

# Towards better flood risk management using a Bayesian network approach

Gert Wessels<sup>1\*</sup>, Natasha Botha<sup>2</sup>, Hildegard Koen<sup>3</sup>, Nicolene Botha<sup>1</sup>, and Beatrice van Eden<sup>2</sup>

<sup>1</sup>Operational Intelligence, Next Generation Enterprises and Institutions, CSIR, Pretoria, South Africa

<sup>2</sup>Center for Robotics and Future Production, Manufacturing, CSIR, Pretoria, South Africa

<sup>3</sup>Command and Control and Integrative Systems, Defence and Security, CSIR, Pretoria, South Africa

**Abstract.** After years of drought, the rainy season is always welcomed. Unfortunately, this can also herald widespread flooding which can result in loss of livelihood, property, and human life. In this study a Bayesian network is used to develop a flood prediction model for a Tshwane catchment area prone to flash floods. This causal model was considered due to a shortage of flood data. The developed Bayesian network was evaluated by environmental domain experts and implemented in Python through pyAgrum. Three what-if scenarios are used to verify the model and estimation of probabilities which were based on expert knowledge. The model was then used to predict a low and high rainfall scenario. It was able to predict no flooding events for a low rainfall scenario, and flooding events, especially around the rivers, for a high rainfall scenario. The model therefore behaves as expected.

## 1 Introduction

While drought can have a disastrous effect on a country, to a larger extent, so can flooding as flooding can lead to loss of life, livelihoods, and infrastructure damage. Urban areas of Tshwane regularly experience flooding [1] that leads to the loss of human life.

It is noted that in academia there is no set standard to quantify flood risk and it is open to interpretation. The traditional way in which disaster damage is measured involves examining separately the number of fatalities, injuries, and people affected, and the financial damage that disasters, triggered by natural hazards such as earthquakes or floods, cause. Kourgialas and Karatzas [2] provide a flood management strategy and determine hazardous areas of interest in the region. This is said to be the key to flood preparedness. The Flood Forecasting-Warning System uses data from rainfall, meteorological parameters, and river flow to predict floods [3]. This can be combined with the evacuation of risk areas.

With new sensing technology, some organisations provide inundation maps based on satellite images and remote sensing [4]. These maps give a good indication of the depth of a body of water. This information aids experts in understanding water shortages or flooding

---

\* Corresponding author: [gwessels@csir.co.za](mailto:gwessels@csir.co.za)

and other variables. One such new methodology is the Floodwater Depth Estimation Tool (FwDET) which estimates floodwater depth based on an inundation map and a digital elevation model (DEM). Another open-source toolbox for floodwater mapping is the Flood Mapping Python toolbox (FLOMPY) [5]. This is an easy-to-use Python tool, even for less experienced users. The current iteration of this tool is more accurate for rural areas than urban and is dependent on Sentinel-1 data which is limited to the satellite path. Dealing with geospatial information can be demanding on computing resources but in the last decade processing capabilities have improved significantly. The implementation of traditional probabilistic techniques on geospatial data have become an attainable goal.

In a recent study by Nkwunonwo et al. [6], a review of the literature around current flood modelling techniques was conducted where the authors adopted a model classification scheme. Flood modelling methods can be classified based on spatial extent, dimensionality, or mathematical complexity. The weaknesses and strengths of these methods are highlighted to gain some insight on the methodologies and assumptions on how the applications in developing countries are constrained. The aim was to place emphasis on the status of flood modelling for urban flood risk management in developing countries. The big takeaway was that developing countries are more constrained in the use of some of the approaches due to the limitation of available data. Interestingly, Nkwunonwo [6] state that the lack of data might be due to technology and resources but also because of political influence that sometimes prohibits some of the required data from either being collected or accessed.

A study by Cvetkovic and Martinović [7] reviews innovative approaches to flood risk management which include some of the most prominent approaches in the field. The work is presented in four parts: flood prevention, flood preparedness and mitigation, flood response, and flood recovery. A view of how the flood areas can be modelled and the different factors that impact floods in urban areas and flood management were not included.

Lately, there has been a rise in the use of machine learning (ML) techniques in flood risk management, leading to more data-driven approaches. Most of these models rely on sufficiently large and good quality datasets. Chen *et al.* [8] compares six machine learning models (MLMs) for flood risk assessment. Of the six models, Support Vector Machine (SVM), Random Forest (RF), and Multi-layer Perceptron (MLP) were the most prevalent with the most accurate results. Mosavi [9] provides a summary of ML methods used for flood prediction and highlights four major trends, namely integration of two or more machine learning methods; data decomposition techniques for the purpose of improving the quality of the dataset; an ensemble of methods and using add-on optimiser algorithms to improve the quality of machine learning algorithms. It was concluded that ML will be the cornerstone of advancements in flood mitigation and management for the foreseeable future.

More recently other types of neural networks have also been employed successfully in real-time river flood forecasting [10]. Between SVM, MLP, and Long Short-term Memory (LSTM), the LSTM model was the most accurate and robust. Convolutional neural network-based (CNN) methods can also produce more reliable and practical flood susceptibility maps compared to methods such as SVM [11]. While these ML techniques produce promising results, this is not feasible in a data scarce environment due to the extensive data requirements of ML models.

