



**ADVANCES IN  
FOREST FIRE  
RESEARCH**

**DOMINGOS XAVIER VIEGAS**

**EDITOR**

**2014**

# Estimating daily fire risk in the mesoscale by means of a Bayesian network model and a coupled GIS

Panagiota Papakosta, Anke Scherb, Kilian Zwirgmaier, Daniel Straub

*Engineering Risk Analysis Group, Technische Universität München, Munich, Germany  
panagiota.papakosta@tum.de*

## Abstract

A probabilistic spatio-temporal model for fire risk prediction based on Bayesian networks is presented. It predicts daily fire risk on houses and vegetated areas in the mesoscale (1 km<sup>2</sup> spatial resolution). The BN model consists of three parts. (1) The fire occurrence model, which involves as predictive variables the weather conditions (expressed by the Fire Weather Index - FWI), land cover types, population and road density. It predicts the probability of a fire occurring daily in each 1km<sup>2</sup>. (2) The fire behavior model, triggered by the occurrence model, which includes the influence of actual and past weather conditions, fire behavior indices and topography. (3) The damage model, which predicts the expected losses relevant to houses and vegetated areas conditional on a fire hazard. Vulnerability (resistance capacity) and exposure (values at risk) indicators are used to quantify the damage, which also depends on the fire suppression efficiency. The final output of the model are the expected house damage costs (the risk to houses) and the restoration costs for the vegetation (the risk to vegetated areas). The BN model is exemplarily established for Cyprus. The conditional probability distributions of the BN variables are populated with data from the time period 2006-2009, regression model results and expert knowledge. Data from 2010 for Cyprus are used as verification dataset. The results are shown in daily maps with 1 km<sup>2</sup> spatial resolution.

**Keywords:** wildfire risk, Bayesian networks, GIS

## 1. Introduction

Fire risk prediction can support the planning of measures for fire prevention (e.g. thinning, prescribed burning) and risk mitigation (e.g. allocation of suppression resources, raising preparedness). In order to quantify fire risk, the predictive model must include models for fire occurrence, fire behavior and fire consequences. Due to the randomness inherent in the wildfire process and because the modeling is subject to uncertainty in all three stages (occurrence, behavior, damages), fire risk prediction is ideally carried out in a probabilistic format. Fire occurrence is influenced by weather conditions, human presence and vegetation types. Fire behavior is influenced mainly by local weather conditions (e.g. wind speed), topography and fuels (Forestry Canada 1992). Fire damages depend to a large extent on fire severity and the vulnerability of assets (Cohen 2000).

In this study, we focus on the Mediterranean area, where most fires occur due to human intervention as a result of arson or negligence, favored by long dry summer periods (Keeley *et al.* 2012). We propose a probabilistic spatio-temporal model based on Bayesian networks, which predicts daily fire risk on houses and vegetated areas in the mesoscale (1 km<sup>2</sup> spatial resolution). Bayesian Networks (BN) are graphical probabilistic models that can effectively represent complex processes with multiple random variables, their interdependencies and the associated uncertainties.

## 2. Methods

The predictive BN model consists of three parts: (1) The fire occurrence model, which includes as predictive variables weather conditions (expressed by the Fire Weather Index - FWI of the Canadian Forest Fire Weather Index System - CFFWIS), land cover types, population and road density, and

which predicts the probability of a fire occurring in each spatial-temporal unit, i.e. daily and per 1km<sup>2</sup>. (2) The fire behavior model, triggered by the occurrence model, which is a function of the actual and past weather conditions, the fire behavior indices of the CFFWIS and the topography. (3) The damage model, which predicts the expected losses relevant to houses and vegetated areas conditional on fire hazard. Vulnerability (resistance capacity) and exposure (values at risk) indicators are used to quantify the damage, which also depends on the fire suppression efficiency. The final results of the combined model are the risk to houses (expected house damage costs [€]) and the risk to vegetation (vegetated area damage costs [€]). After a short introduction to BN in section 2.1, the three parts of the model are presented in sections 2.2 – 2.4, and the combined model is provided in 2.5.

### 2.1. Bayesian networks

Bayesian Networks (BN) are directed acyclic graphs and consist of nodes, arcs and conditional probability distributions attached to the nodes (Jensen and Nielsen 2007). In a discrete BN, which is used in this study, each node represents a discrete random variable, whose sample space consists of a finite set of mutually exclusive states. The arcs qualitatively describe the dependence structure among the random variables.

