Available online at www.sciencedirect.com
ScienceDirect
Energy Procedia 114 (2017) 3607 – 3612
Energy Procedia

13th International Conference on Greenhouse Gas Control Technologies, GHGT-13, 14-18 November 2016, Lausanne, Switzerland

Using Bayes Theorem to Quantify and Reduce Uncertainties when Monitoring Varying Marine Environments for Indications of a Leak

Guttorm Alendal[a,*], Jeremy Blackford[b], Baixin Chen[c], Helge Avlesen[d], Abdirahman Omar[d]

[a] *Department of Mathematics, University of Bergen, Bergen, Norway*
[b] *Plymouth Marine Laboratory, Plymouth, UK*
[c] *Heriot-Watt University, Edinburgh, UK*
[d] *Uni Research, Bergen, Norway*

Abstract

Monitoring the marine environment for leaks from geological storage projects is a challenge due to the variability of the environment and the extent of the area that migrating CO_2 might seep through the seafloor. Due to the environmental risk associated leaks should not be allowed to continue undetected. There is also a cost issue since marine operations are expensive, so false alarms should be avoided. The main question is then: how large a deviation in the monitoring data should cause mobilization of confirmation and localization procedures? Here Baye's theorem and Bayesian decision theory is suggested as a tool for quantifying certainties and to implement costs for false positives (false alarms) and false negatives (undetected leaks) in the decision procedure. The procedure is exemplified using modeled natural CO_2 content variability and the predicted CO_2 signal from a simulated leak.

Keywords:
Monitoring design, leak detection, Bayesian decision making.

1. Introduction

Geological CO_2 storage project is by regulations such as the London Convention, OSPAR and EU directives, required to have an adequate monitoring program. With proper selection and operational procedures CO_2 geological storage projects will be designed not to leak and a number of different trapping mechanisms will keep the injected CO_2, being buoyant, inside the intended formation [1]. The injection well is believed to be the most probable leakage pathway but transport of the CO_2 within the formation might cause other pathways to the surface to become possible, or the CO_2 might create new pathways [2], possibly far away from the injection well.

Hence, even if geological monitoring of the reservoir, complex and overburden will be the primary monitoring strategy, there is a need and requirement for a surface monitoring program with the objectives

*Corresponding author.
Email address: Guttorm.Alendal@uib.no ()

Peer-review under responsibility of the organizing committee of GHGT-13.
doi:10.1016/j.egypro.2017.03.1492

to 1) maximize assurance of storage integrity, 2) assure that a leak will likely be detected, 3) continue to build an accurate baseline to capture trends and natural variability, and 4) to prevent unjustified accusations of adverse effects from the storage project [3].

For offshore storage projects such a monitoring program can be costly, and the marine environment is hostile for instrumentations. It is therefore suggested that the monitoring program has three levels of modus operandi; 1) anomaly detection modus, 2) confirmation and location modus, and 3) seep quantification modus. Some suggest a fourth step; impact assessment [4].

The focus here is the detection phase, in which the monitoring program looks for anomalies in the environment. A map of probable leak locations, preferably quantifying the internal relative probability between the different sites will govern where it will be most important to search for leaks. This can only be achieved through a thorough site characterization of the overburden.

Equally important will be an understanding of how a leak can be recognized. Probabilistic footprint predictions of a seep have to be achieved through modeling CO_2 entering the water column which may materialise in the dissolved phase, or as individual bubbles, bubble trains, or bubble plumes if the leakage flux is high enough [5]. The dynamics of these regimes are different, with the plume dynamics being the most challenging to model [5, 6, 7, 8]. Detection of bubbles can be made from sonars [9, 10], another indication of a leak might be environmental impact, caused by elevated CO_2 concentration the vicinity of the source [11] possibly as materializing as new occurrences of bacterial mats [12].

Apart from approaches relying on a thorough understanding of processes, such as the vadose zone gas monitoring approach suggested in Romanak et al. [13], a proper environmental baseline is required in order to detect changes in the environment caused by a leak from the storage complex. Such statistical baseline of important environmental parameters will include currents, natural gas seeps and biogeochemical parameters. Historical data are important in combination with new data collected during site characterization. Long time series will capture natural variability, such as seasonal changes and long-term trends. In particularly it will be important to capture the expected acidification caused by increase of CO_2 concentration [14].

