



RISC-KIT

Resilience-Increasing Strategies for Coasts – Toolkit

www.risckit.eu

D3.3 Bayesian Decision Support System

Deliverable No: D.3.3 – Bayesian Decisions Support System

Ref.: WP3 - Task 3.3

Date: March 2015

Grant Agreement No. 603458
DG Research –FP7-ENV-2013-two-stage



This project has received funding from the European Union's Seventh Programme for Research, Technological Development and Demonstration under Grant Agreement n° [603458]. This publication reflects the views only of the author's, and the European Union cannot be considered liable for any use that may be made of the information contained therein.

This document contains information, which is proprietary of the RISC-KIT consortium. Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any third party, in whole or in parts, except with the prior written consent of the RISC-KIT coordinator.

Deliverable Title	D.3.3 – Bayesian Decision Support Tool
Filename	RISCKIT_D.3.3_BayesianDSS_v1
Authors	Wiebke Jäger (Delft University of Technology)
Contributors	Kees den Heijer (Delft University of Technology) Annelies Bolle (International Marine and Dredging Consultants) Steven Smets (International Marine and Dredging Consultants)
Date	30/03/2015

Prepared under contract from the European Commission

Grant Agreement No. 603458

Directorate-General for Research & Innovation (DG Research), Collaborative project, FP7-ENV-2013-two-stage

Start of the project: 01/11/2013

Duration: 42 months

Project coordinator: Stichting Deltares, NL

Dissemination level

<input checked="" type="checkbox"/>	PU	Public
<input type="checkbox"/>	PP	Restricted to other programme participants (including the Commission Services)
<input type="checkbox"/>	RE	Restricted to a group specified by the consortium (including the Commission Services)
<input type="checkbox"/>	CO	Confidential, only for members of the consortium (including the Commission Services)

Deliverable status version control

Version	Date	Author	Review
1.0	18/03/2015	Wiebke Jäger (Delft University of Technology)	Simone de Kleermaeker (Deltares); Kees den Heijer (Delft University of Technology)
2.0	25/03/2015	Wiebke Jäger (Delft University of Technology)	Ap van Dongeren (Deltares)

Table of Contents

1	Introduction.....	3
1.1	Project objectives.....	3
1.2	Project structure	4
1.3	Deliverable context and objective.....	6
1.4	Approach.....	7
1.5	Outline of the report	7
2	Research Gap and Objective of Task.....	8
3	Framework	10
3.1	The Variable Categories.....	10
3.2	Set-up of System Relations.....	12
3.2.1	Configuration Step 1.....	13
3.2.2	Configuration Step 2.....	14
3.2.3	Configuration Step 3.....	14
3.2.4	Configuration Step 4.....	15
4	Demonstration of the Bayesian DSS	16
4.1	Brief Introduction to the Zeebrugge Site	16
4.2	Input Data	17
4.2.1	Hazard Boundary Conditions	17
4.2.2	Receptors	18
4.2.3	Hazards	18
4.2.4	Impacts	19
4.3	Examples of Use	19
4.3.1	The Basic State of the DSS	20
4.3.2	Predicting Impacts.....	21
4.3.3	Diagnosing Storms.....	22
5	Conclusion	24
	References	25

List of Figures

Figure 1.1: Conceptual drawing of the CRAF (top panel), the EWS (middle panel) and the DSS (bottom panel).....	5
Figure 1.2: Case study sites (stars), RISC-KIT case study site partners (blue solid dots) and non-case study site partners (red open circles).....	6
Figure 2.1: Demo Bayesian network.....	8
Figure 3.1: Structure of Bayesian DSS (compare with Figure 3.7 to see more detail)	10
Figure 3.2: Example of BC node	10
Figure 3.3: Example of R node	11
Figure 3.4: Example of H node.....	11
Figure 3.5: Example for C node	12
Figure 3.6: Example of DRR node.....	12
Figure 3.7: Illustration of system relations of the DSS for two receptor types. BC1, BC2 and BC3 represent hazard boundary conditions; R1 and R2 represent receptor densities; H1_R1, H2_R1 and H3_R1 represent hazards affecting receptor type 1; H1_R2 and H4_R2 represent hazards affecting receptor type 2; C1_R1 and C2_R1 represent impacts on receptor type 1; C3_R2 represents an impact on receptor type 2; DDR1 and DDR2 represent DRR measures.....	13
Figure 4.1: Example of a maximum inundation depth map of Zeebrugge.....	18
Figure 4.2: Residential houses and subdivision in areas	19
Figure 4.3: Depth-damage curve for Flanders, Belgium.....	19
Figure 4.4: Expected storm impacts in the hot spot area	21
Figure 4.5: Predicting impacts in the hot spot area of a storm with $Z \geq 7.8\text{m}$ and uncertain wave height.....	21
Figure 4.6: Predicting impacts in the hot spot area of a storm with $Z=7.8\text{m}$ and $H=5.9\text{m}$	22
Figure 4.7: Predicting impacts of a storm with $Z=7.8\text{m}$ and $H=5.9\text{m}$ specifically for area 1	22
Figure 4.8: Diagnosing which storm scenarios can cause maximum damage.....	23

List of Tables

Table 3.1: Example table for the DSS specific georeferencing of receptors (format may be subject to change)	14
Table 3.2: Example format of CVI table for DSS. (This table corresponds to the DSS shown in Figure 1.1)	15
Table 3.3: Example of variables names	15
Table 4.1: Wave heights and water levels for the test matrix	17
Table 4.2: Number of houses per area.....	18

Executive Publishable Summary

This report describes the framework of the Bayesian-based decision support system (DSS) for coastal multi hazards (RISC-KIT deliverable 3.3). By connecting hazard intensities with socio-economic, environmental and cultural properties of different receptor types, the DSS allows for the prediction of impacts resulting from coastal multi-hazards for various hot spot areas.

A Bayesian network (BN), the modeling approach used, is a data driven model that describes system relations in probabilistic terms. Generally, BNs are pre-fed with data from which they “learn” relationships and can then be used to predict or diagnose events. The graphical user interface is very similar to causal diagrams.

The DSS is based on data from hydrodynamic storm simulations, information on land use and so-called vulnerability relationships. The approach can easily be applied to any hot spot area. This report introduces the general framework and explains how the system can be configured at individual sites. Moreover, a demo DSS is presented for the Belgian town of Zeebrugge, which can predict spatially varying inundation depths and damages to buildings for different storm scenarios. Furthermore, it can be used to diagnose under which storm conditions and where on the site the highest damages occur.

1 Introduction

Recent and historic low-frequency, high-impact events such as Xynthia (impacting France in 2010), the 2011 Liguria (Italy) Flash Floods and the 1953 North Sea storm surge which inundated parts of the Netherlands, Belgium and the UK have demonstrated the flood risks faced by exposed coastal areas in Europe. Typhoons in Asia (such as Typhoon Haiyan in the Philippines in November 2013), hurricanes in the Caribbean and Gulf of Mexico, and Superstorm Sandy, impacting the northeastern U.S.A. in October 2012, have demonstrated how even larger flooding events pose a significant risk and can devastate and immobilize large cities and countries.

These coastal zone risks are likely to increase in the future (IPPC, AR5) which requires a re-evaluation of coastal disaster risk reduction (DRR) strategies and a new mix of prevention (e.g. dike protection), mitigation (e.g. limiting construction in flood-prone areas; eco-system based solutions) and preparedness (e.g. Early Warning Systems, EWS) (PMP) measures. Even without a change in risk due to climate or socio-economic changes, a re-evaluation is necessary in the light of a growing appreciation of ecological and natural values, which drive ecosystem-based or Nature-based flood defense approaches. In addition, as free space is becoming sparse, coastal DRR plans need to be spatially efficient, allowing for multi-functionality.

1.1 Project objectives

In response to these challenges, the RISC-KIT project aims to deliver a set of open-source and open-access methods, tools and management approaches to reduce risk and increase resilience to low-frequency, high-impact hydro-meteorological events in the coastal zone. These products will enhance forecasting, prediction and early warning capabilities, improve the assessment of long-term coastal risk and optimise the mix of PMP-measures. Specific objectives are:

1. Review and analysis of current-practice coastal risk management plans and lessons-learned of historical large-scale events;
2. Collection of local socio-cultural-economic and physical data at case study sites through end-user and stakeholder consultation to be stored in an impact-oriented coastal risk database;
3. Development of a regional-scale coastal risk assessment framework (CRAF) to assess present and future risk due to multi-hazards (Figure 1.1, top panel);
4. Development of an impact-oriented Early Warning and Decision Support System (EWS/DSS) for hot spot areas consisting of: i) a free-ware system to predict hazard intensities using coupled hydro-meteo and morphological models and ii) a Bayesian-based Decision Support System which integrates hazards and socio-economic, cultural and environmental consequences (Figure 1.1, centre panel);
5. Development of potential DRR measures and the design of ecosystem-based and cost-effective, (non-)technological DRR plans in close cooperation with end-users for a diverse set of case study sites on all European regional seas and on one tropical coast (Figure 1.1; bottom panel);

6. Application of CRAF and EWS/DSS tools at the case study sites to test the DRR plans for a combination of scenarios of climate-related hazard and socio-economic vulnerability change and demonstration of the operational mode;
7. Development of a web-based management guide for developing integrated plans along Europe's coasts and beyond and provide a synthesis of lessons learned in RISC-KIT in the form of policy guidance and recommendations at the national and EU level.

The tools are to be demonstrated on case study sites on a range of EU coasts in the North- and Baltic Sea Region, Atlantic Ocean, Black Sea and Mediterranean Sea, and one site in Bangladesh, see Figure 1.2. These sites constitute diverse geomorphic settings, land use, forcing, hazard types and socio-economic, cultural and environmental characteristics. All selected regions are most frequently affected by storm surges and coastal erosion. A management guide of PMP measures and management approaches will be developed. The toolkit will benefit forecasting and civil protection agencies, coastal managers, local government, community members, NGOs, the general public and scientists.

1.2 Project structure

The project is structured into seven Work Packages (WP) starting with WP1 on 'Data collection, review and historical analysis'; WP2–4 will create the components of the RISC-toolKIT containing an 'Improved method for regional scale vulnerability and risk assessment' (WP2), 'Enhanced early warning and scenario evaluation capabilities for hot spots' (WP3) as well as 'New management and policy approaches to increase coastal resilience' (WP4). The toolkit will be tested through 'Application at case study sites' (WP5). WP6 will be responsible for 'Dissemination, knowledge transfer and exploitation' and 'Coordination and Management' are handled in WP7.