A conditional probability table (CPT) is attached to each of the nodes, defining the probability distribution of the variable conditional on its parents. If we consider a BN with discrete random variables  $\mathbf{X} = [X_1, \dots, X_n]$ , then the full (joint) probabilistic model of these variables is the joint Probability Mass Function (PMF),  $p(\mathbf{x}) = p(x_1, \dots, x_n)$ , which can be specified with the help of the chain rule:

$$p(\mathbf{x}) = p(x_n|x_{n-1}, \dots, x_1)p(x_{n-1}|x_{n-2}, \dots, x_1) \dots p(x_2|x_1)p(x_1) \quad (1)$$

By making use of the independence assumptions encoded in the graphical structure of the BN, this chain rule reduces to

$$p(\mathbf{x}) = \prod_{i=1}^n p(x_i|pa(x_i)) \quad (2)$$

wherein  $pa(x_i)$ , are realizations of the parents of  $X_i$  (i.e. the variables with a link pointing to  $X_i$ ). In words, the joint probability mass function (PMF) of all random variables in the BN is the product of the conditional PMFs of all random variables given their parents. Therefore, the graphical structure of the BN, together with the conditional PMFs  $p(x_i|pa(x_i))$ , are sufficient for specifying the full (joint) probabilistic model of  $\mathbf{X} = [X_1, \dots, X_n]$ .

Inference in the BN model is performed through updating. When one or several variables are observed, this information (evidence  $\mathbf{e}$ ) is propagated through the network and the joint prior probability of all nodes is updated to its posterior. The posterior joint probability of a set of variables  $\mathbf{Y} \in \mathbf{X}$  in the network given the evidence  $\mathbf{e}$  is:

$$p(\mathbf{y}|\mathbf{e}) = \frac{p(\mathbf{y}, \mathbf{e})}{p(\mathbf{e})} \quad (3)$$

The joint probabilities  $p(\mathbf{y}, \mathbf{e})$  and  $p(\mathbf{e})$  are computed following Eq. 2. Efficient algorithms for performing these computations in rather complex networks exist, which are implemented in software such as GeNIe (Decision Systems Laboratory 2013) or Hugin (HUGIN EXPERT 2012).

In the context of the proposed model for wildfire risk assessment, the main advantage of the BN is not its computational effectiveness but that it facilitates the combination of information from various sources in a single model.

## 2.2. Fire occurrence model

The fire occurrence model, which is described in detail in (Papakosta and Straub 2014), is based on the assumption that fire occurrences are independent of each other for a given fire rate. Hence the number of fires in an area of size  $\alpha$  during one day,  $N$ , has the Poisson distribution, with PMF:

$$p(n|\lambda) = \frac{(\lambda\alpha)^n}{n!} \exp(-\lambda\alpha), n = 0,1,2, \dots \quad (4)$$

wherein  $\lambda \left[ \frac{\text{Nr.Fires}}{\text{day}\cdot\text{km}^2} \right]$  is the mean occurrence rate and  $\alpha = 1\text{km}^2$  is the area of the cell.

$\lambda$  is calculated by a regression model with explanatory variables *FWI*, *population density*, *road density* and *land cover types*. The predictive model based on the regression results of (Papakosta and Straub 2014) is

$$\log(\lambda) = -10.90 + 0.0329 \cdot \text{FWI} + 0.3217 \cdot \text{Road density} - 0.0234 \cdot (\text{Road density})^2 - 0.0010 \cdot \text{Population density} \quad (5)$$

The resulting BN for fire occurrence is shown in Figure 1. The node *fire occurrence rate*  $\lambda$  is a parent of the node *Fires N*, whose CPT is defined by the Poisson PMF, Eq. (4). The fire occurrence rate is a child of the nodes *FWI*, *population density*, *road density* and *land cover types*, following the deterministic relation of Eq. (5). The CPTs of the explanatory variables are learnt from data of the test bed area.

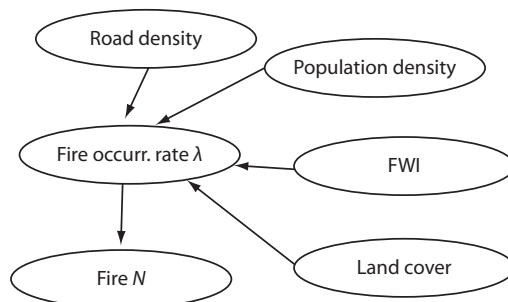


Figure 1. Bayesian network for Fire occurrence. The fire occurrence rate  $\left[ \frac{\text{Nr.Fires}}{\text{day}\cdot\text{km}^2} \right]$  is a function of road density, population density, Fire Weather Index (FWI) and land cover types.

## 2.3. Fire behavior model

The fire behavior model presented in Figure 2 is adapted from (Zwirgmaier *et al.* 2013). The variables are chosen from a wider range of potential variables to represent the processes influencing the resulting burnt area of a fire. In the fire behavior model, the target/predicted node is *Burnt area*. *Burnt area* is influenced by *fire occurrence*, *land cover types*, *topography*, *recent weather conditions*, and *fire behavior indices*. The latter three are hidden variables, meaning that they cannot be observed. They are included to reduce the number of parents of the node *Burnt area*. *Topography* combines the influence of the variables *Wind Speed* [km/h], *Slope* [°] and *Aspect Minus Wind Direction* [same, opposite, perpendicular]. *Recent weather conditions* summarizes the effect of *Relative humidity (RH)* [%], *mean RH over the last 3 days* [%], *mean RH over the last 21 days* [%], *accumulated precipitation over the last 21 days* [mm], *mean temperature of the last day* [°C] and *mean temperature over the last*

7 days [°C]. Initial fire behavior influences the three Indices of the CFFWIS, namely *Fine Fuel Moisture Code (FFMC)*, *Initial Spread Index (ISI)* and *Build-up index (BUI)*.

