Safe storage and effective monitoring of CO₂ in depleted gas fields

Charles R. Jenkins,1,2,a Peter J. Cook,1 Jonathan Ennis-King,1,3 James Undershultz,2,a Chris Boreham,4,a Tess Dance,4,a Patrice de Caritata,e, David M. Etheridge,4,a Barry M. Freifeld,4,a Allison Hortle,6,a Dirk Kirst,3,a Lincoln Paterson4,a, Roman Pevzner1,e, Ulrike Schacht1,e, Sandeep Sharma,5,a Linda Stalker,4,a and Milovan Urosevic1,a

1Cooperative Research Center for Greenhouse Gas Technologies (CO2CRC), National Farmers’ Federation House, 14-16 Brisbane Avenue, Canberra 2600, Australia; 2Earth Science and Resource Engineering, Commonwealth Scientific and Industrial Research Organization, Black Mountain, Canberra 2601, Australia; 3Earth Science and Resource Engineering, Commonwealth Scientific and Industrial Research Organization, Ian Wark Laboratory, Bayview Avenue, Clayton, Victoria 3168, Australia; 4Earth Science and Resource Engineering, Commonwealth Scientific and Industrial Research Organization, Ian Wark Laboratory, Bayview Avenue, Clayton, Victoria 3168, Australia; 5Department of Exploration Geophysics, Curtin University, 26 Dick Perry Avenue, Technology Park, Kensington, Perth 6151, Australia; 6Geoscience Australia, GPO Box 378, Canberra 2601, Australia; 7Marine and Atmospheric Research, Commonwealth Scientific and Industrial Research Organization, 26 Dick Perry Avenue, Technology Park, Kensington, Perth 6151, Australia; 8Lawrence Berkeley National Laboratory, MS 90-1116, One Cyclotron Road, Berkeley, CA 94720; 9Earth Sciences, Simon Fraser University, Burnaby, BC, Canada V5A 1S6; 10Department of Exploration Geophysics, Curtin University, 26 Dick Perry Avenue, Technology Park, Kensington, Perth 6151, Australia; 11Australian School of Petroleum, University of Adelaide, Adelaide 5005, Australia; and 12Schlumberger Carbon Services, 256 St. Georges Terrace, Perth 6000, Australia

Edited by E. Ronald Oxburgh, University of Cambridge, United Kingdom, and approved October 25, 2011 (received for review May 19, 2011)

Carbon capture and storage (CCS) is vital to reduce CO₂ emissions to the atmosphere, potentially providing 20% of the needed reductions in global emissions. Research and demonstration projects are important to increase scientific understanding of CCS, and making processes and results widely available helps to reduce public concerns, which may otherwise block this technology. The Otway Project has provided verification of the underlying science of CO₂ storage in a depleted gas field, and shows that the support of all stakeholders can be earned and retained. Quantitative verification of long-term storage has been demonstrated. A direct measurement of storage efficiency has been made, confirming that CO₂ storage in depleted gas fields can be safe and effective, and that these structures could store globally significant amounts of CO₂.

Increasing atmospheric CO₂, and the resulting climate risk, is a critical issue. Fossil fuels will continue to be burned for decades (1), thus capture and geological storage are vital to reduce the current approximately 30 Gty⁻¹ of CO₂ emitted to atmosphere (2, 3). Many aspects of carbon capture and storage (CCS) are well-understood in chemical engineering and the oil and gas industries. Globally, there appears to be sufficient storage volume for decades to come (2) with depleted oil and gas reservoirs being obvious early targets for CCS projects.

CO₂ has been injected into oil reservoirs for decades to enhance recovery (4). Large United States operations of this type at Weyburn (5), Cranfield (6), and Rangely (7) are monitored as CCS case studies. The Sleipner (8), Snøvit (9), and In Salah (10) projects store 1–3 Mt CO₂ each year from gas processing, and the similar Gorgon project in northwest Australia is under construction (11). Smaller research and development (RD) storage projects have been completed or are in progress (12–15) (www.netl.doe.gov/technologies/carbon_seq/partnerships/validation.html). Subsurface storage of natural gas has a long and successful history (16). Hazardous waste, in large volumes (currently 30 Mt y⁻¹ in the United States) is injected into deep saline aquifers (17), and in Canada approximately 5 Mt of acid gas (CO₂ and H₂S) has been safely stored, in several cases into depleted gas reservoirs (18).

Despite a successful record, CCS remains controversial. Technical concerns are long-term leakage, global capacity, engineering feasibility, and the scale of deployment. Public opposition focuses on perceived risks from leakage. Development of some onshore sites for commercial CCS has been blocked. Pathfinding projects in the Netherlands (Shell, Barendrecht) and Germany (Vattenfell, Almork), which aimed to use onshore depleted gas fields for storage, have foundered on political opposition at many levels. The higher costs of offshore storage, or transport to remote areas, may mean that an issue with the viability of CCS is emerging. However, smaller demonstration projects in Germany (European Union-funded, Ketzin, 15) and France [Total, Lacq, (19)], the latter in a depleted gas field, are progressing.

Deposited gas fields are an important target for RD because they could store many years of emissions from the some of world’s largest point sources (2, 20–22). Deposited gas fields thus may represent a globally significant storage resource, but there have been few direct measurements to date to support this conclusion.

In this environment, noncommercial storage demonstration projects are important in building public confidence, confirming and extending our scientific understanding, and building technical capacity. The CO2CRC Otway Project, located in southeast Australia, is a midscale demonstration utilizing a depleted gas field for storage. Being in a populated area, a key objective was to set an example of working successfully with all stakeholders, with an open approach to communication at all levels and extensive monitoring to provide the fullest assurance that risks are low, well-understood and manageable. The monitoring is also used to test models and develop robust measurement procedures.

There have been 65,445 t stored at Otway, with no safety issues arising and monitoring results showing that there is sound understanding of the storage process. Good relationships have been built and maintained with all stakeholders. Direct measurements of storage capacity have provided an early confirmation of important estimates of large global capacity in depleted gas fields.

The Project Site
The CO2CRC (www.co2crc.com.au) began searching in 2004 for a CO₂ source close to geological storage. A site was located in a mature hydrocarbon province, the Otway Basin of Victoria (23, Fig. 1). Two adjacent petroleum tenements were purchased and extensively explored.

Carbon dioxide | geosequestration | carbon dioxide | climate change | energy policy
and multi-Darcy permeability. Estimated capacity, 150 kt of CO
neous, much has excellent reservoir quality, with 20% porosity
term risk of wellbore leakage. Although the Waarre is heteroge-
gration into overlying aquifers. Naylor-1 presents the only long-
Bounding faults terminate within the Belfast, preventing fluid mi-
the structure. Overlying the reservoir is the low-permeability
contains postproduction immobile gas with free gas at the crest of
25
Cretaceous Waarre-C formation) is fault-bounded on three sides
the site thus had excellent capacity, injection against the 300-m thick Belfast Mudstone, forming a structural
structure. Overlying the reservoir is the low-permeability
Belfast Mudstone (SI Appendix). Bounding faults terminate within the Belfast, preventing fluid migration into overlying aquifers. Naylor-1 presents the only long-
wellbore leakage. Although the Waarre is heterogene-
ous, much has excellent reservoir quality, with 20% porosity and multi-Darcy permeability. Estimated capacity, 150 kt of CO2
comfortably exceeds project targets. The structural trapping
would ensure that the stored CO2 plume would be confined within
in a 0.5 km² footprint. The site thus had excellent capacity, inject-
tivity, and containment.
The dip of the reservoir structure suggested a design with gas from Buttress (referred to as injected CO2) (SI Appendix) being injected at CRC-1 and migrating up-dip by buoyancy to the Naylor-1 monitoring well (Fig. 3). It was vital to site CRC-1 so
migration to Naylor-1 would happen within the project lifetime. Initial geological models were created from available well and seismic data (SI Appendix) (24). These incorporated estimates of porosity, permeability, pressure, and geometry including faults, sedimentary layers, and facies distribution. Preliminary dynamic models were then created and calibrated against the Naylor-1 production history (25). Migration times were estimated to be 4–8 mo. No problems were expected from pressure build-up in the reservoir, and the planned injection volumes were too small to exceed the reservoir’s spill point.

