Using Dempster-Shafer Theory to model uncertainty in climate change and environmental impact assessments

Nadia Ben Abdallah, Nassima Mouhous-Voyneau, Thierry Denoeux

Using Dempster-Shafer Theory to model uncertainty in climate change and environmental impact assessments

Nadia Ben Abdallah
Heudiasyc, UMR CNRS 7253
Université de Technologie de Compiègne
Compiègne, France
Email: nadia.ben-abdallah@hds.utc.fr

Nassima Mouhous-Voyneau
Avenues-GSU, EA 7284
Université de Technologie de Compiègne
Compiègne, France
Email: nassima.voyneau@utc.fr

Thierry Denoeux
Heudiasyc, UMR CNRS 7253
Université de Technologie de Compiègne
Compiègne, France
Email: thierry.denoeux@hds.utc.fr

***Abstract*—We present a methodology based on Dempster-Shafer theory to represent, combine and propagate statistical and epistemic uncertainties. This approach is first applied to estimate, via a semi-empirical model, the future sea level rise induced by global warming at the end of the century. Projections are affected by statistical uncertainties originating from model parameter estimation and epistemic uncertainties due to lack of knowledge of model inputs. We then study the overtopping response of a typical defense structure due to (1) uncertain elevation of the mean water level and (2) uncertain level of storm surges and waves. Statistical evidence is described by likelihood-based belief functions while imprecise evidence is modeled by subjective possibility distributions. Uncertain inputs are propagated by Monte Carlo simulation and interval analysis and the output belief function can be summarized by upper and lower cumulative distribution functions.**

***Keywords*—Dempster-Shafer Theory, likelihood-based inference, imprecision, possibility theory, structural design, sea level rise, Monte Carlo simulation.**

I. Introduction

Defense structures such as dikes (Figure 1) are designed to defend from flooding under extreme hydrographic conditions. Among the functional performance criteria, wave overtopping discharge is commonly chosen as a design parameter. Wave overtopping is mainly caused by conjunctions of water levels and waves whose extreme combinations at the toe of the structure lead the water to reach and pass over the crest of the wall. This causes flooding with undesirable economic, environmental and social consequences. Commonly, the design critical threshold is the centennial return period discharge, defined as the overtopping discharge level that occurs on average once every 100 years.

Flood exposure may significantly increase in the future due to the mean sea water level (SWL) rise, a direct consequence of global warming. Indeed, the climate scientific community expects the global elevation of the seas by the end of the 21st century to range between 30 cm and 200 cm [21], [26]. Given that the typical design life of a coastal defense structure is 100 years, the projected elevations could have a critical impact on the performance of the existing dikes. Long term planning, mitigation schemes and comprehensive risk analyses thus need to take the sea level rise (SLR) into account to minimize future impacts on coastal infrastructures [27].

The design wave overtopping in a context of changing climate cannot be deterministically predicted due to many uncertainties. The inherent variability of some of the hydraulic parameters (waves, surges) and the imprecision of some others due to limited knowledge (mean SLR) or to poor samples are the primary sources of uncertainty. One can further mention the approximations in modeling the overtopping process. The aggregation of these uncertainties may lead to severe imprecision in the model output. If not accounted for, such uncertainty may result in unreliable design and, consequently, unexpected losses.

Uncertainty analysis is becoming a fundamental part of flood risk analyses [2], [25], [37]. Aleatory (or objective) uncertainty arises from the variability of natural phenomena while epistemic (or subjective) uncertainty results from incomplete knowledge about the system; it can then be reduced by collecting additional data. Whereas probabilistic models can naturally handle aleatory uncertainty, their use for modeling epistemic uncertainty is more debatable. In particular, there exist many situations where the available information is incomplete and not rich enough to allow a full probabilistic analysis. A probabilistic representation is then general unfaithful as it hides the imprecision pervading the data and the true state of knowledge by making subjective assumptions (the choice of a uniform the prior, for instance, in the case of total ignorance). Lately, alternative uncertainty models have been advocated in situations of vagueness and imprecision. One can mention Imprecise Probabilities [38], Possibility theory [12], [40] and Dempster-Shafer (DS) theory [8], [31]. All these well-established theories have proved suitable for modeling uncertainty in diverse types of applications and risk analyses [33].

