Stochastic Multi-Hazard Occurrence Model

Mark Bebbington, Alex Dunant, David Harte

Draft Version: 2023-04-08 13:59:29 (UTC+12:00)

Abstract

Abstract required ???

The title and so on are just space fillers. They do not reflect a strong preference on my part.

Multiple question marks indicate problems or missing content (easy to search in PDF).

Contents

List of Figures

1 Introduction

"What is going to happen when disaster strikes?" is the first order question in risk science. While modelling of disaster scenarios is possible, it poses many challenges; natural hazards are inherently complex, meaning hazard magnitude (i.e. the scale of the event) and the impact (i.e. the damage) are unlikely to follow straightforward relationships (B. Liu et al., 2016). Firstly, hazard systems include many dependencies and other types of connections between their parts (Dunant, 2021; Z. Liu et al., 2015; Mignan et al., 2014)(Mignan et al., 2014). For example, an earthquake can create thousands of landslides and landslide dams, which themselves, cause disastrous outburst floods, such as has been recently observed in New Zealand during the 2016 Kaikoura earthquake (Dellow et al., 2017) (Hughes et al., 2015; Jibson et al., 2018). Current "timeless" risk assessment frameworks lead to inaccurate inferences for dynamic processes, underestimating the potential for extreme events and belittle significant risk. Recently, a potential solution to the shortcoming of complexity in risk systems has been proposed (Dunant et al., 2021a,b), whereby hazards and their interactions can be abstractly represented by networks. However, this is yet to be incorporated into a long term temporal dynamic framework for risk assessment. We aim to incorporate time as a key dimension of the systemic risk assessment and provide novel, more realistic insight into natural disaster resilience.

Time-dependant networks are widely used in other disciplines (Holme & Saramäki, 2012) including social communication (Pereira et al., 2016), biology (Berdondini et al., 2009), infrastructure networks (Pan & Saramäki, 2011) and ecology (Blonder et al., 2012). Drawing from those disciplines, a holistic solution to the complex problem of "motionable" risk would be to use dynamic temporal networks to predict the systemic multi-hazard threat.

Using existing empirical relationships and process stochastic distributions, we will use a temporal network model to propose a statistical framework for cascading hazard forecasts. Following a temporal simulation of potential cascading hazards interactions, disaster scenarios could be simulated over periods of decades while providing quantifiable uncertainty.

The temporality brings an added complexity to the multi-hazard simulation as nodes are interacting at different times. Using the algorithm, each process cascades and interacts with other nodes, following the initial concept by Dunant et al. (2021b), using a global time and a process time. Global time would be viewed as time moving forward in a global sense and encompasses all processes. A process time is the local time of each of the processes, which may be ahead of the global time by the means of local delays (Fig. 5). When the process time is in line with the global time, the process "occurs". The proposed stochastic framework is intended to map out the long term consequences of an initial stimuli across the landscape and to study its potential for complex cascading hazard impacts over an extended period of time.

1.1 Physical and Statistical Models

1. Physical natural hazard models tend to focus on a specific hazard: earthquakes, floods, landslides, volcanic eruption, storms, etc.

2. Physical models tend to focus on the physics of the process with little or no attention to stochastic uncertainty.

3. Many of these physical processes have also been described by statistical models, with considerably less computational effort.

4. The statistical models are generally based on observed empirical relationships in historical data. Ultimately, one wants the observed empirical relationships to have a physical

explanation, and the physical models to produce simulated data that are consistent with observed historical data. Hence both approaches support the development of the other.

5. The above leads into the notions of epistemic and aleatory uncertainty.

1.2 Quantifying Uncertainty in Forecasts

1. The purpose of the statistical model is to retain credibility with the known physics of the process, but to also provide a measure of the associated uncertainty. Physical models without an uncertainty component will only produce average (expected) behaviour.

2. We present a stochastic framework for the modelling of multiple hazards, in particular, where infrastructure may be already damaged due to a previous event.

3. Our model treats the recurrence of each individual natural hazard, and the effect of cascading multiple hazards with probability distributions.

4. Forecast probability distributions (empirical) can be determined by simulating the model multiple times.

2 The Network Model

We use the mathematical concept of a graph to formalise the structure of the network model.

2.1 Requirements

1. Make model dynamic.

2. Make model fully stochastic, providing a description of the underlying uncertainty.

2.2 Nodes

1. The nodes are based on Davies et al (2020). An example of potential hazards leading to flooding is in Fig 1.

2. Proposed node locations in a virtual catchment are shown in Fig 2.

2.3 Connections Between Nodes – Edges

1. We consider a more general system of nodes and connections as in Fig 3 to that for flooding in Fig 1.

2. We need to distinguish between spatial connections (edges), see Fig 4; and temporal connections, see Fig 5.

3 Examples

1. Result of a simulation using plausible parameter values for probability distributions described in the Appendix.

2. The above may show little triggering between hazards?

3. Can we give another example with a few fiddled event rates where we get triggering? We may be able to show that the fiddled rates are quite implausible, hence the likelihood of such a sequence happening is quite small. I suppose it depends on the relationship between the aleatory and epistemic uncertainty.