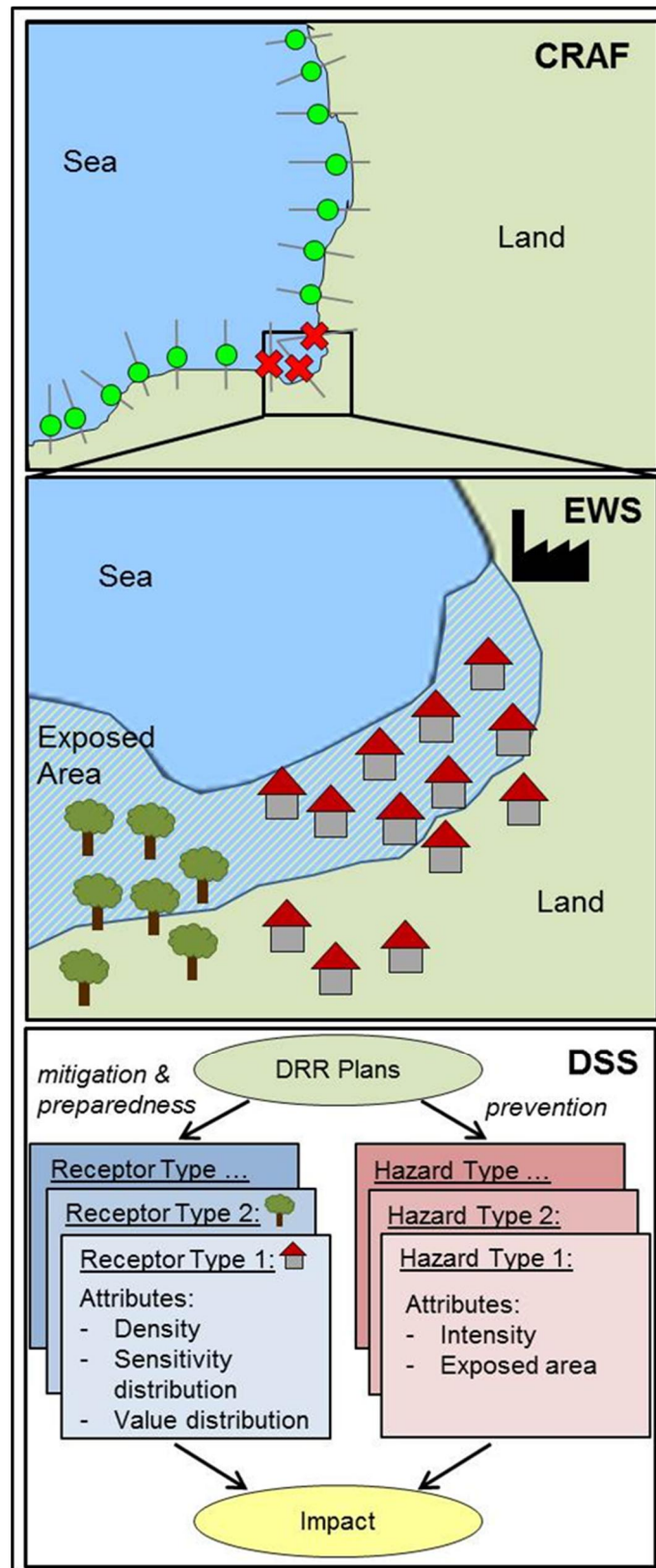


Figure 1.1: Conceptual drawing of the CRAF (top panel), the EWS (middle panel) and the DSS (bottom panel)



Figure 1.2: Case study sites (stars), RISC-KIT case study site partners (blue solid dots) and non-case study site partners (red open circles).

1.3 Deliverable context and objective

The Description of Work states the objective and context of this Deliverable as follows:

“The Bayesian decision support system (DSS) connects physical hazards to socio-economic, environmental and cultural properties of the receptor types, and the effects of the DRR plans. The DSS will allow for the prediction of impacts of coastal multi-hazards for the case study sites, including uncertainty bands. In this task we will design the structure (nodes and relations) of the Bayesian network using the freeware Netica package and a supporting tool that generates impact data from hazard intensities and receptor attributes using the vulnerability relations compiled under WP2 (Task 2.2). This tool will be used in WP5, Task 5.3 to train the system for each case study site

The deliverable of this task is a demo Bayesian DSS and supporting tool (D.3.3, this deliverable [added]), including configuration manuals which will be made available through the project website. We will organize a first training session (Milestone 9) for the partners in WP5 to set-up the DSS for their case study sites. Once the scenarios are run in WP5, Task 5.4, a second training session will be given on how to incorporate the

scenario data in the Bayesian network and on how to visualize the results.”

This deliverable is a demonstration Bayesian DSS, based on the freeware Netica (Norsys 2014) package and a supporting tool that generates impact data from hazard intensities and receptor attributes using the vulnerability relations compiled under WP2. The tool and configuration manual (contained in this document) will be made available through the project website.

This deliverable addresses the objective of WP 3 and Project Objective 4 and the quoted text from the “Description of Work” above, by providing the framework for the DSS part of the EWS/DSS for hot spot areas.

1.4 Approach

The idea is that a Bayesian DSS can be built for an arbitrary site. In RISC-KIT all case study site partners (CSPs) will set up an individual DSS for the local hot spot area of their case study site (cf. Figure 1.2) to exemplify its general applicability (RISC-KIT project Milestone 10 resulting from task 5.2: EWS/DSS systems set up for each case study site).

While this deliverable report addresses the RISC-KIT CSPs, its guidelines can be consulted by anyone to set up a DSS for any case study site.

1.5 Outline of the report

This report is structured as follows. Section 2 gives a more detailed account on the objective of the task and describes the knowledge gaps it addresses. Section 3 introduces the framework of the DSS. It describes the general structure and types of variables of the DSS and contains configuration instructions. Section 4 presents a demo version of the tool for the hot spot Zeebrugge. Section 5 concludes the report.

2 Research Gap and Objective of Task

Researchers across many disciplines are dedicated to developing methodologies that identify risks and to helping decision makers design effective risk reduction plans. On the one hand they apply numerical hydrodynamic process models to assess the natural coastal response and the extent of flooding due to storms. In this project among others XBeach, TELEMAC and LISFLOOD are used. On the other hand they use vulnerability relations to estimate potential impacts for those receptors that are exposed to coastal hazards, e.g. depth – damage curves. Many more examples can be found in the Coastal Vulnerability Indicator (CVI) Library (Deliverable 2.2) that has been developed within WP2 and addresses socio-economic, environmental and cultural impacts (Viavattene et al. 2015).

Integrated assessment methods are still rare. Ideally, integrated models depict the entire cause-effect chain from multiple coastal hazards to their diverse consequences reflecting the complexity of the issue. They offer deeper insight in the design of short, middle, or long-term DRR measures than single discipline models and help to avoid unforeseen consequences of those measures (Rotmans et al. 1994). Recently, Jonkman et al. (2008) proposed a GIS-based approach to describe a spatially varying flood hazard and resulting estimates of direct physical damages to various objects, indirect economic damage and the loss of life. The model presents the results of one typical low probability-high impact flood scenario in the Netherlands with the help of maps. Task 3.3 continues on the trend of integrating the separate modeling approaches into a homogeneous framework. The purpose is to compile impact estimates of many different storm scenarios in a Bayesian network (BN).

BNs are data based models that describe system relations in probabilistic terms. They are pre-fed with data from which they “learn” relationships and then used to predict or diagnose events. The graphical user interface is very similar to causal diagrams. An example of such a model is given in Figure 2.1 (and also in Figure 4.4 in section 4.3 where it is described in detail).

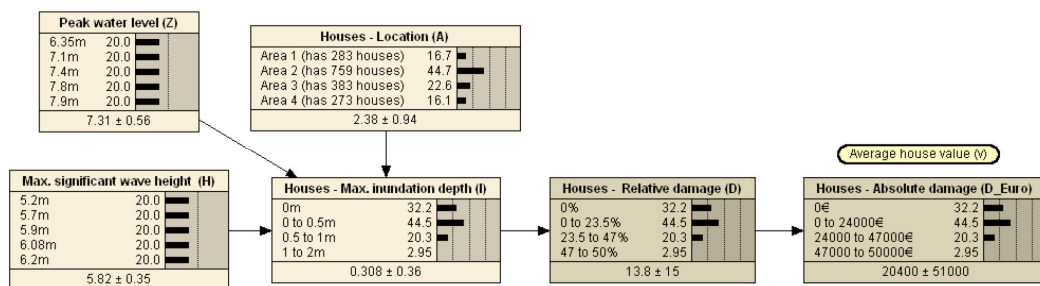


Figure 2.1: Demo Bayesian network

BNs have been applied a number of times as tools for decision-making under uncertainty. Hendriksen et al. (2007) conclude that they are very valuable for negotiations and discussions between managers, experts, stakeholders and representatives of the general public, among others, because they are transparent and flexible tools. In the context of floods, Garrote et al. (2007) combine BNs and deterministic rain run-off models to forecast flooding, and Vogel et al. (2012, 2013) use BNs to estimate damages resulting from river floods. In coastal environments they

have been applied to predict erosion and shoreline retreat (Den Heijer 2013; Gutierrez et al. 2011; Hapke & Plant 2010).

An advantage of BNs is their ability to handle various sources and types of data. The BN framework developed for RISC-KIT combines information on the topography and assets of the potentially affected area with simulation data of flood scenarios and damage estimations from the CVI relations. More precisely, the framework relates coastal storm impacts not only to erosion or flood extent, but also to near-shore storm conditions, such as peak water level, maximum significant wave height or storm duration. A second major benefit, and generally speaking the main use of BNs, is the possibility to update predictions according to the latest available information. In the RISC-KIT context, the site-specific BNs are fed with numerous storm scenario computations, based on which it will give a general prediction of expected inundation and impacts in the hotspot area. This prediction can be updated and tailored to an approaching storm with the latest field observations or model predictions, without the need for doing computationally expensive (near shore) simulations in real-time.

Because BNs are very light, they can be installed on stand-alone computers, laptops or smart phones and still make consequence predictions for impending storms in real-time. In contrast, the hydrodynamic process models, which are currently used to make inundation predictions, are implemented on high performance computers due to computational demand. Warnings and information on events thus need to be distributed to decision makers in the field through ICT networks, which have high vulnerability during extreme events. Also, because of the processing time for computations, it is a challenge to include local observations and rapidly changing information that becomes available towards the peak of an event. Therefore, BNs are more suitable than hydrodynamic process models to provide urgent decision support, possibly in the field. Moreover, the BN enables planners to test risk reduction measures for robustness against a variety of storm scenarios, as long as these measure have been included in the model set up.