In this paper, we address the modeling of uncertainties on hydrographic inputs and analyze their effect on a typical structure's response in a context of changing climate. Uncertainty mainly arises from partial knowledge of the future mean SLR induced by the global warming and from statistical estimation of waves and surges parameters. The DS framework will be shown to be suitable for representing both kinds of uncertainties. The rest of the paper is organized as follows. In the next section, we first lay down the theoretical foundations of this work by recalling basic definitions of DS theory and discussing in some details its application to statistical infer-

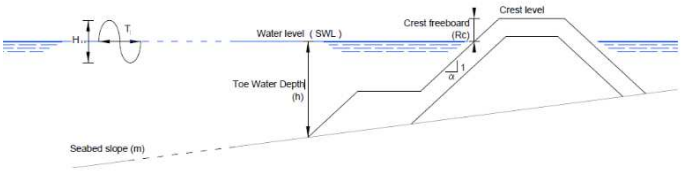


Fig. 1. A typical defense structure and the main overtopping inputs.

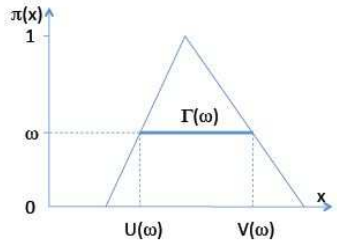


Fig. 2. Random closed intervals induced by a contour distribution.

ence via likelihood-based-belief functions. The semi-empirical model proposed by Vermeer and Rahmstorf [36] to derive SLR projections is briefly presented in Section III. In section IV, we describe in more details the overtopping process, its drivers and the methodology used to estimate the design parameter given input uncertainty. Section V concludes the paper.

II. The Dempster-Shafer framework

The reader is assumed to be familiar with belief functions on finite domains [31]. The background on random intervals and likelihood-based belief functions will be recalled in Subsections II-A and II-B, respectively.

A. Closed random real intervals

Random closed intervals on the real line constitute a simply yet sufficiently general framework for expressing beliefs on a real variable. Let (Ω, A, P) be a probability space and $(U, V) : \Omega \to \mathbb{R}$ a two dimensional random vector such that $U \leq V$ almost surely. Let Γ be the multi-valued mapping that maps each $\omega \in \Omega$ to the closed interval $[U(\omega), V(\omega)]$. This setting defines a random interval, as well as belief and plausibility functions defined, respectively, by:

$$Bel(A) = P(\{\omega \in \Omega; [U(\omega), V(\omega)] \subseteq A\}) \tag{1}$$

and

$$Pl(A) = P(\{\omega \in \Omega; [U(\omega), V(\omega)] \cap A \neq \emptyset\}), \tag{2}$$

for all measurable subset A of the real line. The contour function pl is defined by $pl(x) = Pl(\{x\})$ for all $x \in \mathbb{R}$. The intervals $[U(\omega), V(\omega)]$ are referred to as the focal intervals.

A special case of interest is that of consonant random closed intervals defined as follows (see Figure 2). Let $\Omega = [0, 1]$ and $\pi : \mathbb{R} \to [0, 1]$ be a function, such that, for every $\omega \in \Omega$:

$$\Gamma(\omega) = \{x \in \mathbb{R} | \pi(x) \geq \omega\} \tag{3}$$

is a closed interval $[U(\omega), V(\omega)]$, called the ω-level cut of π. Let P denote the Lebesgue measure on Ω. Then, $[U, V]$ is a random closed interval and π is its contour function, i.e., $pl(x) = \pi(x)$ for all $x \in \mathbb{R}$. Such a random interval is said to be consonant because its focal intervals are nested. The intervals $\Gamma(1)$ and $\{x \in \mathbb{R}, \pi(x) > 0\}$ are called, respectively, the *core* and the *support* of π.

B. Likelihood-based belief functions

Assume that we have observed a realization x of random variable $\mathbf{X}$ with probability density function $p(x, \theta)$, where $\boldsymbol{\theta} \in \Theta$ is an unknown parameter. The parameter inference problem can be addressed using a Bayesian approach, which assumes the existence of a prior probability distribution on Θ, or using a frequentist approach based on the concept of repeated sampling, which relies on hypothesis tests and confidence intervals. A third approach based only on the likelihood function was first suggested by Fisher [17] and later developed in [6], [35] and [13]. A comprehensive survey of likelihood-based methods and their applications was provided by Severini [30].

