GUIDE TO FLOOD MODELLING

Authors: Stefan Rimkus, Satya Patnaukuni, Henry Bovy, Claire Souch, Paul Nunn - SCOR

OVERVIEW

Floods occur all over the world, and alongside windstorms are the most frequent cause of natural hazard losses. Worldwide, about a third of all reported events and a third of the economic losses resulting from natural catastrophes are attributable to floods, with regional concentrations in losses as shown in Figure 1 Global Flood Economic Loss Risk (Source: CHRR &CIESIN)Figure 1.

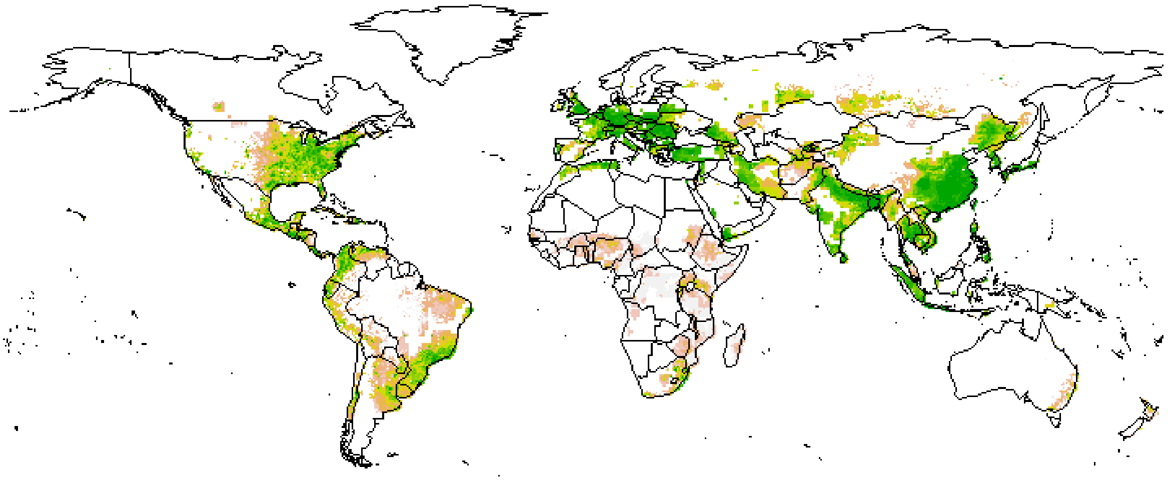

Figure 1 Global Flood Economic Loss Risk (Source: CHRR &CIESIN)

The predominant sources of flooding are as follows, also illustrated in Figure 2**Error! Reference source not found.**.

1. Inland flooding: excess precipitation, including snowmelt, which can trigger flooding from
    - River flooding: Rise of a river to an elevation such that the river overflows its natural banks (termed “fluvial” flooding)
    - Flash flood: Sudden flood of great volume, usually caused by heavy rain where the quantity of water overwhelms the infiltration or drainage capacity or from overflowing streams and rivers or from a dam burst
    - Groundwater flooding: when the water table rises to the surface, for example after prolonged periods of high rainfall
2. Coastal flooding:
    - Storm surge: Caused primarily by high winds pushing on the ocean's surface, especially severe in combination with high tide. The wind causes the water to pile up higher than the ordinary sea levelTsunami: Ocean waves generated by the displacement of water typically by a large underwater earthquake, volcano or landslide.