With probabilistic graphical models (PGM) such as Bayesian networks (BNs), models can be constructed from data, expert knowledge, or both, which is ideal for a data scarce environment [12-14]. Balbi *et al.* [15] use commonly available information to supplement quantitative and semi-quantitative data. Flood risk was successfully modelled in the Sihl valley (Switzerland), including the city of Zurich, using a spatially explicit BN model calibrated on expert opinion. A BN was employed by Joo *et al.* [16] to incorporate weights reflecting the characteristics of six different statistical methods creating the Index for Flood

Risk Assessment (InFRA). In the past, different flood risk methods have been presented to decision-makers, causing confusion, and making it difficult to make informed decisions [16]. Incorporating these characteristics of the different methods into one model ensures a unified representation which minimises confusion. Hosseini *et al.* [17] compares ensemble models of the boosted generalised linear model (GLMBoost), random forest (RF), and Bayesian generalised linear model (BayesGLM) methods. It was found that all the compared models yielded good and close performance.

In South Africa (SA) there is limited data available for flood hazard mapping. Els [18] focuses on determining if flood hazard maps can be created for the areas where there are existing data. To achieve this goal, the information on different flood modelling methodologies, data requirements, and flood hazard mapping for SA were reviewed. Even with the limited availability of data sources, flood modelling could be performed. In later years the Greenbook [19] was established as an online tool to aid local municipalities with flood risk assessments. Due to the size of the area, this map has a limited resolution, but work is still ongoing.

A BN approach allows for model construction to be based on data and/or expert opinion which is ideal for the data scarce environment in South Africa. For this paper, we identified a Tshwane suburb that is notorious for experiencing flash floods. We proceed to develop a BN with the help of environmental experts and implement this BN to determine the flood risk prediction for this suburb for low and high rainfall. The overall aim of this project is to develop a generalised model of South Africa, with the assistance of domain experts, that requires only the most basic data inputs. The model developed in this paper is a first step towards better flood mitigation and management through an early warning system, leveraging domain experts and open-source datasets.

## 2 Methodology

In this section we discuss the design and development of the BN as well as the identification of data and variables that will be considered in the model. The generation of the topographical map is also discussed and will be used to visualise the prediction obtained from the Bayesian model. This workflow is illustrated in Figure 1 in three distinct phases namely, retrieving data, data processing, and data visualisation. The activities flow is illustrated by black arrows.

The geospatial data, as well as the results from the study, are visualised using a topographical map with a 20 m resolution of the Hennops river catchment area in the city of Tshwane, Pretoria. The map is created using Python and includes shape files for the rivers and main roads in the area of interest. The study's results are visualised using the retrieved data and processed data from the BN model as described in the workflow in Figure 1. The maps are used to visualise land cover classification, topography, and the BN results (e.g., probability of a flooding event) in such a way that infrastructure, cities, and landmarks can be used to orient the reader.

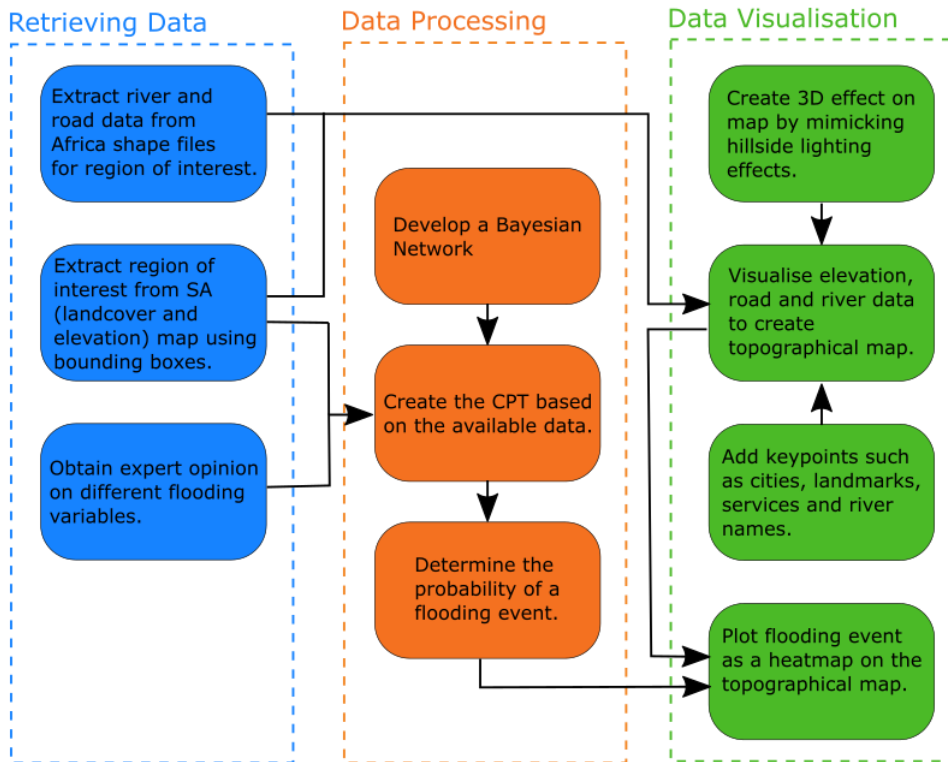


Fig. 1. Workflow considered in the development of the flood risk map.