3 Framework

Each hot spot area (with those considered in RISC-KIT given in Figure 1.2) will have its individual Bayesian DSS with a different choice of variables. However, all models have the same structure and link coastal hazards to impacts. In general, five categories of variables can be distinguished: (1) Hazard boundary conditions, (2) receptors, (3) local hazards, (4) impacts, and (5) DRR measures, defined below. The yellow boxes in Figure 3.1 indicate (groups of) variable categories.

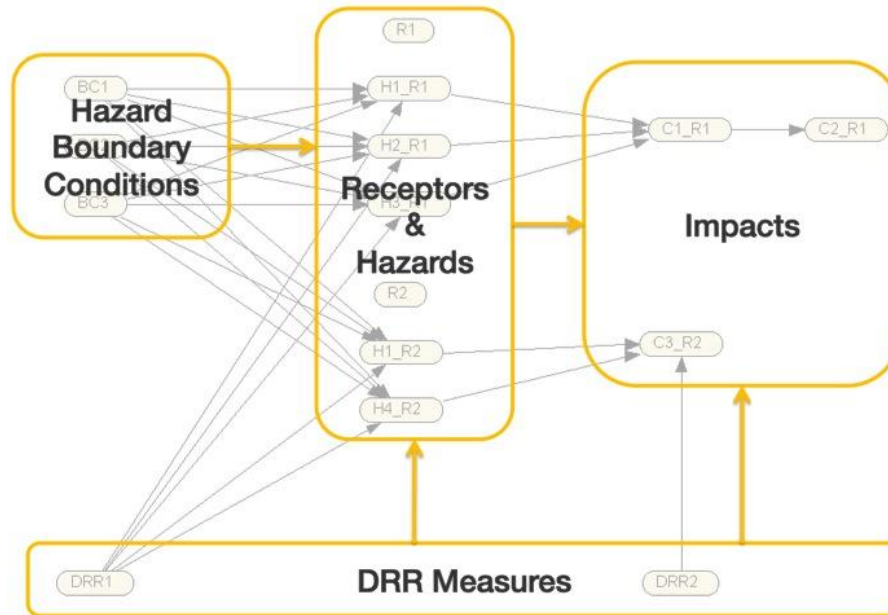


Figure 3.1: Structure of Bayesian DSS (compare with Figure 3.7 to see more detail)

3.1 The Variable Categories

This section explains the type of variables that belong to each box and provides examples.

Hazard boundary conditions (BC)

This category contains hydraulic variables that characterize different storm scenarios. Typically, they coincide with the maxima of the boundary conditions used in the hydro- or morphodynamic models of the coastal FEWS-system. Choices can be e.g. maximum significant wave height and period, peak water level (depicted in Figure 3.2) predominant wave angle or storm duration.

Peak water level [m]		
5 to 5.5	45.0	█
5.5 to 6	30.0	█
6 to 6.5	20.0	█
6.5 to 7	5.00	█
5.68 ± 0.48		

Figure 3.2: Example of BC node

Receptors (R)

In order to represent the spatial distribution of receptors the hot spot sites are subdivided into several areas. The receptor variables indicate the receptors densities in each area of the hot spot. Each receptor type is represented by one variable. The CSPs decide which types are relevant for their site and which they wish to include in the DSS. All receptor types for which vulnerability relationships are available in the CVI library can be included. Examples are (different kinds of) buildings, ecosystems and people. Figure 3.3 gives an example for the receptor type residential houses. Note that 283 houses are 16.7% of the total number of houses.

Houses - Location	
Area 1 (has 283 houses)	16.7
Area 2 (has 759 houses)	44.7
Area 3 (has 383 houses)	22.6
Area 4 (has 273 houses)	16.1
2.38 ± 0.94	

Figure 3.3: Example of R node

Hazards (H)

For each receptor type the DSS contains those hazard variables that are needed to determine impact with the vulnerability relationships. For instance, scour may be a relevant variable for damage to buildings, but less so for loss of life. Other possible choices of variables are e.g. maximum inundation depth, maximum flow velocity, maximum flux or flood duration. Each variable indicates the proportions of receptors that are (or the likelihood that an arbitrary receptor is) threatened by certain hazard intensities, see Figure 3.4 for an example.

Houses - Max. inundation depth	
0m	32.2
0 to 0.5m	44.5
0.5 to 1m	20.3
1 to 2m	2.95
0.308 ± 0.36	

Figure 3.4: Example of H node

Impacts (C¹)

Similar to a hazard variable, an impact variable indicates the proportions of receptors that experience (or the likelihood of an arbitrary receptor to be experiencing) certain impact intensities. Single or multiple hazard variables can be translated into impacts using the transfer functions of the CVI library. Figure 3.5 shows a node for relative damage to residential houses derived from the maximum inundation depth node shown in the previous figure.

¹ “C” stands for consequence, because “I” will be used to abbreviate inundation later on.

Houses - Relative damage	
0%	32.2
0 to 23.5%	44.5
23.5 to 47%	20.3
47 to 50%	2.95
13.8 ± 15	

Figure 3.5: Example for C node

Disaster Risk Reduction Measures

If data is available on the effect of DRR measures they can be included in the DSS. Each measure is represented by one variable, which has at least two states. The first state represents the “do nothing” option, while each additional state represents alternatives. For example the node in Figure 3.6 reflects the DRR measure “increase dike height”. Here, the original dike height is 3m and the proposed DRR options are to increase the dike by either 1m or 2m.

Dike height [m]	
3m	33.3
4m	33.3
5m	33.3

Figure 3.6: Example of DRR node

3.2 Set-up of System Relations

As mentioned above the implementations of the DSS will vary per location. The case study site partners will determine a configuration that best suits their own hot spot area. Figure 3.7 gives a more detailed account of the generic DSS structure than Figure 3.1. It demonstrates the principal of how the individual variables are linked to each other. The variable names are generic and can be considered “place holders” that will be substituted for real variables by the CSPs.

The configuration steps that need to be taken by the partners are outlined in the following sections. A separate configuration manual is not needed (and will not be provided), because the configuration process will be semi-automated and integrated in Delft-FEWS (D3.1) by means of a model adaptor for the BN.

The model adaptor makes use of the Netica API so that the CSPs will not need to interact with the Netica software directly to set up the system. In configuration step 1 - 3, they will provide the necessary information to define the model structure and supply data on receptors and vulnerability relationships, while in configuration step 4 the model adaptor uses this information to set up the system and the data from storm simulations are integrated in the network.

While Figure 3.7 depicts the structure of the generic system, which can be extended for more boundary conditions, receptors, local hazards, impacts or DRR measures, the demonstration DSS in Figure 4.4 can be regarded as “minimal working example”. The latter is published on the project website (“<http://www.risckit.eu/np4/8/>”) together with this deliverable report.

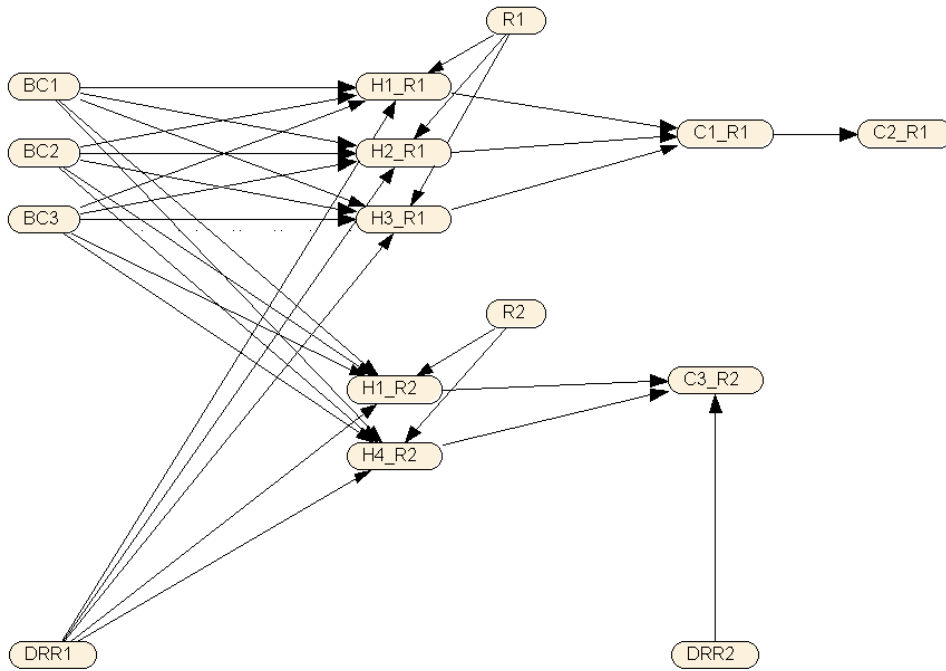


Figure 3.7: Illustration of system relations of the DSS for two receptor types. *BC1*, *BC2* and *BC3* represent hazard boundary conditions; *R1* and *R2* represent receptor densities; *H1_R1*, *H2_R1* and *H3_R1* represent hazards affecting receptor type 1; *H1_R2* and *H4_R2* represent hazards affecting receptor type 2; *C1_R1* and *C2_R1* represent impacts on receptor type 1; *C3_R2* represents an impact on receptor type 2; *DDR1* and *DDR2* represent DRR measures.

3.2.1 Configuration Step 1

The first step deals with the hazard boundary conditions and the DRR plans which affect the hydro- or morphodynamics (e.g. dikes or beach nourishments). The more hazard boundary conditions are incorporated and the finer the discretization (e.g. the example node in Figure 3.2 has four bins and a bin size of 0.5m.), the more storm simulations are necessary. Moreover, each storm scenario needs to be simulated with and without DRR measure in place. In other words, there is a trade-off between simulation effort and detail of the DSS. To find the right configuration, all CSPs are asked to answer the following questions:

1. How many hazard boundary conditions do you wish to include and which ones?
2. For each condition, what are the minimum value, the maximum value and the desired bin size?
3. What is the joint probability distribution of the hazard boundary conditions (in discretized format according to the specified bins)? If this information is not available an assumption can be made, e.g. like in the demonstration BN described in section 4.2.1.
4. How many DRR options do you wish to implement (total of all DRR measures including zero options)?

Based on the answers a minimum number of model runs are determined and hazard boundary conditions for each run are proposed. This process is semi-automated and ensures that a sufficient number of data points fall within each bin. If the number of

model runs does not match the simulation capacity of the CSPs, they can reduce their model complexity or try to expand their computational capacity.

The number of nodes, their names and bins define the BN model and are saved as part of the so-called “ModelDataSetFiles” of Delft-FEWS by the CSPs. When the DSS is started from Delft-FEWS, these files are used to set up the DSS in the software Netica. This will be explained at the last configuration step.

