ISBN 978-0-620-72719-8

Validation & Verification of a Bayesian Network Model for Aircraft Vulnerability

Sunelle Schietekat
sschietekat@csir.co.za
Optronic Sensor Systems (OSS)
DPSS, CSIR, Pretoria, South Africa

Alta de Waal
alta.dewaal@up.ac.za
Statistics Department
University of Pretoria, South Africa

Kevin G. Gopaul
kgopaul@csir.co.za
Optronic Sensor Systems (OSS)
DPSS, CSIR, Pretoria, South Africa

Abstract. This paper provides a methodology for Validation and Verification (V&V) of a Bayesian Network (BN) model for aircraft vulnerability against Infrared (IR) missile threats. The model considers that the aircraft vulnerability depends both on a missile's performance as well as the doctrine governing the missile's launch. The model is a Knowledge Based System (KBS) and therefore has a knowledge base which consists of both expert knowledge and simulated data which acts as input to the model and is used during inferencing to understand how variables interact. A widely accepted process to certify that a model is suitable for use is the Verification, Validation and Accreditation (VV&A) procedure and is followed in this paper. Throughout the V&V procedure, similarities are drawn between this VV&A process and the well-known Vee-model.

1. Introduction

"We demand rigidly defined areas of doubt and uncertainty!"
\- Douglas Adams -

A model has been developed by Optronic Sensor Systems (OSS) in the defense field to evaluate aircraft vulnerability against Infrared (IR) missile threats. The model considers that the aircraft vulnerability depends both on a missile's performance as well as the doctrine governing the missile's launch (Willers *et al.*, 2014). Data about the probability for a missile launch is captured as expert knowledge during work sessions with domain experts. Data about the probability that the missile will hit the aircraft (miss distance) is captured from simulation data generated by Countermeasure Simulation 2 (CmSim2), an evaluation tool that forms part of the Optronics Scene Simulator (OSSIM) developed by the Council for Scientific and Industrial Research (CSIR) and Denel Dynamics (Willers & Willers 2011). The purpose of the vulnerability model is to use it for inference: calculating the posterior probability distribution for a variable given a certain input scenario selected. The user can then evaluate different scenarios or use cases to understand how the different variables interact. In other words, the user reasons about the problem domain. This is known as a knowledge based system.

A Knowledge Based System (KBS) is a system with two components: a knowledge base and an inference engine. The knowledge base contains information about the system, whereas the inference engine contains logical rules about the system typically in the form of IF-THEN and

ISBN 978-0-620-72719-8

WHAT-IF statements. For example, *VAT is calculated at 14%* would be information about the system and *VAT is not added to the price of brown bread* would be reasoning about information within the system. KBSs are very powerful and used in many applications which require reasoning about outcomes of complex systems that often contain incomplete information and for which learning forms a part of the system. Expert systems are good examples of KBSs, such as medical diagnosis (Lauritzen & Spiegelhalter, 1988). One popular fictional character that can be considered the human epitome of a KBS is Sherlock Holmes. His careful observations of the world form his knowledge base and with that he is able to make brilliant chains of inference as a detective.

One way to practise KBS is to employ Bayesian Networks (BNs). BNs thrive in a world of uncertainty, underpinned by causality. It is a network with variables that are linked by cause and effect; where not only historical information can be captured, but new information can be trained; and the result, an integrated tool that can predict future outcomes (Pearl, 2014). One of a BN's notable advantages is data fusion, capable of integrating data from several sources even if this data is incomplete (Koen *et al.,* 2014). Especially so in this unique field of work, where the most fitting solutions need to be found while acknowledging and embracing the uncertainties associated with these threat scenarios.

The vulnerability model serves as a proof of concept, and is still incomplete as it has not yet undergone formal Validation & Verification (V&V). V&V is an important aspect of Systems Engineering that aims to confirm that a system meets its requirements and is fit for its intended use in its intended environment (Walden *et al.*, 2015). The V&V process is a continuous activity and should be applied at every applicable stage throughout the model development and not reserved as a final check (Balci, 1995).

A simple version of the model was presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) conference in 2014. This model shall be referred to as the "vulnerability model" (Willers *et al.*, 2014). A slight variation on this model is used to describe the V&V process which will form the basis for future model V&V. Note that the vulnerability model uses information available on the internet as the authors are not permitted to disclose any information on specific aircraft and missiles. This paper starts with an overview of the V&V procedure, then proceeds to actual Model V&V and concludes with model assessment.