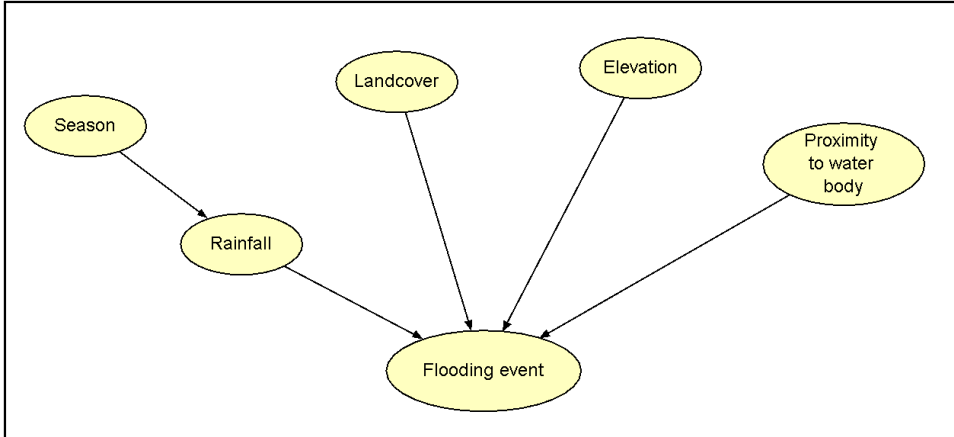
## 2.1 Developing the Bayesian Network

BNs are graphical models consisting of nodes and edges. The nodes represent the variables of the problems, and the edges correspond to the causal relationships between the variables. BNs are DAGs (Directed Acyclic Graphs) because the edges have direction and there are no cycles or loops present in the network. BNs are very flexible and elegant at handling new information, and they can use empirical data, expert knowledge, or a combination of both. Owing to a lack of available flooding data for our area of interest, an expert-driven approach was used.

While experts have a lot of knowledge, it is difficult to translate that into tangible information or data. The difficulty with using expert knowledge is that two people might not have the same opinions or remember the same event differently. They might also influence one another (for example, at a workshop), thereby introducing bias. Most people also find it difficult to quantify their knowledge in terms of probabilities [20]. The traditional way of extracting expert knowledge is through workshops or workgroups. A “clean slate” approach can be used where the participants go into the workshop with a blank canvas and develop the model from scratch. In this case, however, a different approach was used. Owing to time and financial constraints, a first draft model was developed by the researchers themselves after which the model was shown to environmental research experts. One of the pros of going into a workshop or expert elicitation session with a first draft model, is that it is easier to point out mistakes or possible additions to a model than it is to develop one from scratch [20]. A

mixture of expert knowledge and geographical information system (GIS) data is thereafter used to populate the Conditional Probability Tables (CPTs) of the network.

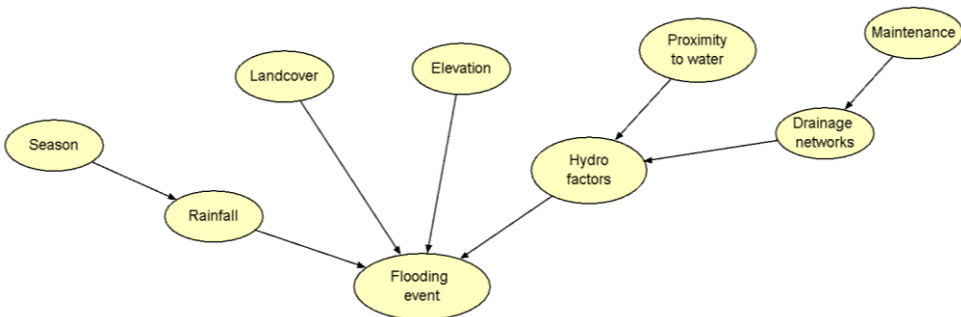
The first draft BN that was developed by the researchers is shown in Figure 2. It is a very basic model of a very complex problem, but it serves to start and facilitate discussions around the problem of flood mapping.



**Fig. 2.** First draft BN developed by researchers

After evaluation by the environmental domain experts, the BN underwent several adjustments, and the current network can be seen in Figure 3. A variable for historic flooding was added in a previous iteration of the network but owing to not being able to find suitable data for that, it was omitted from this version of the network. The assumption is that floods tend to occur in the same areas. The number of historic flooding events in a specific area thus increases the probability of a future flooding in that same area.

Nodes for drainage networks and the maintenance performed on the drainage networks were added to the model. A hidden variable named “Hydro factors” was added to group the nodes together pertaining to hydrological factors, and to ensure that the ensuing CPTs do not get too large (the more directly connected nodes and states there are, the bigger the CPT to populate).