Here signals of elevated CO_2 concentration away from the seep location is used to illustrate the use of Bayes theorem to decide whether a measurement of CO_2 increases our belief that a leak is in progress and quantify our certainty if the decision is that there is no leak occuring. The seep footprints are mainly governed by the varying current conditions, both spatially and temporally, such as the tidal signal or local topography [15, 16, 17].

The leak scenarios used here are discussed in previous publications [18, 19]. The scenarios were simulated in the near zone by the HWU bubble plume model [7, 8] and on a larger scale by an 800m-grid resolution North Sea setup of the three-dimensional terrain-following Bergen Ocean Model (BOM) [19].

Previously these model results have been used to find optimal locations for chemical sensors [20, 21]. It was shown that placing the sensors successively at the location of highest probability is not necessarily the best option; one sensor might detect seeps at several potential leak locations. A threshold for detection, based on a stoichiometric approach [22], was used and an excess concentration above this level immediately concluded that a leak was present.

However, given the cost of mobilizing the resources need to confirm and locate a leak, it will be preferable to have a treatment of any data stream from a monitoring program to quantify with what certainty the alarm of an ongoing leak is based. The main question remains; what level of certainty will be required in order for the monitoring program to sound the alarm? This study argues that Baye's Theorem and Bayesian decision theory offer that opportunity and it is exemplified using the same model data as in the aforementioned studies.

2. Bayesian Decision Theory

If our belief or probability that a leak is ongoing, the prior $p(L)$, the objective is to obtain an updated belief, the posterior $p(L|x)$, after taking a measurement, x. Bayes theorem reads [23]:

$$p(L|x) = \frac{p(x|L)p(L)}{p(x|L)p(L) + p(x|\neg L)(1 - p(L))} \tag{1}$$

$$p(\neg L|x) = \frac{p(x|\neg L)(1 - p(L))}{p(x|L)p(L) + p(x|\neg L)(1 - p(L))} \tag{2}$$

where the probability of measuring x if a leak is present is $p(x|L)$ and similar $p(x|\neg L)$ is the probability to measure x in the natural environment, i.e. no leak on going. It is assumed that $p(L) + p(\neg L) = 1$.

The environmental variability will have to be accounted for in the $p(x|\neg L)$ distribution and will have to be achieved through a thorough analysis of baseline statistics. The probability $p(x|L)$, i.e. likelihood measuring x in the presence of a leak will have to be based on predictions from models [19] and possibly in-situ experiments [24].

After measuring x we can either decide that there is a leak or remain assured that there is none, with an estimate on our uncertainty. Subsequent measurements update our belief. At what probability of a leak being present should the alarm be raised? The four possible outcomes when taking a decision, $(\alpha_1, \alpha_2) = (L, \neg L)$, when the true nature $(\omega_1, \omega_2) = (L, \neg L)$ are illustrated in Tab. 1. The false positive situations will be false alarms that will mobilize unnecessary resources for location of a nonexistent seep, while the false negatives results in undetected seeps that might cause environmental risks.

decision\nature	L	$\neg L$
L	true	false positive
$\neg L$	false negative	true

Table 1. The four different outcomes of taken a decision with respect to the real conditions. The two situations in which a leak is correctly detected or ruled out will both be true conclusions. False positives, i.e. deciding that leak is ongoing when it is not, or false negatives, when leaks go on undetected represent wrong decisions and should be avoided.

In general, let $\lambda(\alpha_i|\omega_j)$ be the cost, or loss, involved in deciding α_i while the true nature is ω_j. The risk of deciding α_i given the measurement x is now

$$R(\alpha_i|x) = \sum_{j=1}^{n} \lambda(\alpha_i|\omega_j)p(\omega_j|x) \tag{3}$$

and the decision rule is to select the α_i that gives lowest risk.

In the "leak-no leak" classification scheme addressed here this translates to

$$R(L|x) = \lambda_{11}p(L|x) + \lambda_{12}p(\neg L|x) \tag{4}$$

$$R(\neg L|x) = \lambda_{21}p(L|x) + \lambda_{22}p(\neg L|x) \tag{5}$$

and the least risk is chosen. This can be translated into decide that a leak is present if likelihood ration exceeds a treshold:

$$\frac{p(x|L)}{p(x|\neg L)} > \frac{\lambda_{12} - \lambda_{22}}{\lambda_{21} - \lambda_{11}} \frac{p(L)}{p(\neg L)}. \tag{6}$$

The cost parameters, λ_{ij}, allows to balance the need to detect a leak with the cost involved with false alarms. The threshold for mobilizing the confirmation and localization procedures can hence be made dependent on the cost involved.

