Combining statistical and expert evidence within the D-S framework: Application to hydrological return level estimation

Nadia Ben Abdallah[1] Nassima Mouhous Voyneau[2] and Thierry Denoeux[3]

Abstract Estimation of extreme sea levels and waves for high return periods is of prime importance in hydrological design and flood risk assessment. The common practice consists of inferring design levels from the available observations and assuming the distribution of extreme values to be stationary. However, in the recent decades, more concern has been given to the integration of the effect of climate change in environmental analysis. When estimating defense structure design parameters, sea level rise projections provided by experts now have to be combined with historical observations. Due to limited knowledge about the future world and the climate system, and also to the lack of sufficient sea records, uncertainty involved in extrapolating beyond available data and projecting in the future is considerable and should absolutely be accounted for in the estimation of design values.

In this paper, we present a methodology based on evidence theory to represent statistical and expert evidence in the estimation of future extreme sea return level associated to a given return period. We represent the statistical evidence by likelihood-based belief functions [7] and the sea level rise projections provided by two sets of experts by a trapezoidal possibility distribution. A Monte Carlo simulation allows us to combine both belief measures to compute the future return level and a measure of the uncertainty of the estimations.

[1] Nadia Ben Abdallah
Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc
nadia.ben-abdallah@hds.utc.fr

[2] Nassima Mouhous-Voyneau
Université de Technologie de Compiègne, Avenues-GSU
nassima.voyneau@utc.fr

[3] Denoeux Thierry
Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc
thierry.denoeux@hds.utc.fr

1 Introduction

Comprehensive uncertainty analysis is a key part of design and safety assessment procedures for reliable results and optimal decision. In the hydrological field, communicating the uncertainty about future flood risk to the decision makers is becoming the rule rather than the exception [1, 12]. If there is a general consensus about the relevant sources of uncertainty within a flood risk analysis, there is an increasing debate among risk analysts about the framework to use for quantifying it. The commonly used probabilistic framework has been strongly criticized for treating in the same way aleatory uncertainty that characterizes natural variability and epistemic uncertainty resulting from limited knowledge [1]. Given that, in environmental risk analysis, these uncertainties usually arise from different sources (statistical evidence, expert opinion...), the need for alternative frameworks to address differently both kinds of uncertainty emerged. Intensive works are investigating the appropriateness of approaches such as fuzzy set theory, imprecise probability or Dempster-Shafer theory in assessing reliability and risk analyses.

In this paper, we are interested in modeling the uncertainty pertaining to the design parameters of flood defense structures in the context of future climate change. Evidence for estimating the parameter of interest and its uncertainty originates from two sources of different natures. The first one is related to statistical evidence commonly expressed by frequentist or Bayesian approach, the relevance of which has been increasingly criticized [5, 8]. The second one concerns projections of climate change and its impacts in terms of sea level rise, which have to be assessed by climate experts. Partial disagreement about the future climate change within the climate community leads, as will be showed later, to a high level of uncertainty attached to the projections available in the literature.

We propose to represent and combine the two different sources of evidence (data and experts) using the DS framework, given its ability to address in a unified mathematical context different sources of evidence and the tools it offers to combine them.

The paper is organized as follows. In the first section, we review the use of extreme value statistics in hydrology and the characteristics of hydrologic extremes in flood design. We briefly address the issue of climate change impacts and present the main projections on the future sea level rise existing in the literature. In the second part, we justify and explain the use of likelihood-based inference to represent statistical evidence and briefly address its connection with the DS framework. Finally, the last part describes the application of the methodology and summarizes the main results.

2 Key elements on hydrology and climate change

Flood structures have to withstand exceptional sea events and their design has thus to be based on extreme sea level and waves. The main tool for modeling extreme events in environmental applications such as floods, droughts or rainfalls is Extreme Value Theory (EVT), which has emerged giving the limit of the conventional frequency analysis in fitting the tails of probability distributions. The block maxima approach is the original and best known method in EVT. It is based on the assumption that the maximum of an independent and identically distributed (i.i.d.) sample has asymptotically a generalized extreme value (GEV) distribution [11]. The cumulative distribution function of the GEV distribution is given by:

$$F(z,\mu,\sigma,\xi) = \begin{cases} \exp\left(-\left[1-\xi\frac{z-\mu}{\sigma}\right]^{\frac{1}{\xi}}\right) & \text{for } \xi \neq 0 \\ \exp\left(-\left(\exp\left[-\frac{z-\mu}{\sigma}\right]\right)\right) & \text{for } \xi = 0, \end{cases} \tag{1}$$

where $\mu, \sigma > 0, \xi$ are, respectively, location, scale and shape parameters. According to the sign of $\xi,$ the distribution is called Fréchet ($\xi > 0$), Weibull ($\xi < 0$) or Gumbel ($\xi = 0$).