Based on initial modeling, CRC-1 was drilled 308 m down-dip from Naylor-1, with injection at a depth of 2,003–2,014 m (sub-
level) into the Waarre-C. At this depth the density of the in-
jected CO2 is about 500 kg m⁻³. Geological uncertainties were later reduced significantly with data obtained during the drilling of CRC-1 (26), and the dynamic model was refined. This initial phase of the project followed well-established methods in the oil and gas industry.
Injection of the Buttress-1 gas into the Waarre-C commenced on March 18, 2008, and 65,445 t were stored between March 2008 and August 2009, when injection ceased.

Risk Assessment
Leakage rates of less than 0.1% y⁻¹ to the atmosphere are needed to ensure effective climate abatement (27, 28) by CCS. The Intergovernmental Panel on Climate Change (IPCC) assessed 0.001% y⁻¹ as likely (2) for well-designed storage; this leakage rate was the benchmark for the risk assessment. Commercial-scale storage will also have to manage the risk of financial penalties for leakage, such as those foreshadowed in regulations by the European Union. An expert panel considered both the engi-
neered and natural systems (29). Probabilities were assigned to a range of scenarios and the associated leakage rates or volumes were estimated. Scenarios were combined by Monte Carlo simu-
lation. The combined leakage rate from all risks was estimated to be below 0.001% y⁻¹. The risk assessment was repeated using data from the drilling of CRC-1, with little change (SI Appendix).

Regulation and Community
At the start of the project no regulations existed in Australia for storage of CO2 and it was necessary to work with regulators using existing petroleum, water, environmental, and planning legis-
lation to assemble an appropriate regulatory regime. A crucial
approval was obtained under the research provisions of Environ-
mental Protection Agency (EPA) Victoria. Key performance
indicators were agreed with EPA, based on comprehensive mon-
itoring. The process of permitting the Otway Project was an example that had an important influence on the development of CCS legislation in Australia (30).
Securing and maintaining the consent of the community was vital. The CO2CRC developed a communications strategy, based on market research of the area, and was proactive in engaging with communities and decision makers, both face to face and also through leafleting, media releases, and a comprehensive web site (31). Key principles were a willingness to listen to the public, to be open about all aspects of the project, and to ensure no surprises—any news about the project was communicated first by the CO2CRC directly to those affected. This outreach effort was aided by familiarity in the area with oil and gas operations. Evaluation shows a generally positive attitude has persisted in the local community, and wider media coverage has been generally balanced and positive over the period where the CO2CRC has been active (31).

Monitoring Design

A wide range of monitoring was carried out in the Otway Project as part of its research objectives: A commercial project might rely on a smaller number of well-developed techniques. Measurements were necessary to confirm containment of injected CO2 in the reservoir and provide assurance that groundwater, soil, and air are unaffected (32). These measurements were often technically challenging, and complicated by access to farmers’ land requiring frequent negotiation and compromise.

The assurance measurements compare pre- and postinjection properties, and began well before CO2 injection commenced. Seismic measurements investigate an overlying aquifer, whereas groundwater, soil, gas, and atmospheric monitoring provide assurance at increasing distances from the reservoir. These measurements will only show changes if leakage out of the reservoir occurs. Leakage is very unlikely and it is impractical to characterize the approximately 2 km of overburden and model hypothetical leak pathways through it. Interpretation of results is therefore based on conditional sensitivities—if CO2 were to enter this zone, then the following effects are predicted. Simple leakages, usually point sources, are modeled to understand the sensitivity of the assurance measurements. Even if the sensitivity is ill-defined, an important public assurance objective may be achieved if no changes are detected in important assets such as ground water. Anomalies in an assurance measurement would not be decisive in isolation, but would initiate a cascade of investigations by successively more precise and expensive techniques (SI Appendix).

It was vital to the community to confirm that potable aquifers were unaffected by CO2. The composition and chemistry of water in deep and shallow water wells was measured twice a year (33). Samples were collected from 24 existing wells, mostly within a 5 km radius of CRC-1. Vadose zone soil-gas composition was also measured annually during summer (34). Typically 150 samples were collected on a 4 × 3 km grid over the expected location of the subsurface plume, the area in which major faults terminate close to surface, and natural CO2 accumulates (SI Appendix).

Atmospheric concentrations and fluxes of CO2, isotopic composition (δ13C CO2), and tracers of injected gas (SF6, CH4) or of combustion (CO) were monitored 700 m northeast of the injection well (35) and compared to the undisturbed background measured at the long-established Cape Grim site (Tasmania). CO2 atmospheric analyzers were also located near the injection site to monitor for larger anomalies and deep soil fluxes (36).

Direct measurements of reservoir fluids at Naylor-1 are the primary confirmation of containment in this project. Naylor-1 was instrumented with a complex bottom-hole assembly (Fig. 3, 32), including U-tube fluid sampling apparatus (37) that recovered pressurized reservoir fluids. Three sampling points in the wellbore straddle the preinjection gas-water contact, marking the boundary between residual and free CH4. Shortly after injection commenced at CRC-1, tracers (deuterated methane CD4, Kr, and SF6) were added to ensure unambiguous detection of injected CO2 at Naylor-1 (38). The small amounts of CO2 originally present in the Naylor reservoir have a distinct 13C signature, so this isotope is also a tracer. The reservoir models are tested directly through a comparison with measured molecular, isotopic, and tracer compositions (SI Appendix).

To detect injected CO2 within the reservoir, or leakage that might occur into overlying aquifers, conventional time-lapse seismic surveying methods were used. Detection of changes in the acoustic reflectivity caused by injected CO2 at around 2 km depth requires a significant difference in reflectivity, excellent survey repeatability, and a high signal-to-noise ratio (39, 40). Modeling predicted that injection into the depleted Waarre-C reservoir would produce small changes because of the significant amounts of residual CH4 (41). Seismic surveys on land generally have poor repeatability, and it was judged unlikely that any seismic effects would be detected in the Waarre-C. If injected CO2 entered overlying aquifers, reflectivity changes should however be pronounced (40). A baseline survey was shot in January 2008, and repeats were acquired in early 2009 and 2010 after 35,000 and 65,445 t of CO2 had been injected. These surveys produced data that could be combined to form time-lapse images of the Waarre-C and overlying aquifers (SI Appendix).

Fig. 3. Schematic of the injection and monitoring wells, indicating wellbore perforations and U-tube inlets. U-tube 1 accesses the free gas cap (red) through leaks in a casing patch, installed during production. Free natural gas is bounded below by the gas-water contact. In the light orange zone natural gas is immobile and the pore space contains mostly water. (Reprinted from Int J Greenhouse Gas Control, Vol 5, Underschultz et al., CO2 storage in a depleted gas field: An overview of the CO2CRC Otway Project and initial results, 922–932, copyright 2011, with permission from Elsevier.)
Reservoir Modeling
The core and wire-line logs obtained from CRC-1 (26) showed that the Waarre-C reservoir comprised stacked sandstone bodies of varying grain sizes, with thin shale baffles and streaks deposited in near-shore, tidally influenced channel settings. Four revised geological models were created stochastically, representing short- and long-range correlations in the shale baffles. Hysteric and nonhysteretic relative permeability curves were estimated from laboratory core tests (42) to give a total of eight models. These models are equiprobable, and capture the likely ranges in geological structure and rock properties. A set of dynamical fluid-flow models for the reservoir was then computed (SI Appendix). Comparison with these models is the metric for evaluating predicted storage performance. The relevant data are the seismic response and details of arrival of injected CO₂ and tracer at Naylor-1. The models predict that the injected CO₂ will remain above the pre-production gas-water contact, which is presumed to be the spill point for the reservoir. Thus assurance measurements are not expected to show injected CO₂. Fig. 4 shows vertical cross-sections of the reservoir model. The high saturation region is the remaining gas cap, with the injected CO₂ accumulating below. The injected gas is cooler than the reservoir as it leaves CRC-1, causing a region of localized cooling. Tracers are most concentrated beneath the gas cap.

Monitoring Results
Aggregated groundwater data from the widely-used shallow limestone aquifer show no statistically significant pre-to-post-injection changes in bicarbonate or electrical conductivity, with a small (but statistically significant) increase (0.05 units, p < 0.01) in median pH. These small changes demonstrate that overall water quality is unaffected by injection to within natural variability (SI Appendix). Aggregated soil-gas data show a consistent correlation between δ¹³C and CO₂ concentration. Data were obtained for three summers before injection and two after, and all follow this correlation, which results from decomposition of organic matter (SI Appendix) (34). Most δ¹³C values are also lower than those of the injected CO₂, so there is no indication of changes that could be attributed to injection.