According to Fisher [17, p. 70], "in the theory of estimation, it has appeared that the whole of the information supplied by a sample is comprised in the likelihood, as a function known for all possible values of the parameter". This view suggests that the information about the unknown parameter θ is entirely comprised in the likelihood, defined by $L(\theta; x) = p(x; \theta)$ for all θ in Θ. More precisely, the likelihood principle states that, within the framework of a statistical model, all the information provided by the observations concerning the relative merits of two hypotheses is contained in the likelihood ratio of these hypotheses on the data. In statistical parlance, the relative likelihood ratio is often referred to as the *relative plausibility* which suggests translating it in the belief function framework as follows:

$$\frac{pl(\theta_1; x)}{pl(\theta_2; x)} = \frac{L(x; \theta_1)}{L(x; \theta_2)} \tag{4}$$

for all $(\theta_1, \theta_2) \in \Theta^2$. Equivalently, the likelihood principle implies that:

$$pl(\theta; x) = cL(x; \theta) \tag{5}$$

for all $\theta \in \Theta$ and some positive c. The least commitment principle [11] then leads us to giving the highest possible value to constant c, i.e., defining pl as the relative likelihood:

$$pl(\theta; x) = \frac{L(\theta; x)}{\sup_{\theta \in \Theta} L(\theta; x)}. \tag{6}$$

The associated plausibility measure is given for every $A \subseteq \Theta$ by:

$$Pl(A; x) = \sup_{\theta \in A} pl(\theta; x) = \frac{\sup_{\theta \in A} L(\theta; x)}{\sup_{\theta \in \Theta} L(\theta; x)}. \tag{7}$$

The corresponding dual measure is referred to the likelihood-based belief function. Equation (7) was first proposed by

Shafer in [31]. It was later justified by Wasserman [39] and in a different way by Denœux [11], on an axiomatic basis.

Let us now assume that the parameter vector can be written as $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \boldsymbol{\theta_2}) \in \Theta_1 \times \Theta_2$, where $\boldsymbol{\theta_2}$ is considered as a nuisance parameter. The likelihood-based plausibility of $\boldsymbol{\theta_1}$ is in this case equal to the relative profile likelihood obtained by marginalizing out the nuisance parameter:

$$pl(\theta_1; x) = \sup_{\theta_2 \in \Theta_2} pl(\theta_1, \theta_2; x). \tag{8}$$

III. Sea level rise assessment

A challenging issue when taking into account future SLR in coastal risk analyses is to address the large uncertainty on climate projections delivered by the scientific community. Indeed, the scientific evidence regarding the SLR is very disparate and highly uncertain, due to the large number and the complexity of inherent processes (thermal expansion, ice sheets melting, mountain glaciers and ice caps melting) involved in the elevation of oceans. Projections range from 0.18 m to 2 m based on physical arguments. The last Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) [21] produced estimates ranging from 0.18 to 0.59 m by the end of the century, relative to the two last decades of the previous century. These projections were based on Atmosphere-Ocean General Circulation Models (AOGCMs), which do not include all relevant dynamic ice processes. In the last decade, the semi-empirical approach, pioneered by Rahmstorf [28], was proposed as a pragmatic alternative to physics-based models. Semi-empirical models provide higher projections than the IPCC: the last dual model proposed by Vermeer and Rahmstorf [36] linking changes in global average near surface temperature and mean sea level yields the range [0.7 m, 1.7 m]. Physics-based studies such as Pfeffer [26] have proposed 2 m as an upper bound for SLR in 2100.

Two major approaches for integrating climate change in flood risk analyses are advocated in the literature. The first one is a typical scenario approach: it consists in estimating exposure and risk for a defined set of scenarios that cover the main situations. Commonly, a mean or an extreme (pessimistic) scenario [7] is considered. However, this deterministic approach may not consider the most relevant scenarios and may lead to erroneous extreme decisions [27]. A more robust approach tries to identify a plausible probability distribution describing the range of sea level projections existing in the literature. By inferring a simple triangular probability distribution over the IPCC last projections [21], Purvis [27] compared both approaches and showed that assessing the risk using the most plausible SLR scenario may significantly underestimate floods monetary losses as it fails to account for the impact of low probability, high consequence events. Purvis pointed out the sensitivity of the risk analysis results to the (subjective) choice of the probability distribution. We may also mention the sensitivity of this approach to the chosen piece of evidence. A more rigorous analysis would have combined different experts' projections and accounted for more extreme scenarios than the moderate IPCC ones. Indeed, in the last decade, measurements have shown that oceans were rising 60 percent faster than the IPCC's latest best estimates: while the IPCC projected sea-level rise to be at a rate of 2 mm per year in the previous decade, satellite data recorded a rate of 3.2 mm per year [29]. Recently, an expert elicitation study [3] has confirmed that the last IPCC projections are underestimated; it provided projections closer to semi-empirical models responses.

Considering these two remarks, it seems more rigorous to address the issue of the projections in the SLR in risk analyses by considering (1) semi-empirical models and (2) a more flexible uncertainty framework than probabilities. In this paper, we consider as SLR evidence the semi-empirical model proposed in [36].

