Developments since 2005 in understanding potential environmental impacts of CO₂ leakage from geological storage

D.G. Jones¹,*, S.E. Beaubien b, J.C. Blackford c, E.M. Foekema d, J. Lions e, C. De Vittor f, J.M. West a, S. Widdicombe c, C. Hauton g, A.M. Queirós c

¹ British Geological Survey, Keyworth, Nottingham NG12 5GG, UK
b Sapienza Università di Roma, Dip. Scienze della Terra, P.le A. Moro 5, 00185 Roma, Italy
c Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, UK
d IMARES Wageningen UR, Postbus 57, 1780AB Den Helder, The Netherlands
* BRGM (Bureau de Recherche Géologique et Minière), 3 Avenue Claude Guillemin, BP 36009, 45060 ORLEANS Cedex 2, France
e DGs (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) Oceanography Section, Vittorio Emanuele II University, Italy
g Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, European Way, Southampton SO14 3ZH, UK

A R T I C L E  I N F O

Article history:
Received 17 March 2015
Received in revised form 19 May 2015
Accepted 22 May 2015
Available online xxx

Keywords:
CO₂ storage
Environmental impacts
Onshore
Offshore
Aquifers

A B S T R A C T

This paper reviews research into the potential environmental impacts of leakage from geological storage of CO₂ since the publication of the IPCC Special Report on Carbon Dioxide Capture and Storage in 2005. Possible impacts are considered on onshore (including drinking water aquifers) and offshore ecosystems. The review does not consider direct impacts on man or other land animals from elevated atmospheric CO₂ levels. Improvements in our understanding of the potential impacts have come directly from CO₂ storage research but have also benefitted from studies of ocean acidification and other impacts on aquifers and onshore near surface ecosystems. Research has included observations at natural CO₂ sites, laboratory and field experiments and modelling. Studies to date suggest that the impacts from many lower level fault- or well-related leakage scenarios are likely to be limited spatially and temporarily and recovery may be rapid. The effects are often ameliorated by mixing and dispersion of the leakage and by buffering and other reactions; potentially harmful elements have rarely breached drinking water guidelines. Larger releases, with potentially higher impact, would be possible from open wells or major pipeline leaks but these are of lower probability and should be easier and quicker to detect and remediate.

© 2015 Published by Elsevier Ltd.

Contents

1. Introduction .......................................................... 00
2. Onshore near surface ecosystem impacts .......................... 00
   2.1. Introduction .................................................. 00
   2.2. Baselines for near surface gases ............................ 00
   2.3. Research developments since 2005 .......................... 00
      2.3.1. Introduction ............................................. 00
      2.3.2. Injection test sites ..................................... 00
      2.3.3. Natural systems ........................................ 00
      2.3.4. Modelling .............................................. 00
   2.4. Implications for monitoring .................................. 00
   2.5. Significant gaps requiring further research ............... 00
3. Drinking water (aquifer) impacts ................................. 00
   3.1. Introduction .................................................. 00

* Corresponding author. Tel.: +44 115 9363576.
E-mail address: dgg@bgs.ac.uk (D.G. Jones).

http://dx.doi.org/10.1016/j.ijggc.2015.05.032
1750-5836/© 2015 Published by Elsevier Ltd.

Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO₂ leakage from geological storage. Int. J. Greenhouse Gas Control (2015), http://dx.doi.org/10.1016/j.ijggc.2015.05.032
1. Introduction

One of the concerns regarding Carbon Capture and Storage (CCS) is that CO₂ might leak out of the storage reservoir towards the ground surface, with possible adverse impacts on underground drinking water supplies or ecosystems either onshore or offshore.

The IPCC Special Report on Carbon Dioxide Capture and Storage (IPCC, 2005) concluded that, with appropriate controls in place, the local health, safety and environmental risks of geological storage would be comparable to the risks of current activities such as natural gas storage, EOR, and deep underground disposal of acid gas. They envisaged two possible leakage scenarios that might give rise to environmental impacts: (1) abrupt leakage, through injection well failure or leakage up an abandoned well, and (2) gradual leakage, through undetected faults, fractures or wells. These remain the two most likely leakage scenarios (e.g. Paulley et al., 2013; Paulley et al., 2013; RISCS, 2014).

In terms of possible onshore environmental impacts, the IPCC considered that elevated CO₂ concentrations in the shallow subsurface could include lethal effects on plants and subsoil animals and the contamination of groundwater, while high fluxes in conjunction with stable atmospheric conditions could lead to local high CO₂ concentrations in the air that could harm animals or people. However, the IPCC report said little about offshore impacts of CCS, dealing more with the possible impacts of ocean storage of CO₂, which is currently excluded from consideration under both the London Protocol/Convention and the Convention for the protection of the marine environment of the North-East Atlantic (OSPAR). A significant proportion of storage capacity is offshore, particularly in Europe (Geocapacity, 2009).

With regard to soil, the impacts could be the result of increased CO₂ and/or the displacement of oxygen by CO₂. The IPCC recognised that high quality baseline data would be needed to detect low rates of leakage. The IPCC 2005 report mentioned effects on microbes at depth but said little about soil microbes. With regard to plants it stated that: ‘While elevated CO₂ concentrations in ambient air can accelerate plant growth, such fertilization will generally be overwhelmed by the detrimental effects of elevated CO₂ in soils, because CO₂ fluxes large enough to significantly increase concentrations in the free air will typically be associated with much higher CO₂ concentrations in soils’. The effects of elevated CO₂ concentrations would be mediated by several factors: the type and density of vegetation; the exposure to other environmental stresses; the prevailing environmental conditions like wind speed and rainfall; the presence of low-lying areas; and the density of nearby animal populations. The most obvious characteristic of long-term elevated CO₂ zones at the surface is the lack of vegetation; ‘New CO₂ releases into vegetated areas cause noticeable die-off. In those areas where significant impacts to vegetation have occurred, CO₂ makes up about 20–95% of the soil gas, whereas normal soil gas usually contains about 0.2–4% CO₂. Carbon dioxide concentrations above 5% may be dangerous for vegetation and as concentration approach 20%, CO₂ becomes phytotoxic’.

According to the IPCC report, groundwater impacts could be directly from CO₂ or indirectly through displaced brine as a result of pressure increases attendant on CO₂ injection. If CO₂ were to leak into an overlying potable aquifer it would dissolve into the water to form carbonic acid, which in turn could react with the aquifer mineral phases and potentially mobilize in situ (toxic) metals, such as Pb or As, SO₄²⁻ or Cl⁻. Instead, if brine associated with the storage reservoir were to migrate into an aquifer it may increase the groundwater’s salinity, or add toxic elements, that may otherwise exist in low concentrations in the aquifer mineral phases. Although it was known that plume migration and element mobility in either case will be a complex function of interacting chemical reactions (adsorption–desorption, dissolution–precipitation) and site specific characteristics (e.g. aquifer mineralogy, groundwater versus leakage flow rates, redox conditions) that could greatly influence any potential impact, almost no research focussed on these issues had been published when the IPCC report was originally written.

Gaps that were recognised in the IPCC report included:

- The temporal variability and spatial distribution of leaks that might arise from inadequate storage sites.
- Microbial impacts in the deep subsurface.

Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO₂ leakage from geological storage. Int. J. Greenhouse Gas Control (2015), http://dx.doi.org/10.1016/j.ijggc.2015.05.032
• Environmental impact of CO₂ on the marine seafloor.
• Methods to conduct end-to-end quantitative assessment of risks to human health and the local environment.

They considered that research was required on:
• Further knowledge of the history of natural accumulations of CO₂.
• Effective and demonstrated protocols for achieving desirable storage duration and local safety.

There have been quite significant developments in research into potential environmental effects of leakage from CO₂ storage since the IPCC report in 2005. Onshore impacts have been studied in the last 10 years through a wide range of projects. Experimental injection sites have been set up to study both soil/plant/microbial and aquifer impacts on four continents from depths of less than a metre to several hundred metres. Research has also continued on natural CO₂ release sites and laboratory experiments have been conducted to assess effects under more closely controlled conditions with modelling to further understanding. Research into potential near surface (essentially soil) impacts of CO₂ leakage is described further in Section 2, while drinking water aquifer impacts are considered in Section 3. The lack of studies of seafloor impacts has been addressed through a number of recent projects, especially in Europe and Japan, including the world’s first offshore injection experiment off the coast of Scotland for the QCS project. Meso-cosm experiments have considered offshore impacts on individual species, whole communities and ecosystem level processes at a variety of scales and observations have been made at offshore natural CO₂ sites. Experiments have also been conducted in situ through the use of benthic chamber lander systems. These developments are addressed in Section 4.

Effects on man and other terrestrial animals directly from atmospheric CO₂ are well understood and have not been a major focus of research by the CO₂ storage community. There have, however, been some studies into fatalities from the large number of natural CO₂ occurrences in Italy, which concluded that the risk of accidental death from 286 documented seepage sites was significantly lower than many socially accepted risks (Roberts et al., 2011). Their modelling suggested that seepage from storage would be less than that of the natural seeps. This aspect of impacts work will not be considered further here, as we concentrate on impacts to other parts of the ecosystem.

CCS regulations have developed significantly since 2005 (Dixon et al., 2015 this volume). Legislation to allow sub-seabed storage has been enshrined in the London Protocol and Convention. Offshore storage requirements have been produced for the NE Atlantic through OSPAR (Dixon et al., 2009; OSPAR, 2007) and much of the OSPAR regulations taken forward into the EU Directive on geological storage of CO₂ (European Union, 2009), which also covers onshore storage. Regulation has also been developed outside Europe, most notably in the US, Canada, Australia, and Japan. Of particular relevance to this account are the requirements for CO₂ injection Class VI wells in the USA (United States Environmental Protection Agency, 2012), as these set out to protect underground sources of drinking water.

OSPAR requires characterisation of site specific risks to the marine environment and collection of baseline data for monitoring. The site operator must consider the risk of adverse impacts and assess possible effects of leakage on the marine ecosystem, including human health and impacts on legitimate users of the marine environment. Site selection should consider the risk of adverse impacts on sensitive, or endangered, habitats and species and natural resources. This includes possible effects from the CO₂ itself, impurities within the injected CO₂ and any fluids that might be displaced as a result of CO₂ injection. Monitoring should be linked to the risk assessment and ‘impact hypothesis’ which includes evaluation of potential ecosystem impacts.

OSPAR (2007) also recognised the need for further research into the effects of CO₂ on marine ecosystems to consider more species at different life stages, and include microbial communities, experimental field studies and models.

Similarly, marine environmental protection is enshrined in Japanese legislation through an amendment to the marine pollution prevention act (Carbon Dioxide Capture and Storage (CCS) Study Group, 2009) based on the London Protocol. Guidance on the safe operation of a demonstration storage site includes a consideration of the potential environmental impacts (both onshore and offshore) and collection of baseline data against which to compare monitoring results gathered during CO₂ injection. Storage must not harm the conservation of the marine environment and the pollution status of the site must be monitored. Baseline environmental surveys have been carried out at the Tomakomai CCS demonstration site (Tanaka et al., 2014). Offshore regulations in Australia also follow the London Protocol and require a plan that demonstrates that environmental impacts and risks will be at an acceptable level (Office of Parliamentary Counsel, 2014). The Australian state of Victoria is the only one to currently have both onshore and offshore regulations and storage there must not cause significant risk to the environment or human health (State of Victoria, 2008, 2010).

Whilst the requirement for Environmental Impact Assessments for CCS projects is still under review in the Canadian province of Alberta, the federal government did require them for projects such as QUEST (Alberta Energy, 2013).

The EU Directive on CCS (European Union, 2009) and associated guidance reflects the OSPAR FRAM requirements but also includes onshore concerns. Thus one of the purposes of monitoring is for ‘detecting significant adverse effects for the surrounding environment, including in particular on drinking water, for human populations, or for users of the surrounding biosphere’. Guidance following the Directive (European Commission, 2011) states that ‘valuable natural resources in proximity to a potential storage complex have to be documented and the risk linked to the exposure to CO₂ leakage has to be carefully assessed’. This includes consideration of conservation areas, and potable groundwaters in risk assessments and environmental monitoring as part of baseline surveys.

US regulation, in particular the designation of Class VI wells for CO₂ injection for storage (United States Environmental Protection Agency, 2012) stresses the protection of underground sources of drinking water (USDWs). It requires the monitoring of water quality in an agreed network of monitoring wells, above the primary seal on the reservoir, within a defined area of review that is likely to expand as injection proceeds and the CO₂ plume spreads in the subsurface. Water quality has to be tested prior to, during and after injection, initially on a quarterly basis. QA/QC requirements are stipulated as are a minimum list of likely analytes with options to broaden that list. Data are to be reported every 6 months. The regulators may also insist on surface gas monitoring to further evaluate any threat to USDWs or to meet additional state or federal legislation.

Any environmental impacts from CCS need to be considered relative to the benefits to be gained from reduction in greenhouse gas emissions and in the context of the natural variability of the ecosystem and other environmental impacts, including the effects of global climate change (e.g. changing weather patterns, increased frequency of extreme events, ocean acidification), industrial contamination and trawling.

Something of a geographical split has developed between onshore and offshore storage. In North America emphasis is on onshore projects (e.g. Decatur, Cranfield, Weyburn, Quest, Aquistore), although offshore sites are being considered off the Gulf Coast. In contrast, most current (Sleipner, Snøhvit) and proposed

Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO₂ leakage from geological storage. Int. J. Greenhouse Gas Control (2015). http://dx.doi.org/10.1016/j.ijggc.2015.05.032
(ROAD, Peterhead-Goldeneye and White Rose) sites in Europe are offshore, where much of the storage capacity is (e.g. Geocapacity, 2009), and onshore projects have faltered, largely because of public opposition. The Tomakomai project in Japan involves offshore storage. Developments in China (e.g. Shenhua Ordos) and Australia (Gorgon, Otway) are, so far, onshore whilst proposals in South Korea may involve a mix of onshore and offshore sites.

Comparatively little quantified data has been published describing possible leakage scenarios due to high degrees of uncertainty, especially in predicting geological flow mechanisms and rates. Whilst there are some analogues in the form of well blowouts, when drilling into natural CO2 occurrences, that may provide upper limits, there is no direct evidence of significant leakage from existing storage sites. Given this, risk assessments have tended to investigate a range of theoretical leakage scenarios, from a minimum inconsequential leak up to a plausible maximum, invoking operational or geological mechanisms of leakage in each case. Upper limits to possible leakage from transportation can be easily constrained as pipeline flow rates are known (e.g., ~3 ktonnes d−1 at the Sleipner field), and it is assumed that such leaks could be stemmed in a matter of hours to days. Leakage from pipelines at lower rates, and from storage, is more speculative, with fluxes estimated, for example, from <1 tonne d−1, to 10–100 tonne d−1, to >1 tonne d−1 being associated with seepage, abandoned wells/geological discontinuities and catastrophic operational failures, respectively (IEA, 2008; Klussman, 2003). A similarly wide range of possible flux rates was also felt to be plausible by a more recent assessment (Pauley et al., 2012) with areas of emission varying from the diameter of a borehole (e.g. 0.03 m2) up to several point sources in an area of 50,000 m2 for fault-related leakage.

2. Onshore near surface ecosystem impacts

2.1. Introduction

There has been significant research into potential near surface onshore impacts of CO2 leakage from geological storage since 2005, with projects in North and South America, Europe and Australia in particular. Laboratory experiments allow conditions to be closely controlled and have permitted effects from both higher CO2 and lower O2 to be investigated. Field injection tests allow injection rates to be controlled but introduce real world complexity. However, the impact of CO2 may be overain to varying degrees by the effects of other sources of stress such as weather conditions or a variety of pests. Observations at natural CO2 sites may be even closer to the conditions of CO2 leakage from deep storage but the ecosystem has usually adapted over long periods to the presence of the gas. Modelling has helped to extend the findings of the above research and to understand processes.

A parallel stream of research on the effects of rising atmospheric CO2 concentrations on ecosystems appears to have little relevance to the potential impacts of the leakage being considered here. Recent studies (e.g. RISCs, 2014) appear to back up the conclusion of the IPCC report that increased concentrations of CO2 in the soil have a much more marked effect than any associated, and transitory, slight rises in atmospheric levels.

2.2. Baselines for near surface gases

A good understanding of natural/background variability in soil gas, flux and atmospheric gases is important for identifying anomalous values of gas concentration or flux that might be indicative of seepage, as part of near surface monitoring above CO2 storage sites (Jenkins et al., 2015 this volume). The baseline range of these parameters is also useful for identifying areas with enhanced CO2 levels that might be impacted as a result of leakage. Work at natural and experimental CO2 release sites suggests that soil gas concentrations are more important in terms of ecosystem impacts than flux rates (which can vary quite considerably for a given concentration) or atmospheric gas contents (where dilution and dispersion is generally rapid except in unusually calm conditions).

