DYNAMIC PROBABILITY FAILURE USING BAYESIAN NETWORK FOR HYDROGEN INFRASTRUCTURE MODELING

by

Mohamad Faizal Bin Nurdin

Dissertation submitted in partial fulfilment of
the requirement for the
Bachelor of Engineering (Hons)
(Chemical Engineering)
JANUARY 2012

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Dynamic Probability Failure Using Bayesian Network for Hydrogen Infrastructure Modelling

by

Mohamad Faizal Bin Nurdin

A project dissertation submitted to the

Chemical Engineering Programme

University Teknologi PETRONAS

In partial fulfilment of the requirement for the

BACHELOR OF ENGINEERING (Hons)

(CHEMICAL ENGINEERING)

Approved by,

____

(Dr. Risza Binti Rusli)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

January 2012

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the reference and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

MOHAMAD FAIZAL BIN NURDIN

ABSTRACT

To produce large scale hydrogen production, it requires adequate and efficient risk control. For decades, fault tree analysis was the most widely used tool for risk assessment for industrial sector generally and hydrogen infrastructure particularly in terms of risk and consequences associated to it. The limitation to this tool is it tends to be static and do not develop over time which can give unreliable estimation of risk.

The purpose of this project is to study the suitability and efficiency of dynamic Bayesian Networks in terms of projecting the risk probability failure that develop over time for hydrogen infrastructure as the alternative of the fault tree analysis. In this study, only the risk probability failure is covered without further exploration on the consequences of the risk. The process involved by the conversion of fault tree to Bayesian Networks model by using appropriate framework. Then, the conditional probability table is assigned to each node where the numbers of CPT depend on the numbers of relationship between nodes. Finally the temporal reasoning is done to show the time-invariant between each node and the beliefs is updated to get the results.

The ways of inference use for this study are filtering and smoothing. The results show that generally, the OR gates contribute to higher risk probability compare to AND gates. Besides that, the probability for hydrogen activities increase from year to year with the assumption the accident did not happen the previous year. In addition, the instantaneous release incident is relatively low and unlikely to happen compare to the continuous release.

ACKNOWLEDGEMENT

This dissertation was completed during my final year of study at Department of Chemical Engineering, PETRONAS Institute of Technology, Malaysia, in the period of January 2012 to September 2012.

Foremost, I would like to express my deep gratitude to my supervisor, Dr. Risza Binti Rusli. She gave me the suggestions to this work and opportunity to carry out this research work under her field of research which is process safety. Her continual support and constructive criticism in numerous valuable discussions always inspired me in my technical and personal development.

Furthermore, my personal sincere thank goes out to all my colleague at the PETRONAS Institute of Technology especially to all Chemical Engineering Department students of batch July 2007 for inspiring discussions, idea, and above all, for the pleasing work environment which I really enjoyed during my stay here.

A very important part of this work goes to my families and my parents Nurdin Abdullah and Siti Sapura Azmi, who have given me on going and continuously encouragement, trust, and supports.

Last but not least is special thanks to the team of Decision Systems Laboratory of the University of Pittsburgh (http://dsl.sis.pitt.edu/) for making the Genie 2.0 modeling software freely available to the public.

Finally, this work was financially supported by my sponsorship, Majlis Amanah Rakyat (MARA). Thank you for this support.

6th August 2012

Mohamad Faizal Bin Nurdin

Table of Contents

Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS

2TBN	two time-slices Bayesian Network
BN	Bayesian Network
BNT	Bayes Net Toolbox for Matlab
CPD	conditional probability distribution
CPT	conditional probability table
DAG	directed acyclic graph
DBN	Dynamic Bayesian Network
EM	expectation maximisation
FTA	fault tree analysis
GeNIe	Graphical Network Interface
GH2	gaseous hydrogen
GMTK	graphical model toolkit
GUI	graphical user interface
HMM	Hidden Markov Model
LH2	liquid hydrogen
LI	level indicator
PCS	process control system
PCV	pressure control valve
pdf	probability distribution function
PIA	pressure indicator and alarm
PSA	pressure swing adsorption
PT	pressure transmitter
RD	rupture disk
SV	safety valve
TE	top event

1.0 INTRODUCTION

1.1 Background Studies

When the issue about the future of our energy supply is rise, particularly in connection with renewable energy sources, hydrogen is considered as a suitable energy carrier. Hydrogen proves to be the most environmentally friendly energy carrier because the only waste gas released when using it is water vapour. Unlike fossil fuels such as crude oil or natural gas, hydrogen will never run out because hydrogen is the element most commonly found in nature. Besides, the stored hydrogen can be used both to generate electricity or directly as a fuel which makes it highly suitable for stationary as well as mobile applications. However, it is important to notice that hydrogen is not an energy source by itself because it must be obtained from water or hydrocarbons by separation.

In addition, applications of hydrogen in energy sectors especially for road vehicle and household uses are a promising avenue that can lead to an increased use of hydrogen infrastructure. Hydrogen used in fuel cells or as a fuel in an internal combustion engine would result in reduced pollution. A rapid development of end-use technologies of hydrogen today will put hydrogen close as a future energy carrier and fuels. A significant increase of hydrogen use as an energy carrier is only possible if the risks of an accident in a production plant, during storage, transport, or end-use are controlled in order to avoid an increase of risk to the public as compared with well established procedures.

For a long time, hydrogen was use in the chemical, manufacturing and utility industries which are predominantly operated by experienced people. However, when dealing with a large-scale hydrogen production, it may create safety issues due to lack of knowledge since it is still in the early stage and not many trained people have the experienced on hydrogen technologies.

In order to make hydrogen available at a large-scale, an infrastructure covering the following steps must be built up: production, transportation, storage, filling station, and end-use. Furthermore, the possibility of handling incidents may occur in many places and therefore it is reasonable to determine the safety technological conditions and associated operating procedures for the realization of the hydrogen infrastructure at an early stage

Besides that, for hydrogen infrastructure to become commercial in the future, it requires adequate risk control by establishing efficient risk estimation method. Since the technologies of hydrogen infrastructure is still in the development stage, it really need judgement in terms of risk measured so that the investors will have more confidence to invest on the hydrogen infrastructure.

1.2 Problem Statement

For decades, fault tree and event tree analysis has been widely used and commonly applied modelling tools in risk failure analysis particularly on the hydrogen infrastructure. However, it is sometimes found challenging to make this traditional approach sufficiently transparent since the results obtained tend to be static and do not develop over time. Besides that, in terms of risk failure estimation, it always involved with probability and uncertainty and for fault tree, it do not have the ability needed to deal with this uncertainty. This can lead to the results being underestimate or overestimate. Therefore, this is the goal of the present work in which Dynamic Bayesian Networks (DBN) is proposed as the tool for risk estimation that is time-invariant and used to evaluate the risks failure quantitatively, to identify possible weak points, and to make suggestions for improvement.

1.3 Objectives

The main objective of this study is to develop and establish the safety modelling tool for the safe use of hydrogen as an energy carrier by using Dynamic Bayesian Networks which describe the risk failure estimate in time varying pattern.

1.4 Scopes of Study

The main scope of this study is to mapping the fault tree diagram of hydrogen infrastructure to Dynamic Bayesian Networks model by assigning appropriate temporal reasoning in order to estimate the risk over time. Other scopes of study for this project are:

a) To understand the principle of Bayesian Networks and Dynamic Bayesian Networks
b) To select the most suitable modelling software tool that can deal with Dynamic Bayesian Networks
c) To suggest maintenance action based on the risk failure estimate

2.0 LITERATURE REVIEW and/or THEORY

2.1 Fault Trees

Fault Tree Analysis (FTA) is a very popular and diffused technique for the dependability modelling and evaluation of large, safety-critical systems [1,2]. The technique is based on the identification of a particular undesired event to be analysed (e.g. system failure) called the Top Event (TE). The construction of the Fault Tree (FT) proceeds in a top/down fashion which starts from the TE to their causes until failures of basic components or events are reached. The FT works on the following assumptions:

i) Events are binary events (working/not-working);
ii) Events are statistically independent;
iii) Relationship between events and causes are represented by logical AND and OR gates.

Appendix 1 shows a typical fault tree; it consists of events corresponding to functional or hardware failures connected with edges indicating causal propagation of failures from component failures to subsystem functional failures. The leaf events of a FT, typically representing component failures are often annotated with reliability values. The inner events of a FT to which multiple lower level events are connected by edges are annotated with logic operations and are referred to as gate events [3]. The logic operations specify the conditions of the lower level failures which are connected with edges to the higher level events. The most common operation is OR gate which means that any failure at the lower level will result in a failure at the higher level. Also common is AND gate operation which needs all the failure of the lower level to occur to result the failure of higher level. The component reliability values and FT structure make it possible to evaluate system reliability.

2.2 Bayesian Networks (BN)

Bayesian Networks (BN), also known as probabilistic networks (PN) or belief networks or causal networks is a well-established tool for the representation of domains involving uncertain relations among a group of random variables. It is a graphical model that is represented as a Directed Acyclic Graph (DAG) where each node in DAG is representation of objects and events that are usually termed variables or states. Causal relations between nodes are shown by drawing an arc or edges between them. The edge will be directional, leading from the cause variable to the effect variable [4].

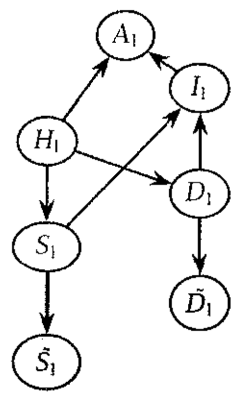


Fig. 1: Simple Bayesian Networks

Figure 1 illustrates simple BN. From the figure, the edge point from the parent to children for example from H_1 to A_1 and H_1 to D_1. Here, H_1 is the parent of A_1 and D_1 and A_1 and D_1 itself are the children of H_1. So, it can be concluded that A_1 and D_1 are conditionally independent and the state of A_1 and D_1 depend on the variable H_1.

Probability theory is based on three basic axioms:

i) $0 \leq P(X) \leq 1$;
ii) P(X) = 1 if and only if X is certain;
iii) If X and Y are mutually exclusive, then P(X U Y) = P(X) + P(Y).

And a fundamental rule of probability calculus:

$$P(X, Y) = P(X|Y)P(Y) \quad (1)$$

Where P(X,Y) is the probability of the joint event $X \cap Y$.

Based on above fundamental rule, the Bayes' rule is call out for computing posterior probability (P(X|Y)), given the prior one (P(X)) and the likelihood P(Y|X) that Y will materialize if X is true;

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \qquad (2a)$$

Here X represent hypothesis while Y represents evidence and P(Y) denotes normalizing factor which can be determined by;

$$P(Y) = P(Y|X)P(X) + P(Y|-X)P(-X) \qquad (2b)$$

Which can be computed by requiring that P(X|Y) and P(-X|Y) sum to unity. If X and Y are conditionally independent of each other like in Figure 2, the following term can be used;

$$P(X|Z,Y) = P(X|Z) \qquad (3)$$

BN can be thought of as a knowledge base [5], which explicitly represents our beliefs about the elements in the system and the relationships that exist between these various elements of the system. The purpose of such knowledge base is to infer some belief or to make conclusions about some processes or events in the system. If there is no link between two nodes, that indicates independence between them given that the value of their parents are known. There is also one rule associated with independence: If a node is observed, then the parents become dependent since the child value is explained away by the parents. Such rule is known as explaining away [6]. In addition to the network topology and for purposes of preserving the computational ability, the prior probability for each state of a root node is required.

One important property of the BN is that the graph can be considered as representing the joint probability distribution for all the variables. The chain rule is used to express this joint probability as the product of the conditional probabilities that need to be specified for each node. The chain rule is given below;

$$P(Xi, \ldots, Xn) = \prod_{i=1}^{n} P\big(Xi \big| Pa(Xi)\big) \qquad (4)$$

Where Pa(Xi) is the parent set of a node X. Taking the graph as a whole, the conditional probabilities, the structure of the BN and joint probability distribution can be used to determine the marginal probability or likelihood of each node holding one of its state. This procedure is called marginalisation.