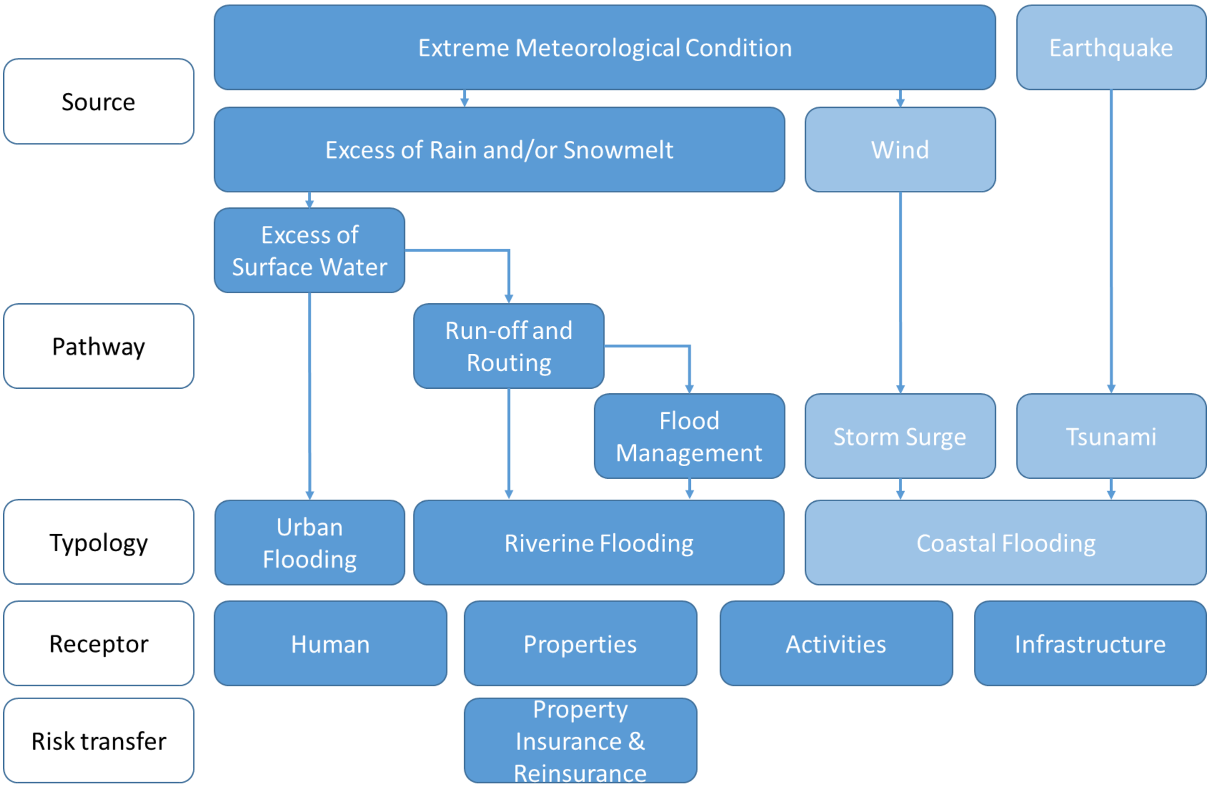

Figure 2 Schematic representation showing sources of flood losses

Of course insured flood losses only occur if the excess water reaches properties, and if that property has flood coverage included under an insurance policy. Property flood coverage often cover coastal and inland flood. Tsunami losses are typically covered by earthquake policies. This article will focus on the dark blue boxes ofFigure 2, namely rainfall-driven inland flooding, and property as the receptor and subject of insurance risk transfer.

Flood coverage varies extensively around the world, and by line of business. The U.K., for example, has traditionally provided flood cover as part of the standard property coverage. Most other countries offer flood as an additional cover. In some countries, federal schemes such as the U.S. National Flood Insurance Program (NFIP) may exist to provide flood cover for homeowners and small businesses.

This document will not provide a comprehensive view on flood insurance coverages globally, but it is helpful for those building catastrophe flood models to understand the coverages in place by line of business in the country's for which they are building models. This will help with defining the scope of a model, knowing where to focus most effort and attention in the development, and to estimate the potential demand for a model: bearing in mind that interest might be high for a region of the world where flood cover may be about to change, even if coverage is currently low.

TYPES OF FLOOD MODELS

Scenario models:

These models are useful to find out how an area may flood in different scenarios such as a dam failure, dike breach, and 100 year return period flood etc. This can be done with a computer simulation model which models how the water spreads over the area and water depth at different locations on the basis of the terrain height. They do not attempt to represent the full suite of plausible flood events.

Stochastic models:

The process of generating thousands of plausible flood events and estimating their return periods, flood extents and water depths at different locations is known as stochastic flood modelling. The first step is the creation of a stochastic event set. These events can be generated using historical (observed) flood data and RP-flow relationships or historical rainfall and rainfall-runoff modeling. There are pros and cons of each methodology, and an ongoing debate about which is best.

The first approach estimates RP-flow relationships across the river network based upon either rainfall or flow gauge observations. From these, a series of return-period floodplain maps along with a synthetic set of maximum river flows, also at defined return-periods, at points along the river network are generated. These two components are linked together to form the event set of flood footprints.

The main benefit of rainfall-runoff modelling is that it’s a more realistic representation of how the hydrological system works, and therefore is technically a more robust way of modelling pluvial as well as fluvial flooding with realistic correlations both spatially and temporally. Additionally, rainfall records may extend further back in time, and be more extensive, than river-gauge records. However, they are mainly point measurements and it’s difficult to spatially extrapolate them.