There have been a number of studies of onshore baseline gas concentrations and fluxes. These broadly fall into two categories: extensive measurements at different times, which provide spatial and seasonal variability, and continuous measurements at a relatively small number of sites that give more detailed information on temporal variability. Extensive measurements at Weyburn over a large grid of points that were repeated from 2001–2005 and again in 2011 are an example of extensive datasets (Beaubien et al., 2013). They show the seasonal and year on year variability linked to temperature and rainfall/soil moisture levels, which can be quite marked. For example, some extreme summer CO2 soil gas concentrations exceeded 10%, with highest recorded values nearing 20% (Trium Inc., and Chemistry Matters, 2011). The importance of comparing CO2 values against those of O2 and N2 to determine the origin of anomalies has also been highlighted at this site (Beaubien et al., 2013; Romanak et al., 2014). In other examples a mix of spatial and continuous monitoring was carried out for the Sitechar project in Denmark and the Sulcis site in Sardinia (Beaubien et al., 2014b). These also showed seasonal and land-use variations, with soil gas concentrations above 5% at both sites and over 10% at Sulcis. Whilst soil gas concentrations rise in response to increased biological activity at higher temperatures, higher CO2 values can also result from the retention of gas below low permeability soils. High summer CO2 concentrations, reaching 12-16% were also seen in soil gas surveys at the Lacq-Rousse pilot site in France (Lescanne et al., 2011). Baseline maxima of up to around 10% were measured in extensive coverage at the Otway site in Australia (Schacht and Jenkins, 2014). All the above concentrations were measured at 80–100 cm depth.

Low baseline fluxes from 5 to 13 g m−2 d−1 were recorded by extensive surveys at Hontomin in northern Spain (Elío et al., 2013) with few outliers. There was greater variability at other sites with a summer range of around 5–55 g m−2 d−1 at Lacq-Rousse (Lescanne et al., 2011) and much higher July 2001 results from Weyburn (median = 48 g m−2 d−1; max = 468 g m−2 d−1) although autumn values at this site were all below 60 g m−2 d−1 (Beaubien et al., 2013).

Long term continuous soil gas monitoring was carried out at the Altmark site in Germany by Schlömer et al. (2013) who reviewed its use for CCS projects. They recommended at least 3 years of data to define the baseline. Continuous monitoring below the biologically active soil zone (50 cm in their case) was recommended to increase sensitivity to leakage detection (Schlömer et al., 2014) but measurements within the near surface zone are required to assess potential ecosystem impacts.

Ecosystem baselines have been evaluated in different climates at a few CO2 storage sites, at natural sites where there has been long-term exposure to elevated CO2 soil gas concentrations, and at experimental injection sites. Baselines have included both botanical surveys and various aspects of microbiology as well as acquisition of hyperspectral remote sensing data prior to CO2 injection at experimental sites, such as ASGARD. Faunal and floral diversity was evaluated prior to injection at Lacq-Rousse and will be re-surveyed annually (Lescanne et al., 2011). Floral and microbiological baseline studies were undertaken at In Salah (Jones et al., 2011). The Quest project in Canada includes ecosystem studies and hyperspectral image analysis in its monitoring plan (Shell Canada Limited, 2010). The requirement for ecosystem baselines is likely to become a part of environmental impact assessments as new
projects have to meet the requirements of legislation such as the EU Directive (European Union, 2009).

2.3. Research developments since 2005

2.3.1. Introduction

Since the IPCC report appeared in 2005, research on near surface onshore impacts of CO2 seepage has moved forward using observations from natural CO2 seepage systems, shallow CO2 injection tests (which were deliberately designed to produce leakage through the soil and into the atmosphere), and modelling.

Shallow injection tests assessing potential impacts of CO2 on near surface ecosystems have been carried out at the following sites, which are described in more detail below: ZERT in Montana, USA, ASGARD in the UK, CO2 Field Lab and Grimsrud Farm in Norway, PISCO2 in Spain, Ginninderra in Australia and Ressacada Farm in Brazil.

Many of these sites were set up primarily to test monitoring techniques rather than to study impacts per se, but impacts were monitored so results relevant to this review were generated. The converse is also true, in that sites where the main objective was the study of ecosystem impacts had, of necessity, to monitor the concentration and flux of CO2 and other parameters to assess the impacts and so have helped to develop monitoring methods (Jenkins et al., 2015 this volume).

2.3.2. Injection test sites

Controlled injection of CO2 has been undertaken at the ASGARD field site near Nottingham, England periodically from March 2006 until 2013. The site comprises a number of experimental plots measuring 2.5 m x 2.5 m with injection at 60 cm depth into half of the plots, the remaining being controls (West et al., 2015 In submission; West et al., 2009). The impacts of elevated CO2 on pasture, grass/clover and a number of other crops including barley, wheat, oil seed rape and beetroot, grown in the sandy clay loam soil resting on clay and marl, were examined. Detailed botanical, microbiological, soil gas concentration, and flux measurements were performed plus limited mineralogical and geochemical analyses, (Smith et al., 2013; West et al., 2015 In submission; West et al., 2009).

In brief, when the CO2 injection rate was higher for early experiments at ASGARD (3 L min⁻¹ at 60 cm depth into each gassed plot) there was a rapid response in plant health with a change in leaf colour in a variety of plants within 7–10 days (Smith et al., 2013; West et al., 2009). After 2 years of intermittent gassing on pasture plots, with an injection rate of 3 L min⁻¹ from May 2006 until April 2007 and 1 L min⁻¹ thereafter into each gassed plot (resulting in CO2 soil gas concentrations between 10–30% at 20 cm depth across the transect), monocotyledonous plants (grasses) increased proportionately compared to dicotyledenous (broad-leaved) species on the gassed plots when compared to the controls. When the injection rate into each gassed plot was 1 L min⁻¹ at 60 cm depth over a 24 month gassing period from June 2010, with a resulting overall soil gas concentration below 10%, this increase in monocotyledonous plants across the gassed pasture plots was not observed.

In addition to visible signs of plant stress at ASGARD, such as leaf discoloration and reduced stem height, biomass at harvest was lower in most crops grown in high CO2 conditions (>10% at 30 cm) (Lake et al., 2013; RISCS, 2014; Smith et al., 2013). Autumn sown barley showed a decrease in yield of over 50%. Beetroot leaf weight showed a similar reduction but there was no apparent effect on the biomass of beets. This may reflect the shallow rooting of this plant, which thus avoids exposure to the higher CO2 concentrations (and lower O2) at greater depth. Whilst autumn sown oilseed rape showed no visual damage (unlike spring sown rape) the yield of seed pods decreased by 34%. This suggests that although the vegetative stage of autumn oilseed rape was not affected by the increased CO2 the transfer of resources to the reproductive stage was disrupted leading to a decrease in yield. Grass/clover plots showed a reduction in biomass of both species but with clover more greatly affected and, unlike the grass, showing no signs of recovery after injection ceased. The clover may in part have been affected by very wet weather in 2012. However, there is some evidence that leguminous plants like clover may be more strongly affected by high soil CO2 due to the high oxygen requirement of the symbiotic bacteria in their root nodules (Lake et al., 2013). Leaf discoloration is commensurate with nutrient deficiency and this may result from reductions in root biomass preventing the plant from accessing sufficient nutrients and water.

The microbiological studies in the pasture plots at ASGARD showed that after 16 weeks of CO2 injection at 3 L min⁻¹ and where soil gas concentrations reached 90% at 70 cm depth, no active biomass was detected. However, after 2 years of CO2 injection at 1 L min⁻¹, results were not as clear. There was no obvious impact on total numbers or biomass where CO2 soil gas concentrations were below 10% at 20 cm (Smith et al., 2013). However, where soil gas concentrations vary between 20–40% at 20–50 cm depth, total numbers of organisms appeared to increase when compared to those observed in the ungaussed plot. 16S rRNA analyses of Bacteria and Archaea using quantitative real time PCR showed variations in microbial copy numbers over time and depth. However, these variations may not result from CO2 injection but from changes in weather conditions; the experimental period included spells of unusually wet and dry weather.

An experimental injection site was set up by Bioforsk at Grimsrud Farm in Norway as part of the RISCS project (Moni and Rasse, 2013a; Moni and Rasse, 2014). A longitudinal CO2 gradient was produced by injection of CO2 at a rate of 2 L min⁻¹ at one end of 6 × 3 m plots at a depth of 85 cm in a permeable sand layer beneath 40 cm of top silt loam topsoil. The simulated CO2 seepage caused visible discoloration of oats with reduced plant growth, chlorophyll and canopy water contents at the end of the growing season. This appeared to be caused only by changes in soil gas concentration (increased CO2 and/or reduced O2) as there was almost no change in atmospheric CO2 concentrations. The changes in the plants were spatially limited such that estimated losses in yield for a typical European field were below 1% and less significant than those due to other common factors such as waterlogging. The impact on the oats could be readily detected using hyperspectral indices such as NDVI705 and the Agricultural Stress Index.

Laboratory experiments to investigate the effect of enhanced soil gas CO2 and reduced O2 on oats and wheat grown in sandy soil were also undertaken by Bioforsk for RISCS (Moni and Rasse, 2013b; Moni and Rasse, 2013b; RISCS, 2014). Higher CO2 affected the development of oats; this could be mitigated at 40% CO2 by maintaining normal O2 levels but this was not effective at higher CO2 concentrations (75%). This suggests that both CO2 toxicity and anoxia have an effect. However, repeat experiments with better established oats and wheat plants did not show any effects at up to 40% CO2, suggesting that the size and/or physiological condition of the plants may also influence the impact.

The PISCO2 site in northern Spain was set up to examine microbiological, botanical, and biogeochemical effects of CO2 seepage in a range of different soil types from across Spain. During the first year of operation the CO2 injected at 0.33 and 0.66 L min⁻¹ into sandy loam and loam soils, had no significant effects on the soil microbial communities (Montiel and Mantecón, 2013; Fernández-Montiel et al., 2015).

The ZERT site in Montana was established primarily to develop and test near surface monitoring methods using shallow release of CO2 from 2 m depth through an undisturbed organic silt and clay soil with some sand resting on sandy gravel (e.g. Spangler et al., 2010). However, parts of the work conducted there are relevant to

Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO2 leakage from geological storage. Int. J. Greenhouse Gas Control (2015). http://dx.doi.org/10.1016/j.ijggc.2015.05.032
environmental impacts. They include a study of plant stress effects, the use of CO₂-induced plant stress observed by hyperspectral imagery, and microbiological studies over several seasons of injection. Sharma et al. (2014) examined the physiological responses of dandelion and orchard grass to high CO₂ concentrations in the soil. Circular areas of leaf senescence and dieback formed after one to two weeks of injection that were about 3 m in diameter. These areas showed more pronounced effects on dandelion leaves than those of orchard grass with clear reduction in chlorophyll levels and assumed increase in pigments such as anthocyanins, causing purple colouration of dandelion leaves, and increased stomatal conductance. The greater tolerance of orchard grass to CO₂ seen at ZERT is paralleled in studies at other injection and natural sites. Changes in reflectance spectra were observed within 1 m perpendicular from the 100 m long CO₂ release pipe in alfalfa (Keith et al., 2009). A wider area of plant stress of up to 2.5 m from the injection well was observed after 10 days of injection of 300 kg d⁻¹ (≈100 L min⁻¹ spread over 6 packered zones) of CO₂ in the following year and comparison with soil CO₂ measurements suggested a lower limit for plant stress of 4–8% at 30 cm depth (Male et al., 2010). Plant stress was indicated by visible purple leaf discoloration and changes in several indices indicating reduction in both chlorophyll A and B (Lakkaraju et al., 2010). Bellante et al. (2013) developed a red edge index from hyperspectral data at ZERT that was significantly lower for CO₂-stressed vegetation than healthy vegetation. This appeared to show a progressive response to CO₂ injection and subsequent recovery once injection ceased. A low cost multispectral system was also able to identify plant stress but as with the hyperspectral work, unable to unequivocally define the cause of that stress (Hogan et al., 2012a, b). Enhanced CO₂ in the soil at ZERT also affected both the abundance and activity of microbial functional groups mediating C and N transformations (Morales and Holben, 2013). However, these effects were not equal suggesting there may be two separate mechanisms. The results differ in some respects from studies on elevated atmospheric CO₂ suggesting that microbial response to subsurface seepage may be different.

The Ginninderra site in Australia is based on the ZERT model, with releases over a 100 m length of horizontal pipe at about 2 m depth into undisturbed soil, and mainly focusses on monitoring methods. Early experiments involved injection of 144 and 218 kg d⁻¹ (≈50 L min⁻¹ and ≈75 L min⁻¹) of CO₂ through 6 packered intervals. Hyperspectral investigation of plant stress was combined with microbial soil genomics at Ginninderra (Feitz et al., 2014a). Crops of field pea, wheat and barley were visibly impacted by the CO₂ with patches of yellowing or drying out of the plants. Leaves of oil seed rape (Canola) turned purple. These impacts were restricted to circular areas 5–15 m in diameter. Crops generally showed lower chlorophyll contents and dried out as they matured and suffered from lack of water, but this was more pronounced where CO₂ fluxes were higher. Changes in chlorophyll C were one of the strongest indicators of CO₂ impact. It was difficult to distinguish the effects of drought and CO₂ in hyperspectral data for mature field pea and wheat. However, preliminary analysis does suggest some promise for vegetation indices such as the chlorophyll normalized difference index, although this does not appear to be specific for CO₂-induced stress so there remains a danger of false positives. Microbial surveys showed a clear switch in the community, post-injection in higher CO₂ areas, towards anaerobic respiration. This could be distinguished statistically from larger seasonal shifts.

The Ressacada Farm site in Brazil became operational more recently, with injection at 3 m depth from a vertical borehole, and includes work on plant stress (Moreira et al., 2014a, b) but full initial results have yet to be published.

2.3.3. Natural systems

Hyperspectral assessment of plant stress using airborne data has been undertaken at natural CO₂ seepage sites including Latera in Italy and Laacher See in Germany (Bateson et al., 2008; Govindan et al., 2013, 2011). These studies build on previous work such as that summarised by Pickles and Cover (2005) from the CO₂ Capture Project, which was cited in the IPCC report.

Studies at natural CO₂ seepage sites have also included direct assessment of impacts on plants and microbes, for example those carried out at Latera (Beaubien et al., 2008; Oppermann et al., 2010), Laacher See (Frerichs et al., 2013; Krüger et al., 2011 Krüger et al., 2009) and Florina (Ziogou et al., 2013). The results of these studies are detailed in West et al. (In submission). Each of these sites has yielded considerable information on the impacts of elevated CO₂ soil gas concentrations and soil CO₂ surface flux rates on a range of parameters, including botany and soil microbiology, for pasture in different European climatic conditions. It has to be borne in mind, however, that the ecosystem at these sites has had time to adapt to higher CO₂ conditions over many years and thus shorter term responses to leaking CO₂ may not be the same. Results from botanical research have shown that where pasture plants have been exposed to high CO₂ soil gas concentrations for many years, species-specific responses to high CO₂ concentrations will be dependent on a number of factors, including temporal and spatial soil CO₂ concentration and flux patterns. Site specific plant ‘biomarkers’ were observed for concentrations >35% at 15–20 cm depth, such as Polygonum arenastrum Boreau at Laacher See and P. aviculare L at Florina. Polygonum spp. can be regarded as a biomarker of elevated CO₂ concentrations at both sites. At Latera, the acid tolerant grass, A. capillaris L. can also be regarded as a biomarker tolerating between 15–40% CO₂ concentrations at 10 cm depth (Beaubien et al., 2008). At Florina and Latera, monocotyledonous plants (grasses) become increasingly dominant where soil gas CO₂ concentrations at 20 cm depth exceed 10% up to a maximum of ≈30–40%. Interestingly, this was also observed at ASGARD after the pasture plot had been exposed to elevated CO₂ soil gas concentrations between 10–30% for 24 months. Similar changes in grass coverage have also been observed at the ZERT controlled CO₂ injection site in grassland in Montana, USA (Zhou et al., 2013). The results from all the sites suggest that a concentration of between 10–15% CO₂ soil gas at 20 cm depth, which is within the root zone, is an important threshold level for observing changes in plant coverage. Flux rates will also be important because these influence soil gas concentrations. However, flux rates are also influenced by other factors, such as soil moisture content, atmospheric pressure and wind speed so determining a threshold will be dependent on site characteristics. Additionally, it must be recognised that plants are not exposed to constant fluxes and/or concentrations of CO₂ so these values can only be regarded as indicative and must be considered in the context of a particular site.

Microbiological analyses at all the natural CO₂ vent sites suggest significant changes in the soil microbial community caused by high CO₂ levels in the soil gas (West et al., 2015 In submission). For example, there is a shift in the microbial community composition in the vent cores towards anaerobic and acid tolerant microorganisms as well as an ecosystem adaption to the CO₂ induced soil biogeochemistry. These differences may be related to long-term adaptation of the microbial ecosystems at the natural sites because they are not observed at ASGARD. Observations at ASGARD are unclear but may suggest that when CO₂ soil gas concentrations are <10%, there is little impact on the microbial ecosystem. Soil gas concentrations between 15 and 40% may enhance biomass and total numbers of organisms (as seen at Latera and Laacher See). Taken together, the results from all the sites suggest that microbiological changes can be observed when CO₂ soil gas concentrations are between 15–40%.
However, it is not clear how rapidly these changes occur. A relatively short exposure (approximately 24 months) appears to have had little immediate impact on the microbial community composition at ASGARD. Such changes are, however, observed at Latera, Laacher See, and Florina where exposure to elevated CO₂ concentrations has been taking place over much longer periods.