The power of the belief calculations in BN comes to light whenever we change one of these marginal probabilities. The effects of the observation are propagated throughout the network and in every propagation step the probabilities of different neighbouring nodes are updated. According to [7], in simple networks the marginal probabilities or likelihood of each state can be calculated from the knowledge of the joint distribution using the product rule and Bayes' theorem.

Therefore, for defining the whole structure of a BN, we must specify the Conditional Probability Distribution (CPD) of each node that has parents and as mentioned before, prior probability of a root nodes need to be specified.

2.2.1 Conditional Probability Table (CPT)

The quantitative part of the BN is represented by the assignment of the conditional probability distributions to the nodes. Each node in a BN has a CPT that defines the conditional probability distributions (CPD) of the represented discrete random variables. The entries in the CPTs tell the probabilities of a hidden node given its parents.

A potential problem of using CPTs for defining the conditional probability distribution is their size. The size of the CPT of variable $X_i \in X$ depends on the number of states r_i of the variable and the number of states r_j of each parent $X_j \in Pa(X_i)$ as follows;

$$size(CPT)i = r_i. \prod_{r_j = r(X_j \in Pa(X_i))} r_j \qquad (5)$$

From the equation, the size of CPTs grows exponentially with the number of parents

2.2.2 *Inference*

In BN, one significant characteristic is that the conditional dependencies between variables can be inferred by visually inspecting the network's graph. Thus, we can divide set of BN nodes into non-overlapping subsets of conditional independent nodes. This decomposition is important when doing the probability inference. Inference is the task of computing the probability of each state of a node in a BN when other variables are known. To perform inference, one must be familiar with belief propagation. It is the action of updating the beliefs in each variable when observations are given to some of the variables. As also stated in [8], inference is the task of efficiently deducing the belief distribution over a particular subset of random variables given that we know the states of some other variables in the network.

Generally, variables in BN can be divided into groups depending on their position in BN and taking into account the meaning of real world state that they represents including their observability. Consider the partition;

$Z_i = X_N \cup Y_M$;

and let $X_N = \{x_0, x_1, \ldots, x_{N-1}\}$, $Y_M = \{y_0, y_1, \ldots, y_{M-1}\}$;

and $L = N + M$ denote two subsets as the sets of hidden and observed variables respectively. Let U_K be an arbitrary subset of Z. The goal of inference is to find the conditional probability distribution function (pdf) over U given the observed variables Y which can be written as $\Pr(U_K|Y)$;

If $U_K \subseteq Y$, we can easily find that pdf is trivially equal to;

$\Pr(U_K|Y) = \prod_{k=1}^{k} \partial(uk - yk)$ where $\delta x = 1$ for $x = 0$ and $\delta x = 0$ otherwise. A nontrivial case arises when $U_K \subseteq X$. The desired pdf can now be obtained using the famous Bayes rule;

$$\Pr(Uk|Y) = \frac{\Pr(Uk, Y)}{\Pr(Y)} = \frac{\Pr(Uk, Y)}{\sum_{Uk} \Pr(Uk, Y)} \quad (6)$$

It can be concluded that it is sufficient to find the joint pdf $\Pr(U_k|Y)$ and then marginalize over U_k . Moreover, the joint pdf over U_k and Y is obtained by marginalizing Pr(Z) over the set of hidden variables $X \backslash U_K$;

$$\Pr(Uk|Y) = \sum_{x \in X \setminus Uk} \Pr(x, Uk, Y) = \sum_{x \in X \setminus Uk} \Pr(Z) \qquad (7)$$

Inference in arbitrary BN is in general NP-hard [9]. However, there are special cases of BN topologies that allow for more efficient inference algorithm. Inference can be done by:

i) Exact probability propagation in a singly connected network. To do this, network need to be transform into singly connected structure;
ii) Approximate inference (Monte Carlo inference technique, inference by ancestral simulation, Gibbs sampling, Helmholtz machine inference, variation inference technique, etc.).

2.2.3 *Learning*

Sometimes, it often happens that the conditional probabilities are not known throughout the network. In order to overcome this problem, some learning techniques need to be employ to be able to complete the missing beliefs in the network. According to [8], the role of learning is to adjust the parameters of the BN so that the pdfs defined by the network sufficiently describes statistical behaviour of the observed data.

Let M be a parametric BN model with parameters θ of the probability distribution defined by the model. Let Pr(M) and Pr(θ|M) be the prior distributions over the set of models and the space of parameters in these model respectively. Given some observed data assumed to have been generated by the model as defined by the goal of learning in BN, the θ parameters have to be estimate such that the posterior probability of the model given data Z_i become maximized;

$$\Pr(M|Z_L) = \frac{\Pr(M)}{\Pr(Z_L)} \int_{\theta}^{\theta=n} \Pr(Z_L|\theta, M) \Pr(\theta|M)\, d\theta \qquad (8)$$

To make the equation above in more appropriate form, it is assumed that the pdf of the parameters of the model, Pr(θ|M) is highly peak around maximum likelihood (ML) estimates of those parameters. Thus, the equation is transformed into;

$$\Pr(M|Z_L) \approx \frac{\Pr(M)}{\Pr(Z_L)} \Pr(Z_L|\theta_{ML}, M) \Pr(\theta_{ML}|M) \qquad (9)$$

Where the ML estimate θ_{ML} for a given model M is obtained from next term;

$$\theta_{ML} = argmaxlogPr(Z_L|\theta) \qquad (10)$$

In addition, consider a case where not all of the variables Z in the model of a BN M are observed (represented by X). For this case, the goal of learning is describe as follows;

$$\hat{\theta} = argmaxlog \sum_X P(Y, X|\theta) \qquad (11)$$

Where P denotes specific joint pdf defined by the network and alternatively, the cost function defined can be minimizing as;

$$J(\theta) = -log \sum_X P(Y, X|\theta) \qquad (12)$$

2.3 Dynamic Bayesian Networks (DBN)

A BN is useful for problem domains where the state of the world is static. In such a world, every variable has a single and fixed value. Unfortunately, this assumption of a static world is not always sufficient. A Dynamic Bayesian Networks (DBN) which is the extension of BN with time dimension can be used to model dynamic system [10].

2.3.1 Network Structure

Firstly, dynamic extension does not mean that the network structure or parameters changes dynamically but that a dynamic system is modelled. A DBN is a directed acyclic graphical model of a stochastic process. It consists of time-slices (or time-steps) with each time-slice containing its own variables. A DBN is defined as the pair (B_1,$B_{\rightarrow}$) where B_1 is a BN that defines the prior or initial state distribution of the state variables $P(Z_1)$ [11]. Typically, $Z_1 = (U_t, X_t, Y_t)$ represents the input, hidden and output variables of the model. $B_{\rightarrow}$ is a two-slice temporal BN (2TBN) that defines the transition model $P(Z_t|Z_{t-1})$ as follows:

$$P(Z_t|Z_{t-1}) = \prod_{i=1}^{N} P(Z_t^l \middle| Pa(Z_t^l)) \qquad (13)$$

Where Z_t^l is the l-th node at time t and could be a component of X_t, Y_t, or U_t. $Pa(Z_t^l)$ are the parents of Z_t^l which can be in the same or the previous time-slice. The nodes in the first slice of the 2TBN network do not have parameters associated with them. The nodes in the second slice do have a CPT. The structure repeats and the process is stationary, so the parameters for the slices t = 2,3,... remain the same. This means that the model can be fully described by only giving the first two slices. In this way, an unbounded sequence length can be modelled using a finite number of parameters. The joint probability distribution for a sequence of length T can be obtained by unrolling the 2TBN network.

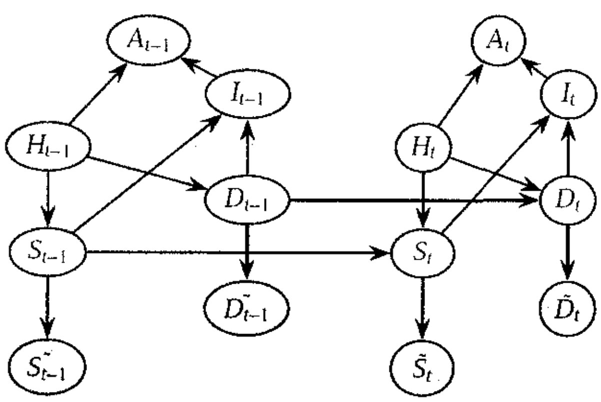


Figure 2: Extension of BN (Fig. 2) to 2TBN

$$P(Z_{1-T}) = \prod_{t=1}^{T} \prod_{i=1}^{N} P(Z_t^i \middle| Pa(Z_t^i)) \qquad (14)$$

2.3.2 *Inference*

In the DBN, only a subset of states can be observed at each time slice. So, all the unknown states in the network need to be calculated. This is done by inference. The problem of inference in DBN can be represented as the problem of finding $\Pr(X_0^{T-1}|Y_0^{T-1})$ where Y_0^{T-1} denotes a finite set of T consecutive observations, Y_0^{T-1} = $\{y_0, y_1, \ldots, y_{T-1}\}$ and X_0^{T-1} is the set of the corresponding hidden variables X_0^{T-1} = $\{x_1, x_2, \ldots, x_{T-1}\}$. Figure 3 shown graphical representation of this problem. The shaded circle indicates the states that need to be estimate (x_t) based on observations y_t

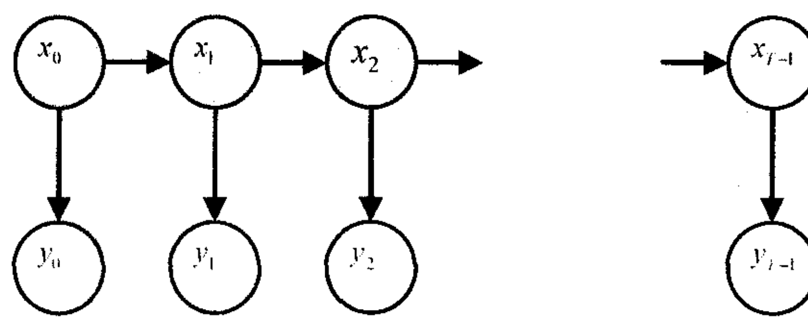


Figure 3: Inference in DBN

There exist several ways of performing inference on DBN. The most common types of inference are discussed below:

1) Forward propagation

We will used $\alpha_t(x_t)$ to denote the forward probability distribution that describes the joint probability observations gathered upon time t and at time t:

$$\alpha_t(x_t) = \Pr(Y_0^t, x_t) \qquad (15)$$

if we rely on network structure (Figure 3) , it can be concluded that

$$\alpha_{t+1}(x_{t+1}) = \Pr(y_{t+1}|x_{t+1}) \sum_{x_t} \Pr(x_{t+1}|x_t)\, \alpha_t(x_t) \qquad (16)$$

with the initial condition $\alpha_0(x_0) = \Pr(x_0)$

One of the interesting results of forward propagation is the term for likelihood of the observation data sequence Y_0^{T-1}. From the definition of the forward factor $\alpha_t(x_t)$ in equation (15), it can be determine that:

$$\Pr(Y_0^{T-1}) = \frac{\alpha_{T-1}(x_{T-1})}{\sum_{x_t} \alpha_t(x_t)} \qquad (17)$$

The probability of the observation sequence is proportional to the forward factor of the last hidden state. The probability from equation (17) can also be used to determine how well different DBN models corresponds to a data sequence in the framework of maximum likelihood estimation.