4. Are there other combinations of parameters where the methodology breaks, giving bizarre results?

A Details of Probability Distributions

A.1 Volcano Tephra

We model the explosion process with a renewal process (Bebbington & Lai, 1996). Use Weibull distribution with distribution

$$F(r) = 1 - \exp\left[-(\alpha r)^{\beta}\right] ,$$

where α and β need to be estimated. Then, if last explosion occurred at time s, the time of the following explosion is at $s + R$ where

$$R = \frac{1}{\alpha}(\log_e X)^{1/\beta} ,$$

and X is a uniform $[0, 1]$ random variable.

The amount (thickness) of ash that arrives in a catchment at location x (using the centroid???) can be modelled using Kawabata et al. (2013) where

$$\overline{T}(d, \theta) = \frac{\gamma}{d^{-\eta}} \exp\left\{-vSd[1 - \cos(\theta - \phi)]\right\}$$

is the mean tephra thickness at a location a distance d from the volcano at an angle of q off the dispersal axis ϕ. The actual tephra deposit can be considered a lognormal random variable with mean given in (1.2)??? and a coefficient of variation (standard deviation divided by mean) of approximately 0.5 (Kawabata et al., 2013). So the actual thickness can be generated as

$$T = \exp(0.22Z + \overline{T}) ,$$

where Z is a normal (0,1) random variate. Note the following:

1. The parameter γ is the expected thickness at 1km along the dispersal axis, which can be obtained from the volume and the expected wind via numerical integration (or just summation) of (1.3)???.

2. The combination vS is a combination of diffusivity (grainsize) and wind speed. Typical values are approximately 1.5 for a 50km/hr wind (Kawabata et al., 2013).

3. The decay parameter η can be estimated by

   $$\eta = 2.5 - 0.05H$$

   (González-Mellado & De la Cruz-Reyna, 2010), where H is the column height in km.

So we need wind speed/direction, column height and eruptive volume. The column height (km) is consistent with a Weibull distribution (1.1)?? with $\alpha = 0.93$, $\beta = 0.19$, so H can be generated from (1.2)??. The bulk volume (cubic km) of ash (tephra) can be generated from the relationship

$$V = 2.6 \exp(0.065H - 1.69) ,$$

(Mastin et al., 2009). If we want to add an error term, multiply (1.5)??? by a lognormal random variable, derived as $M = \exp(2Z)$, where $Z \sim \mathrm{N}(0, 1)$.

Wind speed and direction: get bivariate distribution for $(u, v) = (\mathrm{toeast}, \mathrm{tonorth})$ components of wind velocity at circa 12km altitude. The model is bivariate normal, i.e.

$$(u, v)' = \mathrm{N}(\mu, \Sigma) ,$$

Season		μ_u	μ_v	σ_u	σ_v	ρ
Spring	Sep–Nov	21.1	0.69	10.8	10.0	-0.11
Summer	Dec–Feb	20.8	0.55	13.6	14.1	-0.28
Autumn	Mar–May	23.7	0.88	13.7	12.5	-0.22
Winter	Jun–Aug	27.6	0.26	12.1	10.3	-0.17
Overall Average		23.3	0.60	12.9	11.9	-0.21

Table 1: Seasonal volcano parameter values used in simulations. Given the low variability, we could just use the overall average.

where $\mu = (\mu_u, \mu_v)'$ is the mean, and

$$\Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{uv} \\ \sigma_{uv} & \sigma_{vv} \end{pmatrix}$$

is the covariance matrix. We can reparametrize the latter as

$$\Sigma = \begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix}$$

where $u \sim \mathrm{N}(\mu_u, \sigma_u^2)$, $v \sim \mathrm{N}(\mu_v, \sigma_v^2)$ and ρ is the correlation between u and v. It is then easy to simulate a draw (U, V) from the bivariate distribution via the equations

$$U = \mu_u + Z_1\sigma_u$$

and

$$V = \mu_v + \rho\frac{\sigma_v}{\sigma_u}(U - \mu_u) + Z_2\sqrt{(1-\rho^2)\sigma_v^2}\,,$$

where $Z_i \sim \mathrm{N}(0, 1)$ are standard normal random variables.

Seasonal variability is modelled through parameters values (in m/s) as in Table 1.

Then

$$\phi = \arctan(-U/V)\,,$$

(clockwise from East), and

$$vS = \frac{5.1}{50}\sqrt{U^2 + V^2}\,.$$

A.2 Debris Avalanches

Stuart??? Assume that dome growth may or may not occur between significant explosions. Time between explosions already modelled. If dome growth, distribution for rate of growth. Random dome collapse as function of volume, or deterministically at next explosion. Given volume of debris, where does it run to (probability distribution – does it hit nearby stream or infrastructure nodes.)