3.2.2 Configuration Step 2

Configuration step 2 focuses on georeferencing the receptors in a specific way. Foregoing is the subdivision of the hot spot into smaller areas. CSPs will decide on the division themselves. It can be based on administrative districts or zones of similar hazards. The latter is recommended and can rest on expert judgment or analysis of flood patterns in multiple storms.

The following information is needed for each individual receptor:

1. What type of receptor is it?
2. To which area of the hot spot does it belong?
3. Which grid point(s) of the hydrodynamic model correspond(s) to its location? (This can for example be done in a GIS software by comparing a shapefile containing the locations of the receptor with the model grid.)

This information will be provided by the case study site partners in a format similar to the one given in Table 3.1 and is saved as part of the ModelDataSetFiles.

Table 3.1: Example table for the DSS specific georeferencing of receptors (format may be subject to change)

Receptor ID	Receptor Type	Grid point ID	Area ID
1	Residential House	[931, 932, 935]	4
...			

3.2.3 Configuration Step 3

In this step the hazard and impact variables are defined. First, CSOs consult the CVI library and identify which vulnerability relationships are applicable at the hot spot level to the receptor types identified in step 2. The relationships define which impacts can be computed for each receptor and which hazard variables are needed as input. For instance, the maximum flux may be needed to compute risk of life. Second, CSOs decide on the desired bin sizes for the impact variables, which typically dictate the bin sizes for the hazard variables as well. Third, the CSOs write a table for each relevant vulnerability relationship as exemplified in Table 3.2.

Again, the information on the choice of variables and bin sizes is stored as part of the ModelDataSetFiles in Delft-FEWS.

Table 3.2: Example format of CVI table for DSS. (This table corresponds to the DSS shown in Figure 1.1)

Inundation depth [m]	Relative damage to residential buildings [%]
0	0
0	0
0.5	23.5
1	47
2	50

3.2.4 Configuration Step 4

When the BN is started from Delft-FEWS, the ModelDataSetFiles are used to set up the DSS in the software Netica. The BN model adapter reads the ModelDateSetFiles and sets up the DSS up for each hot spot (RISC-KIT project Milestone 10 resulting from task 5.2: EWS/DSS systems set up for each case study site). Thereafter, the Bayesian network will be trained and used with data from storm simulations (RISC-KIT project tasks 5.3 and 5.4).

In this context it is also important to adhere to a naming system for the netcdf output files. The data in the netcdf together with its name and the ModelDataSetFiles need to contain all information about the scenario that has been run and if a DRR measure has been applied which one. The names should be as follows

YYYYMMDD_modelIX_mAbinN_mBbinM.netcdf

where mA stands for measure A, mB stands for measure B and N and M are bin numbers. At the same time a table is supplied to the ModelDataSetFiles, which translates mA, mB, ... to their official variable name in binN, binM, ... to their official bin names. *modelIX* can be anything and is chosen by the CSPs. For example, the name 20150331_modelIX_mAbin1_mBbin2.netcdf together with Table 3.3 indicates that the model has been run with a 3m high dike and beach nourishment.

Table 3.3: Example of variables names

Variable code	Variable name	Bin code	States	State titles
mA	Dike height	bin1	3	3m
		bin2	4	4m
		bin3	5	5m
mB	Beach nourishment	bin1	No	No
		bin2	Yes	Yes

4 Demonstration of the Bayesian DSS

This section describes the Bayesian DSS in its simplest form and functionality. The demo DSS has been developed for the site of Zeebrugge, Belgium. This model has been developed in (Jäger et al. 2015), which contains more technical details and can be found in Appendix A.

4.1 Brief Introduction to the Zeebrugge Site

Zeebrugge is a residential town and seaside resort on the North Sea Coast and location of Belgium's second largest harbor. The harbor consists of three parts. The outer and inner ports are shown in



Figure 4.. The outer port is constructed on reclamation ground and is protected by two breakwaters. It has direct access to the sea and sufficiently deep water for roll-on/roll-off and container traffic. Also, LNG vessels moor here. The Pierre Vandamme lock (East/ right in the figure) and the Visart lock (West/ left in figure) connect the tidal outer port with the non-tidal inner port. The docks of the inner port accommodate logistic centers for handling, storage and distribution of e.g. new cars, break bulk or food products. The Baudouin canal (South/ bottom of the figure) leads to the seaport of Bruges. The activities in this part of the port mainly involve bulk and conventional cargoes.

The hot spot is the old town of Zeebrugge, which is mainly residential. It is located between the inner and outer port and the two locks. A scheme can be seen in Figure 4.1.



Figure 4.2: Port and old town of Zeebrugge (© Luchtfotografie Henderyckz)

4.2 Input Data

This section describes the data that has been used to train the Bayesian network for each of the variable categories. A DRR measure is not yet included in the demo version.

4.2.1 Hazard Boundary Conditions

The demo DSS uses a simple synthetic storm climate that consists of 25 realistic storm scenarios. They are combinations of five water level time series with different peak water levels, z , varying between 6.35m and 7.9m and five wave time series with different maximum significant wave heights, h , varying between 5.2m and 6.2m (see Table 4.1). This choice covers a range of storms with return periods from about 100 years to more than 10000 years. For each combination a 46 hours storm has been simulated, which corresponds to three high tides. For simplicity, Z and H are assumed to be independent and have discrete uniform probabilities of occurrence in 100 years, i.e. 20%, where the time frame is chosen arbitrarily. Hence, each storm scenario occurs within the next 100 years with a probability of 4%.

Table 4.1: Wave heights and water levels for the test matrix

Water level [m TAW]	6.35	7.1	7.4	7.8	7.9
Wave height [m]					
5.20					
5.70					
5.90					
6.08					
6.20					

4.2.2 Receptors

The demo DSS is limited to residential houses as receptors. Because the depth-damage curve to be used in the impact estimation (see section 4.2.4) does not distinguish between different kinds of houses, all residential houses are modeled as the same receptor type.

Of relevance to the DSS are the spatial distributions of residential houses, which is captured by dividing the hot spot area into four areas (Figure 4.2) and by counting the number of houses for each area (Table 4.2). This information stems from the cadastral map that has been collected in WP1.

Table 4.2: Number of houses per area

Area	1	2	3	4	Total
# Houses	283	759	383	273	1698

4.2.3 Hazards

Maximum inundation depth patterns are obtained through numeric simulation of storm scenarios. The simulations focus on overtopping North of the old time, close to the Vandamme lock, and do not take into account flooding from the basin in the West. Because NNW is the most critical wave direction for this effect, it is used in all scenarios. The overtopping discharge time series is input for a TELEMAC 2D model, which calculates the dynamic behavior of the flooding on land and from which the maximum inundation for each grid point can be inferred. An example of an inundation map is given in Figure 4.1.

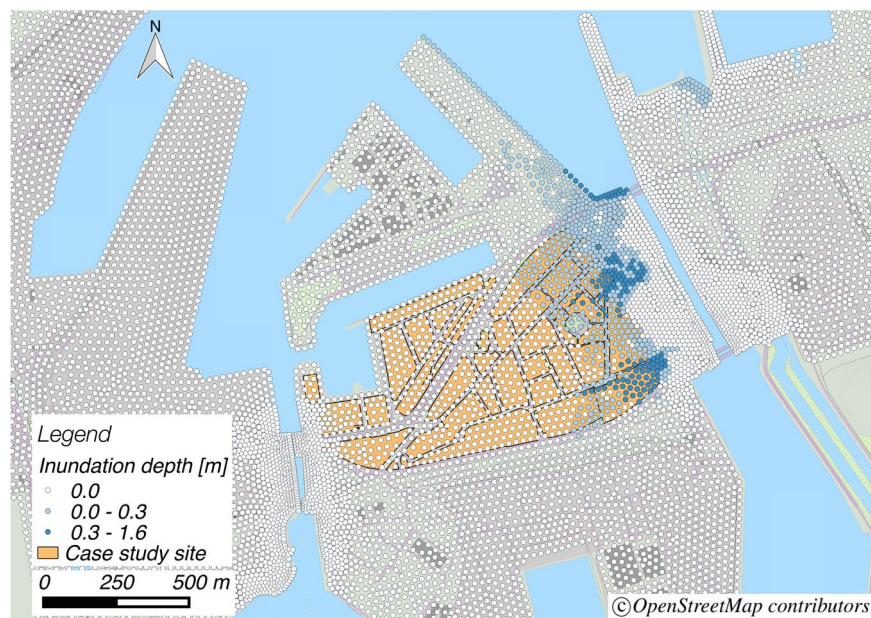


Figure 4.1: Example of a maximum inundation depth map of Zeebrugge

The individual locations of residential houses are known from the data collection in task 1.2 and can be matched to the model grid points. For the Zeebrugge case, this has been done by comparing a shape file with the gridded output of the hydrodynamic

model. For every storm scenario, each house is associated with a maximum inundation depth. For houses that are located in between grid points the inundation depths are interpolated. The maximum inundation depths never exceeded 2m for any house.

4.2.4 Impacts

The relative damage per house (in terms of the maximum possible damage), is calculated with the depth damage curve for residential houses in Flanders, Belgium, by (Vanneuville 2003). This curve is also part of the CVI library (Viavattene et al. 2015). It is depicted in Figure 4.3 and provides a functional relationship between maximum inundation depth and relative damage. If an average house value is known, an indication can be given for the absolute damage per house. For this demo DSS, an average house value of 100,000€ is assumed. Because the maximum inundation depths never exceeded 2m, the maximum attainable relative and absolute damage in this demo case are 50% and 50,000€, respectively.

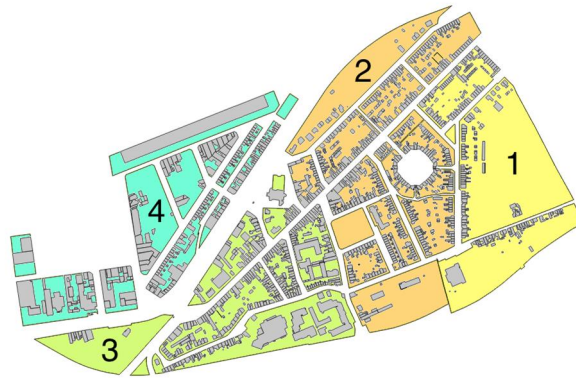


Figure 4.2: Residential houses and subdivision in areas

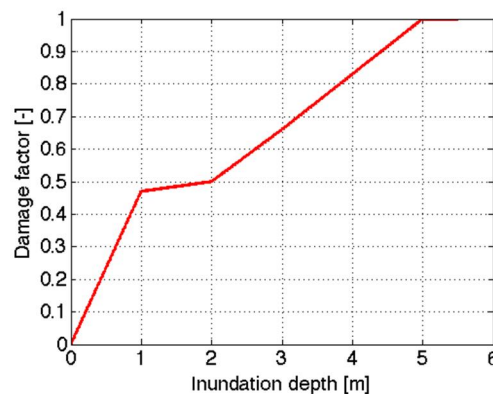


Figure 4.3: Depth-damage curve for Flanders, Belgium

4.3 Examples of Use

Before turning to specific examples of use, it is important to understand that the reliability of the Bayesian DSS results leans on the quality of the underlying data and models. Note especially that while the demo DSS is based on the hot spot Zeebrugge, it

makes a strong simplifying assumption about the storm scenarios and does not reflect real storm climate at the site.

