available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ijggc

Constraints on the *in situ* density of CO_2 within the Utsira formation from time-lapse seafloor gravity measurements

Scott L. Nooner [a,*], Ola Eiken [b], Christian Hermanrud [b], Glenn S. Sasagawa [a], Torkjell Stenvold [c], Mark A. Zumberge [a]

[a] Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, United States
[b] Statoil Research Center, Rotvoll, N-7005 Trondheim, Norway
[c] The Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 28 July 2006
Accepted 10 January 2007
Published on line 12 March 2007

Keywords:
CO_2
Seafloor gravity
Reservoir monitoring
Sleipner
Utsira
Geologic storage

ABSTRACT

At Sleipner, CO_2 is being separated from natural gas and injected into an underground saline aquifer for environmental purposes. Uncertainty in the aquifer temperature leads to uncertainty in the *in situ* density of CO_2. In this study, gravity measurements were made over the injection site in 2002 and 2005 on top of 30 concrete benchmarks on the seafloor in order to constrain the *in situ* CO_2 density. The gravity measurements have a repeatability of 4.3 μGal for 2003 and 3.5 μGal for 2005. The resulting time-lapse uncertainty is 5.3 μGal. Unexpected benchmark motions due to local sediment scouring contribute to the uncertainty. Forward gravity models are calculated based on both 3D seismic data and reservoir simulation models. The time-lapse gravity observations best fit a high temperature forward model based on the time-lapse 3D seismics, suggesting that the average *in situ* CO_2 density is about to 530 kg/m^3. Uncertainty in determining the average density is estimated to be ±65 kg/m^3 (95% confidence), however, this does not include uncertainties in the modeling. Additional seismic surveys and future gravity measurements will put better constraints on the CO_2 density and continue to map out the CO_2 flow.

1. Introduction

1.1. *The Sleipner Project*

The Sleipner Project is the world's first commercial application of emissions avoidance through the use of carbon capture and sequestration technologies for geologic storage of CO_2. The Sleipner field is a natural gas production area located about 240 km off the coast of Norway in the North Sea (Fig. 1) and operated by Statoil. In order for natural gas drawn from the site to meet commercial specifications, its CO_2 content must be reduced from about 9 to 2.5%. In gas fields worldwide, this excess CO_2 is typically vented into the atmosphere, but at Sleipner the CO_2 is compressed and injected into a porous saline aquifer known as the Utsira formation (Fig. 2). Injection began in 1996; now about 1 million tonnes (MT) of CO_2 are being separated from the natural gas and injected into the Utsira formation each year.

Because this is the first industrial-scale project of CO_2 injection into a geologic formation for environmental sequestration, monitoring the CO_2 is useful in confirming that this is a safe and reliable sequestration option. Previously, time-lapse 3D seismic surveys have been successfully employed to image the underground CO_2 (Arts et al., 2003; Chadwick et al., 2004).

* *Corresponding author at*: Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, United States. Tel.: +1 845 365 8944; fax: +1 845 365 8156.
E-mail address: snooner@ldeo.columbia.edu (S.L. Nooner).
1750-5836/$ – see front matter
doi:10.1016/S1750-5836(07)00018-7

Fig. 1 – A map of the southern Norwegian coastline. The location of the Sleipner platform is indicated by a square in the lower left corner of the map.

In this study, we obtain seafloor gravity measurements in 2002 and 2005. We also construct a series of gravity forward models using the results from both the seismic surveys and reservoir simulation models that were computed by the independent research company SINTEF. These are used to help interpret changes in gravity from 2002 to 2005 in order to place bounds on the *in situ* density of CO_2.

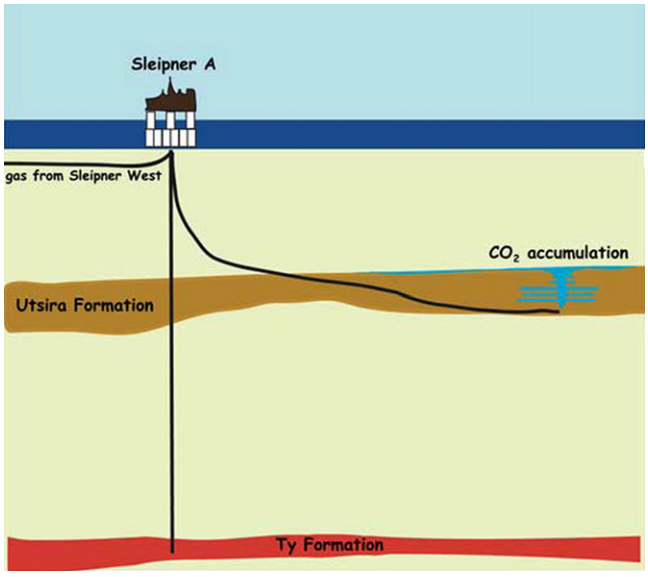

Fig. 2 – Cartoon illustrating the CO_2 injection operation at Sleipner. CO_2 is separated from the incoming gas, then injected into the Utsira formation.

1.2. *The Utsira formation*

The Utsira formation is a long, narrow sandstone formation spanning a large portion of the central North Sea (Chadwick et al., 2004; Zweigel et al., 2004). Near the injection site the aquifer extends from a depth of about 1100 m below sea level (bsl) to about 800 m bsl, where it is capped by a 200–300 m thick shale caprock. It consists of fine to medium grained, moderately well sorted sand, cut by intra-reservoir shale layers with an average thickness of about 1 m and vertical separation of 30 m (Zweigel et al., 2004). Porosity of the sand was determined to range from 35 to 40% (Holloway et al., 2000; Zweigel et al., 2004), and the sand is almost completely unconsolidated. The Utsira sand is about 300 m thick in the Sleipner area, but the shale layers segment the sand into 30 m sections, on average. The direct overburden consists of clay rich sediments with a thickness of about 250 m. The injection point is at a depth of 1012 m bsl and the water depth is about 80 m.

1.3. *Time-lapse reflection seismic surveys*

In addition to a pre-injection 3D seismic survey in 1994, 3D seismic data were collected in 1999, 2001, 2002 and 2004. The seismic monitoring surveys all clearly show a signal from the injected CO_2 (Fig. 3). By 1999, the CO_2 had reached the top of the Utsira sand and has since been spreading laterally as more of the CO_2 has migrated upwards. High amplitude sub-horizontal reflections are caused by accumulation of CO_2 under the thin inter-reservoir shale layers (Arts et al., 2003; Chadwick et al., 2004), which act as temporary barriers to buoyantly driven CO_2 flow. The decrease in P-wave velocity due to the presence of CO_2 causes seismic pushdown, as events beneath the CO_2 layers are delayed in travel time. Pushdown can be seen on the seismic data in Fig. 3 as an apparent downward dip in the reflective layers, increasing towards the center. The area-integrated pushdown (Chadwick et al., 2005), which depends on both the amount of CO_2 and the CO_2 saturation, has increased proportionally to the amount of injected CO_2.

Chadwick et al. (2000) modeled seismic velocity as a function of CO_2 saturation using Gassmann's relationships (e.g. Han and Batzle, 2004; Mavko and Mukerji, 1998; Nolen-Hoeksema, 2000; Wang et al., 1998). Fig. 4 shows the P-wave velocity versus CO_2 saturation for the Utsira sand as predicted by Gassmann's relationships for different values of reservoir temperature (compare the uniform saturation curves). Because the rock matrix in the Utsira sand is weak, the compressional velocity is sensitive to the compressibility of the fluid, which varies with temperature. The CO_2 density is assumed to be either 550 or 700 kg/m^3 in Fig. 4.

By assuming the density of CO_2 within the reservoir is 700 kg/m^3 Chadwick et al. (2005) used seismic reflection amplitudes and pushdown to estimate an *in situ* mass of 2.01 MT compared to the known injected mass of 2.35 MT for 1999. In their model, the CO_2 within the reservoir was partitioned between high saturation thin layers and a low saturation volume existing in a diffuse form between the layers. The diffuse CO_2 is presumably a consequence of CO_2 percolating upwards from the layers through the overlying

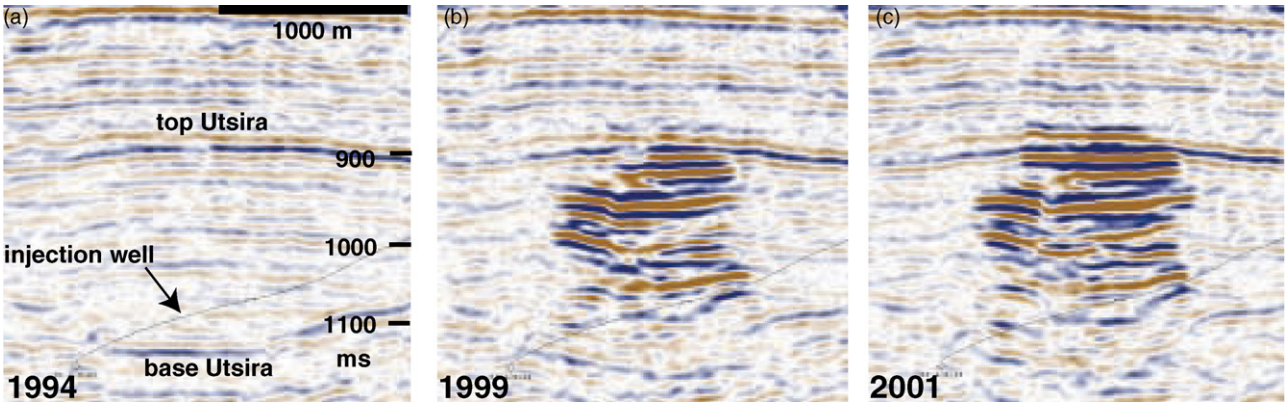

Fig. 3 – This figure shows the time-evolution of an east–west slice through the 3D seismic reflection data. (a) Seismic profile before CO_2 injection, (b) after 3 years of injection, and (c) after 5 years of injection. Data for this figure were provided by the SACS Consortium.

shales. However, reservoir flow simulation models indicate that extensive clouds of diffuse CO_2 are difficult to produce physically. The diffuse mass was assumed to have a vertically uniform distribution, yielding a minimum value for the CO_2 mass required to fit the data, since non-uniform or 'patchy' CO_2 distributions require more mass for a given seismic pushdown (Mavko and Mukerji, 1998). However, this makes it impossible to determine the mass of contained CO_2 without first determining the CO_2 saturation and density. In fact, two of the three quantities must be known in order to determine the third. Examples of two velocity versus average CO_2 saturation curves for patchy saturation models are shown in Fig. 4. Furthermore, dissolution of up to 15 and 20% of the CO_2 into the formation water is a fundamental process of the injection (Johnson and Nitao, 2003) which decreases the observed mass.

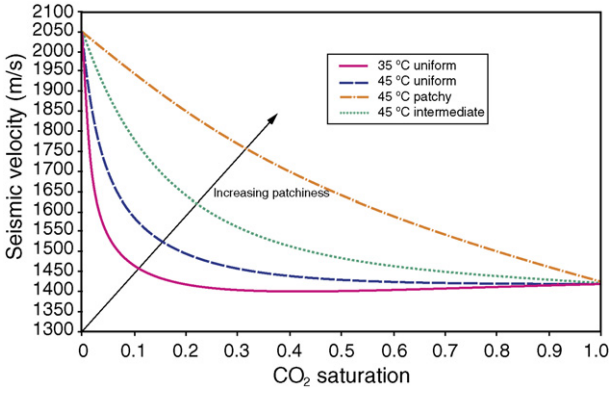

Fig. 4 – Velocity vs. CO_2 saturation curves derived from Gassman's relationships for the Utsira formation. The solid curve is for a uniform saturation with the reservoir temperature of 35 °C and $\rho_{CO_2} = 700\,kg/m^3$. The dashed curve is for a uniform saturation at 45 °C and $\rho_{CO_2} = 550\,kg/m^3$. The dash-dot curve is for a patchy saturation at 45 °C and $\rho_{CO_2} = 550\,kg/m^3$, and the dotted curve is for the same reservoir temperature but with an intermediately patchy distribution of CO_2. The two temperatures represent the expected end member scenarios for the reservoir.