The BN for fire occurrence (Figure 1) is included in the BN for fire behavior. The variable land cover, which is used for predicting the fire occurrence rate, is also influencing the burnt area directly. However, the node *LC grouped* is introduced to group the land cover types grouped in fewer classes. This reduces the number of free parameters of the variable burnt area and facilitates the parameter learning with hidden variables.

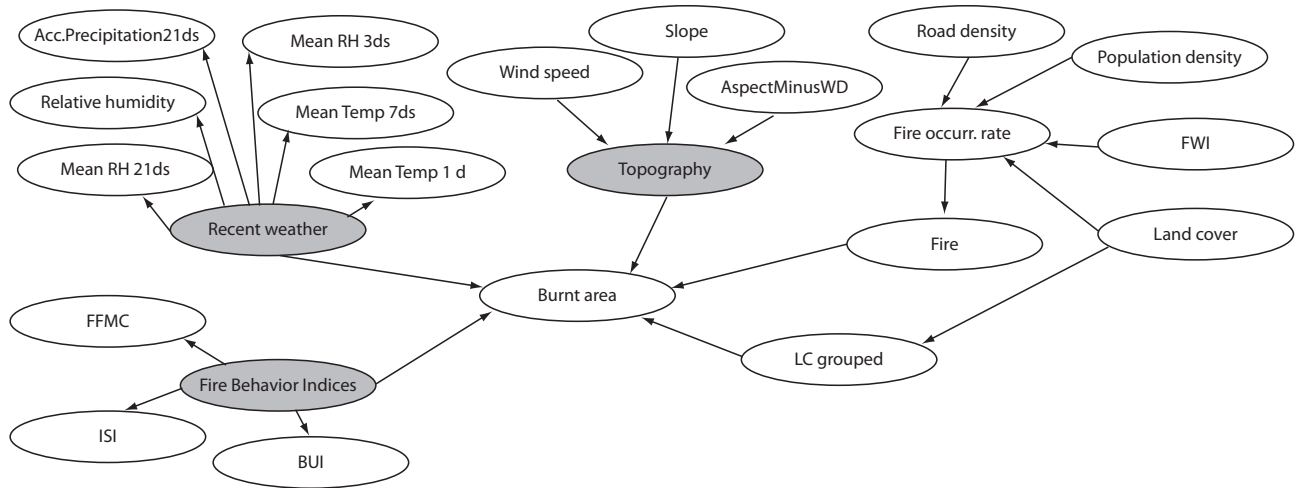


Figure 2. Bayesian network for the prediction of Burnt area. Hidden variables are shown in grey. Indices of the CFFWIS - FWI: Fire Weather Index, FFMC: Fine Fuel Moisture Code, ISI: Initial Spread Index, BUI: Built-Up Index.

### 2.4. Fire consequences model

The fire consequences model estimates the house damage cost and the restoration cost for vegetated areas. The cost is a function of the hazard characteristics, and the consequences that might occur. Wildfire consequences are a function of vulnerability and exposure of the affected biotic and abiotic systems (e.g. human properties, infrastructure, soil and air quality). Vulnerability describes the degree of expected damage as a function of hazard intensity (UNDRO 1991, Thywissen 2006). Exposure refers to the items at risk, such as house density.

Risk is the expected consequences of wildfires and can be formulated as a function of the hazard  $H$ , the resulting damages  $D$  and the consequences  $C$  as,

$$R = E_{H,D}[C] = \int_H p(H) \int_D p(D|H) C(D, H) dD dH \tag{6}$$

$E_{H,D}$  denotes the expected value with respect to  $H$  and  $D$ .  $p(D|H)$  is the probability distribution of damage  $D$  conditional on the hazard  $H$ , i.e. it describes the vulnerability, and  $C(D, H)$  is the cost as a function of damage and hazard. The inner integral in Eq. (6) describes the expected consequences for a given hazard:

$$E_D[C | H] = \int_D p(D|H) C(D, H) dD \tag{7}$$

In the consequence model, the expected costs conditional on hazard characteristics  $E_D[C | H]$  are estimated.

We introduce a model to estimate wildfire consequences on house damage and vegetated area damage (Figure 3). The model is based on the assumption that consequences are influenced by the exposure of the items to fire, the flammability of the items and the suppression effectiveness. As with the fire occurrence and behavior models, the modeling is conducted at the mesoscale.

The variables that describe the hazard  $H$ , and which are outputs of the fire occurrence and behavior models, are *Burnt area detailed*, *Fire type* and *FWI*. *Fire type* can be a surface fire with flame length <3.5m, surface fire with flame length >3.5m and crown fire. Burnt area is here expressing wildfire severity.