The groundwater and soil-gas results provide assurance by showing that water and soils are practically unaffected; this assurance was the objective of this monitoring. Considerably more data-gathering and modeling would be needed to bound quantitatively the amounts of CO₂ that could have entered these zones of investigation and yet remain undetected. For example, modeling shows that leakage of CO₂ might only affect small volumes of the aquifer and so is unlikely to reach the monitoring wells (SI Appendix). More detailed interpretation of soil-gas data is difficult because of the very variable permeability of the vadose zone. Attempting to characterize the sensitivity of soil-gas and groundwater monitoring may not be cost effective.

The atmospheric concentrations and fluxes of CO₂ show large diurnal and seasonal variations, reflecting the effects of plant respiration, photosynthesis, and atmospheric dispersion. Much of this variability can be modeled (SI Appendix) (43). By day, in windy conditions, the CO₂ concentrations often settle at a steady baseline. In these favorable conditions small emissions of CO₂ (equivalent to about 2 kt y⁻¹) from the diesel engines used while drilling of CRC-1 were identified, confirmed by tracers and quantified with dispersion modeling (Fig. 5). Other nearby industrial sources have also been detected. These measurements confirm useful sensitivity to spatially small emitters.

Injected gas at reservoir level was detected in the 2008–2009 seismic survey (40) but not confirmed in 2010. Modeling predicts that changes would be below the noise level on the time-lapse seismic images (SI Appendix). From an assurance perspective, simulations for the immediately overlying Paaratte aquifer show that 5 kt of injected CO₂ in a small accumulation would have been detectable (40) but no changes were observed (Fig. 6).

None of the assurance-monitoring techniques have detected any anomalies that would indicate the presence of injected CO₂ outside the storage reservoir. The sensitivity of some assurance
techniques to leakage has been estimated, conditional upon there being a leakage route to aquifer or atmosphere, where effects of CO$_2$ become apparent. For the seismic and atmospheric monitoring, such conditional sensitivities to CO$_2$ of a few kilotonnes, or kilotonnes y$^{-1}$, have been demonstrated for point sources. Such sources might result from a small fault acting as a conduit into an aquifer, or an unmapped wellbore transferring CO$_2$ to the surface. Reservoir fluid samples were collected weekly to fortnightly at the Naylor-1 monitoring well. Pressurized fluids were brought to surface by U-tube by reservoir pressure within the gas cap or initially by using N$_2$ pressure for formation water (Fig. 3) from U-tubes 2 and 3. Pressure assistance was initially required for water samples because of the depressurization of the reservoir during earlier commercial natural gas production. The arrival of injected CO$_2$ caused a transition to self-lift, as the composition changed to lighter injected fluid, which flowed to surface under reservoir pressure during sampling. The upper sampling point (U-tube 1, in the remnant gas cap) always returned gas. The lower sampling points (U-tube 2 and U-tube 3) initially returned reservoir water later transitioning to a CO$_2$/CH$_4$ mixture with very little water. Chemical analysis was done after controlled depressurization. Fluid samples were analyzed for concentrations of CO$_2$, hydrocarbons, tracers, and carbon isotopic composition (SI Appendix) (37, 38, 44).

The arrival of injected CO$_2$ at the monitoring well is a key indicator of the progress of the storage. Fig. 7 shows the measured arrival of CO$_2$ and tracers at U-tube 2, where both show a rapid breakthrough and then a plateau in concentrations. Comparison with the range of predictions (SI Appendix), with uncertainty dictated by uncertain geology and rock physics, shows a broadly satisfactory agreement although breakthrough is slightly earlier than predicted. Data and predictions from U-tube 3 are similar, with breakthrough slightly early but not significantly. The timing discrepancies, multiplied by the injection rate, correspond to uncertainty of only 5–10 kt in the implied amount of injected CO$_2$ that is needed to match models to data. These small discrepancies in timing, and the steady CO$_2$ concentrations at later times are consistent with containment in the reservoir.

Water samples from U-tube 2 and U-tube 3 were collected from preinjection to self-lift. The pH decreased sharply with increasing CO$_2$ content and changes in the water composition are consistent with minor dissolution of carbonate and silicate minerals. However, the dominant process affecting the water composition was determined to be mixing within the reservoir and wellbore. There was no evidence for significant mobilization of trace metal species.

The geochemical results show that the behavior of CO$_2$ in the subsurface is understood so that the storage model is realistic. The range of uncertainty from reservoir heterogeneity was adequately characterized and the main additional uncertainties are from poorly known relative permeabilities. These could have been more constrained with core analysis techniques better adapted to the highly permeable samples from CRC-1. Containment is demonstrated by the consistency of the geochemical data with forward models at CO$_2$ mass levels close to those that were actually injected. The clear absence of CO$_2$ in the overlying aquifer (shown by repeat seismic) adds credibility to this conclusion. Because of the simple geometry of the reservoir, lateral and vertical movement of the CO$_2$ plume are predicted to be very limited.

Direct confirmation of containment is difficult because the amount of CO$_2$ stored is relatively small. A minimalist climate abatement target (24, 25) of 0.1% y$^{-1}$ leakage to atmosphere is 65 ty$^{-1}$ for the amount we have stored. This leakage rate is comparable to the nighttime respiration rate for 1 km$^2$ of nearby pasture. However, the sensitivity of our seismic or atmospheric techniques would remain the same for industrial scale storage and so correspond to much smaller fractional sensitivity. The sensitivity of our fluid sampling method (injection rate x timing uncertainty) would scale with storage size. An industrial site might store 100x more CO$_2$ than Otway. A point leakage rate of 6,500 ty$^{-1}$ (0.1% y$^{-1}$) from such a site would be detectable by the atmospheric and seismic methods we have tried. However the IPCC estimates (2) of likely leakage rates (1% in 1,000 y) would be very hard to confirm, amounting again to only 65 ty$^{-1}$.

**Reservoir Storage Efficiency**

There is uncertainty regarding the extent to which the pore space previously occupied by gas (primarily CH$_4$, accumulated over
geological time) can be reoccupied by CO₂ injected in only a few years. Global estimates of capacity depend on applying average assumptions to a wide range of different circumstances and need to be checked by experience (22). The International Energy Agency Greenhouse Gas RD Program estimated that 160 Gt of practical capacity in depleted gas fields, matched to point sources, will be cumulatively available by 2050 (21). Emissions from such sources are 11 Gt yr⁻¹, accumulating to 279 Gt by 2050 at present growth rates (2.5% yr⁻¹). This potential capacity is globally significant.

With our sampling system we can measure the dynamic storage capacity of the reservoir as it was refilled with injected CO₂ (25). Self-lift occurs when our U-tube samples become mostly CO₂ and CH₄, and hence the contents of the U-tube are of sufficiently low density to be forced to surface by reservoir pressure alone (SI Appendix). Detailed modeling of the sampling process verified that the transition to self-lift reliably indicates the passage of the gas-water contact (GWC) (Fig. 2) (SI Appendix). The amount of injected CO₂ that was stored between self-lift at U-Tube 2 and U-Tube 3 is then compared to the available pore space, which is estimated both from the geological model and from the production data of natural gas. Detailed calculations (SI Appendix) predict that 56–84% of the space originally occupied by recoverable CH₄ is reoccupied by CO₂, taking account of uncertainties in the pore volume, postproduction residual CH₄ saturation and the possible range of movement of the GWC. Generic estimates used to calculate global capacity in depleted gas fields (21, 22) have assumed 75%.

The determination of storage efficiency is a unique measurement of this type in CO₂ storage, and is consistent with extensive experience of natural gas storage (45). Although dependent on detailed circumstances (SI Appendix), for example rock type or timing of storage after production, our data adds weight to the conclusion that depleted gas fields have enough storage capacity to make a significant contribution to reducing global emissions (21).

Fig. 7. Measured arrival and accumulation of injected CO₂ at the Naylor monitoring well’s U-tube 2 (orange points) for SF₆ (parts per million) and CO₂ (mole fraction). The background color coding superimposes predictions of eight reservoir models, each blurred by the experimental error levels. Lighter shades indicate higher probability of a reservoir model occupying a point in the concentration, time plane. The scatter in SF₆ around day 400 appears to be real, but is unexplained.

Conclusions
The CO₂CRC Otway Project has demonstrated that the storage of CO₂ in a depleted gas field can be designed and safely achieved. Monitoring showed that there has been no measurable effect of stored CO₂ on soil, groundwater, or atmosphere. Good relations with the local community, and decision makers at various levels of government, were established and maintained. Seismic imagery and fluid sampling confirmed dynamic and geochemical models. Sensitivity of monitoring techniques to surface leakage rates at the few kilotons yr⁻¹ level, or subsurface accumulations at the kilotons level, was demonstrated. Achieving this sensitivity shows that commercial-scale storage programs could be effectively monitored to ensure climate abatement was being achieved.