Soil microbial studies have also been carried out at the Campo de Calatrava natural CO₂ site in Spain (Sæz de Miera et al., 2014). As CO₂ flux increases the relative abundance of Chloroflexi increased, whereas the relative abundance of Acidobacteria, Verrucomicrobia and Gemmatimonadetes phyla decreases. Within the Chloroflexi phylum, operational taxonomic units related to the genera Thermogemmatispora, Ktedonobacter and Thermomicrobium dominated bacterial communities sampled in sites with the highest CO₂ flux. It should be noted that the highest fluxes quoted at this site are over 500 kg m⁻² d⁻¹ and therefore 2 orders of magnitude higher than those at other studied sites and may not be representative of likely seepage rates from CO₂ storage.

Diversity of the edaphic protozoan communities also decreased with increasing CO₂ flux at this site (Gablonzo and Bécares, 2014). A shift from Polyhymenophorea (Spirotrichea and Heterotricha) to Colpodea dominated communities and a decrease in the percentage of rapacous ciliates was observed as CO₂ increased.

2.3.4. Modelling

Models have been developed to represent and improve our understanding of the soil/plant system at Latera (Maul et al., 2009) and the ASGARD and Grimstrup Farm sites (Bond et al., 2013). These show that, at moderate CO₂ levels, there can be CO₂ utilisation effects as well as the more negative impacts at higher gas concentrations. These are apparent at natural and experimental CO₂ sites as zones of much lower growth surrounding the main areas of gas venting. The system-level models have been able to reproduce the main features of the sites, suggesting that the key features are understood. The ASGARD and Grimstrup models could only mimic the actual patterns of CO₂ flux if flow was channelled along narrow pathways rather than diffusely through the bulk of the soil (RISCS, 2014). This suggests that flow is concentrated in more permeable zones such as fractures, which were seen at the experimental sites after dry periods. This may help explain the observed restriction of impacts to small patches of ground whilst most of the wider area containing the vents does not experience higher CO₂ concentrations.

2.4. Implications for monitoring

It has been observed at both natural and experimental injection sites that impacts appear to be spatially restricted, generally metres to tens of metres in scale. This means that leakage monitoring has to be capable of detecting such small targets, whilst the surface footprint of areas above CO₂ storage sites is large (up to hundreds of km²) (Feitz et al., 2014b; RISCS, 2014).

Plant stress effects occur relatively rapidly (in days to a few weeks) so there is the potential to see leaks fairly quickly, but not immediately as, for example, might be the case for pressure monitoring. Some plants are also potential ‘biointicators’ of elevated gas concentrations so could, in themselves, be considered as a possible monitoring tool.

It would not be possible to use plant-stress monitoring at sites with little vegetation (e.g. In Salah). Leaks might be picked up through changes in microbial activity or community structure but simple, cheap and rapid analytical tools would have to be developed to make this into an effective tool (Noble et al., 2012). Monitoring of the microbial populations and community make-up could also provide early indications of increasing CO₂ impacts on soil ecosystems and may also indicate where there has been long-term undetected exposure.

There are a number of drawbacks to using hyperspectral or multispectral remote sensing techniques for seepage monitoring. They may only work well in certain seasons with, for example, winter snow cover or lack of crop cover after harvest or tillage limiting its usefulness to the main growing season. Also effects such as seasonal drought may cause, for example, summer dieback in areas with hot dry summers. Arable areas may also be more difficult to monitor with crop rotation adding additional complicating factors for interpretation of stress.

The amount and duration of baseline near surface gas data required remains a point of debate. Some authors argue that baseline gas monitoring may not be necessary at all because the origin of any gas can be identified (e.g. Romanak et al., 2012a), whilst others suggest at least three years of data are needed (Schlömer et al., 2013). Baselines are important to help define trigger levels where investigation of the source of CO₂ is needed, and may represent a useful tool for site assessment and public awareness at proposed sites (Beaubien et al., 2014b). Legislation requires environmental impacts are assessed and that suggest the acquisition of baseline data of this type will be needed in most jurisdictions. The precise extent of baselines is likely to be site specific.

2.5. Significant gaps requiring further research

Whilst there have been a number of detailed baseline studies of CO₂ in soil gas, fluxes from the soil to the atmosphere and near ground atmospheric concentrations, these have covered only a limited range of environments. Further research is needed to broaden this range and to help define possible anomalous (threshold) values that might indicate leakage in wide area monitoring of storage sites.

To be really effective hyperspectral or multispectral remote sensing approaches to monitor plant stress and detect possible leakage need a better success rate with fewer false positives. It would be advantageous if indices could be developed that are unique to CO₂-induced stress and not affected by other stressors. The use of ground-based or unmanned aerial vehicles (UAVs) for cheaper more frequent surveys of plant stress could be investigated. Impacts need to be studied on a wider range of plant types and geographical settings and compared in a more quantitative way to the effects of other stresses, such as drought, waterlogging or pests and diseases, in order to put the impacts of CO₂ in a broader context.

This could involve laboratory work, field injection tests and further work on natural CO₂ seepage sites. There have been virtually no studies on the effects of impurities in the CO₂, such as CH₄, H₂S and SO₂ although these do occur at some natural CO₂ sites. Field injection experiments have not yet explored to any great extent the impact of the timing of leakage i.e. in what season the event occurs. However, the so far limited evidence suggests that this could be significant. Some natural sites (e.g. those associated with faulting in sedimentary basin settings) are much better analogues of leakage and could provide valuable data to help constrain the likely nature (flux and area) of leakage from storage through faults or fractures as well as giving information on impacts.

Modelling has provided useful insights into processes such as transport through the soil and is now able to link transport and plant impact modules. However, simulations of near surface leakage and its impacts onshore has so far been limited and there is scope for further study not only of the soil–plant system, but also of the circumstances leading to, and likelihood of, possible build ups of CO₂ in the atmosphere that could affect man and other animals. Further research is needed on microbial responses to higher CO₂ to supplement the relatively small number of studies to date and, if possible, to identify changes specific to CO₂ rather than other factors.
The dependence of impacts on soil type has yet to be explored and the mobilisation of potentially harmful elements, whilst a focus of research on aquifers, has not been considered for soils. There needs to be more work on the recovery of different ecosystems after seepage ceases including research on any long term effects on soil fertility (i.e. soil quality and value, which are especially relevant for agriculture) and the need for remediation or mitigation strategies, such as user of fertilisers or deep tilling.

3. Drinking water (aquifer) impacts

3.1. Introduction

Despite the fact that the potential risk to potable groundwater resources has often been cited as one of the major concerns that the public has regarding large-scale CCS deployment, very little focussed research had been conducted on this topic before 2005. In fact, in the IPCC report (IPCC, 2005), the section dedicated to “Hazards to groundwater from CO2 leakage and brine displacement” consists of half a page of text and makes reference to only one published article (Wang and Jaffe, 2004).

At that time the level of dedicated published research on chemical reactions and impacts within the deep injection reservoir (e.g. Emberley et al., 2005) was much more advanced than that for shallow aquifers, given the need to understand their influence on practical issues like CO2 injectivity and storage integrity. While some processes can be extrapolated to shallow systems, the different physical characteristics of these two environments means that this is limited. For example, a deep storage reservoir is characterised by high pressures (which influence CO2 solubility), multi-phase liquid components (particularly super critical CO2 but also hydrocarbons in EOR settings), and host water that is typically saline, whereas shallow aquifers have higher groundwater flow rates, often have cooler temperatures (which affect kinetics), can be chemically oxidising, and typically host very low salinity waters.

In many cases research conducted in shallow environments for issues unrelated to CCS has been found to be more appropriate, and thus, such data was (and still is) used to help predict potential impacts of CO2 intrusion (e.g. Aiuppa et al., 2005). This has included research into such diverse topics as, for example, acid rain, mine tailings waste, and regional groundwater chemistry in volcanic/geothermal areas. While such “proxy studies” provide invaluable data, various specific issues related directly to CO2 and/or brine leakage from a CCS storage reservoir are better addressed by targeted studies.

Since 2005 the situation has changed significantly, with a large amount of focussed CCS-related groundwater research having taken place. This is illustrated in Fig. 1, which shows how scientific productivity in this area of research has increased dramatically over the last five years, using laboratory and field experiments, natural analogue studies, and geochemical modelling to address such issues as trace element release, natural attenuation processes, plume dynamics, MMV (Monitoring, Measurement and Verification), and health and economic risks.

Over this period three excellent review articles have also been published (Harvey et al., 2012; Lemieux, 2011; Lions et al., 2014a). The following will build on these works by summarizing the large number of new articles that have recently been published (Fig. 1), by addressing some topics that have received less focus in the past, and by using this 10th anniversary since the publication of the IPCC report to outline possible knowledge gaps and areas for future research.

3.2. Baselines: natural groundwater chemistry and natural variability of dissolved CO2

Edmunds and Shand (2009) define the natural baseline concentration of a substance in groundwater as its range of concentrations derived entirely from natural, geological, biological or atmospheric sources under conditions not perturbed by anthropogenic activity. These values are often used to identify ‘anomalous’ concentrations relative to typical values. When initial conditions may include some anthropogenic component the terms ‘background’ or ‘threshold’ are preferred. Precise determination of baseline or background concentrations is needed to distinguish any water quality changes induced by CO2 and/or brine leakage from deep CO2 storage.

Groundwater is in equilibrium with dissolved and/or free-phase gases (e.g. CO2, O2) that are naturally present (Appelo and Postma, 2005). The main sources of CO2 in groundwater are plant-root respiration and oxidation of modern organic carbon in the unsaturated zone (especially the soil), oxidation of old organic carbon in the aquifer matrix, acid neutralization reactions with carbonate minerals, and upward leakage of magmatic or petroleum-reservoir CO2.

Groundwater CO2 partial pressures are typically ~10–100 times higher than atmospheric. Groundwater CO2 is not uniformly distributed, being most variable near the water table, lowest at intermediate depth, and highest in deep aquifers where non-potable water is typical (Coudrain-Ribstein et al., 1998; Macpherson, 2009). This results from thermodynamic equilibrium between the gas and water that depends on the pressure and temperature conditions of the system. For example, in deep systems, the dissolution of CO2 in the aqueous phase is notably enhanced due to higher pressures.

Vesper and Edelborn (2012) illustrate that a range of dissolved CO2 concentrations occur in different hydrochemical...
facies (e.g., coal mine discharge: 0.23–4.7 mM; carbonate waters: 0.30–1.4 mM; thermal-mineral waters: 0.31–10.6 mM). Higher CO₂ concentrations (>20 mM) have been reported for groundwater affected by experimental CO₂ leaks (e.g. Kharka et al., 2010), travertine depositing springs (Pentecost, 2005), coal mine drainage, mineral springs, within faulted rocks (Keating et al., 2010), and volcano-associated springs.

Baseline surveys have to define natural variability as baseline chemistry of groundwaters varies both spatially and temporally due to many complex factors, including climate, soil type, geology and residence time, which control hydrochemical evolution. For example, baseline groundwater studies in aquifers above active or proposed injection sites show a very wide range of pH values, such as 6.6–7.5 at Ottway, Australia (de Caritat et al., 2013), 5.4–7.5 above Cranfield, USA (Yang et al., 2013b), and 6.7–8.2 above Hontomin–Huermeces, Spain (Nisi et al., 2013); other parameters, like Total Dissolved Solids (TDS) and electrical conductivity (EC), vary as much if not more at these sites. Groundwater geochemistry is a function of water–rock–gas reactions occurring along the flow path between the recharge and discharge areas. Moreover, the chemistry of natural waters is mainly related to chemical thermodynamics and to reaction mechanisms and rates. All these factors result in a wide range of water chemistries associated with different lithologies (from carbonate-poor silicilastics to carbonate dominant) and geochemical processes, such as water–gas interaction, mineral dissolution/precipitation, redox reactions, ion exchange, and mixing.

Macpherson and Sophocleous (2004) noted that the chemistry of deep groundwater in unconfined aquifers and water in confined aquifers is typically relatively stable in time at any one location. But vertical stratification in groundwater chemistry exists, especially in very shallow, unconfined aquifers where recharge water is chemically different from aquifer water, and/or where strong redox zonation exists. Whilst the assumption is often made that the water chemistry at any single location in an unconfined aquifer is virtually unchanging, observations show that composite parameters (e.g. specific conductance) may change significantly over relatively short time periods. For example, these authors demonstrated that the water chemistry of a shallow, unconfined, floodplain aquifer in Kansas changed significantly and in an oscillatory manner over two and a half years of monitoring. At this site, periodic vertical chemical stratification in the water alternated with periods of chemical homogeneity.

Macpherson et al. (2008) quantified the increasing alkalinity and dissolved CO₂ in a thin, shallow limestone aquifer in a mid-continental temperate-climate setting. Their results suggest that shallow groundwater, unaffected by urbanization or other changes in land use or surface-water engineering, is associated with an increase in net chemical weathering driven by increasing subsurface CO₂. Consequently, the annual cycle and long-term increase in atmospheric CO₂ is also observed in shallow groundwater. CO₂. Annually, natural pCO₂ values range between around 10⁻².6 and 10⁻¹.5 atm. Consequently, weathering is impacted and the chemistry of groundwater oscillates annually with a range of approximately 0.7 pH units, up to 1 mmol/L alkalinity, 0.15 mmol/L Ca, and 0.1 mmol/L Mg.

3.3. Range of potential Impacts (amplitude, spatial footprint)

Impact on an overlying groundwater resource could come about in two ways. First, by leakage of a portion of the injected CO₂ into an overlying aquifer, resulting in acidification and in situ chemical reactions that change water quality. Second, by the migration of deep reservoir brines into an overlying aquifer, resulting in the addition of high concentrations of dissolved elements and in situ reactions induced by the brine's different pH, Eh and water composition. While in the first case anomalous element concentrations are dependent on the aquifer mineralogy/chemistry itself, the second case can “import” species that may not occur in that aquifer. Clearly these two mechanisms could also occur together.

Aside from the far-field, up-dip pressure displacement of brine out of the reservoir unit, gas or brine leakage into a fresh-water aquifer, if it occurs, would predominately be from individual points/small areas. This could be from a leaking well or from single or multiple points along a fault due to channelled flow. The resultant plumes of impacted water within a shallow potable aquifer will tend to be long and narrow due to a constant groundwater flow direction, although plumes can be deviated if they fall within the capture zone of a pumped well or can coalesce with other plumes if there are multiple, nearby leakage points (Fig. 3a). While the spatially restricted nature of a single plume will limit its impact on the groundwater resource as a whole, it will also make monitoring and the discovery of its presence much more difficult (Carroll et al., 2014). For example, well 1 in Fig. 3a may not observe the nearby leak because its capture zone is up-gradient of the source, while the down-gradient well 2 may not observe it because of distance, slow groundwater flow velocity, and/or plume attenuation. The size and chemical composition of any plume will be a function of site specific parameters, such as:

- Aquifer mineralogy. The types of minerals (carbonates, silicates, oxides, etc.) will control geochemical reactions and will provide pH and Eh buffering capacity within the aquifer. These primary and secondary mineral phases are also sources and sinks for trace elements, whose mobilization could result in changes to groundwater quality. Because different species will migrate at different velocities, as a function of their reactivity, plumes will be zoned such that conservative elements like Cl or stable water isotopes will migrate at the velocity of groundwater flow while reactive species will be attenuated (Fig. 3b).
- 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 only accumulate directly along the vertical leakage path thus limiting its impact (Fig. 3c). In contrast, gas leaking into a confined aquifer will accumulate as a separate gas phase at the top of the aquifer (e.g. Lamert et al., 2012), create a long-term source over a much wider area, and produce a larger plume (Fig. 3d).
- Leakage rate versus aquifer groundwater flow rate. This will influence plume size and impact severity through mass balance effects, particularly in terms of dilution and its effect on mineral solubility and buffering/ scavenging capacity (Carroll et al., 2009).
- Chemical composition of leaking gas. The main impact of a CO₂ gas–oil leak on an aquifer will be a lowering of the pH, however if a significant percentage of the leaking gas consists of reduced species like CH₄ or H₂S there is also the potential to change the Eh of the system. This will affect the stability of various mineral phases, making them more or less soluble and changing the potential for trace element release and/or scavenging.
- Chemical composition of leaking brine. The level of specific trace elements that may pose a health hazard, the concentration of major species that can influence ion exchange or form complexing agents, and the pH and Eh levels will all influence the potential impact of a brine plume within a potable aquifer.
- Leakage pathway. As alluded to above, the lateral displacement of brine as a function of the injection-induced pressure pulse could potentially impact a larger area far from the CO₂ injection area if the up-dip extension of the saline reservoir is hydraulically connected with a fresh water aquifer, whereas leakage along wells or faults will likely result in a more spatially restricted impact.
3.4. Research developments since 2005

The various approaches taken in the literature can be thought of as a continuum that covers a wide spectrum of both spatial and temporal scales. At one end are laboratory batch and column experiments which study pore-scale processes for relatively short time periods under controlled conditions. Intermediate are field experiments where issues related to geological heterogeneity and scale can be addressed, but again for limited time periods. Finally, the study of naturally leaking CO₂ sites permits an examination of long-lived systems that may have attained steady state. Results from all of these approaches can be interpreted using geochemical models. Although each method has its own advantages and shortcomings, results from the various approaches are often different due to different kinetic rates and levels of complexity for the studied scales.