3. An example using CO_2 concentration baseline and signal.

To illustrate the use of the theorem CO_2 concentration distributions have been created based on limited sets of model results. Time series from the Norths Sea model evaluation set up from Plymouth Marine Laboratory [25] has been used to fit a nonparametric distribution function using standard MatLab routines. These are as continuous curves in Figs. 1 and 2, the distribution based on all data, i.e. a yearly distribution, is in grey colour while the blue curves represent the respective monthly distributions.

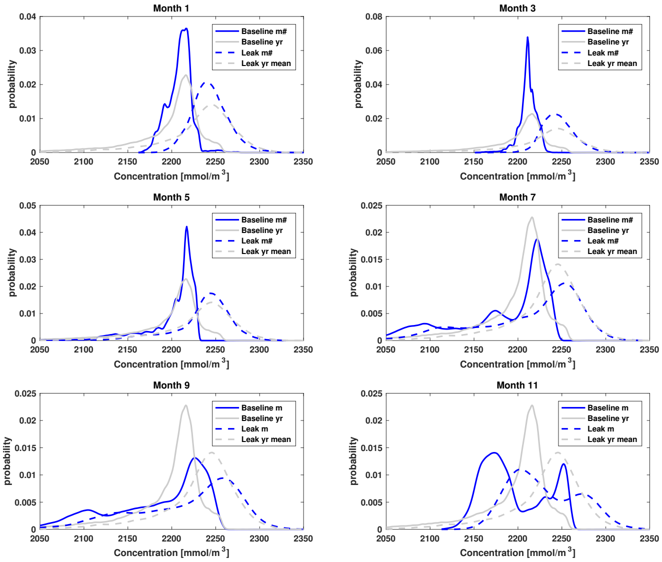

Fig. 1. The resulting distribution functions for different months (blue lines) and the annual mean distribution (grey) for in the leak grid cell. The stapled lines represent the leak situation while the continuous lines represent baseline distribution.

Notice the two maximums in concentration present for November (month 11) indicating that the environment has two modes then. Also notice the elevated tails for smaller concentrations for summer and autumn. There is too few data, from one single realization, to draw any conclusions from these distributions. But it illustrates that there will be seasonal dependency in the environmental statistics. These differences will most likely influence the ability to detect leaks.

To simulate signals from a leak, the time series from Ali et al. [19] has been used, indicated by the stapled curves, for the leak location in Figs. 1 and the adjacent grid cell just north of the location in Fig. 2.

Since these simulations represent excess CO_2 content the resulting distributions have been convolved with the respective baseline distributions causing the baseline features to recognized in the leak signal.

The shift in distribution between the baseline (continuous line) and the leak situation (stapled line) toward higher concentration is what assists in detecting a leak. As expected this shift is highest close to the source in Fig. 1 compared to some distance away Fig. 2.

To simulate streams of measurement data a series of time series has been produced by randomly pulling a starting point in the respective time series used to find the distribution functions. For each of these time series the time to detect using Eq. 6 is calculated, as presented as box plots in Fig. 3. Notice the different range along the y-axis since the right hand figure represents a measurement taken further away from the source.

Not surprisingly more measurements will be needed to detect a leak further away from the measurement location, and there are more outliers present. For both locations it seems like May (month 5) is a good month to detect a leak, the median time to detection is low and the variability is also small. It seem that the baseline distribution levels to zero for higher concentrations and that the tail for low concentrations is limited, combined with a reasonable shift when adding the leak signal. In both locations the use of annual mean results in the need for more measurements to detect the leak.

4. Discussion

The example illustrates the importance of capturing the tails of the distributions, i.e. incorporate rare events, in the baseline statistics. It is shown that the monitoring program will benefit from resolving seasonal

