International Journal of Disaster Studies and Climate Resilience
https://doi.org/10.64866/j.ijdscr.2025.10002

Vol. 1 No. 1 Jan–Jun 2025, 8–21

Bayesian Networks and GIS-based analysis for flood risk assessment in agriculture during heavy rainfall events

Paul Gräfe*, Angela Blanco-Vogt, and Franz-Josef Behr

Faculty of Surveying, Mathematics and Informatics, Stuttgart University of Applied Sciences, Stuttgart, Germany

Abstract

The increasing frequency and intensity of heavy rainfall events due to climate change pose a growing risk to agriculture, and its prediction remains a substantial challenge for the scientific community. Leveraging the availability of high-quality open geospatial data provided by governmental institutions and existing probabilistic risk models, this study proposes a systematic approach for analyzing the risks associated with heavy rainfall events in agriculture. The objective is to develop and test a methodology that integrates open geospatial data, considers uncertainties, and conducts scenario analyses based on the historical heavy rainfall event of June 2, 2024, in the Rems-Murr district, Germany. Initially, variables such as *Rainfall Intensity*, *Temperature*, *Land Use* (from land use maps), *Soil Type*, *Soil Moisture*, *Slope*, *Elevation*, and *River Discharge*—obtained from official institutions—along with variables like *Proximity to River*, *Road Density*, and *Proximity to Forest* (derived from GIS analysis), were conceptually integrated into a Bayesian Network (BN). This integration was based on theoretical foundations and quantified using conditional probability distributions (CPD). The results demonstrate that the methodology combining BN and GIS analyses, along with scenario analysis, sensitivity analysis, optimization, and model validation, was successfully applied to the wider area of the Rems-Murr district. When tested on the heavy rainfall event of June 2, 2024, in the study area of Rudersberg, it provided qualitatively convincing results for flood risk assessment in agriculture. Validation in Rudersberg yielded a Root Mean Square Error (RMSE) of 23%. The methodology was also successfully applied across regions in Miedelsbach, where validation with official data collected by the county resulted in an RMSE of 30%. These findings indicate that the methodology is applicable not only within the study area but also across different regions. It is recommended to improve the model by incorporating additional variables such as surface parameters, roughness values, and drainage systems to improve accuracy. Furthermore, integrating meteorological forecasts could provide a basis for forward-looking risk predictions.

Keywords: Flood risk assessment, Bayesian Networks, GIS analysis, Heavy rainfall events, Agricultural risk modeling, Geospatial data, Scenario analysis, Disaster risk management, Extreme weather events, Machine learning in hydrology, Climate change adaptation

Received: March 31, 2025
Revised: April 10, 2025
Accepted: April 11, 2025

Citation
Gräfe, P., Blanco-Vogt, A., and Behr, F.-J. (2025). Bayesian Networks and GIS-Based Analysis for Flood Risk Assessment in Agriculture During Heavy Rainfall Events. *Int. J. Disast. Studies Climate Resil.* 1(1):8–21.

CC BY SA

1 Introduction

In recent years, the frequency of heavy rainfall events has increased globally, largely due to climate change and the associated rise in extreme weather events (Lehmann et al., 2015). The German Weather Service (DWD) and the Federal Environment Agency (UBA) predict further increases in precipitation levels (Becker et al., 2016; Wilke, 2024). Similarly, data from the international disaster database EM-DAT show a growing number of extreme weather events, including heavy rainfall, in Central Europe (CRED, 2025).

Risk analysis identifies, assesses, and prioritizes risks to support decision-making (Aven, 2015). It typically involves risk identification – recognizing risk components, and risk assessment – evaluating probabilities and impacts. Measures for risk mitigation follow the analysis. Methods range

*Corresponding author. Email: 12grpa1bvg@hft-stuttgart.de (PG), angela.blanco-vogt@hft-stuttgart.de (AB-V), franz-josef.behr@hft-stuttgart.de (F-JB)

ISSN: 3107-605X (Online)

from qualitative (e.g., expert surveys) to quantitative (e.g., Bayesian Networks for flood risk modeling).

Floods caused by heavy rainfall pose a major risk to agriculture, leading to soil oversaturation and erosion, which can significantly impact crop yields (Li et al., 2019; Muthiah et al., 2025). This underlines the need for robust risk assessment models. However, approaches that integrate uncertainties and rely on open geospatial datasets at local level remain scarce.

This study investigates how Bayesian Networks (BN) and Geographic Information System (GIS) analyses can be combined to assess and quantify flood risk in agricultural areas. The research focuses on the heavy rainfall event of June 2024 in the Rems-Murr district (see Section 4.1) and aims to develop a methodology that integrates open geospatial data, accounts for uncertainties, and enables scenario analyses.

To achieve this, BN variables were conceptually defined, and location-specific GIS data were processed for BN configuration.

BN has proven to be effective in risk modeling, particularly in capturing uncertainties in flood events and generating meaningful results even with incomplete data (Kreibich et al., 2018).

While BN and GIS have previously been used for spatial risk analysis (Harris et al., 2022; Lu et al., 2024; Wu et al., 2019), their application in assessing agricultural flood risk due to heavy rainfall remains largely unexplored.

Although conceptual hydrological models are used successfully, they reach their limits when modeling local flash floods and Horton's surface runoff formation. Studies show that the accuracy of the model is strongly dependent on data availability and that more complex models do not necessarily provide better predictions (Fatichi et al., 2016; Hrachowitz and Clark, 2017). Therefore, probabilistic models such as BN offer a promising alternative as they explicitly account for uncertainties and can provide robust results even with limited data (Cooper and Herskovits, 1992). The use of open data ensures cost-efficiency and facilitates broader application, particularly for smaller municipalities in Baden-Württemberg.

The BN was implemented in Python using the 'Probabilistic Graphical Models in Python' (pgmpy) library, which enables causal inference modeling through directed acyclic graphs (Ankan and Panda, 2015). Data preparation and result visualization were performed using QGIS.

2 Theoretical background

The section begins with an overview of risk analysis concepts, focusing on its objectives and methods. It explains the components of heavy rainfall and its associated risks to agriculture. The section lays the foundation for the investigation by detailing the characteristics, classification, and relationships of heavy rain and flood events. It then explores how agricultural risk is composed and the key factors influencing it. The theory behind the BN is introduced, followed by a discussion on sensitivity analysis related to BN. The final section covers GIS basics and the use of open geospatial data in the study.

2.1 Heavy Rainfall and Flooding

This section distinguishes between surface runoff from heavy rainfall and flooding caused by rising water levels in rivers and streams.

Heavy rainfall events are localized, intense precipitation events that are difficult to predict (DWD, 2024). As noted by Assmann (2023), distinguishing between surface runoff and flooding in small catchment areas is complex due to their gradual transitions. Surface runoff refers to surface water flow, while flooding describes overflow directly from a river or stream.

Definitions of surface runoff vary depending on discipline: some focus on precipitation thresholds (e.g., DWD classification), while others emphasize urban drainage metrics. Process-oriented approaches highlight areas that contribute disproportionately to runoff.

In this study, flood risk primarily refers to surface runoff from heavy rainfall, acknowledging that rising rivers can also result from such events, blurring the distinction between both processes.

The DWD (DWD, 2024) classifies and forecasts heavy rainfall based on intensity and duration:

- **Significant weather warning**: 15–25 l/m² in 1 h or 20–35 l/m² in 6 h.
- **Severe weather warning**: 25–40 l/m² in 1 h or 35–60 l/m² in 6 h.
- **Extreme weather warning**: >40 l/m² in 1 h or >60 l/m² in 6 h.

River flooding, on the contrary, is classified using HQ values (e.g., HQ10, HQ100), which indicate probabilities of recurrence based on hydrological data.

In summary, heavy rainfall events lead to surface runoff, which can transition into flooding when rivers rise. Their distinction is fluid, making classification and assessment challenging.

2.2 Risk Components in Agriculture

As explained in the previous Sections, breaking down risk components is essential for further analysis. This Section examines the key components contributing to agricultural risk assessment.

Agriculture faces particularly high risks from heavy rainfall events, which increase both in frequency and intensity due to climate change. These events can cause severe flooding, threat crop yields and cause long-term soil degradation.

Disaster risk is generally understood as the interaction of three key components: Hazard, vulnerability, and exposure (Rana and Routray, 2017). This concept is also applied in the following risk analysis of agriculture concerning heavy rainfall events.

The hazard refers to the potential event, such as heavy rainfall leading to flooding. Vulnerability describes how susceptible a system or area, such as agricultural land, is to this hazard. Finally, exposure encompasses the degree to which an area is affected by the hazard, for example, due to its geographical location or environmental conditions (Huang et al., 2021).