A.3 Precipitation/Storms

Precipitation feeds streams, sediment mobilization, landslides and surface flooding

Storms (i.e. wind etc.) are direct hazards. Can be ignored.

Probability distribution for 24 hour(???) precipitation at location. Then do random distribution for reference cell, probably seasonal.

Each day, with probability p there is rain (probability $1-p$ of no rain). This varies by season:

Season		p
Spring	(Sep–Nov)	0.43
Summer	(Dec–Feb)	0.40
Autumn	(Mar–May)	0.35
Winter	(Jun–Aug)	0.39

On days when there is rainfall, the amount (in mm), can be modelled by a lognormal random variable with parameters:

Season		μ	σ
Spring	(Sep–Nov)	1.62	1.01
Summer	(Dec–Feb)	1.56	1.06
Autumn	(Mar–May)	1.73	1.09
Winter	(Jun–Aug)	1.80	1.07

The amount of rainfall per day is then simulated as

$$R = I(U < p)\exp(\mu + \sigma Z)$$

where U is a uniform(0,1) random variable, and Z is a normal(0,1) random variable.

Comment: There is little to no spatial variation in rainfall within a 24 hour period, so we will ignore it and use a spatially constant rainfall rate.

A.4 Sediment Mobilization

State of catchment node = thickness of tephra.

Assuming a thickness of sediment T, and an amount of rainfall R, we need a function for how much sediment is mobilized into the stream segment (along with the rainfall) and how much remains. There are three equations needed: one for how much rainfall runs off (the rest is absorbed), one for the saturation level of the ground, and one for how much sediment (read tephra) is carried with the runoff.

Tarawera Rainfall Runoff

Runoff into rivers is calculated using a non-linear reservoir hydrological model (NLRRM). Each subcatchment (attached shapefiles for Rangitaiki and Tarawera) are represented as a reservoir with non-linear discharge. The discharge (q) is a function of the catchment storage depth (S) and coefficients K and p, where

$$q = \left(\frac{S}{K}\right)^{1/p} .$$

The storage depth is calculated using the F1-RSA method. It consists of a primary runoff rate (F1) and a saturated runoff rate (FSA). Essentially, some rainfall infiltrates into the ground until saturated, which reduces the amount of runoff. Once the enough rainfall has accumulated to saturate the soil, the runoff rate increases. The effective rainfall, r, is given as a proportion to rainfall intensity (I, mm/h):

$$r = ??? ,$$

where S_{sat} is the total storage of water and R_{sa} is the storage amount for saturation.

The total storage, S, is then the sum of total storage and effective rain, minus the runoff

$$S = S_{\text{sat}} + r - q ;$$

are mm the units for S_{sat}???

In the algorithm, r is calculated first, then q (using S values from previous timestep), then S. The timestep is the duration between rainfall data points, see Fig 6.

The total discharge from each catchment is then the discharge (q) multiplied by catchment area, divided by the time increment plus any baseflow.

In the case study model, this discharge is then provided to the morphodynamic model as an inflow to the river. We take this approach to model the effect of sedimentation.

However, in NLRRM, the water is routed downstream using a kinematic routing method to provide/forecast streamflow heights and flowrates. This is a very simple method, but provides good quantitative results to downstream flow gauge records. The NLRRM approximation uses a simplified parabolic shape for the channel cross section.

So we have an effective rainfall equation

$$R_{\text{eff}} = \begin{cases} c_1 R & S_{\text{sat}} \le R_{\text{sa}} , \\ c_1 R + (c_2 - c_2)(S_{\text{sat}} - R_{\text{sa}}) & S_{\text{sat}} > R_{\text{sa}} , \end{cases}$$

where $R = R(t)$ is the rainfall in time step t, R_{sa} is a constant (the level of saturation at which rainfall runs off at a higher rate), and $S_{\text{sat}} = S_{\text{sat}}(t)$ is the current saturation level of the ground.

In each catchment, water runs off at a (lower) rate c_1 until the ground is saturated, then runs off at a higher rate c_2. So the model needs to keep track of the saturation level in each catchment,

$$S_{\text{sat}}(t + 1) = S_{\text{sat}} + R_{\text{eff}} - q$$

where the discharge to the stream Q is a non-linear function of the effective rainfall and coefficients K, p:

$$Q = \left(\frac{R_{\text{eff}}}{K}\right)^{1/p}$$

For our purposes (see attached data), $p = 1$, and so K can be absorbed in c_1 and c_2.