A. Semi-empirical SLR model

The semi-empirical approach is based on linking a driver that can be predicted with a relative confidence (Temperature, CO_2 projections...) to the response of interest, here the mean sea level or the rising rate, via a simple model. The relationship is tested and the model fitted on the observed data. Rahmstorf [28], [36] proposed a semi-empirical model, denoted RV09 in the following, that links the mean global temperature (one of the more accurate and best assessed variables in AOGCMs) to the mean SLR: it is a dual model with two time scales describing long and instantaneous responses:

$$\frac{dH}{dt} = a(T - T_0) + b\frac{dT}{dt}, \tag{9}$$

where H stands for the global mean sea level, a is a proportionality coefficient, b is the instantaneous response proportionality coefficient, T is the global temperature anomaly with respect to 1950-1980 and T_0 the previous equilibrium temperature.

Rahmstorf derived SLR projections for the coming century by applying his fitted model to IPCC temperature projections. His projections range between 70 cm to 170 cm, for different emission scenarios defined by the IPCC. Two major sources of uncertainty pervade these sea level projections: the first one is related to the statistical error of the fit and the second is related to the model input, i.e., the future temperature projections. We propose to address both uncertainties in the DS framework.

1) Constructing a belief function on temperatures projections: The IPCC last Assessment Report used 19 AOGCMs which differ in their formulations of physical processes in the atmosphere, ocean, sea ice and land components and three carbon cycle feedback schemes. Projections are represented by a global average trajectory computed as the mean of the AOGCMs' responses and using the standard carbon cycle setting, as well an upper and lower bounds of the envelope of all the circulations models and carbon cycles. Figure 3 shows a reproduction of the IPCC temperature projections, shifted wrt to 1950-1980 average, for the specific A1B emission scenario. This scenario describes a future world of very rapid economic growth, global population that peaks in the mid-century and declines thereafter, and the rapid introduction of new and more efficient technologies, with a balance across all energy sources.

Since no information is available on the likelihood of each of the AOGCMs, a deterministic or probabilistic combination of these projections would not be justified. To avoid selection error, a comprehensive characterization of the projections should consider the outer envelope of the IPCC published estimations for every scenario. Hence, for every year t in the 2000-2100 time span, the information available is composed of mean (T_t), lower ($\underline{T}_t$) and upper ($\overline{T}_t$) values. The nature

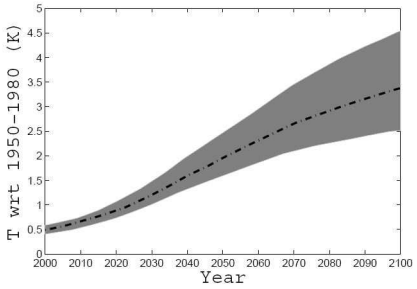


Fig. 3. Reproduction of IPCC temperature anomalies (wrt 1950-1980) projections for A1B scenario.

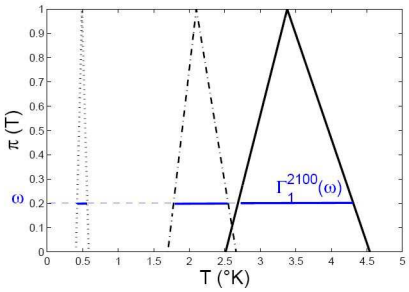


Fig. 4. Membership functions and ω-level cuts of the IPCC projected temperature anomalies (wrt 1950-1980) in 2000 (in dotted line), 2060 (dashed line) and 2100 (continuous line).

of the available evidence suggests representing the annual temperature projections T_t by a belief function verifying $Bel([\underline{T}_t, \overline{T}_t]) = 1$ and $Pl(\{T_t\}) = 1$ (see previous section). In the absence of more precise information, consonant random intervals with core T_t and support $[\underline{T}_t, \overline{T}_t]$ are a natural representation. The associated contour function π is defined by:

$$\pi(x) = \begin{cases} 1 & x = T_t \\ \phi(\frac{x-\underline{T}_t}{T_t-\underline{T}_t}) & \underline{T}_t \leq x \leq T_t \\ \phi(\frac{\overline{T}_t-x}{\overline{T}_t-T_t}) & T_t \leq x \leq \overline{T}_t \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

where ϕ is a continuous, non-decreasing function from $[0,1]$ into $[0,1]$ such that $\phi(0) = 0$ and $\phi(1) = 1$.