The proponents of RP-flow based modelling note that the uncertainty associated with rainfall-runoff modelling is extremely high, and may outweigh the benefits of the approach. The full hydrological system is very complex, and there are a lot of processes to model with such a methodology, requiring sophisticated approaches, high resolution modelling and a lot of computing power, and each with their own uncertainty involved.

Thus, a flood model developer will need to choose which method to use. The decision will be influenced by the availability and quality of historical data, the availability of the required compute power, the modeller’s particular specialism, and the resolution of input data at which the various processes can be modelled.

HAZARD MODELLING

Hazard modeling is the process of estimating the flood extent and depth associated with each event within the stochastic event set, or for specific scenarios. Flood is a complex hazard to model and very data intensive. The interaction between precipitation and river discharge is dependent on many attributes, such as catchment size and shape, topography, soil type, land use – land cover, and antecedent conditions (whether the ground conditions are already wet or dry depending on recent weather conditions). Flood is also high gradient peril, meaning that flood depths can vary significantly over short distances depending on terrain. Therefore, a model with a large spatial modelling grid (aka "low resolution") may not adequately model this variability, and may overestimate the flood risk by averaging flood depths over larger areas

There are two main components in flood hazard modelling:

- Hydrologic modelling: Modelling a drop of water from where it falls on the land, to the stream.
- Hydraulic modelling: Modelling of the movement of water, for example downstream along river channels or across the land surface

Hydrologic Modeling

Simple Runoff Modeling Methods

a. **The Rational Method:**

Estimates only the peak flow using rainfall intensity (I), catchment area (A), and a land use factor (C). There is no timing associated with the peak flow. Simple method with limited applicability for flood catastrophe models.

Complex Hydrologic Models

There are three main categories of complex hydrologic models, as illustrated inFigure 3.

a. **lumped models**

Simple & minimal data required and easy to use. Provides estimates of flow values at the basin outlet. However, given that in reality basin parameters vary spatially, this method has limited practical applicability for catastrophe modelling.

b. **semi – distributed models**

Parameters are partially allowed to vary in space. River basin is divided into a number of smaller sub basins. Less demanding on input data than distributed models. Model structure is more physically based than lumped. Eg. SWMM, HEC HMS, SWAT

c. **distributed models:**

Parameters are fully allowed to vary in space at a resolution chosen by the user. Requires large amount of data, high computational time, and experts. Governing physical processes are modelled in detail. Results available at any location & time. Eg. MIKE11/SHE, WATFLOOD

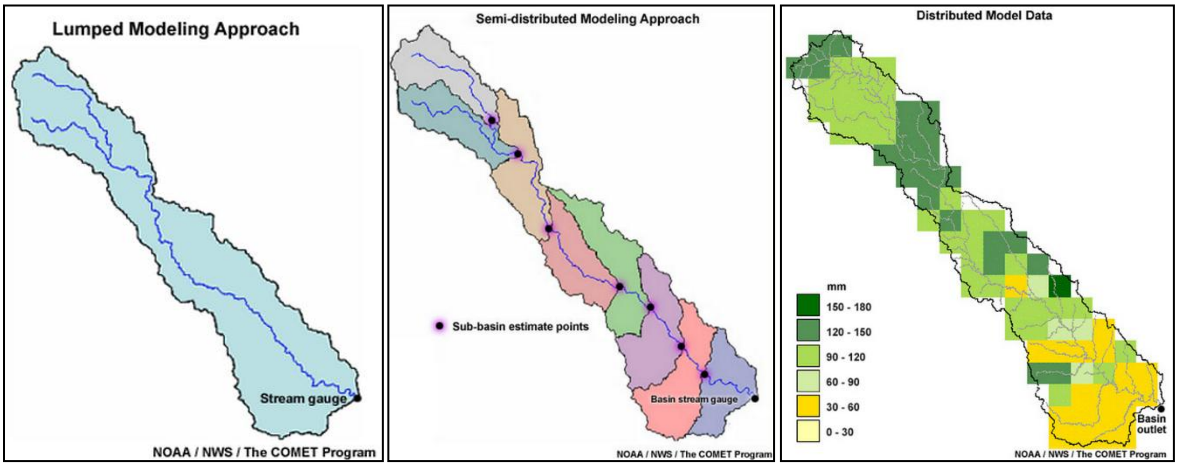

Figure 3 Illustration of three approaches for complex hydrological modelling

Complex Statistical Methods

A more recent approach to flood modelling is to skip the actual hydrological modelling and instead estimate possible stream flows using statistical methods such as extreme value theory. With these flows a flood extent can be estimated given the channel properties at a given location. The method is restricted to riverine flood considerations.