3.4.1. Laboratory experiments

Batch experiments are the most common, involving the continuous flow of CO₂ gas into a small volume of homogenized aquifer material and the monitoring of the major and trace element aqueous geochemistry to determine impact on water quality, controlling processes, and kinetic rate constants. In some cases, the systems are agitated to minimize reaction times. Another approach involves column experiments, where CO₂ charged water and/or brine is made to flow along a 1D column packed with aquifer material. This addresses better spatial plume evolution but is limited by slower kinetics and by a relatively short flow path within a homogeneous lithology.

Montes-Hernandez et al. (2013) studied sorption of trace elements onto calcite and goethite, two minerals frequently found in aquifers. They showed that adsorption could prevent the remobilization of Cu, Cd, Se(IV) and As(V) if CO₂ intrudes into a freshwater aquifer. In fact the adsorbed ions on calcite were not remobilized when CO₂ intruded into the system. This means that dissolution of calcite induced by CO₂ is dissociated from potential remobilization of sorbed Cd and Cu. The role of calcite in Cd immobilization onto sediment (illite and calcite) was also confirmed by Frye et al. (2012), who concluded that calcite contents as low as 10% can mitigate the effect of pH reduction and that Cd release and sorption is strongly pH-dependent.

Montes-Hernandez et al. (2013) showed that the oxyanions Se(IV) and As(V) are red-adsorbed on goethite. As(III), which is reputed to be highly mobile, is only slightly remobilized by CO₂ intrusion. In contrast, goethite-adsorbed Cd and Cu were largely re-mobilized by CO₂ injection. This can be explained by a significant change in the surface charge of the goethite after CO₂ intrusion, assuming that surface charge is strongly dependent on pH as it is widely described in the literature.

Frye et al. (2012) investigated the effect of salinity on desorption behaviour in a column experiment. Besides a higher solubility of CO₂ in low ionic strength solutions with a lower pH, they showed that desorption is enhanced by high ionic strength. For example, Cd desorption was to a large extent controlled by ion exchange processes.

Batch experiments on natural sediments (Little and Jackson, 2010; Lu et al., 2010; Mickler et al., 2013; Viswanathan et al., 2013) highlight two different metal-behaviours after CO₂ intrusion: (i) a fast initial release driven by CO₂ intrusion (carbonate dissolution and surface processes); and (ii) a slow release trend driven by kinetically constrained processes, such as mineral dissolution (mainly silicates and aluminosilicates). Varadhharajan et al. (2013) describe two possible mechanisms for the initial fast release, one driven by pH and one by the formation of metal-carbonate complexes that enhance the release of some trace elements (As, Mo, Sr, Mn).

Cahill et al. (2013) carried out a series of CO₂ exposure batch experiments to study sediments of varying composition. They concluded that carbonate contents control pH and alkalinity. Carbonate sediments with clays showed the most severe change in water chemistry, in all major and trace elements, coupled to minimal reductions in pH due to carbonate buffering capacity. Silicate-dominated sediments exhibited small changes in major ion concentrations and the greatest reduction in pH, inducing a higher mobilization of trace elements sensitive to pH.

Little and Jackson (2010) concluded that solid-phase metal mobility, carbonate buffering capacity and redox state in shallow aquifers influence the impact of CO₂. According to these authors, the best approach to study the influence and impact of redox state on the groundwater system is to conduct field experiments to ensure that in situ conditions are maintained. In fact, interactions with the atmosphere have to be carefully controlled to avoid any subsequent modification of redox conditions that can bias laboratory results obtained for the behaviour of trace elements, such as Fe, Mn and As (e.g. sulphide oxidation and oxide precipitation).

All these studies underline that potential mobilization of hazardous elements may be self-mitigated because adsorption occurs when pH rebounds, meaning that in flow system scavenging of elements in solution will occur down-gradient as pH increases during mineral buffering.

3.4.2. Field experiments

Controlled release field experiments extend knowledge gained in the laboratory by addressing the influence of natural geological complexity, heterogeneity, and scale on plume size, migration, temporal/spatial evolution, and impact. While previously applied in other areas of contaminant hydrogeology (Jellali et al., 2001; Rivett et al., 2001), no such experiments had been performed prior to 2005 to assess the potential impact of CO₂ leakage on groundwater quality. This has recently changed, however, as the results from at least 12CO₂ injection experiments have been published since 2007; to date no brine injection experiments have been performed. Experimental design can be broadly divided into “push-pull” and “down-gradient monitoring”. The former consists of using a single well to inject CO₂-charged water into an aquifer and then pumping it out after a given incubation time, while the latter consists of injecting CO₂ gas (or CO₂-charged water) into an aquifer using one or multiple wells and then monitoring the developing plume down-gradient using multiple piezometers or multi-level wells.

Significant differences have been found between field and laboratory results, highlighting the importance of recreating conditions as close as possible to those which would occur for a real leak. In particular, numerous studies have found that reaction and element-release rates in laboratory batch experiments are 2 to 100 times higher than those observed in field experiments (Matter et al., 2007; Mickler et al., 2013; Yang et al., 2013a). These authors postulate that the unrealistically high water / rock ratios, reactive surface area (due to disaggregation), and mixing in batch experiments maintain mineral phases under saturated and thus mass transfer rates artificially high. Only Yang et al. (2014b) found comparable release rates for both experimental types.

Field experiments also give critical data regarding temporal and spatial variability. For example, significant chemical differences have been observed prior to injection even within the small areas encompassed by these experiments. This includes, amongst others, pH variations of 1 to 1.5 units (Cahill and Jakobsen, 2013; Humez et al., 2014b; Peter et al., 2012; Trautz et al., 2013), salinity stratification (Humez et al., 2014b), and strong EC differences (e.g. 420–1,060 µS/cm) [Peter et al., 2012]. This illustrates how background variability must be well understood (see section 3.2) to define an appropriate monitoring strategy for a given site and to avoid interpreting natural anomalies as being due to leakage (i.e. false positives).
Spatial variability has also been observed after injection has commenced as a function of the hydrogeological characteristics of the studied aquifer. Cahill and Jakobsen (2013) found no CO₂ in an aeolian sand overlying the glacial sand in which injection occurred due to its slightly lower permeability. Cahill et al. (2014), working on the same site but with injection in both sand units, observed rapid and more uniform changes of CO₂ and other parameters in the glacial sand as compared to slower, erratic, and longer persistence in the aeolian sand. Clearly channelled flow, lower overall flow rates, and perhaps diffusion-controlled storage and release in low-permeability lenses affect mass transfer in the latter unit.

Temporal variability occurs in a number of the studies where high frequency sampling has been conducted. In particular, both Trautz et al. (2013) and Cahill et al. (2014) observed a short-lived pulse in EC, alkalinity, Ca, Ba, Sr, Mg, and other elements, with values rising rapidly then decreasing afterwards to levels slightly higher than the original background; in contrast pH decreased and remained low. As both sites have little or no measured carbonates, the authors postulate that this pulse could be due to rapid release from the complete dissolution of the trace carbonates present or from pH-moderated desorption from, for example, oxi-hydroxides (Cahill et al., 2014; Newell et al., 2014). The subsequent, slightly higher than background values are likely controlled by the persistent, much lower pH values (due to exhaustion of the carbonate buffering capacity) and related, slower silicate dissolution reactions. Such temporal events illustrate the advantage of continuous monitoring systems (e.g. Peter et al., 2012).

In terms of impact on drinking water quality, aside from a limited number of exceptions, trace element concentrations remain well below national and international drinking water limits in all field experiments in which they were measured (Cahill and Jakobsen, 2013; Cahill et al., 2014; Kharaka et al., 2010; Mickler et al., 2013; Peter et al., 2012; Trautz et al., 2013b; Yang et al., 2013a,b). Exceptions included Al values over 10x the WHO guidelines (Cahill et al., 2014), two samples that exceeded Ni drinking water levels (Peter et al., 2012), and some pH, Fe, Mn and Zn values that exceeded secondary standard levels where effects are deemed to be cosmetic or aesthetic (Yang et al., 2014b). Toxic elements like As and Pb were often below the detection limit of the analytical methods used, however some authors stressed the low levels of these elements in their aquifer material.

3.4.3. Natural systems

The migration of deep-origin brines and/or geologically produced CO₂ into near-surface aquifers is a relatively common occurrence in geologically active areas throughout the world, with those associated with CO₂ often being exploited for sparkling drinking water. The study of such sites can yield valuable information regarding local to regional scale processes and complexities occurring in a real-world setting over long time periods.

In an effort to overcome the kinetic-related limitations of relatively brief injection tests, Beaubien et al. (2014a) performed detailed groundwater sampling through two different natural leaks in silicate and carbonate sediments, showing how pH and major and trace elements of the respective plumes were controlled by aquifer mineralogy. Another detailed study looked not directly at ground-water but rather at aquifer material through which CO₂-charged brines once flowed, causing dissolution of hematite grain coatings (bleaching) and concurrent mobilization of trace elements (Wigley et al., 2013). This work clearly illustrates plume zonation and potential mechanisms of trace metal attenuation, as species liberated in the bleached interval (such as Cu, Pb, Zn) were re-deposited over a 5-10 cm interval at the reaction front where sediment buffering capacity was still intact. The authors state that the observed attenuation mechanism could potentially be important in leak-impacted aquifers if flow rates are slow enough relative to reaction rates.

Other studies are of a more regional nature, sampling existing wells and/or springs over 100's of km² to examine large-scale processes, such as the control of faults as migration pathways or determining the origin of groundwater anomalies. One site that has received particular attention is an area of active CO₂ degassing located near Chimayó, New Mexico (Keating et al., 2010; 2013a,b; Viswanathan et al., 2013), where physical-chemical data from about 30 shallow (<60m) drinking water wells around the Robert’s Fault have been studied and modelled. Initial work used major and trace element concentrations and associations, major ion ratios (e.g. Br/Cl), and stable carbon isotopes to distinguish gas only leakage along the northern part of the fault from gas plus brackish water leakage to the south (Keating et al., 2010). The fact that elevated U, As, and Pb concentrations are associated with brine migration in the south led the authors to state that this mechanism may have a greater potential for impact compared to that of CO₂-induced in situ reactions in a buffered aquifer. Subsequent studies on this site focussed on the behaviour of anomalous trace elements, As and U, with the former being found to be controlled by sorption on clay and oxide surfaces (Viswanathan et al., 2013) and the latter by cation exchange and dissolution / precipitation of calcite (Keating et al., 2013a).

The mixing of upwelling brines with shallow fresh waters and the approaches used to determine groundwater origins and evolution have been addressed in various other regional studies. These issues are particularly important from the point of view of CCS monitoring, as natural variability and chemical changes during mixing and migration can complicate leak recognition. For example, Choi et al. (2012) showed how the mixing of a deep Na-HCO₃ water with a shallow fresh-water in the Kangwon District, Korea, created a third water that further evolved to a Ca-HCO₃ composition via reverse cation exchange. Kirk et al. (2009) studied the mixing of an upwelling deep brine with shallow groundwater in the Rio Grande floodplain, New Mexico, USA, showing that the major impact relates to redox mobilization of Mn and Fe, but that trace elements like As were sequestered in oxides. Keating et al. (2014) found evidence of simple mixing between an upwelling brine (with high Cl, As, and other trace metal concentration) and fresh water at Springfield, Arizona, USA, however δ¹³C isotope data off the mixing trend indicated that the associated CO₂ migrated in the dissolved phase rather than as bubbles. Isotopes, together with trace element analyses and modelling, were also used to show that there is, instead, no leakage from a natural CO₂ reservoir into the overlying fresh-water aquifers at the Montmiral site in France (Lions et al., 2014b). This study demonstrates that all DIC (Dissolved Inorganic Carbon) originates from oxidation of organic material in the aquifer and that the bulk of measured δ¹³C values were compatible with open and closed carbonate dissolution while a small sub-set of anomalous, enriched samples could be explained by incongruent dissolution of Mg–Sr–calcite or dolomite.

3.4.4. Geochemical modelling

The prediction of geochemical reactivity in space and time via modelling is crucial for the assessment of potential impacts of leakage from a CCS reservoir into an overlying shallow aquifer. This can take two main forms, batch geochemical modelling that addresses the complex, thermodynamically controlled reactions between water, gas, and rock, or reactive-transport modelling that couples the former with fluid flow. In particular, the coupling of both chemical and hydrodynamic processes is still a numerical challenge for the computational capacity of present-day simulators. Still, the integration of fluid displacement (due to either the regional flow or CO₂ injection) is needed to track plume development and to make recommendations regarding monitoring and detection techniques (Siirila et al., 2012).
Geochemical modelling has been applied thus far to assess monitoring options (Carroll et al., 2009) for the interpretation of observed changes at natural analogue sites (Keating et al., 2010; Keating et al., 2013a; Lions et al., 2014b; Viswanathan et al., 2013), for the prediction of potential impact on groundwater chemistry (Carroll et al., 2014; Humez et al., 2011; Navarre-Sitchler et al., 2013; Wilkin and DiGiulio, 2010), the fate of trace metals in shallow groundwater (Wang and Jaffe, 2004; Birkholzer et al., 2008; Zheng et al., 2009, 2012; Siirila et al., 2012), and to better understand the significance of microbially mediated reactions (West et al., 2011). It is also being used to interpret observed groundwater changes at CO₂ injection field experiments (Peter et al., 2012 Yang et al., 2013a,b; Zheng et al., 2012). For these studies, the most commonly applied models are TOUGHREACT (Xu et al., 2004) and PHREEQC (Parkhurst and Appelo, 1999).

While modelling has the advantage that it can be applied at any temporal or spatial scale, it is limited by the quality of the thermodynamic and kinetic data used and by the complexity of the natural systems modelled. For example, Fahrner et al. (2012) demonstrate that the numerical simulation of CO₂ intrusion in a freshwater aquifer needs to consider pressure and temperature effects. Pressure is important to assess parameters such as pH and inorganic carbon through the dissolution of CO₂(g), while temperature is important to correctly assess kinetically constrained mineral precipitation/dissolution and sorption processes. In particular, the authors state that present models may underestimate the in situ release of trace elements in shallow aquifers due to an imprecise correction of the temperature effect on sorption processes.

The accuracy of a modelled geochemical response of freshwater aquifers to CO₂ intrusion is also dependent on the adequacy of the developed model and the objectives of the study, and thus one can choose to model different levels of complexity. As the main pH control is provided by carbonates, several studies consider a simplified system using only calcite and aqueous carbonate species to predict impacts on pH and alkalinity in shallow groundwater (Carroll et al., 2014; Navarre-Sitchler et al., 2013; Yang et al., 2013a). In other cases, complex mineralogical characterization of geological formations has been obtained from core analyses (e.g. XRD, SEM-EDS) (Keating et al., 2010, 2013b; Viswanathan et al., 2013; Zheng et al., 2012, 2009). Such complex mineralogy allows one to predict impacts on water composition and the precipitation / dissolution of the main minerals (Humez et al., 2011; Zheng et al., 2012, 2009). In this case, incorporation of kinetic rates helps to describe the evolution of the system. However kinetic rates and associated parameters (specific surface area, temperature and pH dependency) given in the literature are highly variable for a number of minerals, such as silicates, and this can decrease the accuracy of long-term model predictions.

The fate of trace elements is controlled both by their source phases and by the processes involved in their mobility. Lead and arsenic have received the most attention due to their natural presence in US drinking waters and the relatively high simulated release compared with legislated maximum contaminant levels (MCLs) (Navarre-Sitchler et al., 2013; Siirila et al., 2012; Viswanathan et al., 2013; Wang and Jaffe, 2004; Zheng et al., 2009).

Regarding trace elements release, the first step is to identify the mineral hosts. Sulphides and carbonates are mainly considered in models (Keating et al., 2010; Navarre-Sitchler et al., 2013; Siirila et al., 2012; Wang and Jaffe, 2004) but sorption on clays and oxides may also affect trace element mobility as they can be a seal or a source (Humez et al., 2013; Viswanathan et al., 2013; Zheng et al., 2012, 2009). Although sorption processes (ion exchange and complexation surfaces) are not systematically implemented in all codes, they can significantly reduce metal mobility and thus should be considered. Ion exchange is important in defining the fate of major and trace ions, however this mechanism is not pH sensitive. Therefore surface complexation, which is pH dependent, is generally the only mechanism included in models (Navarre-Sitchler et al., 2013; Zheng et al., 2009). Ideally, both should be considered to correctly simulate all surface processes (Zheng et al., 2012; Viswanathan et al., 2013). Siirila et al. (2012) used a simplifying approach, assuming metal mobilization only at the leakage source and spatially constant linear retardation downstream from the source. This simplification allows for a quantitative approach to exposure risk assessment. These authors show that hydrologic flow and transport parameters (horizontal stratification, local scale dispersion) are also sensitive parameters in assessing overall risk.