2. Validation and Verification procedure

The INCOSE handbook notes that a model can be considered a product or system in its own right and it therefore too needs to undergo V&V (Walden *et al.*, 2015). Not all V&V procedures will work for all systems and it consequently needs to be tailored to match the nature and complexity of the specific problem. A widely accepted procedure for certifying that a model is acceptable for use is the Verification, Validation and Accreditation (VV&A) procedure (M&SCO, 2013). The framework proposed by De Waal *et al.* (2013) adopts this VV&A procedure for simulation models and is adapted slightly when used to perform V&V on the BN vulnerability model discussed here. By showing that the model has undergone successful V&V, the user gains confidence that the knowledge resulting from using the model is credible (Walden *et al.*, 2015). Even so, keep in mind that within the military context the model can never provide 100% fool proof answers, but simply equips the client with solutions that attempt to align the odds in their favour when encountered with a threat situation. It is in this context that the following words resonate with our approach (Walden *et al.*, 2015): *"...the risks of not using the model or simulation are greater than the risk of using the model or simulation…"*

ISBN 978-0-620-72719-8

It is important at this stage to define some of the terms that will be used throughout this paper (M&SCO, 2013), (Walden *et al.*, 2015):

> Verification – the process of determining that a model implementation and its associated data accurately represent the developer's conceptual description and specifications. *Did I build the thing right?*
>
> Validation – the process of determining the degree to which a model and its associated data provide an accurate representation of the real world from the perspective of the intended use of the model in its intended environment. *Did I build the right thing?*
>
> Accreditation[1] – the official certification that a model, simulation, or federation of models and simulations and its associated data is acceptable for use for a specific purpose. *Is it believable enough to be used?*
>
> VV&A procedure - the procedure to gather and evaluate evidence to determine whether a simulation's capabilities, accuracy, correctness, and usability are sufficient to support its intended use.

A well-known model used to explain V&V as applied in Systems Engineering is the Vee-model as shown by Fig.1 and was first introduced by what was then known as the National Council On Systems Engineering (NCOSE) (Forsberg & Mooz, 1991).

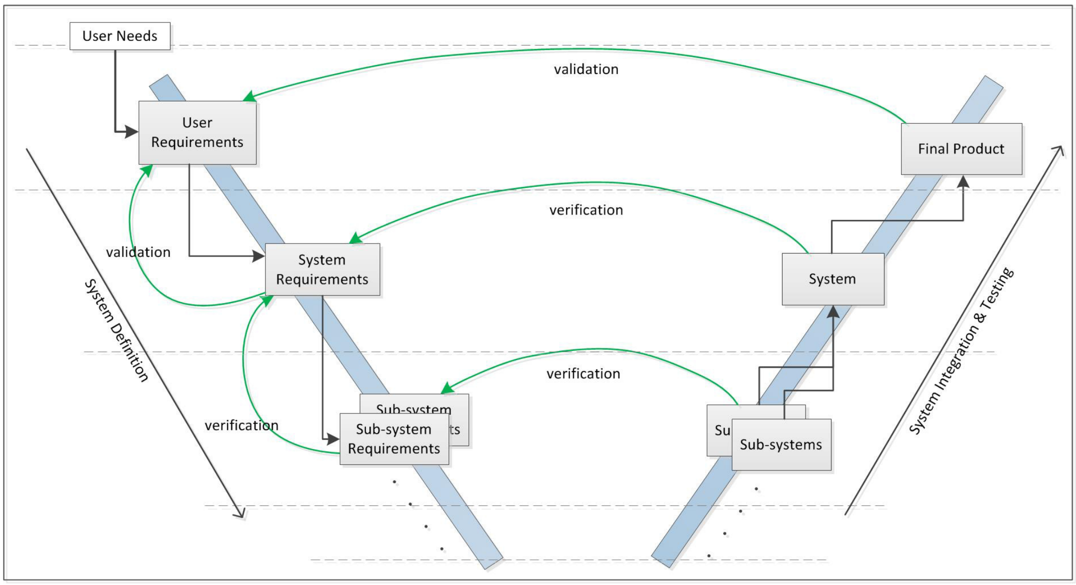

Figure 1. Vee-model used in Systems Engineering.

The process starts on the left upper side of the diagram with a user need that is captured as user requirements and which in turn is further developed into system requirements. The system requirements need to be validated against the user requirements to see whether it is consistent with what the user intended. This process is repeated for sub-systems lower down the left side of the Vee until all requirements have been defined. From here we move to the bottom right side of the diagram where the product or system starts development. Sub-systems are integrated and tested to become larger systems as we move up the right side of the Vee until we reach the final product. Each time, an integrated product needs to be verified against its corresponding system requirements on the left side of the Vee. The final product is validated against the user

[1] It is in the authors' understanding that this model will not formally be accredited as there is no organisation authorised to do so in South Africa.

ISBN 978-0-620-72719-8

requirements. The rule to follow is that verification is performed against technical requirements, whereas validation is performed against user needs.

The VV&A procedure for models as based on Schlesinger's framework is depicted by Fig.2 (Schlesinger *et al.*, 1979), (Sargent, 1981). In this paper the VV&A procedure will be followed as given in this figure and at the same time comparisons will be drawn with the Vee-model from Fig.1.

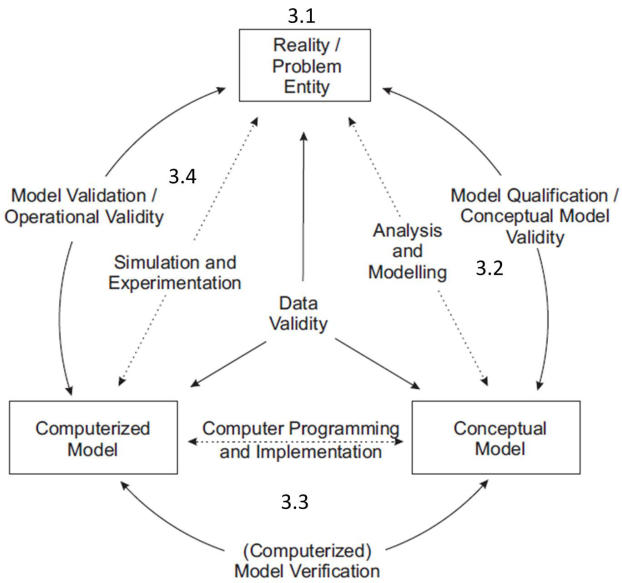

Figure 2. Schlesinger's framework of the VV&A procedure.