Hydraulic modelling

To model the flow of water along the channel, hydraulic modelling is used to estimate the flow parameters of volume, velocity, and depth of flow at different locations along the channel. At each location the capacity of the channel is calculated, along with the flow of water. Flooding will occur if the capacity of the channel is exceeded. Hydraulic modelling is also used to calculate inundation depths over land.

There are 1D and 2D (one dimensional and two dimensional) hydraulic models which can be applied to modelling river channel flows as well as inundation modelling. In 1D models, the parameters are estimated in a single direction. The flow is assumed to be unidirectional - generally happening in the direction parallel to the main channel flow. 2D models are based on shallow water equations, and offer a more realistic representation of water flows, particularly over relatively flat areas for inundation modelling. The down side is that 2D models are computationally very demanding, and require high quality and high resolution terrain data inputs. A number of commercial software packages include the possibility to link a 1D river model to 2D floodplain grids. This has become popular in recent years because it allows the modeller to take advantage of the established tradition of 1D river modelling while at the same time modelling floodplains in two dimensions.

DATA SOURCES

Flood modeling is very data intensive. A list here is provided of the main data sources required, and some links to publicly available data.

One of the biggest discussions in flood modelling is that of the resolution and quality of the Digital Elevation Model (DEM), or Digital Terrain Model (DTM), as illustrated in **Error! Reference source not found.** .The DTM provides a digital 3D representation of the earth's surface elevation, over which the flooding is modelled, without the influence of buildings, man-made structures, vegetation etc (which are incorporated within a Digital Surface Model). DTMs suffer from a number of large number of problems, including gridding artifacts, contour artifacts, and mismatches with other data sources. This can result in false sinks or pits, or obstructions caused by remnants of bridges and buildings which interfere with the hydrologic and hydraulic flow calculations of river and surface water flows. Often the DEM must be scrubbed manually in order to remove traces of buildings and bridges from the data before being used for flood modelling.

DTMs are derived from various sources such as photogrammetry, LiDAR, IFSAR and other less commonly used sources. Some large scale surface models have been made available free of charge to the public such as the NASA Shuttle Radar Topography Mission DEM at 3 second grid (approximately 90m) and ASTER DEM 1.5 second grid which cover most of the planet.

Other important data inputs are listed here, but not discussed in detail, with some sample sources:

Watershed boundaries	http://www.ncgc.nrcs.usda.gov/products/datasets/watershed
Soils	http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo
	http://soildatamart.nrcs.usda.gov
Land use and land cover	http://seamless.usgs.gov
	http://glcf.umd.edu/data/aster
Current and historic water records	http://waterdata.usgs.gov/nwis
	http://www.epa.gov/STORET/index.html
Climate, weather, rainfall	http://www.ncdc.noaa.gov/oa/ncdc.html
	http://www.nws.noaa.gov/ndfd

RESOLUTION

Flood model resolution is an important consideration. As noted previously in this guide, flood is a ”high gradient” peril, which benefits from high-resolution modelling – that is modelling on a small spatial scale. Flood model resolution depends on the resolution of the grid that the model development is conducted at, the resolution of the input terrain data, and the resolution that the hazard is ultimately compiled and stored at in the final software files. It is of little benefit to the end user to have a flood model which is output at 10m, if the input files and model development is conducted at 90m. This can lead to a false sense of accuracy on the part of the end-user.

Flood catastrophe model resolutions today vary from 10m in the absolute best case for the U.K., to 90 m and more. Models may be hybrid, or use a “variable resolution grid”, that is a grid which varies from high resolution where the hazard gradients are high, or there is a lot of property exposure concentrated, to low resolution where there is less property, or low hazard gradients, or insufficient quality of terrain data.

FLOOD DEFENCES AND HUMAN INTERACTION

A significant influence on the development of floods, and a source of significant uncertainty in flood models, is that of human-made flood control systems and other constructions that influence the flood propagation. Dams, sewage systems, river engineering, flood defenses are examples that influence where water flows.

Flood defences are built to protect the very property that is of interest in a catastrophe risk flood model. Thus including the influence of flood defences is critical. However, it’s also very challenging as information on flood defences is often incomplete or not available. Also, flood defences may fail, and human intervention often plays a part in terms of managing flood defences real time and other control mechanisms, such as deliberately allowing one area to flood in order to “save” another. For flood modelling, we need to know both the location and the standard of protection (SOP) provided, i.e. the return period of flood up to which the defence is designed to protect against. Floods above this level will overtop the flood defences.