Modified Universal Soil Loss Equation

The Modified Universal Soil Loss Equation (MUSLE) is used to estimate sediment yield per catchment. The equation is:

$$Y = 11.8(Q_T Q_{\max})^{0.56} \times K \times \text{LS} \times C \times P\,,$$

where Y is the soil loss within a period (often expressed in tons), Q_T is the total discharge (m^3) throughout the period, $Q_{\max}$ is the maximum discharge (m^3/s rate??) during the studied period, K is the soil erodability LS is a slope length and steepness factor (dimensionless), C is a land usage factor (dimensionless), P is an erosion control factor (dimensionless, often assumed as 1).

The parameter K is a soil erodibility factor, representing soil loss in mass/area. Technically, it represents soil loss (tons/ha) for a specific slope length and gradient, and specific land cover, which results in weird units. The dimensionless LS, C and P parameters are then modifications on this coefficient depending on the different catchment characteristics.

In reality, each catchments sediment yield from a rain event will be represented by:

$$Y = 11.8(Q_T Q_{\max})^{0.56}\alpha$$

where $\alpha = K \times \text{LS} \times C \times P$ is a catchment specific coefficient for sediment yield.

Our K values will be found from the literature, and will then use LS parameter that is proportional to catchment area and slope. One example of LS is (Chalov et al., 2017):

$$\text{LS} = (m+1)\left(\frac{A_c}{a_0}\right)^{0.4}\left(\frac{\sin\beta^{1.6}}{b_0}\right)$$

where A_c is catchment area/catchment width (units of m), β is the slope of the catchment, a_0 and b_0 are the normalised length and slopes used on the plots where K was measured. The land management factor C is proportional to vegetation. More vegetation results in a lower factor.

So basically, $?_? = ?_{\max}Q$ (??? missing vaiable names). We then need the area and width of each catchment, and its slope. Then the tephra loss Y into the river (provided there is still some present) is proportional to

$$cQ^{??}\left(\frac{??}{??}\right)^{??}(??)^{??} = c(i)Q(i)^{1.12}\,,$$

for catchment i, where A is the catchment area, w its width, b its slope, and c is a constant depending on vegetation. Allowing for tephra presence,

$$Y_{\text{tostream}} = \min(Y, T)\,,$$

where T is the current subaerial tephra in the catchment.

A.5 Streamflow Segments/Catchments

How many segments? (1 per catchment). Stuart

Catchments have centroid (for tephra/rain location), area, and anything else which affects sediment mobilization (average lope, vegetation type?)

Stream segments. Discrete or continuous states? Need to track bed height (cross section), sediment load, water cross section. Is flow rate a function of water height? Cross section (i.e. stopbank height) is a static parameter

Need relationship for how fast a 'wave' propagates downstream.

A.6 Flood Nodes

One per stream segment. Infrastructure is either in or not. Or perhaps a probability?

A.7 Infrastructure List

From ME (type, locations)

A.8 Dams

Matahina – special case

Landslide / DA / lava dams – potential failure at overtop (calculated from dam height, flow rate). Otherwise becomes a weir.

Dam failure – generate lahar with dam volume + water. How to model? Stuart

A.9 Landslides

Track only specific landslide nodes (can either hit infrastructure or dam a stream).

Static properties – slope, material (volume distribution / runout)

Keep track of water content (precipitation decay). Landslide becomes possible with excessive saturation OR moderate saturation PLUS external trigger (e.g. earthquake)

A.10 Earthquakes

David

References

Bebbington, M. & Lai, C., 1996. On nonhomogeneous models for volcanic eruptions, *Math. Geol.*, **28**(5), 585–600, doi: 10.1007/BF02066102. 6

Chalov, S. R., Tsyplenkov, A. S., Pietron, J., Chalova, A. S., Shkolnyi, D. I., Jarsjö, J., & Maerker, M., 2017. Sediment transport in headwaters of a volcanic catchment – Kamchatka Peninsula case study, *Frontiers of Earth Science*, **11**(3), 565–578, doi: 10.1007/s11707-016-0632-x. 10

Dellow, S., Massey, C., Cox, S., Archibald, G., Begg, J., Bruce, Z., Carey, J., Davidson, J., Pasqua, F. D., Glassey, P., Hill, M., Jones, K., Lyndsell, B., Lukovic, B., McColl, S., Rattenbury, M., Read, S., Rosser, B., Singeisen, C., Townsend, D., Villamor, P., Villeneuve, M., Godt, J., Jibson, R., Allstadt, K., Rengers, F., Wartman, J., Rathje, E., Sitar, N., Adda, A.-Z., Manousakis, J., & Little, M., 2017. Landslides caused by the Mw7.8 Kaikōura earthquake and the immediate response, *Bulletin of the New Zealand Society for Earthquake Engineering*, **50**(2), 106–116, doi: 10.5459/bnzsee.50.2.106-116. 3

