International Journal of Greenhouse Gas Control 35 (2015) 150–159

Contents lists available at ScienceDirect

International Journal of Greenhouse Gas Control

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

Bi-directional risk assessment in carbon capture and storage with Bayesian Networks

M.C. Gerstenberger[a,b,*], A. Christophersen[a,b], R. Buxton[a,b], A. Nicol[a,b]

[a] GNS Science, 1 Fairway Drive, Avalon, Lower Hutt, New Zealand
[b] Cooperative Research Centre for Greenhouse Gas Technologies, Canberra, Australia

ARTICLE INFO

Article history:
Received 17 March 2014
Received in revised form 15 January 2015
Accepted 19 January 2015
Available online 12 March 2015

Keywords:
CCS
Bayesian Networks
Risk assessment

ABSTRACT

The complex system required for carbon capture and storage (CCS) encompasses numerous sub-systems with inter-dependencies and large parameter uncertainties that propagate throughout the system. Exploring and understanding these inter-dependencies and uncertainties is invaluable for developing robust risk information. Bayesian Networks (BN), a decision support tool, are being increasingly used in the broader risk assessment community and show promise for use in CCS. BNs explore the dependencies and uncertainties within a system and have the potential to provide a better understanding of risk than more traditional tools such as logic trees or other less integrated approaches. Working with experts from within the Cooperative Research Centre for Greenhouse Gas Technologies (CO2CRC), we have developed a generic BN structure for the storage sub-system of CCS which can be used to guide the development of BNs for other CCS applications and for use in both diagnostic and predictive analysis. This bi-directionality provides one of the more important benefits of BNs; it allows for (1) traditional bottom-up risk assessment where the likely consequences based on the expected state of the system can be calculated and also (2) top-down, or outcome oriented risk, where the state of the system leading to a particular outcome, such as the likelihood of 2% leakage in 1000 years, is determined. This allows for a comprehensive sensitivity analysis which highlights important contributors to the risk and also where additional knowledge may benefit the project and reduce uncertainty. A robust expert elicitation procedure, for both the development of the network structure and the determination of event probabilities, is an integral part of the use of any such BN tool in CCS. Finally, we show the direct application of a smaller CCS BN by the CO2CRC.

1. Introduction

The science behind Carbon Capture and Storage (CCS) is maturing rapidly with an increasing number of projects in operation, or planned, where CO_2 will be injected into deep saline aquifers or depleted gas and oil reservoirs which are viewed as providing long term safe storage of CO_2. "Long term" in this case is assumed to mean thousands of years and "safe" means with minimal risk to multiple Health, Safety and Environment (HS&E) concerns including leakage of harmful concentrations to humans or the environment or contamination of natural resources such as potable groundwater or petroleum deposits. Although the physical processes involved with CO_2 storage are known, there are gaps in the detailed knowledge that is required to produce reliable simulations or predictions of the processes over the time scales that stakeholders (e.g., the public and legislators) are interested in. The effects of these gaps in knowledge on HS&E are explored via risk analysis and assessment. A detailed summary of these knowledge gaps is described by Bachu et al. (2007).

Risk is often defined in probabilistic terms as the likelihood of an event occurring combined with the negative consequences of that event occurring. Risk assessment in this context is taken to be the analysis of the risks and measures to reduce or control those risks. Risk assessment is common in many domains outside of CCS which include the nuclear and aerospace industries and earth sciences where risk due to hazard from earthquakes, volcanoes, tsunami and landslides are considered. A number of risk assessment tools have been proposed as being applicable to CCS including the use of Features, Events and Processes databases (FEPS). FEPs originated in the nuclear industry (Savage et al., 2004) and provide scenarios for risk assessments. Since 2004 Quintessa has provided an online database

* Corresponding author at: GNS Science, 1 Fairway Drive, Avalon, Lower Hutt, New Zealand.
E-mail address: m.gerstenberger@gns.cri.nz (M.C. Gerstenberger).

http://dx.doi.org/10.1016/j.ijggc.2015.01.010
1750-5836/

for CCS that has recently been extended by the European Union funded project RISCS (2014). FEPs can be used in the initial stages of a risk assessment to prioritise possible scenarios, as well as for auditing of a particular risk assessment to ensure no possible scenarios were omitted (Gerstenberger et al., 2008). RISQUE (Bowden and Rigg, 2004) is a method which has demonstrated its utility as a site selection tool and which has been applied for example at the In Salah CCS Project (Dodds et al., 2011), Weyburn (Bowden et al., 2013) and a number of Australian CCS projects. It is based around the assessment of six Key Performance Indicators (KPI) that relate to containment, effectiveness, viability of the project, wider community benefits, community safety and community amenity. Each of the KPIs are considered separately with each KPI having its own risk metric. Expert panels are used to assign probabilities to events contributing to each of the KPIs and then Monte Carlo simulations is used to calculate the risk metric for each KPI. The end result is a risk index for all KPIs that is used to rank a set of sites. Another risk assessment tool is the decision support software TESLA (Benbow et al., 2006) which is based on evidence support logic. This is done via the use of 3-value logic and an algorithm based on interval probability theory. A systematic approach to risk assessment based on logic trees, a method popular in seismic hazard and risk has also been proposed (Gerstenberger et al., 2009). A different, but systematic approach to risk assessment was used in Weyburn CCS project (Wilson and Monea, 2004) where aspects of the system were analysed using many different independent tools but giving a comprehensive assessment of the system.