3.5. Implications for monitoring

For aqueous geochemical methods to be successful in finding a leak various criteria must first be met, including that wells exist or are created, that a leakage-induced plume intersects one of those wells, and that it must be possible to recognize the chemical changes within the plume as resulting from leakage and not natural site variability.

Regardless of other options, existing drinking water wells will definitely be sampled due not only to cost effectiveness but also to safeguard public health. That said, Carroll et al. (2014) illustrate that the probability of defining a leakage-induced plume using exist-
Fig. 3. A model derived climatology of the pH range for two regions in the North Sea. Southern North Sea (reproduced from Blackford et al., 2015).

ing wells is extremely low. While a higher density of newly drilled observation wells would increase the likelihood of detection, the numbers required would be unrealistic. Instead, a limited number of wells could be drilled in areas having a higher risk of leakage (e.g., along fault traces or near abandoned wells) or potential impact (e.g., up-gradient from a major municipal water well), or could be drilled ad hoc should other techniques infer leakage in a specific area. Well completion issues will also affect monitoring: very long screened intervals may result in a diluted leakage signal, while monitoring the top of a confined aquifer may be better for CO2 leakage (Carroll et al., 2009; Fig. 2d) but the bottom may be better for dense brine leakage (Fig. 3b).

Potential impact from a leak will vary site specific, depending on leakage characteristics and potable aquifer characteristics. These factors must be taken into consideration to ensure that a leakage anomaly can be defined within a potentially complex background (Humez et al., 2014a). As a first step, pre-injection spatial/temporal characterization of the aquifer must be undertaken to determine mineralogy and the range of natural baseline values for aqueous geochemistry, in order to define appropriate monitoring parameters and thresholds (Carroll et al., 2014). Site-specific batch and reactive transport geochemical modelling can greatly aid in this work.

Based on these initial results, a number of parameters could be chosen for monitoring. Those related to the carbonate system (pCO2, pH, DIC, alkalinity) are the most commonly recommended, however the specific choice will be dependent on aquifer mineralogy. In carbonate rich aquifers pH changes will be limited due to buffering but alkalinity and DIC will increase significantly with the dissolution of calcite or dolomite, whereas in carbonate poor aquifers pH will decrease sharply. Thus authors have suggested the use of DIC (Romanek et al., 2012b), DIC and pH (Trautz et al., 2013), or DIC, pH, and δ13CDIC (Yang et al., 2013a,b), while Yang et al. (2014b) used modelling results to suggest that pCO2 and DIC may be the best compromise in all mineralogical settings.

Some major ions, such as Ca, Mg, Fe, and Mn, have shown CO2 induced changes in controlled injection experiments (Cahill et al., 2014; Kharaka et al., 2010; Trautz et al., 2013), however their reactivity means that data interpretation is problematic. Similarly, most trace elements of interest from a health perspective, such as As, Pb, or U, should be avoided for leakage monitoring due to scavenging and attenuation (Keating et al., 2013a), although they may be measured for public health reasons. EC has been observed to travel at the groundwater velocity in carbonate poor settings (Cahill et al., 2014; Trautz et al., 2013), however aside from providing an early warning signal it was found to subsequently drop probably due to the transformation of HCO3− to H2CO3 at lower pH values.

Considering the transient changes observed in numerous studies, continuous monitoring may be helpful to recognise leakage and to give early warning. Various sensors have been used or are being developed for long-term deployment and monitoring of pH, pCO2, EC, etc. (Gal et al., 2012; Graziani et al., 2014; Peter et al., 2012; Yang et al., 2014a). Isotopes are another powerful tool which can help distinguish the source of an anomaly and give important information about flow-path reactions (Humez et al., 2014a). The stable isotopes of C (Schulz et al., 2012), C and Sr (Lions et al., 2014a; Newell et al., 2014), and B, Li, Sr, O, H, and S (Humez et al., 2014b) have all been used to help interpret complex hydrochemical data.

In the case of brine leakage, original aquifer chemistry may have less of an impact on monitoring strategies while pH values may not be affected (Wunsch et al., 2013). Instead, brines contain high concentrations of Cl and Br, and because these anions are conservative they will migrate at the groundwater velocity (potentially giving an early warning) while their ratio can be used to define their source. Other elements, such as B and Li, are also typical brine tracers while the high salinity means that EC could be used as a simple, robust monitoring tool in wells.

Considering the difficulty of finding groundwater impacted plumes using water well sampling only, other methods have been suggested to improve the chances of success. In confined aquifers a leak may result in a pressure pulse that can travel much farther and more uniformly than the impacted plume, giving early warning of an anomalous event (Wiese et al., 2013). Soil gas or atmospheric monitoring can cover large areas in much more detail than groundwater sampling (e.g. Beaubien et al., 2013), and thus discovery of a leak using these methods would focus groundwater sampling at that location. Finally geophysical methods show promise for imaging the shallow sub-surface over larger areas to look for leakage-induced plumes. For example the higher conductivity associated with such plumes, especially those from brine leakage, can potentially be delineated using electrical resistivity methods (Pettinelli et al., 2010; Wagner et al., 2013), although there is a risk of false positives (Trainor-Guitton et al., 2013).

3.6. Significant gaps requiring further research

The large amount of research conducted over the last 10 years has significantly improved our understanding of the potential impacts of CCS-related leakage on groundwater quality. Starting from relatively simple scenarios, work has evolved to better assess and quantify ever more complex systems in terms of temporal and spatial scale, kinetics, complex chemical mechanisms, aquifer heterogeneities, etc. Future work must continue on this path, focussing in particular on integrated laboratory, field, and modelling studies that characterise a wider range of representative aquifer types and leakage styles.

To date a large number of laboratory batch experiments have been performed. While this approach is useful, the often reported discrepancy between lab and field reaction rates implies that batch experiments do not mimic natural systems. To address this it is recommended that more 1D column experiments be conducted, preferably on un-disturbed core, to obtain more realistic rock ratios and reactive surface areas. Such experiments have great potential to study a wide range of issues (e.g., scavenging and attenuation processes along the flow path) and are particularly well suited to coupling with reactive transport modelling. Further laboratory experiments should also address different aquifer
mineralogy, plume chemistry (e.g. redox, pH, and composition), temperature, etc.

Recent field injection experiments have yielded important results that address natural complexity and large scale processes; similar experiments are needed to study different types of aquifers, depths, flow rates, and leakage rates. In addition, field experiments involving the injection of brine or CO2 with impurities (e.g. CH4, H2S, SO2) could give valuable information regarding geochemical reaction pathways, element mobility and attenuation, and eventual impacts in these different scenarios.

The various studies conducted thus far on natural analogues have shown that drinking water limits are typically not exceeded at these sites. More such studies are required, however, in different geological settings to ensure that such preliminary results are valid under a wider range of conditions (mineralogy, flow rates, leakage rates, etc.). In addition, it is recommended that additional detailed studies be conducted on individual leakage points, as this will be similar to injection experiments in terms of scale and natural complexity but on a system that has been leaking for much longer time periods. Non-leaking sites, including aquifers above potential storage sites, need to be studied as well to better define natural background variability and, thus, thresholds for leakage detection. To improve this threshold, “fingerprinting” tools like isotopes should be better tested.

Finally, more integrated experimental/modelling studies are required to improve numerical tools. The coupling of multiphase flow with reactive transport modelling is very demanding in CPU time and memory, which often requires simplification of the proposed scenarios (e.g. single phase approach, 2D or 1D modelling, homogeneous properties, reduction of chemical processes and data, etc.). Although high performance computing is becoming an important support that allows for the simulation of ever more complex 3D heterogeneous systems, such models still need to be validated against real field data. Model calibration and the reduction of uncertainties in the associated thermodynamic and kinetic data bases is therefore crucial to ensure continued progress in modelling prediction. We also highlight the need to model the potential impact of impurities, something that has received little attention to date.

4. Offshore ecosystem impacts

4.1. Introduction

The IPCC report of 2005 primarily discussed marine environmental impacts in the context of open ocean storage. Since that time, with the effective exclusion of ocean storage as an option, attention has turned to environmental impacts associated with leakage of CO2 from geological storage or transportation. Also around 2005 a huge international research effort started on the impacts of ocean acidification (OA) on marine systems as a result of rising atmospheric CO2 levels. OA is a long term projected reduction of global marine pH by up to 0.7 units over the next decade (Caldeira and Wickett, 2003) which is hypothesised to be potentially damaging for marine systems. This has been highly informative for CCS issues. In the last decade the most striking outcome is the revelation of significant complexity in the environmental impact at the organism and community level, such that the hoped for ecosystem sensitivity synthesis is only just beginning to emerge (Section 4.3.5).

The physiological mechanisms by which excess CO2 affects marine organisms identified a decade ago (Pörtner et al., 2004, 2005) still hold. Briefly, additional CO2 in marine systems affects carbonate chemistry, increasing acidification and bicarbonate ion concentration and decreasing carbonate ion concentration and pH. The level of acidity controls many biochemical processes and both bicarbonate and carbonate are substrates for key biological processes. More recent, however, is an understanding of the role of (nutritional) resources in mediating impact (e.g. Thomsen et al., 2013) and the ability of some species to acclimatise to altered chemistry, at least in the short term (e.g. Schlüter et al., 2014) or even exhibit adaptation over longer periods (e.g. Pespeni et al., 2013).

Whilst there is still much debate regarding the likelihood and potential rates of leakage, the scientific community has made advances in understanding the spatial impact from a range of hypothetical leakage scenarios (section 4.3.1), via models. It can be argued that this understanding is vital, as a first order impact assessment needs to quantify the proportion of a system that is affected.

Another advance is a far better understanding of the natural physical/chemical variability of the marine system (e.g. Hofmann et al., 2011; Thomas et al., 2005; Artioli et al., 2012), especially with respect to carbonate chemistry in coastal and shelf seas which will predominantly host storage sites. This is crucial for both monitoring for leakage signals but also for understanding the resilience of populations to chemical changes. As a consequence it is now possible to design the baseline acquisition strategies that would be necessary for both detection and impact assessment (Blackford et al., 2015).

4.2. Baselines

The fundamental chemistry of dissolved CO2 in seawater is well understood (Zeebe and Wolf-Gladrow, 2001) but is latterly better characterised with respect to coastal environments and CO2 seepage. Although seeping CO2 is generally thought likely to appear at the sea floor in bubble form (or liquid droplet in cold, deep environments) direct seepage in the dissolved phase cannot be ruled out. Gas bubble plumes, however buoyant, will dissolve rapidly such that they are unlikely to penetrate more than a few metres from the sea floor (Dewar et al., 2013). Consequently all forms of seepage are likely to form plumes of dissolved CO2 in the bottom few metres of the water column and mixing processes are likely to disperse CO2 rapidly both horizontally and vertically. Hence the focus for impact research is on sea floor or benthic ecosystems.

Prior to 2005, less was understood about the variability of carbonate chemistry in coastal and shelf seas, which are the location for the majority of offshore carbon storage projects in operation or planning. It has now been established that marginal marine ecosystems have highly variable carbonate chemistry which arises from complex interactions of biological, terrestrial and oceanic processes (Hofmann et al., 2012). Consequently marine organisms are adapted to a range of pH conditions and variability. Understanding this complex natural variability, the factors which drive it, how it influences biological processes and what constitutes a deviation from this natural range are key to understanding impact potential.

Oceanic CO2 equilibrates with atmospheric CO2 over timescales of the order of a month, consequently marine CO2 is showing a gradual increase, and pH a gradual decrease, in line with increasing atmospheric CO2. In contrast, the influences of biological processes on CO2 and the carbonate system, are far more dynamic. Uptake of Dissolved Inorganic Carbon (DIC) via photosynthesis is a feature of nutrient-rich well lit waters and occurs on relatively predictable seasonal and diurnal cycles. The subsequent respiration of organic matter by heterotrophs releases DIC throughout the water column and within the surface sediment layer. Predator – prey interactions create less predictable sub-seasonal dynamics, which are exacerbated by the availability of nutrients, mediated by both biological regeneration and physical mixing. In coastal regions influenced by rivers, terrestrially derived nutrient loads fertilise biological production leading to large swings in DIC concentration, and hence pH, which may range over 1 unit or more. Away from direct coastal
influences, in regions of more oceanic water, variability over an annual cycle may be 0.3 pH units or less. In deeper, sometimes stratified, water columns the lower layers tend to exhibit less variability due to the lack of in-situ photosynthesis in the darkness. However shallower water columns that are well mixed can show significant variability at the sea floor. Whilst the surface ocean is relatively well sampled for carbonate chemistry, there are a lack of high frequency time series and a dearth of observations near the sea bed. However hydrodynamic-ecosystem models, which encode all of the processes responsible for natural DIC variability, described above, can reveal, with reasonable skill (e.g. Artioli et al., 2012), the spatio-temporal dynamics of relevant marine systems (e.g. Fig. 3).

Within sediments very strong gradients of carbonate chemistry exist, typically with sediment layers a few centimetres deep exhibiting pH levels about 0.5 units or more below that of the overlying seawater. This is due to the vertical profile of redox chemistry and infaunal activity, including microbial respiration. An example sediment pH image is illustrated in Fig. 4. There is however much less seasonality within the sediment carbonate system, although there is significant spatial variability associated with different sediment types and on a smaller scale associated with burrows and similar structures created by benthic organisms (Stahl et al., 2006).

Moving beyond carbonate chemistry, the physical, chemical and resource environment within which marine organisms live varies considerably both through space and time. Consequently it is hardly surprising that marine communities, and the processes mediated by these communities, exhibit large natural variability expressed as a temporal and spatial patchwork of communities and life-cycles. This is driven by short term seasonal dynamics (Reiss and Kröncke, 2005), spatial changes associated with gradients (e.g. sediment type, food availability, Dauwe et al., 1998), long-term change associated with environmental drivers (e.g. ocean warming, Hiscock et al., 2004; OA, Kroeker et al., 2013), and direct human induced change (e.g. fisheries, Queirós et al., 2006), the characterisation of which develops continually. The conclusion from recent CCS related studies is that a baseline from which to assess impact must account for this heterogeneity in both community structure and the underlying chemical environment (Section 4.4, Blackford et al., 2015).

4.3. Research developments since 2005:

Offshore environmental research relating to CCS is challenging, given that the major focus is on the relatively inaccessible sea floor. Traditional laboratory based experiments, where individual species or small sediment samples are exposed to a range of concentrations of CO₂ enable testing of detailed physiological and limited behavioural responses. Whilst some such experiments have been specifically targeted at CCS issues, a huge body of research surrounding the chemically similar phenomena of OA has provided much information relevant to CCS impact assessment. Significant research has used natural analogues, generally sub-sea volcanically derived CO₂ seeps, which allows an investigation of whole system effects along with an assessment of impact footprint; although these can be slightly undermined by the thermal characteristics of, and impurities in, the gas flow. More recently direct injections into the water column and sediments build on analogue studies by enabling a study of both onset and recovery associated with known rates of CO₂ injection. Finally modelling has enabled the prediction of the spatial and chemical extent of CO₂ plumes for a wide range of leakage rates and are being used in ecological impact quantification studies, although the latter are in their infancy.

4.3.1. Modelling

Given that leakage is hypothetical, and that there is minimal data that directly provides evidence of leakage characteristics, modelling research over the last 10 years has played a significant role in advancing understanding. Primarily, models have quantified the spatial and temporal dynamics of a range of leakage scenarios (Chen et al., 2005; Blackford et al., 2008, 2009, 2013; Dewar et al., 2013; Phelps et al., 2015), initially to understand impact potential but increasingly to address detectability and monitoring strategies (Hvidevold et al., 2015; Greenwood et al., 2015 in press). Models can also be used to aid leakage quantification (Mori et al., 2015), to underpin baseline understanding and in the future to directly assess impacts on ecosystems, as has been done for OA (Artioli et al., 2014). Model based research is complex and computationally demanding, generally requiring parallel machines or large clusters of computers, and requires a suite of model components:

Hydrodynamic models which characterise the 3D movement and mixing of marine systems are the key component for understanding the dispersal of dissolved CO₂. Realistic atmospheric, tidal and geostrophic forcing are essential in order to correctly estimate dispersion characteristics. Initial studies of leakage used available models, often with a relatively coarse resolution (e.g. Blackford et al., 2008, ~7 km horizontal resolution), and were only able to address large scale leakage events. In the last decade the resolution of hydrodynamic models has improved, as a result of advances in computational systems. Shelf wide models can now reach resolutions of 1 km horizontally (e.g. Phelps et al., 2015) whilst local models can resolve at least part of the domain with resolutions of a few metres (e.g. Blackford et al., 2013).

Bubble plume models (Chen et al., 2005; Dewar et al., 2013) are necessary to properly understand the characteristics of the leak epicentre, in terms of the gas phase plume and the near field dissolved plume. Bubble size (larger bubbles are more buoyant and dissolve slower) is the key determinant of the elevation of a plume from the sea floor and consequently the vertical profile of chemical change and the patterns of dispersion; with implications for both monitoring strategy and environmental impact. As with the hydrodynamic models, emission form, topology and currents impart a large variability on the outcome of a given emission scenario. Plume models have made progressive improvements to the parameterisation of processes, based on observations of natural and man-made leakage analogues (Dewar et al., 2015).