Dunant, A., Bebbington, M., & Davies, T., 2021a. Probabilistic cascading multi-hazard risk assessment methodology using graph theory, a New Zealand trial, *International Journal of Disaster Risk Reduction*, **54**, art102018, doi: 10.1016/j.ijdrr.2020.102018. 3

Dunant, A., Bebbington, M., Davies, T., & Horton, P., 2021b. Multi-hazards scenario generator: A network-based simulation of natural disasters, *Risk Analysis*, **41**(11), 2154–2176, doi: 10.1111/risa.13723. 3

González-Mellado, A. O. & De la Cruz-Reyna, S., 2010. A simple semi-empirical approach to model thickness of ash-deposits for different eruption scenarios, *Natural Hazards and Earth System Sciences*, **10**(11), 2241–2257, doi: 10.5194/nhess-10-2241-2010. 6

Kawabata, E., Bebbington, M. S., Cronin, S. J., & Wang, T., 2013. Modeling thickness variability in tephra deposition, *Bull. Volcanol.*, **75**(8), art738, doi: 10.1007/s00445-013-0738-x. 6

Mastin, L., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, A., Ewert, J., Neri, A., Rose, W., Schneider, D., Siebert, L., Stunder, B., Swanson, G., Tupper, A., Volentik, A., & Waythomas, C., 2009. A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions, *J. Volcanol. Geotherm. Res.*, **186**(1-2), 10–21, doi: 10.1016/j.jvolgeores.2009.01.008. 6

Mignan, A., Wiemer, S., & Giardini, D., 2014. The quantification of low-probability-high-consequences events: Part I. A generic multi-risk approach, *Natural Hazards*, **73**(3), 1999–2022, doi: 10.1007/s11069-014-1178-4. 3

Unprocessed References

1.

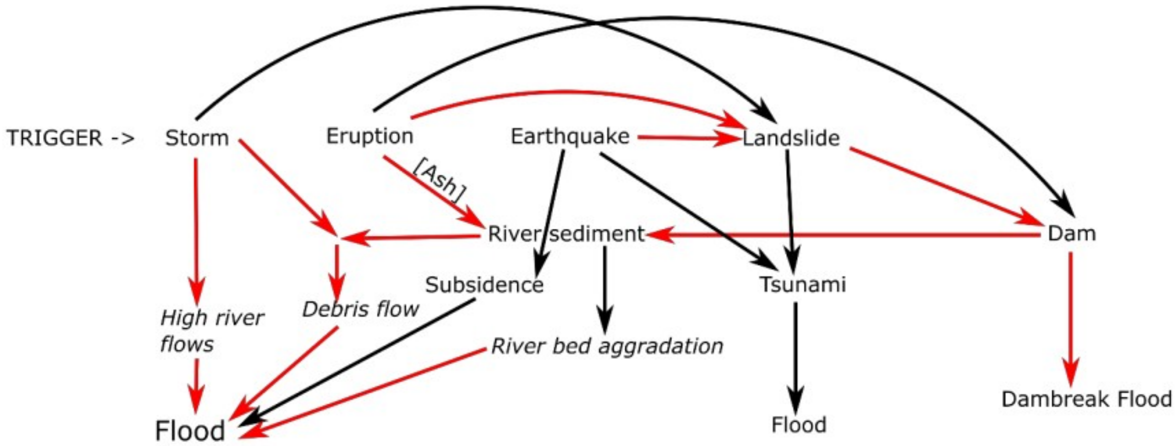

Figure 1: Potential hazard cascades leading to flooding: red arrows indicate cascades in Case Study (Davies et al., 2020).