Bayesian Networks (BN) (Pearl, 1985) are a common tool in other fields of risk assessment and show promise for providing a robust risk assessment in CCS (e.g., Kvien et al., 2013; Dahm et al., 2010; Yang et al., 2012). BNs are a type of flexible and probabilistic graphical modelling where the key variables in the system to be modelled are described by nodes and the relationships between nodes by arrows. With this structure the BN captures the key variables of the system and their relationships. Each variable can have different states, and the relationship between states of related variables are quantified by conditional probabilities, with a mathematical base underpinned by Bayes' Theorem (Price, 1763). In this paper we explore the applicability of BNs to CCS by introducing a BN structure for CO_2 storage in a saline aquifer that has been developed with experts from within the Cooperative Research Centre for Greenhouse Gas Technologies (CO2CRC). The BN structure provides a template for modelling a variety of different risk questions. As an example, we present a fully quantified BN that has been applied by the CO2CRC as a pilot project. We first discuss BNs in more detail in Section 1.1. A potentially useful aspect of BNs for understanding risk and optimising the risk analysis procedure comes from the bi-directionality of BNs. Bi-directionality describes the ability to propagate the effects of conditional probabilities both backward and forward in a system.

With a system as complex as CCS, expert elicitation is a critical part of the risk assessment procedure. Expert elicitation comes with its own technical difficulties and uncertainties, such as a potentially large elicitation burden on the experts; while this is not a problem unique to BNs there is a body of literature focussed on this topic with a key focus on reducing the number of probabilities to elicit (e.g., van Engelen, 1997; van der Gaag et al., 1999; Tang and McCabe, 2007; Wisse et al., 2008). What is clear is that adequate time and resources need to be allowed in CCS projects for the development of a robust and defensible risk analysis without hindering the project in an unnecessary way.

1.1. Bayesian Networks

Bayesian Networks (BN) are a type of probabilistic, graphical model, where the system being modelled is represented by a Directed Acyclic Graph (DAG) and which provides the means to encode the joint probability distribution for a set of variables representing the system (Heckerman, 1995). BNs, also called causal or probabilistic networks, are largely developed by the artificial intelligence community and they have been applied in a number of diverse problem domains including ecological modelling (Uusitalo, 2007), medical diagnosis (Wiegerinck et al., 1999), image classification (Malka and Lerner, 2004) and fraud detection (Kirkos et al., 2007). A number of publications have used BNs as a method of addressing risk assessment, such as in nuclear waste disposal (Lee and Lee, 2006), neural tube defects (Liao et al., 2010) and in seismic risk (Bayraktarli et al., 2006). Recently BNs have started to make their way into the CCS community for analysing safety risk related to loss of containment in CO_2 transport (Kvien et al., 2013), for combining evidence from multiple CO_2 leak detection technologies in geological storage (Yang et al., 2012) and for discriminating between natural, triggered and induced earthquakes in areas that have geo-engineering operations (Dahm et al., 2010).

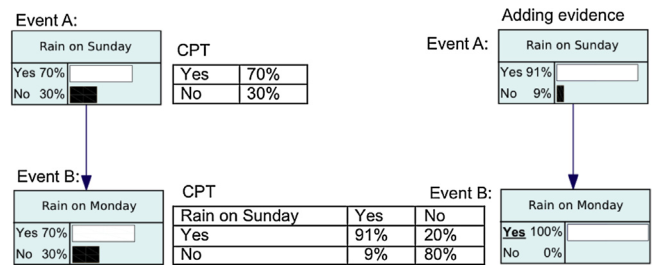

Fig. 1. An example of a simple two node BN where each node has two states. The right hand part of the graph show the effect of adding evidence, i.e., saying a particular event "Rain on Monday" occurred.

Although probabilistic models based on DAGs have been in existence since the 1920s (Wright, 1921) it was the need for a mathematically robust method capable of top-down (causal) and bottom-up (diagnostic) reasoning in the presence of evidence (Pearl and Russell, 2001) that gave researchers the impetus to develop the approach that became BNs. Pearl (1993) describes these motivations in detail and thorough introductions to BNs are given in Pearl (1991, 1985) and Neopolitan (1989). Here we give a basic introduction with a very simple example.

When considering a BN model of a system of interest, the components, or parameters of interest, of the system are represented by nodes and the conditional dependencies between the components are represented by edges (arrows). Nodes that are not connected are conditionally independent of each other, i.e., they do not directly effect one another. The direction of the arcs intuitively follows the flow of influence from parent to child; however, this does not limit the BN to uni-directional reasoning as is discussed later. Each of the nodes has two or more discrete states that fully describe the possible ranges of values of that node within the BN model. States represent discrete probability ranges and not point values. Conditional Probability Tables (CPT) attached to each node determine the causal effects of changes in the state of a parent node on a child node and propagate these effects through to all of its descendants. The CPTs express the probabilities that a node will assume a particular state given the states of all parent nodes. The exact mechanism for propagating the effect of a change of node state through a BN is based on Bayes' theorem (Price, 1763) which was named after Reverend Thomas Bayes (1702–1762).

Fig. 1 shows a basic example of a BN with only two nodes: Event A; "Rain on Sunday"; and Event B; "Rain on Monday". Each node has two states "Yes" and "No". Without any prior knowledge of what actually occurs on either of these days, the probability of having rain is the same for both days. The node "Rain on Sunday" has no parent nodes, and therefore the CPT has only two entries,