volcano, Greece: Earth and Planetary Science Letters, v. 377-378, p. 180-190.

Pearce, J.M., Czernichowski-Lauriol, I., Lombardi, S., Brune, S., Nador, A., Baker, J., Pauwels, H., Hatziyannis, G., Beaubien, S.E., and Faber, E., 2004, A review of natural CO_2 accumulations in Europe as analogues for geological sequestration, in Baines, S.J., and Worden, R.H., eds., Geological Storage of Carbon Dioxide, Volume Special Publication, 233: London, The Geological Society of London, p. 29-41.

Pettinelli, E., Beaubien, S.E., Lombardi, S., and Annan, A.P., 2008, GPR, TDR, and geochemistry measurements above an active gas vent to study near-surface gas-migration pathways: Geophysics, v. 73, p. A11-A15.

Pettinelli, E., Beaubien, S.E., Zaja, A., Menghini, A., Praticelli, N., Mattei, E., Di Matteo, A., Annunziatellis, A., Ciotoli, G., and Lombardi, S., 2010, Characterization of a CO_2 gas vent using various geophysical and geochemical methods: Geophysics, v. 75, p. B137-B146.

Phillips, R.L., Wick, A.F., Liebig, M.A., West, M.S., and Daniels, W.L., 2012, Biogenic emissions of CO_2 and N_2O at multiple depths increase exponentially during a simulated soil thaw for a northern prairie Mollisol: Soil Biology and Biochemistry, v. 45, p. 14-22.

Polak, A., Elsworth, D., Liu, J., and Grader, A.S., 2004, Spontaneous switching of permeability changes in a limestone fracture with net dissolution: Water Resources Research, v. 40, p. W03502.

Ramnarine, R., Wagner-Riddle, C., Dunfield, K.E., and Voroney, R.P., 2012, Contributions of carbonates to soil CO_2 emissions: Canadian Journal of Soil Science, v. 92, p. 599-607.

Ranjith, P.G., Choi, S.K., and Fourar, M., 2006, Characterization of two-phase flow in a single rock joint: International Journal of Rock Mechanics and Mining Sciences, v. 43, p. 216-223.

Riley, P.R., Goodwin, L.B., and Lewis, C.J., 2010, Controls on fault damage zone width, structure, and symmetry in the Bandelier Tuff, New Mexico: Journal of Structural Geology, v. 32, p. 766-780.

Risk, D., Kellman, L., and Beltrami, H., 2002a, Carbon dioxide in soil profiles: Production and temperature dependence: Geophys. Res. Lett., v. 29, p. 1087.

—, 2002b, Soil CO_2 production and surface flux at four climate observatories in eastern Canada: Global Biogeochem. Cycles, v. 16, p. 1122.

Risk, D., Nickerson, N., Phillips, C.L., Kellman, L., and Moroni, M., 2012, Drought alters respired d13CO_2 from autotrophic, but not heterotrophic soil respiration: Soil Biology and Biochemistry, v. 50, p. 26-32.

Rissmann, C., Christenson, B., Werner, C., Leybourne, M., Cole, J., and Gravley, D., 2012, Surface heat flow and CO_2 emissions within the Ohaaki hydrothermal field, Taupo Volcanic Zone, New Zealand: Applied Geochemistry, v. 27, p. 223-239.

Riveros-Iregui, D.A., and McGlynn, B.L., 2009, Landscape structure control on soil CO_2 efflux variability in complex terrain: Scaling from point observations to watershed scale fluxes: J. Geophys. Res., v. 114, p. G02010.

Rochelle, C.A., and Milodowski, A.E., 2013, Carbonation of borehole seals: Comparing evidence from short-term laboratory experiments and long-term natural analogues: Applied Geochemistry, v. 30, p. 161-177.

Romanak, K.D., Bennett, P.C., Yang, C., and Hovorka, S.D., 2012a, Process-based approach to CO_2 leakage detection by vadose zone gas monitoring at geologic CO_2 storage sites: Geophys. Res. Lett., v. 39.

Romanak, K.D., Smyth, R.C., Yang, C., Hovorka, S.D., Rearick, M., and Lu, J., 2012b, Sensitivity of groundwater systems to CO_2: Application of a site-specific analysis of carbonate monitoring parameters at the SACROC CO_2-enhanced oil field: International Journal of Greenhouse Gas Control, v. 6, p. 142-152.

Romanak, K.D., Wolaver, B., Yang, C., Sherk, G.W., Dale, J., Dobeck, L.M., Spangler, L.H., 2014. Process-based soil gas leakage assessment at the Kerr Farm: Comparison of results to leakage proxies at ZERT and Mt. Etna. International Journal of Greenhouse Gas Control 30, 42-57.

Rosenbaum, J.M., 1994, Stable isotope fractionation between carbon dioxide and calcite at 900°C: Geochimica et Cosmochimica Acta, v. 58, p. 3747-3753.

Rouchon, V., Magnier, C., Miller, D., Bandeira, C., Goncalves, R., and Dino, R., 2010, The relationship between CO_2 flux and gas composition in soils above an EOR- CO_2 oil field (Brazil): A guideline for the surveillance of CO_2 storage sites: Energy Procedia, v. 4, p. 3354-3362.

Sabatelli, F., and Mannari, M., 1995, Latera development update., Proceedings world geothermal congress: Florence

Sano, Y., and Marty, B., 1995, Origin of carbon in fumarolic gas from island arcs: Chemical Geology, v. 119, p. 265-274.

Sauer, U., Schutze, C., Leven, C., Schlomer, S., and Dietrich, P., 2014, An integrative hierarchical monitoring approach applied at a natural analogue site to monitor CO_2 degassing areas: Acta Geotechnica, v. 9, p. 127-133.

Scanlon, B.R., Nicot, J.-P., and Massmann, J.W., 2002, Soil gas movement in unsaturated systems, in Warrick, A.W., ed., Soil Physics Companion Boca Raton, CRC Press, p. 297-341.

Schutze, C., Dietrich, P., and Sauer, U., 2013, Diagnostic monitoring to identify preferential near-surface structures for CO_2 degassing into the atmosphere: Tools for investigations at different spatial scales validated at a natural analogue site: International Journal of Greenhouse Gas Control, v. 18, p. 285-295.

Schutze, C., Sauer, U., Beyer, K., Lamert, H., Brauer, K., Strauch, G., Flechsig, C., Kampf, H., and Dietrich, P., 2012, Natural analogues: a potential approach for developing reliable monitoring methods to understand subsurface CO_2 migration processes: Environmental Earth Sciences, p. 1-13.

Schweizer, M., Fear, J., and Cadisch, G., 1999, Isotopic (^{13}C) Fractionation During Plant Residue Decomposition and its Implications for Soil Organic Matter Studies: RAPID COMMUNICATIONS IN MASS SPECTROMETRY, v. 13, p. 1284-1290.

Serrano-Ortiz, P., Roland, M., Sanchez-Moral, S., Janssens, I.A., Domingo, F., Godderis, Y., and Kowalski, A.S., 2010, Hidden, abiotic CO_2 flows and gaseous reservoirs in the terrestrial carbon cycle: Review and perspectives: Agricultural and Forest Meteorology, v. 150, p. 321-329.

Shipton, Z.K., Evans, J.P., Dockrill, B., Heath, J., Williams, A., Kirchner, D., and Kolesar, P.T., 2005, Natural leaking CO_2-charged systems as analogs for failed geologic storage reservoirs, in Benson, S.M., Oldenburg, C.M., Hoversten, M., and Imbus, S., eds., Carbon Dioxide Capture for Storage in Deep Geologic Formations - Results from the CO_2 Capture Project, Volume 2, Elsevier Science, p. 699-712.

Sibson, R.H., 2000, Fluid involvement in normal faulting: Journal of Geodynamics, v. 29, p. 469-499.

Simei, S., Acocella, V., Palladino, D.M., and Trigila, R., 2006, Evolution and structure of Vulsini calderas (Italy): Geophys. Res. Abs., v. 8, p. 09302.

Sinclair, A.J., 1991, A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited: J. Geochem. Explor, v. 41, p. 23.

Smith, B.N., and Epstein, S., 1971, Two Categories of $^{13}C/^{12}C$ Ratios for Higher Plants: Plant Physiology, v. 47, p. 380-384.

Stevens, S.H., 2005, Natural CO_2 fields as analogs for geologic storage, in Benson, S.M., Oldenburg, C.M., Hoversten, M., and Imbus, S., eds., Carbon Dioxide Capture for Storage in Deep Geologic Formations - Results from the CO_2 Capture Project, Volume 2, Elsevier Science, p. 699-712.

Strazisar, B.R., Wells, A.W., Diehl, J.R., Hammack, R.W., and Veloski, G.A., 2009, Near-surface monitoring for the ZERT shallow CO_2 injection project: Int. J. Greenhouse Gas Control, v. 3, p. 736-744.

Stumm, W., and Morgan, J.J., 1995, Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, 3rd Edition, Wiley, 1040 p.

Takahashi, M., 2003, Permeability change during experimental fault smearing: Journal of Geophysical Research: Solid Earth, v. 108, p. 2235.

Tenthorey, E., and Fitz Gerald, J.D., 2006, Feedbacks between deformation, hydrothermal reaction and permeability evolution in the crust: Experimental insights: Earth and Planetary Science Letters, v. 247, p. 117-129.

Trautz, R.C., Pugh, J.D., Varadharajan, C., Zheng, L., Bianchi, M., Nico, P.S., Spycher, N.F., Newell, D.L., Esposito, R.A., Wu, Y., Dafflon, B., Hubbard, S.S., and Birkholzer, J.T., 2013, Effect of Dissolved CO_2 on a Shallow Groundwater System: A Controlled Release Field Experiment: Environmental Science & Technology, v. 47, p. 298-305.

Uehara, S.., and Shimamoto, T., 2004, Gas permeability evolution of cataclasite and fault gouge in triaxial compression and implications for changes in fault-zone permeability structure through the earthquake cycle: Tectonophysics, v. 378, p. 183-195.

Verkerke, J.L., Williams, D.J., and Thoma, E., 2014, Remote sensing of CO_2 leakage from geologic sequestration projects: International Journal of Applied Earth Observation and Geoinformation, v. 31, p. 67-77.

Vezzoli, L., Conticelli, S., Innocenti, F., Landi, P., Manetti, P., Palladino, D.M., and Trigila, R., 1987, Stratigraphy of the Latera Volcanic Complex: proposals for a new nomenclature: Per. Mineral., v. 56, p. 89–110.