2) Backward propagation

Conditional probability of observations from time t+1 until the last observations at time T-1 conditioned on the values of the state at time t is describe as backward probability distribution:

$$\beta_t(x_t) = \Pr(Y_{t=1}^{T-1} | x_t) \qquad (18)$$

We can infer the next relation from the backward factor definition:

$$\beta_{t-1}(x_{t-1}) = \sum_{x_t} \beta_t(x_t) \Pr(x_t | x_{t-1}) \Pr(y_t | x_t) \qquad (19)$$

With $\beta_T(x_{t-1}) = 1$ as the final value.

3) Smoothing

Given the expressions for forward and backward probability propagation (factors), one can come out to the smoothing for inference and learning in the DBN. It is form from the equation (17) and (19):

$$\gamma_t x_t = \Pr(x_t | Y_0^{T-1}) = \frac{a_t(x_t)\beta_t(x_t)}{\sum_{x_t} a_t(x_t)\beta_t(x_t)} \qquad (20)$$

where $\gamma_t(x_t)$ is the smoothing operator. These terms is used for easier calculation of probability values of node's state from neighbouring nodes, and for distributing this evidence to neighbouring nodes.

4) Prediction

Another interesting inference problem deals with predicting future observations or hidden states based on the past observation data. Namely, a prediction of hidden states in the next time slice can be described as the following inference calculation task of $\Pr(x_{t+1}|Y_0^t)$ or $\Pr(y_{t+1}|Y_0^t)$. It is easy to show that:

$$\Pr(x_{t+1}|Y_0^t) = \frac{\sum_{x_t} \Pr(x_{t+1}|x_t)\, \alpha_t(x_t)}{\sum_{x_t} \alpha_t(x_t)} \qquad (21)$$

And also in the same way we can write:

$$\Pr(y_{t+1}|Y_0^t) = \frac{\sum_{x_{t+1}} \alpha_{t+1}(x_{t+1})}{\sum_{x_t} \alpha_t(x_t)} \qquad (22)$$

2.3.3 Learning

A representation of real world problems by a DBN structure often requires introduction of several nodes which conditional probabilities cannot be exactly determined. Even the expert knowledge cannot offer us solution for some conditional relationship in particular domain. In such situation, it is essential to learn this CPDs. This process is complex and is based on the Expectation Maximisation (EM) algorithm for DBN. The joint probability distribution is described below:

$$\log \Pr(X_0^T, Y_0^{T-1}|\theta) = \sum_{t=1}^{T-1} \log \Pr(x_t|x_{t-1}) + \sum_{t=0}^{T-1} \log \Pr(y_t|x_t) + \log \Pr(x_0) \qquad (23)$$

Where θ denotes the model parameter vector. The maximisation step now intends to find parameters θ that satisfy next condition:

$$\frac{\partial B(P,\hat{Q})}{\partial \theta} = \sum_{t=1}^{T-1} \langle \frac{\partial \log \Pr(x_t, x_{t-1})}{\partial \theta} \rangle + \sum_{t=0}^{T-1} \langle \frac{\partial \log \Pr(y_t x_t)}{\partial \theta} \rangle + \frac{\partial \log \Pr(x_0)}{\partial \theta} = 0 \qquad (24)$$

This learning problem can be expressed through a gradient-based learning procedure. It can be equivalently used and is implemented to perform utilization of next term $\frac{\partial B(P,\hat{Q})}{\partial \theta}$.

2.4 DBN Extension Formalism

In the original definition [11], a DBN is defined as the pair (B_1,$B_{\rightarrow}$) where B_1 is the prior or initial state distribution of the state variables $P(Z_1)$. $B_{\rightarrow}$ is a 2TBN that defines the transition model $P(Z_t|Z_{t-1})$. Only the nodes in the second time-slice of the 2TBN have a conditional probability distribution.

However, there is some limitation to this definition. One of them is that it is only possible to model first-order Markov processes. Another one is that when unrolling the network for inference, every node is copied to every time-slice even if it has a constant value for all time slices. Finally, although it is possible to introduce a different initial state using the B_1 part of the definition, it is not possible to define a different ending state, which can be useful for modeling variables that are only interesting after the end of the process.

For this project, only the second observations is used for the extension of the DBN formalism since all the variables is treat as first-order Markov processes and there are no nodes that is constant through time and had different ending state.

2.4.1 Temporal Plate and Contemporal Nodes

By using original definition by Murphy's, every node is copied to every time-slice when unrolling the DBN for inference. For nodes that have constant value at every time slice, this is a real waste of memory and computational power.

To represent the notion of nodes that are constant for every time-slice graphically, graphical model theory that was originally developed independently by both Spiegelhalter and Buntine [17] inspired the notion of a plate. In the original definition, the plate stands for N independent and identically distributed (i.i.d) replicas of the enclosed model. Similar way was decided for DBN. However, the plate was originally introduced without the notion of temporal arcs. So, it has been extended to include temporal arcs which introduce the concept of temporal plate.

By definition, temporal plate can be defines as the area of the DBN that holds the temporal information of the network. It contains the variables that develop over time (and are going to be unrolled for inference) and it has an index that denotes the sequence length T of the process. The part of the DBN that is going to be unrolled is represented within the temporal plate, so all variables with temporal arcs need to be inside the temporal plate. Variables outside the temporal plate cannot have temporal arcs. The length T of the process is denoted as an index. All not outside the temporal plates are not copied during unrolling but their arcs are. The temporal arcs are set to the appropriate nodes in the future time-slices according to their temporal order.

The temporal plate has the effect that no matter how many time-slices the DBN is unrolled to, the nodes outside the temporal plate are unique. During unrolling, the unique node itself is not copied but its outgoing arcs are copied to the corresponding nodes in all time-slices. This is useful for modeling variables in the network that remain constant over time, for instance the gender of a patient. We will call these nodes contemporal nodes.

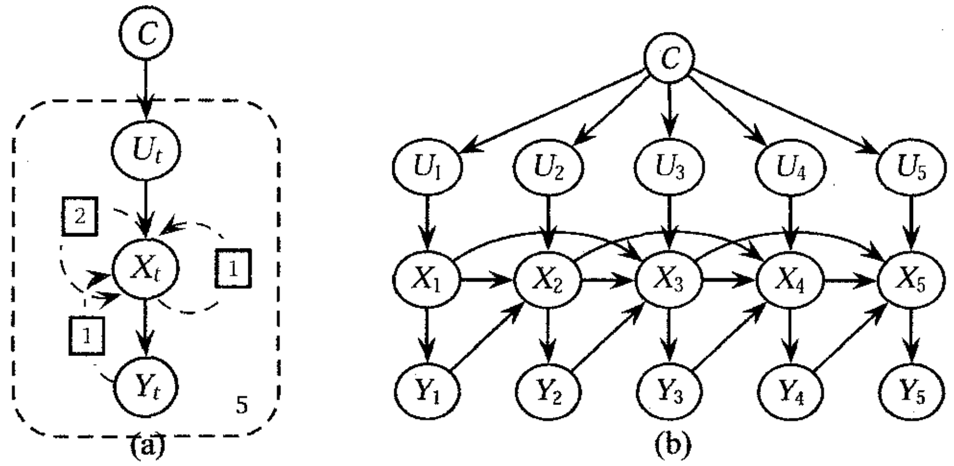


Figure 4: (a) The visualization of a second-order DBN with one contemporal variable. The temporal plate holds the part of the network that is unrolled for inference while the contemporal node is stored outside the temporal plate. The index of temporal plate denotes that the DBN is going to be unrolled for t = 5 time-slices. (b) The unrolled DBN for 5 time-slices.

Therefore, a contemporal node is a node outside the temporal plate whose value remains constant over time. It can have children inside the temporal plate but no parents.

2.5 Software

In this section, libraries and modelling tools that support temporal reasoning for DBN are discussed. The software with DBN functionality that is currently on the market can be divided into two classes: software that have a GUI and software that does not have GUI. The software with GUI is mostly commercial and is relatively easy to use for a user with only average knowledge of BN. The software without GUI is mostly academic oriented which means that it is a result of research in the area of (D)BN. Academic oriented is very flexible but this can make the application to specific problems very time consuming. In this chapter, only the software tools that are public domain are covered.

2.5.1 DBN Modelling Tools

1) GeNIe 2.0

GeNIe is the BN software toolkit developed by the Decision System Laboratory of the University of Pittsburgh. It has two modes: a modelling mode for designing the BN and a validation mode for inference. In the modelling node, it is possible to add temporal arcs to a BN to indicate that the parent node of the two nodes is in the previous time-slice. Only first-order Markov model can be designed this way. Temporal arc have different colour. The user needs to specify the initial state and the temporal probabilities of the nodes. After specification, the validation mode can be selected to follow the changes of the system over time by browsing between time steps after setting the number of time-slice. The DBN model created using this tool can be unroll graphically.

2) BayesiaLab and Netica

BayesiaLab and Netica are also the modelling tools that both support temporal reasoning and have a Graphical User Interface (GUI). BayesiaLab is developed by the French company Bayesia [12] while Netica is a product of Norsys which is located in Canada [13]. However, the use of these programs is not further explained due to them being a license domain and cannot be freely downloaded.

2.5.2 *DBN Libraries*

Three libraries exist that have DBN functionality and are freely downloadable. All three have a work-in-progress status which means that much functionality is missing and/or has to be tested and optimized.

1) Bayes Net Toolbox (BNT) for Matlab

The BNT is an attempt to build such a free, open-source, and easy-to-extend library that can be used for research purposes. Until the introduction of BNT for Matlab [14], the field lacked a free general purpose software library that able to handle many different variants of graphical models, inference and learning technique. The author chose to implement the library in Matlab because of the ease of handling Gaussian random variables. It is still widely used today but it will not be easy for someone who has little knowledge in Matlab and BN particularly to create the code. Besides, much functionality still needs to be added to make it a real general purpose tool such as Bayesian modelling, online inference and learning, prediction, more approximate inference algorithm and tree-structured CPDs. Moreover, the scripting part that is needed to do and implement a model in the BNT can be really difficult.

2) Graphical Models Toolkit (GMTK)

The GMTK is a freely-available and open-source toolkit written in C++ that is specialized in developing DBN-based automatic speech recognition (ASR) systems. The GMTK has a number of features that can be used for a large set of statistical models. These features include several inference techniques, continuous observation variables and discrete dependency specifications between discrete variables. Besides that, the DBN model needs to be specified in a special purpose language. In this language, a DBN is specified as a template that contains several time-slices.

2.5.3 *Choice of Modeling Tool Software*

After reviewing the pros and cons of the available software, GeNIe 2.0 has been selected as the modeling tool software for this project due to it user-friendly which provide GUI and also one do not need to have deep knowledge of DBN in order to use it.