4.3.1 The Basic State of the DSS

Figure 4.4 shows the Bayesian DSS in its basic state giving a default prediction. It has seven probabilistic nodes, *Peak water level* (Z), *Max. significant wave height* (H), *Area* (B), *Maximum inundation depth* (I), *Number of Houses* (N), *Relative damage* (D) and *Absolute damage* (D_Euro), and one constant node, *Average house value* (v). Each probabilistic node shows the states that the variable can be in, the probabilities of being in these states as well as the mean and standard deviation, while the value of the constant node is fixed. The nodes are connected with arrows, which indicate a probabilistic influence on the node at the arrowhead from the node at the other end. Most of the time these influences can be regarded as cause and effect relationships.

Z and H characterize a storm event. Based on our assumption (cf. section 4.2.1), a future storm has any of the five peak water levels with a probability of 20% and independently thereof any of the five maximum significant wave heights with a probability of 20%.

B and N summarize the case study site. B indicates that if we randomly select a house, just like drawing a ball from an urn, the probability of selecting a house that is within area 1 is 16.7%, because this percentage of houses are located there. To be precise 283 houses out of a total of 1698 are situated in area 1 according to N . Similarly 44.7% (759) are in area 2, 22.6% (383) are in area 3, and 16.1% (273) are in area 4. Naturally, B and N are linked, as expressed by the arrow. But even more N is completely determined by B , in other words their relationship is deterministic, which is indicated by the slightly darker shade of this node.

I is an indicator for the hazard intensity at the hot spot. It is influenced by Z and H , because greater storm magnitudes are more likely to cause flooding and by B , because the different areas in the hot spot are not equally prone to flooding. The incoming arrows indicate this. In the base state of the DSS node I represents, intuitively speaking, the average proportion of houses that are inundated up to a certain level²: 32.3% of the houses are expected to never experience inundation, 44.5% of the houses are up to 0.5m inundated, 20.3% of the houses are up to 1m inundated, and 2.95% of the houses are up to 2m inundated. This is equivalent to saying that an arbitrary house in the hot spot area has a chance of never being inundated of 32.3% and so on.

D and D_Euro are translations of inundation depths to relative damage estimates based on the damage curve and absolute damage estimates using an average house value of 100,000€³, respectively. The relationship is deterministic and, hence, the nodes are darker. By modifying node v the average value can easily be changed.

²If we consider, for example, two equally likely storm scenarios, the first one causing 1% of the houses to inundate at least 2m and the second one 5%, then the average proportion would be $0.5 \cdot 1\% + 0.5 \cdot 5\% = 3\%$.

³ Note that Netica uses the dot both as decimal and thousands separator which can be confusing.

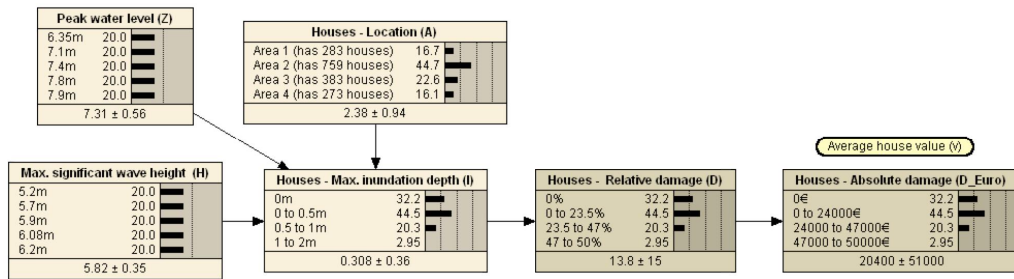
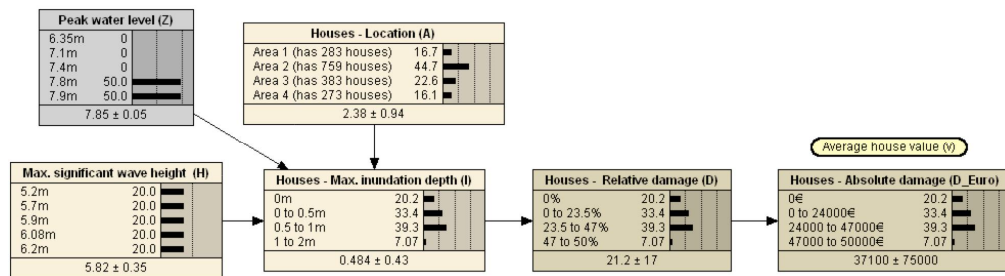


Figure 4.4: Expected storm impacts in the hot spot area

4.3.2 Predicting Impacts

Making impact predictions in real time is relevant during the event phase (from several days before the storm until it has passed) as well as the planning phase. Some examples for the event phase are given below, but the use of the DSS is not limited to them. In the planning phase this mode can for instance be used to explore and discuss the impacts of different storm scenarios and DRR measures during round table discussions.

A few days ahead the magnitude of a storm, which is estimated with large-scale meteorological models, can be somewhat uncertain. For instance, high water levels reaching at least 7.8m are expected, but estimates on wave heights are not available. In this case the DSS can be “conditioned” or “updated” on what is known, which is shown in Figure 4.5. Because the demo storm climate contains only two peak water level scenarios greater or equal to 7.8m, the user has set the distribution of Z to a 50% chance for $Z=7.8m$ and $Z=7.9m$ each. This choice means that he/she is entirely uncertain as to if one or the other will occur. As no additional information is available for wave heights, node H maintains his prior distribution. Updating on one node changes the probability distributions of all nodes that are influenced by it. Here, updating H changes I, D, and D_Euro. Compared to the default prediction, the probabilities for higher inundation depths and damages have increased.


 Figure 4.5: Predicting impacts in the hot spot area of a storm with $Z \geq 7.8m$ and uncertain wave height

If, towards the peak of the storm event, the user gains more accurate knowledge on the hazard boundary conditions, he/she can update the corresponding nodes to reduce the uncertainties. Figure 4.6 exemplifies this for $Z=7.8m$ and $H=5.9m$, where the peak water level and maximum wave height are more constrained. The BN reacts to this decreased uncertainty in the hazard boundary conditions by shifting the distribution of node I to lower values. On the contrary, the distribution is shifted to higher values, when conditioning on $Z = 7.9m$ instead of $Z=7.8m$ (not illustrated in this

report). The expected impacts are less severe than in the previous, more uncertain scenario.

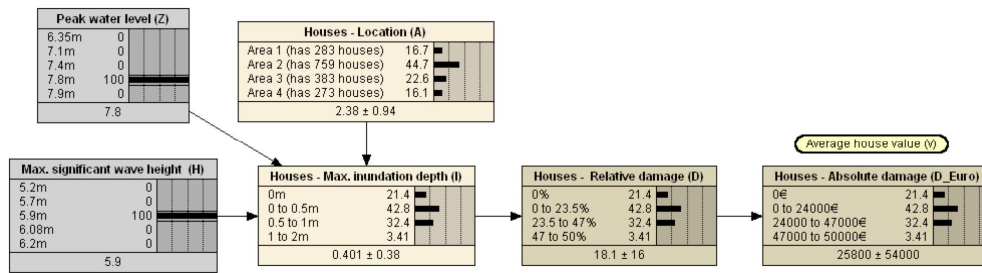


Figure 4.6: Predicting impacts in the hot spot area of a storm with $Z=7.8\text{m}$ and $H=5.9\text{m}$

By conditioning on an area, i.e. node A, the user can obtain more specific information on the individual areas within the hotspot. In Figure 4.7 the demo DSS is conditioned on the same storm as before and on area 1. It demonstrates that houses in area 1 are especially at risk. For instance, 20.5% of the houses in this area are expected to be inundated up to 2m, whereas the proportion of all houses in the hot spot area is only 3.41%.

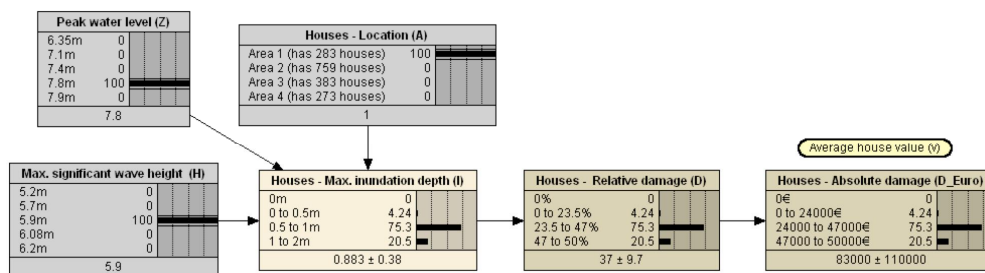


Figure 4.7: Predicting impacts of a storm with $Z=7.8\text{m}$ and $H=5.9\text{m}$ specifically for area 1

4.3.3 Diagnosing Storms

Especially in the planning phase, it can be useful to diagnose the conditions under which impact occurs or does not occur. The DSS in Figure 4.8 is conditioned on the topmost absolute damage, which coincides with the highest states of the relative damage and maximum inundation depth variables, to understand where in the hot spot area and at which storm magnitudes it occurs. A house is most likely to suffer this high damage if it is located in area 1 and the more severe the storm scenario is, foremost the peak water level. The almost uniform distribution of the wave height suggests that this hazard boundary condition is not important for this case. Moreover, houses in area 4 will never incur maximum damage under any storm and no house at all will incur maximum damage, if the peak water level does not exceed 6.35m.