Vulnerability is expressed by the suppression effectiveness and the damage on houses and vegetated areas (nodes *House damage*, *Area damage* and *Construction type*). The node *Fire Containment in 24 hrs* and its parents represent the probability of a fire being contained by the firefighting crews within 24 hrs. *Fire containment in 24 hrs* is influenced by *FWI*, *Time for ground attack*, *Air suppression* and *Vegetation type*. The nodes influencing the vulnerability variable *Fire containment in 24 hrs* are based on (Plucinski *et al.* 2012, Plucinski 2012), where fire containment is the dependent variable and the parameters are defined as a result of logistic regression analysis (Table 1). The probability  $P$  of fire containment in 24hrs is defined as,

$$\ln\left(\frac{P}{1-P}\right) = b_0 + b_1FFDI + b_2groundtime + b_3airtime \tag{9}$$

wherein *FFDI* is the McArthur Forest Fire Danger Index, *groundtime* is the time needed for ground suppression crews to reach a fire spot and *airtime* is the corresponding time for air suppression crews.

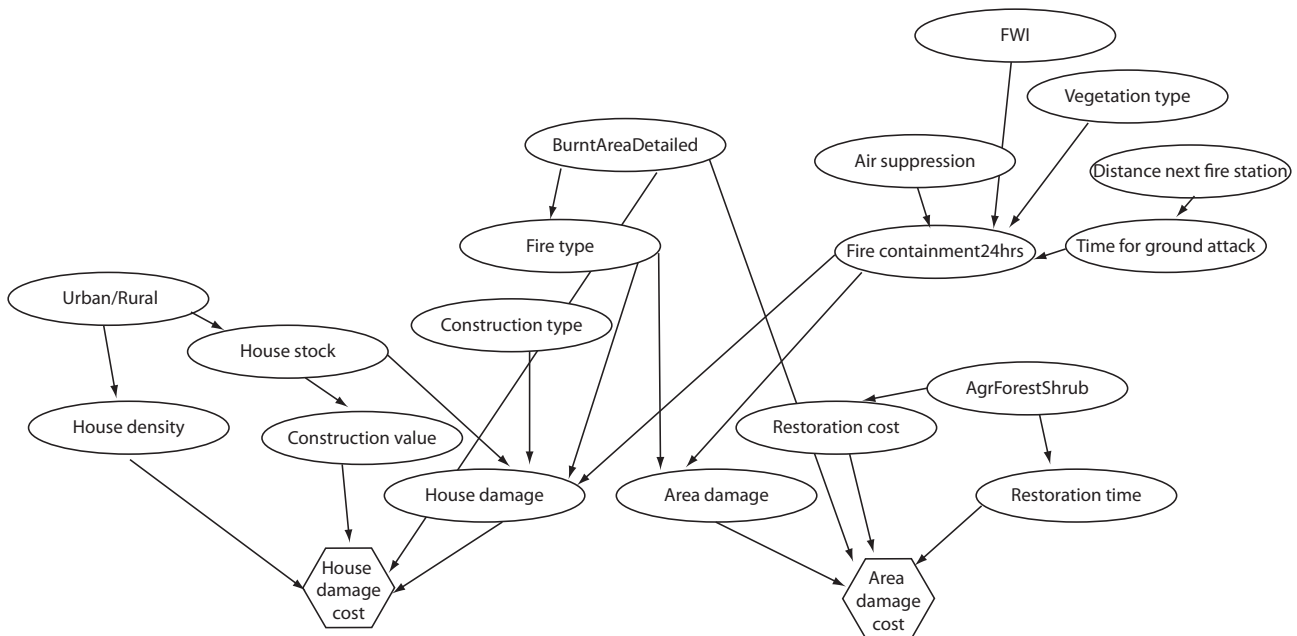


Figure 3. Bayesian network for fire consequences

The vulnerability node *House damage* represents the degree of damage, i.e. the susceptibility of the house portfolio in the cell. The vulnerability is influenced by *Fire type*, *Fire containment in 24 hrs*, *Construction type* and *House stock*. It is expressed as percentage of houses totally damaged in 1 km<sup>2</sup>.

The node *Area damage* is the corresponding vulnerability node for the vegetated area damage. It represents the degree of damage in 1 km<sup>2</sup>. It is expressed as a percentage of the total area (1km<sup>2</sup>).

*Urban/Rural*, *House Stock*, *Construction Value*, *House Density*, *Land Cover types*, *AgricultureForestShrub*, *Restoration cost*, *Restoration time* are considered as exposure indicators. *Urban/Rural* discriminates urban from rural areas, which influences the house density [house/km<sup>2</sup>] and the house stock. *House stock* accounts for the house type portfolio in the mesoscale. It describes the combination of house types in 1km<sup>2</sup>, which include single houses, semi-detached/row houses, and apartments. The node *House damage cost* is calculated as

$$\begin{aligned} & \text{House damage cost} \\ & = \text{Burnt area} \cdot \text{House damage} \cdot \text{Construction value} \cdot \text{House density} \end{aligned} \quad (10)$$

The node *Land Cover types*, represents the land cover types as classified in the nomenclature of the Corine Land cover types. The node *AgricultureForestShrub* classifies the land cover types based on the vegetation type. *Restoration cost* [€] is the cost for restoring previous land cover based on premiums paid country wise by the EU Rural Development Programs. *Restoration time* [yrs] is the time needed for restoring previous land cover type (Oehler *et al.* 2012). The node *Area damage cost* is calculated based on results in (Oehler, Oliveira *et al.* 2012) as

$$\begin{aligned} & \text{Area damage cost} = \\ & \text{Burnt Area} \cdot \text{Area damaged} \cdot \text{Restoration cost} \cdot (1 + r)^{\text{Restoration time}} \end{aligned} \quad (11)$$

where  $r = 3\%$  is the discount rate.