**Fig. 3.** Graphical representation of the current BN after evaluation by experts

## 2.2 Flood related variables

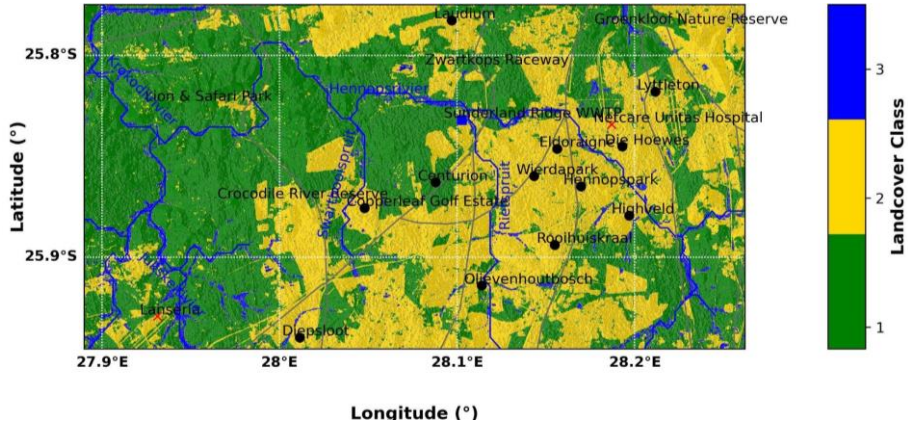
Variables that may have an influence on the probability of an area to undergo a flooding event (listed in Table 1) were identified through a literature study and by consulting domain experts. The variables that were included in the BN (Figure 3) are discussed in this section.

**Table 1.** Variables and descriptions

Node name	States	Description	Data
Season	Spring, summer, autumn, winter	Season directly influences rainfall. Southern parts of South Africa have winter rains whereas Centurion and surrounds have summer rains.	As defined by the date in question, e.g., 01 September - 30 November = Spring.
Rainfall	Low (0-54 mm per month), Medium (55-80 mm per month), High (more than 80 mm per month).	Rainfall is an important driver for flooding.	No empirical data was used, rainfall maps for SA were consulted.
Landcover	Prone to flooding, not prone to flooding, water	Floods directly influenced by the type of landcover. Assumption is that built-up areas are more prone to flooding, and agricultural and open areas are less prone to flooding. Areas containing water are the most prone to flooding. Landcover is illustrated in Figure 4 showing the initial and simplified datasets.	South African National Land Cover (SANLC) 2020 map. Department of Environment, Forestry & Fisheries (DEFF).
Elevation	Low (1100-1300 m), Medium (1301-1500 m), High (1501+ m)	Important variable when considered alongside other variables. For instance, low elevation coupled with heavy rains. Elevation is illustrated in Figure 5. Note that elevation refers to the height above mean sea level.	30 m Shuttle Radar Topography (SRTM) DEM. NASA's Shuttle Radar Topography Mission (SRTM) Land Elevation Dataset.
Proximity to water	Close (0-40 m), Medium (41-80 m), Far (80+ m)	The closer an area is to a water body, the higher its risk for flooding. The states were chosen such that it equates to two (0-40m), four (41-80m), and more than four (80+m) cells on our grid.	Proximity to water pixel - calculated from the land cover map. Department of Environment, Forestry & Fisheries (DEFF).
Maintenance	Less than 6 months ago, more than 6 months ago	The drainage network has a dependency on the regularity of maintenance, e.g., how regularly the drains are cleared of debris.	No information was available for this variable.



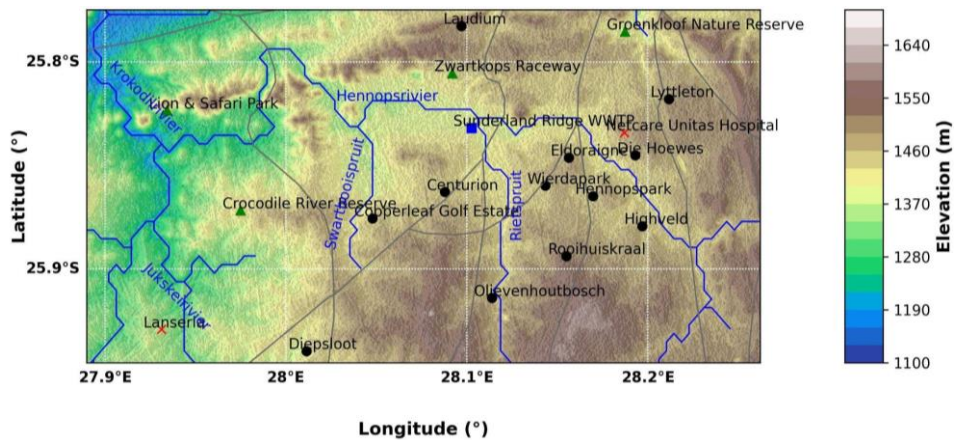




(b) Simplified data (3 classes)

**Fig. 4.** Land Cover map for Centurion (a) original from DAFF and (b) simplified for model with all vegetation/open land (prone to flooding) set to 1, build-up (not prone to flooding) set to 2 and waterbodies (water) set to 3.

The elevation is displayed on a topographical map, shown in Figure 5 below, highlighting the lower elevations towards the upper left of the map and higher elevations towards the bottom right.



**Fig. 5.** Topographical Map of the Hennops River Catchment Area

### 2.3 Implementing the BN

The BN from Figure 3 is implemented in Python using the pyAgrum [21] module. pyAgrum is a C++ and Python library that can be used to develop probabilistic graphical models such as BNs and perform relevant computations.