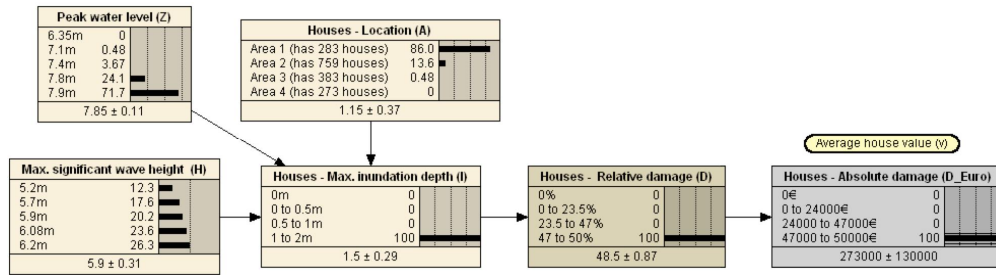


Figure 4.8: Diagnosing which storm scenarios can cause maximum damage

5 Conclusion

This report presented RISC-KIT deliverable 3.3, the Bayesian DSS, which can link offshore storm conditions to hazards and impacts at selected hot spot areas. The main advantage of such a system is threefold. First, the DSS can make impact predictions in real time and it can thus support emergency managers in urgent decision-making. In contrast a new simulation with a hydrodynamic process model would be computationally expensive and time consuming. Second, the DSS enables planners to test DRR measures for robustness against a variety of storm scenarios, as long as these measures have been considered in the model set up. Third, the DSS can diagnose under which storm conditions and where on the site the highest, medium or no damages occur.

The framework of the DSS is generic and suitable for any case study site, also for sites outside the RISC-KIT project. The underlying modeling approach is using a Bayesian network (BN). These types of models describe the relationships between system variables in terms of probabilities and have an interactive user interface which resembles a casual diagram.

RISC-KIT's DSS links five categories of variables to each other:

1. Hazard boundary conditions, e.g. peak water levels during a storm;
2. Spatial distributions of different receptor types, e.g. of residential houses;
3. Spatial distributions of hazard intensities, e.g. maximum inundation depths at the locations of residential houses;
4. Spatial distribution of impact intensities, e.g. relative damages to residential houses; and
5. DRR measures, e.g. alternative height options for a new dike.

In order to set up the DSS, CSPs need:

- A discrete joint probability distribution of hazard boundary conditions at their site;
- Detailed land use information of their site;
- Vulnerability relationships for the receptor types of interest, i.e. transfer functions from hazard to impacts such as depth-damage curves; and
- Sufficient computational capacity to simulate multiple storms and their hydrodynamic and morphological effects on the hinterland.

As illustrated by the demo DSS, a fully functional system can be built for a single receptor type with very few data on the storm climate. In general, one can make a start with such a limited system for any case study site and extend the network as soon as more data on hazard conditions, receptors or vulnerabilities becomes available.

References

- Garrote, L., Molina, M. & Mediero, L., 2007. Probabilistic forecasts using Bayesian Networks calibrated with deterministic rainfall-runoff models. *Extreme Hydrological Events: New Concepts for Security*.
- Gutierrez, B.T., Plant, N.G. & Thieler, E.R., 2011. A Bayesian network to predict coastal vulnerability to sea level rise. *Journal of Geophysical Research*, 116(F2), p.F02009.
- Hapke, C. & Plant, N., 2010. Predicting coastal cliff erosion using a Bayesian probabilistic model. *Marine Geology*, 278(1-4), pp.140–149.
- Den Heijer, C., 2013. *The role of bathymetry, wave obliquity and coastal curvature in dune erosion prediction*. Delft University of Technology. Available at: <http://discover.tudelft.nl:8888/recordview/view?recordId=TUD:oai:tudelft.nl:uid:824df068-8046-414c-a1cc-7d159718918e&language=nl>
- Henriksen, H.J., Rasmussen, P. & Brandt, G., 2007. Public participation modelling using Bayesian networks in management of groundwater contamination. *Environmental Modelling & Software*, 22(8), pp.1101–1113.
- Jäger, W.S., Bolle, A., Den Heijer, C. & Hanea, A.M., 2015. A Bayesian Network Approach to Coastal Storm Impact Modelling. *12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12*.
- Jonkman, S.N., Bočkarjova, M. & Kok, M., 2008. Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecological Economics*, 66(1), pp.77–90.
- Norsys, 2014. Netica. Available at: www.norsys.com.
- Rotmans, J., Asselt, M. & Bruin, A. 1994. *Global change and sustainable development: a modelling perspective for the next decade*, Bilthoven. The Netherlands.
- Vanneuville, W., 2003. Bepalen van de gevolgen van overstromingen door numerieke cartografische modellen. *Colloquium Numerieke oppervlaktewater modellering-* Available at: <http://www.vliz.be/imisdocs/publications/02/236802.pdf> [Accessed December 3, 2014].
- Viavattene, C., Priest, S., Micou, P., Owen, D., Parker, D., Penning-Roswell, P.C., Spencer, T., Möller, I., McIvor, A. & Jiminéz, J., 2015. *RISC-KIT Deliverable 2.2. Coastal Vulnerability Indicator Library. Guidance Document*. Available at: <http://www.risckit.eu/np4/8/>
- Vogel, K., Riggelsen, C., Scherbaum, F., Schröter, K., Kreibich, H. & Merz, B., 2013. Challenges for bayesian network learning in a flood damage assessment application. In *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures - Proceedings of the 11th International Conference on Structural Safety and Reliability, ICOSSAR 2013*. pp. 3123–3130.

Vogel, K., Riggelsen, C., Merz, B., Kreibich, H. & Scherbaum, F., 2012. Flood damage and influencing factors: A Bayesian network perspective. In *Proceedings of the 6th European Workshop on Probabilistic Graphical Models, PGM 2012*. pp. 347–354.

Appendix A

*12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015*

A Bayesian Network Approach to Coastal Storm Impact Modeling

Wiebke S. Jäger

PhD Student, Department of Hydraulic Engineering, Delft University of Technology, The Netherlands

C. (Kees) den Heijer

Researcher, Department of Hydraulic Engineering, Delft University of Technology, The Netherlands

Annelies Bolle

Engineer Advisor, International Marine & Dredging Consultants, Antwerp, Belgium

Anca M. Hanea

Research Fellow, Centre of Excellence for Biosecurity Risk Analysis, The University of Melbourne, Australia

ABSTRACT: In this paper we develop a Bayesian network (BN) that relates offshore storm conditions to their accompanying flood characteristics and damages to residential buildings, following on the trend of integrated flood impact modeling. It is based on data from hydrodynamic storm simulations, information on land use and a depth-damage curve. The approach can easily be applied to any site. We have chosen the Belgian village Zeebrugge as a case study, although we use a simplified storm climate. The BN can predict spatially varying inundation depths and building damages for specific storm scenarios and diagnose under which storm conditions and where on the site the highest impacts occur.

Coastal zones are very attractive to develop social, industrial and recreational infrastructure. They have rich natural resources, impressive landscapes and excellent navigation possibilities. In 2003 an estimated 23% of the world population lived in low-lying¹ coastal areas (Small and Nicholls, 2003). The ongoing trend is a disproportionately rapid expansion of economic activity, urban areas and tourist resorts. At the same time coasts are affected by various hydro-meteorological phenomena, such as wind, waves, tides and precipitation which can reach extraordinary magnitudes during storm surges, hurricanes, typhoons or tsunamis. Resulting floods threaten people, cause land loss, damage property, infrastructure and ecological habitats, and destabilize economic activities.

While coastal zone managers cannot influence

the occurrence of extreme events, they can apply measures to reduce the accompanying risks in the short, middle and long term.

Researchers across many disciplines are dedicated to developing methodologies that identify risks and to helping decision makers design effective risk reduction plans. They apply numerical hydrodynamic process models to assess the natural coastal response and the extent of flooding due to storms, e.g. XBeach (Roelvink et al., 2009), TELEMAC (Hervouet, 2000) or MIKE21 (Warren and Bach, 1992), and use separate models to estimate economic, political, social, cultural, environmental and health-related impacts. Comprehensive reviews have been written on assessment methods for economic damage (Merz et al., 2010), on flood-related health impacts (Ahern et al., 2005; Hajat et al., 2005), and on estimation methods for loss of life (Jonkman et al., 2008b).

¹By low-lying coastal areas we mean areas both within 100km of the shoreline and less than 100m above sea level.

Ongoing research is on the one hand directed towards improving the various consequence models and comparing them with each other (e.g. Schröter et al., 2014). On the other hand there is a trend towards integrating the separate modeling approaches into a homogeneous framework. A GIS-based approach to describe a spatially varying flood hazard and associated estimates of direct physical damages to various objects, indirect economic damage and the loss of life has been proposed by (Jonkman et al., 2008a).

We continue on the trend of model integration. While Jonkman's model presents the results of one typical low probability-high impact flood scenario with the help of maps, we attempt to compile impact estimates of many different storm scenarios in a discrete Bayesian network (BN). BNs are graphical models that describe system relations in probabilistic terms. They can handle various sources and types of data enabling us to combine information on the topography and assets of the potentially affected area with simulation data of flood scenarios and damage estimations from single discipline models.

More precisely, we relate flood impacts not only to flood characteristics, but also to offshore storm conditions, such as peak water level and maximum significant wave height. This has two advantages. First, the BN can make spatially varying consequence predictions for an impending storm in real-time and it can thus support emergency managers in urgent decision making. In contrast a new simulation with a hydrodynamic process model would be computationally expensive and time consuming. Second, the BN can facilitate round table discussions of e.g. planners. It enables them to instantly compare the effect of risk reduction measures for a variety of storm scenarios, as long as these measure have been included in the model set up.

In this article we develop and describe a prototype of this BN and apply it to a case study site. We use the implementation of the software Netica (Norsys, 2014).

Our study site is the old town of Zeebrugge, located on the North Sea coast of Belgium, which is mainly residential. The storm scenarios, however,

are synthetic due to data limitations. While the network structure can be applied to any site, the quantitative component is site specific. It implicitly contains site topology or other unique features, such as flood defenses, which determine if flooding occurs and, if so, the spatial extent of the flooding.

As a first step, we focus on the prediction of physical damage to residential buildings that have been in direct contact with floodwater. We plan to add other damages to the network later on in the same manner.

1. BASIC CONCEPTS OF DISCRETE BAYESIAN NETWORKS

BNs have been applied numerous times as tools for decision-making under uncertainty. Henriksen et al. (2007) conclude that they are very valuable for negotiations and discussions between managers, experts, stakeholders and representatives of the general public, among others, because they are transparent and flexible models. In the context of floods, Garrote et al. (2007) combine BNs and deterministic rain run-off models to forecast flooding, and Vogel et al. (2012, 2013) use BNs to estimate damages resulting from river floods. In coastal environments they have been applied to predict erosion and shoreline retreat (Den Heijer, 2013; Gutierrez et al., 2011; Hapke and Plant, 2010). At the moment of writing we are not aware of applications to coastal flooding.