In conclusion, risk is the intersection of all three key components. The more a region or area (e.g., agriculture) is affected by both a hazard (e.g., heavy rainfall), high vulnerability (e.g., lack of protective measures), and high exposure (e.g., proximity to river), the greater the risk for that area.

As discussed above, risk in agriculture arises from the interplay of three key components: hazard, referring to the potential event such as heavy rainfall; vulnerability, which describes the susceptibility of agricultural land to the hazard; and exposure, which refers to the degree of exposure, for example, due to geographical conditions. The interaction of these components determines how strongly agricultural areas are affected by heavy rainfall events and their consequences, defining the overall risk.

2.3 Bayesian Network and Sensitivity Analysis

A BN is a graphical model representing cause-effect relationships through a directed acyclic graph (DAG) (Pearl, 2014). Nodes represent variables, while edges define causal dependencies, quantified by conditional probability distributions (CPD) (Koski and Noble, 2011). BNs typically use discrete probability states.

The Bayes' theorem governs probability updates in a BN:

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{i=1}^{n} P(A_i)P(B \mid A_i)} \tag{1}$$

where:

- $P(A_i)$ are prior probabilities,
- $P(B \mid A_i)$ represents conditional probabilities,
- $\sum_{i=1}^{n} P(A_i)P(B \mid A_i)$ applies the law of total probability (Huang et al., 2021).

A BN enables probabilistic inference, allowing decision-making even with incomplete data (Cooper and Herskovits, 1992). Its flexibility supports modeling flood risk while accounting for uncertainties (Li et al., 2023).

Sensitivity analysis examines how variations in the input parameters influence BN results (Dinkelbach, 1969). It serves to:

1. Identify key influencing variables,
2. Test model robustness,
3. Optimize variable selection for enhanced accuracy.

These concepts form the theoretical basis for the subsequent methodology and application.

2.4 GIS and Open Geodata

Geographic Information Systems (GIS) are crucial for capturing, analyzing, and visualizing spatial data. They enable the examination of geographical relationships and spatial modeling, including raster/vector processing, buffer creation, and terrain analysis. Open-source tools like QGIS and Python allow efficient spatial analyses (Bill, 2010).

Open geospatial data from platforms such as OpenStreetMap or Copernicus provide freely accessible spatial information. The State Office for Geoinformation and Land Development (LGL) Baden-Württemberg offers official datasets like terrain models and land use maps. While open data enhances accessibility, challenges such as data gaps and accuracy inconsistencies exist, especially in community-driven sources like OpenStreetMap (Weigell, 2023).

This study utilizes GIS analyses and LGL open geospatial data to model spatial variables (e.g., *River Proximity*, *Slope*, *Proximity to Forest*) and classify them into fixed states for risk assessment in agriculture.

3 Methodology

While BN and GIS have previously been used for spatial risk analysis (Harris et al., 2022; Lu et al., 2024; Wu et al., 2019), their application in assessing agricultural flood risk due to heavy rainfall remains largely unexplored. This chapter discusses how to set up and configure the model to use BN and GIS combined for flood risk quantification.

At the beginning, the workflow is presented, forming the core of this study. It integrates the two central pillars: GIS analysis and BN, demonstrating the connection between probabilistic modeling and spatial data analysis, showing how the investigation's steps interlink.

For assessing agricultural risk due to heavy rainfall, the BN methodology first requires model configuration, including defining the network's structure and representing it through a conceptual diagram (Furlan et al., 2020). This begins with identifying the key variables and risk factors needed to simulate heavy rainfall risk.

Next, data sources for the BN's key parameters are selected, ensuring they are standardized and accessible for cross-territorial use.

In addition, GIS analysis is used to process the data and prepare it for BN analysis. Due to the varying scales of the data, classification and rescaling are necessary (Huang et al., 2021).

The next step involves configuring the probability distributions within the BN to establish relationships and conditions. The final stage focuses on advanced analyses like scenario analysis, optimization, and BN validation, which are integral part to this methodology.

3.1 Connection Between Bayesian Networks and GIS Analyses

The integration of BN and GIS enables a structured approach to assessing agricultural flood risk from heavy rainfall. Fig. 1 illustrates the methodological workflow, highlighting the interaction between GIS analysis and probabilistic modeling.

The process begins with the definition of key risk variables, which form the basis for GIS analysis and BN modeling. The GIS block involves selecting and processing spatial data, classifying variables into discrete states, and visualizing cartographic results. This processed data feeds into the BN block, where a conceptual model is structured, probabilities are assigned using expert knowledge and empiri-

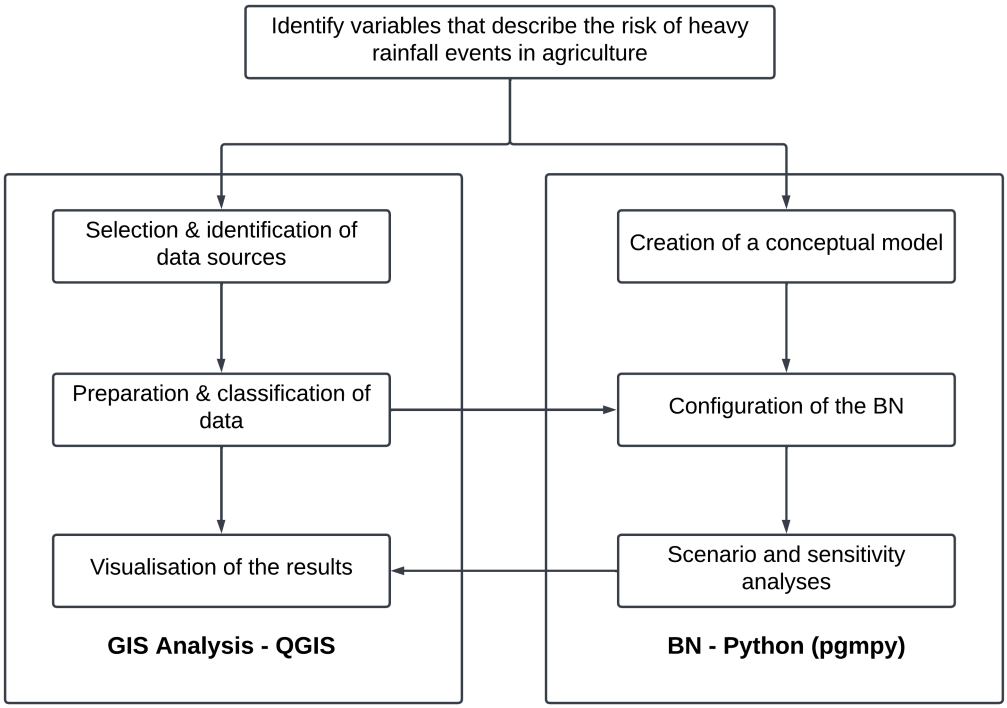

Fig. 1. Methodical workflow (modified after Abebe et al. (2018)).

cal data, and sensitivity analysis is performed to refine the model.

GIS and BN are interdependent: GIS analysis provides essential input for the BN configuration, while BN results can be spatially visualized and validated using GIS. This iterative workflow ensures a comprehensive assessment of flood risk.

3.2 Variables and Model Structure

When modeling a BN to assess flood risk for agricultural areas due to heavy rainfall, it is important to select variables that precisely represent the three core components – hazard, vulnerability, and exposure – while reflecting region-specific characteristics.

For the hazard component, three crucial variables are identified: *Rainfall Intensity*, which is a key characteristic for distinguishing specific types of precipitation events, such as heavy rainfall (Spektrum.de, 2004). This variable significantly influences both the risk and the derived risk. Another important variable is *Precipitation amount*, which describes the overall amount of rainfall during the event and is derived from the rainfall intensity. The final hazard variable is *Temperature*, as higher temperatures increase humidity and favor the occurrence of heavy rainfall events (Niu et al., 2024).

The vulnerability component can be subdivided into the following variables: *Land use*, which differentiates between arable land and grassland due to their different runoff behaviors (Achleitner et al., 2020). Soil type also plays a role, as different soils (clay, loam, sand) affect runoff behavior (Achleitner et al., 2020). Soil moisture is another critical factor, as dry soils can store more water and thus reduce runoff, whereas wet soils (e.g., after prior rainfalls) have a lower storage capacity and generate more runoff (Borga et al., 2007; Merz, 2008). The runoff coefficient, which is influenced by land use, soil type, and moisture, represents the ratio of runoff to precipitation and is crucial to determining the vulnerability of the area (Adam et al., 2000; Merz, 2008). Additionally, elevation plays a role — lower areas are more prone to flooding due to natural runoff (Huang et al., 2021), and the slope of the land influences how quickly water flows off (Huang et al., 2021).