*Table 1. Fire Containment in 24hrs: regression parameters (Eq.9) (Plucinski *et al.* 2012), (Plucinski 2012)*

Vegetation Type	Suppression	$b_0$ intercept	$b_1$ F/G FDI	$b_2$ ground time	$b_3$ air time
grass	ground	2.41124	-0.02454	-0.51708	
forest	ground	1.168703	-0.024632	-0.20104	
shrub	ground	1.664122	-0.019558	-0.282204	
grass	ground&air	4.80436	-0.042789	-0.66977	-0.3253
forest	ground&air	3.83561	-0.05031	-0.29845	-0.34783
shrub	ground&air	3.75257	-0.03704	-0.3184	-0.08101

## 2.5. Fire risk model

The Fire risk model of Figure 4 is obtained as the combination of the fire occurrence model, the fire behavior model and the fire consequences model. This model predicts the risk to Houses and vegetated areas.

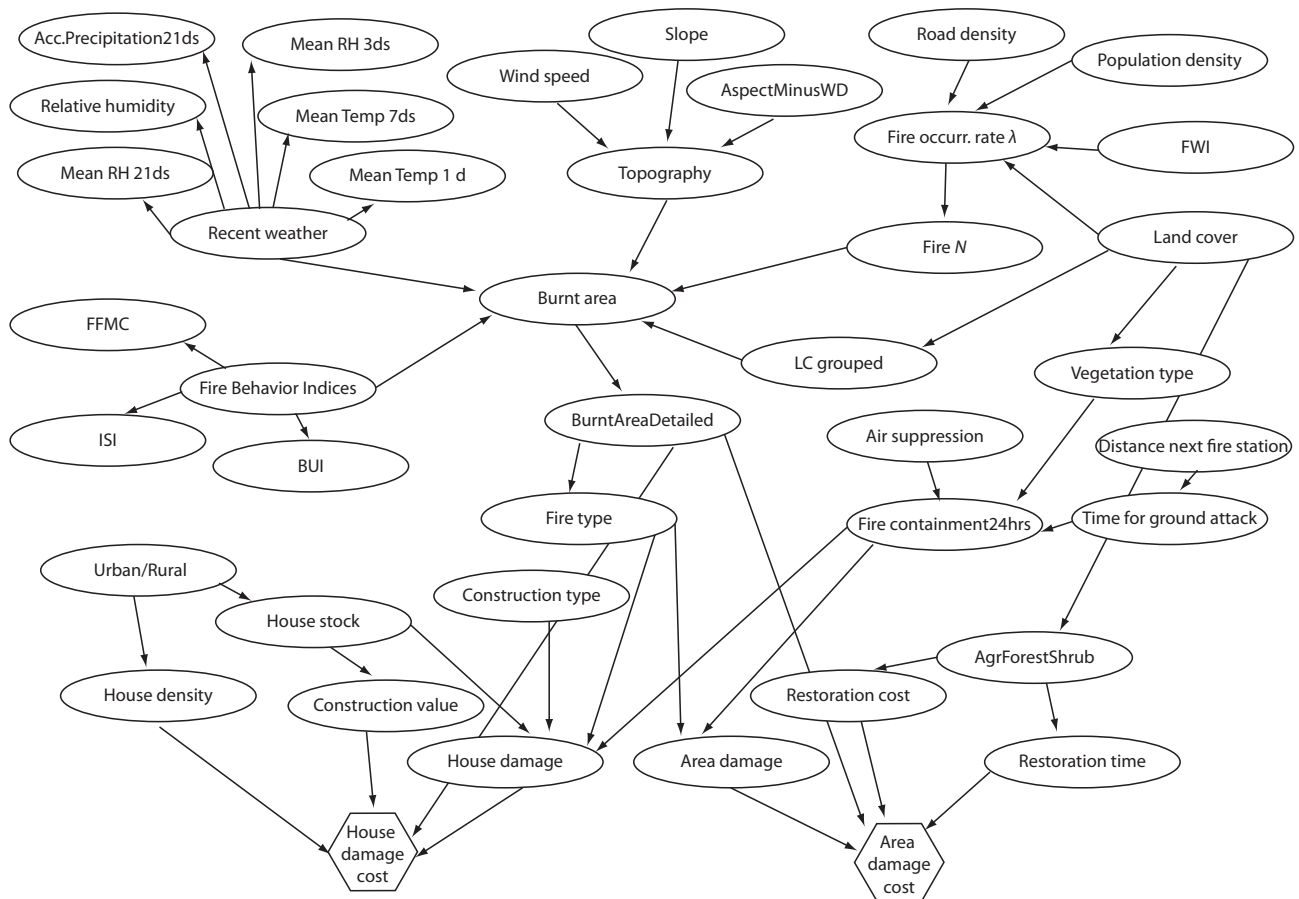


Figure 4. Combined Bayesian network for fire risk.