Viswanathan, H.S., Pawar, R.J., Stauffer, P.H., Kaszuba, J.P., Carey, J.W., Olsen, S.C., Keating, G.N., Kavetski, D., and Guthrie, G.D., 2008, Development of a Hybrid Process and System Model for the Assessment of Wellbore Leakage at a Geologic CO_2 Sequestration Site: Environmental Science & Technology, v. 42, p. 7280-7286.

Walter, L., Binning, P.J., and Class, H., 2013, Predicting salt intrusion into freshwater aquifers resulting from CO_2 injection - A study on the influence of conservative assumptions: Advances in Water Resources, v. 62, Part C, p. 543-554.

Webb, S.W., 2006a, Gas transport mechanisms, in Ho, C.K., and Webb, S.W., eds., Gas Transport in Porous Media: Theory and Applications of Transport in Porous Media: Dordrecht, The Netherlands, Springer, p. 5-26.

—, 2006b, Two-phase gas transport, in Ho, C.K., and Webb, S.W., eds., Gas Transport in Porous Media: Theory and Applications of Transport in Porous Media: Dordrecht, The Netherlands, Springer, p. 55-70.

Wells, A.W., Diehl, J.R., Bromhal, G., Strazisar, B.R., Wilson, T.H., and White, C.M., 2007, The use of tracers to assess leakage from the sequestration of CO_2 in a depleted oil reservoir, New Mexico, USA: Applied Geochemistry, v. 22, p. 996-1016.

White, C.M., Strazisar, B.R., Granite, E.J., Hoffman, J.S., and Pennline, H.W., 2003, Separation and Capture of CO_2 from Large Stationary Sources and Sequestration in Geological Formations - Coalbeds and Deep Saline Aquifers: Journal of the Air & Waste Management Association, v. 53, p. 645-715.

Whiticar, M.J., 1996, Stable isotope geochemistry of coals, humic kerogens and related natural gases: International Journal of Coal Geology, v. 32, p. 191-215.

Wigley, M., Kampman, N., Chapman, H.J., Dubacq, B., and Bickle, M.J., 2013, In situ redeposition of trace metals mobilized by CO_2-charged brines: Geochemistry, Geophysics, Geosystems, v. 14, p. 1321-1332.

Wunsch, A., Navarre-Sitchler, A.K., and McCray, J.E., 2013a, Geochemical Implications of Brine Leakage into Freshwater Aquifers: Ground Water.

Wunsch, A., Navarre-Sitchler, A.K., Moore, J., and McCray, J.E., 2014, Metal release from limestones at high partial-pressures of CO_2: Chemical Geology, v. 363, p. 40-55.

Wunsch, A., Navarre-Sitchler, A.K., Moore, J., Ricko, A., and McCray, J.E., 2013b, Metal release from dolomites at high partial-pressures of CO_2: Applied Geochemistry, v. 38, p. 33-47.

Yang, C., Mickler, P.J., Reedy, R., Scanlon, B.R., Romanak, K.D., Nicot, J.-P., Hovorka, S.D., Trevino, R.H., and Larson, T., 2013, Single-well push-pull test for assessing potential impacts of CO_2 leakage on groundwater quality in a shallow Gulf Coast aquifer in Cranfield, Mississippi: International Journal of Greenhouse Gas Control, v. in press.

Zheng, L., Apps, J.A., Spycher, N., Birkholzer, J.T., Kharaka, Y.K., Thordsen, J., Beers, S.R., Herkelrath, W.N., Kakouros, E., and Trautz, R.C., 2012, Geochemical modeling of changes in shallow groundwater chemistry observed during the MSU-ZERT CO_2 injection experiment: International Journal of Greenhouse Gas Control, v. 7, p. 202-217.

Zheng, L., Apps, J.A., Zhang, Y., Xu, T., and Birkholzer, J.T., 2009, On mobilization of lead and arsenic in groundwater in response to CO_2 leakage from deep geological storage: Chemical Geology, v. 268, p. 281-297.