The exposure component includes several variables: *Road density*, which represents the proportion of impermeable surfaces in an area. Roads have a high runoff coefficient, increasing the risk of flooding during heavy rainfall (U.S. Department of Agriculture, 1986). *River exposure*, which includes proximity to rivers and river discharge, is another important factor. Areas close to rivers are more exposed to flooding, as heavy rainfall and concentrated runoff can lead to higher river discharge and increased flood risk (Zhang et al., 2022). The variable *Proximity to forest* behaves differently. As forests have a low runoff coefficient, the runoff from precipitation seeps away more quickly here, whereby they act as a natural buffer. This variable reduces the impact on the area in a heavy rainfall event (U.S. Department of Agriculture, 1986).

These variables are conceptually linked to their respective components and are shown in a diagram, illustrating the dependencies between them (see Fig. 2).

3.3 Data Sources

In order to configure the BN and make predictions,careful selection and preparation of data sources are

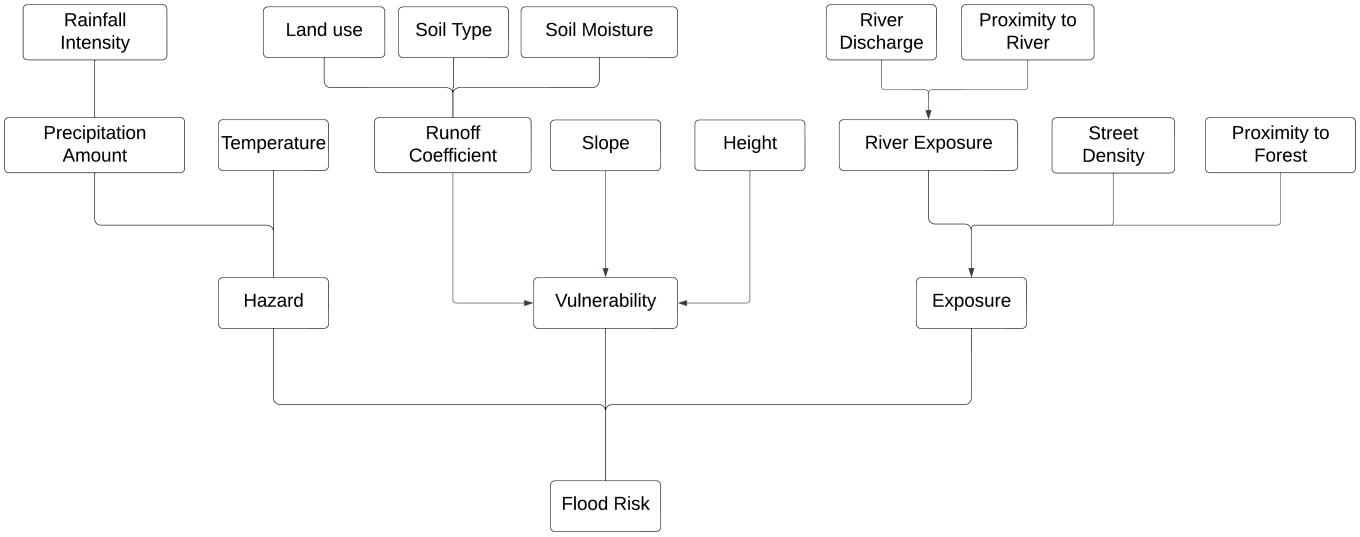

Fig. 2. Conceptual model of the Bayesian network (own illustration).

essential. The variables, as described in the previous chapter, are differentiated into two types for accurate analysis (see Table 1): data-based variables (evidence variables), for which data sources are available, and structural variables, which influence network calculations without having direct data sources.

Table 1. Classification of model variables.

Variable	Type
Rainfall Intensity	Data-based Variable
Precipitation Amount	Structural Variable
Temperature	Data-based Variable
Hazard	Structural Variable
Land Use	Data-based Variable
Soil Type	Data-based Variable
Soil Moisture	Data-based Variable
Runoff Coefficient	Structural Variable
Slope	Data-based Variable
Elevation	Data-based Variable
Vulnerability	Structural Variable
River Discharge	Data-based Variable
Proximity to River	Data-based Variable
River Exposure	Structural Variable
Road Density	Data-based Variable
Proximity to Forest	Data-based Variable
Exposure	Structural Variable
Flood Risk	Structural Variable

Data-based variables require further differentiation. Temporally changing variables (e.g., *Rainfall intensity*, *Temperature*) need periodic updates, which can be obtained from APIs like Open-Meteo, offering historical and up-to-date weather data (Zippenfenig, 2023). Location-based variables, like *Elevation* and *Land use*, can be sourced from freely available geodata, such as the Open GeoData Portal of LGL Baden-Württemberg[1], as shown in Tables 2, 3.

Table 2. Metadata of the variables: Data sources.

Variable	Unit	Data Source
Rainf. Intens.	mm/h	DWD API
Temperature	°C	Open-Meteo API
Land Use	-	Land Use Map - LGL
Soil Type	-	ALKIS - LGL
Soil Moisture	m^3/m^3	Open-Meteo API
Slope	%	GIS Analysis
Elevation	m	DTM - LGL
River Discharge	m^3/s	Open-Meteo API
Prox. to River	m	GIS Analysis
Road Density	%	GIS Analysis
Prox. to Forest	m	GIS Analysis

Table 3. Metadata of the variables: Resolution and uncertainties.

Variable	Resolution	Uncert.
Rainf. Intens.	10 min, single-station	-
Temperature	Hourly, single-station	-
Land Use	per district	±15 m
Soil Type	per district	±5%
Soil Moisture	Hourly, single-station	-
Slope	0,25 m × 0,25 m	±0.15 m
Elevation	0,25 m × 0,25 m	±0.15 m
River Discharge	Daily, single-station	-
Prox. to River	per district	±15 m
Road Density	per district	±15 m
Prox. to Forest	per district	±15 m

This section emphasized that careful selection of data sources is critical for the BN's configuration and predictions. Each variable must be supported by data that accurately represents it, either temporally or spatially.

[1] https://www.lgl-bw.de/Produkte/Open-Data/

3.4 GIS Analyses and Classification

Since the BN works with discrete values, each variable is classified into a set of state values that are associated with a probability. It is important to discretize all variables (Wu et al., 2019) and also to adapt the classification to the topographical conditions on site (Huang et al., 2021). Since the variables in the BN are divided into data-based variables and structural variables, the data-based variables must be considered in their entire data range in order to obtain meaningful threshold values for the classification. Structural variables are also divided into classes, but not on the basis of data, but with regard to the logical network configuration.

In Section 3.3, the subdivision of data-based variables into location-based and time-based variables has already been discussed. Location-based variables describe the topographical conditions in the study area. Variables such as *Altitude*, *Slope*, or *Land Use* must be extracted and processed using GIS analyses before they can be classified. QGIS is used for analyses as it is freely accessible.

The topographical components such as forest, road, and river geometries are extracted from the land use map, which is freely available from the LGL Baden-Württemberg.

The variable *Soil Type* can be taken from the Standardized Exchange Interface (NAS) data format for each parcel. NAS is a nationwide standardized data exchange interface for geoinformation and is used to exchange ALKIS data (LGLN, 2025). The parcels that are relevant for the application of the methodology are also contained in the ALKIS data.

The variable *Height* is taken from a digital terrain model with a grid resolution of 0.25 m x 0.25 m (DGM025) which is also freely accessible from the LGL. For each parcel, the median height is calculated from the height data of the grid cells that lie within the boundaries of the respective parcel. This ensures that the height is representative of the entire parcel and that extreme values have no influence.

The *Slope* variable is obtained from the digital terrain model using a slope analysis and describes the area at the same resolution as the *Height* variable. The median slope is calculated for each parcel in the same way as the *Height* variable.

The variables *Proximity to River* and *Proximity to Forest* are determined according to the following principle. The forest and river geometries are first converted into raster data. The minimum distance to the nearest raster cells of the river and forest geometry is then calculated for each parcel. These distances serve as the basis for deriving the variables that describe the proximity or distance of a parcel to rivers and forests.

A rather unusual approach was chosen for the *Road Density* variable. A 200 m buffer geometry was created for each parcel. Then an overlap analysis with the road geometry was carried out. This produces a percentage value that describes the proportion of the buffer area of the parcels covered by roads.

Usually, a ratio of square meters of road to square kilometers of area is used to calculate the road density. However, this approach is not representative for the study area, as the parcels and their buffer areas are often relatively small and a more meaningful percentage analysis of the area coverage is more suitable here.

Once processed, the data is classified into classes suitable for the BN. Most variables are divided into three classes, while the rainfall intensity classification follows DWD standards (DWD, 2024). Structural variables are classified based on logical relationships rather than data. The *Flood Risk* variable, for example, is divided into two classes, as seen in Table 4, and the classification of structural variables will be discussed in the next section, as these variables are interrelated.