Discrete BNs are probabilistic graphical models that represent a high-dimensional probability distribution over a finite set of discrete variables X_1, X_2, \dots, X_n (Pearl, 1988; Jensen, 1996). The core of the representation is a directed acyclic graph (DAG) whose nodes represent random variables and whose arcs indicate a direct influence from "parent node" to "child node". Because the graph structure stipulates that each variable is conditionally independent of all predecessors given its parents, the joint distribution can be economically factorized using the chain rule:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i)), \quad (1)$$

where $Pa(X_i)$ denotes the set of parent nodes of X_i in the graph. Together, the DAG and a specification

12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015

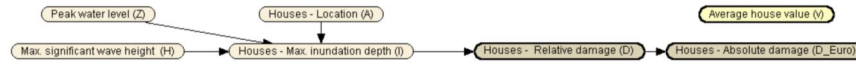


Figure 1: BN structure

of $P(X_i | Pa(X_i))$, for $i = 1, \dots, n$, or $P(X_i)$ in case of no parents, uniquely specify a joint distribution over X_1, X_2, \dots, X_n .

A main use of BNs is updating: Once new evidence on one or more variables is obtained, the effect can be propagated through the network using Bayes' theorem. Evidence can be propagated both forward and backward, which enables predictive as well as diagnostic reasoning.

2. DESIGN OF THE COASTAL STORM IMPACT MODEL

This section motivates and describes the design of the BN, i.e. the definition of random variables and the structure as shown in Figure 1. The parent nodes of the network characterize the hydrodynamic forcing, i.e. peak water level and maximum significant wave height, and the location of buildings in terms of areas. They influence spatially varying inundation depths, which in turn are translated into relative and absolute building damage with a simple depth-damage-curve and by assuming an average building value.

2.1. Storm Scenarios

Extreme hydraulic conditions are commonly characterized in terms of peak water level, maximum significant wave height and period, predominant wave angle, and storm duration. Naturally, data on these hydraulic variables is rare. Since recently, copulas are being used to represent their multivariate distributions at offshore locations (e.g. De Waal and van Gelder, 2005; Corbella and Stretch, 2013; Li et al., 2014). However, the hydrodynamic process model requires near-shore conditions as input. The transformation of the joint distribution of hydraulic variables from offshore to near-shore is complex and has, to our knowledge, been rarely described in the literature up to now (Bolle et al., 2014; Leyssen et al., 2013). Also for our case study site this information is not yet available. Therefore,

we assume a simplistic synthetic storm climate with the intention to extend the model in the future.

This storm climate consists of 25 realistic storm scenarios. They are combinations of five water level time series with different peak water levels, z , varying between 6.35m and 7.9m and five wave time series with different maximum significant wave heights, h , varying between 5.2m and 6.2m. This choice covers a range of storms with return periods from about 100 years to more than 10.000 years. For each combination a 46 hours storm is simulated, which corresponds to three high tides.

For simplicity we assume Z and H to be independent random variables (see the two left nodes in Figure 5) with discrete uniform probabilities of occurrence in 100 years, i.e. 20%, where the time frame is chosen arbitrarily. Hence, each storm climate scenario occurs within the next 100 years with a probability of 4%. This is a strong assumption and does not reflect the storm climate at Zeebrugge realistically. However, this assumption is unproblematic for applications in real-time decision making, because Z and H will be fixed to the (forecasted) values of the impending or occurring storm.

2.2. Residential Buildings on the Site

The case study site is divided into four areas, as illustrated by Figure 2. The parcels correspond to administrative districts, but other division criteria are possible as well, e.g. based on topography. How many residential houses lie within each area can be extracted from a cadastral map and is listed in Table 1. We introduce a node A to the network to represent the location of an arbitrary residential building. If we randomly select a house, just like drawing a ball from an urn, the probability that it is within area a is proportional to the number of houses in a . This defines the probability distribution A . Note that it is independent of the storm scenario.

12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
 Vancouver, Canada, July 12-15, 2015



Figure 2: Residential buildings and areas at case study site

Table 1: Number of residential buildings per area

Area	1	2	3	4
Number of buildings	283	759	383	273

2.3. Maximum Inundation Depth and Damage

We obtain maximum inundation patterns through numeric simulation of storm scenarios. The simulations focus on overtopping North of the old town and do not take into account flooding from the basin in the West. Because NNW is the most critical wave direction for this effect, it is used in all scenarios. The overtopping discharge time series is input for a TELEMAC 2D model, which calculates the dynamic behavior of the flooding on land and from which the maximum inundation depth can be inferred for each grip point, and by interpolation for each house. An example is given in Figure 3.

We introduce a node *maximum inundation depth* (of an arbitrary house under an arbitrary storm scenario, more details in section 3), I , to the BN which is Z , H and A 's child, and discretize its distribution into four intervals $\{i_1, i_2, i_3, i_4\} = \{[0m], (0m, 0.5m], (0.5m, 1m], (1m, 2m]\}$. Then the conditional probabilities can be specified as

$$P(I = i_j | A = a, Z = z, H = h) = \frac{n_{i_j, a, z, h}}{n_a} \quad (2)$$

for $j = 1 \dots 4$, where n_a is the number of houses in area a and $n_{a, i_j, z, h}$ is the number of houses in area a with maximum inundation depth i_j under storm scenario $\{z, h\}$.

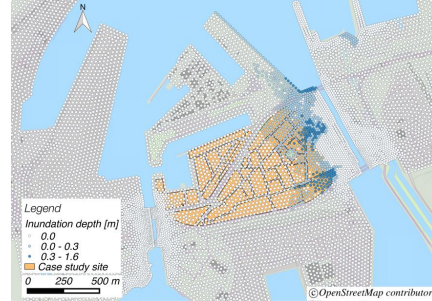


Figure 3: Example of an Inundation map for Zeebrugge and surroundings. The North Sea is to the North.

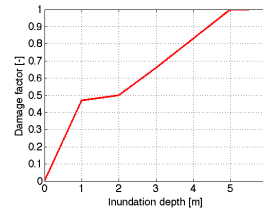


Figure 4: Depth-damage curve for Flanders, Belgium

The relative damage per house (in terms of maximum possible damage), d , is calculated with the depth-damage curve for residential houses in Flanders, Belgium, by Vanneuille et al. (2006). This curve is depicted in Figure 4 and provides a functional relationship between I and D . Assuming an average value per house, v , an indication can be given for the absolute damage per house in \mathcal{B} , $d_{\mathcal{B}}$. As an example, the BN here has $v = 100000 \text{ €}$ and is represented as a constant node in the figures.²

3. INTERPRETATION OF THE COASTAL STORM IMPACT MODEL

The resulting BN is shown in Figure 5. At its heart is node I . This node can be interpreted in two ways, which are described in separate sections below. The same applies to the two damage nodes, which are merely translations from the maximum inundation

²Note that constant nodes do not have arcs in Netica

12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
 Vancouver, Canada, July 12-15, 2015

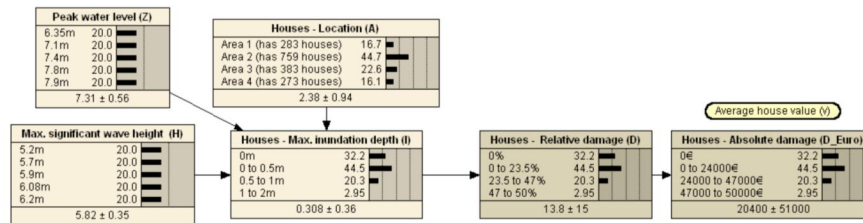


Figure 5: BN with prior distributions

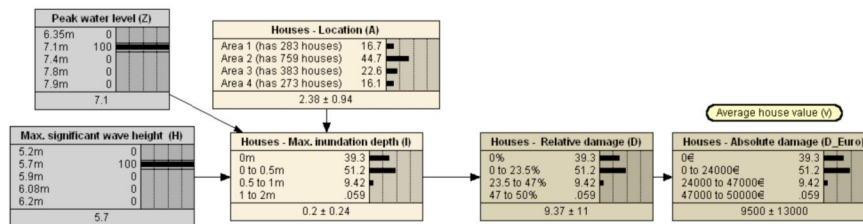


Figure 6: Updated BN for $Z = 7.1m$ and $H = 5.7m$

depth to units of impact and will therefore not be discussed individually.

3.1. In General: (Conditional) Probabilities

One possibility, the conventional one, is to interpret node I as the maximum inundation depth of a single house. The prior probability distribution of this node (Figure 5) represents the uncertainty about the true maximum inundation value for an arbitrary house whose location at the site is unknown as well as the storm scenario by which they are affected. We can reduce this uncertainty by conditioning, for example, on $Z = 7.1m$ and $H = 5.7m$. Now the distribution represents the uncertainty in the inundation for a house at an unknown location due to the storm with peak water level $7.1m$ and maximum significant wave height $5.7m$. This is shown in Figure 6. By conditioning on $A = 2$ (Figure 7) we obtain the distribution for a house under this storm that is located in area 2. It is important to realize that the uncertainty does not stem from the physical modeling. It arises, because the various houses in area 2 experience different inundation

depths. In that sense it reflects the unknown exact location.

Alternatively, we can reason backward, e.g. by conditioning on $D_{€} = [47000€, 50000€]$ (Figure 8) to understand the conditions due to which top-most damage occurs. A house is most likely to suffer maximum damage if it is located in area 1 and the more severe the storm climate is, foremost the peak water level. Moreover, no house in area 4 will incur maximum damage and no house at all will incur maximum damage, if peak water level is $6.35m$.

3.2. In the Special Case of Forward Reasoning: (Conditional) Expectations

Unless we reason backward, we can interpret node I in an alternative manner. Besides representing one random variable with four possible states, it represents four random variables associated with (conditional) expectations.