APPENDIX 1

GasGrid 7.5 code

```
Dim path As String

Option Explicit

Private Sub Cmd_cancel_Click()
End
End Sub

Private Sub Cmd_make_grid_Click()

Randomize

Frm_initial_settings.Enabled = False
Frm_data_type.Enabled = False
Frm_bkg_flux.Enabled = False
Frm_vents.Enabled = False
Frm_sampling_settings.Enabled = False
Frm_contour.Enabled = False
Frm_run.Enabled = False

PauseIt

'*************************************************************
'*****************      file name variables     ***************
'*************************************************************

Dim BkgInputFile As String 'imported file from BG_Format program
Dim BkgPlusVentFile As String
Dim BkgPlusVentGridFile As String
Dim SampleDataFile As String
Dim SampleGridFile As String
Dim SummaryFile As String
Dim vent_info_file As String
Dim sample_spacing_file As String
Dim fixed_num_sample_file As String
Dim BackgroundFluxFile As String
Dim BackgroundGridFile As String
Dim RealFieldDataFile As String 'imported file of real field data
Dim SimulationDescriptionFile As String 'file summarising simulation settings
Dim sim_dir As String 'used to create new root directory for series of simulations
Dim sim_path As String 'used to number simulation directory
Dim ss_dir As String 'sample spacing directory
```

```
'*****************************************************
'******************     variables     ***************
'*****************************************************

Dim anomalies As Integer
Dim avg_bkg_flux As Double
Dim bkg_max As Double
Dim circle_X_min As Integer
Dim circle_X_max As Integer
Dim circle_Y_min As Integer
Dim circle_Y_max As Integer
Dim count_loop As Long 'used to see how many loops are needed before sample point is NOT ""
Dim ContourType As String
Dim core_radius As Integer
Dim CVResults As Variant
Dim distance As Double
Dim EqnVariable As Double
Dim ExtraPoints As Integer
Dim flux As Double
Dim grid_pt_seq_num As Long
Dim grid_X_max As Long
Dim grid_Y_max As Long
Dim num_samples As Long
Dim num_simulations As Integer
Dim num_sample_spacings As Long
Dim fixed_num_samples As Long
Dim num_of_fixed_samples As Long
Dim Num_of_vents As Long
Dim OddEven As Integer
Dim point_to_sample As Variant
Dim random_number As Long
Dim remaining_pts As Long
Dim sample_pt As Variant
Dim sample_pt_X As Integer
Dim sample_pt_Y As Integer
Dim sample_spacing As Integer
Dim SampleSpacingTitle As String
Dim sampling_density As Long
Dim sigma_bkg_flux As Double
Dim starting_point As Long
Dim TotalFlux As Variant 'total flux calculated ONCE using all data (simulated=BG+vent; realdata=realdata)
Dim TotalFluxBackground As Variant 'total flux calculated ONCE for BG
Dim TotalFluxVents As Variant 'total vent flux calculated ONCE using all data: TotalFlux - TotalFluxBackground
Dim TotalFluxSim As Variant 'total flux calculated for each simulation
Dim TotalFluxVentsSim As Variant 'total flux from vents calculated for each simulation
Dim total_num_pts As Long
Dim ventloop As Integer
Dim Vent_Radius_x_title As String 'used to make title of simulation directory
Dim X_length As Integer
Dim X_coord As Integer
Dim X_samples As Integer
Dim Y_length As Integer
Dim Y_coord As Integer
Dim Y_samples As Integer
Dim StartTime As Date 'stores starting time of simulation
```

```
'*************************************************
'*****************     counters     ***************
'*************************************************

Dim counter_1 As Long
Dim counter_2 As Long
Dim count_samples As Long
Dim Count_Row As Integer
Dim Count_Column As Integer
Dim dir_count As Integer
Dim grid_array_counter As Long
Dim grid_check_array_counter As Long
Dim i As Long
Dim j As Long
Dim k As Long
Dim l As Long
Dim m As Long
Dim o As Long
Dim p As Long
Dim s As Long
Dim y As Long
Dim z As Integer

'*************************************************************
'*****************      arrays     ***************
'*************************************************************

Dim CV(1, 6) As Variant
Dim grid_array() As Variant
Dim grid_check_array() As Variant 'used as "parallel" array for random selection of sample points (or for first point for grid sampling)
Dim sample_array() As Variant
Dim vent_info() As Variant
Dim Results() As Double
Dim sample_spacing_array() As Variant
Dim fixed_num_sample_array() As Variant
Dim Imported_BG_array() As Variant
Dim Imported_RealData_array() As Variant

'*********************************************************************
'*****************      collect info from form     ***************
'*********************************************************************

StartTime = Time()

num_simulations = CInt(Txt_num_simulations)
avg_bkg_flux = CDbl(Txt_BG_avg.Text)
grid_X_max = CInt(Txt_X_max.Text)
grid_Y_max = CInt(Txt_Y_max.Text)
total_num_pts = CLng(Txt_num_grid_pts.Text)

ReDim grid_check_array(1, total_num_pts)

Dim fso As New Scripting.FileSystemObject ' Get instance of FileSystemObject.
If fso.FolderExists(Txt_path.Text) = False Then
```

```
   fso.CreateFolder (Txt_path.Text)
End If
path = Txt_path.Text

If Opt_contour.Item(0).Value = True Then ContourType = "srfKriging"
If Opt_contour.Item(1).Value = True Then ContourType = "srfNaturalNeighbor"
If Opt_contour.Item(2).Value = True Then ContourType = "srfInverseDistance"
If Opt_contour.Item(3).Value = True Then ContourType = "srfRadialBasis"

'*************************************************************************
'*****************     create folder for these simulations   *************
'*************************************************************************

sim_path = CStr(path) & "output\sim"
sim_dir = sim_path
dir_count = 1
Do Until fso.FolderExists(sim_dir) = False
   sim_dir = sim_path & CStr(dir_count)
   dir_count = dir_count + 1
Loop

fso.CreateFolder (sim_dir) ' Create a new folder with the FileSystemObject object.

'*****************************************************************
'** import fixed number of samples for purely random sampling  **
'*****************************************************************

If Opt_fixed_num_samples(0).Value = True Then
   fixed_num_samples = 1
   ReDim fixed_num_sample_array(1, 1)
   fixed_num_sample_array(1, 1) = CInt(Txt_single_fixed_num_samples.Text)
ElseIf Opt_fixed_num_samples(1).Value = True Then
   Txt_message.Text = "import fixed number of samples for random sampling"
   PauseIt
   fixed_num_sample_file = Txt_fixed_sample_number_file.Text
   fixed_num_samples = Txt_fixed_num_samples.Text
   ReDim fixed_num_sample_array(1, fixed_num_samples) As Variant
   Call ImportIntoArray(fixed_num_sample_array, 1, fixed_num_samples, fixed_num_sample_file, 1)
End If

'*************************************************************************
'* import "minimum" sample spacings for systematic random sampling *****
'* MAYBE import sample spacings for regular grid sampling       *******
'*************************************************************************

If Opt_sample_spacing(0).Value = True Then
   num_sample_spacings = 1
   ReDim sample_spacing_array(1, 1)
   sample_spacing_array(1, 1) = CInt(Txt_singlespacing.Text)
ElseIf Opt_sample_spacing(1).Value = True Then
   Txt_message.Text = "import sample spacings"
   PauseIt
   sample_spacing_file = Txt_sample_spacing_file.Text
   num_sample_spacings = Txt_num_sample_space.Text
   ReDim sample_spacing_array(1, num_sample_spacings) As Variant
   Call ImportIntoArray(sample_spacing_array, 1, num_sample_spacings, sample_spacing_file, 1)
```

```
End If

'*****************************************************************
'******     import vent locations and characteristics     ******
'*****************************************************************

Txt_message.Text = "import vent locations"
PauseIt

If Opt_vents(0).Value = True Then
   ReDim vent_info(5, 1)
   Num_of_vents = Txt_num_of_vents.Text
   vent_info(1, 1) = CInt(Txt_singlevent_X.Text)
   vent_info(2, 1) = CInt(Txt_singlevent_Y.Text)
   vent_info(3, 1) = CInt(Txt_MaxVentFlux.Text)
   vent_info(4, 1) = CInt(Txt_core_radius.Text)
   vent_info(5, 1) = CInt(Txt_vent_radius.Text)
ElseIf Opt_vents(1).Value = True Then
   vent_info_file = Txt_vent_info_file.Text
   Num_of_vents = Txt_num_of_vents.Text
   ReDim vent_info(5, Num_of_vents) As Variant
   Call ImportIntoArray(vent_info, 5, Num_of_vents, vent_info_file, 1)
End If

'*****************************************************************
'******     create file with description of simulation     *****
'*****************************************************************

SimulationDescriptionFile = sim_dir & "\Simulation Description.txt"
Open SimulationDescriptionFile For Append As #1
Print #1, Txt_comments.Text
Print #1, "X max = " & Txt_X_max.Text
Print #1, "Y max = " & Txt_Y_max.Text
Print #1, "total number of pts = " & Txt_num_grid_pts.Text
If Opt_data_type(0) = True Then
   Print #1, "Simulated data"
ElseIf Opt_data_type(1) = True Then
   Print #1, "real data"
   Print #1, "imported real data file = " & Txt_field_data_file.Text
End If
If Opt_Sim_Type(0) = True Then
   Print #1, "simulation type - find anomaly"
ElseIf Opt_Sim_Type(1) = True Then
   Print #1, "simulation type - calculate flux"
End If
If Opt_BG(0) = True Then
   Print #1, "fixed average = " & Txt_BG_avg.Text
ElseIf Opt_BG(1) = True Then
   Print #1, "calculated normal distribution background, average = " & Txt_BG_avg.Text & ", sigma = " &
Txt_BG_sigma.Text
ElseIf Opt_BG(2) = True Then
   Print #1, "imported log-normal distribution background, average = " & Txt_BG_avg.Text & ", sigma = " &
Txt_BG_sigma.Text
   Print #1, "imported background file = " & Txt_BG_File.Text
End If
If Opt_vents(0) = True Then
```

```
   Print #1, "single vent, X" & Txt_singlevent_X.Text & " Y" & Txt_singlevent_Y.Text & ", vent radius = " & Txt_vent_radius.Text
ElseIf Opt_vents(1) = True Then
   For z = 1 To Num_of_vents
      Print #1, "multiple vents - X" & vent_info(1, z) & " Y" & vent_info(2, z) & " max flux = " & vent_info(3, z) & " vent core diameter = " & vent_info(4, z) & " vent radius = " & vent_info(5, z)
   Next
End If
Print #1, "number of simulations per sample spacing = " & Txt_num_simulations.Text
If Opt_sampling_strategy(0) = True Then
   Print #1, "sampling strategy is purely random"
ElseIf Opt_sampling_strategy(1) = True Then
   Print #1, "sampling strategy is distributed random"
ElseIf Opt_sampling_strategy(2) = True Then
   Print #1, "sampling strategy is systematic random (grid)"
End If
If Opt_sample_spacing(0) = True Then
   Print #1, "single sample spacing = " & Txt_singlespacing.Text
ElseIf Opt_sample_spacing(1) = True Then
   For z = 1 To num_sample_spacings