Overall the project shows that our level of knowledge of subsurface processes involving CO₂, and our ability to forecast its behavior, is adequate to proceed with large-scale geological storage of CO₂ in depleted gas fields, and the capacity of these structures could be of global significance for carbon capture and storage.

Materials and Methods
The design, execution, and interpretation of this project followed proven methods developed for oil and gas extraction, and environmental monitoring. Key aspects of adapting these methods for CO₂ storage are discussed in the main text. Details of standard modeling of the subsurface, and the collection of data both from the surface and subsurface are given in SI Appendix.

ACKNOWLEDGMENTS. The Otway Project was supported by the Australian Government through the Cooperative Research Center program, by the Victorian Department of Primary Industry, the US Department of Energy and by the industrial and institutional partners in the CO₂CRC. The conclusions reported here depend on work, published and unpublished, by many participants in the CO₂CRC.


Supporting Information for

Safe storage and effective monitoring of CO2 in depleted gas fields


correspondence to: charles.jenkins@csiro.au
Gas compositions

Commercially-purchased CO\textsubscript{2} is extremely expensive for a large injection, so that finding a natural source close to an injection site was crucial. The gas from Buttress (the “injected CO\textsubscript{2}”) has a molar composition of 75±2\% CO\textsubscript{2} and 21±2\% CH\textsubscript{4}. The remaining 4\% is mostly heavier hydrocarbons. The CO\textsubscript{2} is of magmatic origin and has an isotopic value $\delta^{13}C = -6.7\pm0.1$‰ VPDB. It proved too expensive to remove methane from the gas stream, but at the pressures and temperatures involved the mixture closely models pure CO\textsubscript{2}. Our detailed modeling takes account of the composition, which affects physical properties such as the viscosity, compressibility, and the equation of state. At Naylor, the composition in the gas cap is 86±2\% CH\textsubscript{4} and 1.4±0.4\% CO\textsubscript{2}, with about 6\% N\textsubscript{2} and the remainder again being heavier hydrocarbons. The isotopic value of the CO\textsubscript{2} in Naylor is $\delta^{13}C = -11\pm1$‰ VPDB (1).

The initial downhole injection pressures and temperatures were 17.8 MPa and 63 \textdegree C (compared to a reservoir temperature of 83 \textdegree C); by the end of injection the injection pressure was 19.3 MPa and the temperature of the injected CO\textsubscript{2} had fallen slightly to 61 \textdegree C (2).

Risk assessment

The updated quantitative risk assessment, conducted after drilling of CRC-1 but prior to injection, is described in (3). The main risks that were identified were leakage along fault planes or up well bores, but all risks were considered to be very low. Key elements in the risk assessment included assessment of the seal by mercury injection capillary pressure analysis (4), geomechanical modeling to bound injection pressures (5), and detailed examination of tectonic risk (6), as well as geological and dynamical modeling.

Monitoring Design and Interpretation

The design and interpretation of assurance monitoring in a CCS context raises some difficult issues. Consider for definiteness the case of groundwater monitoring. Interpretation, and hence design, involves three nested problems. Firstly, there is the question of the statistical significance of a particular measurement – is it real, or is it an artifact of measurement error or natural variability? Secondly, there is an inverse problem. We necessarily make measurements sparsely in time and space. What can we infer about the aquifer as a whole from our limited suite of measurements? This involves assumptions about the location and geometry of the hypothetical leakage of CO\textsubscript{2} into the aquifer, and geological and hydrological modeling. Subsurface information is always limited and so there will be much uncertainty in the inversion from point measurements to aquifer-wide conclusions. Finally there is a second inverse problem; if we are
confident that we have detected leakage of CO\textsubscript{2} into the aquifer, how did it get there? And what can we conclude about where else it may have gone, and how much has left reservoir containment and why? This will involve yet more modeling, subject again to uncertain assumptions and limited knowledge of the sub-surface.

The statistical aspects are of themselves not straightforward. If measurements are subject to uncontrolled fluctuations, there is a risk of a false alarm and ideally the probability of false alarms would be known. This does require good understanding of the statistics of the data and points to the need for a substantial pre-injection campaign to characterize storage sites. There is also the question of the sensitivity of the measurements: what size of contamination or leakage will give a signal that we will regard as being statistically significant? Clearly these false-alarm rates and sensitivities will themselves be subject to significant modeling uncertainty and may even accidentally not include consideration of what turns out to be the actual leakage mechanism.

We therefore distinguish two types of interpretation. Pure assurance means that the asset in question – say groundwater – has not changed, to within natural variability. Knowing sensitivity means that upper limits on changes can be translated into upper limits on the amounts of CO\textsubscript{2} that could have entered the zone of investigation. The notion of a conditional sensitivity is useful: it is the sensitivity given that CO\textsubscript{2} has entered the zone in some simple way, typically as a temporally constant point source. Conditional sensitivities could be chained together for a range of scenarios for leakage up through the overburden.

Very similar discussions could be given of all of the monitoring techniques we have used, both for assurance and containment. Because of the difficulty in interpreting monitoring data, there is safety in numbers, in two ways. Firstly we may hope than an unforeseen leakage event will manifest itself in several ways and the combined information will be easier to assess. Secondly, it is clear that more detailed investigations can be initiated on the basis of initial, albeit uncertain clues. One may envisage a cascade of investigations, becoming successively more detailed, precise, and of course expensive.

In the case of the Otway project, being a research and demonstration project, the monitoring design approach was to monitor as many domains as possible, aiming for a good understanding of natural variability over hours, days, months and seasons. Analytical errors were kept low enough confidently to detect these variations or (in the case of the reservoir measurements) to detect the expected effects of filling with CO\textsubscript{2}. This approach was particularly important for the seismic program: early calculations showed that it would not be possible to detect CO\textsubscript{2} in the reservoir, but nonetheless the highest possible data quality was aimed at, and proved very useful in demonstrating the absence of CO\textsubscript{2} above the reservoir at small mass levels.
The Naylor-1 Bottom-Hole Assembly

The bottom-hole assembly (depicted schematically in Figure 3) that was deployed in Naylor-1 for monitoring included hydrophones and geophones that were intended for microseismic sensing, vertical seismic profiling, and travel time measurements. Many of these failed shortly after deployment, as did the pressure and temperature sensors. This is thought to be due to corrosion at downhole electrical connections. Pressure data were available from the production phase of Naylor-1, and a downhole gauge was installed to monitor temperature and injection pressure in CRC-1. Surface pressures were available at Naylor-1, which could be corrected to reservoir pressure once the composition of the borehole fluids was known. A backup microseismic system was installed in a shallow water well close to Naylor-1. The U-tube sampling system was also part of the bottom-hole assembly.

Naylor Fluid Sampling Results

Injected CO$_2$ (dissolved in the water phase) arrived 121 days after injection at the Naylor U-tube-2 sampling port (1, 7, Figure 6). This is located just below the gas-water contact. By 177 days concentrations had peaked and the arriving CO$_2$ was mostly in the free gaseous phase, the gas-water contact having been displaced downwards by the injected CO$_2$. Tracers arrived with the CO$_2$ and their concentrations followed a similar pattern. The breakthrough curves agree with predictions, although the CO$_2$ arrived about 50 days before the mid-range of predictions. This difference is not significant compared to the spread in the predictions resulting from uncertainties in geology and relative permeabilities. Results from U-tube-3 (located 5 m deeper) are similar to U-tube-2, showing a slower and more irregular progression to arrival of free CO$_2$.

A total of 11 water samples from U-tube 2 and 23 from U-tube 3 were collected covering a period from pre-injection to self-lift. The pre-injection water samples display evidence of KCl (kill fluid) contamination (~15%) introduced during the recompletion of Naylor-1 to re-perforate and install the U-tube bottom hole assembly. The kill fluid content decreased to under 0.5% in the samples collected after injection began and remained low (predominantly <3%) for remainder of the sampling period. Baseline conditions are believed to have been achieved by day 68 after injection began.