3.0 METHODOLOGY / PROJECT WORKS

Figure 5 summarizes the methodology of the project works that is central to this study with the FT (with reliability data) as the existing databases supply

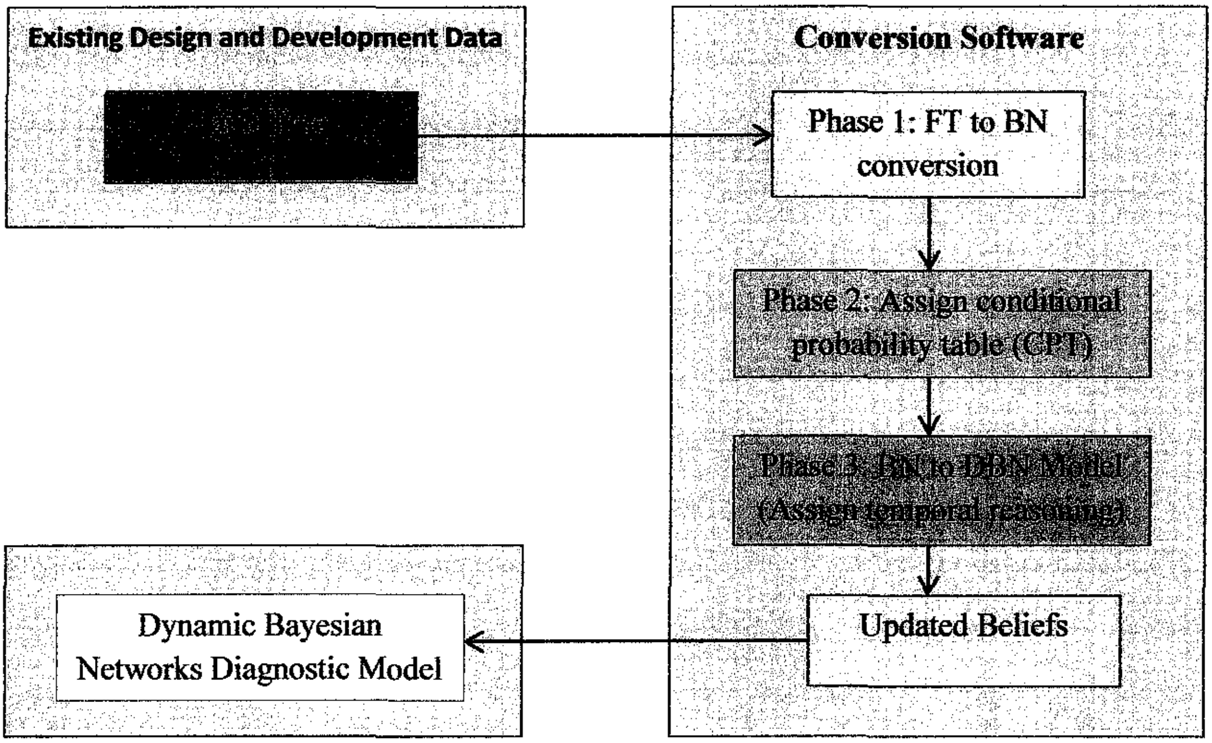


Figure 5: Block diagram of FT to DBN conversion process

Conceptually, it is straightforward to convert a FT into a BN. One only needs to re-draw the nodes and connect them while correctly enumerating reliabilities data. However, in practice this is not easy especially for large systems. The difficulty arises predominantly due to the different utilization and development focus: FTA deals in "truth, whereas BN deal in "observations of the truth". Thus, the conversion process requires a certain degree of semantic checking and adjustment. The following defines the conversion of FT to diagnostic BN in 3 phases as shown in Figure 3. First the FT is used to create the structure and parameters of the BN. The conversion process requires that the FT meet two basic assumptions which are commonly adhered to in FT creation:

i) The leaf nodes (event nodes) of the tree are statistically independent i.e. each node occurs in the tree only once and two different nodes that represent exclusive failures modes of the same component (e.g. valve failed open and valve failed close) are connected with an XOR gate node,

ii) The FT nodes are interconnected by links so that they form a directed tree. Thus, for every two nodes there is a unique path connecting them – loops are not permitted.

3.1 Phase 1: FT to BN Conversion

These steps need to be specify for the first phase of conversion using FT to create the structure and parameters of the BN:

1) Create nodes of the BN as follows:
   a) For each node in the FT, create a node in the BN,
   b) Set the BN node name and identifier using those provided in the fault tree. If there are no unique identifiers available in the FT, then create them,
   c) Assign to each BN node two states e.g. "Failed" and "OK". The former is the failure state whereas the latter represents a nominal condition. If different state names are defined, use those,
   d) Leaf nodes of the FT become ranked nodes of the BN i.e. the nodes whose failure modes are displayed in GUI in the list of failures ranked by probability,
   e) The topologies need to start from the parents (ranked nodes) to children (up events).
2) Create links of the BN as follows:
   a) Connect the nodes in the BN with links as they are connected in the FT,
   b) The direction of the connections in BN should be equivalent to the direction up the FT i.e. from leaves (events) to gates above them and from them to the gates above etc. up to the root gate. Thus the event nodes, which are leaf nodes of the FT will become the root nodes of the BN,
   c) The graphical models need to clearly shown the dependencies between each node and if the two nodes are conditionally independent, no links is connected between them.

3.2 Phase 2: Assign Conditional Probability Table (CPT)

1) Define BN parameters (probabilities) as follows:
   a) The root nodes of the BN (i.e. ranked nodes), which are the event nodes of FT, require prior probability of the failure which is assigned to the failure state ("Failed"). The prior probability of the "OK" state is computed as its complement. If the prior probability is not available, a default component failure rate with a fixed time horizon can be applied. Failure rate can also be computed from other reliability parameters, e.g. $\lambda = 1 / \text{MTTF}$, where MTTF stands for mean time to failure.,
   b) All the nodes of BN that are not root nodes require CPT. The CPT is created using zeros and one and implements the logic operation performed by the gate e.g. AND or OR where Noisy OR probability tables are used for the latter to reduce the number of node linkage parameters i.e. with *n* connections, the Noisy OR probability table is only of order n, versus order n2 for conventional probability table. More discussion of Noisy OR can be seen on [16],
   c) Pay attention to the number of parents connected to the children as an increase will require high number of CPTs. So, try to reduce as much as possible to avoid complexity of CPTs.

3.3 Phase 3: BN to DBN Model (Assign Temporal Reasoning)

1) In order to extend the BN to DBN, temporal reasoning need to be applied to the structure model:

    a) Identify for each node whether it is contemporal, temporal, or terminal conditions,

    b) For variables that change over time, temporal condition is assigned to them and for variables that do not develop over time but the failures data is bring to every time slices, contemporal condition is assigned,

    c) Define for every node whether the "Failed" state in the next time slices is dependent on the "Failed" state in the current time slice i.e. the "Failed" state at time slices t and $t+1$ are probabilistically dependent if only the order of the dependence is 1.

    d) The temporal arcs to the nodes itself suggest that the variables also change with time aside from influencing the beliefs of other nodes.

    e) If there are any observations or evidence for some variables, it need to be defined before the model is updated

3.4 Updated Beliefs

After the temporal reasoning has been assigned and all the evidences have been entered, the model beliefs are ready to be updated and unroll over the required time slices. Then, the results will show the distribution of the variables changes over time.

4.0 CASE STUDY BACKGROUND

The database of the FT is taken from [15] as the background study for this project of hydrogen infrastructure in DBN. The databases include the entire FT diagram for hydrogen activities and the reliability data. Due to the time constraint on this project, only two activities of hydrogen infrastructure are considered which are the hydrogen production and hydrogen storage.

4.1 Hydrogen Production

4.1.1 System Description

The study considers a solar hydrogen plant situated at Neunburg vorm Wald, Germany. It is a hydrogen production plant using solar energy to electrochemically decompose water in an electrolyser to obtain hydrogen and oxygen. In the electrolysis of water, the electric current is passed through an electrolyte solution of water and potassium hydroxide or alkali, decomposing the water into its constituent elements hydrogen and oxygen. Hydrogen is formed in the cathode and oxygen in the anode. A diaphragm separates the two cells to keep two gases from recombining into water. The produced hydrogen is then stored in a pressurized vessel. Energy input required to produce one cubic meter of hydrogen is about 5 KWh [18, 19]. Figure 6 shows a layout of the overall facility from the aerial photograph.

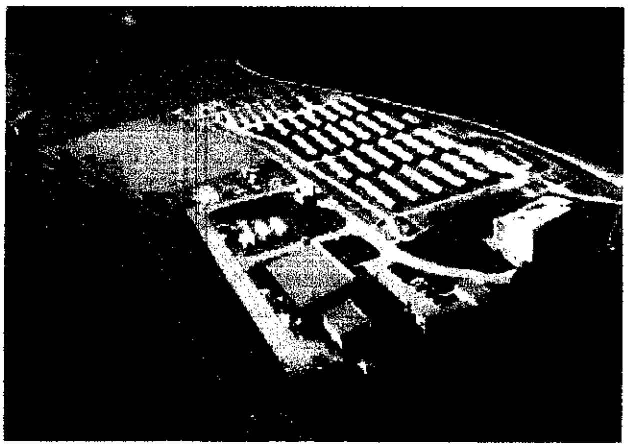


Figure 6: The solar hydrogen plant in Neunburg vorm Wald, Germany

4.1.2 *GH2 Storage*

The high pressure hydrogen tank of the plant stores the largest hydrocarbon inventory of 5000 Nm^3 compared to other components. It may be the largest contributor to societal risk as assessed in this study. Therefore, the study is focused on the two horizontal cylindrical high-pressure hydrogen storage (Figure 7) with an operating pressure of 3 MPa at ambient temperature. The tank is filled directly from the water electrolysis in the plant generated from the two low-pressure electrolysers requiring subsequent compression of the product gases.

Figure 7: GH2 storage at the solar-hydrogen plant [20]

Appendix 2 shows a simplified piping and instrumentation diagram of the high-pressure storage. The tank is filled from electrolysers continuously during the day (e.g. 200/year) through filling valve, V7 and V13. The filling process is stopped when the set point at the pressure control valve, PCV-19 is reached. Pressure indicator and alarms (PIA) are installed to measure and indicate pressure levels of the tank and its piping system. The tank PIAs are equipped with pressure switch or transmitter for remote controllers (e.g. alarm). If the operator fails to observe PIAs or to respond to the alarm the tank pressure increases rapidly and the tank is overfilled. To protect against overpressure, each tank is equipped with two pressure safety valves (SVs) and a rupture disk (RD). One of the SVs is operated exchangeable at the relative pressure of 3.3 MPa. The SVs will automatically re-close if the tank pressure returns to the operating pressure. The rupture disks (RD-1 and RD-2) are provided in case the safety valves should fail.

The ultimate overpressure protection of the tank is provided by stopping the filling line automatically. It is performed by a safety shut-off (PCV-20) which is actuated by PSH signal. The gas is withdrawn from the tank through withdrawal valve V12. The required output pressure is determined by setting pressure at the pressure control (i.e. high pressure = PCV-16, low-pressure = PCV-17).

Table 1: Most important capacities and dimensions of the GH2 storage

H2 Storage/Lines	Dimension	Capacity
1. High-pressure vessel	L = 9.8 m, D = 2.8 m,	2 x 2500 Nm3 (*)
	Vuseful = 50 m^3	30 bar (400 kg)
2. Input line	50.8 mm (2 in)	30 Nm^3/h
3. Output line	50.8 mm (2 in)	30 Nm^3/h

Source: Messer Griesheim GmbH, Linde AG; (*) m^3 H_2 at 15°C, and 1 bar (NTP)

4.2 Hydrogen Storage

The study focuses on liquid hydrogen (LH2) storage situated in the Vonburg-Ingolstadt-Refinery (RVI) as a representative example. The LH2 storage is used to store liquid hydrogen produced from the hydrogen liquefaction plant. The LH2 is delivered to the consumers through an LH2 tanker truck.

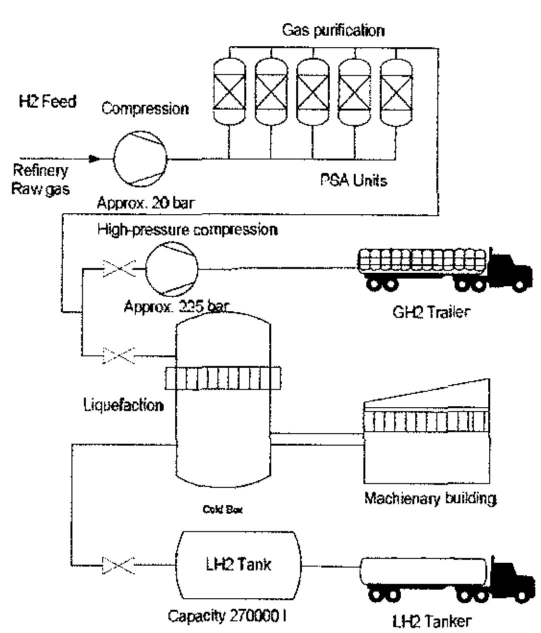


Figure 8: Process flow diagram of the liquefaction plant [21]

4.2.1 *System Description*

The Linde liquefaction plant (Figure 8) mainly consists of compressor units, pressure swing adsorption (PSA) purification, liquefier, and LH2 tank. The hydrogen rich raw gas is supplied from the RVI refinery and has pressures varying between 0.9-1.4 MPa. The gas is then compressed to about 2.1 MPa and is cleaned in PSA purification units. The gas is further purified by low temperature and liquefied into para-hydrogen [21].