The three elements as shown in the figure are *Reality*, *Conceptual Model* and *Computerized Model* which is considered similar to the User Requirements, System Requirements and Final Product respectively from the Vee-model. The three processes that relate these elements are *Analysis and Modelling*, *Computer Programming and Implementation* and *Simulation and Experimentation* and are comparable to the system definition down the left side and system integration and testing up the right side of the Vee-model. The three procedures that evaluate the credibility of the processes are *Conceptual Model Validation*, *Model Verification* and *Model Validation* and are similar to the requirement validation, system verification and final product validation respectively in the Vee-model. Table 1 provides definitions for each. The *Data validity* will not be treated separately, but will form part of the V&V criteria. If we consider that the conceptual model consists of information including designs, requirements and data that forms the blueprint for the model, we can argue that it represents the system requirements as given by the Vee-model.

Table 1: Definitions for the VV&A aspects as per Fig.2.

Reality	This element represents the user requirements including an understanding of how the model will be used and the environment in which it will be used.
Analysis & Modelling	This process describes how the conceptual model is developed and captured in the form of designs, data and requirements.
Conceptual Model	This element consists of all the artifacts resulting from the Analysis & Modelling process in order to adequately describe and capture the model.
Computer programming & Implementation	This process describes all the steps involved with implementing the conceptual model into the software.
Computerized Model	This element represents the model as it is realized in the software.
Simulation & Experimentation	This process describes how the model is exercised and evaluated for its intended use.

ISBN 978-0-620-72719-8

Conceptual Model Validation	Concerned with validating that the conceptual model is consistent with reality (validate system requirements against the user requirements).
Model Verification	Concerned with verifying that the computerized model is representative of the conceptual model (system meets the system specification).
Model Validation	Concerned with validating that the computerized model is representative of reality – fit for its intended use in its intended environment (final product meets user requirements).

2.1 V&V Criteria

A set of criteria is defined that will be under consideration for each step of the V&V procedure:

Model structure (qualitative) - Relates to the physical structure of the model based on the casual relationships of the variables.

Model parameters (quantitative) - Relates to the data used as input to the model and to create the Conditional Probability Tables (CPTs). These include the expert knowledge data and CmSim2 simulation data.

Inferencing - When the model is exercised for the different use cases in a what-if analysis. The different inference modes are predictive, prescriptive and diagnostic mode.

2.2 V&V Methods

Various methods exist for V&V as given by Balci (1995) and Korb & Nicholson (2004). Those that are relevant to this paper are summarized here:

Informal - A subjective, qualitative method that relies on human reasoning, i.e. consulting domain experts (assess face validity or graphical display).

Static - Includes inspecting and analysing the information contained in the model without exercising the model, i.e. data consistency, traceability of information, inspection, or historical validation.

Dynamic – An objective, quantitative method and includes evaluating the parameter variability when exercising the model, i.e. what-if analysis or sensitivity analysis.

Formal - An objective, quantitative method and includes formal proof, i.e. mathematical calculation or logical deduction.

2.3 V&V Methodology

The methodology steps are given here and defined in the sections to follow: a) define the reality, b) develop a conceptual model and perform conceptual model validation, c) develop a computerized model and perform model verification, d) exercise the model by performing model validation.

3. Model Validation and Verification

3.1 Define the Reality

"The formulation of a problem is often more essential than its solution…"
\- Albert Einstein -

The reality element for the model is where the end user is utilizing the model for its intended use in its intended environment. The user need is defined as follows:

ISBN 978-0-620-72719-8

User Need -*"How does a set of variables governing a specific surface-to-air missile engagement influence the aircraft vulnerability?"*

It is helpful for model V&V to answer the following questions:

<u>How will the client use the model?</u> The model will be used as an investigative tool to study the effects of changing the variables within a given scenario. The outcome of this parameter variability study is to gain insight towards the potential future use of an advanced model during mission planning.

<u>In which environment will the client use the model?</u> This version of the model will only be used in a research capacity. Future versions of the model could be used during mission planning and for pilot training.

The user requirement can therefore be formulated as follows:

User Requirement – Develop a method that shall be used in a research capacity to reason about how variables governing a specific surface-to-air missile engagement influence the aircraft vulnerability.

During V&V, the domain experts will represent the user for all practical purposes. This user requirement is derived further as system requirements during the analysis & modelling process (3.2) and validated during the simulation & experimentation process (3.4).

3.2 Develop a Conceptual Model

This section describes how the Conceptual Model is realized through the Analysis and Modelling process (Fig.2). The system requirements that are defined here for each V&V criteria is validated against the user requirement (Fig.1) given in section 3.1.