      Print #1, "multiple sample spacing - " & sample_spacing_array(1, z)
   Next
End If
If Opt_fixed_num_samples(0) = True Then
   Print #1, "single fixed number of samples = " & Txt_single_fixed_num_samples.Text
ElseIf Opt_fixed_num_samples(1) = True Then
   For z = 1 To CInt(Txt_fixed_num_samples.Text)
      Print #1, "multiple fixed number of samples - " & fixed_num_sample_array(1, z)
   Next
End If
If Opt_contour(0) = True Then
   Print #1, "Kriging, anisotropy ratio = " & Txt_AnisotropyRatio & ", Anisotropy angle = " & Txt_AnisotropyAngle
ElseIf Opt_contour(1) = True Then
   Print #1, "Natural Neighbour, anisotropy ratio = " & Txt_AnisotropyRatio & ", Anisotropy angle = " & Txt_AnisotropyAngle
ElseIf Opt_contour(2) = True Then
   Print #1, "Inverse distance to a power, anisotropy ratio = " & Txt_AnisotropyRatio & ", Anisotropy angle = " & Txt_AnisotropyAngle & "inverse distance power = " & Txt_IDPower.Text & "inverse distance smoothing = " & Txt_IDSmoothing.Text
ElseIf Opt_contour(3) = True Then
   Print #1, "Radial Basis Function, anisotropy ratio = " & Txt_AnisotropyRatio & ", Anisotropy angle = " & Txt_AnisotropyAngle & "radial basis function r2 = " & Txt_RBRSquared.Text & "radial basis function type = " & Cbo_RBType.Text
End If
Close #1

'*****************************************************************
'***********       import background dataset       ************
'*****************************************************************

Txt_message.Text = "import background dataset"
PauseIt

If Opt_BG(2).Value = True Then
   BkgInputFile = Txt_BG_File.Text
   ReDim Imported_BG_array(1, total_num_pts) As Variant
```

```
   Call ImportIntoArray(Imported_BG_array, 1, total_num_pts, BkgInputFile, 2)
End If

'*****************************************************************
'*************       import real field data       **************
'*****************************************************************

Txt_message.Text = "import real field data"
PauseIt

If Opt_data_type(1).Value = True Then      'real data
   RealFieldDataFile = Txt_field_data_file.Text
   ReDim Imported_RealData_array(1, total_num_pts) As Variant
   Call ImportIntoArray(Imported_RealData_array, 1, total_num_pts, RealFieldDataFile, 2)
End If

'*************************************************************************
'*** Create background grid with calculated or imported distribution *****
'*** ----- OR ----- input real data into grid ----------------------******
'*************************************************************************

If Opt_BG(0).Value = True Then
   Txt_message.Text = "populate grid with real data"
ElseIf Opt_BG(1).Value = True Then
   Txt_message.Text = "populate grid with calculated BG normal random flux"
ElseIf Opt_BG(2).Value = True Then
   Txt_message.Text = "populate grid with imported BG random flux"
End If
PauseIt

ReDim grid_array(5, total_num_pts)
grid_pt_seq_num = 1
X_coord = 1
Y_coord = 1
bkg_max = avg_bkg_flux
BackgroundFluxFile = sim_dir & "\" & "background data.txt"
Open BackgroundFluxFile For Output As #1

For Y_coord = 1 To grid_Y_max
   For X_coord = 1 To grid_X_max
      grid_array(1, grid_pt_seq_num) = X_coord
      grid_array(2, grid_pt_seq_num) = Y_coord
      If Opt_BG(0).Value = True Then 'fixed avg BG = import real data
         grid_array(3, grid_pt_seq_num) = Imported_RealData_array(1, grid_pt_seq_num) 'populate grid w imported BG
dataset
      ElseIf Opt_BG(1).Value = True Then 'calculate BG
         sigma_bkg_flux = CDbl(Txt_BG_sigma.Text)
         grid_array(3, grid_pt_seq_num) = Round(Normal(sigma_bkg_flux, avg_bkg_flux), 3) 'populate grid w normally-
distributed random flux values based on input average and sigma
         If bkg_max < grid_array(3, grid_pt_seq_num) Then bkg_max = grid_array(3, grid_pt_seq_num)
      ElseIf Opt_BG(2).Value = True Then 'import BG
         grid_array(3, grid_pt_seq_num) = Round(Imported_BG_array(1, grid_pt_seq_num), 3) 'populate grid w imported
BG dataset
         If bkg_max < grid_array(3, grid_pt_seq_num) Then bkg_max = grid_array(3, grid_pt_seq_num)
      End If
```

```
      grid_array(4, grid_pt_seq_num) = grid_pt_seq_num 'sequential number used to select random points and to filter out pts that have been cancelled within sampling radius
      grid_array(5, grid_pt_seq_num) = grid_array(3, grid_pt_seq_num) 'used later to see if pt's BGflux has been over-written with vent value
      If Opt_BG(0).Value = False Then 'ie if simulated data because fixed avg BG = import real data
         Write #1, grid_array(1, grid_pt_seq_num), grid_array(2, grid_pt_seq_num), grid_array(3, grid_pt_seq_num)
      End If
      grid_pt_seq_num = grid_pt_seq_num + 1
   Next X_coord
Next Y_coord
Close #1

'*********************************************************************
'*** Grid Background Data using Surfer to calculate BG flux using **
'***(imported file or normal distribution created by program *******
'** OR calculate BG flux for real data using input average BG flux *
'*********************************************************************
If Opt_Sim_Type(1).Value = True Then 'if calculate flux and NOT find anomaly
   BackgroundGridFile = sim_dir & "\backgroundgrid.grd"
   Txt_message.Text = "grid background data"
   PauseIt
   '++Surfer calculation for simulated data; simple calculation for realdata+
   If Opt_data_type(0).Value = True Then      'simulated data
      Call SurferGrid(BackgroundFluxFile, ContourType, CInt(Txt_AnisotropyRatio.Text), CInt(Txt_AnisotropyAngle.Text), BackgroundGridFile, grid_X_max, grid_Y_max, 2, TotalFluxBackground)
   ElseIf Opt_data_type(1).Value = True Then      'real data
      TotalFluxBackground = avg_bkg_flux * (grid_X_max * grid_Y_max) 'average times total surface area
   End If
End If

'*************************************************************************
'** Opt_data_type(0) - add vents to simulated background grid   **********
'***Opt_data_type(1) - if real data, the array is not touched   **********
'*************************************************************************

If Opt_data_type(0).Value = True Then 'ie for simulated data
If Opt_vents(2).Value = False Then 'skip if no vents added to synthetic data

   Txt_message.Text = "create vents"
   PauseIt

   For ventloop = 1 To Num_of_vents

      EqnVariable = 0.065 * Log(vent_info(3, ventloop)) - 0.1525 'used to calculate flux at a vent point
                                 'created empirically in EXCEL - vent_info(3,ventloop) is MaxVentFlux

      circle_X_min = vent_info(1, ventloop) - vent_info(5, ventloop) 'vent_info(5, ventloop)=vent_radius
      If circle_X_min < 1 Then circle_X_min = 1
      circle_X_max = vent_info(1, ventloop) + vent_info(5, ventloop)
      If circle_X_max > grid_X_max Then circle_X_max = grid_X_max
      circle_Y_min = vent_info(2, ventloop) - vent_info(5, ventloop)
      If circle_Y_min < 1 Then circle_Y_min = 1
      circle_Y_max = vent_info(2, ventloop) + vent_info(5, ventloop)
      If circle_Y_max > grid_Y_max Then circle_Y_max = grid_Y_max

      X_length = (circle_X_max - circle_X_min) + 1
```

```
      Y_length = (circle_Y_max - circle_Y_min) + 1


      counter_1 = ((circle_Y_min - 1) * grid_X_max) + circle_X_min

      For i = 1 To Y_length
         For j = 1 To X_length
            distance = (((vent_info(1, ventloop) - grid_array(1, counter_1)) ^ 2) + ((vent_info(2, ventloop) - grid_array(2,
counter_1)) ^ 2)) ^ 0.5
            If distance <= vent_info(5, ventloop) Then 'vent_info(5, ventloop)=vent_radius
               'fixed max vent flux or calculated flux decrease - for find anomaly or calculate flux options
               If Opt_Sim_Type(0).Value = True Then 'for "Find Anomaly" option
                  flux = vent_info(3, ventloop)
               ElseIf Opt_Sim_Type(1).Value = True Then 'for "Calculate Flux" option
                  'equation based on GVA with 20m radius, therefore the ratio
                  'of (20m / specified vent radius) is used to scale the equation
                  flux = Round(Exp(-EqnVariable * ((20 / vent_info(5, ventloop)) * distance)) * vent_info(3, ventloop), 3)
'vent_info(3,1)=MaxVentFlux; vent_info(5, 1)=vent_radius
               End If
               'the following adds vent flux to BGflux if flux at that point has never been changed
               'or if the flux had been changed (ie adjacent vent) but is lower than new value
               ' - instead if flux had been changed and is higher than new flux, nothing is changed
               If grid_array(3, counter_1) = grid_array(5, counter_1) Then
                  grid_array(3, counter_1) = grid_array(5, counter_1) + flux
               Else
                  If grid_array(3, counter_1) < flux Then
                     grid_array(3, counter_1) = grid_array(5, counter_1) + flux
                  End If
               End If
            End If
            counter_1 = counter_1 + 1
         Next j
         counter_1 = counter_1 + (grid_X_max - X_length)
      Next i

   Next ventloop
End If
End If

'*************************************************************************
'******   copy coord and sample-space files to simulation directory   *******
'*************************************************************************

If Opt_vents(0).Value = True Then
   Open (sim_dir & "\vent coordinates.txt") For Output As #1
   Write #1, CInt(Txt_singlevent_X.Text), CInt(Txt_singlevent_Y.Text)
   Close #1
ElseIf Opt_vents(1).Value = True Then
   fso.CopyFile vent_info_file, sim_dir & "\", True
End If

If Opt_sample_spacing(0).Value = True Then
   Open (sim_dir & "\sample spacings.txt") For Output As #1
   Write #1, CInt(Txt_singlespacing.Text)
   Close #1
ElseIf Opt_sample_spacing(1).Value = True Then
```

```
   fso.CopyFile sample_spacing_file, sim_dir & "\", True
End If

If Opt_fixed_num_samples(0).Value = True Then
   Open (sim_dir & "\fixed_num_sample_spacing.txt") For Output As #1
   Write #1, CInt(Txt_single_fixed_num_samples.Text)
   Close #1
ElseIf Opt_fixed_num_samples(1).Value = True Then
   fso.CopyFile fixed_num_sample_file, sim_dir & "\", True
End If

'*************************************************************
'*****************     output "true grid"     ***************
'*************************************************************

If Opt_data_type(0).Value = True Then
   Txt_message.Text = "output complete background plus vent data"
   PauseIt
   BkgPlusVentFile = sim_dir & "\" & "background plus vent data.txt"
ElseIf Opt_data_type(1).Value = True Then
   Txt_message.Text = "output complete real field data"
   PauseIt