Also, incomplete ability to resolve the pushdown associated with the main chimney can cause the mass estimation to be low.

1.4. *Utsira temperature and CO_2 density*

One of the largest sources of uncertainty in seismic estimates of CO_2 mass comes from uncertainty in the density of CO_2 within the Utsira formation. The density of the injected mixture depends on the amount of trace impurities, temperature, and pressure. The carbon dioxide injected at Sleipner contains nitrogen (0.063%), ethane (0.123%), methane (0.9%), which tend to lower the density, and BTX (butanes, toluenes, and xylenes, 0.667%), which tend to increase the density. For this study, the impact of these impurities are neglected, meaning the thermodynamics and equation of state (EOS) for pure CO_2 are used (e.g. Span and Wagner, 1996).

The temperature profile through the formation is based on a single downhole measurement of 37 °C at a depth of 1058 m bsl (Lindeberg et al., 2000; Zweigel et al., 2004), made at the time of drilling. For a water depth of 80 m and assuming 4.8 °C on the seafloor, this gives a linear temperature gradient of 33 °C/km. However, this single measurement is subject to an uncertainty of up to 10 °C (Hermanrud, 1988; Rider, 1986; Williamson et al., 2001), because the measurement was made before the fluids in the borehole reached equilibrium, something that can take several months.

The problem is minimized when continuous temperature monitoring during drill stem tests is done. At the Sleipner natural gas field a total of 21 different drill stem tests measured a reservoir temperature of 101.7 °C with a standard deviation of 0.5 °C at 2600 m bsl (Hermanrud, 1988). This temperature was used as a basis for thermal modeling of the temperatures in the overburden rocks at Sleipner, including the Utsira formation.

The thermal modeling was done using BasinMod 1D (March 2005 release). Model parameters consisted of a thermal conductivity of 1.2 W/mK for shales and claystones, a thermal conductivity of 2.4 W/mK for sandstones, and a seabed temperature of 5 °C. The heat flow was calibrated to match the observed temperature of 101.7 °C at 2600 m bsl. A porosity of

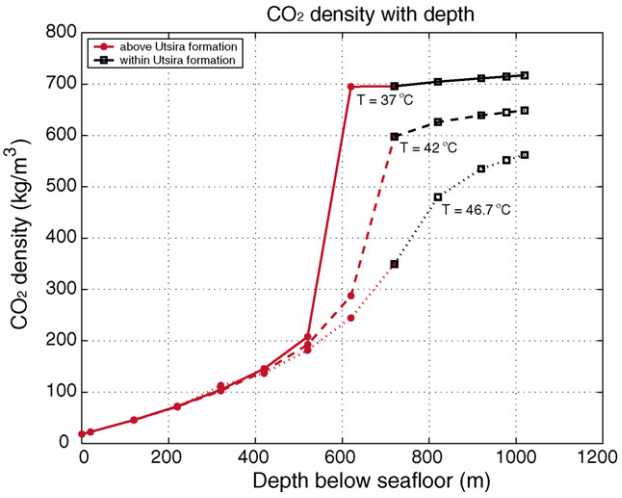

Fig. 5 – CO_2 density vs. depth for three possible temperature profiles in the Utsira formation. Because CO_2 goes through a critical phase transition, the resulting density is highly dependent on temperature. Thus a change in temperature of 5–10 °C can change the density estimate of the CO_2 by a factor of two.

0.37 was used for the sandstone within the Utsira formation. The modeling predicts a temperature of 36.2 °C at the top of the Utsira formation (810 m bsl), with a thermal gradient of 26 °C/km within the Utsira. The average thermal gradient in the overburden section has is 38 °C/km. Sensitivity checks were performed by increasing the sandstone matrix thermal conductivity to 4 W/mK and by increasing the claystone thermal conductivity to 2 W/mK. In both cases, the temperature at the base of the Utsira formation changed by 1.5 °C or less.

Further uncertainty in the calculations is related to how seabed temperature fluctuations during the last million years have influenced temperature in the underlying sediments. According to Mjøen (1988), seabed temperature fluctuations of 7 °C in the Pliocene/Pleistocene have resulted in corresponding temperature fluctuations of 2 °C at 1 km burial depth. By assuming a temperature ranged from 0 °C at the base of the ice sheets to 7 °C, the average seafloor temperature is approximately 4 °C. As the present day temperature of the seafloor is 3 °C above this average, the temperature of the Utsira formation is expected to be 1 °C higher than the average Pliocene/Pleistocene value. Therefore a seafloor temperature of 5 °C was used in the simulations. These arguments and modeling results suggest that the virgin rock temperature of the Utsira formation at 1058 m bsl is 42.5 °C, with an uncertainty (standard deviation) of 1 °C.

However, near the predicted reservoir temperature and pressure conditions, CO_2 goes through a critical phase transition in which the density changes from 200 to over 700 kg/m^3 (Span and Wagner, 1996) (Fig. 5). Thus a slightly higher temperature could result in a much lower CO_2 density. Additionally, the CO_2 will be heated during compression from the wellhead conditions (25 °C, 64 bar) and down through the injection well. Because of the high injection rates, the injected CO_2 may experience close to adiabatic conditions, putting the temperature at a maximum of 57 °C at the bottom of the injection well. This could create an ultra-low-density front or plume of CO_2 surrounded by cooler CO_2. Until recently, most of the work that has been done in reservoir simulations and in estimating the *in situ* CO_2 mass has assumed that the 37 °C measurement is correct, and that the CO_2 density is 650–700 kg/m^3. Therefore, determining the *in situ* CO_2 density is important for long-term modeling and predictions.

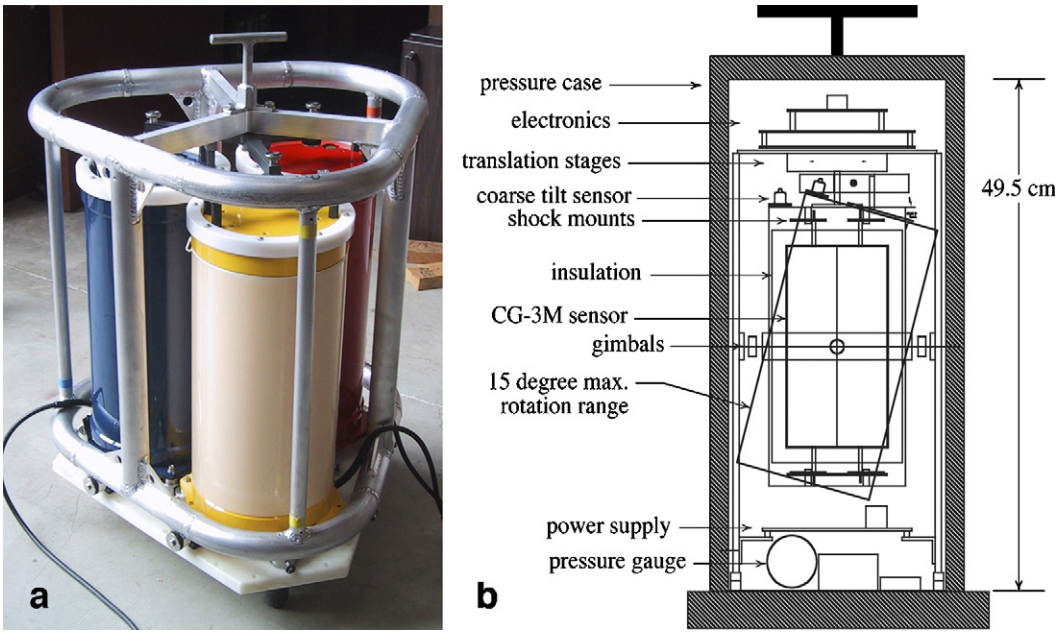

Fig. 6 – (a) Photograph of the ROVDOG instrument package used in the Sliepner gravity survey in 2002. Three instruments are affixed to one frame to increase the number of measurements at each site. (b) Diagram from Sasagawa et al. (2003) illustrating the internal schematics of the ROVDOG.

As CO_2 is injected into the Utsira sand, it displaces the water from the pore space in the sand, causing an effective bulk density decrease within the formation. One reservoir monitoring technique sensitive to changes in density is time-lapse gravity. Seafloor gravity measurements made with an ROV carried instrument have been shown to be capable of measurement accuracies of 18 μGal or less (Nooner et al., 2003, 2004; Sasagawa et al., 2003), comparable to land surveys. This instrument, the ROVDOG (Remotely Operated Vehicle deployable Deep Ocean Gravimeter), is well suited for this type of study (see Fig. 6).

2. Gravity and pressure data acquisition

The procedure for the two surveys followed the method of Eiken et al. (2004). First, the vessel transited to a benchmark location, then the ROV was launched with ROVDOG held in place by the manipulator arm and a mounting bracket (e.g. Fig. 7). The pilot guided the ROV to the benchmark, locating it acoustically and visually with cameras. Short baseline acoustic navigation of the ROV usually enabled benchmark location to within a few meters of its expected location. Upon benchmark location, the pilot maneuvered the ROV into position, placed the ROVDOG on top of the benchmark and released it from the manipulator. The ROVDOG operators then initiated an automatic leveling routine and began the gravity and pressure measurement. During the measurement, the ROV thrust downward to hold its position on the bottom 1–2 m from the benchmark. The only link between the ROV and the sensors during the measurement was a cable, which was weighted to lay on the seafloor, thereby mechanically decoupling the sensors from the ROV. Each observation lasted for about 20 min. At the end of the measurement, data logging was terminated and the ROV pilot retrieved the ROVDOG with the manipulator arm. The ROV then began an under-water transit to the next site (followed by the vessel) at a typical speed of 1–2.4 knots. When the observation sites are separated by less than 1.3 km (as they are in the Sleipner array), this mode of transiting between benchmarks is more time efficient than recovering the ROV to the vessel. It affords the added benefit that the gravity meters remain at seafloor temperature, enhancing the survey repeatability.

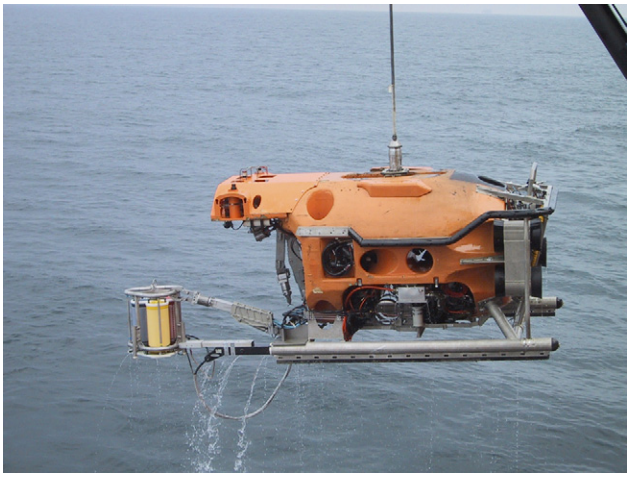

Fig. 7 – HIROV 3000 Mk II ROV deploying the ROVDOG meters in 2002.