The liquefaction process which is designed for a flow rate of 180 kg/h is based on the Claude cycle. The necessary refrigeration is provided at three temperatures levels using: LN2 (from 300K to 80K), expansion turbine (80K to 30K), and Joule Thomson (JT) valve (30K to 20K). The cooling down process from ambient to LN2 temperature levels is operated manually for about 5 hours. Once the LN2 temperature is reached the operating mode of 50% or 100% LH2 can be selected from the monitor screen and the process control system starts automatically. The steady state liquefaction is achieved after a further 3 hours. Opening the JT valve and hence liquefaction capacity is controlled by the outlet temperature of the third turbine.

The liquid hydrogen is then stored in a horizontal vacuum insulated tank at -253°C having a capacity of 270 000 litres. The tank can store hydrogen for several weeks without significant vaporization [21]. The whole plant is operated and controlled by a central process control system (PCS).

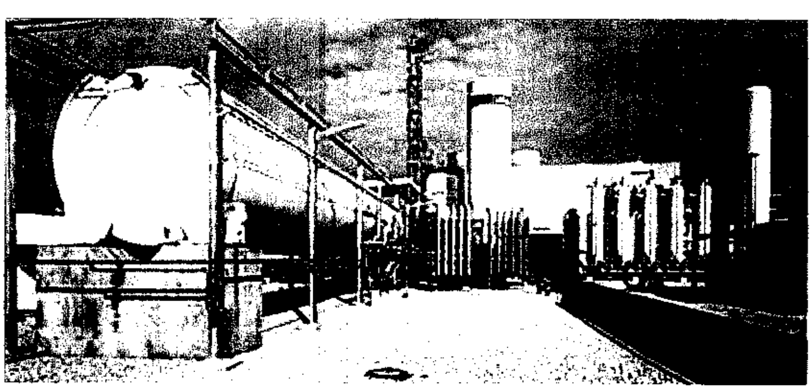


Figure 9: Hydrogen liquefaction plant in Germany [21]

4.2.2 LH2 Storage

The LH2 storage mainly consists of a horizontal cylindrical cryogenic tank with a capacity of 270 000 litres (17 000 kg of LH2) at temperature of -253°C and pressure of 0.13 MPa, pressure building circuits and piping system. Appendix 5 shows a simplified P&ID of the system. The LH2 tank is filled directly from a liquefaction plant. It is equipped with the level indicator (LI), level switch (LSHL) and a trip switch (LSH) at successively higher levels. It has two independent shutoff valves V-1 and V-3, both of which are operator actuated. The LI is simply an indicator which transmitted to control room. The LSHL is connected to an audible/light alarm and LSH to an automatic trip system and close main valve of the plant (PCV-40).

The tank pressure is maintained by a pressure building circuit which mainly consists of a coil (ambient evaporator, D) and its pressure regulator (PCV-1). The circuit vaporizes liquid hydrogen from the bottom of the tank and sends hydrogen in the gas phase to the tank (top). Operation of the circuit is controlled by the PCV-1 triggered by PIC-1 based on the tank pressure (low) obtained and transmitted by pressure transmitter (PT). When the pressure in the tank is lower than the set point of the PCV-1, then the circuit is working.

In order to protect the tank against overpressure, two pressure safety valves (SV-1, SV-2) are installed. One of the safety valves is operated exchangeable at relative pressure of 0.143 MPa. Additionally, the PCV-2 is used as the secondary pressure relief devices. The operation of the PCV-2 is similar to the PCV-1 but it opens if the tank pressure is high. In addition, the tank is equipped with pressure switches (PIS and PSHL) used to protect the tank against excessive lower pressure. The PSHL activates the PCV-3 to close in case of the pressure is very low.

Table 2: The most important capacity and dimension of the LH2 storage

Components	Dimension	Capacity
1. LH2 tank	Horizontal cryogenic tank	V = 270 000 litres (17 000 kg)
2. Liquid lines	Diameter of 101.6 mm (4 in)*	180 kg/h
3. Vapour lines	Diameter of 101.6 mm (4 in)*	180 kg/h

Source: Linde AG; (*) estimated value

5.0 RESULTS AND DISCUSSION

The fault trees [22] include the sub-fault trees representing the instantaneous and continuous (both in liquid and vapour phases) release of hydrogen from the containment systems which describe the hydrogen production and hydrogen storage activities. The sub-fault trees include in this project are as follows:

1) Continuous release of hydrogen from GH2 tank in the production plant (Case G1.2)
2) Instantaneous release of hydrogen from LH2 tank at depot (Case G2.1)
3) Instantaneous release of hydrogen in liquid phase from LH2 tank at depot (Case G2.2)
4) Continuous release of hydrogen in vapour phase from LH2 tank at depot (Case G2.3)

The methodology explained in Chapter 3 has been implemented in the BN modelling software GeNIe 2.0 which unrolls the DBN over the required number of time slices. For example, 10 time-slices represents 10 years. For all the results shown, the posterior probability failure shown will just describe the failure probability of top event for every case in order to show the significant probability failure. For each sub-fault trees, the numbers of time-slices assign to each of them depend on the complexity of the fault tree. Complex FT require low time-slice because higher time-slice will make the network structure too complex and the task to update the beliefs will be impossible.

For the analysis purpose, the computation using the GeNIe 2.0 modeling tool software is based on the temporal reasoning assigned to every parent nodes with the order of 1 since the assumption made is that for all the events that happen in the future, the probability failure during that time is estimate based on the probability failure on the previous year.

5.1 Probability Failure Distribution for System Failure (Top Event) at Time t for all cases

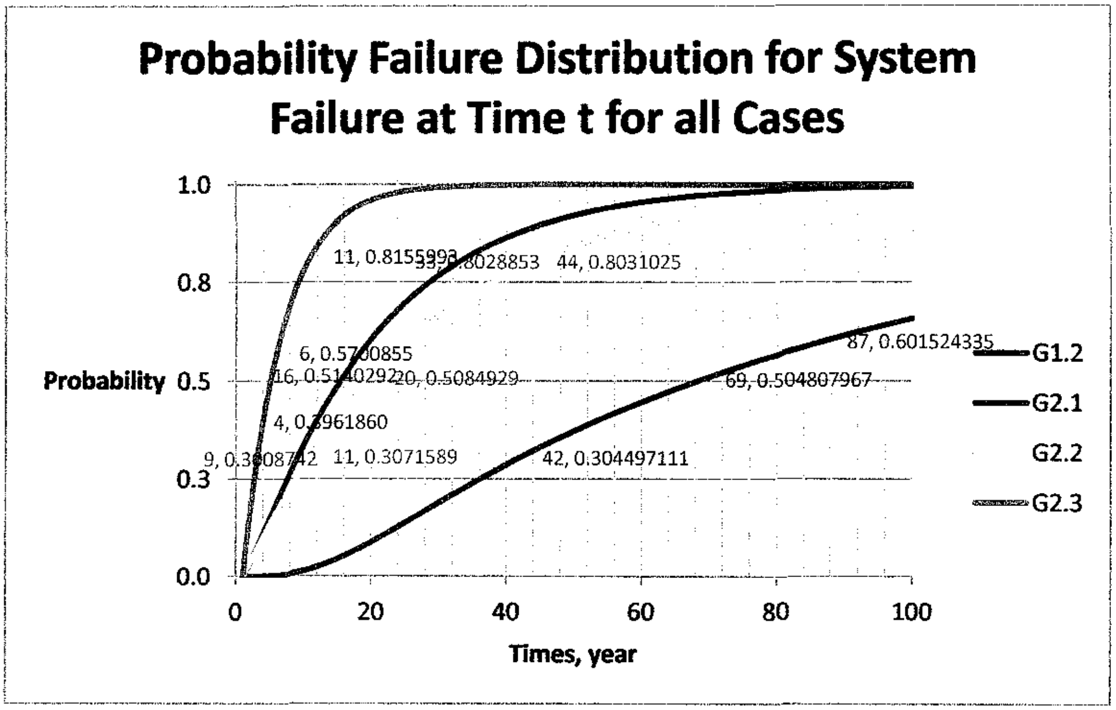


Table 3: Probability failure distribution for system failure at time t for all cases

The graph showed the probability failure distribution for system failure at time t for all cases of hydrogen production and hydrogen storage without maintenance. For the overall observation, it is observed that for all cases, the probability failure increase from year to year but not in a linear way. Besides that, case G2.3 is the case that reach 100% probability failure faster compared to other cases which occur on year 59 followed by case G1.2 with 99.5% on year 100 and case G1.2 with 97.7% on year 100. Case G2.1 is the case with the longest period of time to reach 100% failure probability with the failure probability at year 100 only reaches 65.83%. The selection of the level of probability failure depend on the company's policy in terms of dealing with risk failure and also based on the principle of as low as practicably possible (ALARP) so that to avoid unexpected incidents. For all cases, the probability failure at 30%, 50%, and 80% is observed.

In maintenance aspect, the personnel can decide the level of failure that will be used as the critical point in order to start performing maintenance. If the personnel decide that the critical limit to be 30% of probability failure, then they can observe on what year the 30% level is reached. With this information, the maintenance can be performed the year before the probability failure reach 30%. The same goes to the probability failure of 50% and 80%. This information also helps to increase the operation reliability and effectiveness. In addition, with better decision making on scheduling the maintenance for the operation, it can save a lot of maintenance cost and therefore increase the profitability. The maintenance that is supposed to be performed for every consecutive year can now be reduced.

For case G1.2 which is the continuous release of hydrogen from the GH2 tank at the production plant, it is observed that for the probability failure to reach 30%, it happened on year 9 while for the 50% probability failure, it occurred on year 16. The 80% probability failure is reached on year 33. Based on this information, the time it takes for the system to reach 50% failure is 16 years which will give the personnel more than enough time to identify potential measure for maintenance as well to identify critical events that contribute more to the top event. On the other hand, for case G2.1, instantaneous release of hydrogen from LH2 tank at depot, it is observed that the probability failure reach 30% on year 42 and 50% probability failure on year 69. For 100 years observation, the probability failure for the system did not reach 80% so as the probability failure only reach 65.83%. Based on the information obtain, it takes 42 years for the system to reach 50% probability failure. So, the failure risk is quite low for this case among other case. Next, for case G2.2 which is the instantaneous release of hydrogen in liquid phase from LH2 tank at depot, it is observed that for the 30% probability failure, it happened on year 11 and year 20 for 50% probability failure followed by year 44 for 80% probability failure. For the last case which is the continuous release of hydrogen in vapour phase from LH2 tank at depot (case G2.3), it takes 4 year for the system to reach 30% probability failure followed by 50% probability failure on year 6 and 80% probability failure on year 11.

Based on the probability failure results of the study model, GH2 storage has a lower accident frequency compared with the LH2 storage. The reason is that the LH2 introduces more potential hazards than the one in GH2 (i.e. cryogenic liquid hazards). All these may contribute to modes of potential failure and result in great contributions to the overall release frequency.

The loss of containment events (LOCs) of hydrogen storages and transportation considered in the QRA study include: continuous and instantaneous release. From the results, probability of occurrence per year of the instantaneous release is low compared to continuous release. The instantaneous release of hydrogen mainly results from a catastrophic failure of tank storage (e.g. tank rupture), and release the all inventory contents. Tank rupture is mainly caused by tank overpressure (with the contribution of more than 50%), followed by external events and spontaneous events.