From day 68, the temperature and pressure corrected pH values display an initial increase then a decreasing trend from day 121 to gas lift, coincident with the increasing CO$_2$ content in the exsolved gas phase (Figure S1). The increase in pH, and Ca$^{2+}$ and HCO$_3^-$ reflect the arrival of the front where water displaced by the injected CO$_2$ phase is mixing with the initial formation water. Reactive transport modeling indicates the displaced water has dissolved some carbonate minerals and elevated Ca$^{2+}$, Mg$^{2+}$, Fe$^{2+}$ and HCO$_3^-$ are the result while the early increase in pH is a product of mixing. The dissolved silica content also increases suggesting the possibility that silicate mineral dissolution is taking place. The Waarre C mineralogy consists primarily of quartz and potassium feldspar with very minor amounts of illite, kaolinite and ankerite. The low reactivity of these minerals
(apart from ankerite) results in very little CO$_2$-water-rock interaction and very low likelihood of injectivity or fluid migration being impacted.

**The Seismic Surveys**

To optimize the seismic monitoring program, we conducted several test surveys (pre-baseline) in the period 2006-2008. Since the area is actively farmed there were restrictions on access and the type of seismic source that could be used. The highly variable water table in the surface karst restricted surveys to the dry months. The seismic program was designed to maximize signal-to-noise ratio (minimize non-repeatability) by employing very high fold and keeping source-receiver lines close together (8, 9).

A baseline 3D survey was shot in January 2008, using a weight-drop source. The survey area was 1.6x1.9 km, average fold about 100. The first repetition (the first monitor survey) was acquired in 2009, using a vibroseis source. The second monitor survey was taken in 2010.

**Figure S1.** U-tube 2 water chemistry data showing changes in pH, Ca$^{2+}$, HCO$_3^-$ and dissolved silica (SiO$_2$$_{(aq)}$) with time. Breakthrough of the injected CO$_2$ at the monitoring well was recognized on day 121. The chemical changes reflect the arrival of fluids that have interacted with the CO$_2$ and minerals present as it is displaced by the injected CO$_2$ front. The significant drop in pH after day 150 marks the arrival of the injected CO$_2$-water interface near the U-tube 2 inlet.
acquired in 2010, also with vibroseis. These surveys were shot after 35,000 and 65,445 tonnes of CO\textsubscript{2} had been injected. Despite the use of different sources, processing of all the three surveys produced similar results; the best time-lapse signal-to-noise ratio was achieved with 2009/2010 data (10).

Geological Modelling

The Otway Basin is a large, northwest trending basin located on the southern Australian passive margin. It formed during the Jurassic to Cainozoic (from 140 to 6 million years before the present). During rifting and continental breakup of what is now Australia and Antarctica, large sedimentary depocentres opened up which were subsequently faulted and inverted during phases of rift, sag and compression. These episodes resulted in complex fault structures and compartments providing traps and anticlines for hydrocarbon accumulations where reservoirs are juxtaposed to seals. Commercial gas discoveries include the offshore Thy lacine, Geographe, Minerva and Casino fields, and numerous smaller onshore gas fields. Several of these are rich in CO\textsubscript{2}, derived from a late phase of volcanism during the Pleistocene with CO\textsubscript{2} influx occurring between 2 million to as recently as 5,000 years ago (11). The Waarre Formation is the main gas bearing reservoir in the onshore region.

Geological modeling for the project fell into two broad phases. The first, characterized by limited data, was focused on establishing a suitable location for the injection well. In conjunction with flow modeling, the objective was to find a location that would ensure migration of injected CO\textsubscript{2} to the Naylor-1 monitoring well comfortably within the funded lifetime of the project. The second phase was relatively data-rich, benefiting especially from the drilling of the injector well. Again in conjunction with flow modeling, the objective was to predict the details of the breakthrough at Naylor-1, especially delineating the range of possibilities implied by geological and rock physics uncertainties. Forward modeling of the seismic response, and associated uncertainties, was also an important aspect of this phase.

During the initial site screening process there were limited data available to guide the reservoir modeling. Well information was supplemented from regional data and extrapolated from adjacent fields to develop an initial set of geological models. This phase is described in detail in (12). Adding to the uncertainty, at the time there were two regional paleo-depositional models considered plausible for the Waarre C Formation - a transgressive shoreline model, with the main depositional trends perpendicular to the expected CO\textsubscript{2} flow direction (13), and a braided fluvial model (14) where reservoir sands are highly connected and deposited parallel to the direction of flow. A series of static geological models was created using PETREL (15) to investigate the optimal location for the injector well CRC-1. Along with the two depositional model cases mentioned above, two additional extreme cases for reservoir connectivity were also investigated - high permeability and no shale baffles, and low permeability with major barriers to flow. These were considered geologically implausible but potentially possible. Uncertainty in the seismic interpretation was also investigated by modeling several different possibilities for the structural dip of the reservoir that would result from shifts in the depth conversion
of the horizons. This was important to assess the risk of breakthrough at the monitoring well Naylor-1 occurring too early to record meaningful results, or happening too late in the funded lifetime of the project.

All of the cases were simulated using the ECLIPSE flow modeling package (16) after first history-matching the production data from Naylor-1. History matching to production data revealed little agreement with the extreme cases so these were discounted, but a good match to the two intermediate plausible cases was found. Results for breakthrough for these cases were mainly impacted by the distribution and size of shale baffles (times ranged from 6-14 months). The detailed processes leading to the choice of location of the injection well CRC-1, and the associated refinements of the models, are described in detail in (7).

Attention then turned to the improvement of the geological model. Extensive log and core data were then collected during the drilling of CRC-1 (17) in order to reduce the geological uncertainties. In addition new data were acquired in Naylor-1, including Vertical Seismic Profiling (VSP), contributing to an improved database for understanding the velocities for time to depth conversion of the horizons. New wire-line logs were run (reservoir saturation RST and total porosity TPHI) to define the level of the post-production/pre-injection gas-water contact. A detailed sedimentological description of the new core identified the heterogeneous nature of the Waarre-C reservoir which contains sandstone bodies of varying grain sizes and thin (1 m - 3 m) shale baffles. The core was then correlated to the down-hole logs and six depositional facies were identified on the basis of electro signature, sedimentary setting, and distinctly different reservoir potential related to grain size, composition and sorting (17).

Existing pre-production 3D seismic was reinterpreted, with the new well tie at CRC-1 confirming the top and base of the reservoir and the geometry of bounding faults. The new CO2CRC baseline 3D seismic survey was also interpreted and was able to resolve the reflectors of the top and base of the reservoir confirming the well tie. The new structural model was well constrained by the new VSP data in the closely spaced wells and check-shot data from the nearby Naylor South-1 well just to the south of the main field (Figure S2). The structural dip was estimated at around 14°. The throw of the main Naylor fault was examined at the point where juxtaposition of sand on sand turned to sand on shale and so an estimate of the maximum depth for the structural spill point could be obtained.
Stress estimation was carried out to assess the risk of fault activation resulting from the expected pressure increases from injection (18). Extended leak-off test data, a borehole wall electrical image and dipole sonic log data in the CO₂ injector CRC-1 were used to constrain principal horizontal stress orientation and magnitudes. Consistency of the stress model was then checked against the occurrence of breakouts using a mechanical earth model built along the CRC-1 well. Results indicate minimum horizontal and maximum horizontal stress gradients on average equal to 15.98 and 18.13 MPa/km, respectively, corresponding to a normal stress regime. This information then fed into the geomechanical model to ascertain the likelihood of fault reactivation under the current regime (5).

After including data from CRC-1, new static models were created which incorporated improved geological information such as porosity, permeability, pressure, and the geometry of the reservoir including faults, sedimentary layers, and facies (rock
types) distribution. These are the primary characteristics controlling the behavior of stored CO$_2$. The underlying PETREL geo-cellular grid was based on a UTM (Universal Transverse Mercator) projection of a 20 m x 20 m optimally oriented (for the direction of flow) grid. Layering was 0.5 m to 2 m thick and upscaling of the petrophysical logs to this resolution seems to capture the vertical variation in the data and adequately represent vertical permeability. The PETREL model and subsequent ECLIPSE model used for pre-injection modeling (7), used an irregular cell geometry which honors the geometry of stratigraphic bedding, i.e. on-lap and erosional surfaces between the incised valley fill and the overlying transgressive sands. The ECLIPSE grid maintained this geometry and also incorporated local grid refinement to investigate near-well bore effects. The ECLIPSE cornerpoint simulation grid was converted to an equivalent PEBI grid for use with the TOUGH2 simulation software, but without the local grid refinement.