Table 4. Classification of the variables.

Classes	***Low***	***Medium***	***High***
Rainf. Intens.	see 2.1	see 2.1	see 2.1
Temperature	< 10.0	10.0 - 21.0	> 21.0
Soil Moisture	< 0.2	0.2 - 0.4	> 0.4
Slope	< 2.9	2.9 - 5.0	> 5.0
Elevation	< 275.0	275.0 - 307.0	> 307.0
River Discharge	< 0.16	0.16 - 0.5	> 0.5
Prox. to River	< 100.0	100.0 - 200.0	> 200.0
Road Density	< 3.7	3.7 - 5.7	> 5.7
Prox. to Forest	< 100.0	100.0 - 200.0	> 200.0
Hazard	x	x	x
Prec. Amount	x	x	x
Runoff Coeff.	x	x	x
Vulnerability	x	x	x
Exposure	x	x	x
River Exposure	x	x	x
Flood Risk	x		x
Land Use			
Grassland		Arable Land	
Soil Type			
Clay	Loam	Clayey Loam	Sandy Loam

In summary, classification of data sources is essential, as a BN can only be configured or evaluated with classified data.

3.5 Configuration of Conditional Probabilities

As already mentioned in Section 2.3, the dependency and relationship between the individual influencing factors is influenced by the CPD. The following discusses the configuration of the CPD for different variables.

These CPDs can be obtained through parameter learning or by configuring them using empirical statistics and relationships (Huang et al., 2021). For parameter learning, a certain amount of data for each variable in the network is required, so that the CPD can be "learned" using various parameter learning algorithms. Since publicly available data for parameters such as hazard, vulnerability, precipitation amount, or runoff coefficient are not available in the required data quantity, the CPD are configured based on empirical statistics and relationships.

3.5.1 Configuration of Data-based Variables

The configuration of data-based variables follows empirical statistics. Classification is applied to a data-based

variable data set with 365 days, resulting in a classified data set to represent the variable for one year.

Table 5. Example data from the DWD CSV file.

Station-ID	Date	Rain. Am. (mm)
Station 1	2023-01-01	5.2
Station 2	2023-01-02	3.5
Station 3	2023-01-03	7.1

Table 5 displays a snippet from a CSV dataset, which serves as the data source for the rainfall intensity variable. *Rainfall Amount* indicates the amount of rain in mm for each station. Using a Python script, the data is classified for one year. The results are as follows:

- ***Low***: 358 days out of 365
- ***Medium***: 6 days out of 365
- ***High***: 1 day out of 365

These results are incorporated into the CPD distribution, as shown in Table 6.

Table 6. CPD for the variable Rainfall Intensity in the Bayesian network.

Rainfall Intensity	Probability
Low	0.9808
Medium	0.0164
High	0.0027

This approach is applied to all temporally variable variables in the BN. For site-related variables, such as *Height*, GIS data is classified similarly.

The classification results for this variable over the entire study area are:

- ***Low***: 773 out of 3659 parcels
- ***Medium***: 1757 parcels out of 3659
- ***High***: 1129 parcels out of 3659

The results are incorporated into the CPD distribution for the *Height* variable, as shown in Table 7.

Table 7. CPD for the *Height* variable in the Bayesian Network.

Height	Probability
Low	0.2113
Medium	0.4802
High	0.3086

3.5.2 Configuration of Structural Variables

For structural variables, it is not possible to rely on datasets during configuration; relationships between variables must be established, as discussed in Section 3.2. These relationships are then logically incorporated into the CPD distributions as probabilities.

As shown in Table 8, the CPD distribution of the *Precipitation Amount* variable depends on the *Rainfall Intensity* variable, following the logic: the higher the rainfall intensity, the more likely a high precipitation amount is, and vice versa. The categories *High*, *Medium* and *Low* correspond to the rainfall intensity levels, and they are represented as columns in the table. The values in column 1 show the conditional probability distribution for each combination of precipitation amount and rainfall intensity.

Table 8. CPD of Precipitation Amount depending on Rainfall Intensity.

	Rainfall Intensity Level		
Prec. Am.	**High**	**Medium**	**Low**
High	0.9	0.5	0.01
Medium	0.09	0.3	0.09
Low	0.01	0.2	0.9

- 1st row: The probability for a high precipitation amount is 90% (0.9) when the Rainfall Intensity variable is in the *High* state.
- 2nd row: The probability for a medium precipitation amount is 9% (0.09) when the Rainfall Intensity variable is in the *Medium* state.
- 3rd row: The probability for a high precipitation amount is 1% (0.01) when the Rainfall Intensity variable is in the *Low* state.

In the above section, it was shown that the configuration of a BN is based on data-based and structural variables. Data like rainfall intensity or height are classified based on empirical statistics, and their classification directly determines the probabilities in the CPD. Logical relationships, such as those between *Rainfall Intensity* and *Precipitation Amount*, define the CPD of the structural variables.

3.6 Advanced Analysis: Scenario Analysis, Sensitivity Analysis, Optimization, and Validation

Advanced analyses in flood risk assessment using a BN include scenario analysis, sensitivity analysis, optimization, and model validation.

Once the BN is configured, scenario analyses can be performed. For an initial analysis, a day with heavy rain is simulated, setting evidences for the data-based variables within the BN. Probabilities are derived for the target variable, *Flood Risk*, based on the reference event and calculated through exact inference in the BN (Sperotto et al., 2017).

In sensitivity analysis, the effects of stepwise changes in variable states are assessed to predict impacts on other variables or target variables (Korb and Nicholson, 2010; Kragt, 2009). The *prior probabilities* indicate the probability of the target variable based on the CPD. The BN is tested with empty priors to analyze dependencies and variable influences on the target variable.

Two methods are used for sensitivity analysis. First, the influence of each variable on the target's probabilities is analyzed. The extent to which each variable impacts the target's probability distribution is measured by calculating the average change in probabilities. This shows the variable's influence on the target in the BN.

In the second step, the influence of specific states of the influencing variables, as classified, on the target variable is examined. This helps identify if certain states have too large or small an effect on the target variable.

For optimizing a BN, based on sensitivity analysis, steps include reducing the states of important variables (e.g., *Hazard*, *Vulnerability*, *Exposure* from three to two states) to focus on extreme flood risk assessments, and reconfiguring CPD by adjusting variables with similar influences.

To validate a BN, the root mean square error (RMSE) is calculated to measure how well the predicted probabilities match real events. Validation data, such as from a heavy rainfall event, is required. RMSE evaluates the model's fit to actual data, but notably is sensitive to outliers, as errors are squared. A high RMSE indicates poor model fit, while a low RMSE shows better accuracy (Chai and Draxler, 2014).

The **RMSE** here is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{2}$$

Where:

- N: Number of data points
- y_i: Actual flood risk measurement
- $\hat{y}_i$: Predicted flood risk

4 Testing the Methodology

The results of an initial BN are then presented, as discussed and configured in Section 3.

In the following, a first scenario analysis is conducted on the BN, which has been configured based on literature and empirical statistics. It is expected that the scenario analysis will provide results that do not yet correspond to real events. Therefore, the application of advanced analyses, as presented in Section 3.6, is suitable to critically scrutinise the first results and strive for improvements in the accuracy of flood risk assessment.

4.1 Study Area and Heavy Rainfall Event of June 2, 2024

Rudersberg and the district of Schlechtbach are located in the Rems-Murr district in Baden-Württemberg, Germany. The municipality covers 39.37 km² with a population of 11,319 and is situated in the Wieslauftal valley at elevations between 270 and 536 meters.

The Wieslauf River flows through the region and is a tributary of the Rems. The area's hilly terrain and narrow valleys make it prone to flooding, particularly during intense rainfall events.

Land use consists mainly of forests (45.2%) and agricultural land (40.8%), which are vulnerable to heavy rainfall. The temperate climate has an average annual precipitation of 930 mm, mainly in the summer months (Climate-Data.org, 2025).

On June 2, 2024, an intense rainfall event occurred with 65 mm (65 l/m²) of rain falling within a few hours, causing significant flooding. This affected agricultural land, infrastructure, and residential areas, especially in Schlechtbach, with an economic damage of 332 million euros (Baden-Württemberg, 2024).

This event illustrates the vulnerability of the area to heavy rainfall, highlighting the combination of topography and agricultural use that makes the region particularly prone to flooding.

4.2 Results of an Initial Scenario Analysis

In the following Section, the initial results of the configured BN are presented. A scenario analysis of June 2, 2024, was conducted for the study area.

A total of 3,659 agricultural parcels in the Rudersberg / Schlechtbach area were evaluated using the BN, and a probability for the target node *flood risk* was calculated for each parcel based on relevant influencing factors. These influencing factors consist of location-based variables and temporally changing variables, as shown in Table 9. All variables, except for *Temperature*, are set to the state in which they exert the highest risk influence.