Looking back at equation (2), we notice that the right hand side is not just a conditional probability. It is also simply the *fractions of houses in area a with maximum inundation depth i_j under storm*

12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015

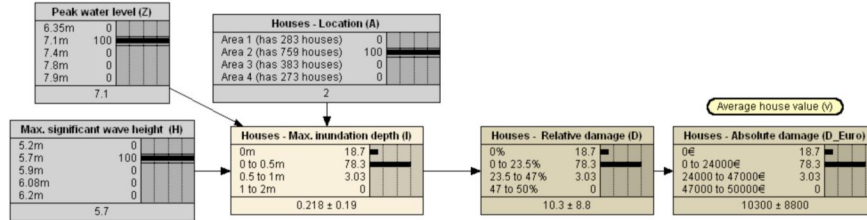


Figure 7: Updated BN for $B = 2$, $Z = 7.1m$ and $H = 5.7m$

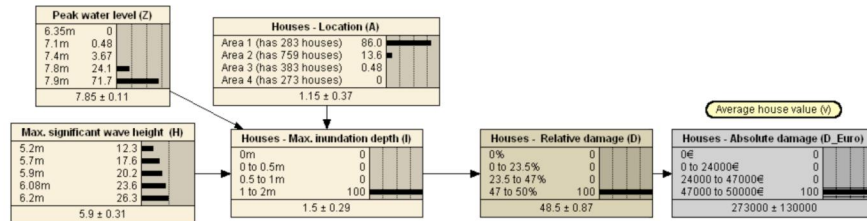


Figure 8: Updated BN for $D_e = [47000€, 50000€]$

scenario $\{z, h\}$. If we define four new random variables, the fractions of houses that are inundated by i_j , F_j , $j = 1 \dots 4$, then

$$\{F_j | A = a, Z = z, H = h\} = \frac{n_{i_j, a, z, h}}{n_a}. \quad (3)$$

Figure 7 indicates that area 2 has 759 houses, of which 18.7% are not flooded, 78.3% are inundated up to 0.5m, and 3.03% are inundated between 0.5m and 1m.

If we remove evidence for node A, as in Figure 6, Netica uses the law of total probability and computes the distribution of I with its conditional probability table and the marginal distribution of A:

$$\begin{aligned} P(I = i_j | Z = z, H = h) & \quad (4) \\ &= \sum_{a=1}^4 P(I = i_j | A = a, Z = z, H = h) \cdot P(A = a). \end{aligned}$$

This equals

$$\sum_{a=1}^4 \frac{n_{a, i_j, z, h}}{n_a} \cdot P(A = a) \quad (5)$$

and, using that (3) is a constant,

$$\sum_{a=1}^4 \mathbb{E}[F_j | A = a, Z = z, H = h] \cdot P(A = a). \quad (6)$$

This can be rewritten, using the law of total expectation, as

$$\mathbb{E}[F_j | Z = z, H = h]. \quad (7)$$

Hence, each bin j in Node I also represents the conditional expectation of the fraction of houses with maximum inundation depth i_j over all areas given storm scenario $\{z, h\}$. Note that because $P(A = a)$ is proportional to the number of houses in area a , this coincides with

$$\{F_j | Z = z, H = h\} = n_{i_j, z, h}, \quad (8)$$

where $n_{i_j, z, h}$ is the total number of houses with maximum inundation depth i_j under storm scenario $\{z, h\}$. This reasoning with conditional expectations can easily be extended to different conditioning sets. For example, for the network in Figure 5 we have

$$\mathbb{E}[F_j]. \quad (9)$$

12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015

Thus, the bins $j = 1 \dots 4$ of node I in the BN with prior probabilities provide a summary of the distribution of each F_j in terms of the expected value.

We can explore F_j 's distribution by conditioning on storm scenarios $\{z, h\}$: we find the values f_j that correspond to the probability $P(Z = z, H = h) = P(Z = z) \cdot P(H = h)$. Admittedly, the usefulness of this information depends on how realistically the storm climate is quantified. In our case it is completely synthetic. Moreover, we can zoom in and out in space: we can obtain information per area or for the entire case study site by conditioning node A , or not.

4. CONCLUSION

In this article we proposed a BN approach to coastal flood impact modeling. The BN links various offshore storm conditions to flood depths and building damages.

To understand the implications of a specific storm scenario it seems very useful to interpret the bins of nodes I , D and D_E as the (conditional) expectation of the fraction of houses that are inundated by i_j , have relative damage d_j or absolute damage d_{E_j} , respectively. Conditioned on a scenario the BN indicates corresponding spatially varying inundation depths and building damages.

Because we distinguish just four areas, the spatial detail is significantly less than the one of an inundation or damage map: We can predict how many houses within an area have a specific flood depth, but we do not know which ones. If desired, the resolution can be increased by adding bins to node A , the area in which a house is located.

Nevertheless this BN approach has a couple of advantages over map-based approaches. The variables of interest can be seen simultaneously, while one map per variable is needed. They can easily be compared across storms, by conditioning on different water and wave heights, or across areas, by conditioning on areas. Admittedly, as yet, we have treated only maximum inundation depth, relative damage and absolute damage, but this quality grows when more flood consequences are integrated. Additionally, the BN presents the exact percentage of inundated and damaged houses, an information which is not apparent after a quick glance on

a map.

We can also interpret the BN results in the conventional way: I , D and D_E are the maximum inundation depth, relative damage and absolute damage for a single house. Then we can diagnose under which conditions the highest flood depth and damage occur, which may help decision makers to design risk reduction measures.

Finally, we would like to point out that the BN can be built gradually and improved continuously, according to data availabilities and simulation capacities. Naturally its prediction and diagnosis value depends on the quality of underlying models. Here it has to be noted that many consequence models, including damage curves, are "simple approaches [...] to complex processes [...]" (Merz et al., 2010) and are associated with large, and often unknown, model uncertainties.

In the future, we aim to use a realistic joint probability distribution of hydraulic storm conditions, being represented by continuous nodes³ and possibly including storm duration and wave period or angle as additional variables. Then, the BN could give an indication of the flood risk to residential buildings, as it links the damage extend to its probability of occurrence. And again, we have the ambition to extend the approach to a wider range of damage categories. Another step could be to take model uncertainties into account, for both the damage model as well as the hydraulic model.

ACKNOWLEDGMENTS

This work is funded by the RISC-KIT project (EU contract 603458, www.risckit.eu). Special thanks go to Ap van Dongeren, Robert McCall and Oswaldo Morales Nápoles for the fruitful discussions in preparation of this article.

5. REFERENCES

- Ahern, M., Kovats, R. S., Wilkinson, P., Few, R., and Matthies, F. (2005). "Global health impacts of floods: epidemiologic evidence." *Epidemiologic reviews*, 27(1), 36–46.
- Bolle, A., Blanckaert, J., and Leyssen, G. (2014). "Une méthode probabiliste pour la détermination de

³We mean continuous according to the definition of Netica.

*12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015*

- l'événement de référence et ses conséquences pour les ppl et ppli." *Congrès SHF: Modélisation opérationnelle de grands transitoires hydrauliques*.
- Corbella, S. and Stretch, D. D. (2013). "Simulating a multivariate sea storm using Archimedean copulas." *Coastal Engineering*, 76, 68–78.
- De Waal, D. and van Gelder, P. (2005). "Modelling of extreme wave heights and periods through copulas." *Extremes*, 8(4), 345–356.
- Den Heijer, C. (2013). "The role of bathymetry, wave obliquity and coastal curvature in dune erosion prediction." Doctoral dissertation, Delft University of Technology, Delft, The Netherlands.
- Garrote, L., Molina, M., and Mediero, L. (2007). "Probabilistic forecasts using bayesian networks calibrated with deterministic rainfall-runoff models." *Extreme Hydrological Events: New Concepts for Security*, Springer, 173–183.
- Gutiérrez, B. T., Plant, N. G., and Thieler, E. R. (2011). "A Bayesian network to predict coastal vulnerability to sea level rise." *Journal of Geophysical Research*, 116(F2), F02009.
- Hajat, S., Ebi, K., Kovats, R., Menne, B., Edwards, S., and Haines, A. (2005). "The human health consequences of flooding in europe: a review." *Extreme weather events and public health responses*, Springer, 185–196.
- Hapke, C. and Plant, N. (2010). "Predicting coastal cliff erosion using a Bayesian probabilistic model." *Marine Geology*, 278(1–4), 140–149.
- Henriksen, H. J., Rasmussen, P., Brandt, G., von Bülow, D., and Jensen, F. V. (2007). "Public participation modelling using Bayesian networks in management of groundwater contamination." *Environmental Modelling & Software*, 22(8), 1101–1113.
- Hervouet, J.-M. (2000). "TELEMAC modelling system: an overview." *Hydrological Processes*, 14(13), 2209–2210.
- Jensen, F. (1996). *An introduction to Bayesian networks*. University College London Press, London.
- Jonkman, S. N., Bočkarjova, M., Kok, M., and Bernardini, P. (2008a). "Integrated hydrodynamic and economic modelling of flood damage in the Netherlands." *Ecological Economics*, 66(1), 77–90.
- Jonkman, S. N., Vrijling, J. K., and Vrouwenvelder, A. C. W. M. (2008b). "Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method." *Natural Hazards*, 46(3), 353–389.
- Leyssen, G., Merceles, P., De Schoesitter, P., and Blanckaert, J. (2013). "Generation of multivariate near shore extreme wave conditions based on an extreme value copula for offshore boundary conditions." *EGU General Assembly Conference Abstracts*, Vol. 15, 10094.
- Li, F., van Gelder, P., Ranasinghe, R., Callaghan, D., and Jongejan, R. (2014). "Probabilistic modelling of extreme storms along the Dutch coast." *Coastal Engineering*, 86, 1–13.
- Merz, B., Kreibich, H., Schwarze, R., and Thieken, A. (2010). "Review article "Assessment of economic flood damage"." *Natural Hazards and Earth System Science*, 10(8), 1697–1724.
- Norsys (1995-2014). *Netica*. <www.norsys.com>.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.
- Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., and Lescinski, J. (2009). "Modelling storm impacts on beaches, dunes and barrier islands." *Coastal Engineering*, 56(11-12), 1133–1152.
- Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., and Merz, B. (2014). "How useful are complex flood damage models?" *Water Resources Research*, 50(4), 3378–3395.
- Small, C. and Nicholls, R. J. (2003). "A global analysis of human settlement in coastal zones." *Journal of Coastal Research*, 19(3), 584–599.
- Vanneuville, W., Maddens, R., Collard, C., Bogaert, P., de Maeyer, P., and Antrop, M. (2006). *Impact op mens en economie t.g.v. overstromingen bekeken in het licht van wijzigende hydraulische condities, omgevingsfactoren en klimatologische omstandigheden*. Ghent University.
- Vogel, K., Riggelsen, C., Merz, B., Kreibich, H., and Scherbaum, F. (2012). "Flood damage and influencing factors: A Bayesian network perspective." *Proceedings of the 6th European Workshop on Probabilistic Graphical Models*, 347–354.
- Vogel, K., Riggelsen, C., Scherbaum, F., öter, K., Kreibich, H., and Merz, B. (2013). "Challenges for bayesian network learning in a flood damage assessment application." *Proceedings of the 11th International Conference on Structural Safety and Reliability*, 3123–3130.
- Warren, I. and Bach, H. (1992). "MIKE 21: a modelling system for estuaries, coastal waters and seas." *Environmental Software*, 7(4), 229–240.