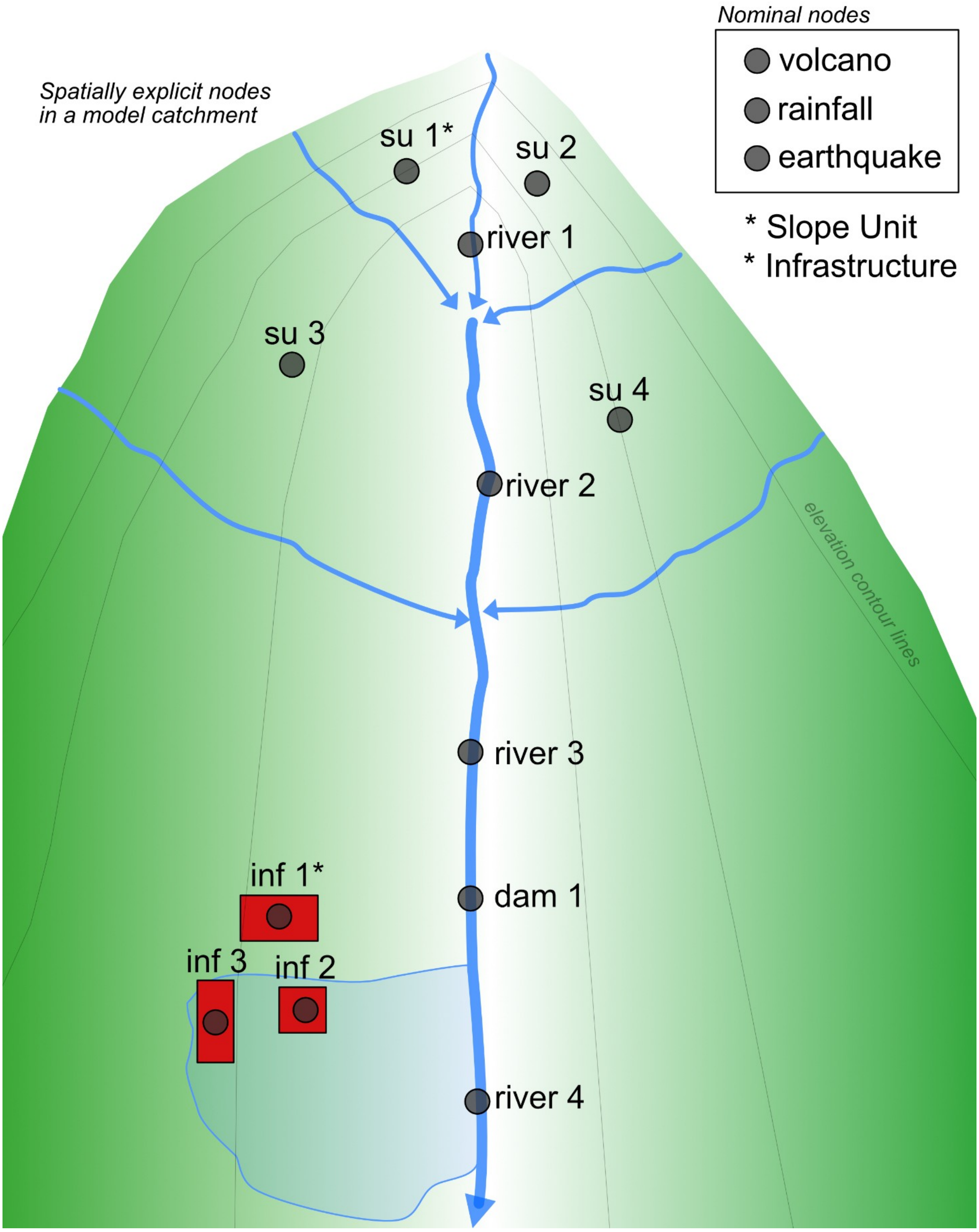

Figure 2: Node proposed locations in a virtual catchment. The blue lines represent river streams of different orders. The blue transparent outline represents potential flooding.