The location-based variables vary per parcel and were collected through GIS analyses for each parcel, as discussed in Section 3.4.

Table 9. Overview of the time-varying variables on June 2, 2024, in Rudersberg.

Variable	Value	Classification
Rainfall Intens.	63 l/m^2 in 6 h	*High*
Temperature	15.4°C	*Medium*
Soil Moisture	0.428 m³/m³	*High*
River Discharge	7.71 m³/s	*High*

To visualize the scenario analysis of June 2, 2024, and the results of the first model, all flood risk values of the parcels in the study area were mapped using GIS in a graduated map.

The model produces reasonable results for parcels that exhibit the following characteristics:

- High road density,
- Located in low-lying areas,
- Low slope inclination,
- Close proximity to river,
- Far from forested areas,
- Soils causing a high runoff coefficient.

Parcels with these characteristics tend to be assessed with a high probability of flood risk.

Parcels tend to be assessed with a low probability of flood risk if these variables have opposite values, like low road density.

The distribution of the model's probability values across the entire study area shows that most agricultural parcels are evaluated with a probability of 60%, concentrated in the range of moderate to slightly elevated flood risk. The mean probability of the first model is 63%, with a minimum of 54% and a maximum of 78%. Considering the minimum and maximum values, it can be concluded that the model

evaluates flood risk from heavy rainfall events for agricultural land in a rather conservative way.

Additionally, probability values close to the minimum and maximum are not sufficiently low or high. This suggests that flood risk probabilities predominantly cluster in a moderate range and extreme values are underrepresented.

As discussed in the previous Section, an initial scenario analysis allows the evaluation of the prior BN configuration. The model produces reasonable results based on the risk map. However, by examining the histogram of the probability distribution, it can be concluded that the BN is too conservatively assessed.

For further analysis and subsequent optimization of the BN, scenario analysis is crucial to derive insights and critically evaluate a conservatively assessed BN and its results.

4.3 Sensitivity Analysis

A sensitivity analysis was conducted to better understand the initial scenario analysis of the BN and assess its prior configuration. The analysis examines the influence of individual variables on the probabilities of the target variable.

The table below shows the influence of various variables on flood risk:

Hazard (36.14%), *Precipitation Amount* (25.85%), and *Rainfall Intensity* (23.01%) have the highest influence values, strongly impacting flood risk. Moderate influence of other variables: *Vulnerability* (17.26%) and *Exposure* (9.18%) also show significant influences. Although they affect the target variable noticeably, their impact is lower than the dominant factors above. *Slope* (6.31%) has a moderate influence on flood risk compared to many other variables.

Variables such as *Soil Moisture* (0.02%), *Land Use* (0.06%), *Soil Type* (0.57%), and *River Discharge* (0.48%) have minimal impact on the target variable. This means that changes in these variables hardly affect the target variable, making them less relevant for the current scenario analysis.

Table 10. Influence of individual variables on the target variable Flood Risk.

Variable	Influence (%)
Rainfall Intensity	23.01
Precipitation Amount	25.85
Temperature	3.21
Soil Moisture	0.02
River Discharge	0.48
Elevation	3.59
Slope	6.31
Proximity to River	0.93
Proximity to Forest	1.02
Road Density	2.59
Land Use	0.06
Soil Type	0.57
Hazard	36.14
Vulnerability	17.26
Exposure	9.18
River Exposure	1.91
Runoff Coefficient	3.71

Elevation (3.59%), *Runoff Coefficient* (3.71%), and *Temperature* (3.21%) show moderate influences. They are relevant but not dominant. *Proximity to River* (0.93%), *Proximity to Forest* (1.02%), and *Road Density* (2.59%) have small but not negligible influences. The influence of variables on the main components vulnerability and exposure is significantly lower than in the main component hazard because more variables are subordinate in those categories. As a result, multiple variables share a smaller influence, leading to lower individual impacts on flood risk.

The analysis reveals that *Rainfall Intensity* and *Precipitation Amount* are the most influential variables, while others such as *Soil Moisture* and *Land Use* are less significant. These insights can guide the optimization of the BN in the next step.

4.4 Optimization

The BN, which was applied and analyzed in Sections 4.2 and 4.3, is now optimized based on the results of the sensitivity analysis. The optimization is carried out in two steps, as discussed in Section 3.6. First, the reduction of the states of the three main components - hazard, vulnerability, and exposure - from three to two is performed. The middle state is removed, leaving only the states *High* and *Low*. The goal is to make the CPD distribution of the main components, as well as the flood risk variable more adaptable in order to address the problem of the model's conservative assessment.

The evaluation of the BN can be made more precise by adjusting the minimum and maximum ranges, making it more meaningful for assessing the risk of agricultural land due to heavy rainfall events.

The second step involves a new configuration of the CPD distributions, particularly for variables where the sensitivity analysis showed that all states had the same influence on the flood risk variable. The process was carried out as follows: influence variables that showed a low impact in the sensitivity analysis were given greater weight in the subsequent variables. Additionally, greater focus was placed on the states of variables where the influence of the states was nearly equal.

For example, closer proximity to a river increases water exposure, while being farther from a river leads to lower water exposure.

As demonstrated in this Section, a critical examination of the initial results from a scenario analysis and the insights from a sensitivity analysis are essential to improving the BN configuration. These analyses make it possible to identify weaknesses in the model structure and make targeted adjustments. This allows for the optimization of the network configuration and the modification of probability distributions.

5 Result of the Methodology

The optimization, as discussed in Section 4.4, was implemented in the BN. A renewed analysis of the same scenario for the Rudersberg / Schlechtbach area took place on June 2, 2024, the day of the heavy rainfall event. The same 3,659 agricultural parcels were assessed using the optimized

BN, and a probability for the target node flood risk was calculated for each parcel based on the relevant influencing factors.

Compared to the first scenario analysis with the initial BN, the results in the map, as seen in Fig. 3, are more clearly interpretable and can be linked to the aforementioned characteristics. This is mainly because the optimized model evaluates more progressively and distinctly classifies agricultural areas as either high-risk or low-risk.

The optimized model highlights differences between risk categories and assesses fewer areas in the medium-risk range. The optimized model presents a different probability distribution for the target variable *Flood Risk* compared to the initial model. Here as well, most parcels cluster around the middle probability range of approximately 50%. The mean value is 55%, with a minimum of 39% and a maximum of 86%. Considering the minimum and maximum values, it can be observed that the BN no longer evaluates as conservatively as before due to the optimization. The probabilities for the flood risk node now fluctuate more towards the minimum and maximum ranges, and the flood risk probabilities are no longer concentrated in the middle range.

As the results of the optimized BN demonstrate, it enables a more precise and differentiated assessment of flood risk for agricultural areas in the Rudersberg / Schlechtbach region. The model optimization leads to clearer classifications of high-risk and low-risk areas, significantly reducing the medium-risk category. The probability distribution now shows a greater spread with clearly distinguishable differences between minimal and maximal risks.

5.1 Results on Influencing Variables

Flood risk is strongly influenced by topographical and hydrological variables. Low slopes are predominantly associated with high flood risk, while steep slopes reduce it. *Proximity to River* alone is not a decisive factor; it must be assessed with elevation and slope. Areas near forests tend to have lower flood risk, emphasizing the protective effect of vegetation. High road density correlates with higher flood risk, likely due to increased surface runoff.

Table 11. Comparison of variable influences before and after optimization.

Variable	Before (%)	After (%)
Rainfall Intensity	23.01	21.33
Precipitation Amount	25.85	23.96
Temperature	3.21	3.68
Soil Moisture	0.02	0.13
River Discharge	0.48	0.35
Elevation	3.59	3.01
Slope	6.31	2.35
Proximity to River	0.93	0.97
Proximity to Forest	1.02	1.49
Road Density	2.59	1.78
Land Use	0.06	0.01
Soil Type	0.57	0.30
Hazard	36.14	26.58
Vulnerability	17.26	5.38
Exposure	9.18	5.41
River Exposure	1.91	1.78
Runoff Coefficient	3.71	0.77

Hazard, rainfall intensity, and precipitation amount remain the dominant factors, despite a slight decrease in influence after optimization. Elevation and slope show strong correlations with flood risk, with low-lying areas and gentle slopes being particularly vulnerable. Proximity to river and forests plays a role, but depends on other conditions. Road density is moderately relevant, while soil type and land use have only minor influence.

5.2 Validation of the Results

As discussed in the previous Section 5, the optimized model produces meaningful probability values in areas where the flood risk due to heavy rainfall events in agriculture is typically high.

To establish a solid benchmark for the reliability of the model, the predictions of the model are now validated using validation areas. These areas were selected based on the heavy rainfall event of June 2, 2024, in the Rems-Murr district. They represent agricultural areas that were mostly or at least partially flooded.