Carbonate system models are an essential component of all leakage simulations as they can derive pH, pCO₂, CO₃²⁻ and HCO₃⁻ ion concentration and saturation state from given concentrations of dissolved CO₂. These parameters are necessary to understand both impact and detectability. Whilst carbonate system models have been available for decades, since 2005 international agree-
ment on the parameterisation of reaction constants (Dickson et al., 2007) and a far better treatment of alkalinity (e.g. Artioli et al., 2012) has improved the realism of these models, especially when applied to shelf and coastal systems.

**Ecosystem models** such as ERSEM (Blackford et al., 2004), which resolve the marine carbon cycle, have a dual purpose when coupled with hydrodynamic and carbonate system models. By including processes, such as photosynthesis and respiration, which affect natural DIC concentrations, these models provide an ability to estimate natural variability (Artioli et al., 2012) or the baseline against which leakage must be detected (Blackford et al., 2015) (Fig. 3). This is particularly important because there are very few direct observations of the sea floor carbonate system. Ecosystem models have also been used to examine biological impacts from high CO2 (e.g. Artioli et al., 2014), however as the organism response, as demonstrated by a wealth of experiments (Widdicombe et al., 2013a) is complex and species specific, this is not yet well resolved by model systems. In the last few years models are attempting projections of impact, but these are not yet sufficiently developed for publication.

Because of the large range of hypothetical seepage scenarios considered (<0.1–10000 tonnes d\(^{-1}\)) no one model system has an over-arching forecast capability, rather various combinations of modelling components are used as appropriate. For small scale leakage scenarios (<1 tonnes d\(^{-1}\)) a detailed treatment of the bubble plume is the primary modelling requirement. For larger scenarios (>10 tonnes d\(^{-1}\)) the length scale of the gas phase plume is significantly less than that of the dissolved plume and hydrodynamic models are the primary model component. Choosing an appropriate resolution for the physical structure of the model depends on the length scale of the modelled plumes.

Both modelling and experimental work indicate that the maximum length scale of potentially biologically harmful CO2 plumes increases with the leakage rate, being of the order of 10 metres for emission rates around 0.1 tonnes d\(^{-1}\) (Shitashima et al., 2015), of the order of a few hundred metres for 10 tonnes d\(^{-1}\) (Dewar et al., 2013) and around 50 km for 10,000 tonnes d\(^{-1}\) (Phelps et al., 2015).

What actually constitutes a biologically harmful exposure to high CO2 is complex and is discussed below.

Modelling has shown two key things; firstly that the dynamics of leaked plumes can be very complex, for example oscillating around an epicentre due to tidally driven mixing (Blackford et al., 2013). As well as making detection more complex, the resulting on-off exposure of fixed benthic communities to high CO2 water is likely to have implications for their response and poses challenges for experimental design. Secondly, models suggest that the evolution of leakage plumes is highly dependent on a range of factors, including the spatial pattern and form of sea floor emissions, the emission rate, local topography, and seasonal and weather driven conditions. Consequently leakage projections are not transferable from one storage site to another or between seasons.

**4.3.2. Laboratory experiments**

In the past few years there have been several major changes to the way in which laboratory experiments are used, and will be used in the future, to assess the potential environmental impacts of CO2 leakage from sub-seabed CCS reservoirs.

For many years standard ecotoxicological studies have used laboratory based exposure experiments to assess the likely impacts of environmental toxicants on marine organisms or communities. The major benefit of this approach has been that the scale, longevity and nature of any exposure could be tightly controlled, and any confounding effects from other factors minimised. Unsurprisingly, therefore, the recognition of CCS as a potential environmental threat caused a rapid growth in such studies with new data produced describing the effect of CO2 exposure on a range of marine species and hence the potential impacts of CCS leakage (see Widdicombe et al., 2015). However, unlike many toxicants, CO2 is an element to which marine organisms are already exposed; many have strong physiological responses to CO2 or to other, CO2-related, chemical changes (e.g. pH or carbonate saturation state). Consequently, CO2 exposure experiments have needed to evolve beyond the traditional assessment of simplistic estimates of mortality or lethal limits (e.g. LCSV values). In particular, it is essential for studies not to be limited to a few traditional response variables but instead to consider the likely trade-offs and compensatory mechanisms used by marine organisms to cope with short-term variations in the levels of environmental CO2. This represents perhaps the biggest advance in laboratory-based CCS experiments with new studies incorporating elements of physiological plasticity (Ellis et al., 2015) and energy allocation (e.g. scope for growth or dynamic energy budget modelling) (Kloek et al., 2014).

CCS laboratory-based research has moved from simple, single stressor exposure experiments, measuring a limited number of response variables under a constant level of CO2 stress, to more complex, multi-endpoint studies which consider a host of potential interacting factors. The relevance of physiological mechanisms that allow certain species to cope with high levels of CO2 (hypercapnia) was identified and differences between species in this respect were revealed (Kamenos et al., 2013; Hammer and Pedersen, 2013; Donohue et al., 2012; Small et al., 2010). More subtle effects of elevated CO2 concentrations were shown in experiments with multiple stressors, such as pathogens (Ellis et al., 2015). Experimentation in more complex systems also revealed indirect impacts on marine ecosystem functions such as bioturbation and nutrient cycling (Burdett et al., 2014; Widdicombe et al., 2013b; Murray et al., 2013; Widdicombe and Needham, 2007).

In addition, studies are now beginning to consider the potential for interactions between CO2 and other stressors (e.g. toxic metals) as well as the potential for CCS activities to facilitate the release of other potentially harmful fluids such as methane or reservoir/formation water (hypoxic brine). Early consideration of the latter (Queirós et al., 2014) suggests that they may cause greater changes in marine ecosystem components than leaking CO2.

Many experimental approaches have been deployed to investigate the potential consequences of OA, but because of practical limitations in experimental procedures and the desire to establish response across a gradient of pH values, many OA experiments have used relatively short term, sudden onset exposures to CO2 and have explored reductions in pH exceeding 1 pH unit. Such experiments are therefore arguably better analogues for CCS leakage than OA, although only they only consider increased CO2 in the water column and do not address the complex chemical transformations and biological impacts that may occur when CO2 flows vertically through sediments. It is beyond the scope of this paper to review the literature on OA, currently growing by several hundred papers a year, and periodically summarised (e.g. Gattuso and Hansson, 2011). In synopsis this research has revealed a large range in response to lowered pH which depends on species, class, life-stage, resource level and other stressors, nevertheless has contributed significantly to our understanding of the ecological sensitivities of the marine system to leakage from geological storage (see Section 4.3.5).

**4.3.3. Natural systems**

Natural CO2 seeps that occur over a wide depth range, from intertidal to the abyss, can be studied as a large-scale, real-world analogue of what might occur at a leaking offshore CCS site in order to achieve a more complete (and realistic) understanding of the possible consequences of a seabed leak of stored CO2.

In the last few years, in the framework of several CCS-related research projects, natural CO2 gas seeps have been widely studied to validate methods used for the detection and monitoring of CO2 leakage from geological storage. Int. J. Greenhouse Gas Control (2015), http://dx.doi.org/10.1016/j.ijggc.2015.05.032
migration, to perform in situ investigations into a number of key processes controlling leakage pathways, and to evaluate impacts on marine ecosystems.

For example, research into the potential effects of CO₂ leakage on benthic and pelagic systems was performed near Panarea Island (Southern Tyrrhenian Sea, Italy) within the EU-funded CO₂GeoNet, RISCS and EC02 projects, as well as at other EC02 sites like the Southern Okinawa Trough (NE off Taiwan), Jan Mayen Vent Field (North Atlantic) and Just Salt Dome (Southern North Sea). Russell et al. (2013) used three separate tropical volcanic CO₂ seeps in Papua New Guinea as natural "laboratories" to assess whether the observed increase of seagrass productivity and biomass due to locally increased CO₂ concentrations could be used as a proxy for what may occur near a CCS leak. The purpose was to assess if increased benthic primary productivity could capture and store CO₂ leakage in areas targeted for CCS. Karuza et al. (2012) investigated the interrelationships between planktonic prokaryotes and viruses in the Panarea hydrothermal system; de Beer et al. (2013) focused on biogeochemical processes and microbial activity in deep sea sediments of the Yonaguni Knoll IV hydrothermal system (Japan).

Volcanic CO₂ vents have also been extensively utilized to investigate the potential long term impacts of greenhouse gas-induced OA on marine organisms and ecosystems. Over recent years, CO₂ vent sites in areas off Ischia Island (Italy) have been explored to evaluate the effects of acidification on benthic organisms and ecosystems at shallow coastal sites and to examine ecological tipping points along gradients of increasing pCO₂ (Hall-Spencer et al., 2008; Hall-Spencer and Rodolfo-Metalpa, 2009; Cigliano et al., 2010; Dias et al., 2010; Martin et al., 2008; Rodolfo-Metalpa et al., 2010; Rodolfo-Metalpa et al., 2010, 2011; Kroeker et al., 2011; Porzio et al., 2011). Studies into the response of seagrass ecosystems to long-term high CO₂ were performed at a shallow volcanic CO₂ vent near Vulcano Island (Italy) (Apostolaki et al., 2014). Limpet shell modifications have been observed at intertidal hydrothermal vents in the Azores Archipelago (Couto et al., 2012). The consequences of long-term exposure to high CO₂ on coral-reef-associated macroinvertebrate communities were analysed around three shallow volcanic CO₂ seeps in Papua New Guinea (Fabricius et al., 2011, 2013) while calcifying coral abundance near low-pH springs were studied off the coast of Mexico near the Mesoamerican Reef (Crook et al., 2012).

Outcomes from these studies should help us understand the potential impact of CCS leakage because the magnitude, duration, and localized gradients of CO₂ exposure are more likely to mimic industrial CO₂ leakage scenarios than the more gradual but globally elevated CO₂ associated with OA (Carroll et al., 2013).

Organisms and ecosystem responses in naturally acidified environments depend upon site characteristics and species, which vary in relation to water depth and light penetration, to the strength of the leakage and its persistence, to the hydrodynamic conditions at the site and other environmental complexities. Many of these vent sites show a wide range of pH values (Kroeker et al., 2011) and temporal and spatial variations in emission rates and chemical composition. Such different natural conditions make it possible to assess a wider range of potential long-term (multi-decadal) impacts on entire marine ecosystems and provides a good opportunity to document species-species and species-environment interactions under low pH conditions. Moreover, in situ transplantation and colonization experiments along pH gradients can be designed at natural CO₂ vent sites to define species sensitivity to increasing pCO₂ values.

Results obtained from studies performed at these natural laboratories (within CCS or OA applications) show that the effects of increasing CO₂ levels on marine organisms are variable and species-specific but can also have cascading effects on communities and ecosystems (Carroll et al., 2013). Typically, species diversity decreases near CO₂ vents.

At shallow-water hydrothermal vents (off Panarea, Italy), Karuza et al. (2012) observed a lower abundance of free viruses and lower virus-to-prokaryote ratio in close proximity to the vents, suggesting that prokaryotes were more tolerant of the CO₂. Manini et al. (2008) had also found viruses to be more affected by the CO₂ and found higher benthic bacterial ribotype richness and diversity near the vents, highlighting a positive effect of the fluid emissions on prokaryote diversity and suggesting that the vents might influence benthic prokaryote community composition.

In the same area, transplantation experiments to assess the short to mid-term (1 year) CO₂ leakage effect on bacterial assemblages indicate that CO₂ leakage causes a shift in bacterial diversity (Queirós et al., 2014). Regarding the planktonic community, preliminary results suggest that CO₂ seeps do not seem to have any clear influence on phytoplankton and microzooplankton communities, whose structure appeared instead to be mainly linked to seasonal variations (Queirós et al., 2014).

Among primary producers, many seaweeds can clearly continue to photosynthesise even at extremely high dissolved CO₂ levels. Various Rhodophyta, Ochrophyta and Chlorophyta are able to grow at high dissolved CO₂ levels, even if there is evidence of decreased reproductive capacity in some species (Porzio et al., 2011). Seagrasses seem to benefit from high CO₂ conditions, increasing productivity (Cymodocea serrulata, Halophila ovalis and Cymodocea nodosa) (Russell et al., 2013; Apostolaki et al., 2014) and biomass (only C. serrulata); this may positively influence the diverse range of organisms which rely on these primary habitats, thus partially mitigating the effect of CO₂ increase (Russell et al., 2013). Negative effects of increasing CO₂ concentrations on habitat forming species has been reported instead for reef building corals (Fabricius et al., 2011; Crook et al., 2012), with reductions in coral diversity, recruitment and abundances of structurally complex framework builders and shifts in competitive interactions between taxa. This loss of habitat complexity is also associated with losses in many macroinvertebrate groups, especially predation-prone mobile taxa, including crustaceans and crinoids (Fabricius et al., 2013). Other calcified species adversely affected by reductions in pH are calcified algae, snails, sea urchins (Hall-Spencer et al., 2008; Hall-Spencer and Rodolfo-Metalpa, 2009) and calcareous seagrass epibions such as epiphytic coralline algae (Martin et al., 2008). Transplant experiments show that, at moderate temperatures, the bryozoan Myriapora truncata are able to up-regulate their calcification rates and survive in areas with higher levels of pCO₂, although this ability breaks down below a mean pH value of 7.4 (Rodolfo-Metalpa et al., 2010). Also invertebrate taxonomic richness appears to be reduced in extremely low pH zones (below 7.8; Cigliano et al., 2010; Kroeker et al., 2011; Dias et al., 2010). In general it appears that increased levels of CO₂ can profoundly affect the settlement of a wide range of benthic organisms, thus decreasing the diversity, biomass, and trophic complexity of benthic marine communities.

Deep-sea organisms may be particularly sensitive to CO₂ perturbations compared to their shallow-water counterparts due to lower metabolic rates, and variability in natural CO₂ concentrations and pH in the deep-sea compared to shallow water (Carroll et al., 2013). Results from natural deep-sea CO₂ seepage areas demonstrate that deep-sea benthic communities may fundamentally change in response to the release of CO₂ in the context of CCS (Neumann, 2012). Hypercapnia associated with CO₂ leakage may profoundly affect the physiology of deep-sea animals (Carroll et al., 2013). Combined effects of the release of other, maybe toxic, fluids may further alter community composition and functioning at the seafloor. Microbes might be able to survive at low pH values but most likely will suffer mortality as a consequence of the dissolving and uncoupling characteristics of liquid and supercritical.
CO₂ (Neumann, 2012). This limits life to the surface sediment horizons above the liquid CO₂ phase where less extreme conditions prevail (de Beer et al., 2013). In general, the number of microbes decreased sharply with increasing sediment depth and CO₂ concentration. Some members of the anaerobic methanotrophs and sulphate reducers can adapt to the CO₂–seep sedimentary environment; however, CO₂ and pH in the deep-sea sediments were found to severely impact the activity and structure of the microbial community (Yanagawa et al., 2013). Looking at the higher trophic levels, the macrofauna appeared to be more impacted than the meiofauna (Neumann, 2012). In the Okinawa Trough, the area of the Yonaguni Knoll, where liquid CO₂ is being released and CO₂ hydrates are forming, background fauna found at ambient CO₂ levels are completely absent in areas of high CO₂ and diversity of infauna is reduced. Echinoderms are present in the area but in high CO₂ conditions they are replaced by organisms associated with CO₂ vents (Wallman, 2011).

4.3.4. Injection tests

Whilst natural analogue sites and laboratory based work have been extremely valuable, they can have significant limitations. Analogue sites are often too well established, with an ecosystem that has adapted to higher CO₂ levels, to reveal initial impacts. Also, they may be compromised by an environmental setting, and chemical and/or thermal characteristics, that differ significantly from those above a CO₂ storage site. In laboratory work it is difficult to recreate the complex vertical structure of sediments or a realistic community structure, or address behavioural responses or recovery processes, all of which are key to understanding impacts. Consequently direct injection of limited amounts of CO₂ into the marine environment are an attractive if challenging research mechanism, with early work focussing on deep-sea sequestration (Brewer et al., 1999; Tamburri et al., 2000; Thistle et al., 2006). There have also been small-scale in situ injection experiments using benthic chamber landers (Ishida et al., 2013 and the RISCS project).