In extreme-values studies, the probability of exceedance of a certain value *z* is usually expressed in terms of the return period *T*, defined as the average number of years between two successive exceedances of the corresponding return value *z*. Within the annual maxima method, the return period of a given level *z* is directly related to its annual non exceedance probability *p* by the relation: $T = 1/(1-p)$. Therefore, we get from (1) the following expression of the return level z_T associated to a given return period *T*:

$$z_T = \begin{cases} \mu - \frac{\sigma}{\xi}[1-\left(-\mathrm{Log}\left(1-\frac{1}{T}\right)\right)^{-\xi}] & \text{for } \xi \neq 0 \\ \mu - \sigma\,\mathrm{Log}(-\mathrm{Log}(1-\frac{1}{T})) & \text{for } \xi = 0, \end{cases} \tag{2}$$

The only available evidence when estimating extreme quantiles is derived from the historical observations.

Commonly, flood defenses in coastal areas are designed to withstand at least 100-year events. However, due to climate change, they will be subject during their life time to higher loads than the design estimations. The main impact is related to the increase of the mean sea level which affects the frequency and intensity of surges. For adaptation purposes, the present statistics of extreme sea levels derived from the observations should be combined with the projections of the future sea level rise (SLR)

Future SLR projections provided by the IPCC's (International Panel of Climate Change Experts) last Assessment Report [10] assess the likely range of values for sea-level rise over the 1990-2095 period as 0.18 to 0.79 m; it is indicated in this report that higher values should not be excluded. This range takes into account uncertainties associated to future emissions of greenhouse gases (GHGs) corresponding to the SRES (Special Report Emission Scenarios) (scenarios that cover a wide range of possible economic, technological and energetic states of the world), global circulation models used to estimate future temperature projections and impacts models (melting of the Antarctic and Greenland, oceans expansion, etc.).

Since the release of the last IPCC report, other sea level rise assessments based on semi-empirical models have been undertaken, proposing more pessimistic sea level rise scenarios for 2100. For example, based on a simple statistical model, Rahmstorf [15] suggests [0.5m, 1.4 m] as a likely range of values for sea-level rise at the end of this century. However, recent studies showed that there is a physical limit to the sea level rise in the coming years: the threshold of 2 m could not be exceeded by the end of this century [13].

Current methods for integrating future SLR in flood risk or design analysis have considered a deterministic particular sea level rise scenario since there is no information to quantify the probability of any given sea level magnitude within the IPCC range. However, as shown by Purvis [14], who undertook a flood risk analysis under climate change, using the most plausible sea level rise scenario may significantly underestimate effective consequences and lead to erroneous decisions.

For estimating design level under climate change, we proceed in two steps: we first infer the current design level from statistical evidence (available sea level measurements). In a second step, we integrate expert judgment on future sea level rise.

3 Likelihood-based representation of statistical evidence

The estimated level is usually obtained from (2) by replacing the probability distribution parameters by their best estimates. Commonly, parameters are estimated using frequentist methods. However, these methods are based on asymptotic properties and their performance turns to be quite poor when we deal with small samples. As for the estimations, the confidence intervals supposed to inform about the level of uncertainty within the estimations are quite unreliable because of the very crude approximations in the calculation of the upper and lower bounds of the confidence interval [18]. In fact, confidence intervals are based on the repeated sampling hypothesis which consists of hypothetically repeating the particular experiment and derive accordingly the confidence bounds. In cases such that the repetition is not possible, this approach can be questioned and alternative approaches to effectively represent the available evidence are needed.

Authors such as Fisher [8], Cox [5], Barnard *et al.* [3] and Edwards [7] have criticized the frequentist approach for its inappropriate use of significance levels, confidence intervals and other repeated-sampling criteria to represent evidence and have advocated a new, more ‘evidential’ approach to statistical inference that uses only the likelihood function.

The likelihood-based inference approach relies on the likelihood principle, which states that given an observation X, the relevant information about an unknown parameter $\theta(\theta \in \Theta)$ (possibly a vector) is all contained in the likelihood function for the observed sample X, denoted $L(\theta; \mathrm{X})$. Recall that the likelihood is a function of the parameters of a statistical model $f(x; \theta)$ defined as follows: given some observed outcomes, the likelihood of a set of parameters is equal to the probability of the observations given those parameters. Thus $L(\theta; \mathrm{X}) = f(X; \theta)$.