Unlike simulations of saline aquifer storage, in which it is necessary to refine the grid vertically beneath the top seal in order to resolve thin layers of CO$_2$, here the injected CO$_2$ fills downwards beneath the methane-rich gas cap. Thus the vertical grid spacing was more regular, starting with 0.5 m at the top and going at 2 m at the bottom of the Waarre C.

Because of the strong relationship identified between the depositional facies and reservoir quality, the new geological models used facies objects (sand channels and shales) to constrain the spatial arrangement of permeability streaks and low flow baffles between the wells. These were generated by sequential indicator simulation, as implemented in PETREL (15). Similarly a method of incorporating facies based permeability anisotropy was developed. Ratios of vertical to horizontal permeability $K_v/K_h$ from core measurements were applied as multipliers for each facies - this is an improvement on just using one ratio for the whole reservoir. Depositional analogues, appropriate to this type of setting, indicated that the permeability conduits within channels would be expected to be highly connected over the distance of 300 m between the injector and monitoring well. Because the variogram lengths for sands exceeded the correlation distance there was less uncertainty to cover in the stochastic modeling.

On the other hand, there was still uncertainty in the distribution of shale barriers as analogues suggest that they are not expected to be greater than 80 m to 100 m wide due to down-cutting channels eroding the fine grained sediment. Both Naylor-1 and CRC-1 intersected at least two 1 m-3 m thick shales. The main uncertainty was whether these were continuous or truncated between the wells. This would have implications for interpreting vertical connectivity between the injection perforations and the U-tubes which spanned these shales in both wells.

Two geological cases were considered to quantify this key uncertainty. Case 1 used a small correlation length for the shales (60 m-80 m), and case 2 used a long correlation length (120 m-240 m). Five equiprobable realizations were generated for each case giving a total of 10 models, using the Sequential Gaussian Simulation algorithm implemented in Petrel (15). In this model, cells in the model intersected by the wells always honor the well data. A code is given for the depositional facies that were
observed in logs and core. Cells between the wells are then assigned codes according to rules governing the probability of variance depending on their spatial distance away from the wells using Kriging to a normal distribution. The co-variance is defined by a variogram that allows the geologist to condition the data to a range for length and width of each facies based on observations from natural analogues. This method is used to add a random component between the wells; however the shapes and geometries of the facies still must be geologically plausible. The petrophysical properties are then conditioned to the facies model. This means that porosity and permeability values are populated according to the depositional facies codes, for example low values for the shale baffles and so on. The two cases that were generated are illustrated in Figure S3. Two realizations were randomly selected for each case and for each of these, two relative permeability curves were considered, giving a total of eight models that were finally investigated in detail.

**Figure S3** Two example realizations for the sand and shale distribution for the two geological cases eventually considered. Sand is yellow and shale is grey. The overlying Flaxmans Formation and a portion of the Belfast Mudstone is shown as well as the underlying Waarre B shale unit at the base of the reservoir. The position of the wells and gamma ray logs colored to represent sand and shale lithology is also shown; the wells are shown as a 3D representation of a conventional gamma-ray well log, with the injection well CRC-1 on the left.
Dynamic Modeling

The multiphase flow simulation software TOUGH2 (19) was used for the dynamic model (20). TOUGH2 has been widely used in modeling CO₂ storage projects, and the results of two code comparison efforts (21,22) indicate that the quality of the numerical solutions obtained with TOUGH2 are quite comparable to those from commercial packages and other research codes. The TOUGH2 simulation results on this project are quite similar to the earlier ECLIPSE work (7).

The areal extent of the reservoir model was about 800m by 800m, with a 20 m grid spacing laterally. The average reservoir thickness was about 25m, with variable vertical spacing (0.5-2 m, as discussed above), and the grid had in total 45,000 blocks. Earlier studies on lateral grid refinement found that this was adequate given the modest lateral extent of the model. Given the computational burden of doing history matching on each realization, with multiple runs on each to adjust parameters, a much more detailed grid would have been hard to manage. The reservoir model was bounded by faults on three sides, which were considered as no-flow boundaries. The other side was attached to a large aquifer, and the parameters of this were determined by the history match to the pressure.

As noted, for detailed modeling eight models were used. These were believed to span the range of uncertainty in geology and rock physics that remained after the extensive data gathering of this phase following the drilling of the injection well CRC-1.

Production rates and wellhead pressure were available for the period when the production was in progress at Naylor-1. The location of the post-production gas-water contact was obtained by logging this well. During the injection phase, surface data on CO₂ injection rates, and reservoir data from downhole pressure and temperature gauges in CRC-1 were also used (2). Each dynamical model was successfully adjusted to match pressure history from the production and post-production phases. Models were mainly sensitive to the aquifer parameters, which largely determine pressure recovery after depletion. For the injection phase the model was calibrated with downhole pressure data from CRC-1. The main sensitivity was to the bulk permeability, which had to be adjusted separately for each permeability realization considered. Bulk reservoir permeability had to be constrained by pressure history and core observations, because of the failure of an injectivity test. This test failed because the wellbore was not successfully perforated on the first attempt, and could not be repeated because of cost.

Figure S4 shows the underlying detail of the reservoir model predictions that are represented in Figure 7 of the main text. The eight curves correspond to four geological models: two realizations of case 1 with short correlation lengths (60 – 80 m) in the shale baffles, and two realizations of case 2 with long correlations (120-240m) in the shale baffles. For each geological model, two models for reservoir relative permeability (hysteretic and non-hysteretic) were used, based on core measurements (23). The absolute permeability of all the grid blocks in each realization was adjusted by means of a
single multiplier in order to provide a good match to the downhole pressure record at the injection well. It is noticeable that the peak observed values of the SF₆ tracer exceed the predicted concentrations. This is probably due to a combination of the particular heterogeneity and grid effects in the model causing the tracer plume to disperse slightly more than in reality. However given the large dilution that already occurs due to injection, and the typical variability between successive tracer measurements, the agreement is still encouraging. We note that the predicted values for SF₆ are not fitted to the tracer data at all and follow entirely from the pressure history matching and the geological and dynamical models.
Figure S4. Detailed comparison of measured and predicted breakthrough curves, for CO$_2$ (a) and SF$_6$ (b). The flow models are based on geological models with long and short correlation lengths (c1 and c2 in the key) with r1… indexing different statistical realizations. Two relative permeability curves were used, non-hysteric (nh) and hysteretic (hy). The renditions in the main text were created by blurring the model curves with a Gaussian of standard deviation three times the nominal standard deviation (5%). This estimate accounts only for analytical error, not error associated with the sampling process.
Groundwater, soil gas, and atmospheric measurements

The locations at which these assurance measurements were made are shown in Figure S5a and S5b.

**Figure S5** (a) General map of the Otway site area, showing the location of soil gas surveys, monitored water bores, and atmospheric monitoring stations.
Groundwater was sampled twice yearly from 21 shallow (<100 m) bores in the unconfined Port Campbell Limestone and from 3 deeper bores (>800 m) in the confined Dilwyn Formation. A wide range of measurements were made, including pH, electrical conductivity, redox potential, and the concentration of reduced iron (Fe$^{2+}$, bicarbonate (HCO$_3^-$) and various major, minor and trace inorganic species. Isotopic compositions of hydrogen and oxygen in water, carbon in dissolved bicarbonate and sulfur in dissolved sulfate were also determined. Vadose zone soil gas composition (CO$_2$, CH$_4$, He, H$_2$, $\delta^{13}$CO$_2$, $^{14}$C – in 2007) was also measured annually during summer. It proved impractical for the soil gas sampling to revisit the same grid of sampling points, although substantially the same area was covered each season. Both groundwater and soil gas samples were checked for the presence of the added tracers.