   BkgPlusVentFile = sim_dir & "\" & "real field data.txt"
End If

Open BkgPlusVentFile For Output As #1
For k = 1 To total_num_pts
   Write #1, grid_array(1, k), grid_array(2, k), grid_array(3, k)
   grid_check_array(1, k) = k      ' populate array used to select random sample points;
                        'this array is condensed when needed to decrease the random selection of points that no longer have
values
Next k
Close #1

'*********************************************************************
'**** Grid BG+vent or RealData using Surfer to calculate *************
'**** total flux using all data for simulated BG+vent ****************
'**** OR all RealData. An estimate of "true" vent leakage ************
'**** is calculated for either by subtracting BG from total flxu *****
'*********************************************************************

If Opt_Sim_Type(1).Value = True Then 'if calculate flux and NOT find anomaly
   If Opt_data_type(0).Value = True Then      'simulated data
      BkgPlusVentGridFile = sim_dir & "\BkgPlusVentGrid.grd"
   ElseIf Opt_data_type(1).Value = True Then      'real data
      BkgPlusVentGridFile = sim_dir & "\RealFieldData.grd"
   End If
   Txt_message.Text = "grid total dataset"
   PauseIt

   Call SurferGrid(BkgPlusVentFile, ContourType, CInt(Txt_AnisotropyRatio.Text), CInt(Txt_AnisotropyAngle.Text),
BkgPlusVentGridFile, grid_X_max, grid_Y_max, 2, TotalFlux)
   TotalFluxVents = TotalFlux - TotalFluxBackground
End If
```

```
'*********************************************************************
'**********     create arrays for exporting summary tables     *****
'*********************************************************************
If Opt_sampling_strategy(0).Value = True Then
   ReDim TotalFlux_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim TotalVentFlux_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim R_RootMeanSquare_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim R2_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim ZE_Correlation_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim ZR_Correlation_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim ZE_RankCorrelation_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
   ReDim ZR_RankCorrelation_Array(1 To fixed_num_samples, 1 To num_simulations) As Variant
Else
   ReDim TotalFlux_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim TotalVentFlux_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim R_RootMeanSquare_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim R2_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim ZE_Correlation_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim ZR_Correlation_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim ZE_RankCorrelation_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
   ReDim ZR_RankCorrelation_Array(1 To num_sample_spacings, 1 To num_simulations) As Variant
End If

Txt_sim2.Text = num_simulations
Txt_ss2.Text = num_sample_spacings
PauseIt

'++++ put headers for summary file +++++++++
SummaryFile = sim_dir & "\" & "simulation summary.txt"
Open SummaryFile For Output As #5
   Write #5, "simulation number", "number of samples", "sample spacing", "sampling density (samp/km2)", "anomalous samples", "Total Flux (all data)", "Total Flux background (all data)", "Total Vent Flux (all data)", "Simulated Total Flux (sub-sampled data)", "Simulated Vent Flux  (sub-sampled data)", "Estimation Error Statistics - Root Mean square", "Spatial Regression Statistics for Estimation Error - Error Coefficient of multiple determination", "Inter-Variable Correlations - correlation between Z and estimated statistics", "Inter-Variable Correlations - correlation between Z and estimation error", "Rank correlation between Z and E", "Rank correlation between Z and R"
Close #5

'*********************************************************************
'*****************     sub-sampling simulations     ***************
'*********************************************************************

If Opt_sampling_strategy(0) = True Then 'purely random point selection based on a fixed number of points
   Txt_ss2.Text = fixed_num_samples
   For y = 1 To fixed_num_samples
      num_of_fixed_samples = fixed_num_sample_array(1, y)
      ' Create a new sub-folder for each fixed number of samples for this simulation
      ss_dir = (sim_dir & "\SN_" & CStr(num_of_fixed_samples) & "\")
      fso.CreateFolder (ss_dir)

      For s = 1 To num_simulations
         Txt_message.Text = "simulations"
         Txt_ss1.Text = y
         Txt_sim1.Text = s
         PauseIt
         SampleDataFile = ss_dir & "sampledata" & CStr(s) & ".txt"
```

```
        ExtraPoints = ((2 * (CInt(grid_X_max / 20) + 1)) + (2 * (CInt(grid_Y_max / 20) + 1)))
        ReDim sample_array(3, num_of_fixed_samples + ExtraPoints) 'formula to put boundary pts around 4 sides
        num_samples = 0
        anomalies = 0
        Open SampleDataFile For Output As #3

        For count_samples = 1 To num_of_fixed_samples
           sample_pt = grid_array(4, ((UBound(grid_array, 2) - 1) * Rnd + 1))
           If sample_pt = "" Then
              Do Until sample_pt <> ""
                 sample_pt = grid_array(4, grid_array(1, Int((UBound(grid_array, 2) - 1) * Rnd + 1)))
              Loop
           End If
           sample_array(1, count_samples) = grid_array(1, sample_pt) 'X
           sample_array(2, count_samples) = grid_array(2, sample_pt) 'Y
           sample_array(3, count_samples) = grid_array(3, sample_pt) 'flux
           Write #3, sample_array(1, count_samples), sample_array(2, count_samples), sample_array(3, count_samples)
           If sample_array(3, count_samples) > bkg_max Then anomalies = anomalies + 1
        Next count_samples

'***NEW***  'adds a value of 20 g/m2/d at 50m intervals around edge to define boundary value
        'to avoid problem of contouring beyond the boundary
        For k = 0 To grid_X_max Step CInt(grid_X_max / 20)
           Write #3, k, 0, avg_bkg_flux
           Write #3, k, (grid_Y_max + 1), avg_bkg_flux
        Next k
        For k = 0 To grid_Y_max Step CInt(grid_Y_max / 20)
           Write #3, 0, k, avg_bkg_flux
           Write #3, (grid_X_max + 1), k, avg_bkg_flux
        Next k
'***NEW***
        Close #3

        sampling_density = CLng(num_of_fixed_samples / ((grid_X_max * grid_Y_max) / 1000000))

        '+++++++++ IF stmt for "find anomaly" vs "calculate flux". For "find anomaly", Surfer is bypassed
        If Opt_Sim_Type(0) = True Then
           '++++++++++ output results for simulation
           Open SummaryFile For Append As #5
              Write #5, s, num_of_fixed_samples, sample_spacing, sampling_density, anomalies
           Close #5
        ElseIf Opt_Sim_Type(1) = True Then
           '+++++++grid sub-sampled data for each simulation
           SampleGridFile = ss_dir & "samplegrid" & CStr(s) & ".grd"
           Txt_message.Text = "grid simulated data - simulation " & CStr(s)
           PauseIt
           Call SurferGrid2(SampleDataFile, ContourType, CInt(Txt_AnisotropyRatio.Text),
CInt(Txt_AnisotropyAngle.Text), SampleGridFile, grid_X_max, grid_Y_max, 5, TotalFluxSim, CV)
           TotalFluxVentsSim = TotalFluxSim - TotalFluxBackground
           '++++++++++ output results for simulation
           Open SummaryFile For Append As #5
              Write #5, s, num_samples, sample_spacing, sampling_density, anomalies, TotalFlux, TotalFluxBackground,
TotalFluxVents, TotalFluxSim, TotalFluxVentsSim, CV(1, 1), CV(1, 2), CV(1, 3), CV(1, 4), CV(1, 5), CV(1, 6)
           Close #5
           '+++++++put various parameters in arrays for later export into summary files
           TotalFlux_Array(y, s) = TotalFluxSim 'total flux of sub-sampled data
```

```
            TotalVentFlux_Array(y, s) = TotalFluxVentsSim 'total flux (minus Background) of sub-sampled data
            R_RootMeanSquare_Array(y, s) = CV(1, 1)
            R2_Array(y, s) = CV(1, 2)
            ZE_Correlation_Array(y, s) = CV(1, 3)
            ZR_Correlation_Array(y, s) = CV(1, 4)
            ZE_RankCorrelation_Array(y, s) = CV(1, 5)
            ZR_RankCorrelation_Array(y, s) = CV(1, 6)
          End If
        Next s
      Next y
    ElseIf Opt_sampling_strategy(1) = True Then 'distributed random sampling based on a minimum sampling distance
      SampleSpacingTitle = "simulation number"
      For y = 1 To num_sample_spacings

        sample_spacing = sample_spacing_array(1, y)
        SampleSpacingTitle = SampleSpacingTitle & "," & CStr(sample_spacing)

        ' Create a new sub-folder for each sampling spacing for this simulation
        ss_dir = (sim_dir & "\SS_" & CStr(sample_spacing) & "\")
        fso.CreateFolder (ss_dir)

        For s = 1 To num_simulations

          Txt_message.Text = "simulations"
          Txt_ss1.Text = y
          Txt_sim1.Text = s
          PauseIt

          SampleDataFile = ss_dir & "sampledata" & CStr(s) & ".txt"
          remaining_pts = total_num_pts
          ExtraPoints = ((2 * (CInt(grid_X_max / 20) + 1)) + (2 * (CInt(grid_Y_max / 20) + 1)))
          ReDim sample_array(3, total_num_pts + ExtraPoints) 'formula to put boundary pts around 4 sides
          num_samples = 0

          Do Until remaining_pts < 2
            'select random point and test if that point in the array has a value or has already been sampled/removed (ie "")
            count_loop = 1
            sample_pt = grid_array(4, grid_check_array(1, Int((UBound(grid_check_array, 2) - 1) * Rnd + 1)))
            If sample_pt = "" Then
              Do Until sample_pt <> ""
                sample_pt = grid_array(4, grid_check_array(1, Int((UBound(grid_check_array, 2) - 1) * Rnd + 1)))
                count_loop = count_loop + 1
              Loop
            End If

            'if the previous loop loops too many times to look for a value, the check array is condensed
            grid_check_array_counter = 0
            If count_loop > 5000 Then
              For grid_array_counter = 1 To total_num_pts
                If grid_array(4, grid_array_counter) <> "" Then
                  grid_check_array_counter = grid_check_array_counter + 1
                  grid_check_array(1, grid_check_array_counter) = grid_array(4, grid_array_counter)
                End If
              Next
            ReDim Preserve grid_check_array(1, grid_check_array_counter)
            End If
```

```
            num_samples = num_samples + 1

            'put selected sample in array of sampled points
            sample_array(1, num_samples) = grid_array(1, sample_pt) 'X
            sample_array(2, num_samples) = grid_array(2, sample_pt) 'Y