Based on the user requirement, it was decided that the model will take on the form of a KBS as it enables one to reason about variables under uncertainty. The KBS will be implemented as a bayesian network. BNs models contain nodes that represent variables and arcs which indicate the causal relationships between nodes. Variables contain sets of values or states (Korb & Nicholson, 2004).The construction of a BN model under a variety of circumstances is defined as Knowledge Engineering with Bayesian Networks (KEBN). The major modelling issues that arise are the following and will be answered in section 3.2.1, 3.2.2 and 3.3.3 as the model is developed (Korb & Nicholson, 2004):

- What is the graph structure, the variables and their values/states?
- What are the parameters (probabilities)?
- What inference modes will be used?

3.2.1 Structure

The model structure consists of the physical representation of the model as well as the variables with their states. In this section the variables with states are first defined as given by Table 2, and then the structure is created from these variables.

Table 2: The variable requirements.

Requirement - The model shall consist of the following variables with their description and states*:	
Launch	Whether an IR surface-to-air missile will be launched assuming an experienced operator is presented with a specific set of conditions.
Miss distance	In a simulation, the closest that the missile will pass (in meters) by a specific point defined on the aircraft, assuming that the aircraft continues with a constant speed

ISBN 978-0-620-72719-8

	and altitude, as the missile engages with the aircraft.
Aspect angle	The angle in degrees between the missile and the aircraft.
Range	The distance between the missile and the aircraft in kilometres.
Aircraft altitude	The altitude of the aircraft in feet.

*Note that even though the states for the variables are not given in this paper, it should also be defined.

Some of the variables were chosen as they already exist as part of the CmSim2 environment. These variables include the *miss distance, angle, range* and *altitude.* The validation method is therefore static: historical validation as CmSim2 variables have been validated before. The launch variable is included based on discussions with domain experts and the validation method is therefore informal: domain expert. The values of the variables should also be defined as a requirement. For this model the variables are all discrete and the values are defined as states that are mutually exclusive and exhaustive (Korb & Nicholson, 2004). When dividing the variable value into states or intervals, it is important to understand how it will be interpreted by the software. In our case, the lower bound of each interval is inclusive, while the upper bound of each interval is exclusive (except for the last interval)[2]. The states are validated by either static: historical validation or informal: domain expert. The variables *angle*, *range* and *altitude* will be referred to as the input variables, while *launch* and *miss distance* will be the output variables.

Requirement – The model structure shall be based on a Bayesian Network that captures the intended relationship of the variables.

As described before, a Bayesian Network is a useful technique to model a KBS and is suited to use for the model under consideration. A BN is a casual structure which means that the variables in the network are associated with cause and effect. The design of the structure depends on the relationship between the variables. Different versions of the model were proposed and after careful consideration one was chosen based on several arguments supporting it and the resulting conceptual model given by Fig.3.

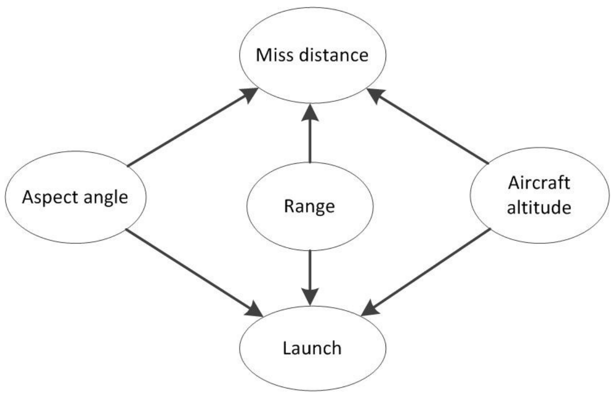

Figure 3. The conceptual model of the Bayesian Network.

What follows is a description of the reasoning for this model structure which can be used as the formal (logical reasoning) validation method for this requirement. Two types of structures are presented in Fig. 4, common cause (left) and common effect (right) (Korb & Nicholson, 2004). The common cause structure has a parent node feeding into the feature nodes where the feature nodes are conditionally independent. This structure is typically used for classification where the aim is to predict certain outcomes and is often referred to as a naïve Bayes structure (De Waal *et al.*, 2016). The common effect structure has feature nodes feeding into a child node where the

[2] Hugin GUI Manual. `http://download.hugin.com/webdocs/manuals/Htmlhelp/`

ISBN 978-0-620-72719-8

feature nodes are conditionally dependant. This type of structure is used for reasoning about variables.

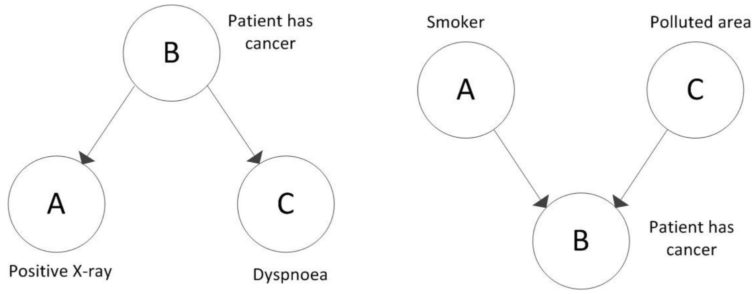

Figure 4. Common cause (left) and common effect (right) structures.