The representation of statistical evidence in the belief function framework is motivated by the fact that belief functions form a richer set of functions than probability measures: it is thus expected that inference, when based on belief functions, would allow us to model a wider range of uncertainty than probabilities. Shafer [17] was the first to propose to represent likelihood information as a consonant belief function about the parameters. Shafer’s method was later justified axiomatically by Wasserman [19]; additional arguments for its use in statistical inference were provided by Aickin [2]. Fisher [8] interprets the likelihood function as an expression of the relative plausibility of the parameters when no additional information, except the observations, is available. It thus seems reasonable to define the plausibility contour function (or credibility function), when the likelihood is bounded, as:

$$pl(\theta; X) = \frac{L(\theta;X)}{L\,(\hat{\theta};X)} \tag{3}$$

where $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ. The associated plausibility is easily computed for every subset $A \in \Theta$ as :

$$Pl(A) = \sup\,\{pl(\theta; X);\ \theta\ \in A\ \} \tag{4}$$

The contour function (3) has a simple interpretation: $pl(\theta; x)$ represents the probability of observing x if the true parameter value is θ, relative to the maximum probability of observing x for any value of θ. A parameter value with low plausibility, say 0.001, indicates that there are other values of θ which ensure a 1000 times higher probability to observe x.

The set $\{ \theta \in \Theta \,/ pl(\theta, X) \geq \alpha\}$, called the α-level likelihood region, allows us to characterize ranges of implausible values (for example, values ranging outside 5% likelihood region) and very plausible values.

4 Application and results

As a case study, we applied the likelihood-based inference method described above to infer the design variable z_{100} from the sample of observations X corresponding to 15 years of hourly records of sea level (observations from tide gauges at le Havre harbor, France). This measure was estimated under the assumption that the annual maxima have a Gumbel distribution (2); here, μ is the structural parameter and σ the nuisance one. The latter was eliminated through a profile likelihood approach. The corresponding contour function is shown in Figure 1.The most plausible value characterized by a plausibility level equal to 1 corresponds to the maximum likelihood estimate.

In a second step, we integrated the uncertain effect of climate change in terms of SLR (in meters) to estimate the future return level (at the end of the century) associated to the same return period. As sources of information about the SLR, we considered projections by the IPCC and Rahmstorf [15] estimations provided above as the current best available evidence. We formalized each of these pieces of evidence by a trapezoidal possibility measure that represents our best interpretation of the expert's estimations (Figure 2). Since both sources are reliable, a conjunctive aggregation is applicable. Among the conjunctive rules, the minimum is the most appropriate when the opinions of the experts are based on a common knowledge: we thus applied this rule to derive the aggregated SLR possibility distribution.

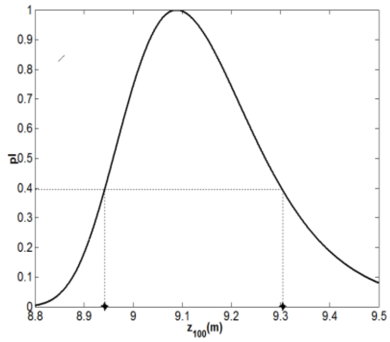

Fig.1 Plausibility measure of the design parameter z_{100}

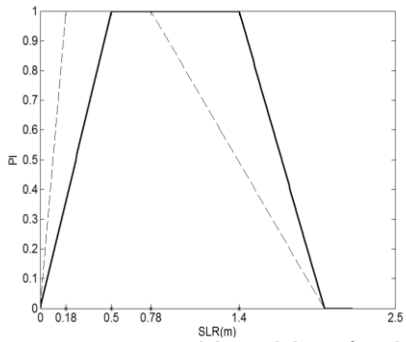

Fig.2 SLR Trapezoidal Possibility inferred measures (in continuous bold line: inference based on Rahmstorf evidence. In dashed line: inference based on IPCC evidence)

Finally we computed the belief function on the future design level $z_{100}^{f} = z_{100} + SLR$ using a Monte Carlo sampling procedure. This procedure consists in randomly drawing plausibility levels α and possibility levels ω using independent uniform distributions For every random α and ω, we associate the α and ω likelihood regions [$\underline{z_{100}^{\alpha}}$, $\overline{z_{100}^{\alpha}}$] and [$\underline{SLR^{\omega}}$, $\overline{SLR^{\omega}}$]; the corresponding design level z_{100}^{f} is within [$\underline{z_{100}^{\alpha}} + \underline{SLR^{\omega}}$, $\overline{z_{100}^{\alpha}} + \overline{SLR^{\omega}}$]. This procedure was repeated a thousand times. From these simulated intervals, we can calculate for a fixed level z_{100}^{f}, the cumulative plausibility and belief. The cumulative plausibility corresponds to the relative frequency, over the simulations, of the event "the lower bound is less than the fixed level", whereas the cumulative belief corresponds to the relative frequency of the event "the upper bound is less than the fixed level". Figure 3 shows the cumulative plausibility and belief functions of the current and future return level (respectively in dashed and solid line). The upper curve corresponds to the plausibility function and the lower one to the belief measure.