            sample_array(3, num_samples) = grid_array(3, sample_pt) 'flux

            sample_pt_X = grid_array(1, sample_pt)
            sample_pt_Y = grid_array(2, sample_pt)

            'define block to be erased around the sample point, based on sample spacing
            'and on whether the point is near a border of the grid
            circle_X_min = sample_pt_X - sample_spacing
            If circle_X_min < 1 Then circle_X_min = 1
            circle_X_max = sample_pt_X + sample_spacing
            If circle_X_max > grid_X_max Then circle_X_max = grid_X_max
            circle_Y_min = sample_pt_Y - sample_spacing
            If circle_Y_min < 1 Then circle_Y_min = 1
            circle_Y_max = sample_pt_Y + sample_spacing
            If circle_Y_max > grid_Y_max Then circle_Y_max = grid_Y_max

            X_length = (circle_X_max - circle_X_min) + 1
            Y_length = (circle_Y_max - circle_Y_min) + 1

            counter_1 = ((circle_Y_min - 1) * grid_X_max) + circle_X_min

            'cancel all points in a radius (ie sample distance) around the sampled point
            For l = 1 To Y_length
               For m = 1 To X_length
                  distance = (((sample_pt_X - grid_array(1, counter_1)) ^ 2) + ((sample_pt_Y - grid_array(2, counter_1)) ^
2)) ^ 0.5
                  If distance <= sample_spacing Then
                     If grid_array(4, counter_1) <> "" Then
                        grid_array(4, counter_1) = ""
                        remaining_pts = remaining_pts - 1
                     End If
                  End If
                  counter_1 = counter_1 + 1
               Next m
               counter_1 = counter_1 + (grid_X_max - X_length)
            Next l
         Loop

         anomalies = 0
         Open SampleDataFile For Output As #3
         For o = 1 To num_samples
            Write #3, sample_array(1, o), sample_array(2, o), sample_array(3, o)
            If sample_array(3, o) > bkg_max Then anomalies = anomalies + 1
         Next o

'***adds a value of 20 g/m2/d at 50m intervals around edge to define boundary value
         'to avoid problem of contouring beyond the boundary
         For k = 0 To grid_X_max Step CInt(grid_X_max / 20)
            Write #3, k, 0, avg_bkg_flux
            Write #3, k, (grid_Y_max + 1), avg_bkg_flux
```

```
        Next k
        For k = 0 To grid_Y_max Step CInt(grid_Y_max / 20)
          Write #3, 0, k, avg_bkg_flux
          Write #3, (grid_X_max + 1), k, avg_bkg_flux
        Next k
'***

        Close #3

        sampling_density = CLng(num_samples / ((grid_X_max * grid_Y_max) / 1000000))

        'get grid_array ready for next simulation by resetting sampled/not sampled control column
        'get grid_check_array ready for next simulation by resetting values to sequential numbers
        ReDim grid_check_array(1, total_num_pts)
        For p = 1 To total_num_pts
          grid_array(4, p) = p
          grid_check_array(1, p) = p
        Next p

        '+++++++++ IF stmt for "find anomaly" vs "calculate flux". For "find anomaly", Surfer is bypassed
        If Opt_Sim_Type(0) = True Then
          '++++++++++ output results for simulation
          Open SummaryFile For Append As #5
            Write #5, s, num_samples, sample_spacing, sampling_density, anomalies
          Close #5
        ElseIf Opt_Sim_Type(1) = True Then
          '+++++++grid sub-sampled data for each simulation
          SampleGridFile = ss_dir & "samplegrid" & CStr(s) & ".grd"
          Txt_message.Text = "grid simulated data - simulation " & CStr(s)
          PauseIt
          Call SurferGrid2(SampleDataFile, ContourType, CInt(Txt_AnisotropyRatio.Text),
CInt(Txt_AnisotropyAngle.Text), SampleGridFile, grid_X_max, grid_Y_max, 5, TotalFluxSim, CV)
          TotalFluxVentsSim = TotalFluxSim - TotalFluxBackground
          '++++++++++ output results for simulation
          Open SummaryFile For Append As #5
            Write #5, s, num_samples, sample_spacing, sampling_density, anomalies, TotalFlux, TotalFluxBackground,
TotalFluxVents, TotalFluxSim, TotalFluxVentsSim, CV(1, 1), CV(1, 2), CV(1, 3), CV(1, 4), CV(1, 5), CV(1, 6)
          Close #5
          '+++++++put various parameters in arrays for later export into summary files
          TotalFlux_Array(y, s) = TotalFluxSim 'total flux of sub-sampled data
          TotalVentFlux_Array(y, s) = TotalFluxVentsSim 'total flux (minus Background) of sub-sampled data
          R_RootMeanSquare_Array(y, s) = CV(1, 1)
          R2_Array(y, s) = CV(1, 2)
          ZE_Correlation_Array(y, s) = CV(1, 3)
          ZR_Correlation_Array(y, s) = CV(1, 4)
          ZE_RankCorrelation_Array(y, s) = CV(1, 5)
          ZR_RankCorrelation_Array(y, s) = CV(1, 6)
        End If
      Next s
    Next y
ElseIf Opt_sampling_strategy(2) = True Then 'gridded sampling
    SampleSpacingTitle = "simulation number"
    For y = 1 To num_sample_spacings
      sample_spacing = sample_spacing_array(1, y)
      SampleSpacingTitle = SampleSpacingTitle & "," & CStr(sample_spacing)
      ' Create a new sub-folder for each sampling spacing for this simulation
```

```
ss_dir = (sim_dir & "\SG_" & CStr(sample_spacing) & "\")
fso.CreateFolder (ss_dir)

'create array from which to select first point_to_sample
Dim Temp_file As String
Temp_file = ss_dir & "start_pt_block_spacing_" & sample_spacing & ".txt"
Open Temp_file For Append As #5
ReDim grid_check_array(1, (CLng(sample_spacing)) ^ 2)

counter_1 = 1
counter_2 = 1
For o = 1 To sample_spacing
   For p = 1 To sample_spacing
      grid_check_array(1, counter_2) = counter_1
      Write #5, counter_2, counter_1, grid_array(1, counter_1), grid_array(2, counter_1)
      counter_1 = counter_1 + 1
      counter_2 = counter_2 + 1
   Next p
   counter_1 = counter_1 + (grid_X_max - sample_spacing)
Next o
Close #5

If CInt(Txt_num_simulations.Text) > (sample_spacing ^ 2) Then
   num_simulations = (sample_spacing ^ 2)
Else
   num_simulations = CInt(Txt_num_simulations.Text)
End If

For s = 1 To num_simulations
   Txt_message.Text = "simulations"
   Txt_ss1.Text = y
   Txt_sim1.Text = s
   Txt_sim2.Text = num_simulations
   Txt_ss2.Text = num_sample_spacings
   PauseIt

   'select random point and test if that point in the array has a value or has already been sampled/removed (ie "")
   count_loop = 1
   point_to_sample = ""
   If point_to_sample = "" Then
      Do Until point_to_sample <> "" Or count_loop > 1000
         random_number = Int((UBound(grid_check_array, 2) - 1) * Rnd + 1)
         point_to_sample = grid_check_array(1, random_number)
         If point_to_sample <> "" Then grid_check_array(1, random_number) = ""
         count_loop = count_loop + 1
      Loop
   End If

   If count_loop > 1000 Then
      count_loop = 1
      Do Until point_to_sample <> "" Or count_loop > (sample_spacing ^ 2)
         point_to_sample = grid_check_array(1, count_loop)
         If point_to_sample <> "" Then grid_check_array(1, count_loop) = ""
         count_loop = count_loop + 1
      Loop
   End If
```

```
Dim Temp_file_2 As String
Temp_file_2 = ss_dir & "start_pts_" & sample_spacing & ".txt"
Open Temp_file_2 For Append As #5
   Write #5, grid_array(1, point_to_sample), grid_array(2, point_to_sample)
Close #5

starting_point = point_to_sample
point_to_sample = point_to_sample - sample_spacing

SampleDataFile = ss_dir & "sampledata" & CStr(s) & ".txt"
remaining_pts = total_num_pts

X_samples = Int(grid_X_max / sample_spacing)
Y_samples = Int(grid_Y_max / sample_spacing)
If (grid_array(1, starting_point) + (X_samples * sample_spacing)) < grid_X_max Then
   X_samples = X_samples + 1
End If
If (grid_array(2, starting_point) + (Y_samples * sample_spacing)) < grid_Y_max Then
   Y_samples = Y_samples + 1
End If
ExtraPoints = ((2 * (CInt(grid_X_max / 20) + 1)) + (2 * (CInt(grid_Y_max / 20) + 1)))
ReDim sample_array(3, (X_samples * Y_samples) + ExtraPoints) 'formula to put boundary pts around 4 sides
num_samples = 0
OddEven = 2
For Count_Row = 1 To Y_samples
   For Count_Column = 1 To X_samples
      point_to_sample = point_to_sample + sample_spacing
      num_samples = num_samples + 1
      'put selected sample in array of sampled points
      sample_array(1, num_samples) = grid_array(1, point_to_sample) 'X
      sample_array(2, num_samples) = grid_array(2, point_to_sample) 'Y
      sample_array(3, num_samples) = grid_array(3, point_to_sample) 'flux
   Next Count_Column
   point_to_sample = (starting_point + (Count_Row * sample_spacing * grid_X_max)) - sample_spacing

   'shifts every second row by 1/2 sample spacing for offset grid option
   If chk_grid_offset.Value = 1 Then
      If OddEven = 2 Then
         If grid_array(1, starting_point) > (sample_spacing / 2) Then
            point_to_sample = point_to_sample - (sample_spacing / 2)
         ElseIf grid_array(1, starting_point) <= (sample_spacing / 2) Then
            point_to_sample = point_to_sample + (sample_spacing / 2)
         End If
         OddEven = 1
      ElseIf OddEven = 1 Then
         OddEven = 2
      End If
   End If

Next Count_Row

anomalies = 0
Open SampleDataFile For Output As #3
For o = 1 To num_samples
   Write #3, sample_array(1, o), sample_array(2, o), sample_array(3, o)
```

```
            If sample_array(3, o) > bkg_max Then anomalies = anomalies + 1
        Next o

'***adds a value of 20 g/m2/d at 50m intervals around edge to define boundary value
        'to avoid problem of contouring beyond the boundary
        For k = 0 To grid_X_max Step CInt(grid_X_max / 20)
            Write #3, k, 0, avg_bkg_flux
            Write #3, k, (grid_Y_max + 1), avg_bkg_flux
        Next k
        For k = 0 To grid_Y_max Step CInt(grid_Y_max / 20)
            Write #3, 0, k, avg_bkg_flux
            Write #3, (grid_X_max + 1), k, avg_bkg_flux
        Next k
'***

        Close #3

        sampling_density = CLng(num_samples / ((grid_X_max * grid_Y_max) / 1000000))

        '+++++++++ IF stmt for "find anomaly" vs "calculate flux". For "find anomaly", Surfer is bypassed
        If Opt_Sim_Type(0) = True Then
            '++++++++++ output results for simulation
            Open SummaryFile For Append As #5
                Write #5, s, num_samples, sample_spacing, sampling_density, anomalies
            Close #5
        ElseIf Opt_Sim_Type(1) = True Then
            '+++++++grid sub-sampled data for each simulation
            SampleGridFile = ss_dir & "samplegrid" & CStr(s) & ".grd"
            Txt_message.Text = "grid simulated data - simulation " & CStr(s)
            PauseIt
            Call SurferGrid2(SampleDataFile, ContourType, CInt(Txt_AnisotropyRatio.Text),