The aim of the model is to infer whether a flooding event will occur. The Centurion area is divided into a grid of 20 m by 20 m cells which correspond to the granularity of the Land Cover map that was used. The BN's structure (nodes and edges) is created in pyAgrum as



well as the CPTs. The probabilities used to populate the CPTs are elicited from experts and the probabilities of the root nodes are mostly calculated from real-world data. For instance, you have an equiprobable chance of being in a certain season at any given time. The probabilities for Land Cover, Elevation, and Proximity to water were calculated by counting the number of cells that fall into each state and dividing that by the total number of cells. There was no information available for Maintenance, thus an arbitrary number was used. Figure 6 shows a snapshot of the target variable (flooding event) CPT that illustrates some of the states and their corresponding probabilities calculated with Bayes’ theorem.

The BN is executed for each of the cells in the grid. Thus, a probability for Flooding event is obtained for each of the cells. This probability is used against a colour map to determine where it lies on the heatmap.

				FloodingEvent	
Rainfall	Landcover	Elevation	HydroFactors	True	False
Low	Prone_to_flooding	Low_1100_1300m	Conductive_to_flooding	0.25	0.75
			Not_conductive_to_flooding	0.10	0.90
		Medium_1301_1500m	Conductive_to_flooding	0.15	0.85
			Not_conductive_to_flooding	0.06	0.94
		High_1501m	Conductive_to_flooding	0.07	0.93
			Not_conductive_to_flooding	0.03	0.97
	Not_prone_to_flooding	Low_1100_1300m	Conductive_to_flooding	0.16	0.84
			Not_conductive_to_flooding	0.06	0.94
		Medium_1301_1500m	Conductive_to_flooding	0.05	0.95
			Not_conductive_to_flooding	0.02	0.98
		High_1501m	Conductive_to_flooding	0.01	0.99
			Not_conductive_to_flooding	0.01	0.99

**Fig. 6.** Snapshot of target variable (Flooding event) CPT showing some of the states and corresponding probabilities, calculated with Bayes’ theorem

### 3 Results and Discussion

Once the model has been constructed, the input data supplied, and the CPTs populated, the flooding event probability can be inferred from the BN model. A what-if analysis is considered to test the model viability. The model can be used to predict the probability for flooding for a given date and rainfall, in the area of interest. The probability of a flooding event to occur is visually illustrated in the form of a heatmap.

As mentioned, the area of interest is divided into 20 m x 20 m cells and the probabilities for the applicable variable states for each individual pixel is used to calculate the probability for the pixel to flood, using the constructed model, resulting in a flooding probability heatmap.

#### 3.1 What-if analysis

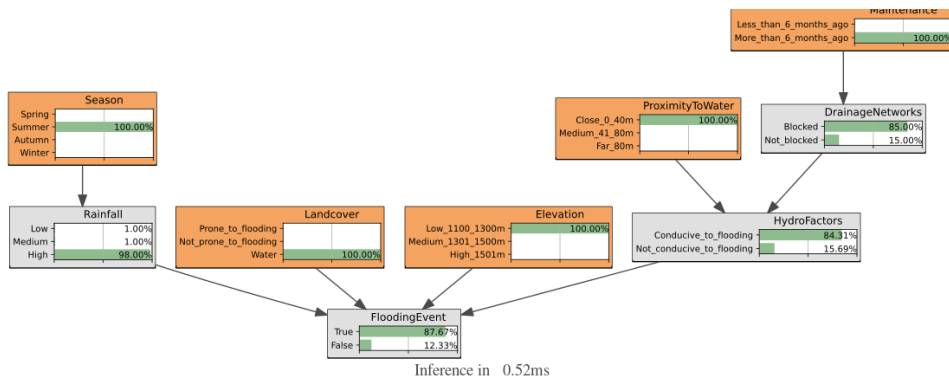
To test the validity of a developed BN model, three what-if analyses are considered. The first is a best-case scenario for a flood to occur, which is the ideal conditions that allow for a

flooding event. The second is a worst-case scenario for a flood event to occur, which is when a flooding event does not (or should not) occur. Both the best- and worst-case scenario conditions are listed in Table 2, where the states for these conditions were provided by domain experts. Lastly, a bottom-up scenario is shown where it is inferred what the conditions would need to be for a flooding event to occur. That is, only the flooding event variable state is set to True, no other variables are predefined.

**Table 2.** States assigned to the nodes for the best- and worst-case scenarios to have a flooding event.

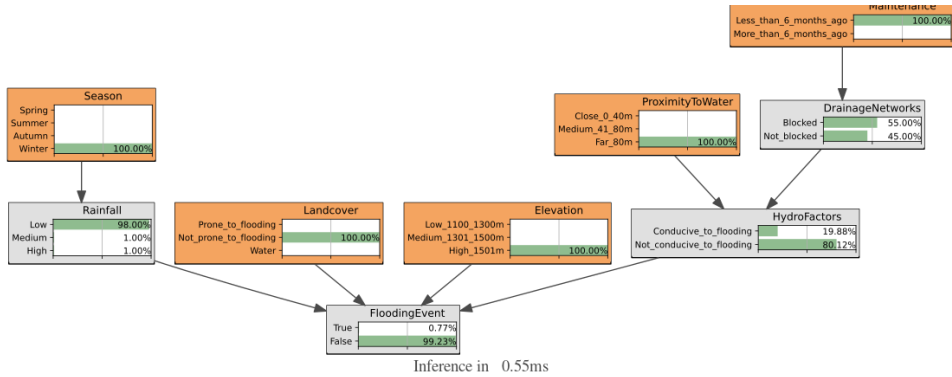
Parameters	Best Case	Worst Case
Season	Summer	Winter
Land cover	Water	Not prone to flooding
Elevation	Low (1100-1300)	High (1501+)
Proximity to water	Close (0-40m)	Far (80m+)
Maintenance	More than 6 months ago	Less than 6 months ago

The first scenario tested is the best-case scenario, where the parameters for a pixel are set to the values in the second column in Table 2. Top-down inference is used to calculate the flooding probability for the pixel using Bayes' theorem and the result can be seen in Figure 7. The probability for a flooding event to occur under these conditions is 87.67%. The high probability of a flooding event occurring corresponds to our intuition of the outcome of the target variables given the factors conducive to a flooding event.