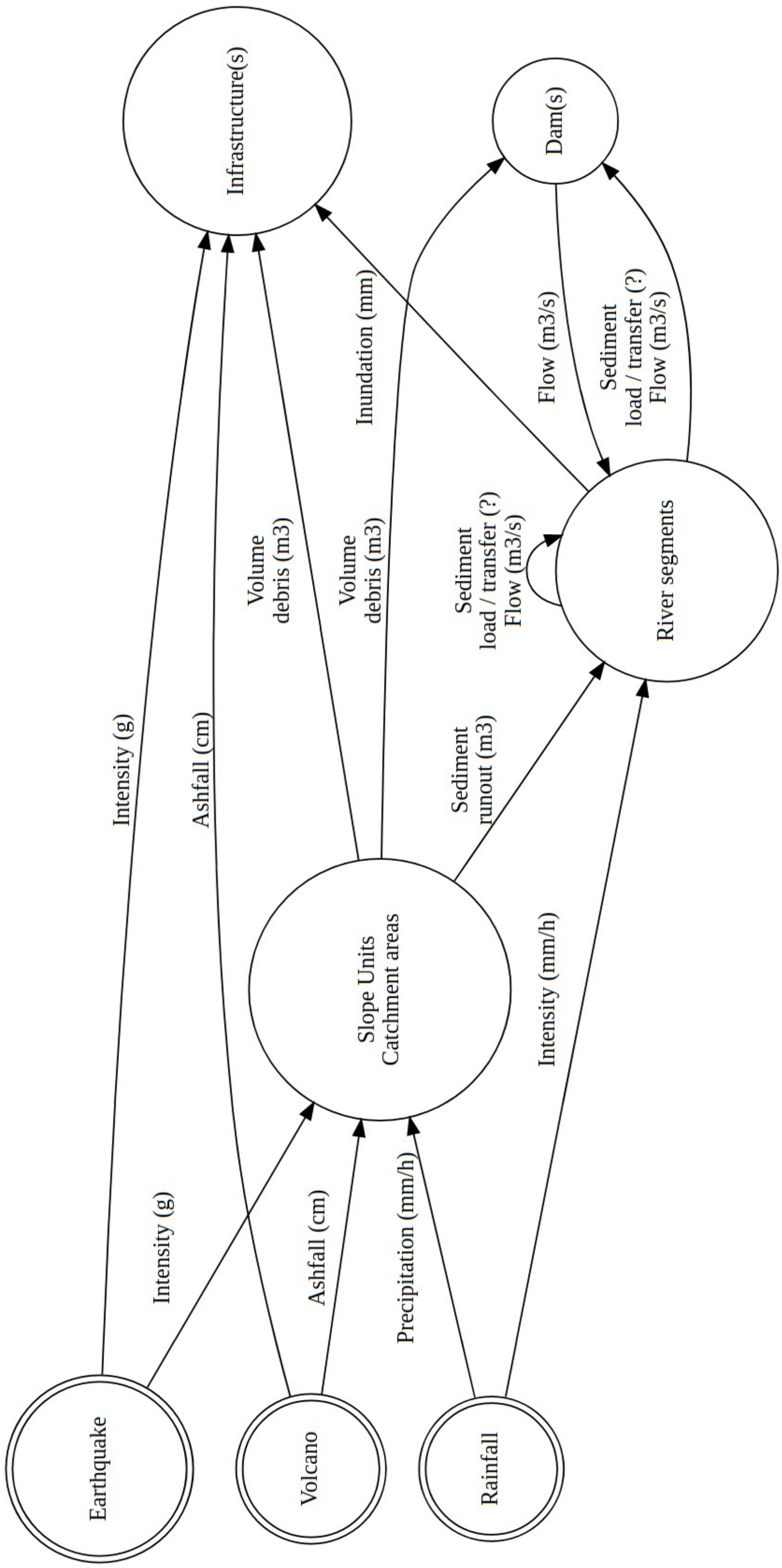

Figure 3: Connections between the various components types of the multi-hazard system (the double circles represent the nominal nodes). In this diagram, dams refer to natural dams and man-made dams.