CInt(Txt_AnisotropyAngle.Text), SampleGridFile, grid_X_max, grid_Y_max, 5, TotalFluxSim, CV)
            TotalFluxVentsSim = TotalFluxSim - TotalFluxBackground
            '++++++++++ output results for simulation
            Open SummaryFile For Append As #5
                If Opt_sampling_strategy(0) = True Then 'random
                    Write #5, s, num_of_fixed_samples, sample_spacing, sampling_density, anomalies, TotalFlux,
TotalFluxBackground, TotalFluxVents, TotalFluxSim, TotalFluxVentsSim, CV(1, 1), CV(1, 2), CV(1, 3), CV(1, 4),
CV(1, 5), CV(1, 6)
                ElseIf Opt_sampling_strategy(1) = True Then 'distributed random
                    Write #5, s, num_samples, sample_spacing, sampling_density, anomalies, TotalFlux,
TotalFluxBackground, TotalFluxVents, TotalFluxSim, TotalFluxVentsSim, CV(1, 1), CV(1, 2), CV(1, 3), CV(1, 4),
CV(1, 5), CV(1, 6)
                ElseIf Opt_sampling_strategy(2) = True Then ' systematic random (grid)
                    Write #5, s, num_samples, sample_spacing, sampling_density, anomalies, TotalFlux,
TotalFluxBackground, TotalFluxVents, TotalFluxSim, TotalFluxVentsSim, CV(1, 1), CV(1, 2), CV(1, 3), CV(1, 4),
CV(1, 5), CV(1, 6)
                End If
            Close #5
            '+++++++put various parameters in arrays for later export into summary files
            TotalFlux_Array(y, s) = TotalFluxSim 'total flux of sub-sampled data
            TotalVentFlux_Array(y, s) = TotalFluxVentsSim 'total flux (minus Background) of sub-sampled data
            R_RootMeanSquare_Array(y, s) = CV(1, 1)
            R2_Array(y, s) = CV(1, 2)
            ZE_Correlation_Array(y, s) = CV(1, 3)
            ZR_Correlation_Array(y, s) = CV(1, 4)
```

```
            ZE_RankCorrelation_Array(y, s) = CV(1, 5)
            ZR_RankCorrelation_Array(y, s) = CV(1, 6)
          End If
        Next s
    Next y
End If


'*********************************************************************
'*****************     export summary files        ***************
'*********************************************************************

If Opt_Sim_Type(1) = True Then

    Dim f As Integer
    Dim g As Integer
    Dim e As Integer
    Dim EachLine As String
    Dim FileName As String

    For e = 1 To 8
      If e = 1 Then FileName = sim_dir & "\TotalFluxSummary.txt"
      If e = 2 Then FileName = sim_dir & "\TotalVentFluxSummary.txt"
      If e = 3 Then FileName = sim_dir & "\R_RootMeanSquareSummary.txt"
      If e = 4 Then FileName = sim_dir & "\R2Summary.txt"
      If e = 5 Then FileName = sim_dir & "\ZE_CorrelationSummary.txt"
      If e = 6 Then FileName = sim_dir & "\ZR_CorrelationSummary.txt"
      If e = 7 Then FileName = sim_dir & "\ZE_RankCorrelationSummary.txt"
      If e = 8 Then FileName = sim_dir & "\ZR_RankCorrelationSummary.txt"
      Open FileName For Append As #6
      Print #6, SampleSpacingTitle
        For g = 1 To num_simulations
          EachLine = g
          For f = 1 To num_sample_spacings
            If e = 1 Then EachLine = EachLine & "," & TotalFlux_Array(f, g)
            If e = 2 Then EachLine = EachLine & "," & TotalVentFlux_Array(f, g)
            If e = 3 Then EachLine = EachLine & "," & R_RootMeanSquare_Array(f, g)
            If e = 4 Then EachLine = EachLine & "," & R2_Array(f, g)
            If e = 5 Then EachLine = EachLine & "," & ZE_Correlation_Array(f, g)
            If e = 6 Then EachLine = EachLine & "," & ZR_Correlation_Array(f, g)
            If e = 7 Then EachLine = EachLine & "," & ZE_RankCorrelation_Array(f, g)
            If e = 8 Then EachLine = EachLine & "," & ZR_RankCorrelation_Array(f, g)
          Next f
          Print #6, EachLine
        Next g
      Close #6
    Next e
End If
MsgBox "start time is " & StartTime & "; end time is " & Time()
End

Frm_initial_settings.Enabled = True
Frm_data_type.Enabled = False
Frm_bkg_flux.Enabled = False
Frm_vents.Enabled = False
Frm_sampling_settings.Enabled = False
```

```
Frm_contour.Enabled = False
Frm_run.Enabled = False

End Sub
Private Sub PauseIt()
Dim PauseTime, Start
PauseTime = 0.01   ' Set duration.
Start = Timer   ' Set start time.
Do While Timer < Start + PauseTime
   DoEvents   ' Yield to other processes.
Loop
End Sub

Private Sub Cmd_activate_Click()
Txt_path.Enabled = True
Frm_data_type.Enabled = False
Frm_bkg_flux.Enabled = False
Frm_vents.Enabled = False
Frm_sampling_settings.Enabled = False
Frm_contour.Enabled = False
Frm_run.Enabled = False
Cmd_confirm_path.Enabled = True

End Sub

Private Sub Cmd_confirm_path_Click()
Txt_root_length.Text = Len(Txt_path.Text)
path = Txt_path.Text
Txt_path.Enabled = False
Frm_data_type.Enabled = True
Frm_bkg_flux.Enabled = True
Frm_vents.Enabled = True
Frm_sampling_settings.Enabled = True
Frm_contour.Enabled = True
Frm_run.Enabled = True
Frm_Sim_Type.Enabled = True

End Sub

Private Sub Form_Load()
Cbo_RBType.AddItem "Inverse MultiQuadric"
Cbo_RBType.AddItem "MultiLog"
Cbo_RBType.AddItem "MultiQuadric"
Cbo_RBType.AddItem "Natural cubic spline"
Cbo_RBType.AddItem "Thin plate spline"

Cbo_RBType.Text = "MultiQuadric"

Opt_contour.Item(0).Value = True
End Sub

Private Sub Opt_contour_Click(Index As Integer)
If Opt_contour(0).Value = True Then
   Txt_IDPower.Enabled = False
   Txt_IDSmoothing.Enabled = False
   Cbo_RBType.Enabled = False
```

```
    Txt_RBRSquared.Enabled = False
ElseIf Opt_contour(1).Value = True Then
    Txt_IDPower.Enabled = False
    Txt_IDSmoothing.Enabled = False
    Cbo_RBType.Enabled = False
    Txt_RBRSquared.Enabled = False
ElseIf Opt_contour(2).Value = True Then
    Txt_IDPower.Enabled = True
    Txt_IDSmoothing.Enabled = True
    Cbo_RBType.Enabled = False
    Txt_RBRSquared.Enabled = False
ElseIf Opt_contour(3).Value = True Then
    Txt_IDPower.Enabled = False
    Txt_IDSmoothing.Enabled = False
    Cbo_RBType.Enabled = True
    Txt_RBRSquared.Enabled = True
End If
End Sub

Private Sub Opt_data_type_Click(Index As Integer)
If Opt_data_type(0).Value = True Then      'data simulated by program
    Frm_vents.Enabled = True
    Txt_singlevent_X.Enabled = False
    Txt_singlevent_Y.Enabled = False
    Txt_MaxVentFlux.Enabled = False
    Txt_core_radius.Enabled = False
    Txt_vent_radius.Enabled = False
    Txt_vent_info_file.Enabled = False
    Opt_vents(0).Value = False
    Opt_vents(1).Value = False
    Txt_X_max.Enabled = True
    Txt_Y_max.Enabled = True
    Txt_field_data_file.Text = ""
    Txt_field_data_file.Enabled = False
    Opt_BG(0).Enabled = False
    Opt_BG(1).Enabled = True
    Opt_BG(2).Enabled = True

ElseIf Opt_data_type(1).Value = True Then  'real data imported from field data
    Frm_vents.Enabled = False
    Txt_singlevent_X.Text = ""
    Txt_singlevent_Y.Text = ""
    Txt_MaxVentFlux.Text = ""
    Txt_core_radius.Text = ""
    Txt_vent_radius.Text = ""
    Opt_vents(0).Value = False
    Opt_vents(1).Value = False
    Txt_vent_info_file.Text = ""
    Txt_X_max.Enabled = False
    Txt_Y_max.Enabled = False
    Txt_field_data_file.Enabled = False
    Opt_BG(0).Value = True
    Opt_BG(0).Enabled = True
    Opt_BG(1).Enabled = False
    Opt_BG(2).Enabled = False
```

```
    'select file (ie path and file name) for real data input
    '++++++++ file name format - e.g. abcdef_x1234y1234.dat ++++++++++
    Dim fso As New FileSystemObject ' Get instance of FileSystemObject.
    CommonDialog1.FileName = ""
    CommonDialog1.InitDir = path & "field_data\"
    CommonDialog1.DialogTitle = "Select real data file"
    CommonDialog1.ShowOpen
    Txt_X_max.Text = CInt(Mid(CommonDialog1.FileTitle, 9, 4))
    Txt_Y_max.Text = CInt(Mid(CommonDialog1.FileTitle, 14, 4))
    Txt_field_data_file.Text = CommonDialog1.FileName
    Txt_num_grid_pts.Text = Txt_X_max.Text * Txt_Y_max.Text

End If
End Sub

Private Sub Opt_fixed_num_samples_Click(Index As Integer)
 If Opt_fixed_num_samples(0).Value = True Then
    Txt_single_fixed_num_samples.Enabled = True
    Txt_single_fixed_num_samples.Text = 100
    Txt_fixed_num_samples.Text = 1
    Txt_fixed_sample_number_file.Text = ""
    Txt_fixed_sample_number_file.Enabled = False
 ElseIf Opt_fixed_num_samples(1).Value = True Then
    Txt_fixed_num_samples.Text = ""
    Txt_single_fixed_num_samples.Enabled = False
    Txt_single_fixed_num_samples.Text = ""
    Dim fso As New FileSystemObject ' Get instance of FileSystemObject.
    CommonDialog1.FileName = ""
    CommonDialog1.InitDir = path
    CommonDialog1.DialogTitle = "Select sample spacing file"
    CommonDialog1.FileName = "fixed number of samples.txt"
    CommonDialog1.ShowOpen
    Txt_fixed_sample_number_file.Text = CommonDialog1.FileName
    Txt_fixed_sample_number_file.Enabled = False

    ' count number of sample spacings
    Dim num_fixed_sample_numbers As Integer
    Dim DataLine As String
    Open CommonDialog1.FileName For Input As #1
    num_fixed_sample_numbers = 0
    Do Until EOF(1)
       Line Input #1, DataLine
       num_fixed_sample_numbers = num_fixed_sample_numbers + 1
    Loop
    Close #1
    Txt_fixed_num_samples.Text = num_fixed_sample_numbers
 End If

End Sub

Private Sub Opt_sample_spacing_Click(Index As Integer)
If Opt_sample_spacing(0).Value = True Then   'single sample spacing
    Txt_sample_spacing_file.Enabled = False
    Txt_sample_spacing_file.Text = ""
    Txt_singlespacing.Enabled = True
    Txt_singlespacing.Text = 50
```

```
    Txt_num_sample_space.Text = 1
ElseIf Opt_sample_spacing(1).Value = True Then   'import file with multiple sample spacings
    Txt_sample_spacing_file.Enabled = False
    Txt_singlespacing.Enabled = False
    Txt_singlespacing.Text = ""

    'select file (ie path and file name) for sample spacing list
    '++++++++ file name - "sample spacings.txt" ++++++++++
    Dim fso As New FileSystemObject ' Get instance of FileSystemObject.
    CommonDialog1.FileName = ""
    CommonDialog1.InitDir = path
    CommonDialog1.DialogTitle = "Select sample spacing file"
    CommonDialog1.FileName = "sample spacings.txt"
    CommonDialog1.ShowOpen
    Txt_sample_spacing_file.Text = CommonDialog1.FileName
    Txt_sample_spacing_file.Enabled = False

    ' count number of sample spacings
    Dim num_sample_spacings As Integer
    Dim DataLine As String
    Open CommonDialog1.FileName For Input As #1
    num_sample_spacings = 0
    Do Until EOF(1)
       Line Input #1, DataLine
       num_sample_spacings = num_sample_spacings + 1
    Loop
    Close #1
    Txt_num_sample_space.Text = num_sample_spacings


End If
End Sub

Private Sub Opt_sampling_strategy_Click(Index As Integer)
If Opt_sampling_strategy(0).Value = True Then   'random
    Opt_sample_spacing(0).Enabled = False
    Opt_sample_spacing(1).Enabled = False