In order to better replicate conditions similar to the initial stages of a small leak from geological storage the recent QICS experiment (Blackford et al., 2014; Taylor et al., 2015a) engineered a direct injection of CO₂ into marine sediments in a sea bay and observed the physical, chemical and biological impacts. Although expensive and challenging, in terms of both execution and public acceptance, this approach has many benefits. It reveals the interplay between CO₂ flow and the complex physical, chemical and biological structure of the sediment, it allows an assessment of the hydrodynamic effect on CO₂ plumes and it demonstrates a biological response in the context of ecosystem interactions, behaviour and recovery. As in some terrestrial controlled release experiments (Jones et al., 2014) there was a considerable lateral displacement of CO₂ from the sub-sea level release point (up to 20m horizontally in 11m vertical flow). There were also complex interactions and differential flow regimes in each of the sediment layers present (Blackford et al., 2015; Cevatoglu et al., in press). A considerable proportion of the injected CO₂ was buffered by rapid dissolution of calcium carbonate (CaCO₃) in the sediments such that acidification of the pore water was limited in places (Lichtschlag et al., 2015), indicated by increased pore water alkalinity and Ca²⁺. Attempts at achieving a mass balance of the QICS CO₂ flow suggest that between 85% (Blackford et al., 2015) and 50% (Mori et al., 2015) of the injected CO₂ was retained within the shallow sediments, by a combination of dissolution and residual trapping.

Reductions in pH of up to 1.0pH units were seen in parts of the sediment and the overlying water column, with larger reductions associated with specific bubble plumes (Taylor et al., 2015b; Shitashima et al., 2015). There was considerable heterogeneity in chemical changes to both the water column and sediments with significant implications for detection, quantification and impact. Impact on benthic macrofauna, in terms of both number of species (biodiversity) and number of individuals (mortality and emigration) was significant towards the end of the 37 day injection period, but recovery of both chemical parameters and biological communities was rapid, within a few weeks (Blackford et al., 2015; Widdicombe et al., 2015). No impacts were seen on caged surface megafauna near the release site (Pratt et al., 2015) nor on the behaviour of mobile surface megafauna over the release site (Kita et al., 2015) indicating some short term resilience in the selected species for this limited release. Microbial gene expression seemed to be the most sensitive biologically related parameter, with significant but small impacts up to 25m from the release epicentre (Tait et al., 2015) and slight disruption to the microbial mediated nitrogen cycle (Watanabe et al., 2015). The QICS study concluded that a small, short-term leak was unlikely to have substantial impact on the local ecosystem.

4.3.5. Summary of advances in impact understanding

Taking together the insights from modelling, laboratory, analogue, injection and OA research, the understanding of sensitivity to high CO₂ has come a long way since 2005 but gaps, uncertainties and contradictions remain. A seep or leak of CO₂ from geological storage will produce two types of impact zone. Within the area of leakage, CO₂ will enter the marine ecosystem from below, via the sediments and the first to be exposed will be the deeper sediment dwelling fauna. Away from the area of leakage, impacts will be mediated via mobile plumes of CO₂ rich water permeating the water column, which will initially impact sediment surface and pelagic communities. All leak events will produce a gradient of pH and other chemical change between the leak location and the periphery of the affected area. The length scale of the gradient will depend primarily on the leakage rate but will be influenced by other factors associated with the form of leakage and hydrodynamic mixing.

Whilst models show that no two leakage scenarios will produce the same outcome, the area impacted increases with the leak rate. For small leaks, (<1t/d) the spatial impact will be of the order of a few tens of meters radius and arguably not significant on a regional scale, especially when compared to the likely impacts of climate change (Dewar et al., 2013; Blackford et al., 2008, 2013; Phelps et al., 2015). Only a very large leak (>100t/d), would seem to have significant impact potential, in terms of areal footprint (km scale), and the likelihood and impact of such a leak would be greatly reduced through good design and operational practice, including rigorous monitoring procedures.

The potential physiological impacts of CO₂ exposure on the health, behaviour, function and ultimate survival of marine species and communities have been intensively studied and have been detailed in a number of previous reviews (Seibel and Walsh, 2001, 2003; Pörtner et al., 2004, 2005; Fabry et al., 2008; Widdicombe and Spicer, 2008; Kroeker et al., 2013). In summary, when marine organisms are exposed to low pH seawater the primary physiological effect is a decrease in the pH or an “acidosis” of the extracellular body fluids such as blood, haemolymph, or coelomic fluid. In some species this extracellular acidosis is fully compensated by two mechanisms. The concentration of extracellular bicarbonate can be increased by either active ion transport processes in the gills or through passive dissolution of a calcium carbonate shell or carapace (see Widdicombe and Spicer, 2008 and references therein). However, in other species from a variety of different taxa, such as mussels (Michaelidis et al., 2005), crabs (Wood and Cameron, 1985; Pane and Barry, 2007) and sea urchins (Miles et al., 2007) studies have reported only partial, or no, compensation in the extracellular acid-base balance. In these instances the uncompensated acidosis can lead to more or less severe metabolic depression in the affected

Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO₂ leakage from geological storage. Int. J. Greenhouse Gas Control (2015), http://dx.doi.org/10.1016/j.ijggc.2015.05.032
organism (Pörtner, 2008) in turn having a negative impact on that individual’s contribution to the ecosystem.

Perturbations in an organism’s acid base physiology represent one potential impact of elevated CO₂ on marine benthic species. Species with calcified external structures are at risk of dissolution in response to seawater acidification. To further complicate the picture, the biological impacts of CO₂ leakage can also be modified by the prevailing conditions at the leak site. The buffering capacity of high carbonate sediments can be substantial, limiting the net change in ecosystem pH in response to limited CO₂ release (Lichtschlag et al., 2015). This may limit the magnitude of impacts to benthic infauna. However, in non-carbonate sediments, or for large CO₂ releases, the buffering capacity of the sediments might be exceeded. In these situations biogenic carbonate structures (bivalve shells, urchin tests and corals) will undergo dissolution to liberate aqueous carbonate ions. The dissolution of biogenic calcified structures has been widely reported (Gazeau et al., 2007; Gazeau, 2008; Byrne et al., 2014) with effects generally more pronounced in juvenile and larval stages (Talmage and Golber, 2009; Sheppard Brennand et al., 2010; Stumpf et al., 2012; Long et al., 2013; Hu et al., 2014). However, these impacts are not universal, and notable exceptions (normal calcification, hypercalcification) have been reported (e.g. Wood et al., 2008; Miller et al., 2009; Martin et al., 2011; Dorey et al., 2013), especially in situations in which the exposed shellfish are not resource limited (Thomsen et al., 2013).

Ultimately, the variability in response of closely related species and individuals precludes the formation of general predictions of likely in situ impact. As such, it is currently necessary to adopt a precautionary approach to predict the direction and magnitude of calcification responses to limited CO₂ release.

In extreme cases of CCS leakage, severe acidification could result in death of the organisms unable to escape the impacted area. However, this will not be the case for every leak scenario as many marine species, even some heavily calcified taxa, can tolerate shorter periods of more moderate acidification. This is because, unlike other potentially toxic substances, CO₂ is a naturally occurring and fluctuating compound in the marine environment. As a result of millions of years of exposure, marine creatures have incorporated this CO₂, along with other elements of carbonate chemistry, into many of their routine physiological processes. So whilst this means that large changes in seawater carbonate chemistry can potentially affect many aspects of an organism’s physiology, there is also the potential for organisms to temporarily alter or adjust their physiology to cope with these chemical changes. So in addition to the process of extracellular buffering described previously, many species have been seen to change their respiration rates, their activity levels and their reproductive outputs when exposed to high CO₂ (Kroeker et al., 2013). This response, known as physiological plasticity, affords some protection to organisms from rapid changes in their environment and can provide temporary protection against moderate acidification.

Plasticity, however, does not offer permanent protection for any organism against the impact of CCS leakage on water chemistry. This is because an organism’s ability to express plastic responses is to a large extent governed by the energy it has available (Thomsen et al., 2013). To maintain calcification rates under low pH, low carbonate saturation state conditions, some organisms can temporally reallocate more energy to this process and use less energy on other processes such as growth, locomotion or development of reproductive tissues. In the short term this can be an effective strategy to deal with an acidification shock. However, if leakage were to persist the increased energetic demand associated with living in a high CO₂ environment would inevitably lead to reduced growth and lower reproductive output or even death, with potential consequences for population development. The environmental consequences of CO₂ leakage therefore depend on both the severity and longevity of the leak. This means that even if a leak is fairly small, if it were to continue for many years it could ultimately cause some species to die and change the structure and the function of the community living around the leak.

In addition to physiological plasticity described above, there are 3 other important biological factors that impact upon the CO₂ sensitivity of any individual, population or species:

1. Life history stage. Most marine organisms go through a number of different developmental stages throughout their lives. Often, each of these stages displays very different physiological, ecological and behavioural traits. In addition, early life stages in general are less able to compensate for impacts than adult specimens (Hutchinson et al., 1998). It is unsurprising therefore that numerous studies have shown large differences in CO₂ sensitivity between these stages often with larval or juvenile stages showing greater sensitivity than adults. So if CCS leakage were to occur the major impact may not be on adult populations but on juveniles and could have longer term effects on recruitment and future population success. In addition, even as adults the impacts of CO₂ exposure could be greater if it were to occur during periods of high energy demand (e.g. the reproductive season or periods of intense growth such as moulting).

2. Local adaptation in populations. Recent studies have indicated that there is the potential for different populations of the same species to become more resilient to elevated CO₂ levels through adaptation (or acclimatisation) to local conditions. For example, Parker et al. (2011) showed that cultivated populations of oyster that had been selectively bred to increase energy efficiency and reduce food demands were better able to cope with high CO₂ conditions that the wild population. Such adaptation has also been observed in phytoplankton (Schlüter et al., 2014).

3. Variability between individuals within a population. It has also been shown that even individuals from the same population can exhibit a high degree of variability in responses to CO₂ (Pistevos et al., 2011). This is often reflected in experiments by an increase in variance observed in data from high CO₂ treatments when compared to the controls. Chronic exposure of a population to sub-lethal CO₂ concentrations could thus in theory result in selection of the less sensitive individuals. In addition to the adaptation mentioned above this would then increase the population’s resilience but potentially adversely reduce genetic diversity of the species in future generations.

The presence of heavy metals can also affect the impacts of CCS leakage. Acidification of the pore water will increase the mobility of metals bound within the sediment as has been demonstrated in laboratory investigations (Romanao de Orte et al., 2014). Mobile metals including Al, Fe, Zn, Co, Pb and Cu have all been found to increase in porewaters with acidification, compounded by increased time of exposure to a CO₂ leak. Furthermore, acidification also influences the speciation of metals, transforming metals and metalloids, like As, into species much more available and toxic to biota (e.g. Di Toro et al., 2009). Although some mobilisation of heavy metals was seen in the short-term QICS release experiment, these did not exceed environmental impact thresholds (Lichtschlag et al., 2015).

Finally, organisms interact and are commonly confronted simultaneously with multiple environmental stresses (e.g. temperature, hypoxia, food shortage, suspended matter) compounding the impact of any single environmental stress (such as CO₂ exposure from CCS leakage) and severely inhibiting the scope for physiological and behavioural adjustments to overcome that stress (Riedel et al., 2014; Queirós et al., 2015). It is unlikely therefore that hypercapnia will act as a single stressor, and caution is needed when assessing tolerances of species to environmental perturbations from studies that have only factored in one variable. Moreover, marine ecosystems are under increasing pressure from a range of anthropogenic activities like fishing, pollution and habitat destruc-
tion. Benthic trawling has a widespread and significant impact on benthic communities, causing changes in structure (species density and diversity) and functioning (production, bioturbation and biogeochemical processes) (Kaiser et al., 2002; Thrush and Dayton, 2002). It is likely that such modified communities will react differently to increased CO$_2$ concentrations than undisturbed communities. They might be more sensitive to this additional stressor, but could also turn out to be less sensitive as the most sensitive species have already disappeared.

4.4. Monitoring and baseline acquisition for impact assessment

4.4.1. Biochemical baselines

Understanding the natural variability of carbonate chemistry in marine systems is an important aspect of predicting what level of perturbation could lead to impacts and ensuring the accurate attribution of observed changes. Since 2005 a far better evidence base describing the mechanisms of variability of coastal systems has emerged, especially for systems such as the North West European Shelf (e.g. Thomas et al., 2005; Artioli, 2012; Rippeth et al., 2014). However there is a significant bias in observations towards surface waters. It is clear from these studies that significant heterogeneity exists and consequently specific baseline monitoring should be undertaken for each individual storage site. Primarily baselines will be of interest for monitoring and for that purpose will need to resolve changes over relatively short time scales. For impact assessment the key criteria will be to determine the seasonal range of carbonate chemistry (similar to Fig. 3). Measurements will be required at the sea floor throughout an annual cycle perhaps on a monthly basis, with higher frequency observations at selected times, for example corresponding with periods of high biological activity or changes in riverine forcing. Models, if used in combination with observations can provide a comprehensive synthesis of sparse data.

Other parameters such as nutrient levels and oxygen concentration have major implications for marine ecosystems and an understanding of how these relate to natural fluctuations can be important for attribution and monitoring. In particular a process-based approach (using ratios of CO$_2$:O$_2$:nutrients departing from relationships governed by respiration and photosynthesis and rapid changes in CO$_2$ not associated with physical changes signified by changing temperature) similar to that proposed onshore by Romanak et al. (2012a, 2014) may have utility, although no proof of concept yet exists for marine systems. In summary near sea floor measurements of pH, pCO$_2$, temperature, O$_2$ and nutrients, at monthly-weekly timescales, are likely to be needed to define baselines.

4.4.2. Biological baselines

Given the huge amount of temporal and spatial biological variability that exists in any given local system, it can be difficult to define a specific, constant condition or value that would identify a pristine, undisturbed, or “baseline” condition, especially when using ecological parameters such as diversity, abundance or biomass. Rather, in undisturbed conditions these parameters will fluctuate within a dynamic range, with a mean value that varies temporally and spatially, within a measurable interval. This has significant implications for detecting and monitoring potential impacts of CCS on marine communities as well as for the definition of adequate baselines for areas of the seabed where CCS will take place. Ultimately any monitoring or baseline activities will require an ability to capture these local and regional natural biological dynamics and document any existing disturbance of the ecosystem.

While it is unreasonable to suggest that baseline surveys should aim to comprehensively quantify local long-term dynamics of the large areas of the seabed above CO$_2$ reservoirs, some understanding of seasonal dynamics, types of habitats covered and a reasonable mapping of ongoing parallel pressures (e.g. seasonal hypoxia, trawling grounds) is necessary. While some or all of these aspects are already regularly monitored by marine users and academia, it is important to highlight the need for an understanding of local ranges of stressors which may be exacerbated by potential CCS impact scenarios. In particular, given the evidence described in this paper and in the wider CCS and OA literature, determining vulnerability to risk factors associated with CCS will require some degree of understanding of what range for each of the parameters (pH, TA, DIC, O$_2$, salinity) natural communities have adapted to. This is because there is wide evidence to support the perspective that local adaptation determines the response of individual species and populations to environmental change (e.g. Eliason et al., 2011; Morley et al., 2009). As such, baseline characterisation should aim also to cover such parameters.

4.4.3. Monitoring for impact assessment

A fully quantified assessment of impact relating to a seepage of CO$_2$ will be immensely challenging, given both the limited accessibility of the sea floor and the complexity of the marine ecosystem. Two questions are really pertinent; firstly is any impact significant and, secondly, over what time scale does recovery occur? The initial indicator of impact potential is an assessment of the degree and duration of chemical perturbation – how much has pH decreased, for how long and over what area? If pH changes are sufficiently large, long term or extensive then a biological characterisation of the impacted area is appropriate, referenced to similar, nearby unimpacted areas and baseline observations. This monitoring should involve determining the identity, abundance, biomass and distribution of benthic fauna. This can be satisfactorily achieved by standard sediment core sampling, deployed from larger research vessels. In post-injection surveys, observers should be looking, in particular, for changes in the community structure and diversity of infaunal organisms that have been shown to be sensitive to CO$_2$ impact (e.g. Widdicombe et al., 2015; Widdicombe et al., 2009). For CCS leakage, these changes could be a reduction in the presence of, juveniles of calcified taxa or a reduction in the occurrence of calcium carbonate structures (e.g. shells and exoskeletons).

Microbial gene analysis may also be a powerful indicator of CO$_2$ specific changes to microbial populations (Tait et al., 2015); microbially mediated impacts could be assessed by determining changes in chemical profiles within sediments or between the sediments and water column. As with all biological indicators, understanding their natural temporal variability from comprehensive baseline data will be needed to discriminate possible impacts from CO$_2$ leakage. Not only is this important for identifying impact but also in guarding against the false attribution of natural changes, or changes due to other stressors, to CCS activities. By adopting effective biological monitoring operators and regulators could address both potential impacts and boost public confidence. The QICS experiment (Blackford et al., 2014) indicated that impacts are likely to appear within a few weeks of seepage initiation, and recovery may occur relatively soon after seepage cessation. Any temporal sampling strategy would need to be informed by the specific characteristics of any given leakage event, but surveys targeted a month after initial detection and a month after cessation would be an initial recommendation. Whilst biological impact understanding remains less than complete, progress since 2005 is considerable and we can now identify monitoring strategies that are appropriate for CCS activities.

Certain ecosystem changes could be indicative of leakage and its impact. For example the appearance of infauna or microbial mats on the sediment surface or the development of phytoplankton blooms (e.g. Murray et al., 2013; RScS, 2014). However, these may not be
caused by CO2 leakage and are likely to be transient, so may not be a reliable monitoring tool.