In case of LH2, there is an additional incident that may contribute to the tank rupture i.e. tank under-pressure. All hydrogen storages are equipped with redundant safety protection against tank overpressure such as pressure relief valves and rupture disks. The tank overpressure may lead to tank rupture if all pressure relief devices fail close. The tank overpressure is mainly caused by tank overfilling, loss of vacuum (in case of LH2 only), external fire, internal explosion, overheating of the pressure building circuits (in case of LH2 only) and so on. The continuous release gave the greatest contribution to the loss of containment event in this study. It is mainly caused by tank leakage. Although the tank leakage event may be considered as a rare event but it may result severe damage to the environment. In case of LH2 storage, an additional release may be resulted from pressure building circuit (PBC) failure. In addition, for the failure risk obtain, the result will be more effective if the maintenance variables is included as the joint probability for the system which means the results obtain will take into consideration the maintenance action.

6.0 CONCLUSION AND RECOMMENDATION

6.1 Relevancy to the Objectives

Bayesian Networks approaches have become increasingly popular in risk study. It permits the combination of disparate information streams including historical data, expert knowledge and opinion, and noisy observations which make it more reliable to deal with uncertainty compare to the traditional fault tree. However, to further explore the uniqueness of BN, Dynamic Bayesian Networks is used to show more clearly the relationship between the probability failures with time. Until now, lot of study adapt the DBN approach. So, it is also believe that it is also well suited to hydrogen infrastructure modeling.

6.2 Expected Outcome

The risk failure study done on the hydrogen activities shown that the GH2 storage gives lower probability failure compare to LH2 storage. Besides that, this study also proof that instantaneous release is more unlikely to happen compare to the continuous release. The probability failure for all case studies showed an increase from year to year since the data available try to project the future data instead of playing with the past data. Besides that, with the information on failure risk over time, it is convenience to decide the prevention measure based on personnel opinion on managing the risk failure.

6.3 Future Works for Expansion and Continuation

In order to produce more reliable results, the reliability data for basic events should be more accurate. In addition, the future works should also focus on building DBN model based on the past history in order to identify threshold level/time which require additional action before it probability failure rise. Besides that, the scope should be enlarge to include all the hydrogen activities that contribute to the hydrogen economy so that the outsider can have absolute look on the potential of hydrogen economy. The most important is to include maintenance variable as the joint probability in the system so that the results obtain will be more similar to the real life situation.

REFERENCES

[1] Leveson NG. Software: System Safety and Computers. Reading, MA: Addison-Wesley, 1995.

[2] Henley EJ, Kumamoto H. Reliability Engineering and Risk Assessment. Englewood Cloffd, NJ: Prentice Hall, 1981.

[3] Deb, S. Pattipati, K.R. Raghavan, V. Shakeri, M. Shrestha, R. "Multi-signal Flow Graphs: A Novel Approach for System Testability Analysis and Fault Diagnosis," AUTOTESTCON 1994. Annaheim, CA, 20-22 Sept 1994, pp 361-373.

[4] V. Mihajlovic, M. Petkovic. Dynamic Bayesian Networks: A State of the Art. Computer Science Department, University of Twente, Netherlands, 2001.

[5] A. Singhal, C. Brown, "Dynamic Bayes Net Approach to Multimodal Sensor Fusion," Proceedings of the SPIE – The International Society for Optical Engineering, 3209-32020, October 1997.

[6] K. Murphy, S. Mian, "Modelling Gene Expression Data Using Dynamic Bayesian Networks. Technical Report, Computer Science Division, University of California, Berkeley, CA, 1999.

[7] R. Sterritt, A.H. Marshal, C.M. Shapcott, S.I. McClean. Exploring Dynamic Bayesian Belief Networks for Intelligent Fault Management System. IEEE International Conference on Systems, Man and Cybernetics V. Nashville, Tennessee, USA, pp. 3646-3652, 2000.

[8] V. Pavlovic. Dynamic Bayesian Networks for Information Fusion with Application to Human-Computer Interfaces. PhD Thesis, Uniersity of Illinois at Urbana-Champaign, 1999.

[9] G.F. Cooper. The Computational Complexity of Probabilistic Inference using Bayesian belief Networks. Artificial Intelligence, vol. 42, pp. 393-405, 1990.

[10] T. Dean, K. Kanazawa. Probabilistic Temporal Reasoning. In Proceedings of the 7th National Conference on Artificial Intelligence (AAAI-88), pages 524-529, St. Paul, MN, USA, August 1988.

[11] K.P. Murphy. Dynamic Bayesian Networks: Representation, Inference, and Learning. PhD Thesis, Computer Science Division, University of California, Berkeley, CA, USA, 2002.

[12] Bayesia SA. BayesiaLAB 3.0 Tutorial, 2004.

[13] Norsys Software Corporation. Netica Application for Belief Networks and Influence Diagrams – User Guide, March 1997.

[14] K.P. Murphy. The Bayes Net Toolbox for Matlab. Computing Science and Statistics, 33:331-350, October 2001.

[15] K. Wojtek Przytula, Richard Milford. An Efficient Framework for the Conversion of Fault Tree to Diagnostic Bayesian Network Models.2005.

[16] F. Jaier Diez. Parameter Adjustment in Bayes Networks: The Generalized Noisy-OR Gate. In Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence, (UAI-93), pg 99-105, Washington D.C., 1993.

[17] W.L. Buntine. Operations for Learning with Graphical Models. Journal of Artificial Intelligence Research, 2:159-225, December 1994.

[18] Szyska, A. Ten Years of Solar Hydrogen Demonstration Project at Neunburg vorm Wald, Germany. International Journal of Hydrogen Energy, 23, pp. 849-860, Elsevier, 1998.

[19] Szyska, A. Thirteen Successful Years of SWB's solar hydrogen demonstration Project at Neunburg vorm Wald, Germany. Manuscript No. 23, Germany, July 1999.

[20] SWB: Towards the future – The Solar Hydrogen Project in Neunburg vorm Wald, Germany. Solar Wassertoff Bayern GmbH, 1999.

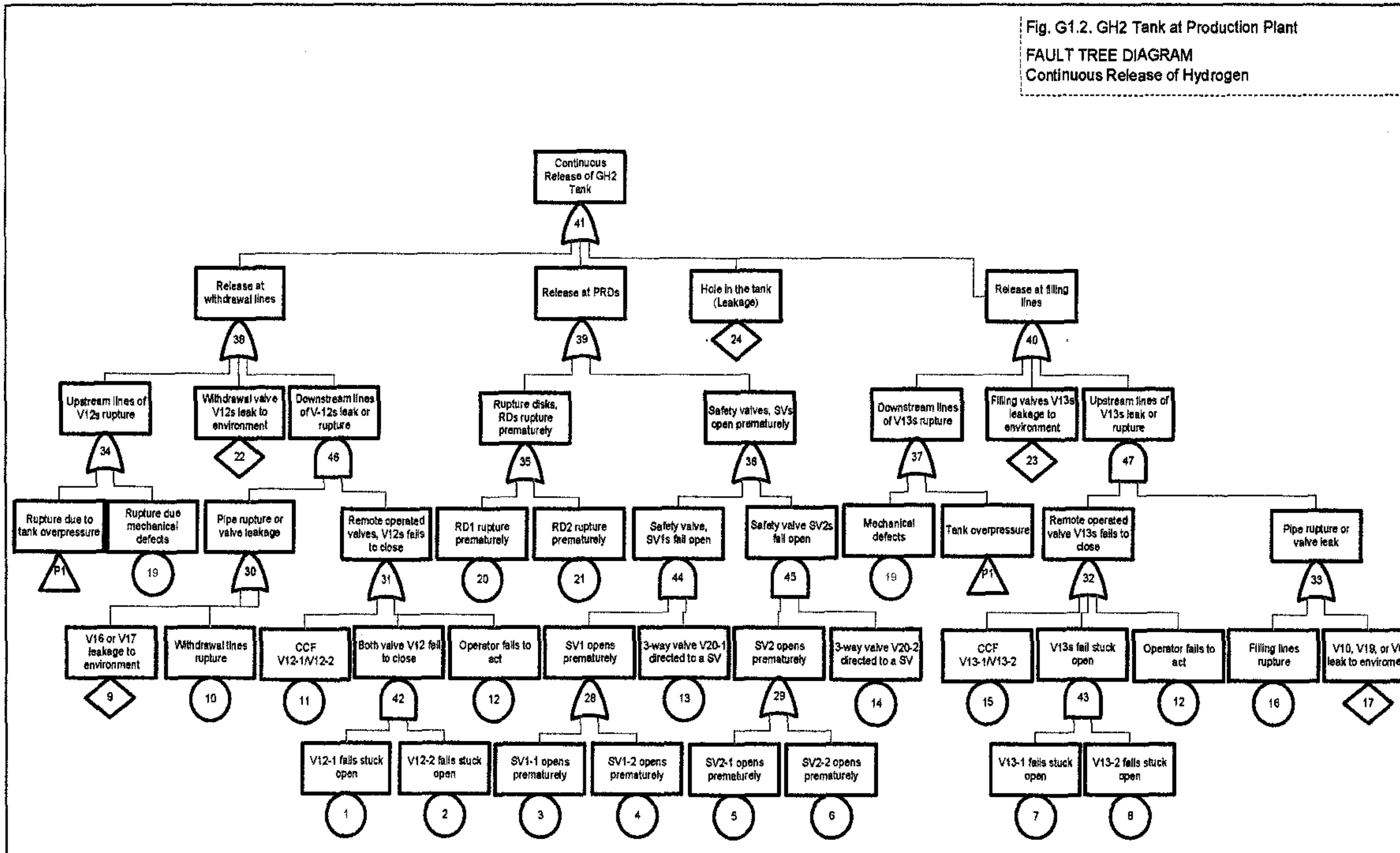
[21] Gross, R., W. Otto, A. Patzeit, M. Wanner. Liquid Hydrogen for Europe – the Linde Plant at Ingolstadt. Report on Science and Technology, Linde AG, 54, pp. 37-43, 1994.

[22] Abdul Rosyid. System-analytic Safety Evaluation of the Hydrogen Cycle for Energetic Utilization. M.Sc., University of Magdeburg, Germany, 2006.