This conditional independence is described by the following example (Korb & Nicholson, 2004). Consider Fig.4 (left): If we don't know anything about B (the patient has cancer), then learning that one symptom is present (A) will increase the chances of cancer which in turn will increase the chances of the other symptom (C). On the other hand, knowing that A has occurred doesn't make any difference to our beliefs about C *if we already know* that the patient has cancer (B). The nodes, A and C, are therefore conditionally independent.

Consider Fig.4 (right): If we observe the effect, B (that the patient has cancer), and we then find out that one of the causes, A, is absent (the patient does not smoke), it raises the probability of the other cause, C (that he lives in a polluted area). In general, the parent nodes are not dependant, but *once we know* that the patient has cancer (B), the parent nodes become dependent, i.e. they are conditionally dependent.

The nature of the model is that it will contain data which are deterministic. It will be used as a lookup table to reason about variable configurations. It is for this reason that the structure chosen for the model under consideration is that of a common effects structure.

3.2.2 Parameterisation

Requirement – Expert data shall be captured from domain experts to determine the probability distribution for the *launch* variable.

Sources for probability parameters include data, domain experts and literature (Korb & Nicholson, 2004). For this model, the expert data is captured during work sessions with domain experts. For each set of variable configurations, a probability is assigned, which results in a Conditional Probability Table (CPT) for the launch variable. These probabilities are captured in a spreadsheet. By acquiring the data through expert knowledge elicitation, uncertainty is introduced in the data. One problem is that people are biased in estimating probabilities. This could be due to factors such as overconfidence, anchoring and availability. Another problem is inconsistent population of large CPTs as described in De Waal *et al.* (2016). One way of handling this uncertainty is to use the visual elicitation technique.

The spreadsheet is converted into a 3D graphical representation and displayed to the experts for further input, similar to the graphs given by Fig.5. The underlying principle is that the expert's cognitive skills are more attuned towards interpretation of a 3D colour image than a 2D numerical table and therefore the uncertainty in the data would be reduced. This serves as validation of the data through the informal: domain expert method.

The information in the spreadsheet is now converted into a usable .txt format that can be read by the software. For this conversion a python script is used. This script is validated by manually inspecting whether the information from the input file is correctly represented by the output file and serves as validation through the static: inspection method.

ISBN 978-0-620-72719-8

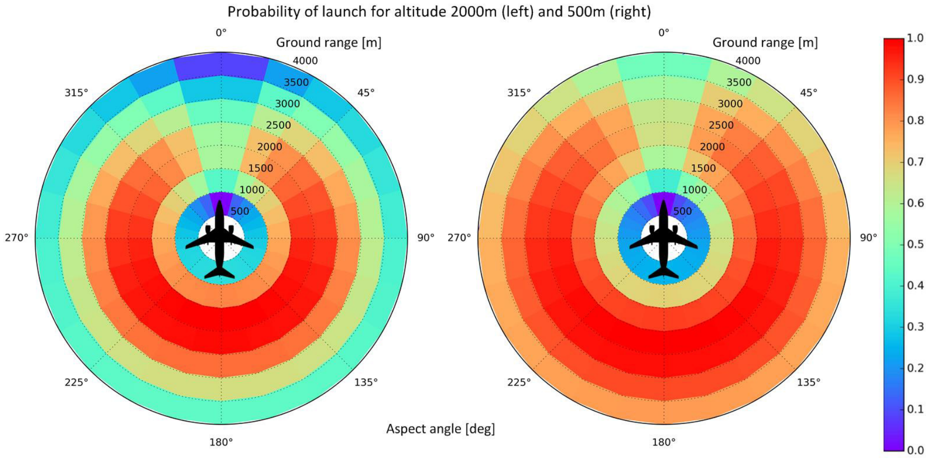

Figure 5. Example of a 3D graphical representation of the probability distribution of the *launch* variable[3].

Requirement – Simulation data shall be used to train the model to create the probability distribution for the *miss distance* variable.

The miss distance data is captured from simulations with CmSim2. The validation method is therefore static: historical validation. These results are converted into a usable .txt format that can be read by the software. For this conversion a python script is used. This script is validated by manually inspecting whether the information from the input file is correctly represented by the output file and serves as validation through the static: inspection method.

The assumption is made that no prior information is known for the input variables and therefore a uniform distribution is adopted for each of these variables. The probability distribution requirements and descriptions are given by Table 3.

Table 3: Probability distribution requirements and description for each node in the model.

Requirement – the probability distribution shall be defined as given by this table.		
Variable	**Probability Description**	**Probability Distribution**
Launch	P (launch I angle, range, altitude) The chances that a ground missile is launched, given evidence of the input variables.	Captured by the expert knowledge spreadsheet.
Miss distance	P (miss distance I angle, range, altitude) The chances the miss distance would fall within a specified state, given evidence of the input variables.	Determined by the software during training with the simulation data.
Aspect angle	P(angle) The chances that the aspect angle falls within a specified state.	Uniform distribution
Range	P(range) The chances that the range falls within a specified state.	Uniform distribution
Aircraft altitude	P(altitude) The chances that the aircraft altitude falls within a specified state.	Uniform distribution

[3] Note that for this specific model example, the altitude variable is defined in meters and not in feet as per Table 2.