The data was provided by the Rems-Murr district office, Department of Agriculture. Communication with the Department of Agriculture confirms that the areas were affected by the heavy rainfall event.

At the beginning of the validation process, the areas must first be divided through GIS analyses so that they match the land parcels originally evaluated by the BN. Thus, 103 land parcels in the Rudersberg / Schlechtbach area are obtained, of which it is known with certainty that they were flooded.

A histogram of the probabilities generated by the BN for the validation parcels shows that all flood risk values, except one outlier, are above 70%, with a mean of 77% and a maximum of 87%.

As discussed in Section 3.6, RMSE is suitable for validating and assessing the accuracy of the BN. The RMSE for the optimized BN is 0.232 which means that the average deviation between the predicted probabilities and the actual values for the flood risk variable is 0.232 units, or 23.2%.

Considering the scale of the target node flood risk, which ranges between 0 and 1, the RMSE is less than a quarter of the scale. In a probabilistic model, where uncertainties are explicitly represented, as in a BN, this value is considered acceptable. According to Korb and Nicholson (2010), BN are specifically suited for modeling uncertainties, making errors in this range acceptable. From these validation results and the results for the entire study area, it can be concluded that the optimized BN successfully evaluates flood risk for agricultural areas during heavy rainfall events.

The validation shows that the optimized BN reliably estimates flood risk and realistically represents the events of June 2, 2024, in the Rems-Murr district. With an RMSE of 23.2%, the model is suitable for flood risk assessment of agricultural areas.

5.2.1 Cross-regional Validation

To further apply the BN in another area and validate it, the model is applied on June 2, 2024, during the heavy rainfall event, for a scenario analysis in another area of the

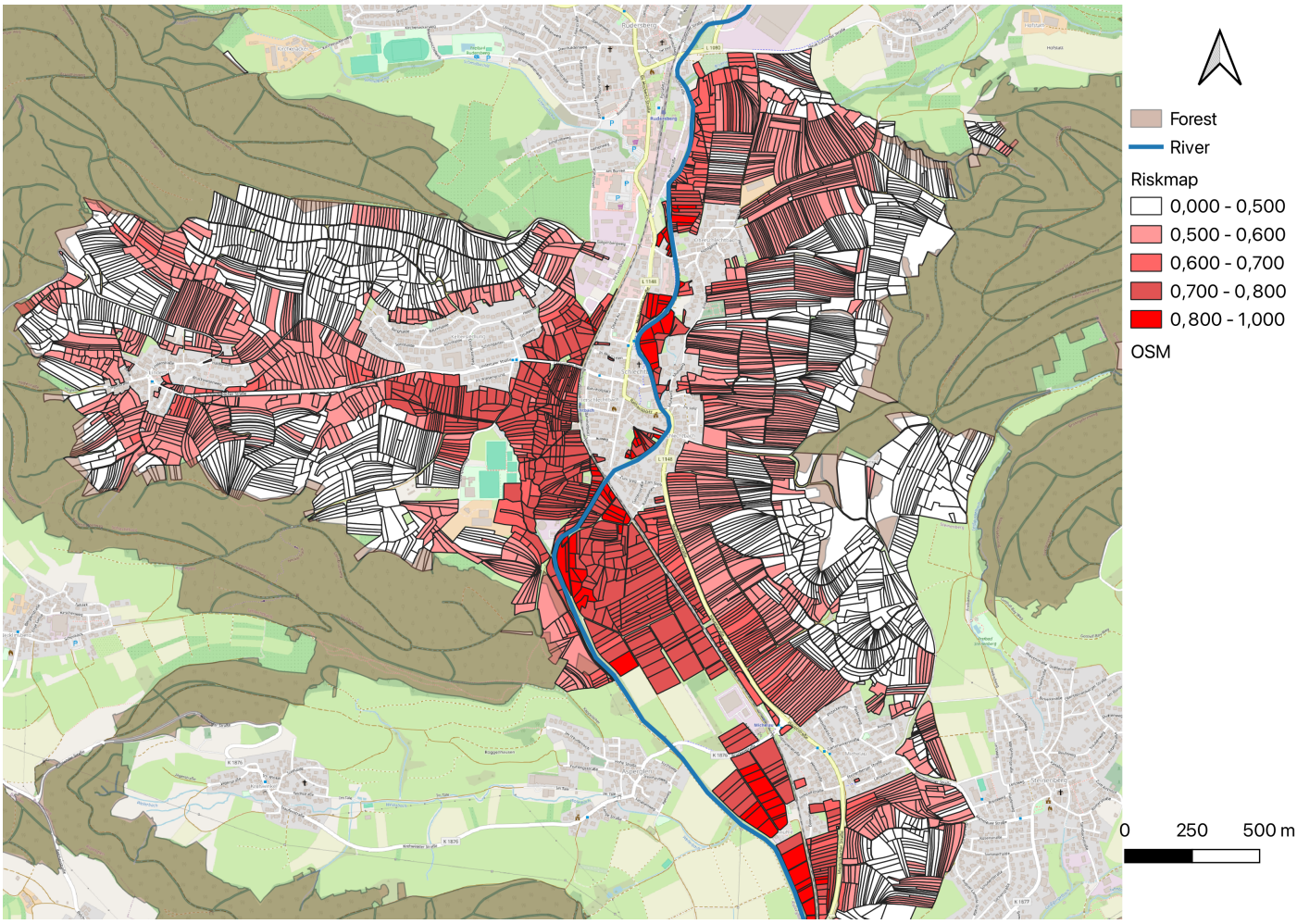

Fig. 3. Results of the optimization: Risk map of the agricultural parcels for the heavy rainfall event on 2 June 2024 in Rudersberg (own presentation).

Rems-Murr district.

Miedelsbach, a district of the city of Schorndorf, is located south of the study area Rudersberg / Schlechtbach and was also affected by the heavy rainfall event on June 2, 2024. The municipality is smaller than the main study area, and a total of 1,174 agricultural land parcels are analyzed here.

Meaningful results are obtained for agricultural land parcels, as described in Section 4.2, and also can be seen in Fig. 4.

The model produces meaningful results on land parcels with the following characteristics:

- high road density
- located in lower-lying areas
- shallow slope
- proximity to river
- far from forests

Land parcels exhibiting these characteristics tend to be rated with a high probability of flood risk.

Land parcels with opposite characteristics tend to be rated with a low probability of flood risk.

A histogram of the probabilities of flood risk for agricultural land parcels in Miedelsbach shows a minimum of 46%, a maximum of 85%, and a mean of 58%. The values in Miedelsbach are similar to those in the original study area of Rudersberg/Schlechtbach.

There are also validation land parcels in this area, which were also provided by the Rems-Murr district office, Department of Agriculture, and represent agricultural areas that were mostly or at least partially flooded during the heavy rainfall event on June 2, 2024.

A histogram of the probabilities generated by the BN for the validation land parcels in the Miedelsbach area shows differences from the original study area. Outliers are evident in the middle probability ranges. As discussed in 3.6, RMSE is sensitive to outliers, and larger deviations are given greater weight. A higher RMSE is expected in the Miedelsbach area, as the model was configured for Rudersberg/Schlechtbach, not Miedelsbach.

With validation, outliers appeared on eight land parcels in the middle range, which were flooded despite high slopes (between 5% and 8%). It should be noted that the slope of these areas was only slightly above the classification threshold for the *High* condition, which may explain the observed deviations. Nevertheless, this contradicts the assumption that steep slopes are less prone to flooding. Possible expla-