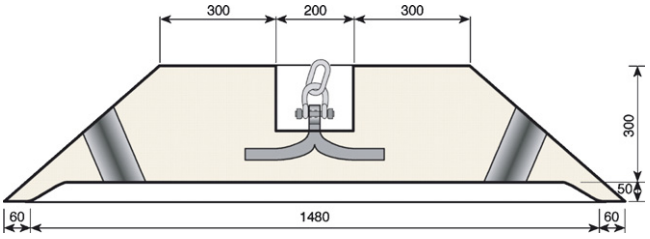

Fig. 8 – Vertical cross-section of the concrete frustum shaped benchmarks used at Sleipner to provide platforms for time-lapse gravity measurements. Dimensions are given in millimeters.

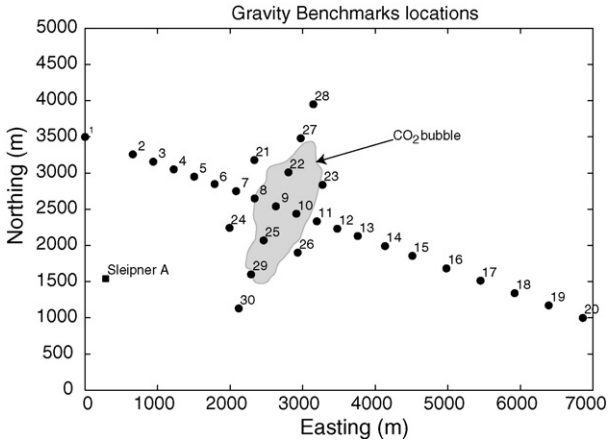

Fig. 9 – Map showing the benchmark locations in relation to the injected CO_2 bubble. The outline of the CO_2 comes from 2001 seismically imaged horizons.

Concrete seafloor benchmarks were used because they are meant to serve as stable platforms to place the instruments in exact registration on the seafloor. These benchmarks are 35 cm tall and are frustum in shape, with a lower diameter of 160 cm and an upper diameter of 80 cm (Fig. 8). This shape minimizes disturbance from trawl fishing. Each benchmark has a mass of about 650 kg. They have proven to be quite stable in a similar experiment in 300 m water depth (Stenvold et al., 2006), with a vertical stability of <1 cm over several years.

Twenty of the benchmarks were placed in a 7.3 km long WNW–ESE profile across the injection point (Fig. 9). The distance between stations increases from about 300 m near the injection point up to 500 m toward the ends. The end points are far from the injection point and are perpendicular to the maximum spreading direction observed from the 1999, 2001, and 2002 seismic surveys. This geometry was chosen to maximize the lateral gravity gradient and minimize the change in gravity over time on the endpoints, which are designed to serve as temporally stable references in our relative surveys. Another 10 locations span the orthogonal

dimension and cover the extent of the CO_2 accumulation in 2002. The benchmarks were lowered to the seafloor with a wire line and acoustic release hooked onto a small chain, which fell into the central hole after release. The deployment operation lasted 10 h for all 30 benchmarks, and was done just before surveying, on 16 August 2002.

2.1. 2002 survey

Gravity measurements were carried out 16–20 August 2002 using the ship *Edda Freya*. One hundred and fifteen measurements were made during this time, at a rate of about 30 per day. Each station was visited at least three times, to give adequate control on drift and survey accuracy. Survey loops were made with benchmark SP09 as the central location (Fig. 9). This site was visited 15 times for a loop duration of about 7 h. The sequence of stations within each loop was alternated in order to separate temporally correlated errors from spatially correlated errors. The six stations with largest scatter (based on onboard processing) received a fourth visit, and the easternmost station (SP20), which is likely to be well outside the area of CO_2 influence and hence serve as a reference location for future gravity changes, received five visits. Table 1 shows the number of visits per station.

ROVDOG Units 1, 2, and 3 were used throughout the survey. Power failure on Unit 3 caused a halt in the operation for repair after 31 measurements had been made. The instrument was brought on deck, the pressure case opened, a loose connector was repaired and it was put in the water again within 2 h. Weather was good during benchmark deployment and at the beginning of the survey but increasingly worsened. This can be observed in the noise level (RMS sample scatter) of the gravity time series (Fig. 10). Significant wave heights were about 3 m towards the end of the survey.

To aid with tide corrections, pressure was continuously recorded over the duration of the survey using portable seafloor instruments (made by Aanderaa Instruments) located at the center of the survey area (at benchmark SP09). Altogether, four reference gauges were deployed strapped together in pairs, two model WLR7 (Water Level Recorder) and two WLR8s. The two WLR7 pressure gauges show good agreement, but the WLR8 gauges deviated by about 20% from each other. Therefore, the data from the WLR8 gauges were discarded. The WLR7 gauges are rated to a depth of 340 m and the WLR8 gauges are rated to 1370 m. The two WLR7 gauges agree to within 36 Pa (standard deviation), corresponding to a depth uncertainty of 3.6 mm. A CTD (Conductivity Temperature Depth) profiler attached to the ROV measured density profiles through the water column at every launch and recovery, for a total of four measurements, all at benchmark SP09.

Table 1 – Details about each station for each year

Station	Latitude	Longitude	Number of visits		Gravity relative to SP20 (mGal)		Pressure relative to SP20 (kPa)	
			2002	2005	2002	2005	2002	2005
SP01	58.3842	1.9012	4	2	4.7644	4.7070	41.5851	43.0740
SP02	58.3821	1.9125	3	2	4.4375	4.4136	33.6682	33.7076
SP03	58.3812	1.9173	3	2	4.3141	4.2764	31.0152	32.3748
SP04	58.3803	1.9222	4	3	4.1420	4.1217	27.9551	27.8403
SP05	58.3794	1.9269	3	4	4.0087	3.9935	25.5377	25.2382
SP06	58.3786	1.9319	3	3	3.8820	3.8750	22.9969	22.7521
SP07	58.3780	1.9351	3	3	3.8002	3.7895	22.1030	22.0779
SP08	58.3768	1.9413	3	3	3.5279	3.5327	19.5980	19.3918
SP09	58.3759	1.9464	15	12	3.3174	3.3046	17.7361	18.4856
SP10	58.3750	1.9513	3	3	3.1505	3.1432	15.7740	16.0283
SP11	58.3742	1.9560	3	3	2.8254	2.8250	13.7904	13.8325
SP12	58.3732	1.9609	3	3	2.6297	2.6274	12.1591	12.4415
SP13	58.3724	1.9657	3	3	2.4402	2.4255	10.4819	11.1114
SP14	58.3712	1.9722	3	3	2.0403	2.0352	9.2097	9.6559
SP15	58.3700	1.9787	3	2	1.7160	1.7237	7.6249	7.2191
SP16	58.3686	1.9868	4	3	1.2853	1.2779	5.1286	5.8440
SP17	58.3671	1.9948	3	3	0.9477	0.9323	3.2265	4.2854
SP18	58.3656	2.0029	3	3	0.4858	0.4879	1.8225	1.8706
SP19	58.3641	2.0110	3	3	0.0589	0.0404	0.5356	1.4732
SP20	58.3627	2.0191	5	5	0.0000	0.0000	0.0000	0.0000
SP21	58.3817	1.9411	3	3	4.1823	4.1891	20.7117	20.3627
SP22	58.3802	1.9492	3	3	3.9437	3.9409	17.5053	17.6863
SP23	58.3787	1.9573	3	3	3.5904	3.6040	14.1962	13.8816
SP24	58.3732	1.9355	3	3	2.9725	2.9541	20.6325	21.1945
SP25	58.3717	1.9436	3	3	2.5839	2.5770	17.5178	18.1675
SP26	58.3702	1.9516	3	3	2.2108	2.2135	14.4556	14.4788
SP27	58.3844	1.9520	4	3	4.4022	4.4398	16.8407	14.4265
SP28	58.3887	1.9548	4	2	4.9743	4.9689	17.0307	17.5382
SP29	58.3674	1.9408	3	3	2.0809	2.0906	16.8690	16.2933
SP30	58.3632	1.9379	4	2	1.5375	1.5399	15.6765	15.5697

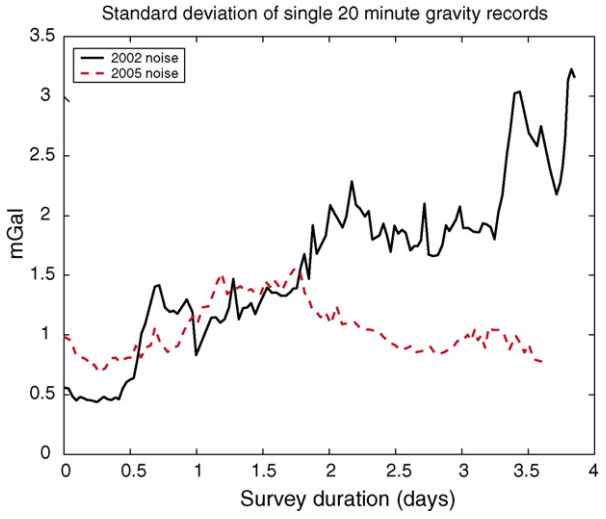

Fig. 10 – As the survey progressed, wind and wave height increased. This is reflected in the RMS scatter of 20 min gravity records, which is shown above for both surveys.

2.2. *2005 survey*

The 2005 survey was done on board the supply vessel *Normand Mjolne*. The gravity survey was carried out 2–6 September 2005. The numbers of measurements at each benchmark are given in Table 1. The reference benchmark, SP20, was measured 5 times, and the central benchmark, SP09, was measured 12 times. A total of seven recoveries of the instruments to the ship were done, for long transits. ROVDOG Units 3, 4, and 6 were used throughout the survey. Units 4 and 6 are new models based on the sensor core of the Sintrex CG5 land gravimeter; Unit 3 is the same as in the 2002 survey and is based on the sensor core of a CG3-M land gravimeter.

The sea state varied from 1 to 4 m maximum wave height, resulting in 1 s sample scatters ranging from 0.6 to 1.8 mGal (Fig. 10). Since most of this noise is periodic, the contribution to the standard errors for a 20 min average reduce to 0.001, 0.005, and 0.009 mGal for single a measurement for Units 3, 4, and 6, respectively. However, instrument drift, changes in the local water density and temperature, tides, benchmark tilt, mechanical disturbance of the gravimeter springs, and other effects reduce this precision.

Five reference tide gauges were deployed, with two at benchmark SP20 and three at benchmark SP09. At benchmark SP20, one gauge was a WLR7 and the other was a WLR8. At benchmark SP09, there were two WLR7s and one WLR8. The data from the WLR8s did not agree well ($\sigma = 0.17$ kPa), so only the WLR7 data was used. Disagreement between the WLR7 gauges was 0.04 kPa (~4 mm in depth), similar to the 2002 survey. A total of 11 CTD measurements were made throughout the survey. The mean value of these was used to determine the average water density at the survey depth.

3. Data processing

Much of the gravity and pressure data analysis was done onboard the ship during the survey for quality control purposes. After a measurement was completed on a benchmark, the raw data were processed immediately to examine data quality. After verifying the data quality, transit to the next site began. The resulting processed data were then put into a spreadsheet where drift and tide corrections were made. Instrument drift was computed based on repeat site measurements.