Comparisons of pre- and syn-/post injection data for soil gas and groundwater are in Figures S6 and S7. The only statistically significant change is a small rise in median pH for the aggregated water samples, combining all bores and comparing data collected before and after injection started.
Figure S6. Histograms of electrical conductivity, bicarbonate concentration, and pH of samples from the shallow (Port Campbell Limestone) aquifer, aggregated into pre-injection samples (left bars) and post-start-injection samples (right bars). There are 78 pre-injection measurements and 77 post-injection. Only the small increase in pH is statistically significant.
Figure S7 CO2 concentrations vs. δ¹³C CO₂ values for the four soil gas surveys (A) March 2007 and (B) February 2008 are pre-injection, (C) February 2009 and (D) March 2010 are during injection. The Otway average natural gas composition is according to (9), and the value of δ¹³C CO₂ for the injected gas is indicated by the horizontal dashed line.
We modeled leakage of CO$_2$ into the Port Campbell limestone as a point source at the base of the aquifer, to gain some insight into the conditional sensitivity of the groundwater measurements. The plume of CO$_2$-rich water that forms is quite narrow and sharp-edged, essentially because of the limited solubility at these shallow depths; most of the transport of CO$_2$ upward is simply in the gas phase. Only at horizontal barriers does the CO$_2$ spread out appreciably. These results are in accord with other modeling (e.g. 24). We know rather little about the Port Campbell Limestone and very little about its heterogeneity, but for a range of plausible parameters (horizontal permeability 1-10 Darcy, vertical permeability 1 – 10 Darcy, porosity 0.3, 1000 tonnes yr$^{-1}$ leakage rate into the aquifer) we find that the plume is about 10 to 20 m in radius. The aquifer’s flow rates are known to below (about 0.14 m/day). If we assume that a leak such as this could be anywhere within a square kilometer of the injection site, and has lasted a year, it follows that the area affected by the plume is only 0.5% of the square kilometer. The chance of a single well intersecting the hypothetical plume is very low, unless there are significant barriers to vertical flow. These would cause much wider spreading of the plume, but in a thin layer that then might not coincide with screened intervals in the wells. The groundwater measurements therefore tells us little about containment although they are very important for assurance. In general an aquifer would have to be well-characterized, and probably also penetrated by a large number of wells, before measurements of water chemistry would be useful guides to leakage.

The atmospheric composition measurements comprise continuous CO$_2$ and concentrations of CO$_2$, quasi-monthly samples for other tracers (including CO and the injected tracers SF$_6$ and HFC134a). Initially only CO$_2$ was measured continuously but the system has been gradually upgraded to include continuous measurements of CH$_4$ and the carbon-13 isotopic ratio of CO$_2$. Ecosystem CO$_2$ fluxes were measured continuously with an eddy-correlation flux tower and during six campaigns with a portable soil flux chamber. Figure S8 gives more detail on the drilling rig detection discussed in the main text.

It is possible to predict much of the variability in the atmospheric composition using an environmental model, which greatly improves the detectability of changes (25). The use of tracers increases the ability to discriminate anomalies in atmospheric composition potentially caused by leakage from emissions from other sources. For example, CH$_4$ is naturally present in the injection stream and SF$_6$ was injected as a tracer. Both have background atmospheric concentrations that are lower and steadier than CO$_2$. CO identifies combustion sources of CO$_2$, unrelated to geological storage. The magmatic origin of the CO$_2$ from Buttress gives it a $^{13}$C isotopic signature that distinguishes it from that produced by the local ecosystem, agricultural and anthropogenic sources (26,27). Such $^{13}$C isotopes in CO$_2$ have the advantage of being naturally occurring and acting as conservative tracers. However, most anthropogenic CO$_2$ (for example, from coal combustion) will be less distinct from ecosystem CO$_2$ in its isotopic signature. Tracers are also useful markers in soil gas and water monitoring, although all the tracers used (except CD$_4$) are present in the environment to some degree.
Our atmospheric measurements are made at a single location and hence can only be inverted to find the strength of a source if its location and geometry are assumed (28); the technique is nonetheless valuable for monitoring the most likely sites of leakage – the wells – and could be extended by using a network of strategically located sensors (29).

**Figure S8.** Detection of emissions from the drilling rig at the injection well CRC-1, as a test of the strategy and sensitivity of the atmospheric monitoring at Otway. CO$_2$ concentrations (left) measured at the atmospheric station 700 m northeast of the emissions were selected for well mixed daytime conditions based on meteorology including wind speed and boundary-layer stability (Monin-Obukhov length) both at right. During daylight ecosystem CO$_2$ fluxes (from the nearby eddy covariance flux station) were near zero or negative indicating uptake due to photosynthesis. Erratic CO$_2$ increases (indicated by vertical lines) above the background (Cape Grim baseline CO$_2$, red) then imply a local CO$_2$ source up wind of the measurement location. This corresponds to wind directions from the general direction of the drilling site of CRC-1 (right centre, horizontal lines). On 15 March the CO$_2$ was elevated by about 2.5 ppm and variable, indicating a point source. This was traced to the drill rig using wind trajectories and confirmed by measurements of flask air samples for CO (emission ratio of 10 ppb CO:1 ppm CO$_2$) and the $\delta^{13}$C of the extra CO$_2$ (-16 per mill) as combustion engine emissions. Emissions estimates (from fuel consumption) were about 6 t CO$_2$/day. Dispersion modeling (e.g 16) simulated the CO$_2$ increases at 1-5 ppm.
Forward modelling of seismic response

Realizations of the numerical flow simulations at the times of the two surveys were used to predict the expected seismic time-lapse result at the reservoir level. These realizations were based on the various equiprobable geological models, and compared to the time-lapse surveys (10). Apart from rock properties and the fluid distribution from simulation, an accurate representation of the bulk modulus of gas mixtures was needed. The open source software Delivery (30) was then used for forward prediction of the time-lapse seismic response. These simulations compared to Otway time-lapse seismic data repeatability analysis showed that the noise level was significantly greater than the predicted signal level, a conclusion that was consistent with earlier results obtained using simple one-dimensional modeling (31).

To evaluate sensitivity of time-lapse seismic for leakage detection a synthetic feasibility study was carried out. Using finite-difference modeling we computed ‘baseline’ and ‘monitor’ seismic surveys with similar parameters to real surveys acquired in the area. To simulate leakage between ‘baseline’ and ‘monitor’ survey a CO$_2$ plume was introduced into the seismic model at the depth of ~1.5 km. This depth interval corresponds to a saline aquifer (the Paaratte) which can hold secondary gas accumulation generated by the leakage. Parameters of the plume, such as size and gas saturation, were obtained from hydrodynamic modeling. Changes in seismic properties of the plume were predicted using a fluid substitution technique based on the Gassmann equation (32). The detection of leakage is not only a matter of the presence of time-lapse signal, it is also limited by the ratio of these signal to noise and by the repeatability of seismic data. To take this into account we introduced to the simulated seismograms the same amount of random noise as was measured in real data, with the same frequency content. By gradually decreasing plume size we determined minimum detectable leakage, assuming a point leakage at this level, to be 5000 tonnes of injected CO$_2$ (Figure 6b).

Fluid sampling details

The results from the fluid samples are robust despite the lack of physical isolation between the three U-tube inlets. Because the reservoir is produced at low rates during sampling (about 25 kg hr$^{-1}$ of gas), buoyancy effects isolate U-tubes from each other when they are producing fluids of different densities. This accounts for the similarity between U-tube-2 and U-tube-3 when both are producing water or gas, and the obvious differences in CO$_2$ and tracer content during the period while U-tube-3 is transitioning to self-lift and U-tube-2 has already done so. U-tube-1 is always isolated, being 15 m above U-tube-2. The predicted isolation of U-tube-2 and U-tube-3 from each other during the self-lift transition gives confidence in the timing of self-lift as an indicator of the location of gas-water contact during this period. This isolation was confirmed by two different (simplified) models of the fluid extraction process. In one, the wellbore was modeled as a very high permeability cylinder, albeit with slow continuous production of reservoir fluids. The second model used near-wellbore parameters from the reservoir models, but explicitly tracked fluid flow and mixing within the wellbore, using a simple advective
model. These models agreed in identifying the buoyancy isolation mechanism, and the second was able to calibrate the location of the gas-water contact in the model against the density of sampled fluid; this makes the connection to the observed phenomenon of self-lift.

U-tube-1 (Figure 3), located in the gas cap, always flowed without assistance. The injected CO$_2$ was detected 233 days post-injection and appears to have reached the upper perforations by a different pathway to that seen at the lower U tubes. This probably results from the variable lateral extent of shale baffles in this reservoir. Extending one of these baffles somewhat in the geological model results in an “attic” effect that channels injected CO$_2$ and tracer into upper compartments of the reservoir, and hence towards U-tube-1, at the observed time and concentration.