The form of function ϕ is not crucial. In a previous paper [5], we have addressed the issue of the choice of the contour function shape in representing different experts' opinions on the SLR. Sensitivity analyses have shown that, provided the core and the support remain the same, the shape of the contour function has a negligible impact on the response of interest. In this analysis, we consider a linear function ϕ. Figure 4 shows the contour functions π describing the available evidence on the temperature anomaly in, respectively, 2000, 2060 and 2100, as well as the ω-level cuts, $\Gamma^t(\omega)$, for a given $\omega \in [0,1]$.

2) Constructing a belief function on the model parameters: The SLR model fitting was processed with an approach similar to the one used in [36]. Annual means of sea level and temperature anomaly data (wrt to 1950-1980) over the 1881-2000 period (i.e. 120 years), were first smoothed with a 15-years smoothing filter in order to remove the low frequency variability (interannual to decadal time scale) unrelated to the climate signal. A statistical regression based on the least square estimations was then performed to infer the best fitting model parameters, a, b and T_0. A subtle difference to mention between our approach and Rahmstorf's is that we fit the model at annual temporal resolution (and not 5-years bins as in [36]). Annual residual errors $r_i (i = 1:120)$ between the observed sea level rates and the estimated ones (via RV09) were first computed. Residuals were assumed, for computational convenience, to be independent and normally distributed with a constant known variance σ^2. Under this hypothesis, the likelihood function defined on the joint parameters space of the model parameters a, b and T_0 is given by:

$$L(a, b, T_0; r_i) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{120} r_i^2\right). \quad (11)$$

Our aim is to infer the model parameters from the available data. We address this inference issue with the likelihood-based approach described in Subsection II-B. By combining (11) and (6), the joint contour function on the RV09 joint parameters space is obtained by the following equation:

$$pl(a, b, T_0) = \frac{L(a, b, T_0; r_i)}{L(\hat{a}, \hat{b}, \hat{T_0}; \hat{r_i})}, \quad (12)$$

where $\hat{a}, \hat{b}, \hat{T_0}$ are the maximum likelihood (ML) estimates of a, b and T_0, respectively.

The joint contour function has an ellipsoidal shape. Figure 5a shows the joint contour function $pl(a, b)$ obtained by marginalizing (12) over T_0. Each ellipse corresponds to a plausibility level. The marginal contour function of parameter b, $pl(b)$, is derived by an additional marginalization of $pl(a, b)$ over parameter a and is plotted in Figure 5b.

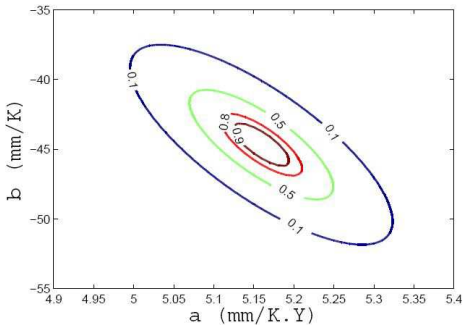
B. Propagating belief functions through the RV09 model

We have constructed two independent random intervals encoding evidence on the RV09 model parameters and the model temperature input, denoted by $(\Omega_1, P_1, \Gamma_1)$ and $(\Omega_2, P_2, \Gamma_2)$, respectively. The combined evidence on the SLR model response at the end of the century is denoted by (Ω, P, Γ) with $\Omega = \Omega_1 \times \Omega_2$. The analytical expressions of the plausibility (Pl) and belief (Bel) functions on Ω are difficult to derive, but they can be approximated using Monte Carlo simulation. The algorithm is as follow. An i.i.d. random sample $(\omega_1, \alpha_1), ...(\omega_N, \alpha_N)$ is generated by sampling a $[0,1]$ uniform distribution. Then, for $i = 1...N$:

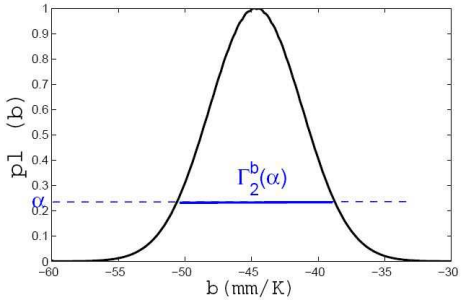
1) Derive the ω_i-level cuts on each of the consecutive annual temperature plausibility: $\Gamma_1^t(\omega_i)$, for every $t \in [2000, 2100]$ (as shown in Figure 4);
2) Derive the α_i-level-cuts on each of the RV09 parameters contour function: $\Gamma_2^a(\alpha_i), \Gamma_2^b(\alpha_i)$ and $\Gamma_2^{T_0}(\alpha_i)$;
3) Propagate the focal elements through (9) and derive, by interval calculations, the SLR response at the end of the century $[U_i, V_i]$.