The steps of processing gravity recordings are as follows: (1) make tilt and temperature corrections to the data, (2) correct the data for solid earth and ocean tides, (3) edit out bad segments of data, (4) estimate instrument drift using a matrix inversion of all repeated sites, then subtract this drift, (5) take the mean of all three gravimeters with appropriate weights for each measurement, and finally (6) calculate the mean value of gravity at each benchmark from all repeats. Individual measurements were evaluated for consistency by examining differences among the meters and recovery effects (viscoelastic relaxation of the quartz spring) for each measurement. Large recoveries prompted the use of the second half of a record, rather than the entire record. A similar processing scheme was used for the pressure analysis: (1) subtract the reference pressure (tide signal) from the raw pressure, (2) estimate and subtract gauge drift, (3) calculate the mean of the three gauges for each site, (4) convert pressure to depth, and finally (5) find the mean depth for each benchmark. Once again, data quality was checked by comparing the differences among gauges for each measurement. Extreme outliers ($>5\sigma$) were not included in further analysis.

3.1. *Pressure*

The reference pressure gauges that were deployed during the survey provide a direct measurement of effects of the ocean tide, changes in air pressure, changing wind conditions, and any other time-varying environmental pressure signal that might contaminate the ROVDOG data. This reference data was subtracted from the 20 min ROVDOG pressure time series. The averages of each 20-min time series were then computed and compared for the three units. Gauge drift during the survey was calculated by fitting a straight line to all repeat measurements. Pressure was then converted to depth using a model with a constant water density of 1028 kg/m^3 (the approximate water density over the range of the survey—determined from the CTD measurements), gravity of 9.82 m/s^2, and air pressure of 101 kPa. The resulting water depths vary from 79.5 to 83.6 m. More details on the processing of the pressure data can be found in Stenvold et al. (2006).

For monitoring relative changes, depths are referenced to locations outside the area of gas injection, such as station SP20. The repeatability of the measurements gives the best indication of the uncertainty in the relative depth values. Fig. 11b shows the residuals after the mean value of a station is subtracted from each measurement at that station. The standard deviation for the 2002 survey is 0.37 cm, which we adopt as the uncertainty in the relative depth estimates. Apart from three outliers, all values repeat to within ±0.8 cm. The depth values for 2005 have an uncertainty of 0.54 cm. The small depth range of 4.1 m contributes significantly to the repeatability, as gauge precision scales with the range of operation, in part due to hysteresis effects.

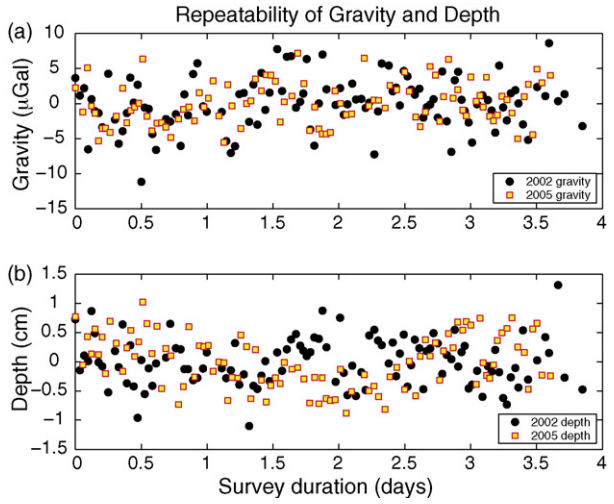

Fig. 11 – (a) Plot showing the scatter of repeated gravity measurements after the mean of each station has been subtracted from each measurement. Each point is the average of the three gravimeters. The standard deviation of the results (4.3 μGal for 2002 and 3.5 μGal for 2005) indicates the precision of the gravity values for the stations. (b) Plot showing the scatter of repeat pressure measurements after the mean for each station has been subtracted. Each point is the average of the three pressure gauges. The standard deviation of the data points is 0.37 cm for 2002 and 0.54 cm for 2005.

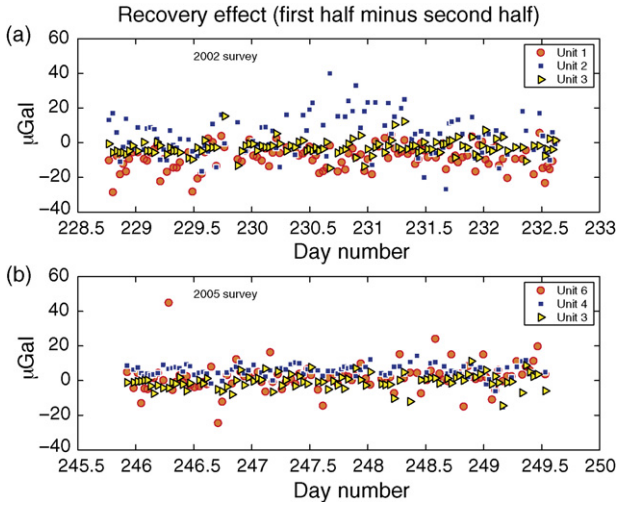

Fig. 12 – Plot showing the recovery effect for each of the three ROVDOG gravimeters. The plot shows the mean of the first half minus the mean of the second half of each 20 min gravity record for (a) 2002 and (b) 2005.

3.2. *Gravity*

For each 20 min long gravity record, noisy samples were eliminated and the time range of good data selected, prior to calculating the average. Narrow-band seafloor accelerations (mostly at 2–3 s period) originate as an interference phenomenon between ocean waves from different directions (Babcock et al., 1994; Longuet-Higgins, 1950). Noise amplitudes in this band were up to 3.5 mGal near the end of the survey, but due to the periodic nature a 20 min average effectively reduced the noise to acceptable levels.

The quality of the relative gravity data is evident in the repeatability of the measurements. Quality control was done by comparing repeated observations in three ways: (1) stability of each measurement was examined by comparing the first and second half of each 20 min gravity record. (2) Agreement among the 3 m was examined for each measurement. (3) Multiple measurements made at each benchmark were compared.

During transit from site to site, the ROVDOG package is subject to ROV motions and vibrations. When a gravity meter is tilted, its quartz spring becomes shorter. During the time the spring has been shortened, visco-elastic deformation of the spring takes place and some time is required for the spring to "recover" from this altered state when making the next measurement. This behavior, called the recovery, can last up to 10 min based on laboratory experiments. The size of this effect is indicated by comparing the mean of the first half and second half of a 20 min record (Fig. 12). The recovery phenomenon is smallest on Units 3 and 4, and largest on Units 1 and 2. Due to the recovery, the first half of a gravity record was frequently discarded.

Gravity values were corrected for solid earth tides and the ocean loading term by using the worldwide model SPOTL (Agnew, 1996). The varying gravity attraction from water tide was compensated using sea level height estimates based on the reference pressure measurements and the average water density determined from CTD measurements.

Instrument drift was estimated individually for each gravity meter by least squares fitting all repeat measurements to a third order polynomial in time. In 2002, a change in drift rate occurred for Units 2 and 3 at the time when Unit 3 was recovered to the surface to replace a faulty connector (decimal day 229.7). Therefore, separate drift polynomials are used before and after the incident (Table 2) for Units 2 and 3. This could be due to temperature fluctuations, as laboratory tests have indicated that instrument drift is sensitive to changes in external temperature.

The drift correction can be quality controlled by plotting unit differences as a function of survey time (Fig. 13). There are no apparent trends left in the plot, which suggests the drift has been removed (drift values are given in Table 2). Unit differences also provide a check if one of the gravimeters is behaving differently or erratically. For example, Unit 6 appears to be behaving erratically for the last day of the 2005 survey. This can be seen as increased scatter in the U6–U3 and U6–U4 data sets.

In 2002, the repeatability (standard deviation of repeat measurements with averages removed) of the units were 8.8 μGal for Unit 1, 9.9 μGal for Unit 2, and 4.7 μGal for Unit 3. Because of the much better performance of Unit 3, it was heavily weighted in the average calculation. After weights of 0.2, 0.1, and 1 were given to Units 1, 2, and 3, respectively, the repeatability is 4.3 μGal (Fig. 11a). The 2005 gravity values have a repeatability of 3.5 μGal (Fig. 11a), after weights of 0.8, 1, and 0.1 were given to Units 3, 4, and 6, respectively.

Table 2 – Gravimeter drift rates for Units 2 and 3 changed mid-survey in 2002 when the instruments were recovered to the surface to replace a faulty connector (at decimal day 229.7)

	Linear term (μGal/day)		Second order term (μGal/day^2)		Third order term (μGal/day^3)	Split time (day)
	First half	Second half	First half	Second half		
2002						
Unit 1	556.5		−1.7		NA	NA
Unit 2	362.8	442.4	0	−0.1	NA	229.7
Unit 3	177.9	192.7	−30.6	0.6	NA	229.8
2005						
Unit 3	417.5		5.9		−2.5	NA
Unit 4	181.0		0.3		−2.7	NA
Unit 6	328.4		1.9		−5.1	NA

4. Time-lapse results

Changes in gravity over time are found by subtracting the 2002 results from the 2005 results. After corrections for tide and drift were made, there remains a long-wavelength gravity trend increasing to the west. This trend has an apparent maximum value at benchmark SP01 (the western most station) of about 0.03–0.04 mGal. The most likely source of this signal is from water influx into the Ty formation, a natural gas reservoir that is being produced well below the Utsira formation and west of the injection point. Production from this reservoir is expected to cause an increase in local gravity due to a rise in the reservoir water as the natural gas is removed. Other possible sources for a long-wavelength gravity trend are massive sediment transport from the west to the east of the survey region or dispersed leakage of CO_2 from the Utsira. Both of these possibilities seem improbable. A forward model was calculated based on the gas reservoir geometry, porosity, temperature, gas production data, and data from monitoring wells (all proprietary information of the Sleipner production license partners).

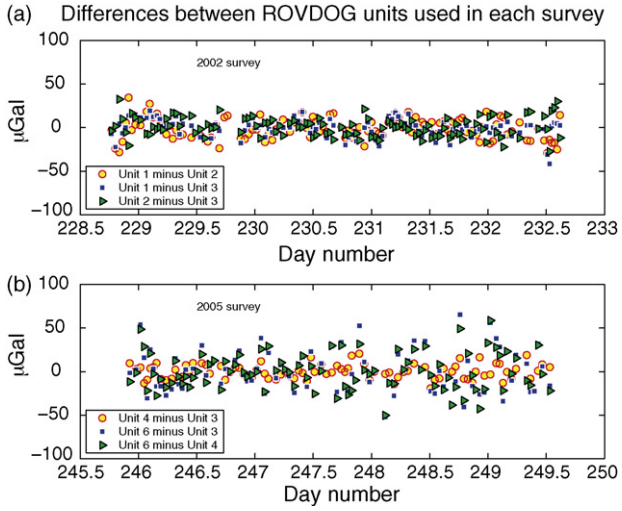

Fig. 13 – Plot showing the differences among the three ROVDOG gravimeters at each site for (a) 2002 and (b) 2005. No apparent trend shows up in the plot, indicating that the drift correction is good.

The time-lapse gravity and depth data are shown together in Fig. 14a. The depth changes have a scatter of ~7 cm, with no apparent spatial correlation. This result is surprising since no subsidence was expected for the area, and we have not observed similar behavior at other sites in the North Sea (Stenvold et al., 2006). Changes in the gravity coincide nicely with the changes in depth, providing assurance that the observed depth changes are real. However, this means that the benchmarks are not as stable at Sleipner as at other North Sea sites in deeper water (Sasagawa et al., 2003; Stenvold et al., 2006). There are several pieces of information that can be used to limit the possible sources of benchmark motion which are explored below.