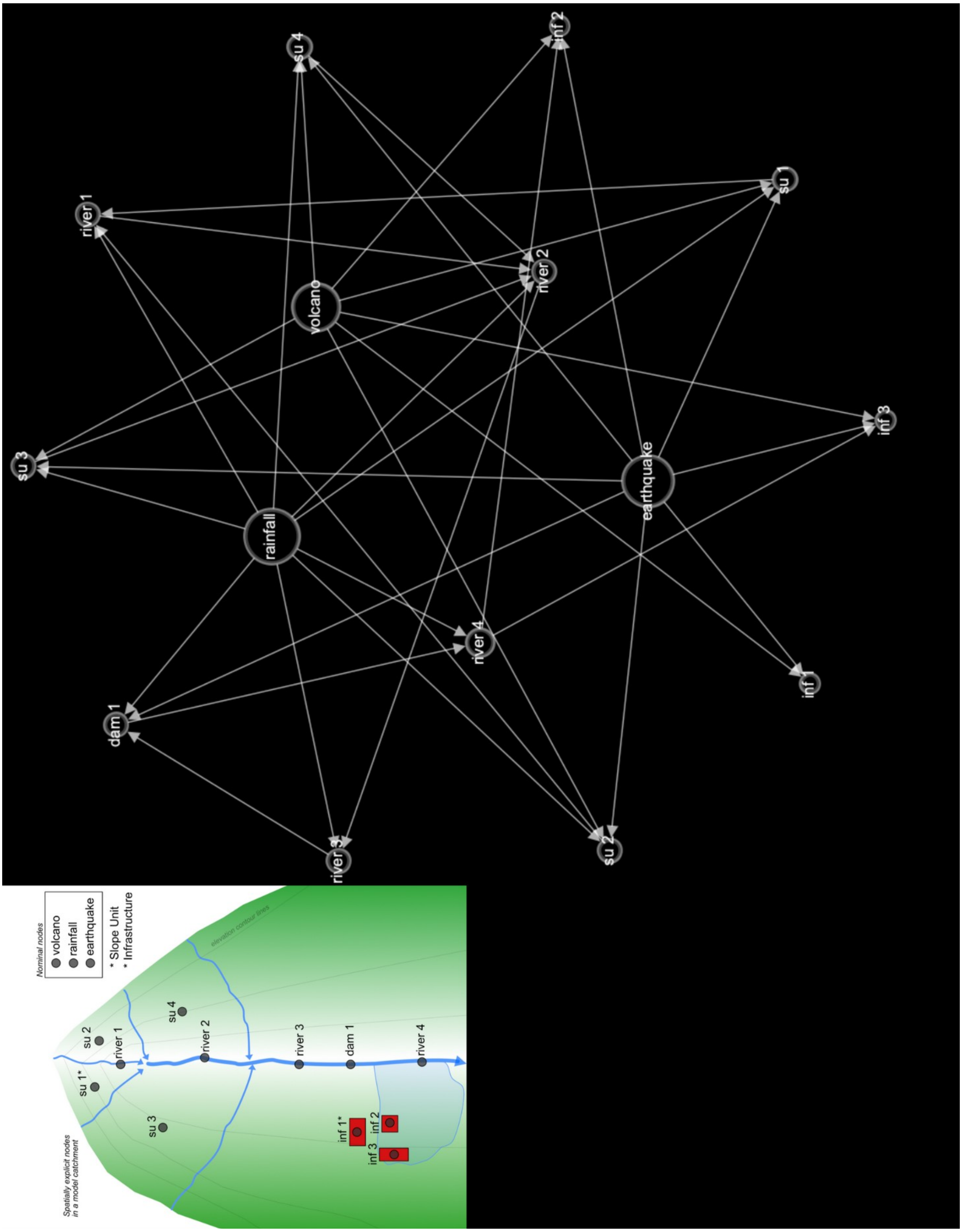

Figure 4: Spatial connections between the various nodes presented Fig 2.

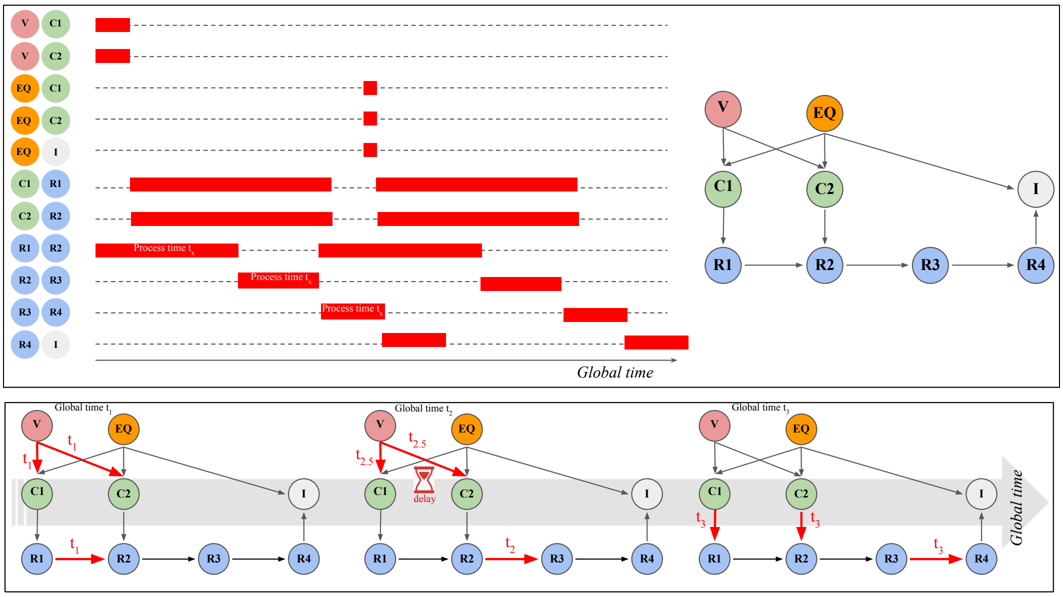

Figure 5: Taking time into account.

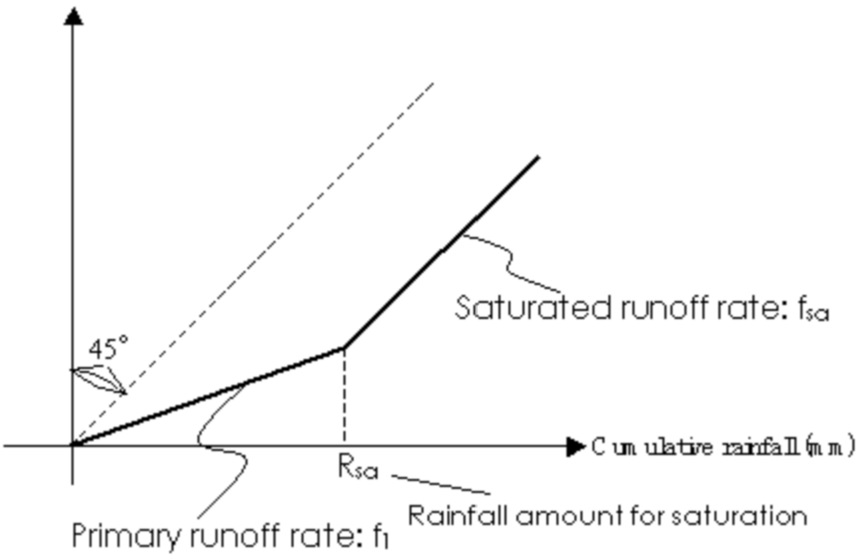

Figure 6: Sediment Runoff.