The BN is coupled with a GIS. Spatial feature groups, such as points, lines and polygons, are processed, stored and managed in a GIS database. Georeferenced spatial features are projected onto a grid with 1 km<sup>2</sup> cell size, which serves as the spatial resolution of the model. The geospatial features are associated with variables in the BN. In each cell, an instance of the BN models the wildfire risk. With the specific geospatial features associated with a cell, the wildfire risk associated to that cell can be calculated from the BN model. Spatial dependence is represented through the dependence of the observed indicator variables, but not through the BN itself. ArcGIS 10.1 is used for geospatial analysis and mapping (ESRI 2012).

The parameters of the exposure indicators are learned with the attribute data of the geospatial features. Evidence (information on the observed variables) is given on the hazard characteristics and exposure indicators and is propagated in the network and the posterior marginal distributions for each variable are estimated.

### 3. Testbed implementation

The BN model is established for Cyprus. The conditional probability distributions of the BN variables are populated with data from the time period 2006-2009, regression model results and expert knowledge. Data from 2010 for Cyprus are used as verification dataset. The results are shown in daily maps with 1 km<sup>2</sup> spatial resolution.

Figure 5 and Figure 6 exemplarily show the estimated expected damage cost (risk) for the two days of 2010 that had the maximum number of registered fires. The figures differentiate between house risk,



vegetated area risk and the accumulated risk. The calculated FWI with 1km<sup>2</sup> spatial resolution for each day is displayed as well. The risk in each spatial-temporal unit is low due to the low probability of fire occurrence (in the order of 10<sup>-5</sup>). In both examples, the highest values of risk occur in wildland-urban interface areas.

The effect of different influencing variables on the probability of fire occurrence is shown in Table 2. For fire behaviour, Table 3 summarizes the effect of the direct influencing variables on the probability of a burnt area > 0.1 km<sup>2</sup>.

*Table2. Effect of influencing variables on fire occurrence probability*

Variable	States	Pr(Fire) Prior: 9.45e-5	Change in Probability [%]
FWI	0-10	4.50e-5	-52
	10-30	7.22e-5	-24
	30-60	1.60e-4	+70
	60-120	2.82e-4	+199
LandCoverTypes	1 : Urban/Wetland/Pastures	2.17e-5	-77
	2: Arable land	2.72e-5	-71
	3: Permanent crops	1.33e-4	+40
	4: Heterogeneous agriculture	8.31e-5	-12
	5: Forests	1.80e-4	+91
	6: Shrubs/Herb. vegetation	9.34e-5	-1
	7: Open spaces	2.21e-4	+133
Population density	0-20	7.81e-5	-17
	20-300	7.78e-5	-18
	300-4000	3.21e-4	+240
Road density	0-0.5	9.36e-5	-1
	0.5-2	7.35e-5	-22
	2-26	1.16e-4	+23

*Table3. Effect of influencing variables on burnt area*

Variable	States	Pr(BurntArea>0.1) Prior=1.70e-5	Change in Probability [%]
LCgrouped	1,4,6	1.65e-5	-3
	2,3,5	1.56e-5	-8
	7	6.54e-5	+284
Topography	1:middle	1.97e-5	+16
	2:gradual	8.59e-6	-50
	3:steep	2.88e-5	+69
Recent weather	1: dry	2.49e-5	+46
	2: moderate dry	1.82e-5	+7
	3: moderate humid	1.28e-5	-25
	4: humid	1.12e-5	-34
Fire behavior indices	1: moderate	1.78e-5	+4
	2: low	1.35e-5	-21