The Bel and Pl functions induced by (Ω, P, Γ) can then be approximated for every subset I by:

$$\hat{Bel}(I) = \frac{1}{N}\text{card}\{1 \leq i \leq N | [U_i, V_i] \subseteq I\} \quad (13)$$



(a)

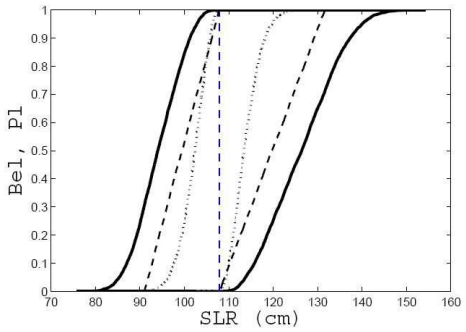


(b)

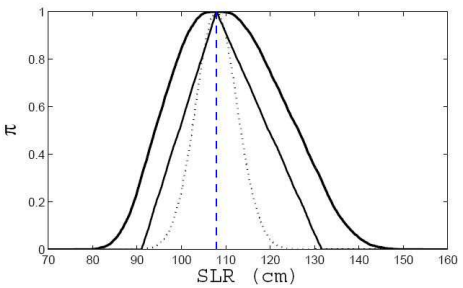
Fig. 5. (a) Joint contour plausibility $pl(a, b)$. (b) Marginal contour function $pl(b)$ and an α-level cut on b parameter

$$\hat{Pl}(I) = \frac{1}{N}\text{card}\left\{1 \leq i \leq N | [U_i, V_i] \cap I \neq \emptyset\right\}. \qquad (14)$$

The cumulative belief and plausibility of the SLR at the end of the century are calculated for intervals $I =]-\infty, x]$. The upper (plausibility) and lower (belief) cumulative distribution functions (cdfs) can be seen as upper and lower envelopes of the set of all cdfs compatible with the available information on the model parameters and inputs. Figure 6a shows the results of three cases: dotted curves represent the case where only the uncertainty in the model parameters is propagated while the model input is deterministic and corresponds to the mean temperature estimates (the bold curve in Figure 3). The upper curve is for plausibility, whereas the lower one is for belief. Dashed lines represent the results when the projections are calculated with the best estimates values of RV09 parameters and random intervals on the temperatures. The bold lines correspond to the results when both uncertainties are combined as explained above. The discontinuous vertical line indicates the deterministic (mean temperature projections, best estimates parameters) results. We can see from these plots that the uncertainty on SLR accounts for most of the uncertainty on the projected temperatures. However, statistical uncertainty related to model fitting is also important. An alternative description of the derived evidence on the SLR at the end of the century, by the means of the contour plausibility, is plotted in Figure 6b. The RV09 deterministic estimation of the future SLR in 2100 is 110 cm. However, by allowing the model inputs and the parameters to take less plausible (but possible) values than the best estimates, the model response can reach much higher values, as shown in Figure 6b.



(a)



(b)

Fig. 6. (a) Cumulative Plausibility and Belief, (b) contour plausibility of the SLR at the end of the century (wrt 1990).

IV. Overtopping hazard under climate change

A. Methodology for overtopping discharge estimation

Commonly, structures are designed to withstand centennial overtopping, i.e., overtopping that occurs on average once every 100 years. This level needs to be determined by extrapolation, as historical measures of the hazards, when available, rarely exceed some decades. In reality, measures exist more often for the sea drivers rather than for the hazard itself. Different methods are available for the prediction of overtopping under given sea conditions. A commonly used one is based on empirical methods that relate the mean overtopping discharge q (the total volume of overtopped water per unit of length in a certain duration ($m^3 s^{-1} m^{-1}$)) to the main sea conditions and structure parameters. The EurOtop Manual [14], a reference guideline in coastal engineering, proposes the following expression:

$$q = 0.21\sqrt{gH_{Toe}{}^3} \times \exp\left(-\frac{R_c}{H_{Toe}\gamma_f\gamma_b(0.33 + 0.022\xi_0)}\right), \qquad (15)$$

where H_{Toe} is the significant wave height at the toe of the structure, obtained by applying propagation models to the offshore wave height (H); ξ_0 is the breaker parameter, which characterizes the breaking wave behavior; g is the gravitational constant; γ_b and γ_f are characteristics the structure and R_c is the crest freeboard (see Figure 1), whose height is defined relatively to the sea water level (SWL). The SWL is composed of the mean sea water level, tide and surge (S) components.