**Fig. 7.** Inference diagram with CPT for the best-case scenario

Similarly, the values in the third column of Table 2 are used as the input for a pixel to test the model's behaviour for a worst-case scenario as can be seen in Figure 8. The probability for a flooding event to occur under these conditions is 0.77%. This low probability corresponds to our intuition of what the outcome should be when the chosen features are not conducive to a flooding event. In both cases the model behaved as expected and gave a low flooding probability for the pixel in the worst case and significantly higher flooding probability for the pixel set to the best-case scenario.

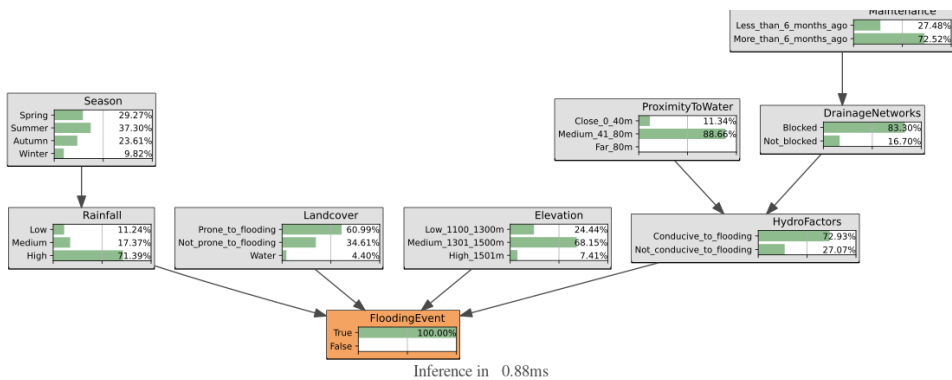


**Fig. 8.** Inference diagram with CPT for worst case scenario

For a third what-if analysis, a bottom-up inference is used to identify the optimal factors for a flooding event to occur, i.e., what conditions should be in place for a pixel to have a 100% probability of flooding. This is done by setting the flooding event state to True, as shown in Figure 9. From this analysis the parameter ranges that would result in a flooding event are laid out. Clearly some of the parameters, such as Season, have a less distinct influence compared to Proximity to water, and Drainage networks.

It is interesting to note that the states within Season contributing to a flooding event are not as clear-cut as we would expect. The probability values are close together (0.2927, 0.3730, 0.2361, and then 0.0982) which could be because the marginal probabilities for Season was set to 0.25 as you have an equiprobable chance of being in any season at a random time. Looking at the individual probabilities we can see that summer still has the highest probability of having high rainfall (and thus contributing to a flooding event), then spring, then autumn, and then lastly winter (with by far the lowest probability). For the Land Cover node the “Prone to flooding” state has the highest probability but the “Not prone to flooding” is also quite high. This is because there are a lot of cells that fall under the “Not prone to flooding” state. The same holds true for Elevation and Proximity to water.

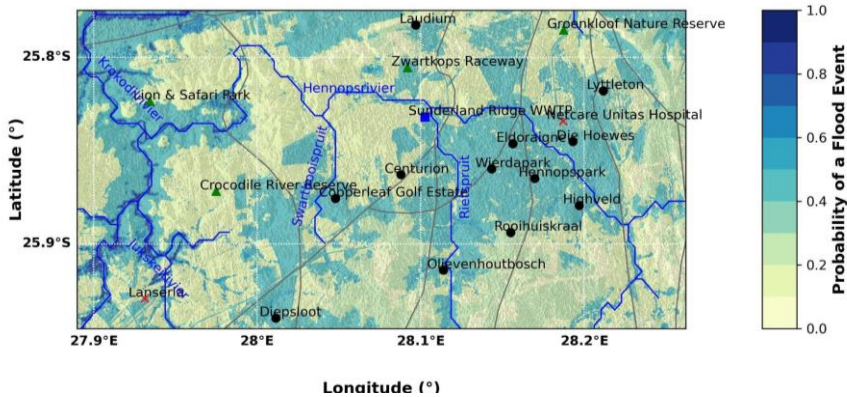
Based on these three analyses, the developed BN model performs as expected based on intuition. The model is therefore validated and can be used to predict the occurrence of a flood event for a defined set of conditions.