[23] GeNIe. Decision Systems Laboratory, University of Pittsburgh. http://genie.sis.pitt.edu/

APPENDIX

Appendix 1: Fault Tree for Continuous Release of Hydrogen for GH2 Tank in Production Plant [22]

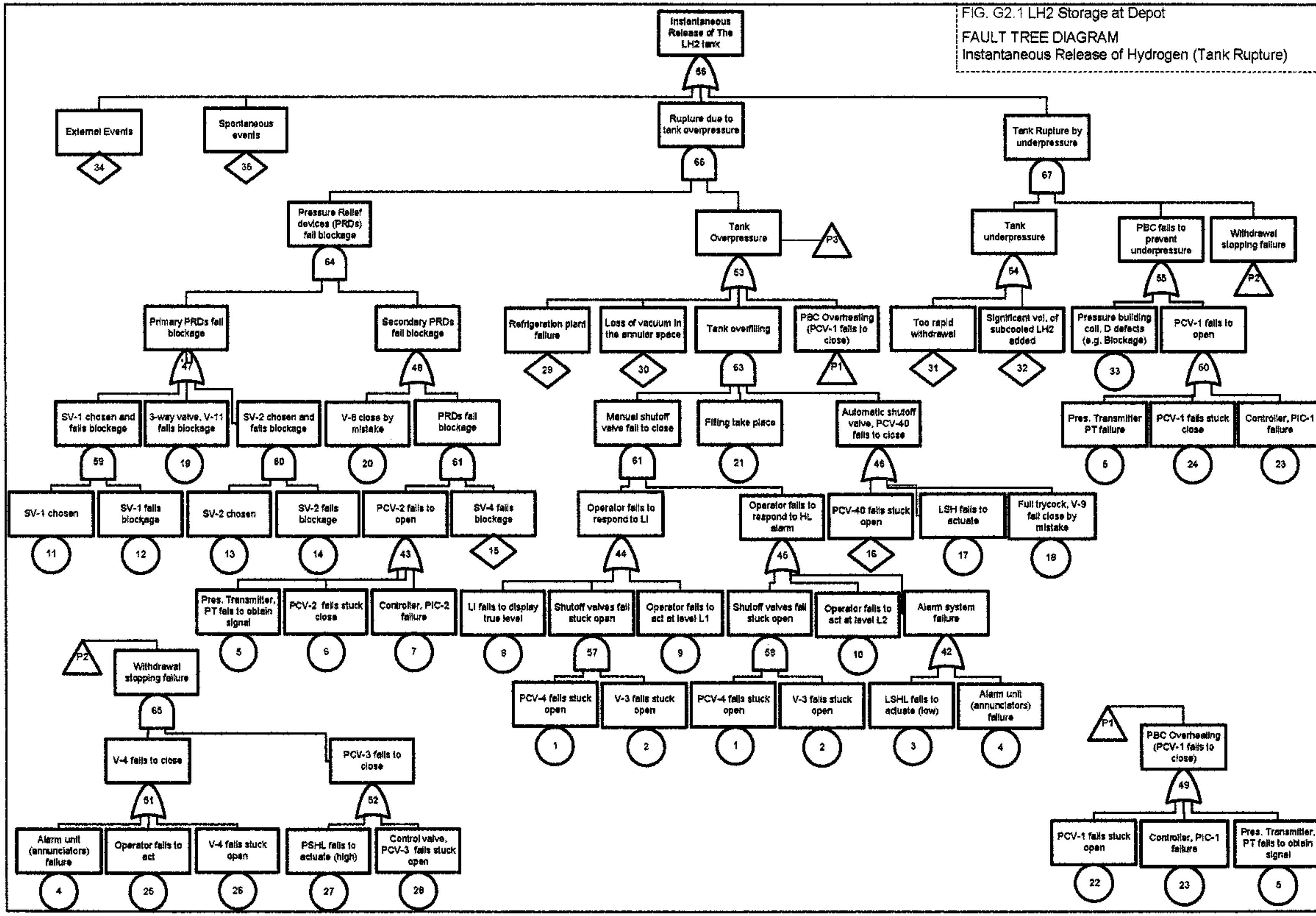


Appendix 2. Estimate of basic event probabilities and frequencies for case G1.2 fault tree

Basic Event No.	Component/ID symbol	Description of the event/failure mode	F/P	Median	Uncertainty factor, EF	Source of justification
1	V12-1	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
2	V12-2	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
3	SV1-1	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
4	SV1-2	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
5	SV2-1	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
6	SV2-2	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
7	V13-1	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
8	V13-2	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
9	PCV-16, 17	Pressure control valves leak to environments (2 units)	F	8,8E-07	5,5	[224], process control valve
10	-	Withdrawal line rupture	F	2,7E-08	10,0	[8], straight section
11	CCF V12	Common cause failures of V12-1 and V12-2	F	1,8E-06	6,7	0.1 x Basic no.1
12	-	Operator fails to act (fails to change valve position)	P	1,0E-03	2,0	[88,137]
13	V20-1	Three-way valve is directed to one safety valve	P	5,0E-01	2,0	Guess estimate, 50%
14	V20-2	Three-way valve is directed to one safety valve	P	5,0E-01	2,0	Guess estimate, 50%
15	CCF V13	Common cause failures of V13-1 and V13-2	F	1,8E-06	6,7	0.1 x Basic no.7
16	-	Upstream line of V13s rupture	F	2,7E-08	10,0	[8], straight section
17	V10	Hand operated (isolation) valves leak to environment	F	1,1E-06	13,8	[184]
18		Pipe rupture due to tank overpressure	F	1,6E-11	35,0	fault tree analysis
19	-	Pipe rupture due to mechanical defect	F	3,6E-10	10,0	[8, 158]
20	RD1	Rupture disk fails break prematurely	F	1,0E-06	10,0	[183]
21	RD2	Rupture disk fails break prematurely	F	1,0E-06	10,0	[183]
22	V12s	Remote operated (isolation) valve leaks to environment	F	3,8E-07	5,5	[224], ESD valve
23	V13s	Remote operated (isolation) valve leaks to environment	F	3,8E-07	5,5	[224], ESD valve
24	-	Hole in the tank	F	2,1E-10	20,8	[8, 116, 159], pressure vessel

F= Frequency [/h], P=Probability [-]

Appendix 5. Fault Tree for Instantaneous Release of Hydrogen from LH2 Tank at depot [22]



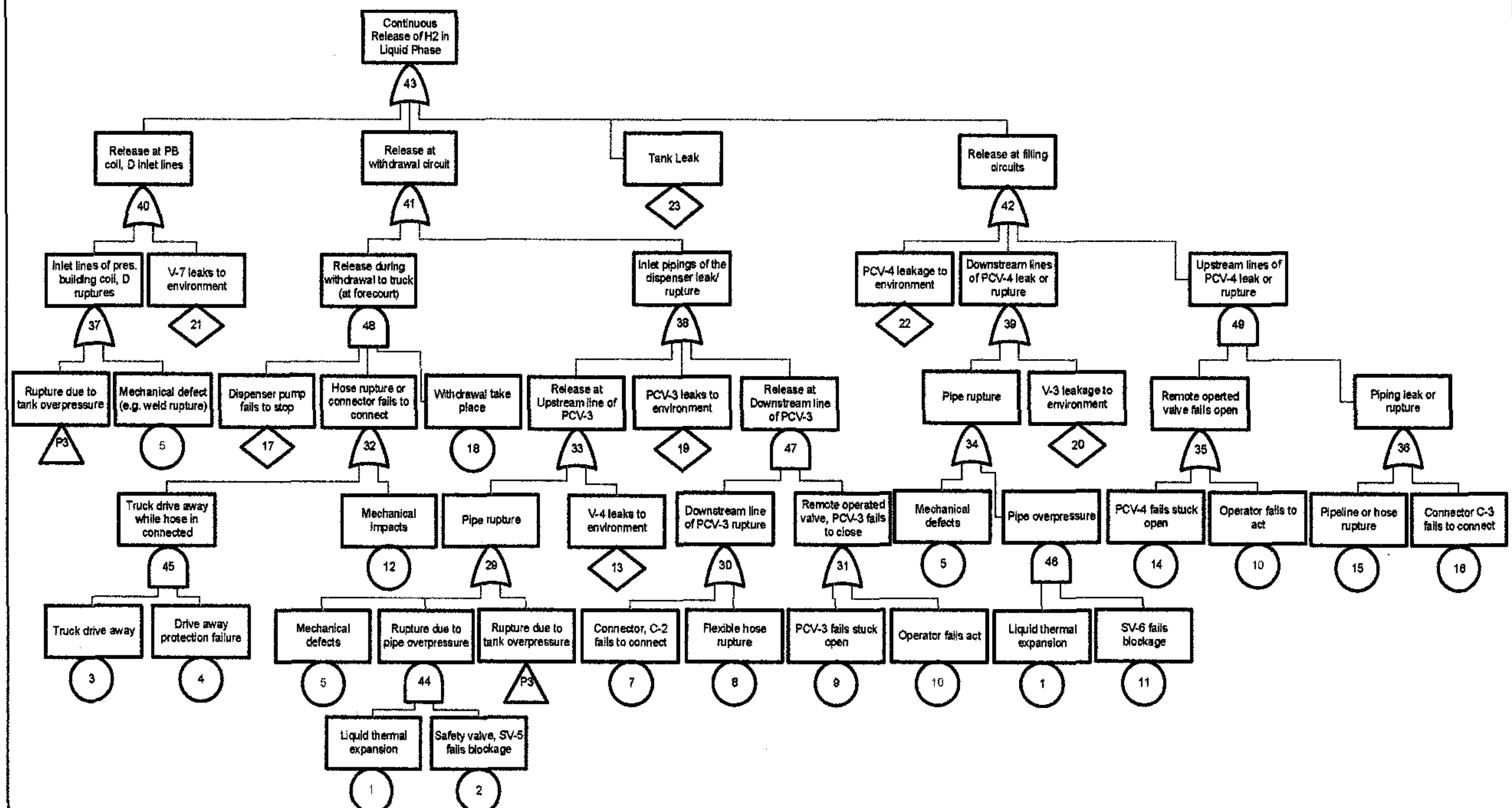
1	PCV-4	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
2	V-3	Hand operated (isolation) valve fails stuck open	F	8,4E-08	6,0	[8], manual valve
3	LSHL	Level switch of high level fails to actuate (low)	F	2,8E-05	13,8	[8]
4		Alarm unit (annunciator) fails to sound	F	2,8E-07	10,3	[8], annunciators
5	PT	Pressure transmitter fails to obtain signal	F	5,8E-07	32,0	[224]
6	PCV-2	Control valve fails stuck close	F	6,1E-07	8,4	[224]
7	PIC-2	Pressure controller fails to operate	F	1,1E-05	14,9	[8]
8	LI	Level indicator fails to display true level	F	7,7E-07	5,6	[184]
9	-	Operator fails to act at level L1	P	1,0E-03	2,0	[88,137]
10	-	Operator fails to act at level L2	P	1,0E-03	2,0	[88,137]
11	SV-1 ch	Pressure safety is chosen to operate	P	5,0E-01	2,0	Guess estimate, 50%
12	SV-1	Spring loaded pressure safety valve fails blockage	F	1,1E-06	8,4	[88]
13	SV2 ch	Pressure safety is chosen to operate	P	5,0E-01	2,0	Guess estimate, 50%
14	SV-2	Spring loaded pressure safety valve fails blockage	F	1,1E-06	8,4	[88]
15	SV-4	Spring loaded pressure safety valve fails blockage	F	1,1E-06	8,4	[88]
16	PCV-40	Plant valve fails to close (emergency shutdown valve)	F	4,4E-06	2,8	[184], ESD valve
17	LSH	Very high level switch fails to actuate (low)	F	2,8E-05	15,1	[8]
18	V-9	Full trycock fails close by mistake	P	1,0E-03	2,0	[88,137]
19	V-11	Three-way valve fails blockage	F	3,3E-06	4,7	[241]
20	V-8	Hand operated valve close by mistake	P	1,0E-03	2,0	[88,137]
21	-	Filling take place	F	8,7E-03	5	76 times/yr, operat. data
22	PCV-1	Control valve fails stuck open	F	2,5E-07	27,7	[224]
23	PIC-1	Pressure controller fails to operate	F	1,1E-05	14,9	[8]
24	PCV-1	Control valve fails stuck close	F	6,1E-07	8,4	[224]
25	-	Operator fails to act	P	1,0E-03	2,0	[88,137]
26	V-4	Hand operated (isolation) valve fails stuck open	F	8,4E-08	6,0	[8], manual valve
27	PSHL	Very low pressure switch fails to actuate (high)	F	5,6E-06	49,0	[8], p.160
28	PCV-3	Control (isolation) valve fails stuck open	F	2,5E-07	27,7	[224], process control valve
29	-	Refrigeration plant failure (e.g. resulting GH2)	F	2,0E-07	3,3	[88], Guess estimate
30	-	Loss of vacuum in the annular space	F	1,4E-08	2,9	Guess estimate
31	-	Withdrawal too rapid (excess withdrawal rates)	F	2,3E-08	10	Guess estimate
32	-	Significant volume of subcooled LH2 added	F	8,7E-08	5	Guess estimate
33	D	Pressure building coil fails to operate	F	6,7E-11	30	[240]
34	-	External events (e.g. Earthquake)	F	1,1E-12	10,0	[116], Table A8.3
35	-	Spontaneous events (e.g. H2 embrittlement)	F	1,0E-11	10,0	[116]