**Timing of passage of gas-water contact**

The rate of filling of the Waarre-C reservoir can be assessed from the timing of self-lift at the two lower sampling points in the Naylor-1 observation well. U-tube-2 transitioned to self lift on September 11th, 2008 after 21,100 tonnes injection. U-tube-3 transitioned to self lift on January 15th, 2009 after 38,100 tonnes injection, although there was a period of some weeks beforehand in which sampling produced a mixture of gas and water. However, 21,100 tonnes is not equal to the effective storage capacity of the reservoir between the GWC and U-tube-2 elevation because there is a portion of that volume which is “in-transit” between the injection and observation wells. Since the daily injection rate is reasonably constant, the volume of gas injected between the times that U-tube-2 and 3 transitioned to self lift (17,000 tonnes) should represent a reasonable estimate of the effective storage capacity of the pore space between the elevations of U-tube-2 and U-tube-3. The elapsed time interval is 126 days, and the sampling interval is 7 days, so the uncertainty in the timing is at least 10-15%, and perhaps more due to the transition period at U-tube-3. Thus the volume of gas has a proportional uncertainty of $17,000 \pm 2000$ tonnes.

Although the depths of the U-tubes are accurately known, each samples from an interval of finite width, of around 0.6m. U-tube-2 and U-tube-3 are not packed off from each other, and are adjacent to the same interval of perforations. Sampling from a U-tube induces a flow into the U-tube at the reservoir level, and the origin of this fluid will depend on the productivity (i.e. the product of permeability and thickness) of the reservoir section near the perforations, as well as possible flow within the wellbore. Since each sampling operation removes a fluid volume comparable to the free volume within the wellbore, there is also some dependence of the history of fluid withdrawal due to sampling. The difference in depth between the two U-tubes is 4.5m. Detailed modeling of the wellbore flow suggests that the movement of the gas-water contact (GWC) between the two transitions to self-lift is in the range $3.5 \pm 0.5$ m. Here the GWC is used as shorthand for the complex distribution of saturation as the reservoir fills, which depends on local porosity and permeability and fluid-rock properties such as capillary pressure and relative permeability.
Details of capacity measurement

The effective storage capacity in a region of the reservoir depends on the residual gas saturation $S_{gr}$ upon imbibition (when the field is depleted and is being repressurised by water influx), and the maximum gas saturation $S_{gmax}$ upon drainage (when the injected gas displaces some of the water that has occupied the pore space). $S_{gmax}$ should be less than the $1-S_{lr}$, where $S_{lr}$ is the original residual water saturation. The efficiency of filling due to hysteresis, which will be denoted by $C_{hyst}$, is then given by

$$C_{hyst} = \frac{S_{gmax} - S_{gr}}{1 - S_{lr} - S_{gr}}$$

If there is no hysteresis, then $S_{gmax}=1- S_{lr}$ and $C_{hyst}=1$. Note that this is only the local hysteresis, and does not take account of reservoir pore volume that is invaded by formation water that is not ultimately displaced. This calculation also implicitly assumes that the endpoints of the relative permeability curves are independent of pressure, and that the reservoir is nearly back to its discovery pressure. In fact there are more complex considerations, where gas that is trapped at residual saturation will increase in density as the reservoir pressure recovers, reducing the saturation further.

From the static geological model, the pore volume that is filled between the two transitions to self-lift can be estimated at $0.7 \times 10^5$ m$^3$ – $1.3 \times 10^5$ m$^3$ using the estimated range of movement of the GWC between the transitions. It has been estimated from logs that $S_{lr}=0.11$ and $S_{gr}=0.20$. Using the lower value for the vertical movement of the GWC (3.0 m) gives $C_{hyst}=0.75$ to 0.84. These in turn imply $S_{gmax}=0.72-0.78$. For the upper value of vertical movement (4.0 m) the range is $C_{hyst}=0.56$ to 0.62, and so $S_{gmax}=0.58-0.63$. Although there is still a considerable range of values, these measurements provide evidence that in this case $C_{hyst}<1$. The values of the relative permeability endpoints ($S_{lr}$, $S_{gr}$ and $S_{gmax}$) are a fluid-rock property, and so specific to particular sites. However it is likely that hysteretic effects, on top of pore volume loss due to aquifer influx, will further reduce the effective storage capacity of depleted gas fields with moderate aquifer drive.

Interpretation of capacity measurement

Estimates of the amount of CO$_2$ that can be stored in depleted gas fields globally are by nature very approximate, since they depend on applying average assumptions to a wide range of different circumstances. The most recent IEA study (33) discusses the methodology in detail.

The reduction in storage capacity from the theoretical capacity to the effective capacity depends on the strength of the aquifer drive, the rate at which formation water is able to re-enter the reservoir. For a field with a weak aquifer drive, the reservoir behaves more or less like a “tank”. In depleted fields with weak or non-existent aquifer support, the same pore space in the reservoir is still available for gas, and so in returning to the original reservoir pressure, one can store the same subsurface volume as originally produced. This would be the case for depletion drive reservoirs, where the gas cap provides the needed pressure for production.
In depleted gas fields with strong aquifer support, two related phenomena limit the storage capacity. The first is that as production proceeds, formation water occupies some of the pore space previously occupied by gas. If the injection pressure exceeds the aquifer pressure, then it should be possible to displace some of this formation water and recover some of the pore space. Also, during gas injection, the residual water saturation after drainage will generally be larger than it was pre-production, and this is observed as hysteresis in the relative permeability curves. Laboratory studies show that with repeated cycles of imbibition (water inflow) and drainage (gas displacement of water), more formation water is trapped in the pore space. Thus not all the pore space previously occupied by gas can be recovered.

The storage efficiencies used (21) attempted to account for the variations in aquifer support by reducing the theoretical capacity by factors derived from modeling.

For a depleted gas field with strong aquifer support, the amount of influx of formation water must be also influenced by the timing of injection. For example, production from the Waarre-C in the Naylor field ceased at the end of October 2003, when the formation pressure was around 10 MPa. For the CO2CRC Otway project, injection began in March 2008, by which time the reservoir pressure had recovered to around 17.8 MPa. This indicates a substantial influx of formation water from the aquifer system, and a consequent reduction in capacity. Thus effective storage capacity of a depleted gas field with a strong aquifer drive will depend upon the timing of injection relative to production, and the limitations on the maximum injection pressure.

For the Waarre-C in the Naylor field, the observations have so far only constrained the efficiency of filling between U-tube-2 and U-tube-3. An estimate of the effective storage capacity of the whole reservoir unit depends on taking some of the ensemble of numerical models matched to the observations, and simulating the effect of continued injection until a suitable stopping criterion is reached. This criterion might be economic or technical in nature.

A more detailed calculation would take into account the possibility of enhanced recovery of natural gas caused by the injection of CO\textsubscript{2}, plus effects such as the increasing pressure and consequent change in size of the free gas cap.
References


(11) M. N. Watson, C. Boreham, P. Tingate (2004), Carbon dioxide and carbonate

cements in the Otway Basin: implications for geological storage of carbon dioxide. 

APPEA Journal, 44, 703-720.


Formation Unit C; Reservoir Static Models. CO2CRC Internal Report RPT05-0123, 


(13) A. J. Buffin, (1989). Waarre Sandstone development within the Port Campbell 


(14) A. Faulkner, (2000), Sequence stratigraphy of the late Cretaceous Sherbrook Group, 

Shipwreck Trough, Otway Basin; B.Sc. (Hons) Thesis; National Centre for Petroleum 

Geology and Geophysics, University of Adelaide.

(15) www.slb.com/services/software/geo/petrel/geomodeling.aspx 

(16) www.slb.com/services/software/reseng/eclipse.aspx 

(18) T. Dance, L. Spencer, Q. Xu, (2009), Geological characterization of the Otway pilot 

site: what a difference a well makes. Energy Procedia 1, 2871-2878


estimation at the Otway CO2 storage site, Australia, in Proceedings of the SPE Asia 


(19) K. Pruess, The TOUGH codes - A family of simulation tools for multiphase flow 


(20) J. Ennis-King, T. Dance, J. Xu, C. Boreham, B. Freifeld et al. (2011). The role of 

heterogeneity in CO2 storage in a depleted gas field: history matching of simulation 

models to field data for the CO2CRC Otway Project, Australia, Energy Procedia 4 3494– 

3501 doi:10.1016/j.egypro.2011.02.276

(21) K. Pruess, J. Garcia, T. Kovscek, C. Oldenburg, J. Rutqvist, C. Steefel , T. Xu, 

(2004). Code intercomparison builds confidence in numerical simulation models for 

gologic disposal of CO2. Energy 29 1431-1444

(22) H. Class, A. Ebigbo, R. Helmig, H. K. Dahle, J. M. Nordbotten et al., (2009), A 

benchmark study on problems related to CO2 storage in geologic formations. 

Computational Geosci. 4, 409-434


scale heterogeneity on CO2 distribution in reservoir rocks, Transport in Porous Media, 

82, 93-109,