Trawl fishing in the area is not uncommon. Some benchmark disturbance by trawlers is evident from the fact that benchmark SP27 was not located in the position it was in 2002. Under each benchmark lived a family of steinbit (also wolf-fish or Atlantic catfish). These fish are large (up to 1.25 m) and feed on sea urchins, mussels, cockles, and crabs. Surrounding each benchmark was a few square meter area littered with shells discarded after feeding. Upon finding the benchmark coordinates, we found a mound of shells but no benchmark. Tracks from the dragged benchmark enabled us to locate it 20 m to the northeast, with no shells in the vicinity, indicating that the benchmark had been moved recently. All other benchmarks were located in the expected place and no biological or morphological evidence suggested movement. Navigation was good to about 1 m, however, so small movements cannot be precluded.

In addition to lateral movements, benchmarks could have been tilted by trawlers. Benchmark tilt was estimated using the coarse motor position voltages for the ROVDOG leveling gimbals. Doing this for both 2002 and 2005, it is evident that some of the benchmarks have become tilted since the first survey. Benchmarks SP01, SP09, SP13, and SP20 appear to be tilted 2–3° more than they were in 2002, while benchmarks SP10, SP18, and SP25 are tilted by 1–2° more. The remaining benchmarks are within a degree of the 2002 tilt values. Benchmark SP20 also had a mound of sediment (a few centimeter thick) on one edge, which could be evidence of some disturbance by trawl fishing. Additionally, in order for at least two large fish to live under each benchmark, there must be a large cavity (maybe 10 cm deep) underlying each benchmark. A 10 cm cavity under each benchmark would cause a

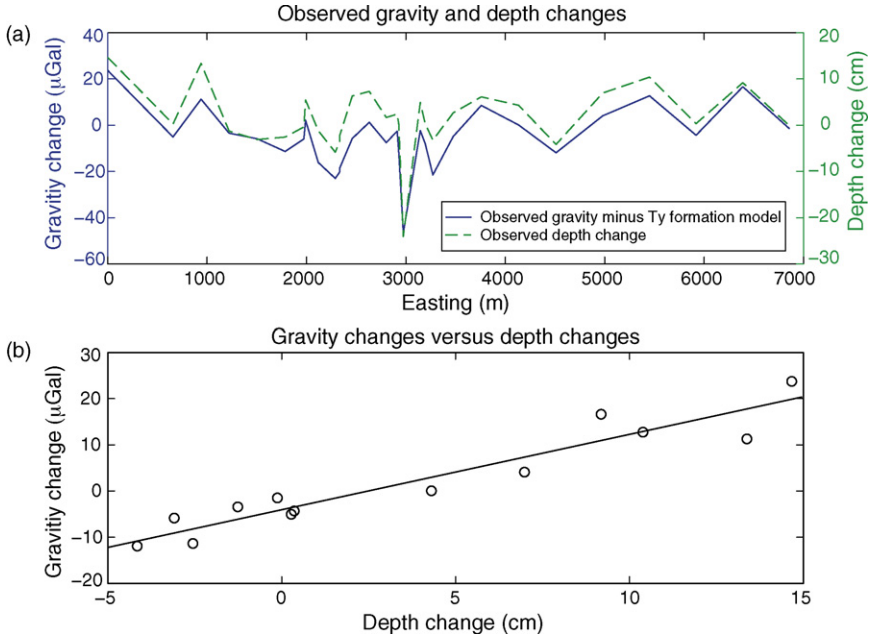

Fig. 14 – The correlation between depth changes and gravity changes is shown in two ways. (a) Variations in gravity (after subtracting a long-wavelength gravity trend due to gas production in the underlying Ty formation) correlate with changes in depth, and appear to be randomly distributed. The fact that the two are so similar provides evidence that the changes are real. (b) Gravity changes are plotted against depth changes for the outermost benchmarks. The slope of the best fitting line is the gravity gradient (0.16 mGal/m). See the text for details.

decrease in observed gravity by 0.002 mGal, but would not affect the gravity gradient. However, a benchmark could sink and/or tilt due to such a large underlying cavity. Trawling could disturb a benchmark enough to shift it by a few centimeter over the underlying cavity, thereby causing it to sink and tilt. Corrections for benchmark tilts were made to the pressure gauges based on an empirical formula (Chadwick et al., 2006; Nooner, 2005). Without this correction, the maximum uncertainty introduced by a tilt of 2° is about 9 mm for depth and 2 μGal for gravity.

Although we did not look for evidence of scouring during the survey, sediment scouring is common in marine environments (e.g. Sumer and Fredsoe, 2002; Whitehouse, 1998), and has been observed extensively in the North Sea (e.g. Heerten, 1981). At 80 m depth the orbital velocity of the water at the seafloor would be enough to move the sediment during storms of 10 m significant wave heights (Soulsby, 1998). Tidal flows might also be enough to move sediments. Biological disturbances (such as from burrowing fish) can also increase the amount of scouring in some cases (Whitehouse, 1998).

The time-lapse gravity data was inverted to simultaneously solve for a scale factor to the Ty formation forward model and the free water gravity gradient. Gravity data from only the outer benchmarks was used, to remove the influence of the injected CO_2 from the inversion. Fig. 14b shows the best fitting line to the dg versus dz data after the Ty formation model has been subtracted. The best fitting value is $dg/dz = 0.16 \pm 0.04$ mGal/m, which is significantly lower than the theoretical value of 0.22 mGal/m.

A combination of factors may account for this. First, the gradient would be reduced by 0.027 mGal/m due to the replacement of sediment with the concrete of the benchmark as the benchmark sinks into the sediments over time; however, not more than a few centimeters of benchmark settling was observed. Second, scouring of sediment surrounding each benchmark and concurrent benchmark subsidence (due to removal of underlying sediments) would cause an additional decrease in the observed gradient of 0.038 mGal/m. The resulting gradient would then be somewhere between 0.182 and 0.155 mGal/m, depending on the amount of benchmark settling. Thus the observed 0.16 mGal/m is a reasonable value.

The horizontal gravity gradient in the area has not been observed to exceed 1.3 μGal/m, meaning that the benchmarks could have moved laterally 5 m at most, assuming that the scatter in the depth gradient corrected gravity data comes from lateral motions. This is unlikely, given the position accuracy of the ROV. From the above arguments, it seems likely that the benchmark motions were primarily due to subsidence from scouring and settling, however, the benchmarks at Sleipner are not as stable as we had hoped.

Fig. 15 shows the resulting depth and Ty formation corrected time-lapse gravity values. Each point has been smoothed by averaging all observations within a 500 m radius of that point. For time-lapse changes, uncertainty is related to determining the reference zero-level, obtained by using stations outside the area of influence of the CO_2 injection. The uncertainty in time-lapse depth changes of each station is 0.9 cm, which maps into an uncertainty in gravity of 1.4 μGal. With its five visits in 2002, the southeastern most station (SP20) has an uncertainty in gravity of 1.9 μGal. SP20 also received five visits in 2005, for an uncertainty of 3.2 μGal. The total time-lapse uncertainty in the gravity measurements,

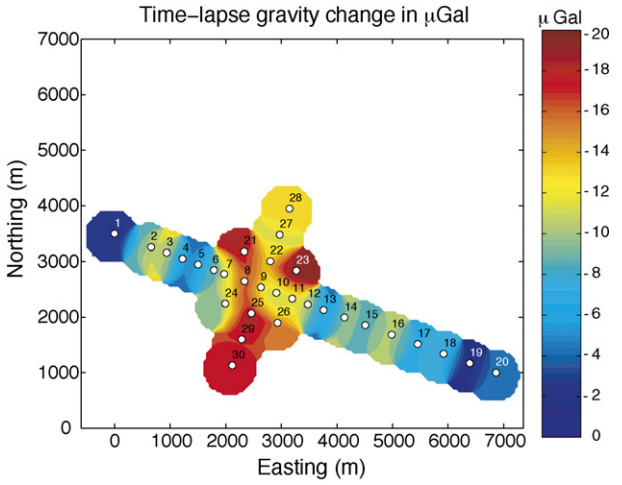

Fig. 15 – A smoothed version of the gravity residuals after correcting for depth and for the modeled gas/water contact rise in the Ty formation. Note the spatially coherent gravity decrease from 2000 to 4000 m easting.

accounting for uncertainty in the reference benchmark depth, is 5.3 μGal. A dip in the gravity with a maximum observed decrease of about 15 μGal can be seen in the data from an easting of ~2000 to 3000 m. This is the region of expected gravity decrease due to CO_2 injection. Benchmark SP03 (located at an easting of about 1000 m) also shows a dip in gravity. However, this is not spatially correlated with surrounding sites, suggesting that it is a spurious point. Benchmarks SP29 and SP30 are similarly low, suggesting spread of CO_2 to the south. This residual time-lapse gravity signal can be compared to forward models to help constrain the average CO_2 density, saturation, and flow geometry.

5. 3D forward modeling

5.1. *Modeling time-lapse gravity changes using seismically imaged CO_2*

Expanding from Chadwick et al. (2005), any viable plume saturation model must satisfy the following conditions: (1) it must reproduce observed seismic reflectivity. (2) It must produce the observed seismic velocity pushdown. (3) The volume of CO_2 in the model must match the known injected volume. (4) It must produce the observed gravity change. (5) It must produce the observed seafloor deformation. Modeling of the expected seafloor deformation has shown that a maximum uplift of 0.01 mm/year is expected (Nooner, 2005), which is far below our capability to resolve.

The seismic data from 1999 to 2001 provide the most complete coverage of the CO_2 bubble, therefore the data from these years were used to build models of injected CO_2 for two scenarios. The first is for an average CO_2 density within the reservoir of 700 kg/m^3, and the second is for an average CO_2 density of 550 kg/m^3, corresponding to low reservoir temperature (35 °C) and high reservoir temperature (45 °C) scenarios, respectively. These models contain supercritical CO_2 in two distinct parts. The first is CO_2 residing in thin, high saturation layers, which have ponded beneath nine thin inter-reservoir shale layers (Fig. 16a). These can be seen as layers of increased reflectivity in the time-lapse seismic data (Fig. 3). The second volume of CO_2 is a low saturation diffuse volume occupying the space between the high saturation layers. This diffuse volume of CO_2 does not cause increased reflectivity, but its existence is indicated by a larger observed seismic pushdown than is expected from the high saturation layers alone. The amount of diffuse CO_2 is uncertain, and depends upon the CO_2 density and upon the details of its distribution. The modeling approach is similar to Chadwick et al. (2005) with the following steps: (1) calculate a thin-layer model for the high saturation CO_2 layers. (2) Calculate the gravity from this model. (3) Calculate the velocity pushdown from this model. (4) Subtract the calculated pushdown from the observed pushdown to obtain the residual pushdown. (5) Use the residual pushdown to calculate the average vertical saturation for the diffuse volume. (6) Calculate the gravity from the diffuse CO_2 model and combine with the gravity calculated from the layers. (7) Compute the total mass and volume of CO_2 in the combined model.