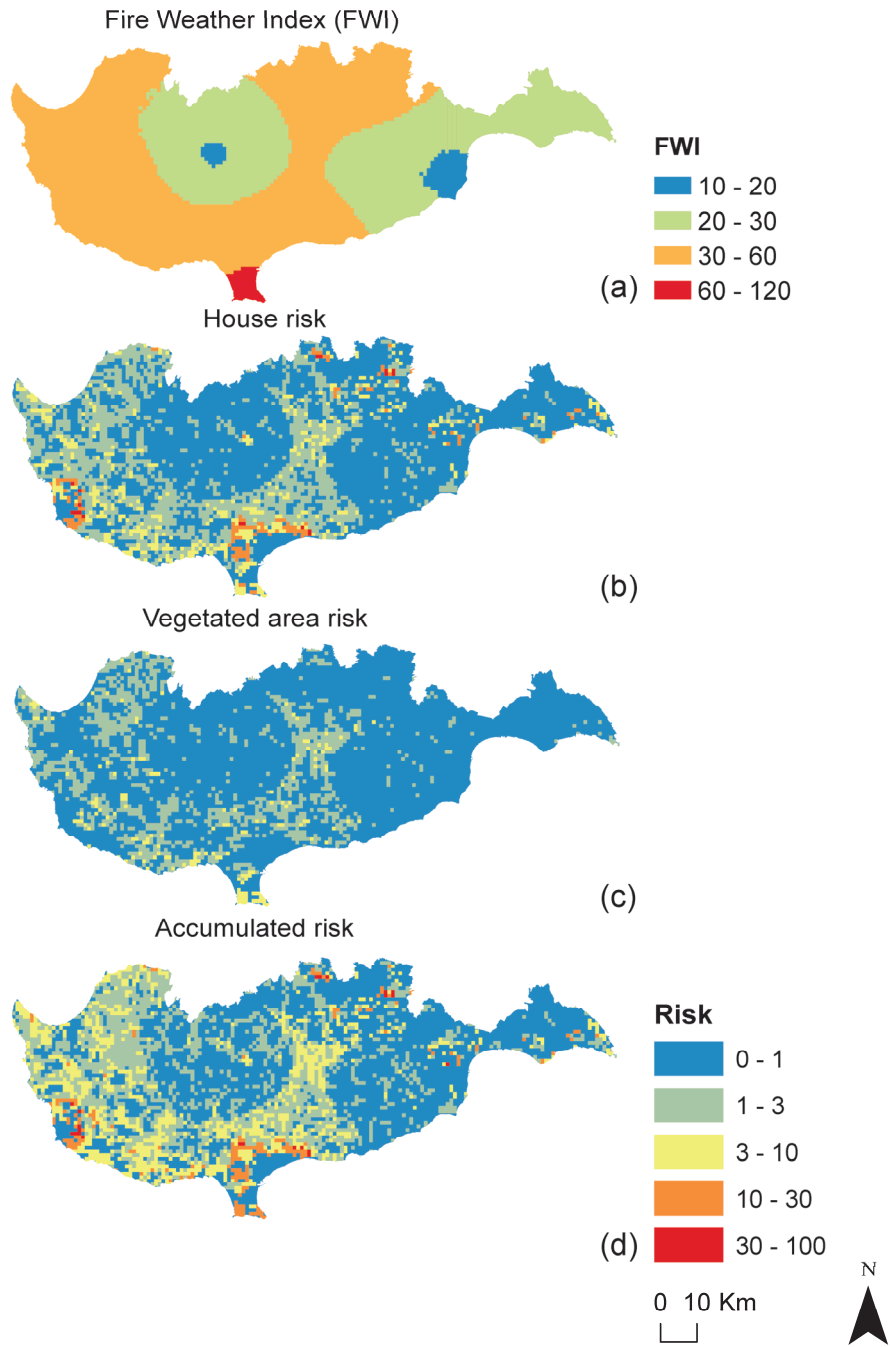


Figure 5. (a) Fire Weather Index (FWI), (b) House risk, (c) Vegetated area risk, (d) Accumulated risk on 08.10.2010