4.5. Significant gaps requiring further research

Whilst the characterisation of ecological sensitivities to high CO2 has advanced significantly over the last decade there are some knowledge gaps that have emerged and are candidates for further research. Arguably the most pressing stems from the identification of saline aquifers as significant storage options, and the possibility for impacts from displaced hyper saline anoxic formation water. Other questions that have arisen from recent research are the lack of knowledge of the effects of long term exposure to high CO2, especially relating to the buffering capacity of sediments and the chronic effects of exposure. In terms of approaches, models that can coherently address a larger range of leakage scenarios and experiments that address whole ecosystem effects - such as injections and analogue work and more realistic exposure scenarios, are most likely to provide new insights.

The injection of gas into sub-seabed aquifers or depleted hydrocarbon fields could lead to the displacement of fluids low in oxygen and highly enriched in ions, which, if they reached the seabed, could bring about a strong change in environmental conditions. The majority of deep aquifers in the North Sea are filled with high salinity formation (Warren and Smalley, 1994; Evans et al., 2003), in some cases in excess of 300 psu, a value similar to that of the Dead Sea (≈340 psu). If these reached the seabed, they could cause an up to ten-fold increase in local salinity, thus representing a potentially severe osmotic shock to organisms. Their ability to cope with such disturbances, as with pH, will depend upon their tolerance to the changing conditions (low oxygen and/or high salinity), and also on the magnitude and duration of the event. The width of that tolerance window will depend on the comparative range of each of the parameters typically experienced by the community in each area of the seabed. Environments of high salinity and low oxygen do exist naturally (Helly and Levin, 2004, e.g. the Red Sea, some areas of the Arabian Sea). However, this combination of stressors as a transient disturbance, is not often observed in nature other than in estuarine environments and coastal lagoons, where strong variation in the flow of rivers is observed seasonally (e.g. Newton and Mudge, 2003). In all cases, marine life inhabiting such extreme environments will have undergone selection and adaptation, and may be well equipped to survive such harsh conditions, at least seasonally. However, benthic organisms inhabiting the comparatively more stable and hospitable seabed areas, where CCS injection is likely to occur, will not have experienced such processes, and would likely exhibit low tolerance to very marked environmental gradients. There is currently limited data availability (although see Hannis et al., 2013) about the plausibility and scale at which formation water release might occur in a commercial CO2 storage operation, specifically its duration, volumes released, extent of impact areas, or dissolution rates. This gap needs to be addressed. However, these parameters are likely to depend on the particular characteristics of the aquifer explored, the nature of the overburden, the rate of injection and local hydrodynamics.

There are currently very few studies published which have exposed natural marine systems to elevated levels of CO2, but when managed (Blackford et al., 2014) have provided significant new insights into geophysical, chemical, biological and social impacts. In order to appreciate the full range of responses observed across different benthic habitats and communities, more of these large exposure experiments need to be conducted on a greater variety of benthic systems and including longer term relatively low exposure levels. Experiments need to incorporate more realistic exposure scenarios. The vast majority of CO2 exposure studies conducted to date have used continuous levels of CO2 exposure. However, results from both models and observations have shown that, due to local hydrodynamics, such as tides and currents, benthic systems will most likely be exposed to CO2 levels which can fluctuate and pulse with time. This periodic exposure to high CO2 could have very different effects on marine organisms and communities and needs to be more fully tested. A first mesocosm experiment indicated that some fluctuating in CO2 levels reduced the impact on adult bivalves compared to continuous exposure. In this pilot study, however, the fluctuation of the pH was relatively small (Klok et al., 2014). Further mesocosm testing of the impact of the fluctuations as predicted by model calculations are required to better assess the potential impact of CO2 released in the marine environment.

The majority of data on high CO2 impacts currently available have come from experiments that increase the CO2 content of the overlying water. Whilst this approach is appropriate for studying the effects of a dense, CO2 - enriched plume, it does not adequately replicate the effects of CO2 migrating up through the sediment and interacting with the pore waters. This is much more difficult to simulate in laboratory conditions but nevertheless would be the primary method of impact at the centre of any leak. The QICs experiment (Blackford et al., 2014) that achieved an injection from below demonstrated that this is an important consideration for understanding impacts and more such experiments are urgently needed.

Finally, whilst great progress has been made with models, very fine scale plume models do not yet include very realistic hydrodynamics, and realistic hydrodynamic models do not adequately resolve plumes. Whilst both have adequate carbonate chemistry, neither sufficiently include ecosystem processes at the same time as simulating leakage scenarios. The individual model tools are all available, but their synthesis is required to optimise the utility of model-driven research.

5. Conclusions

Since the publication of the IPCC Special Report in 2005 there has been a significant focus worldwide on the potential impacts of leakage from geological storage of CO2. Projects have examined possible effects on near surface ecosystems, both onshore and offshore, and on underground drinking water supplies. Much of this work has arisen from developing regulations for CCS. However, relevant information has also come from other avenues of research, most notably from investigation of other impacts on groundwater and from ocean acidification. In 2005, there was in particular a dearth of studies on possible risks to potable groundwater and the marine environment. Major strides have been taken in both research fields in the last ten years, although some gaps remain in all areas of impacts research.

Research has included continued observations at natural CO2 seepage sites, although these may or may not be good analogues of storage sites. It has also involved laboratory experiments and, in particular, the last decade has seen the development of field injection experiments. Whilst only some of these experiments have been primarily designed for impacts studies the majority have provided at least some useful results in that respect. Most field experiments to date have been onshore, but there has been at least one significant offshore injection and release test and a few smaller in situ experiments involving benthic chamber landers. Experimental work has been backed up, in most cases, by modelling, both to extend the scope of the investigations and to help understand processes.

The environmental impacts of leakage appear, in many cases, to be restricted both in spatial and temporal extent. Emissions at the land surface or the sea bed commonly occur at a number of small seeps typically metres to tens of metres across. These cover only a small percentage of the total surface area over which seepage may occur, thus the majority of the ground or seabed above the
CO₂ plume at depth will probably not be affected. This contrasts with other ecosystem impacts such as severe weather, pests and diseases, or anthropogenic effects such as trawling, which may be much more widespread. Modelling suggests that flow follows preferential higher permeability pathways (e.g. fractures) giving rise to the restricted surface expressions of leakage.

CO₂ is dispersed on entering an aquifer, seawater or the atmosphere. It may therefore affect a much larger area/volume than that below ground or the seabed. However, this is offset by the efficiency of dispersion, buffering, sorption/desorption and other processes and their ability to reduce CO₂ concentrations, or associated parameters such as pH, and therefore ameliorate their impact.

The size of the impact will depend on leakage flux rates and areas and, more importantly, their effect on CO₂ concentrations. Flux rates and areas have been examined through consideration of possible leakage scenarios with flow associated with boreholes or faults remaining the likeliest possibilities. They have been assessed from natural CO₂ releases, actual borehole emissions, and through modelling of both styles of leakage. A wide range of values seem plausible, with high rates possible from certain well leakage scenarios. It can be assumed that such cases would be rare and likely to be detected and remediated quickly. Similar assumptions can be made for large scale pipeline releases. The size of impact might therefore be mitigated by the short duration of the event. Rates of fault leakage are likely to be lower but harder to locate and remediate.

The conditions arising from CO₂ leakage need to be viewed in the context of the pre-existing range of values both to identify possible leakage and evaluate its impact. There are far more data on baseline values for near surface gases onshore than offshore. There is a general lack of near seabed measurements of pCO₂ or associated parameters such as pH. It is, however, known that marginal marine ecosystems have very variable carbonate chemistry and pH can vary by more than 1 pH unit. In oceanic settings the range may be 0.3 pH units or less. Observations in aquifers have shown that there can also be quite considerable variability of pH in time and space with ranges of 1-2 pH units being not uncommon. Stratification can occur in both aquifers and the deeper parts of shelf seas.

Visible stress has been observed in a range of land plants when shallow (30 cm depth) CO₂ concentrations exceed 10%. However, this is not always apparent if the plants were well established before exposure to CO₂; for example, autumn sown oilseed rape showed no obvious visible effects but the yield was impacted. Other plants, such as beetroot, had reduced leaf weight but this had no effect on the yield. In pasture vegetation, grasses were more tolerant of CO₂ than most broad-leaved plants with the exception of species of certain very tolerant genera such as Polygonum. There appear to be effects arising from both higher CO₂ concentrations and depletion of oxygen.

The relative anoxia that accompanies higher CO₂ levels is reflected in microbial responses by a switch in the community towards functional groups using anaerobic respiration. However, when CO₂ levels get closer to the upper limits of the natural background range the impact of CO₂ appears to more difficult to separate from effects caused by other sources of stress, such as drought or waterlogging.

Shallow groundwater and marine impacts could be from CO₂ leakage and/or brine displacement from deeper more saline aquifers. Lab experiments suggest overall that mobilisation of potentially hazardous elements can be self-regulating as buffering of pH leads to adsorption of mobilised elements. However reaction and element release rates were far greater in the lab than in field experiments under more realistic conditions. Field experiments highlighted the natural variability of aquifer properties, and, with a few exceptions, most elements remained below drinking water limits.

The impact of both CO₂ and brine leakage on aquifers has been studied over larger areas at a number of natural sites where some redox mobilisation of Fe and Mn has been observed but potentially more toxic elements, such as As, were sorbed onto oxides. Thus the impact of potentially hazardous elements appears limited, as in the smaller scale experiments.

There have been relatively fewer studies of impacts through offshore observations and situ experiments because of the greater logistical complexity and cost. There has therefore been more reliance on lab experiments with modelling to further understanding and extend to a larger scale. This has shown the importance of considering local (site-specific) conditions of topography and water movement as well as temporal variability, including both seasonal and weather-related effects, when considering possible impacts.

Natural CO₂ seeps offshore provide useful information on impacts but may not be true analogues of CO₂ leakage from geological storage. They are often in different geological settings (e.g. volcanic) so styles and rates of CO₂ emission may be different. They may contain significant amounts of other contaminants (e.g. gases such as H₂S, trace metals) that might not be released from CO₂ storage. Although this may allow the effect of impurities to be studied, these other components can complicate assessment of impacts from CO₂ alone. Also, as for onshore examples, the benthic ecosystem has adapted to the presence of CO₂. These sites are useful, however, to understand particular leakage styles and rates, and how the CO₂ is dissolved, transported, and mixed/diluted in the water column (e.g. pulsed exposure), all issues that will influence and control its eventual impact. The recent experimental release of CO₂ through the seabed for the QICS experiment is an important addition to impacts research. Although there were impacts near the bubble plumes at QICS (reduced numbers of individuals and species diversity) recovery was rapid over a matter of a few weeks.

Offshore organisms can, to varying degrees, cope with the acidification of seawater caused by CO₂ escape. This depends on the energy available to them and may come at the expense of growth, reproduction or other processes. The ability to cope with the change also depends on its severity and longevity and the life history stage during exposure. It also depends on whether the ecosystem is under additional stress, for example due to it being a marginal environment for certain species or through external factors such as pollution or trawling.

The restricted spatial extent of surface leakage has implications for monitoring. This needs to be able to cover the large areas of storage sites but be capable of detecting small leakage features. Remote sensing provides the large areal coverage onshore, although only allowing responses to be detected where there is surface vegetation, but so far is unable to distinguish CO₂-specific responses from other causal mechanisms. Wide area coverage offshore is being developed through the use of chemical and acoustic sensors mounted on AUVs or a current UK project (ETI MMV study). The responses of marine organisms to CO₂ leakage, such as surfacing of infauna, development of microbial mats or phytoplankton blooms, might be used as leakage or impact indicators, but are unlikely to be a primary monitoring tool.

Single plumes of impacted groundwater will also be limited in extent in unconfined aquifers and difficult to detect. There will be more of a tendency towards a larger plume of CO₂ at the top of a confined aquifer. There is thus a similar difficulty in detecting small leakage anomalies over wide areas. This is exacerbated by the use of many downhole techniques, which provide very restricted horizontal coverage. It has been shown, however, that the presence of a confined aquifer, or the reservoir for an open aquifer, may be detectable over a wider area than more traditional geochanical or related parameters. The probability of detection may be enhanced by careful siting of new monitoring wells but this cannot be done if only existing wells are available. Indications
from deeper monitoring or near surface techniques could help to focus aquifer investigations. Some geophysical techniques, such as electrical resistivity, show promise for wider area monitoring. Continuous monitoring and modelling suggests that leakage (even at a constant rate) does not give rise to exposure to constant levels of CO₂ (or pH etc.). Rather exposure varies as a result of external influences. Thus a groundwater plume will respond, for example, to groundwater flow and interactions which depend on aquifer mineralogy. Onshore surface releases are affected by soil moisture, atmospheric pressure and wind speed. Offshore leakage is affected by the tide, seasonal and weather related factors as well as topography. There can therefore be quite large swings in CO₂ concentration, and other parameters, that need to be factored in when considering impacts. This has yet to be done in most lab experiments, but of course happens for longer term in situ experiments or at natural sites, although the full extent of the variation and its effect on impacts is not always considered fully.

Observations and modelling of leakage into the marine environment suggest that, although free gas may reach the sea surface in shallow water, rapid dissolution occurs, producing dissolved CO₂ plumes near the seabed.

5.1 Gaps

There remains a need to constrain more tightly the most likely leakage scenarios and define the range of fluxes and areas associated with them. This may be possible through a careful study of the most analogous natural occurrences of CO₂ release through faults/fractures e.g. focussing on sedimentary basins rather than volcanic terrains. It may also be possible to use more sophisticated information on fault properties (e.g. permeability, capillary entry pressure, thickness) from both field observations, and parameters used in petroleum reservoir flow simulators, to provide more realistic models of CO₂ or brine flow rates and areas of release. The outputs from such studies could then be fed into more rigorous experiments, detailed study of truer analogues and further modelling to assess the likely impacts.

Baseline measurements are more numerous onshore than near the seabed. However, there is a need to expand the range of onshore environments covered and for far more offshore observations of the variability of pCO₂ and pH, both spatially and temporally. Baseline variability, and the covariance of other parameters in natural processes, can be used to help recognise leakage.

There has been relatively little consideration of brine leakage and the possibilities of this at different types of storage site and the consequences of any leakage merit further consideration. There has also been little work on the possible impacts of impurities in the injected CO₂ and whether these could contribute to additional impacts.

Further injection experiments in natural onshore and marine systems are required to assess responses and recovery under the most realistic conditions in a wider range of species/ecosystems/aquifer types. Onshore studies have not explored the influence of soil type, mobilisation of potentially harmful elements, effect on soil fertility or need for remediation. Offshore tests and mesocosm experiments need to explore the effect of the pulsing of pH values shown by observations and models and assess possible chronic effects from longer term lower level exposure.

Further synthesis is needed of fine scale plume, hydrodynamic and biogeochemical models offshore. Onshore modelling has been fairly limited for the near surface and further development of the soil-plant system and consideration of whether leakage scenarios could directly impact man or other animals would be desirable. A more integrated approach combining modelling, laboratory and field studies of a wider range of aquifer and leakage types is needed.

Acknowledgements

This paper has drawn on the results of many studies. We would like to acknowledge in particular our involvement in a number of these and the many colleagues who contributed. These include the FP6 European Network of Excellence CO₂GeoNet; Research into Impacts and Safety in CO₂ Storage (RISCS), funded by the EC 7th Framework Programme (Project No. 240837) and by Industry Partners ENEL, IBL, Statoil, Vattenfall AB, EON and RWE; EC2O – Sub-seabed CO₂ Storage: Impact on Marine Ecosystems, also funded by the EC 7th Framework Programme (Project No. 265847) and QICS – Quantifying and monitoring potential ecosystem Impacts of geological Carbon Storage, funded by NERC (NE/H013962/1), the Scottish Government and METI/MEXT of Japan. The comments and suggestions of three anonymous reviewers are gratefully acknowledged.

References


De Waele, M., Wei, W., McNeil, D., Chen, B., 2013. Small scale modelling of the physical-chemical interactions of a CO\textsubscript{2} plume from sub-seabed reservoirs or pipelines within the North Sea and surrounding waters. Mar. Pollut. Bull. 73, 504–515.


Ellis, J., Nisi, R., Ortega, M.F., Mazadiego, L.F., Vallesi, O., Grandia, F., 2013. CO\textsubscript{2} soil flux baseline at the technological development plant for CO\textsubscript{2} injection at Hontomin (Burgos, Spain). Int. J. Greenh. Gas Control 18, 224–236.


Ferries, J., Oppermann, B.L., Gwosdow, S., Möller, I., Herrmann, M., Krüger, M., 2013. Microbial community changes at a terrestrial volcanic CO\textsubscript{2} vent induced by soil acidification and anaerobic microhabitats within the soil column. FEMS Microbiol. Ecol. 84, 60–68.

detecting a footprint of CO2 leak in a varying marine environment. Int. J. CO2 4, 9, 245–278.


Primate Marine Laboratory, 86 pp.


Please cite this article in press as: Jones, D.G., et al., Developments since 2005 in understanding potential environmental impacts of CO2 leakage from geological storage. Int. J. Greenhouse Gas Control (2015), http://dx.doi.org/10.1016/j.ijggc.2015.05.032