ISBN 978-0-620-72719-8

3.2.3 Inferencing

Requirement- The model shall be used in different modes including predictive, prescriptive and diagnostic mode.

It is important to understand and define how the user expects to operate the model so that this can be accommodated in the design of the model. This requirement is validated since the use of a BN model will allow these inferencing modes.

3.3 Develop a Computerized model

This section describes how the Computerized Model is realized though the Computer Programming and Implementation process (see Fig.2). Each part of the model (each V&V criteria) that is implemented in the software is verified against the Conceptual model, or in other words the system is verified against the system requirements (Fig.1). The software used to create the model is *Hugin®*[4] which is software that is specially developed for creating Bayesian Networks.

3.3.1 Structure

The conceptual model is created in the software using the Graphical User Interface (GUI), which is identical to Fig.3. This is manually inspected to be representative of the conceptual model which acts as validation through the static: inspection method. The variables and states are entered into the software according to the requirements defined during the analysis and modelling process. This is visually inspected to be correct which acts as validation through the static: inspection method.

3.3.2 Parameterisation

For the expert data, the prepared .txt file is imported into the software as a CPT. This is manually inspected to be correct and acts as validation through the static: inspection method. For the simulation data, the prepared .txt file is imported into the software and an Expectation-Maximization (EM) algorithm is used to train the model with this data (Korb & Nicholson, 2004). The way in which the training works for this specific type of common effect structure with deterministic data is that it is simply counting the number of occurrences for each configuration of variables. To check that the software is producing the CPT expected, the data can be manually counted and compared with the results which then act as formal validation method: calculation. After data is captured in the software model, it is inspected to have uniform distribution for each of the input variables as defined per requirements in Table 3. This serves as validation through static: inspection method.

3.3.3 Predictive accuracy

The model has now been created and can be evaluated through inferencing which is achieved by exercising the model in different ways. During inferencing, the posterior probability distribution is computed for a query node, given values for some evidence nodes. The query node can be any one of the nodes as information flow is not limited to the direction of arcs in the model (Korb & Nicholson, 2004). One type of evaluation technique called predictive accuracy is now performed, i.e. the model is evaluated for correct inferencing in predictive mode (Korb & Nicholson, 2004). First consider the launch variable as the query node. Set the input variables to different configurations and each time check if the launch probability corresponds

[4] Hugin. http://www.hugin.com/

ISBN 978-0-620-72719-8

to the values expected as given by the expert knowledge spreadsheet. This serves as validation in the form of the dynamic: what-if analysis method. Repeat the exercise with the miss distance as the query node. This time, check if the miss distance probability corresponds to the values as expected as per manually counted simulation data. This serves as validation in the form of the dynamic: what-if analysis method.

3.4 Exercise the model

This section describes how the Computerized Model is validated to be a suitable representative for its intended use in its intended environment (Reality) through the Simulation and Experimentation process (Fig.2). In other words, the final product is validated against the user requirements (Fig. 1). To evaluate the correctness of the model as it is used during inferencing, a case-base evaluation is considered (De Waal *et al.*, 2016).

3.4.1 Case-based evaluation

The model is checked for various cases in predictive, prescriptive and diagnostic mode with a domain expert, similar to the tables given by Fig.6 (Willers *et al.*, 2014). These cases should also include the boundary conditions where the model is tested at the extreme values (Korb & Nicholson, 2004). For each case it should be noted whether the results are as expected and the confidence level in accuracy of the results.

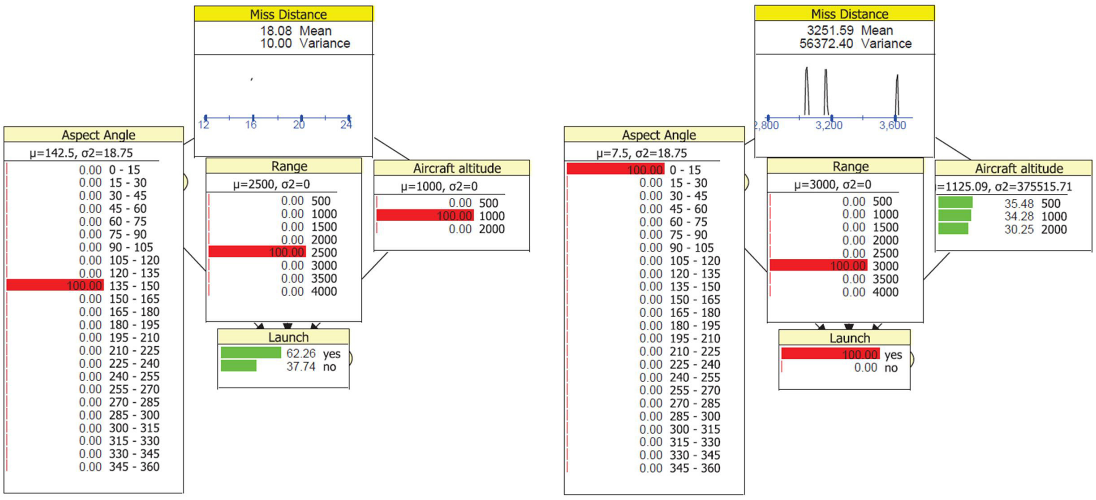

Figure 6. Example of the vulnerability model BN in predictive mode (left) and diagnostic mode (right).