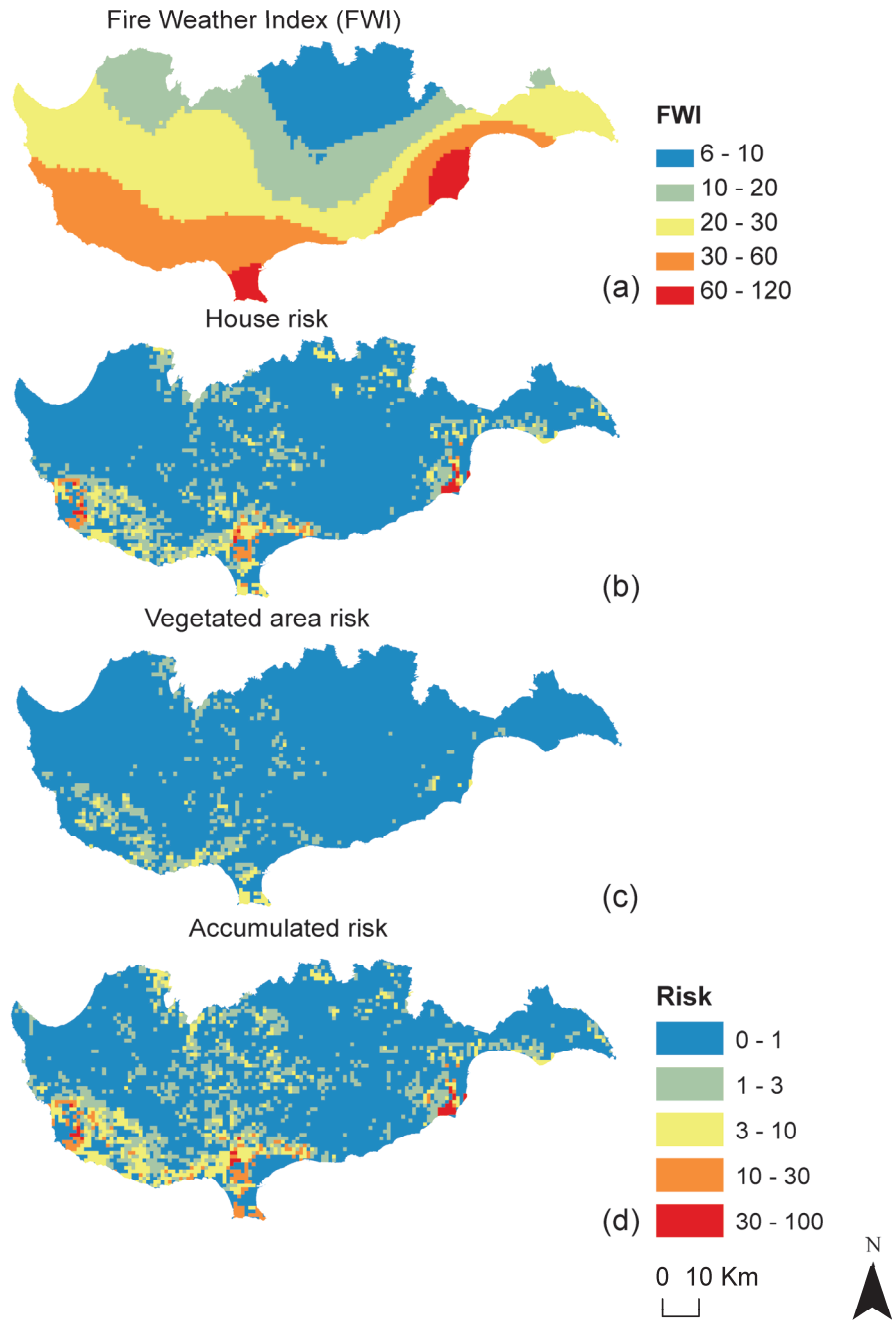


Figure 6. (a) Fire Weather Inx (FWI), (b) House risk, (c) Vegetated area risk, (d) Accumulated risk on 26.06.2010

#### 4. Concluding remarks

While the understanding and prediction of wildfire occurrence has made significant progress in recent years, research on wildfire risk estimation based on exposure and vulnerability quantification is still evolving (Tutsch *et al.* 2010). Wildfire risk estimation may be facilitated by Bayesian Networks, which, as demonstrated in this paper, can be employed to model the wildfire risk in the mesoscale. The mesoscale requires that the indicators included in the model are representative for a 1 km<sup>2</sup> spatial

unit. This has strong implications on the modelling; as an example, it is necessary to identify representative building stocks and construction types not of individual buildings, but of portfolios of buildings. A main advantage of the BN is that different models from different sources can easily be combined. Of the BN for wildfire risk proposed in this paper, some parts are learned from data, some parts are derived from models published in the literature and other parts are based on expert knowledge. The graphical nature of the BN facilitates to consistently combine these model parts probabilistically and, not least, also facilitates the communication of the final model.

## 5. References

- Cohen, J. D. (2000). "Preventing disaster: Home ignitability in the wildland-urban interface." *Journal of Forestry* 98(3): 15–21.
- ESRI (2012). ArcGIS 10.1. ESRI (Environmental Systems Resource Institute), Redlands, California, US.
- Forestry Canada, F. D. G. (1992). Development and Structure of the Canadian Forest Fire Behavior Prediction System. Information Report ST-X-3. S. a. S. D. D. Forestry Canada. Ottawa, Canada: 63.
- Keeley, J. E., W. J. Bond, R. A. Bradstock, J. G. Pausas and P. W. Rundel (2012). *Fire in Mediterranean ecosystems: Ecology, evolution and management*. Cambridge, Cambridge Univ. Press.
- Oehler, F., S. Oliveira, J. I. Barredo, A. Camia, J. SanMiguel-Ayanz, D. Pettenella and R. Mavsar (2012). Assessing European wildfire vulnerability. European Geoscience Union General Assembly 2012. Vienna, Austria.
- Papakosta, P. and D. Straub (2014). "Probabilistic prediction of daily fire occurrence in the Mediterranean based on weather and anthropogenic factors." submitted.
- Plucinski, M., G. McCarthy, J. Hollis and J. Gould (2012). "The effect of aerial suppression on the containment time of Australian wildfires estimated by fire management personnel." *International Journal of Wildland Fire* 21(3): 219-229.
- Plucinski, M. P. (2012). "Factors affecting containment area and time of Australian forest fires featuring aerial suppression." *Forest Science* 58(4): 390-398.
- Thywissen, K. (2006). *Core terminology of disaster reduction: A comparative glossary. Measuring vulnerability to natural hazards*. J. Birkmann. Tokyo, Japan, United Nations University Press: 448–496.
- Tutsch, m., W. Haider, B. Beardmore, K. Lertzman, A. B. Cooper and R. C. Walker (2010). "Estimating the consequences of wildfire for wildfire risk assessment, a case study in the southern Gulf Islands, British Columbia, Canada." *Canadian Journal of Forest Research*(40): 2104–2114.
- UNDRO (1991). *Mitigating natural disasters: phenomena, effects and options.: A manual for policy makers and planners*. Geneva, (UNDRO, Office of the United Nations Disaster Relief Co-Ordinator).
- Zwirgmaier, K., P. Papakosta and D. Straub (2013). Learning a Bayesian network model for predicting wildfire behavior. Proceedings of the 11th International Conference on Structural Safety & Reliability.