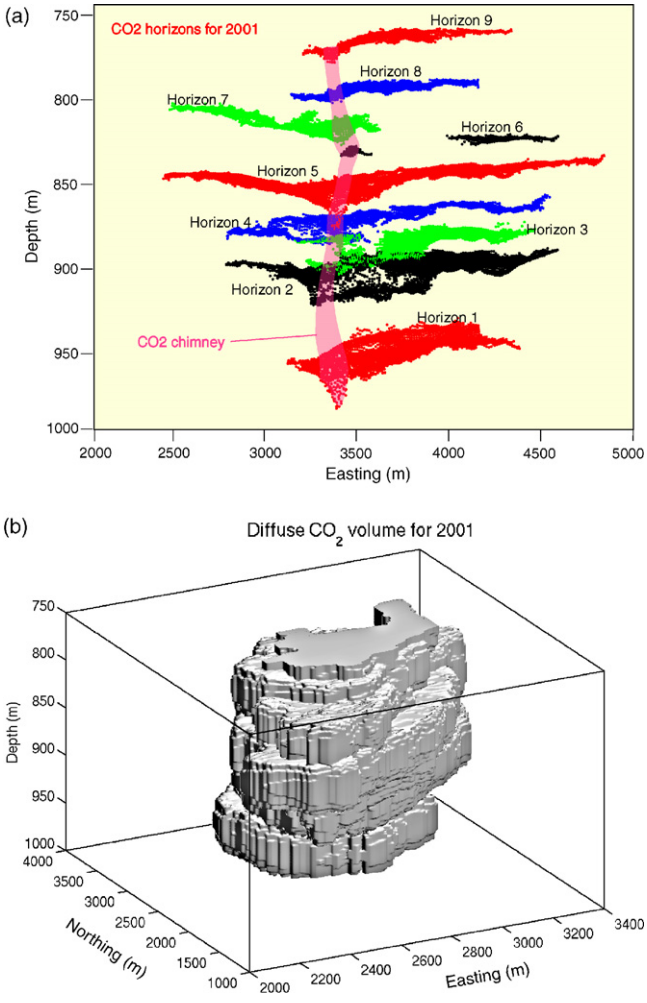

Fig. 16 – (a) A cutaway view of the seismic horizons used to construct the thin layer portion of the seismic models. (b) The volume enclosing the diffuse, low saturation CO_2 in the models based on the time-lapse 3D seismic data for 2001.

The reflection amplitude of the seismic horizons was provided as *xyz* data by the SACS (Saline Aquifer CO_2 Store) consortium. To work with the data, each horizon was first converted into a regularly spaced grid. The reflection amplitudes of the horizons were then linearly related to layer thickness with the maximum reflection amplitude being set equal to 8 m, corresponding to the tuning thickness of the seismic wavelet (Arts et al., 2002). The mass of CO_2 at each grid point is given by

$$m = \rho_{CO_2} S_{CO_2} \phi \, dx \, dy \, dz, \quad (1)$$

where dx and dy are the grid spacings, dz the layer thickness, ρ_{CO_2} the density of CO_2, ϕ the porosity, and S_{CO_2} is the saturation of CO_2. The only unknown is the saturation of CO_2, which varies with height, h, in each CO_2 layer due to capillary pressure, p_c, between the formation brine and injected CO_2. This relationship in SI units was determined by centrifuge experiments on core material from the Utsira Sand (Chadwick et al., 2004, 2005):

$$\Delta\rho g h = p_c = 810.35(1 - S_{CO_2})^{-0.948}. \quad (2)$$

In the above equation p_c is the capillary pressure in Pa, $\Delta\rho$ the difference in density between water and carbon dioxide, and g is the gravity in m/s^2. The mass of CO_2 at each grid point can then be calculated from Eq. (1) using the layer thickness at each gridpoint and the average CO_2 saturation obtained from Eq. (2).

The thin layer mass for the low-density CO_2 case is 0.853 MT in 1999 (36.3% of the injected amount) and 1.5 MT in 2001 (34.0% of the injected amount). For the high-density CO_2 case, the mass is 1.53 MT in 1999 (65.1% of the injected amount) and 2.69 MT in 2001 (61.2% of the injected amount). Uncertainty in these figures comes from uncertainty in the interpretation of the seismic horizons, errors in the simple amplitude to thickness conversion, and reflectivity attenuation in the deeper parts of the plume which are difficult to quantify (Chadwick et al., 2005). An in house 3D gravity modeling code was then used to compute the gravity signal from the thin CO_2 layers. These results are shown in Table 3.

The seismic pushdown from the layers is the difference in two-way travel time (twtt) caused by the presence of the CO_2. This can be found from the change in seismic velocity, determined using Gassmann's relationships. The velocity changes rapidly for low saturations, but for saturations larger than about 0.2 it changes very little, particularly for homogeneous distributions. The pushdown for each density scenario for both years was then calculated and subtracted from the total observed pushdown to give the residual pushdown. This residual pushdown is caused by CO_2 that is not present in the thin layers, and requires the presence of additional CO_2 within the reservoir.

The next step, then, is to use the residual pushdown to estimate the saturation and mass of the diffuse CO_2 volume using the velocity versus saturation curves shown in Fig. 4. The pushdown, ΔT, is defined as follows:

$$\Delta T = 2\left(\frac{1}{V_{S_{CO_2}}} - \frac{1}{V}\right) dz, \quad (3)$$

where $V_{S_{CO_2}}$ is the seismic velocity with CO_2 present, V the velocity without the presence of CO_2, and dz is the vertical thickness of the CO_2. To solve for $V_{S_{CO_2}}$ from Eq. (3), an estimate must be made for the vertical thickness, dz. This introduces a non-unique element into the problem. An infinite number of distributions of diffuse CO_2 can be made to satisfy the residual pushdown. To estimate dz in a meaningful way, a volume enclosing the diffuse CO_2 was defined. A reasonable assumption is that the diffuse CO_2 resides near the higher saturation volumes (the thin layers and the chimney), but is not found some characteristic distance away, which was chosen to be 25 m. This value is two times the grid spacing and is close to the average distance between shale layers, which is 30 m. Therefore, an algorithm stepped through the seismic horizons and every grid point within the characteristic distance from a high saturation grid point was included in the volume. Points that were within 1.75 m of a seismic horizon (corresponding to the average CO_2 layer thickness) were excluded from the volume. This created the gridded 3D volume enclosing the horizons shown in Fig. 16b. The vertical thickness, dz, was then calculated by summing the number of grid points in each vertical column included in the volume and multiplying each resulting number by the vertical grid spacing (12.5 m). The residual pushdown, ΔT, and the vertical thickness, dz, were then used together to solve for the average velocity, $V_{S_{CO_2}}$, through the diffuse CO_2.

Table 3 – Summary of modeling results for seismic horizon models

	Year					
	1999	1999	1999	2001	2001	2001
CO_2 density (kg/m^3)	350	550	700	350	550	700
Total injected mass (MT)	2.35	2.35	2.35	4.26	4.26	4.26
Horizon mass (MT)	0.82	1.28	1.56	1.52	2.25	2.73
Horizon g (μGal)	13.4	8.9	5.9	21.8	14.6	9.7
Diffuse mass (MT)	0.10	0.15	0.38	0.18	0.28	0.75
Diffuse g (μGal)	1.7	1.2	1.6	2.9	2.1	3.0
Total g (μGal)	15.1	10.1	7.5	24.7	16.7	12.7
Total mass (MT)	0.97	1.43	1.94	1.70	2.53	3.49
% of total mass	41.12	60.94	82.75	39.88	59.36	81.85

Only the magnitude of the maximum changes are given.

A density of CO_2 was then chosen and $V_{S_{CO_2}}$ was then used to determine an average CO_2 saturation value at each point using the appropriate velocity versus saturation curve from Fig. 4. Because the seismic velocity changes very little for saturation values greater than 0.2, small errors in the calculation of $V_{S_{CO_2}}$ can lead to large uncertainties in the resulting saturation estimate. These uncertainties are difficult to quantify. Nevertheless, the resulting model for the diffuse volume is a CO_2 distribution in which the CO_2 saturation varies laterally, but is constant vertically. This provides a good estimate of the gravity field, since the shape of the CO_2 bubble is a flat disc with a thickness to diameter ratio of 0.3 (meaning vertical variation in the CO_2 distributions does not affect the gravity much).

The mass in the diffuse volume can then be calculated from Eq. (1). For the low-density case, the diffuse CO_2 contains 0.15 MT for 1999 and 0.28 MT for 2001. Adding this to the thin layer mass gives 1.43 MT for 1999 (60.94% of the known injected mass) and 2.53 MT for 2001 (59.36% of the known injected mass). For the high-density case, the diffuse CO_2 contains 0.384 MT for 1999 and 0.752 MT for 2001. Adding this to the thin layer mass gives 1.94 MT for 1999 (82.75% of the known injected mass) and 3.49 MT for 2001 (81.85% of the known injected mass).

The contribution of the diffuse volume to the gravity signal is then calculated. Table 3 shows the contribution of each part of the model to the maximum gravity signal for each density scenario. Combining the layer contribution with the diffuse volume contribution indicates that the high-density scenario predicts a maximum change of about 2.7 μGal/year, while the low-density scenario predicts a maximum change of about 3.5 μGal/year (Fig. 17).

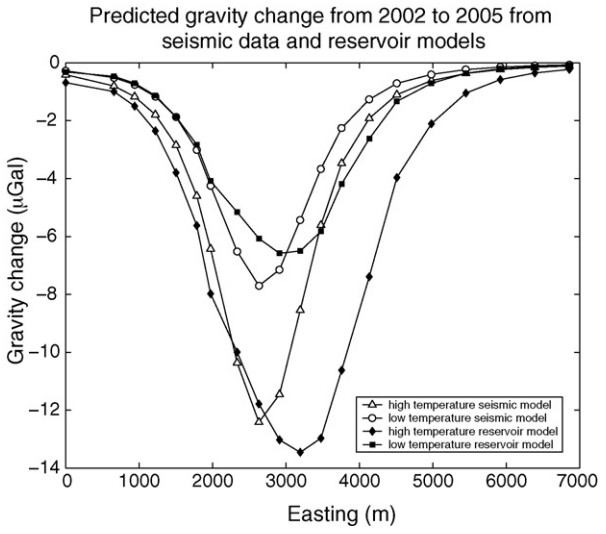

Fig. 17 – The predicted gravity change along the main NW–SE line, calculated from both the 1999 and 2001 seismic data and from the models IIa and IIb reservoir simulation models. The points along the line represent the seafloor benchmarks SP01–SP20 (Fig. 9). Calculations were made using two different densities, 550 and 700 kg/m^3, corresponding to high and low reservoir temperatures, respectively.