F= Frequency [/h], P=Probability [-]

FIG. G2.2 LH2 Storage at Depot

FAULT TREE DIAGRAM

Continuous Release of Hydrogen in Liquid Phase



Appendix 6: Estimate of basic event probabilities and frequencies for case G2.2 fault tree

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	-	Liquid thermal expansion	F	1,1E-07	10	Guess estimate, likely incident
2	SV-5	Spring loaded pressure safety valve fails blockage	F	1,1E-06	8,4	[88]
3	-	Truck drives away	F	1,1E-06	10	Guess estimate, 1 per 100 years
4	-	Drive away protection failure	F	1,1E-08	10	Guess estimate, unlikely incident
5		Weld rupture	F	3,0E-10	31,6	[116], Table A14.4
6		Pipe rupture due to tank overpressure	F	2,1E-10	38,2	FTA
7	C-2	Connector fails to connect	F	1,5E-07	14,9	[8], page. 184
8	-	Flexible hose to dispenser ruptures	F	1,5E-07	14,9	[8], hoses
9	PCV-3	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
10	-	Operator fails to act	P	1,0E-03	2,0	[88,137]
11	SV-6	Spring loaded pressure safety valve fails blockage	F	1,1E-06	8,4	[88]
12	-	Mechanical impact (e.g. cranes etc)	F	1,1E-11	10,0	[116], Table IX.2
13	V-4	Hand operated (isolation) valves leak to environment	F	1,1E-06	13,8	[224]
14	PCV-4	Remote operated (isolation) valve fails stuck open	F	1,8E-05	2,2	[88], pneu. shutoff valve
15	-	Filling lines from the plant ruptures	F	6,9E-10	15,0	[8], p. 183, metal pipe
16	C-3	Connector fails to connect	F	1,5E-07	14,9	[8], page. 184
17		Dispenser valve fails stuck open or LH2 pump fails to stop	P	3,8E-05	10,0	[116], Table A8.2
18		Withdrawal takes place	F	2,3E-02	10,0	200/year
19	PCV-3	Remote operated (isolation) valve leaks to environment	F	1,0E-07	10,0	[183]
20	V-3	Hand operated valve leaks to environment (rupture)	F	1,1E-06	13,8	[224]
21	V-7	Hand operated valve leaks to environment (rupture)	F	1,1E-06	13,8	[224]
22	PCV-4	Remote (control) operated valve leaks to environment	F	3,8E-07	5,5	[224]
23		Serious leakage from inner tank (refrigerated, double wall)	F	2,3E-09	10,0	[116], Table A8.2

F= Frequency [/h], P=Probability [-]

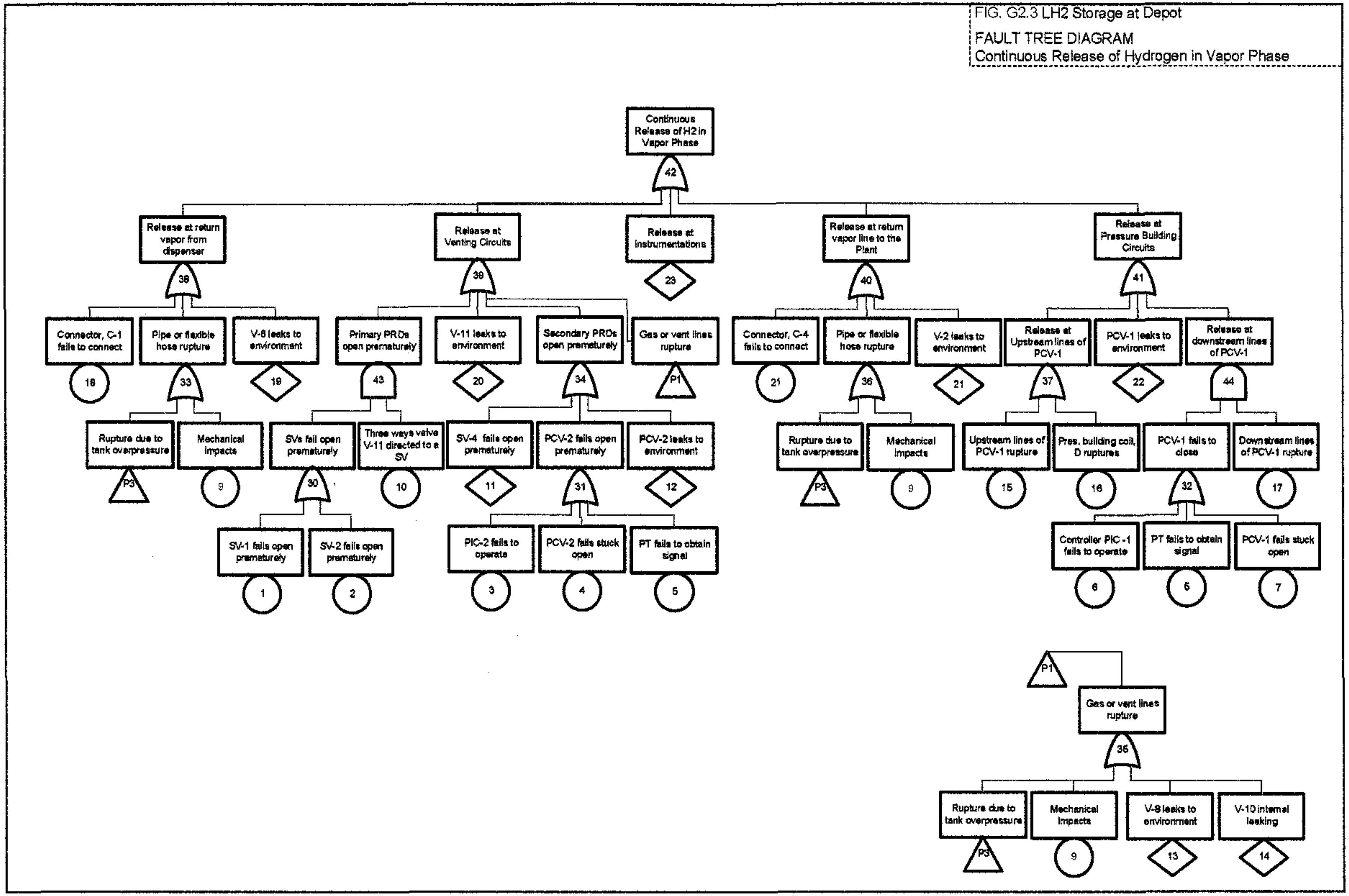


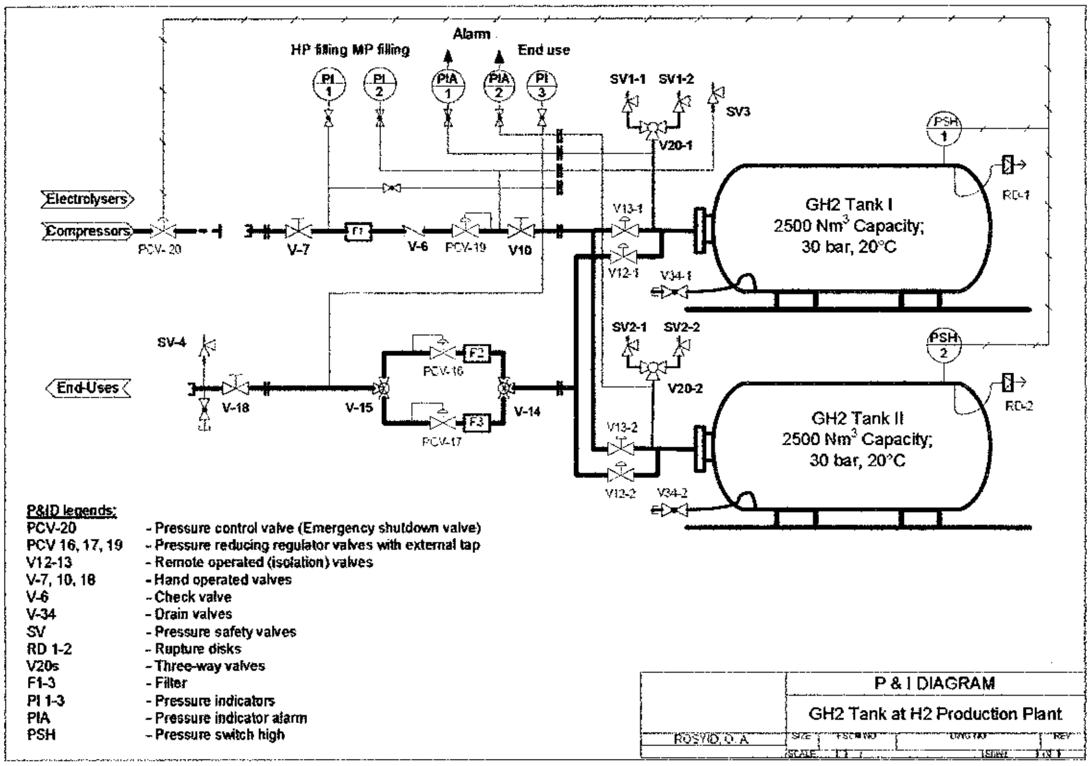
FIG. G2.3 LH2 Storage at Depot
FAULT TREE DIAGRAM
Continuous Release of Hydrogen in Vapor Phase

Appendix 6. Estimate of basic event probabilities and frequencies for case G2.3 fault tree

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	SV-1	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
2	SV-2	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
3	PIC-2	Pressure controller fails to operate	F	1,1E-05	14,9	[8]
4	PCV-2	Control valve fails stuck open	F	2,5E-07	27,7	[224], process control valve
5	PT	Pressure transmitter fails to obtain signal	F	5,9E-07	2,4	[240]
6	PIC-1	Pressure controller fails to operate	F	1,1E-05	14,9	[8]
7	PCV-1	Control valve fails stuck open	F	2,5E-07	27,7	[224], process control valve
8		Pipe rupture due to tank overpressure	F	2,1E-10	38,2	FTA
9		Mechanical impact (e.g. cranes etc)	F	1,1E-11	10,0	[116], Table IX.2
10	V-11	Three-way valve is directd to one safety valve	P	5,0E-01	2,0	
11	SV-4	Pressure safety valve fails open prematurely	F	4,0E-07	3,9	[224], PSV-conventional
12	PCV-2	Control valve leaks to environment	F	4,4E-07	5,5	[224], process control valve
13	V-8	Hand operated valve leaks to environment	F	1,1E-06	13,8	[224]
14	V-10	Hand valve leaks in close position (internal leakage)	F	7,3E-08	11,5	[224], p. 608
15		Upstream line of the PCV-1 (pipe) ruptures	F	6,9E-09	15,0	[8]
16	D	Pressure building coil ruptures	F	9,7E-10	2,5	[116] Table A14.6, [159] Table 3.11
17		Downstream line of the PCV-1 ruptures	F	6,9E-09	15,0	[8]
18	C-1	Connector fails to connect	F	1,5E-07	14,9	[8], page. 184
19	V-6	Hand operated valve leaks to environment	F	1,1E-06	13,8	[224]
20	V-11	Three-way valve leaks to environment	F	1,1E-06	13,8	[224]
21	C-4	Connector fails to connect	F	1,5E-07	14,9	[8], page. 184
22	V-2	Hand operated valve leaks to environment	F	1,1E-06	13,8	[224]
23	PCV-1	Control valve leaks to environment	F	4,4E-07	5,5	[224], process control valve
24		Release at instrumentation system	F	1,0E-09	10,0	Guess estimate

F= Frequency [/h], P=Probability [-]

Appendix 9: Simplified P&ID of the GH2 storage



Appendix 10: Simplified P&ID of the LH2 storage