Any deficiencies picked up in the model should be looked into and corrected where possible or the limitations in the model properly captured. The domain expert can now confirm whether in their opinion the model meets the user requirement for its intended use in its intended environment (3.1). This acts as validation through two methods, i.e. informal: domain expert and dynamic: what-if analysis. After the model is accepted by the domain expert, the model is presented to the user for acceptance. The end user can now confirm whether in their opinion the model meets the original requirement for its intended use as in its intended environment (3.1). This acts as validation through dynamic: what-if analysis method. This concludes the full cycle of the V&V procedure.

ISBN 978-0-620-72719-8

4. Model assessment

The next step in the VV&A process as proposed by De Waal *et al.* (2013) is to create a framework that evaluates each V&V criteria against each element and process in a 2D matrix by means of a subjective confidence score out of 10. It was however found not to be applicable to use in the same manner for this model due to the way in which the elements were defined. After V&V is performed as described in section 3, the process is complete and there is no need to perform further methods like scoring. Note that one does not normally argue that a model is more valid than another as each has their own context of applicability (Sargent, 2007), (Balci, 1995).

This type of framework however could be useful for comparing different models that address the same larger problem space. For this exercise, the definition for reality can therefore be adapted such that it represents the real world. Instead of the user specifying that the model should be used for initial inspection in a research capacity, it is adapted such that the model should be used as part of actual mission planning. If the current model is to be evaluated against this new reality, it would certainly not score very well as the model lacks many aspects that are representative of the real world. Apart from reality, many other factors could be added to the rows of the matrix depending on how the models are to be compared, i.e. accuracy (which is the result of the V&V effort), software, documentation and usability (Sargent, 1981). Table 4 shows an example of a framework that can be used to compare the various models. More work can be done to establish this framework for model assessment in the future.

Table 4: Model assessment framework.

	Model Structure	Model Parameterisation	Inferencing
Reality (real world)	2	3	2
Accuracy (V&V)	9	9	9
Software	8	7	8
Documentation	3	3	5
Usability	4	5	4

The resulting scores can be graphically represented with a spider plot such as given by Fig.7 which are useful for visualising when comparing models (Willers & Wheeler, 2007), (De Waal *et al.*, 2013).

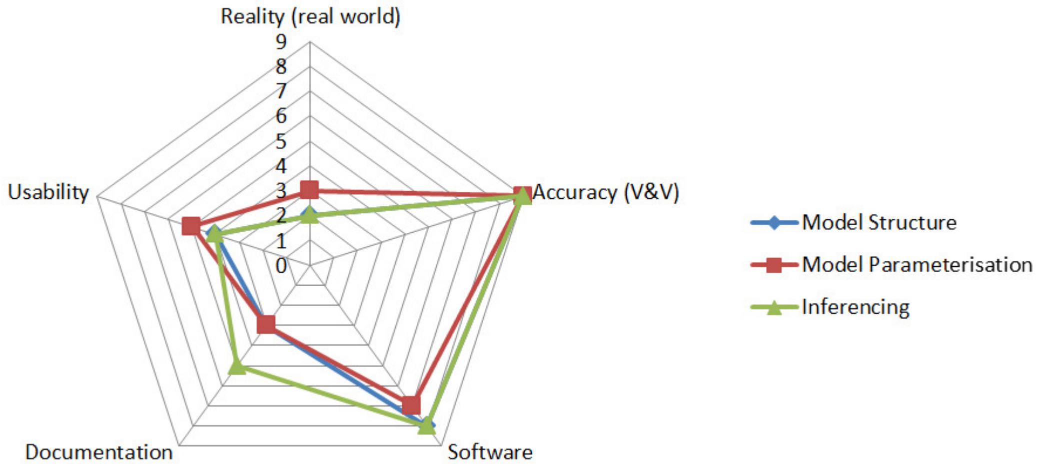

Figure 7. Example of a spider plot for visualisation.

ISBN 978-0-620-72719-8

5. Conclusion

"…we balance probabilities and choose the most likely. It is the scientific use of the imagination…" - Sherlock Holmes -

A methodology was proposed to perform V&V on a Bayesian Network model used for evaluating aircraft vulnerability against infrared missile threats. This work was based on the VV&A framework proposed by De Waal *et al.* (2013) and a BN "vulnerability model" as presented by Willers *et al.* (2014). Throughout the V&V procedure, similarities were drawn between this VV&A process and the well-known Vee-model.

The paper started with an introduction which gave a short background of the model and described KBSs and BNs in general. The next section talked about the V&V procedure where it introduced the VV&A diagram and the Vee-model. It also defined the V&V criteria, methods and general terminology to be used in the paper. The next section formed the majority of the paper as it progressed through the V&V steps which consisted of defining the reality, developing a conceptual model, developing a computerized model and exercising the model. During each step, V&V was performed and the evidence recorded. The last section described how the model can be assessed and compared visually to other models within the greater problem context. The limitations for the methodology are that it does not consider the model for formal accreditation. It also needs to be applied to an actual BN model to see how it behaves in practice.