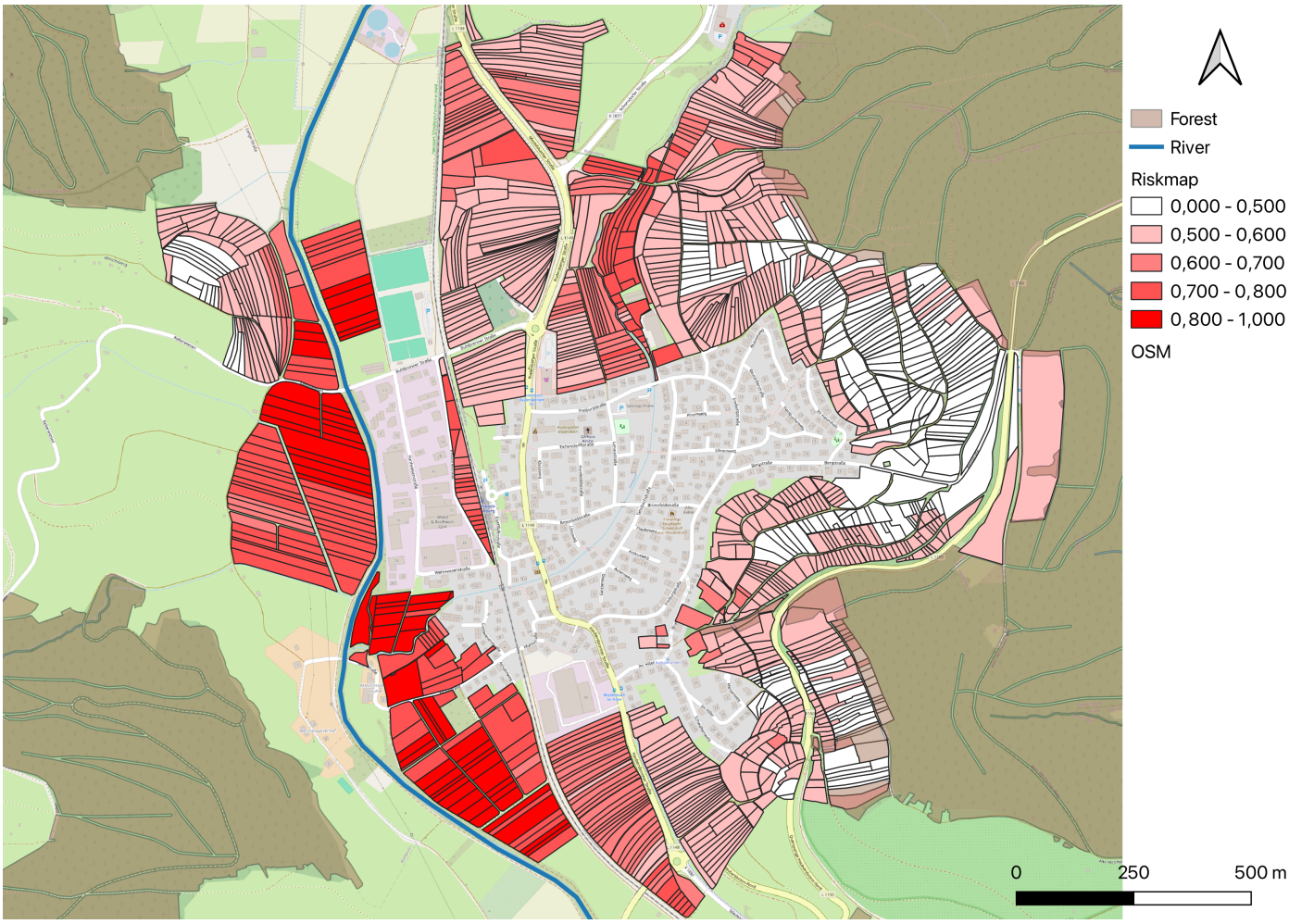

Fig. 4. Risk map of agricultural parcels for the heavy rainfall event on June 2, 2024 in Miedelsbach.

nations include backwater from lower-lying areas or missing variables from the BN model, such as drainage systems or erosion gullies.

A BN models uncertainties in a risk assessment due to a heavy rainfall event. Although the RMSE is slightly higher than a quarter of the flood risk target node scale, it is not significantly higher, at 30%. It is also important to note that this validation applies only to a subset of the total agricultural plots, and as can be seen from the flood risk probability values of the overall histogram, the model provides successful results.

The cross-regional validation shows that the optimized model is also applicable in other regions, as evidenced by the results in the Miedelsbach area. Although the RMSE value of 30% is relatively higher than in the original study area, it is primarily due to the impact of individual outliers, as these are weighted more heavily in the RMSE calculation due to the squared error method.

6 Conclusion

The increasing frequency of heavy rainfall events poses a growing risk to agricultural land, leading to soil saturation, erosion, and reduced crop yields. Given these challenges, precise risk analysis is essential. However, systematic approaches that explicitly model uncertainties and rely on high quality open data sources have been lacking.

To address this, a methodology combining BN and GIS analyses was developed, enabling probabilistic modeling of flood risk based on hydrological, meteorological, and topographical variables. This method effectively represents uncertainties and delivers reliable results even with limited data. Applied to the study area of Rudersberg/Schlechtbach, key variables such as rainfall intensity, precipitation amount, elevation, and slope were integrated. The heavy rainfall event of June 2024 was analyzed using scenario and sensitivity analyses to optimize the model and verify its validity.

The results confirm that BN and GIS analyses can successfully assess and quantify flood risk for agricultural land. The optimized model improved accuracy, enhancing the distinction between high- and low-risk areas and refining probability distributions. Rainfall intensity and precipitation amount remain dominant risk factors, while topographical features such as elevation and slope also exhibit strong correlations with flood risk. Proximity to river, road density, and land use contribute to risk levels in varying degrees. The optimized BN provides a solid foundation for prioritizing risk factors and supporting decision-making.

Model validation using 103 flooded parcels demonstrated high agreement with real flood events, with an av-

erage probability of 77% and an RMSE of 23%. The cross-regional validation in Miedelsbach confirmed the model's transferability, though regional data preparation and reconfiguration are necessary. Despite a slightly higher RMSE of 30%, the model predictions remained robust, underscoring its adaptability.

This study contributes to probabilistic flood risk modeling and highlights the importance of precise risk assessment in agriculture amid climate change. The use of open data sources ensures flexibility and scalability. Furthermore, the combination of Open Geospatial Data, Open Models, and Open Applications such as QGIS makes this analysis transferable. To mitigate future flood risks, targeted adaptation strategies should be implemented, such as optimizing land use in vulnerable areas, reducing road density, and improving protective forestation measures.

Future improvements should integrate additional datasets, including roughness values, detailed surface parameters, and drainage system data, to improve model accuracy. Furthermore, incorporating meteorological forecasts into the BN could provide real-time risk predictions. Expanding the BN into decision-support systems would strengthen agricultural risk management and help mitigate the impacts of heavy rainfall events.

Sources of Support

The authors received no funding for this study.

Conflict of Interest

The authors declare no conflict of interest.

CRediT statement

PG: Investigation, Methodology, Formal analysis, Writing – original draft. **ABV**: Visualization, Writing – original draft. **FJB**: Conceptualization, Data curation, Formal analysis, Writing – review & editing.

References

Abebe, Y., Kabir, G., Tesfamariam, S., 2018. Assessing urban areas vulnerability to pluvial flooding using GIS applications and Bayesian Belief Network model. *Journal of Cleaner Production* 174, 1629–1641. https://doi.org/10.1016/j.jclepro.2017.11.066.

Achleitner, S., Huber, A., Lumassegger, S., Kohl, B., Spira, Y., Weingraber, F., 2020. *Pilotstudie oberösterreich: Modellierung von starkregenoberflächenabfluss / hangwasser (leitfaden).* Association of European Border Regions. https://rainman-toolbox.eu/wp-content/uploads/2020/06/AU_Leitfaden.pdf.

Adam, C., Gläßer, W., Hölting, B. (Eds.), 2000. *Hydrogeologisches Wörterbuch.* Thieme, Stuttgart. https://www.ufz.de/index.php?en=20939&ufzPublicationIdentifier=6960.

U.S. Department of Agriculture, Natural Resources Conservation Service, 1986. *Urban hydrology for small watersheds, technical release 55 (tr-55).* https://www.nrc.gov/docs/ML1421/ML14219A437.pdf. technical release 55, Revised June 1986.

Ankan, A., Panda, A., 2015. pgmpy: Probabilistic Graphical Models using Python. *Proceedings of the Python in Science Conferences*, 6–11. https://doi.org/10.25080/majora-7b98e3ed-001.

Assmann, A., 2023. Starkregen und hochwasser in kleinen einzugsgebieten – auswirkungen, in: Porth, M., Schüttrumpf, H., Ostermann, U. (Eds.), *Wasser, Energie und Umwelt: Aktuelle Beiträge aus der Zeitschrift Wasser und Abfall III.* Springer, Wiesbaden, pp. 123–127. https://doi.org/10.1007/978-3-658-42657-6_13.

Aven, T., 2015. *Risk Analysis.* John Wiley & Sons. https://books.google.de/books/about/Risk_Analysis.html?id=41V_BwAAQBAJ&redir_esc=y.

Baden-Württemberg, Regierungspräsidien, 2024. Gemeinsam gegen die Folgen des Unwetters: Landeshilfe aktiviert / rund 19 Millionen Euro zusätzliche Hilfe für den Stuttgarter Regierungsbezirk. https://rp.baden-wuerttemberg.de/rps/wir/abteilungen/pressemitteilungen-und-aktuelle-meldungen/artikel/hochwasser-landeshilfe-ak-tiviert-19-millionen-euro-zusaetzliche-hilfe-fuer-den-stuttgarter-regierungsbezirk/. zugriff am 19.12.2024, 11:52.