Tides are the fall and rise of sea levels induced by the combined effects of the gravitational forces exerted by the

moon, the sun and the rotation of the Earth. Tidal behavior is said to be semi-diurnal when there are two almost equal high tides (HT) per day, or equivalently, 706 high tides yearly. Surges are induced by the rise of sea level due to high wind and barometric pressure changes during storms; they occur often with high waves. Positive surges and waves when combined with high tides cause extreme sea water levels that may lead to important overtopping (15). Contrary to the astronomical component, which is entirely predictable over a nodal cycle of 18.6 years, the weather components are aleatory. Typically, there is only a modest correlation between waves and total water levels, because the predominant astronomical tidal component of water level is not related to the local weather conditions. On the contrary, dependence between surges and waves is expected, since both are related to the same weather conditions [20]. Dependence is more likely at extreme levels. It is challenging to account for it when estimating the overtopping events.

The procedure to estimate the variable of interest, denoted by q^{100}, is the following: (1) from the available bivariate observations of storm surges and waves, characterize the dependence structure (for this we resort to copulas as a multivariate statistical model), (2) generate in the joint distribution a richer set of conjunctions with more extreme events than those in the original sample, (3) transform the wave-surge conjunctions into wave-SWL conjunctions by shifting the surge level by the tide and the mean water levels, (4) evaluate the associated overtopping for each of the simulated wave-SWL pairs and finally (5) estimate q_{100} as the $(1 - \frac{1}{100})$ quantile of the overtopping distribution. To account for the effect of the climate change on the design criteria, the mean water level needs simply to be shifted positively by the SLR in step (3).

B. Application and results

The design overtopping under climate change was estimated considering the uncertainty on the main drivers: the future SLR (subjective uncertainty) and the storm components (aleatory uncertainty). Both uncertainties were modeled within DS theory. The methodology was applied to a typical dike in Le Havre, France. This site is characterized by a semi-diurnal tide behavior: hence, overtopping is likely to occur when high tides and storm waves and surges occur simultaneously. The dike has the following properties: $CR = 14m$, $\sigma_b = \sigma_f = 1$.

First, storm events were selected from the available datasets. They correspond to wave-surge conjunctions such that both variables exceed a storm threshold. The Havre hydrologic dataset covers 21 years of measurements. Storm waves (H) and surges (S) were sampled from the dataset as the values exceeding, respectively, $2.25\ m$ and $0.27\ m$ storm thresholds. A total of $292(n)$ conjunctions ((H_i, S_i), $i = 1...n$) were picked. Given the sampling method (Peaks Over Threshold), the Extreme Values theory (EVT) [19] suggests the use of the Generalized Pareto Distribution (GPD) for modeling both marginals. Recall that the GP cumulative distribution function of a random variable X exceeding a threshold μ is given by:

$$F_X(x, \beta) = 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}, \qquad (16)$$

where $\beta = (\xi,\ \sigma)$ is the vector of shape and scale parameters.We selected the Gumbel Copula to model the joint probability of storm wave surge conjunctions, as it is directly related to multivariate extensions of extreme value theory. It has the following expression $\forall (u, v) \in [0, 1]^2$:

$$C(u, v) = \exp[-(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}}], \qquad (17)$$

where $\theta > 1$ is the copula parameter that captures the intrinsic dependence between the marginal variables.

To infer the distribution parameters, we used the Inference Function for Margins (IFM, [22]), a parametric inference technique. The IFM method is a two step method that relies on the split of the parameters into specific parameters for marginals and common (or association) parameters for the copula. It was shown that the IFM performs well when marginal distributions fit well the data [41]. Statistical goodness of fit tests have shown that there is no mis-specification of the marginal distributions (see Figure 7a for the wave parameter). The use of the IFM as a estimation technique is thus justified. First, the ML estimates were computed for each margin; then, the association parameter was estimated given the marginal best estimates. The likelihood function in the second step of the IFM is given by:

$$L(\theta, \hat{\beta}_1, \hat{\beta}_2, H_i, S_i) = \prod_{i=1}^{n} f_{H,S}(H_i, S_i), \qquad (18)$$

with:

$$f_{H,S}(H_i, S_i) = c(F_H(H_i, \hat{\beta}_1), F_S(S_i, \hat{\beta}_2)) \cdot f_H(H_i, \hat{\beta}_1) \cdot f_S(S_i, \hat{\beta}_2), \qquad (19)$$

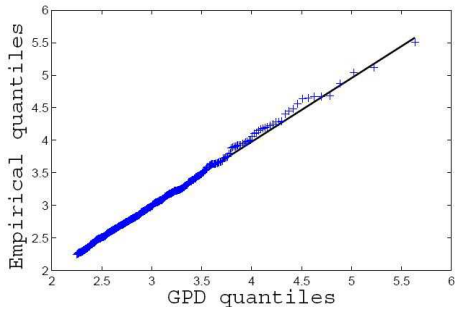
where c is the Gumbel copula density, $\hat{\beta}_1$ and $\hat{\beta}_2$ are the ML estimates of, respectively, waves and surges GP marginals parameters.