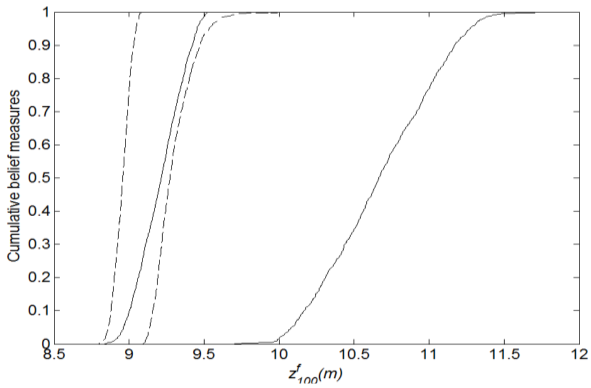

Fig. 3 Cumulative belief functions for the current (in dashed line) and future (in solid line) return level: the lower distribution is the belief; the upper one is the plausibility

The area between the belief and the plausibility dashed curves can be interpreted as a measure of the total uncertainty. When climate change is considered in the estimation of the future level, the area becomes very large, reflecting the important uncertainty associated with the SLR projections.

6. Conclusion

The Dempster-Shafer theory of belief functions places emphasis on the representation of evidence for evaluating degrees of belief. The generality and flexibility of this framework makes it suitable for representing and combining expert judgments and statistical evidence. In this paper, this approach has been applied to the estimation of the centennial sea level at a particular location, taking into account historical data and expert assessments of sea level rise by the end of the century. This work is part of a larger project that aims at defining engineering design processes for the adaptation of coastal infrastructure to climate change.

References

1. Apel, H., Thikeken, A.H.: Flood risk assessment and associated uncertainty. Natural Hazards and Earth Sysem Sciences. **4**, 295-308 (2004)
2. Aickin, M.: Connecting Dempster-Shafer belief functions with likelihood based inference. Synthese. **123**, 347-364 (2000)
3. Barnard, G. A., Jenkins, G. M. and Winsten, C. B. Likelihood inference and time series. Journal of the Royal Statistical Society, 125(3):321{372, 1962.
4. Coles, S. G., Dixon, M. J.: Likelihood based inference for extreme value models. Extremes. **2**, 5-23 (1999)
5. Cox, D.R. : Some problems connected with statistical inference. Ann. Math. Statistics. **29**, 357-372 (1958)
6. Denoeux, T.: Maximum likelihood estimation from uncertain data in the belief function framework. IEEE Transactions on Knowledge and Data Engineering (to appear), doi: 10.1109/TKDE.2011.201.
7. Edwards, A.W.F.: Likelihood. University Press, Baltimore, USA (1972)
8. Fisher, R. A : Inverse Probability and the use of likelihood. Proceedings of the Cambridge Philosophical Society. **28**, 257-261. CP3 (1932)
9. Gumbel, E. J.: The statistics of extremes. New York, Colombia University Press (1958)
10. IPCC Forth Assessment Report (2007), http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml
11. Jenkinson, A.F.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements, Quart. J. Roy. Meteo. Soc.. **81**, 158-171 (1955)
12. Merz, B. and al.: Flood risk curves and uncertainty bounds. Natural hazards. **51**, 437-458 (2009)
13. Pfeffer, W., Harper, J. and O'Neel, S.: Kinematic constraints on glacier contribution to 21st century sea level rise. Science. **321**, 1340-143, doi: 10.1126/science.1159099 (2008)
14. Purvis, M.: Probabilistic methodology to estimate future coastal flood risk due to sea level rise. Coastal Engineering. **55**, 1062-1073 (2008)
15. Rhamstorf, S.: A semi empirical approach to projecting future sea level rise. Science. **315**, 368-370 (2007)
16. Shafer, G. : A mathematical Theory of Evidence, Princeton University Press (1976)
17. Shafer, G.: Belief Functions and Parametric Models, Journal of the Royal Statistical Society.Series B. **44**, 322-352 (1982)
18. Xu, P. and al.: Uncertainty analysis in statistical modeling of extreme hydrological events. Stochastic environmental research and risk assessment. **24**, 567-578 (2010)
19. Wasserman, L.: Belief functions and statistical inference. The Canadian journal of Statistics. **18**, 183-196 (1990)