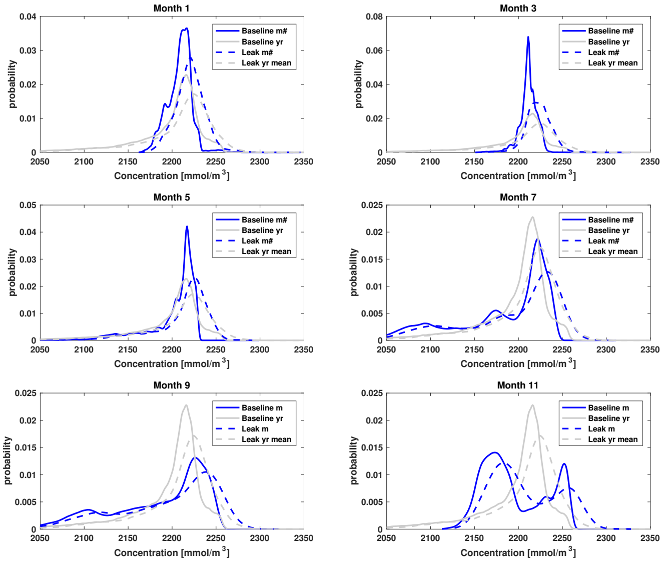

Fig. 2. The resulting distribution functions for different months (blue lines) and the annual mean distribution (grey) for the grid cell North of the leak location. The stapled lines represent the leak situation while the continuous lines represent baseline distribution.

variations, and it might identify the best time of year to supplement any fixed locations with cruises and campaigns.

However, the example used here should be used with care. The data sets used to illustrate the method are all based on model results and in reality they must be supported by in-situ measurements, especially for the baseline acquisition. In addition the ensemble of model realizations should be much higher in a realistic set up.

Acknowledgements

This work has received funding from the Research Council of Norway, trough the CLIMIT program (BayMoDe project no 254711) and the SUCCESS centre for CO_2 storage (193825/S60), and the European Union's Horizon 2020 research and innovation program under grant agreement No. 654462 (STEMM-CCS).

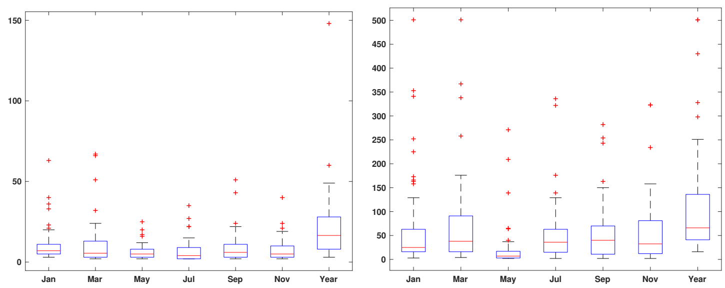

Fig. 3. Box plot for the time to discover a leak when the time series in the leak cell (left) and adjacent cell North of the leak (right) is used. The box represent first and third quartile, top and bottom are the 25th and 75th percentile of the samples, respectively. The median time to detection is shown by the red line inside the box. The whiskers extends to 1.5 the interquartile distance, measurements outside of the whiskers represent outliers and are shown as red crosses.

References

[1] J. Rutqvist, The Geomechanics of CO_2 Storage in Deep Sedimentary Formations, Geotechnical and Geological Engineering 30 (3) (2012) 525–551.

[2] C. M. Oldenburg, J. L. Lewicki, On leakage and seepage of CO_2 from geologic storage sites into surface water, Environmental Geology 50 (5) (2006) 691–705.

[3] A. D. Boyd, Y. Liu, J. C. Stephens, E. J. Wilson, M. Pollak, T. R. Peterson, E. Einsiedel, J. Meadowcroft, Controversy in technology innovation: Contrasting media and expert risk perceptions of the alleged leakage at the Weyburn Carbon dioxide storage demonstration project, International Journal of Greenhouse Gas Control 14 (2013) 259–269.

[4] J. Blackford, J. M. Bull, M. Cevatoglu, D. Connelly, C. Hauton, R. H. James, A. Lichtschlag, H. Stahl, S. Widdicombe, I. C. Wright, Marine baseline and monitoring strategies for Carbon Dioxide Capture and Storage (CCS), International Journal of Greenhouse Gas Control 38 (2015) 221–229.

[5] G. Alendal, H. Drange, Two-phase, near-field modeling of purposefully released CO_2 in the ocean, Journal of Geophysical Research 106 (C1).

[6] T. Sato, K. Sato, Numerical prediction of the dilution process and its biological impacts in CO2 ocean sequestration, Journal of Marine Science and Technology 6 (4) (2002) 169–180.

[7] M. Dewar, W. Wei, D. McNeil, B. Chen, Small-scale modelling of the physiochemical impacts of CO_2 leaked from sub-seabed reservoirs or pipelines within the North Sea and surrounding waters, Marine Pollution Bulletin.

[8] M. Dewar, N. Sellami, B. Chen, Dynamics of rising CO_2 bubble plumes in the QICS field experiment, International Journal of Greenhouse Gas ControlHttp://dx.doi.org/10.1016/j.ijggc.2014.11.003.