**Fig. 9.** Inference diagram with CPT for bottom-up scenario

### 3.2 Flooding event predictions

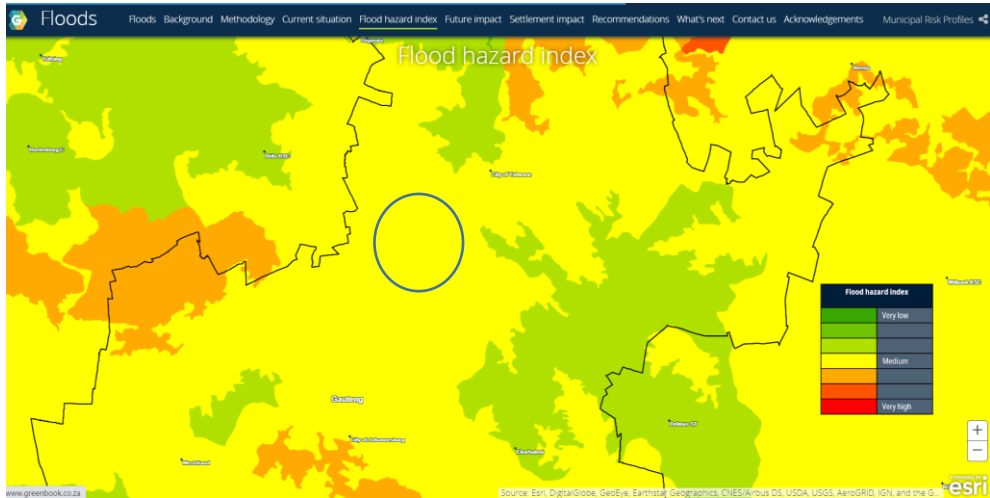
After the model has been successfully validated, it can be used to predict the probability of a flooding event. Two scenarios are considered: one where there is high rainfall with maintenance occurring more than 6 months ago as shown in Figure 10 and one with low rainfall and maintenance occurring less than 6 months ago, see Figure 12. The darker the shade of blue on the map, the higher the probability of a possible flooding event. The lighter and more yellow the colour, the lower the probability of a possible flooding event.



**Fig. 10.** Flood Map of the Hennops River Catchment Area for high rainfall and maintenance more than 6 months ago.

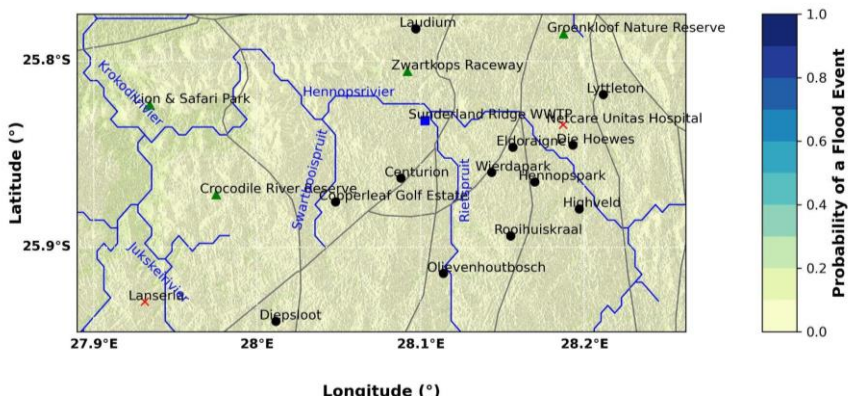
Generally, a high rainfall scenario will result in several flooding events as has been the case in past historical floods in the Hennops river catchment area. From Figure 10 the probability of flooding occurring around the rivers is quite high with some flooding predicted to occur in lower to medium elevation areas as well (see area to the left of Centurion). The highest probabilities correspond to areas along the Krokodil river and the Jukskei river which both have a low to medium elevation and is thus more prone to flooding than areas with higher elevation such as the areas in the middle and to the right of the figure.

When comparing the flooding scenario with the flood risk map of the GreenBook shown in Figure 11 certain similarities are evident. The GreenBook shows a medium flood risk for the whole Centurion area, while our flood map shows a 60% chance of flooding in the flood prone areas, which can also be considered as a medium flood risk. It should be noted that the GreenBook has a coarser resolution and uses different metrics to calculate the flooding risk, and so cannot be compared directly to our flood map (which has as previously mentioned a resolution of 20 m x 20 m).



**Fig. 11.** Flood risk map of Gauteng extracted from the GreenBook [19]. Approximate location of Hennops river catchment indicated by the blue circle.

It is expected that a low rainfall scenario would result in the occurrence of no flooding events across the whole area. This is corroborated in Figure 12, where we see that the probability of flooding events occurring are very low across the entire area.



**Fig. 12.** Flood Map of the Hennops River Catchment Area for low rainfall and maintenance less than 6 months ago.

## 4 Conclusion

This study aims to develop a first approximation model for flash flood prediction, considering a catchment area in the city of Tshwane as a case study. Obtaining flood related data was challenging, for this reason, a BN was developed that was populated with expert knowledge and available data to predict possible areas of flooding. The consulted domain experts, who contributed to the GreenBook, were very helpful in providing information on what variables to include for flood estimates, but it was not possible to obtain absolute flooding probabilities for each variable from them due to the complexity of the relationships.

Three what-if scenarios were considered to test the validity of the developed BN. For the best- and worst-case scenarios for a flood event to either occur or not occur, the model behaved as expected. The bottom-up scenario indicated that for a flooding event to occur the rainfall should be high, drainage networks should be blocked, and therefore a lack of proper maintenance needs to be in place, and hydrological factors should be conducive to flooding.