5.2. *Modeling time-lapse gravity changes using reservoir simulation models*

Reservoir simulation models provide insight into the physical behavior of the injected CO_2 with things such as flow geometry, dissolution into the formation brine, and CO_2 saturation. These models, however, are highly dependent on reservoir characteristics such as temperature and CO_2 density. Therefore, calculating the expected gravity change on the seafloor from reservoir flow models provides a way, independent of seismic data, to use time-lapse gravity to put constraints on the density (hence average temperature) of CO_2 within the Utsira formation. Reservoir modeling at Sleipner has been done by SINTEF, an independent research organization. SINTEF produced 3D saturation grids from CO_2 flow simulations using the commercial reservoir modeling software Eclipse. The models were for a 3D volume with a permeability of 2 darcy and a porosity of 0.38, cut laterally by five impermeable layers (representing shale). The shape of these layers was guided by, but not matched to, the geometry of the seismically imaged CO_2 horizons. The boundaries of the model volume were kept at a constant pressure, simulating an infinite reservoir (Mo, 2003). Two types of simulation models were examined. The first type, model I, has a central chimney and horizontal CO_2 layers like the seismic model; however, it has no low saturation volume (Fig. 18a). The engineers at SINTEF have not been able to produce a CO_2 flow scenario resulting in a low saturation volume as suggested by the seismic pushdown. Therefore, a second model was examined, model II, composed of several micro-chimneys, which, if small enough, might look like a diffuse volume of CO_2 to seismic energy (Fig. 18b). This was created by randomly distributing 640 holes of increased permeability within the impermeable shale layers. Simulations for model II spanning 20 years were computed by SINTEF for average reservoir temperatures of both 37 and 45 °C, corresponding to CO_2 densities of $\rho_{CO_2} = 750\,kg/m^3$ and $\rho_{CO_2} = 550\,kg/m^3$ (call them models IIa and IIb, respectively). The reservoir simulations also predict the amount of CO_2 that dissolves in the brine over time. Therefore, the mass contributing to the gravity signal will be the total injected amount of CO_2 minus the dissolved CO_2.

The output of each flow model is a volumetric CO_2 saturation grid containing 428,400 grid points ($70 \times 85 \times 72$). The thickness of the grid blocks varies from 15.2 to 0.2 m with depth, as the grid is refined below the shale layers (where most of the CO_2 resides). The horizontal grid spacing is constant at $dx = 34.4$ m and $dy = 36.1$ m.

In the current study, saturation, S_{CO_2}, was converted to a change in mass at each grid point using a reservoir porosity of $\phi = 0.37$, a shale fraction of $\nu_{sh} = 0.01$, a cell volume of $V = dx \times dy \times dz$, and a CO_2 density ρ_{CO_2} dependent on the flow model:

$$\Delta M = \Delta\rho V\phi(1 - \nu_{sh})S_{CO_2}, \quad (4)$$

where $\Delta M = M_{CO_2} - M_{H_2O}$ and $\Delta\rho = \rho_{CO_2} - \rho_{H_2O}$. Gravity was then calculated from ΔM at each grid point.

The time-varying gravity spanning 1996–2002 computed on the seafloor benchmarks from model I indicates a maximum gravity change of about 2.2 μGal/year. Model IIa spans the years

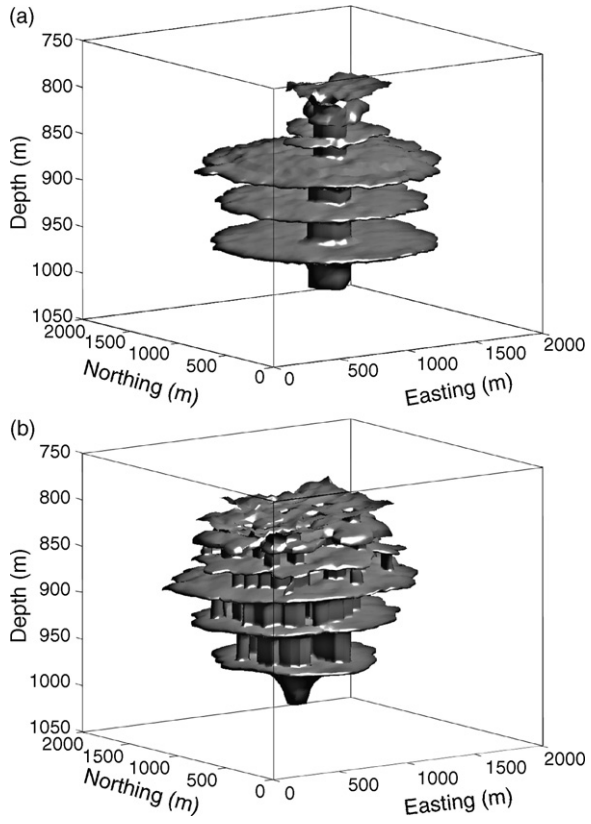

Fig. 18 – The reservoir simulation models. (a) The central chimney model, model I. (b) The random holes model, models IIa and IIb.

1996–2016 and predicts a maximum change of about 2.4 μGal/year. Interestingly, the results are almost identical, in spite of the fact that the models have different dissolution rates for CO_2 into the aquifer brine. Model I is composed of a horizontal layers central chimney, while model IIa is composed of horizontal layers and multiple vertical chimneys. This means that model IIa has a larger surface area of CO_2 in contact with the brine, allowing more dissolution to take place (~4.5% in model I and ~17.5% in model IIa). Therefore, for the same amount of injected mass, model I has more undissolved CO_2.

Model IIb predicts a maximum gravity change of 4.7 μGal/year. The higher temperature of model IIb causes the CO_2 density to decrease, creating a larger density difference between the CO_2 and formation water. The lower density CO_2 also occupies more volume within the reservoir, increasing surface area in contact with the brine. The dissolution into water in this case is more than 23%. The lower density CO_2 is also more buoyant, which tends to increase the gravity driven vertical flow. Therefore, compared to model IIa, more mass in model IIb is located in the shallow layers.

6. Discussion

6.1. Density estimate

By modeling the seismically imaged horizons in 1999 and 2001 as thin, high saturation layers and the residual seismic velocity pushdown as a low saturation non-reflective diffuse volume of CO_2, estimates for *in situ* CO_2 mass can be made. However, not all of the known injected amount of CO_2 is accounted for in these models. The high-density model, $\rho_{CO_2} = 750\,kg/m^3$, was able to account for almost 82% of the known injected mass. From the reservoir simulation models, we expect about 17% dissolution into aquifer water for the high-density case, for a total of 99% of the injected mass. However, if the CO_2 in the aquifer is in the low-density state, with $\rho_{CO_2} = 550\,kg/m^3$, the seismic model accounts for only about 60% of the injected mass, assuming a uniform distribution for the diffuse CO_2. Reservoir simulations indicate that just less than 23% dissolution of CO_2 into water would occur, thus accounting for only about 83% of the injected mass. This leaves 17% of the mass unaccounted for in both 1999 and 2001.

It is useful to revisit the low-density gravity model calculated from the seismic data in terms of a patchy CO_2 volume. It is a straightforward exercise to imagine putting the missing 17% back into the diffuse volume and redistributing the volume in such a way that the pushdown constraint is satisfied. From Table 3 it is apparent that (in view of the maximum gravity predicted) the diffuse volume of CO_2 accounts for 8.0 μGal/MT in 1999 and about 7.6 μGal/MT in 2001. Putting the missing 17% of the injected CO_2 mass back into the model each year at the rates above adds 3.2 μGal to the maximum gravity in 1999 and 5.5 μGal in 2001. This means that the expected change in the maximum gravity increases from 3.3 to 4.5 μGal/year.

Using this value means that the maximum gravity value predicted by the two types of models (seismic and reservoir simulation) have good agreement (Table 4). This indicates that the detailed geometry of the CO_2 bubble does not have a large effect on the observed peak gravity change. This insensitivity to detailed flow geometry suggests that the magnitude of the maximum time-lapse gravity signal will be due primarily to CO_2 density.

As a final exercise, an extreme case in which the reservoir temperature is warmer than 45 °C can be examined. As the temperature increases, the average density of CO_2 within the reservoir decreases rapidly. Taking the average density of CO_2 as $\rho_{CO_2} = 350\,kg/m^3$ and following the procedure outlined

Table 4 – Summary of the maximum magnitude of the gravity change expected per year for each of the different models

Model	ρ_{CO_2} (kg/m^3)	Maximum change (μGal/year)
Seismic horizon	350	~7.4
Seismic horizon	550	4.5
Seismic horizon	700	2.7
Model I	700	2.2
Model IIa	700	2.4
Model IIb	550	4.7

The values given for the seismic horizon models with $\rho_{CO_2} = 350\,kg/m^3$ and $\rho_{CO_2} = 550\,kg/m^3$ are an estimate of the result of redistributing the diffuse CO_2 to match the pushdown and injected mass (see the text).

above, the expected gravity from the seismically imaged horizons can be calculated. The pushdown can be estimated using the velocity versus saturation from the 45 °C curve shown in Fig. 4. This is not completely accurate and tends to under predict the mass of CO_2 by a few percent. However, this illustrates possible expectations for an extreme scenario. Tables 3 and 4 show the results of this calculation. The total mass estimated from the model is only 40% of the total injected mass. The contribution of the diffuse volume to gravity is about 17.8 μGal/MT in 1999 and 16.5 μGal/MT in 2001. Assuming a dissolution of 40% (a very large amount) and distributing the missing 20% of the injected mass into patchy saturation volume, the expected gravity change is to 7.4 μGal/year. For a 3-year time span, the maximum gravity change would be 22.2 μGal. With less dissolution of CO_2 the change in gravity could be even larger. This signal would be clearly discernable in the gravity data. In fact, assuming no dissolution gives a change of ~13 μGal/year applying the same logic. Comparing this to the feasibility study of Williamson et al. (2001), which predicted 15 μGal/year for $\rho_{CO_2} = 350\,kg/m^3$ and no dissolution, it is again apparent that the detailed flow geometry will have minimal effect on the estimate of CO_2 density from gravity. To date, no reservoir simulation models have been calculated for temperatures higher than 45 °C, so the amount of CO_2 dissolution is speculation.

With a time span of 3 years, the expected maximum decrease in the observed gravity is between 7 and 14 μGal, depending on the density of CO_2 within the reservoir (Table 4). Fig. 19 shows the observed gravity along with model predictions. Fig. 19b shows the direct comparison between the measured gravity change and the gravity change predicted from forward modeling. Only the benchmarks on the main line (SP01 through SP20) are shown. Although the maximum decrease in the time-lapse gravity is 15 ± 5.6 μGal, the scatter along the profile make it difficult to pick the best fitting model.

Fig. 19a shows a spatially smoothed version of the forward model predictions and the observed gravity. The smoothing was done by averaging each point with its nearest neighbor to the east. Three points from about 2200 to 3200 m easting, which have smaller error bars, include the nearby points off the northwest–southeast trending main line, so that all the time-lapse gravity information is collapsed onto a single line. This reduces point to point scatter and makes it easier to discriminate areas of spatially correlated changes. The error bars for this plot were calculated as the time-lapse uncertainty (5.3 μGal) divided by the square root of the number of points included in the average. The difference in the shape of the reservoir models and seismic model reflects the differences between the CO_2 flow geometries. The flow in the idealized reservoir simulation models is simplified and has a much larger westward component than the seismic data indicate. The smoothed observations fit the high temperature seismic model the best, suggesting that the CO_2 has an average density near 550 kg/m^3 and that the extent of the CO_2 bubble is imaged accurately by reflection seismics.