    Opt_fixed_num_samples(0).Enabled = True
    Opt_fixed_num_samples(1).Enabled = True
    Txt_singlespacing.Text = ""
    Txt_num_sample_space.Text = ""
    Txt_sample_spacing_file.Text = ""
ElseIf Opt_sampling_strategy(1).Value = True Then   'systematic random
    Opt_sample_spacing(0).Enabled = True
    Opt_sample_spacing(1).Enabled = True
    Opt_fixed_num_samples(0).Enabled = False
    Opt_fixed_num_samples(1).Enabled = False
    Txt_fixed_sample_number_file.Text = ""
    Txt_fixed_sample_number_file.Enabled = False
    Txt_single_fixed_num_samples.Text = ""
    Txt_single_fixed_num_samples.Enabled = False
    Txt_fixed_num_samples.Text = ""
    Txt_fixed_num_samples.Enabled = False
    Opt_sample_spacing(0).Value = False
    Opt_sample_spacing(1).Value = False
```

```
ElseIf Opt_sampling_strategy(2).Value = True Then   'regular grid
   Opt_sample_spacing(0).Enabled = True
   Opt_sample_spacing(1).Enabled = True
   Opt_fixed_num_samples(0).Enabled = False
   Opt_fixed_num_samples(1).Enabled = False
   Txt_fixed_sample_number_file.Text = ""
   Txt_fixed_sample_number_file.Enabled = False
   Txt_single_fixed_num_samples.Text = ""
   Txt_single_fixed_num_samples.Enabled = False
   Txt_fixed_num_samples.Text = ""
   Txt_fixed_num_samples.Enabled = False
   Opt_sample_spacing(0).Value = False
   Opt_sample_spacing(1).Value = False
   chk_grid_offset.Enabled = True
End If
End Sub

Private Sub Opt_vents_Click(Index As Integer)
If Opt_vents(0).Value = True Then    'single vent location and strength
   Txt_vent_info_file.Enabled = False
   Txt_singlevent_X.Enabled = True
   Txt_singlevent_Y.Enabled = True
   Txt_MaxVentFlux.Enabled = True
   Txt_core_radius.Enabled = True
   Txt_vent_radius.Enabled = True
   Txt_singlevent_X.Text = 50
   Txt_singlevent_Y.Text = 50
   Txt_MaxVentFlux.Text = 1000
   Txt_core_radius.Text = 5
   Txt_vent_radius.Text = 20
   Txt_vent_info_file.Text = ""
   Txt_num_of_vents.Text = 1

ElseIf Opt_vents(1).Value = True Then   'import file with multiple vents
   Txt_vent_info_file.Enabled = True
   Txt_singlevent_X.Enabled = False
   Txt_singlevent_Y.Enabled = False
   Txt_MaxVentFlux.Enabled = False
   Txt_core_radius.Enabled = False
   Txt_vent_radius.Enabled = False
   Txt_singlevent_X.Text = ""
   Txt_singlevent_Y.Text = ""
   Txt_MaxVentFlux.Text = ""
   Txt_core_radius.Text = ""
   Txt_vent_radius.Text = ""

   'select file (ie path and file name) for vent locations and strength
   '++++++++ file name - "vent coordinates.txt" ++++++++++
   Dim fso As New FileSystemObject ' Get instance of FileSystemObject.
   CommonDialog1.FileName = ""
   CommonDialog1.InitDir = path
   CommonDialog1.DialogTitle = "Select vent coordinate file"
   CommonDialog1.ShowOpen
   Txt_vent_info_file.Text = CommonDialog1.FileName
   Txt_vent_info_file.Enabled = False
```

```
    Dim Num_of_vents As Integer
    Dim DataLine As String

    Open CommonDialog1.FileName For Input As #1
    Num_of_vents = 0
    Do Until EOF(1)
       Line Input #1, DataLine
       Num_of_vents = Num_of_vents + 1
    Loop
    Close #1
    Txt_num_of_vents.Text = Num_of_vents

End If
End Sub

Private Sub Opt_BG_Click(Index As Integer)
If Opt_BG(0).Value = True Then 'a fixed value is chosen for BG average - multiplied by grid area to calculate total BG flux
        Txt_BG_Dist_Type.Text = "fixed average"
        Txt_BG_avg.Text = 20
        Txt_BG_sigma.Text = ""
        Txt_BG_File.Text = ""
        Txt_BG_Dist_Type.Enabled = False
        Txt_BG_avg.Enabled = True
        Txt_BG_sigma.Enabled = False
        Txt_BG_File.Enabled = False
        Txt_X_max.Enabled = True
        Txt_Y_max.Enabled = True
ElseIf Opt_BG(1).Value = True Then 'program will calculate normal distribution for background distribution
        Txt_BG_Dist_Type.Text = "normal"
        Txt_BG_avg.Text = 20
        Txt_BG_sigma.Text = 5
        Txt_BG_File.Text = ""
        Txt_BG_Dist_Type.Enabled = False
        Txt_BG_avg.Enabled = True
        Txt_BG_sigma.Enabled = True
        Txt_BG_File.Enabled = False
        Txt_X_max.Enabled = True
        Txt_Y_max.Enabled = True
ElseIf Opt_BG(2).Value = True Then 'program will import external file with background distribution
        Txt_BG_Dist_Type.Text = ""
        Txt_BG_avg.Text = ""
        Txt_BG_sigma.Text = ""
        Txt_BG_File.Text = ""
        Txt_BG_Dist_Type.Enabled = False
        Txt_BG_avg.Enabled = False
        Txt_BG_sigma.Enabled = False
        Txt_BG_File.Enabled = False
        Txt_X_max.Enabled = False
        Txt_Y_max.Enabled = False

        'select file (ie path and file name) for background input
        'the file name is read to determine a(average) s(sigma) x (metres along x axis) y (metres along y axis) t(type of statistical distribution)
        '++++++++ file name format - e.g. d123a123s123x1234y1234tLL.txt ++++++++++
        Dim fso As New FileSystemObject ' Get instance of FileSystemObject.
```

```
        CommonDialog1.FileName = ""
        CommonDialog1.InitDir = path & "background\"
        CommonDialog1.DialogTitle = "Select background data file"
        CommonDialog1.ShowOpen
        Txt_BG_File.Text = CommonDialog1.FileName
        If Txt_BG_File.Text <> "" Then
           Txt_BG_avg.Text = CInt(Mid(CommonDialog1.FileTitle, 6, 3))
           Txt_BG_sigma.Text = CInt(Mid(CommonDialog1.FileTitle, 10, 3))
           Txt_X_max.Text = CInt(Mid(CommonDialog1.FileTitle, 14, 4))
           Txt_Y_max.Text = CInt(Mid(CommonDialog1.FileTitle, 19, 4))
           Txt_BG_Dist_Type.Text = Mid(CommonDialog1.FileTitle, 24, 2)
           Txt_num_grid_pts.Text = Txt_X_max.Text * Txt_Y_max.Text
        End If
End If

End Sub

Private Function Normal(Optional Sigma As Double = 1, Optional Mean As Double = 0) As Double
  Normal = GetGausse * Sigma + Mean
End Function


Private Function GetGausse() As Double
  ' This Function returns a standard Gaussian random number
  ' based upon the polar form of the Box-Muller transform.

  ' since this calc is capable of returning two calculations per
  ' call, it's been set up to save the second calc for the next
  ' pass through the function, saving some time.

  ' Call the randomize function once (and ONLY once) in the life of the project.

  Static blReturn2 As Boolean  ' Flag to calc new values, or return
                               ' previously calculated value.  It defaults
                               ' to False on the first pass.
  Static dblReturn2 As Double  ' Second return value

  Dim Work1 As Double, Work2 As Double, Work3 As Double

  Const Two = 2#, One = 1#

  If blReturn2 Then  ' On odd numbered calls
   GetGausse = dblReturn2
  Else
   Work3 = Two
   Do Until Work3 < One
      Work1 = Two * Rnd - One
      Work2 = Two * Rnd - One
      Work3 = Work1 * Work1 + Work2 * Work2
   Loop
   Work3 = Sqr((-(Two) * Log(Work3)) / Work3)
   GetGausse = Work1 * Work3
   ' a second valid value will be returned by Work2 * Work3.
   ' Calculate it for the next pass.  This will save some processing
   dblReturn2 = Work2 * Work3
  End If
```

```
  blReturn2 = Not blReturn2 ' and toggle the return value flag

End Function

Private Sub ImportIntoArray(TempArray As Variant, numcolumns, numrows As Long, InputFileName As String, VarType As Integer)

Dim DataLine As String
Dim Countletter As Integer
Dim Field As String
Dim letter As String
Dim countcolumn As Integer
Dim countrow As Long

Open InputFileName For Input As #1
For countrow = 1 To numrows
   Line Input #1, DataLine
   Countletter = 1
   For countcolumn = 1 To numcolumns
      Field = ""
      letter = Mid(DataLine, Countletter, 1)
         Do Until letter = "," Or letter = ""
            Field = Field & letter
            Countletter = Countletter + 1
            letter = Mid(DataLine, Countletter, 1)
         Loop

         If VarType = 1 Then '1 = integer
            TempArray(countcolumn, countrow) = CInt(Field)
         ElseIf VarType = 2 Then '2 = single
            TempArray(countcolumn, countrow) = Round(CSng(Field), 3)
         End If

      Countletter = Countletter + 1
'      Debug.Print TempArray(countcolumn, countrow)
   Next countcolumn
Next countrow
Close #1
End Sub

Private Sub Txt_X_max_Change()
Txt_num_grid_pts = Txt_X_max * Txt_Y_max
End Sub

Private Sub Txt_Y_max_Change()
Txt_num_grid_pts = Txt_X_max * Txt_Y_max
End Sub

Private Sub SurferGrid(DataFileIn As String, AlgorithmType As String, RatioAnisotropy As Integer, AngleAnisotropy As Integer, OutPutGrid As String, Max_X As Long, Max_Y As Long, node_space As Integer, TotalCalcFlux As Variant)

Dim Results() As Double
Dim SurferApp As Object 'Declares SurferApp as an object
Set SurferApp = CreateObject("Surfer.Application") 'Creates an instance of the Surfer Application object and assigns it to the variable named "SurferApp"
```

```
SurferApp.GridData(DataFile:=DataFileIn, xCol:=1, yCol:=2, _
zCol:=3, Algorithm:=AlgorithmType, AnisotropyRatio:=RatioAnisotropy, _
AnisotropyAngle:=AngleAnisotropy, ShowReport:=False, _
OutGrid:=OutPutGrid, xMin:=1, xMax:=Max_X, yMin:=1, yMax:=Max_Y, _
NumCols:=(Max_X / node_space), numrows:=(Max_Y / node_space)) = True

SurferApp.GridVolume(Upper:=OutPutGrid, Lower:=0, pResults:=Results, ShowReport:=False) = True
TotalCalcFlux = Results(srfGVTrapVol)

SurferApp.Quit

End Sub
Private Sub SurferGrid2(DataFileIn As String, AlgorithmType As String, RatioAnisotropy As Integer, AngleAnisotropy
As Integer, OutPutGrid As String, Max_X As Long, Max_Y As Long, node_space As Integer, TotalCalcFlux As Variant,
CV As Variant)

Dim Results() As Double
Dim CVResults As Variant
Dim SurferApp As Object 'Declares SurferApp as an object
Set SurferApp = CreateObject("Surfer.Application") 'Creates an instance of the Surfer Application object and assigns it to
the variable named "SurferApp"

SurferApp.GridData(DataFile:=DataFileIn, xCol:=1, yCol:=2, _
zCol:=3, Algorithm:=AlgorithmType, AnisotropyRatio:=RatioAnisotropy, _
AnisotropyAngle:=AngleAnisotropy, ShowReport:=False, _
OutGrid:=OutPutGrid, xMin:=1, xMax:=Max_X, yMin:=1, yMax:=Max_Y, _
NumCols:=(Max_X / node_space), numrows:=(Max_Y / node_space)) = True

SurferApp.GridVolume(Upper:=OutPutGrid, Lower:=0, pResults:=Results, ShowReport:=False) = True
TotalCalcFlux = Results(srfGVTrapVol)

SurferApp.CrossValidate(DataFile:=DataFileIn, ShowReport:=False, pResults:=CVResults) = True
        CV(1, 1) = CVResults(srfCV_R_RootMeanSquare)
        CV(1, 2) = CVResults(srfCV_R2)
        CV(1, 3) = CVResults(srfCV_ZE_Correlation)
        CV(1, 4) = CVResults(srfCV_ZR_Correlation)
        CV(1, 5) = CVResults(srfCV_ZE_RankCorrelation)
        CV(1, 6) = CVResults(srfCV_ZR_RankCorrelation)


SurferApp.Quit

End Sub
```