Two prediction scenarios were considered where two nodes were defined, low rainfall with maintenance less than 6 months ago, and high rainfall with maintenance more than 6 months ago. The results were as expected, where the low rainfall scenario resulted in no flooding events occurring, and a high rainfall scenario predicted flooding events, especially around the rivers and low elevation areas.

This paper serves as a first phase in this flood management project. Future phases will include refinement of the variables and states, as well as the incorporation of available flooding data. It is envisioned that a generalised model for South Africa, once fully developed, might assist in flood mitigation and management leading to proactive rather than only reactive disaster management.

## References

1. Eyewitness News, <https://ewn.co.za/2022/02/05/regular-floods-have-devastating-effect-on-livelihoods-centurion-residents#:~:text=livelihoods%20%3A%20Centurion%20residents-,Regular%20floods%20have%20devastating%20effect%20on%20livelihoods%20%3A%20Centurion%20residents,Hennops%20River%20overflowing%20in%20Centurion.> Last accessed on 19-09-2022.
2. N.N. Kourgialas, G.P. Karatzas, *Flood management and a GIS modelling method to assess flood-hazard areas—a case study*, Hydrological Sciences Journal—Journal des Sciences Hydrologiques, 56(2), pp.212-225 (2011)
3. C.H. Green, D.J. Parker, S.M. Tunstall, *Assessment of flood control and management options*, WCD Thematic reviews. World Commission on Dams Secretariat, South Africa (2000)
4. S. Cohen, G.R Brakenridge, A. Kettner, B. Bates, J. Nelson, R. McDonald, Y.F. Huang, D. Munasinghe, J. Zhang, *Estimating floodwater depths from flood inundation maps and topography*, JAWRA Journal of the American Water Resources Association, 54(4), pp.847-858 (2018)
5. K. Karamvavis, V. Karathanassi, *FLOMPY: An Open-Source Toolbox for Floodwater Mapping Using Sentinel-1 Intensity Time Series*, Water, 13, 2943 (2021). Available at <https://github.com/kleok/FLOMPY>.
6. U.C Nkwunonwo, M. Whitworth, B. Baily, *A review of the current status of flood modelling for urban flood risk management in the developing countries*, Scientific African 7, e00269 (2020)
7. V.M. Cvetkovic, J. Martinović, *Innovative solutions for flood risk management*, International Journal of Disaster Risk Management, 2(2), pp.71-100 (2020)
8. J. Chen, G. Huang, W. Chen, *Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models*, Journal of environmental management, 293, p.112810 (2021)
9. A. Mosavi, P. Ozturk, K.W. Chau, *Flood prediction using machine learning models: Literature review*, Water, 10(11), p.1536 (2018)

10. S. Dazzi, R. Vacondio, P. Mignosa, *Flood Stage Forecasting Using Machine-Learning Methods: A Case Study on the Parma River (Italy)*, *Water*, 13, 1612 (2021)
11. Y. Wang, Z. Fang, H. Hong, L. Peng, *Flood susceptibility mapping using convolutional neural network frameworks*, *Journal of Hydrology*, 582, p.124482 (2020)
12. Y. Luo, Z. Dong, X. Guan, Y. Liu, *Flood risk analysis of different climatic phenomena during flood season based on copula-based Bayesian network method: a case study of Taihu Basin, China*, *Water*. 2019 Aug;11(8):1534 (2019)
13. S. Huang, H. Wang, Y. Xu, J. She, J. Huang, *Key Disaster-Causing Factors Chains on Urban Flood Risk Based on Bayesian Network*, *Land*, 2021 Feb;10(2):210 (2021)
14. Z. Sobhaniyeh, M.H. Niksokhan, B. Omidvar, S. Gaskin, *Robust Flood Risk Management Strategies Through Bayesian Estimation and Multi-objective Optimization*, *International Journal of Environmental Research*, 15(6), pp.1057-1070 (2021)
15. S. Balbi, F. Villa, V. Mojtahed, K.T. Hegetschweiler, C. Giupponi, *A spatial Bayesian network model to assess the benefits of early warning for urban flood risk to people*, *Natural Hazards and Earth System Sciences*, 16(6), pp 1323-1337 (2016)
16. H. Joo, C. Choi, J. Kim, D. Kim, S. Kim, H.S. Kim, *A Bayesian network-based integrated for flood risk assessment (InFRA)*, *Sustainability*, 11(13), p.3733 (2019)
17. F.S. Hosseini, B. Choubin, A. Mosavi, N. Nabipour, S. Shamshirband, H. Darabi, A.T. Haghghi, *Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: application of the simulated annealing feature selection method*, *Science of the total environment*, 1 Apr 2020, 711:135161 (2020)
18. Z. Els, *Data availability and requirements for flood hazard mapping in South Africa*, Doctoral dissertation, Stellenbosch University, South Africa (2011)
19. GreenBook (2021) Available at <https://greenbook.co.za/>.
20. H.S. Koen, *Predictive Policing in an Endangered Species Context: Combating Rhino Poaching in the Kruger National Park*. Doctoral Thesis, University of Pretoria, South Africa (2017)
21. aGrUM/pyAgrum (2018) Available at <https://pyagrum.readthedocs.io/en/1.1.0/>.