In order to characterize the density resolving power of this technique, a suite of models was constructed by linearly extrapolating the smoothed seismic forward models to other

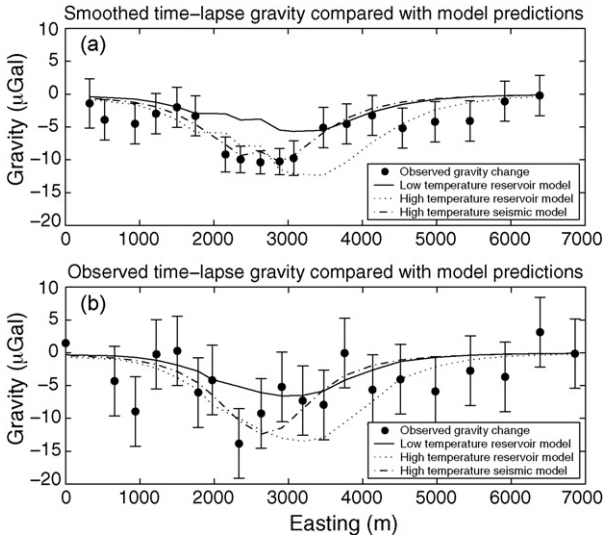

Fig. 19 – (a) Observed time-lapse gravity change plotted along with predicted gravity change for a high reservoir temperature (average CO_2 density of 550 kg/m^3) and a low reservoir temperature (average CO_2 density of 700 kg/m^3) models. Both the models and the observations have been smoothed by averaging neighboring values. (b) Time-lapse gravity with no smoothing for the 20 points on the main NE–SW trending line (benchmarks 1–20). In both plots, the observed gravity changes are most similar to the high temperature seismic model.

density values. This is justified since a simple linear scaling of the high-density model prediction matches the low-density model predictions. Model fit was determined by computing χ^2 for each density, where χ^2 is defined as

$$\chi^2 = \sum_{i=1}^{N} \frac{(\bar{g}_i - \bar{m}_i)^2}{\sigma_i^2}. \quad (5)$$

In Eq. (5), $\bar{g}_i$ is the ith smoothed time-lapse gravity value, $\bar{m}_i$ the ith smoothed model value, σ_i^2 the uncertainty in the ith smoothed time-lapse gravity value, and $N = 19$ is the number of smoothed points. Fig. 20 shows χ^2/N versus CO_2 density. The best estimate of *in situ* CO_2 density is 530 kg/m^3 with a 95% confidence interval of ±65 kg/m^3, shown by the shaded region. The 95% confidence interval indicates that if the experiment were repeated, 95% of the time the best fitting model would fall between 465 and 595 kg/m^3, based on our estimated uncertainties. This demonstrates the resolving power of this technique assuming there is uncertainty only in the gravity; however, there are unaccounted for uncertainties in the modeling, which arise from uncertainties in the seismic data, uncertainties in determining CO_2 saturation from seismic pushdown, and unknown flow geometry from 2002 to 2005. These modeling uncertainties are difficult to quantify. The difference between the seismic and reservoir simulation models gives an idea of the possible uncertainty in the modeling. The maximum difference between the two is about 1.5 μGal over 3 years. Another seismic survey is needed in order to constrain the flow

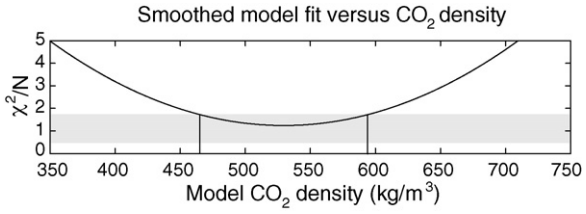

Fig. 20 – A suite of forward models was built by linearly extrapolating the gravity results from the smoothed seismic models. Model misfit (χ^2) normalized by the number of measurements (N = 19) is plotted against model CO_2 density for the smoothed seismic models. A minimum misfit occurs at a density of 530 kg/m^3 with a 95% confidence interval of ±65 kg/m^3. The 95% confidence interval is indicated by grey shading.

geometry from 2002 to 2005 before firm conclusions can be drawn. Future gravity measurements will put better constraints on the CO_2 density and continue to map out the CO_2 flow.

7. Conclusion

This study, has shown that it is possible to measure gravity on the seafloor with uncertainties of <5 μGal, even in a relatively shallow water, high noise environment. Additionally, is has been shown that by simultaneously measuring water pressure, seafloor depth can be determined to sub-centimeter accuracy, relative to a 'fixed' point on the seafloor. These depth measurements are very important for correcting the gravity measurements for anomalous changes in benchmark height, such as from sediment scouring. In the future at shallow high-current environments such as Sleipner, more care should be taken in designing and deploying benchmarks, in order to reduce the effects of scouring and biological disturbances. Techniques such as laying gravel or carpet down prior to benchmark emplacement, or anchoring the benchmarks to the seabed could be employed.

The time-lapse gravity results and modeling presented here support evidence from heatflow measurements and other temperature measurements in the vicinity of Sleipner which suggest the Utsira formation is warmer than previously thought. This is only a beginning step in characterizing the aquifer using time-lapse geophysical measurements. Additional gravity and seismic measurements are needed to further constrain this reservoir property by putting tighter bounds on the *in situ* CO_2 density. Ideally, future 3D seismic measurements and gravity measurements will be made within a few months of each other, so that the geometry of the CO_2 bubble determined from seismics will directly relate to observed changes in gravity.

Time-lapse gravimetric reservoir monitoring may play a role in future CO_2 sequestration efforts, however, this detection technique relies on the density contrast between injected CO_2 and the aquifer fluids, limiting its applicability to fluid filled reservoirs and excluding formations such as depleted coal beds. The best results will be obtained when monitoring shallow reservoirs less than 1000 m deep, where the density of CO_2 is much less than that of the reservoir fluids. In order to slow CO_2 emissions, as is needed to mitigate anthropogenic climate change, hundreds of sites such as Sleipner will be needed along with many other carbon reduction strategies. Undoubtedly, gravity will be a useful tool for monitoring injected CO_2 for a number of these sites.

Acknowledgements

Havard Alnes provided invaluable assistance with the data processing code. We would also like to acknowledge the ship and ROV crew of the *Edda Freya* and the *Normand Mjolne* for their patience and careful treatment of the ROVDOG instruments.

REFERENCES

Agnew, D.C., 1996. SPOTL: some programs for ocean-tide loading. Scripps Inst. Oceanogr. Ref. Ser. 96–98.

Arts, R., et al., 2003. Monitoring of CO_2 injected at Sleipner using time lapse seismic data. In: Paper Presented at Abstracts of the Sixth International Conference on Greenhouse Gas Control Technology (GHGT-6), Kyoto, Japan, October 1–4, 2002.

Arts, R., et al., 2002. Estimation of the mass of injected CO_2 at Sleipner using time-lapse seismic data. In: Paper Presented at EAGE, Annual Meeting, Florence, Italy.

Babcock, J.M., et al., 1994. Relationships between ocean bottom noise and the environment. Bull. Seismol. Soc. Am. 84, 1991–2007.

Chadwick, R.A., et al., 2005. 4D seismic quantification of a growing CO_2 plume at Sleipner, North Sea. In: Dore, A.G., Vining, B.A. (Eds.), Petroleum Geology: North-West Europe and Global Perspectives—Proceedings of the Sixth Petroleum Geology Conference. pp. 1385–1399.

Chadwick, R.A., et al., 2004. 4D seismic imaging of an injected CO_2 plume at the Sleipner Field, central North Sea. In: Davies, R.J., et al. (Eds.), 3D Seismic Technology. Geological Society of London, London.

Chadwick, R.A., et al., 2000. The Utsira sand, Central North Sea—an assessment of its potential for regional disposal. In: Paper Presented at the Fifth International Conference on Greenhouse Gas Control Technologies, Cairns, Australia.

Chadwick, W.W., et al., 2006. Vertical deformation monitoring at Axial Seamount since its 1998 eruption using deep-sea pressure sensors. J. Volcanol. Geotherm. Res. 150, 313–327.

Eiken, O., et al., 2004. Method for monitoring seafloor subsidence and for gravity monitoring an underground hydrocarbon reservoir. Den Norske Stats Oljeselskap A.S. The Regents of the University of California, United States, U.S. Patent # 6813564.

Han, D., Batzle, M.L., 2004. Gassmann's equation and fluid-saturation effects on seismic velocities. Geophysics 69, 398–405.

Heerten, G., 1981. Experiences of different scour protection techniques at offshore-structures in the North Sea. In: Paper Presented at Fifth Australian Conference on Coastal and Ocean Engineering.

Hermanrud, C., 1988. Determination of formation temperature from downhole measurements. Ph.D. thesis. University of South Carolina, Columbia, SC.

Holloway, S., et al., 2000. Work Area 1 (Geology). Saline Aquifer CO_2 Storage (SACS)—Final report, edited.

Johnson, J.W., Nitao, J.J., 2003. Reactive transport modelling of geologic CO_2 sequestration at Sleipner. In: Gale, J., Kaya, Y. (Eds.), Greenhouse Gas Control Technologies. Elsevier Science Ltd., Oxford, pp. 327–332.

Lindeberg, E., et al., 2000. Work Area 2 (Reservoir). Saline Aquifer CO_2 Storage (SACS)—Final report, edited.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms. Philis. Trans. Roy. Soc. Lond. A 234, 1–35.

Mavko, G., Mukerji, T., 1998. Bounds on low-frequency seismic velocities in partially saturated rocks. Geophysics 63, 918–924.

Mjøen, K., 1988. Bassengmodellering - termiske effecter av Pliocene/Pleistocene nedisinger. Diploma thesis. Norges Tekniske Høgskole, Trondheim, Norway, 80 pp.

Mo, S., 2003. Personal communication.

Nolen-Hoeksema, R.C., 2000. Modulus-porosity relations, Gassmann's equations, and the low-frequency elastic-wave response to fluids. Geophysics 65, 1355–1363.

Nooner, S., et al., 2004. A baseline seafloor gravity survey over the Sleipner CO_2 sequestration site. In: Paper Presented at Third Annual Carbon Capture and Sequestration Proceedings. ExchangeMonitor Publications, Washington, DC.

Nooner, S.L., 2005. Gravity changes associated with underground injection of CO_2 at the Sleipner storage reservoir in the North Sea, and other marine geodetic studies. Ph.D. thesis. University of California, San Diego, La Jolla, CA, 188 pp.

Nooner, S.L., et al., 2003. Seafloor micro-gravity survey of the Sleipner CO_2 sequestration site. Eos Trans. AGU 84 Fall Meet. Suppl. Abstract GC31A-01.

Rider, M.H., 1986. The Geological Interpretation of Well Logs. Blackie and Son Limited, Glasgow, 175 pp.

Sasagawa, G.S., et al., 2003. A new sea-floor gravimeter. Geophysics 68, 544–553.

Soulsby, R., 1998. Dynamics of Marine Sands. Thomas Telford Ltd., London, UK, 272 pp.

Span, R., Wagner, W., 1996. A new equation of state for carbon dioxide covering the fluid region from the triple-point to 1100 K at pressures up to 800 MPa. J. Phys. Chem. Ref. Data 25.

Stenvold, T., et al., 2006. High-precision relative depth and subsidence mapping from seafloor water-pressure measurements. SPE J. 11, 380–389.

Sumer, M., Fredsoe, J., 2002. The Mechanics of Scour in the Marine Environment. World Scientific Publishing Co. Pte. Ltd., London, UK, 550 pp.

Wang, Z., et al., 1998. Seismic monitoring of a CO_2 flood in a carbonate reservoir: a rock physics study. Geophysics 63, 1604–1617.

Whitehouse, R., 1998. Scour at Marine Structures: A Manual for Practical Applications. Thomas Telford Ltd., London, UK, 198 pp.

Williamson, J.P., et al., 2001. Work Area 5 (Geophysics)—Gravity Monitoring of the CO_2 bubble, British Geological Survey. Natural Environment Research Council.

Zweigel, P., et al., 2004. Reservoir geology of the Utsira Formation at the first industrial-scale underground CO_2 storage site (Sleipner area, North Sea). In: Baines, S.J., Worden, R.H. (Eds.), Geological Storage of Carbon Dioxide. Geological Society of London, London.