In the end, it always comes back the user and whether they are satisfied with the product. The reality dictates the user requirement and in this case reality is not the real world. The object of the V&V procedure is to gather enough evidence in order to prove that the model is sufficiently accurate to be used as stated in its intended environment. It is with this V&V that you build the necessary confidence in the accuracy of the model and also this quality that is imparted to the user.

6. References

Balci, O. 1995. "Principles and Techniques of Simulation Validation, Verification, and Testing." Proceedings of the 1995 Winter Simulation Conference, pp. 147-154.

De Waal, A., Willers, C. J., Roodt, J. H. S., Mudau, A. 2013. "The Validation of an Infrared Simulation System." Proceedings of the 5th INCOSE SA Conference, Pretoria.

De Waal, A., Koen, H., De Villiers, P., Roodt, H., Moorosi, N., Pavlin, G. 2016. "Construction and Evaluation of Bayesian Networks with Expert-Defined Latent Variables." Fusion 2016, Heidelberg, Germany.

Forsberg, K., Mooz, H., 1991. "The Relationship of System Engineering to the Project Cycle". National Council On Systems Engineering (NCOSE).

Hugin. http://www.hugin.com/ (last accessed April 8, 2016).

Hugin GUI Manual. http://download.hugin.com/webdocs/manuals/Htmlhelp/ (last accessed April 13, 2016).

Koen, H., De Villiers, J. P., Pavlin, G., De Waal, A., De Oude, P., Mignet, F. 2014. "A framework for inferring predictive distributions of rhino poaching events through causal modelling." Fusion 2014:1-7.

Korb, K. B., Nicholson, A. E., ed.2004. *Bayesian Artificial Intelligence.* Boca Raton, Chapman & Hall:CRC Press.

Lauritzen, S. L., Spiegelhalter, D. J. 1988. "Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems." Journal of the Royal Statistical Society: Series B (Statistical Methodology), 50(2), pp. 157-224.

M&SCO. 2013. Verification, Validation, & Accreditation (VV&A) Recommended Practices Guide (RPG). Retrieved from U.S. DoD Modelling & Simulation Coordination Office. http://www.msco.mil/VVA_RPG.html (last accessed April 8, 2016).

Pearl, J. 1988. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann.

Sargent, R. G. 1981. “An assessment procedure and a set of criteria for use in the evaluation of computerized models and computer-based modelling tools.” Final Technical Report RADC-TR-80-409, U.S. Air Force.

Sargent, R. G. 2007. “Verification and Validation of Simulation Models”. Proceedings of the 2007 Winter Simulation Conference.

Schlesinger, S., et al. 1979. SCS Technical Committee on Model Credibility: Terminology for Model Credibility. Simulation, Vol.32, No. 3, pp. 103-104.

Walden, D. D., Roedler, G. J., Forsberg, K. J., Hamelin, R. D., Shortell, T. M., ed. 2015. Sy*stems Engineering Handbook: A Guide for System Life Cycle Processes and Activities.* Version 4. San Diego, CA (US): INCOSE.

Willers, C. J., Wheeler, M. S. 2007. “The validation of models in an imaging infrared simulation.” International Microwave and Optoelectronics Conference, Salvador, Brazil.

Willers, C. J., Willers, M. S. 2011. “OSSIM: Optronics Scene Simulator White Paper.” Technical Report 6700-PG-103400-01 Rev 3, Council for Scientific and Industrial Research (CSIR).

Willers, C. J., Willers, M. S., De Waal, A. 2014. “Aircraft Vulnerability Analysis by Modelling and Simulation.” Proceedings of SPIE Vol. 9251.

7. Biography

Sunelle Schietekat holds a Master’s degree in Electronic Engineering and works as an Infrared Electronic Warfare Simulation Analyst within OSS at DPSS, CSIR in Pretoria. Previously, she was employed by Square Kilometre Array (SKA) South Africa as the Qualification and Verification Engineer on the African Very Long Baseline Interferometry (VLBI) Network (AVN) project where she was based at Hartebeesthoek Radio Astronomy Observatory (HartRAO). Sunelle is certified as an Associate Systems Engineering Professional (ASEP) at INCOSE.

Alta de Waal holds a PhD in Engineering Science from the North West University, South Africa. Previously she worked at the CSIR. Alta has 16 years’ experience in the field of decision support model development. Of special focus is the development of Bayesian networks in order to model multidisciplinary scenarios. This involves the facilitation of experts from these disciplines in order to develop the qualitative and quantitative aspects of these models. Furthermore Alta also specialises in the field of text analytics and unsupervised modelling of natural language text.

Kevin G. Gopaul currently fulfills the role of Research Group Leader: Infrared Electronic Warfare within OSS at DPSS, CSIR. After joining the South African Air Force in 1996 he obtained a Mechanical Engineering degree (with choice subjects in Aeronautics) from the University of Pretoria. He was involved in the acquisition of the Hawk aircraft system and played a key role in the establishment of the operational squadron at Air Force Base Makhado. He left the SAAF in 2012 having reached the rank of Lieutenant Colonel with his final post being that of Staff Officer Engineering Support within Directorate Combat Systems at Air Command. He is a professional engineer and holds a Masters degree in Engineering Management.