Becker, P., Becker, A., Dalelane, C., Deutschländer, T., Junghänel, T., Walter, A., 2016. Die entwicklung von starkniederschlägen in deutschland: Plädoyer für eine differenzierte betrachtung. *Berichte des Deutschen Wetterdienstes* 245. https://www.dwd.de/DE/leistungen/besondereereignisse/niederschlag/20160719_entwicklung_starkniederschlag_deutschland.pdf?__blob=publicationFile&v=3.

Bill, R., 2010. *Grundlagen der Geo-Informationssysteme.* Wichmann.

Borga, M., Boscolo, P., Zanon, F., Sangati, M., 2007. Hydrometeorological analysis of the 29 August 2003 flash flood in the Eastern Italian Alps. *Journal of Hydrometeorology* 8, 1049–1067. https://doi.org/10.1175/jhm593.1.

Chai, T., Draxler, R. R., 2014. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development* 7, 1247–1250. https://gmd.copernicus.org/articles/7/1247/2014/.

Climate-Data.org, 2025. Klimadaten für Rudersberg, Baden-Württemberg, Deutschland. https://de.climate-data.org/europa/deutschland/baden-wuerttemberg/rudersberg-156016/. zugriff am 09.01.2025, 16:05.

Cooper, G.F., Herskovits, E., 1992. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* 9, 309–347. https://doi.org/10.1007/bf00994110.

CRED, Centre for Research on the Epidemiology of Disasters, 2025. *Em-dat: The International Disaster Database.* https://public.emdat.be/data. zugriff am 17.01.2025, 16:07.

Dinkelbach, W., 1969. Sensitivitätsanalysen und parametrische Programmierung. In *Ökonometrie und Unternehmensforschung.* https://doi.org/10.1007/978-3-642-88169-5.

DWD, 2024. *Starkregen: Definition und Klassifikation.* https://www.dwd.de/DE/service/lexikon/begriffe/S/Starkregen.html. https://www.dwd.de/DE/service/lexikon/begriffe/S/Starkregen.html. zugriff am 10.12.2024, 16:56.

Fatichi, S., Vivoni, E., Ogden, F., Ivanov, V., Mirus, B., Gochis, D., Downer, C., Camporese, M., Davison, J., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., Tarboton, D., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology* 537. doi:10.1016/j.jhydrol.2016.03.026.

Furlan, E., Slanzi, D., Torresan, S., Critto, A., Marcomini, A., 2020. Multi-scenario analysis in the adri-

atic sea: A gis-based Bayesian Network to support maritime spatial planning. *Science of The Total Environment* 703, 134972. https://www.sciencedirect.com/science/article/pii/S0048969719349642.

Harris, R., Furlan, E., Pham, H.V., Torresan, S., Mysiak, J., Critto, A., 2022. A Bayesian network approach for multi-sectoral flood damage assessment and multi-scenario analysis. *Climate Risk Management* 35, 100410. https://doi.org/10.1016/j.crm.2022.100410.

Hrachowitz, M., Clark, M. P., 2017. Hess opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences* 21, 3953–3973. https://hess.copernicus.org/articles/21/3953/2017/.

Huang, S., Wang, H., Xu, Y., She, J., Huang, J., 2021. Key disaster-causing factors chains on urban flood risk based on Bayesian Network. *Land* 10, 210. https://doi.org/10.3390/land10020210.

Korb, Kevin B., Nicholson, A.E., 2010. *Bayesian Artificial Intelligence*. In CRC Press eBooks. https://doi.org/10.1201/b10391.

Koski, T., Noble, J., 2011. *Bayesian Networks: An introduction*. John Wiley & Sons. https://books.google.de/books?id=aoBzNiDPfQsC&printsec=frontcover&redir_esc=y#v=onepage&q&f=false.

Kragt, M.E., 2009. A beginners guide to Bayesian network modelling for integrated catchment management. Landscape Logic. https://www.utas.edu.au/__data/assets/pdf_file/0011/588467/TR_9_BNs_for_Integrated_Catchment_Management.pdf.

Kreibich, H., Schröter, K., Shprits, Y., Bedford, J., Tilmann, F., 2018. Maschinelles lernen verbessert gefährdungs- und risikoanalysen von naturgefahren. *System Erde* 8, 10–17. https://gfzpublic.gfz-potsdam.de/pubman/item/item_3539918.

Lehmann, J., Coumou, D., Frieler, K., 2015. Increased record-breaking precipitation events under global warming. *Climatic Change* 132, 501–515. https://doi.org/10.1007/s10584-015-1434-y.

LGLN, Landesamt für Geoinformation und Landesvermessung Niedersachsen, 2025. Nas. https://www.lgln.niedersachsen.de/startseite/wir_uber_uns/hilfe_support/lgln_lexikon/n/nas-190332.html#:~:text=NAS%20ist%20ein%20auf%20XML,mit%20Geobasisdaten%20in%20einem%20Geoinformationssystem.. zugriff am 14.01.2025, 09:10.

Li, G., Wu, X., Han, J.-C., Li, B., Huang, Y., Wang, Y., 2023. Flood risk assessment by using an interpretative structural modeling based Bayesian network approach (ISM-BN): An urban-level analysis of Shenzhen, China. *Journal of Environmental Management* 329, 117040. https://www.sciencedirect.com/science/article/pii/S0301479722026135.

Li, Y., Guan, K., Schnitkey, G.D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Global Change Biology* 25, 2325–2337. https://doi.org/10.1111/gcb.14628.

Lu, Y., Zhai, G., Zhou, S., 2024. An integrated bayesian networks and geographic information system (bns-gis) approach for flood disaster risk assessment: A case study of Yinchuan, China. *Ecological Indicators* 166, 112322. doi:10.1016/j.ecolind.2024.112322.

Merz, R., 2008. Einflussfaktoren auf die form der hochwasserwahrscheinlichkeitskurve, in: *Hochwasser, Wassermangel, Gewässerverschmutzung - Problemlösung mit modernen hydrologischen Methoden. Beiträge zum Tag der Hydrologie am 27./28. März 2008 an der Leibniz Universität Hannover*, DWA, Fachgemeinschaft Hydrologische Wissenschaften, Hennef. pp. 113–106. https://hdl.handle.net/20.500.11970/112870.

Muthiah, K., Arunya, K.G., Sridhar, V., Patakamuri, Sandeep Kumar, 2025. Heavy rainfall impact on agriculture: Crop risk assessment with farmer participation in the paravanar coastal river basin. *Water* 17. https://www.mdpi.com/2073-4441/17/5/658.

Niu, H., Sun, W., Huai, B., Wang, Y., Chen, R., Han, C., Wang, Y., Zhou, J., Wang, L., 2024. Causes of increased compound temperature and precipitation extreme events in the arid region of Northwest China from 1961 to 2100. *Remote Sensing* 16, 3111. https://doi.org/10.3390/rs16173111.

Pearl, J., 2014. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Elsevier. https://books.google.de/books?id=mn2jBQAAQBAJ&printsec=frontcover&hl=de&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Rana, I.A., Routray, J.K., 2017. Integrated methodology for flood risk assessment and application in urban communities of Pakistan. *Natural Hazards* 91, 239–266. https://doi.org/10.1007/s11069-017-3124-8.

Spektrum.de, 2004. *Lexikon der geographie: Regenintensität*. https://www.spektrum.de/lexikon/geographie/regenintensitaet/6508. zugriff am 18.12.2024, 16:39.

Sperotto, A., Molina, J.-L., Torresan, S., Critto, A., Marcomini, A., 2017. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. *Journal of Environmental Management* 202, 320–331. https://doi.org/10.1016/j.jenvman.2017.07.044.

Weigell, P., 2023. Data coverage, richness, and quality of Openstreetmap for special interest tags: wayside crosses – a case study. https://arxiv.org/abs/2306.04752, arXiv:2306.04752.

Wilke, S., 2024. Trends der Niederschlagshöhe. https://www.umweltbundesamt.de/daten/klima/trends-der-niederschlagshoehe#teilweise-sehr-regenreiche-jahre-seit-1965. zugriff am 17.01.2025, 14:26.

Wu, Z., Shen, Y., Wang, H., Wu, M., 2019. Assessing urban flood disaster risk using Bayesian network model and GIS applications. *Geomatics Natural Hazards and Risk* 10, 2163–2184. https://doi.org/10.1080/19475705.2019.1685010.

Zhang, P., Sun, W., Xiao, P., Yao, W., Liu, G., 2022. Driving Factors of Heavy Rainfall Causing Flash Floods in the Middle Reaches of the Yellow River: A Case Study in the Wuding River Basin, China. *Sustainability* 14, 8004. https://doi.org/10.3390/su14138004.

Zippenfenig, P., 2023. Open-meteo.com weather api URL: https://open-meteo.com/, doi:10.5281/zenodo.7970649. creative Commons Attribution 4.0 International.