In some countries, flood defence information can be obtained from the relevant authority, such as the Environment Agency in the U.K. If such a database is not available, it may be possible to identify the presence of flood defences manually from very high resolution LIDAR based DEM data. If the SOP, or the presence, of flood defences is not known, some model developers apply assumptions based on the usage of the land, for example city centres will typically have a higher SOP than rural agricultural land, and nature reserves etc. may have no protection.

The other question with modelling flood defences is their failure. That is a collapse of part or all of the flood defence, at flood severities below its design standard. To model this physically requires very detailed knowledge of the physical condition of every flood defence – and is not feasible. . Some modellers apply a stochastic failure model to account for the probability that this will occur. Other modellers prefer to not apply any assumptions in this regard due to the extreme uncertainty of the problem. There is the additional complexity of deployment of mobile flood defences. Again a modeller may make some assumptions based on research and knowledge of where these exit, or may decide that the uncertainty is too great over their deployment and success in order to incorporate them.

VULNERABILITY AND DAMAGE MODELLING

Damage to a property or physical asset is calculated as a ratio of the loss to the value of the property - via vulnerability curves. Vulnerability curves describe the relationship between hazard – typically relative to the depth of flooding, maybe with some modification for velocity – and damage ratio.

Standard vulnerability curves have been created by organisations such as the U.S. Army Corps of Engineers, the Middlesex University in the U.K. or the Germany HOWAS project based on surveys after floods. Insurance claims data (along with the relevant insured value at the time) is beneficial in order to further calibrate these vulnerability curves to actual loss experience.

Flood losses are dependent on a number of building characteristics, such as:

- Property height/number of floors e.g. contents above the 1st floor may not be damaged
- Type of structure e.g. bungalow, detached house, mobile home, warehouse, etc.

- Occupancy, or the nature of the business, which influences the type and value of contents
- Building material e.g. wood, brick, etc.
- The presence or absence of a basement or cellar
- Local flood defence, the presence of a raised front door step/entry above street level or raised first floor
- Age of the property which determines the building code

In addition, regional variations in building codes and construction practices mean that a property with the exact same characteristics can have a different vulnerably to the same hazard, if in a different region or country. Thus, catastrophe models have many different vulnerability curves to represent the multiple permutations of the above, and other, characteristics. Some of these vulnerability curves may be created based on engineering judgement relative to original curves developed from actual data (damage observations or claims).

The loss of use of the building is typically also covered by an insurance policy, through either business interruption cover or alternative living expenses for residential properties. This is typically modelled as a time-function relative to the original damage severity.

FINANCIAL MODEL

The financial model transforms damage ratios into insured loss estimates. An insurance or reinsurance company is primarily interested in understanding the value of the claims they will be liable for. Thus, the model needs to be able to take into account the impact of the insurance policy coverage terms and conditions in place, for example the impact of deductibles or excess and any reinsurance cover or policies. There are many complex re/insurance policy types that are beyond the scope of this document to cover. By providing models via an open platform such as OASIS means that a model developer will be able to use a common simulation financial loss modelling engine.

EXPOSURE DATA QUALITY

A challenge that catastrophe model developers must be aware of is that of data quality – specifically referring to the data about the insured properties within a re/insurance portfolio. The data held by the insurance/reinsurance company may be incomplete, i.e. with missing or unknown building characteristics from the list above. Thus to be of use, a catastrophe model will typically include some assumptions about the building type and characteristics when attributes are unknown. This is often done by creating a building inventory, i.e. a profile of the building stock built from sources such as census data, LULC data, satellite imagery, public and private records, which is used to create a “default” or “blended” vulnerability curve.

The other element of poor data quality is that of geo-resolution. This is particularly important for high-resolution flood models, where the flood may be modelled on a grid of 1x1 km^2. In reality, insured values may be collected at the individual property level, or at more aggregate levels such as by post code. There are two main ways this can be handled in a catastrophe model. One is to compile the final hazard model at different resolutions. Typically a hazard model will be constructed at the highest resolution possible, such as a 10m grid (if supported by available DEM and modelling methodologies), and then aggregated as average values across lower resolution units such as post codes. If using this approach, then the model developer needs to be cognizant that the model user, the re/insurance company, is primarily interested in flood losses to properties. Thus, over larger areas, the average hazard depth might be weighted based on an understanding of where properties are likely located within that area, so that it is representative of the flooding that will be incurred by the assets of interest.

The alternative method is to output the hazard model on a high resolution uniform grid, and create an algorithm to disaggregate the exposure data from a postal code to a higher resolution grid. This can be based on using Land Use and Land Cover data (LULC) for example, to create informed assumptions about where property is actually located within larger areas.