[9] P. G. Brewer, B. Chen, R. Warzinki, A. Baggeroer, E. T. Peltzer, R. M. Dunk, P. Walz, Three-dimensional acoustic monitoring and modeling of a deep-sea CO_2 droplet cloud, Geophysical Research Letters 33 (23) (2006) 5.

[10] R. R. P. Noble, L. Stalker, S. A. Wakelin, B. Pejcic, M. I. Leybourne, A. L. Hortle, K. Michael, Biological monitoring for carbon capture and storage – a review and potential future developments, International Journal of Greenhouse Gas Control 10 (520–535).

[11] J. C. Blackford, S. Widdicombe, D. Lowe, B. Chen, Environmental risks and performance assessment of carbon dioxide (CO_2) leakage in marine ecosystems, in: Developments and Innovation in Carbon Dioxide (CO_2) Capture and Storage Technology, Volume 2 - Carbon Dioxide (CO_2) Storage and Utilisation., Woodhead Publishing Limited, 2010, pp. 344–373.

[12] G. Wegener, M. Shovitri, K. Knittel, H. Niemann, M. Hovland, A. Boetius, Biogeochemical processes and microbial diversity of the Gullfaks and Tommeliten methane seeps (Northern North Sea), Biogeosciences Discussions 5 (1) (2008) 971–1015.

[13] K. D. Romanak, Bennett, P. C., C. Yang, S. D. Hovorka, Process-based approach to CO2 leakage detection by vadose zone gas monitoring at geologic CO_2 storage sites, Geophysical Research Letters 39 (15).

[14] K. Caldeira, M. E. Wickett, Oceanography: anthropogenic carbon and ocean pH, Nature 425 (6956) (2003) 365–365.

[15] A. M. Davies, G. K. Furnes, Observed and computed M_2 tidal currents in the North Sea, Journal Of Physical Oceanography 10 (2) (1980) 237–257.

[16] G. Alendal, J. Berntsen, E. Engum, G. K. Furnes, G. Kleiven, L. I. Eide, Influence from 'ocean weather'on near seabed currents and events at Ormen Lange, Marine and Petroleum geology 22 (1) (2005) 21–31.

[17] J. Greenwood, P. Craig, N. Hardman-Mountford, Coastal monitoring strategy for geochemical detection of fugitive CO2 seeps from the seabed, International Journal of Greenhouse Gas Control 39 (2015) 74–78.

[18] G. Alendal, M. Dewar, A. Ali, Y. Evgeniy, L. Vielstädte, H. Avlesen, B. Chen, Technical report on environmental conditions and possible leak scenarios in the North Sea., Tech. Rep. D3.4, ECO2 deliverables, http://www.eco2-project.eu (2013).

[19] A. Ali, H. G. Frøysa, H. Avlesen, G. Alendal, Simulating spatial and temporal varying CO_2 signals from sources at the seafloor to help designing risk-based monitoring programs, Journal Of Geophysical Research-Oceans 121 (1) (2016) 745–757.

[20] H. K. Hvidevold, G. Alendal, T. Johannessen, A. Ali, T. Mannseth, H. Avlesen, Layout of CCS monitoring infrastructure with highest probability of detecting a footprint of a CO_2 leak in a varying marine environment, International Journal of Greenhouse Gas Control 37 (2015) 274–279.

[21] H. K. Hvidevold, G. Alendal, T. Johannessen, A. Ali, Survey strategies to quantify and optimize detecting probability of a CO2 seep in a varying marine environment, Environmental Modelling & Software 83 (2016) 303–309.

[22] H. Botnen, A. Omar, I. Thorseth, T. Johannessen, G. Alendal, The effect of submarine CO_2 vents on seawater: implications for detection of subsea Carbon sequestration leakage, Limnology and Oceanography 60 (2).

[23] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006, iSBN 978-0-387-31073-2.

[24] J. Blackford, H. Stahl, J. Kita, T. Sato, Preface to the QICS special issue, International Journal of Greenhouse Gas Control 38 (2015) 1–1.

[25] Y. Artioli, J. C. Blackford, M. Butenschön, J. T. Holt, S. L. Wakelin, H. Thomas, A. V. Borges, J. I. Allen, The carbonate system in the North Sea: Sensitivity and model validation, Journal of Marine Systems 102-104 (2012) 1–13.