To model statistical uncertainty on the waves and surges conjunctions, we applied the likelihood-based inference method described above to estimate the copula parameter. By combining (18) and (6), we derived a contour function on θ (Figure 7). We thus constructed a random interval $(\Omega_3, P_3, \Gamma_3)$ encoding evidence on the copula parameter (or equivalently, on the joint behavior of waves and surges). Each of the uncertain inputs of interest was constrained by a random set: $(\Omega_3, P_3, \Gamma_3)$ for the aleatory hydrological variables and (Ω, P, Γ) for the imprecise SLR (in previous section).

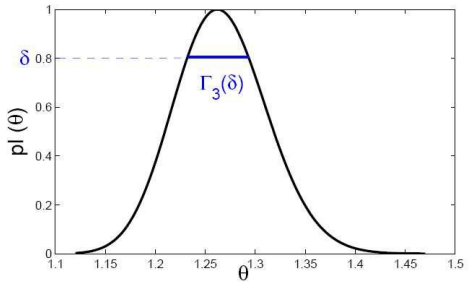
Both random sets were combined and propagated through (15) using the Monte Carlo technique described in Subsection III-B. Belief and plausibility measures on q^{100} were then approximated by (13)-(14).

Two cases were considered: first, we accounted only for the statistical uncertainty of the overtopping input variables, and we forced the model with a constant SLR of 1m. The cumulative plausibility and belief of the derived centennial overtopping are plotted in Figure 8 (dotted lines). In bold, we represent the same measures in the case where the SLR is uncertain and described by the DS structure constructed in the previous section. Clearly, most of the imprecision on the design variable estimate is due to the SLR component rather than the statistical uncertainty of waves-surges conjunctions.

Let us consider the critical value: $q^* = 0.1m^3s^{-1}m^{-1}$ and analyze the effect of the representation of the SLR component



(a)



(b)

Fig. 7. (a) QQ-plot for wave parameter, (b) contour marginal plausibility of the Copula parameter θ and the δ-level cut.

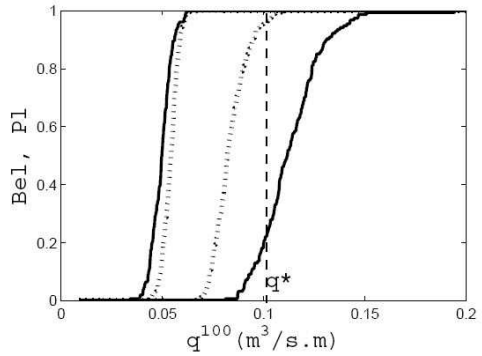


Fig. 8. Cumulative Plausibility and Belief of the centennial overtopping discharge under climate change for two SLR scenarios: constant SLR=1m (in dotted), RV09 projections (in continuous).

on the probability of exceeding this value. We can see from Figure 8 that a deterministic approach integrating the SLR can highly underestimate the likelihood of exceeding a critical extreme threshold: if the 1m scenario is used, the probability of non exceeding of q^* is comprised between 0.99 and 1, or equivalently, the probability of exceeding this threshold is below 0.01. Now, if the uncertainty in the SLR projections is accounted for and is described by a DS structure, the probability of non exceeding the critical threshold ranges between 0.22 and 1, i.e., the probability of exceeding this threshold can be as high as 0.78.

V. Conclusions

The hydrologic parameters controlling the future response of a coastal defense structure to climate change are uncertain and are affected by both randomness (with possible dependence structure) and imprecision. In this paper, we illustrated a promising methodology based on Dempster-Shafer theory to quantify and propagate both uncertainties through an overtopping model. The described tools and the developed method can easily be extended to consider other sources of uncertainty such as the observed data or the models, or to update the results when additional evidence is available. The latter extension, in particular, is possible for the SLR module. A shortcoming of our approach is that the SLR projections were derived based on a unique semi-empirical model. The predictive capability of such models is viewed with some reluctance within the climate science community. Projections would be more consistent if they were combined with experts' opinions on the validity of the models or on their estimations of future rise of the oceans. Recently, a formalized pooling of expert views on uncertainties in ice sheet contribution, the main uncertain component in the SLR, has been reported by Bamber [3]. The elicitation of experts judgements and the combination of the new evidence with the semi-empirical model based results would allow higher confidence in the SLR projections.

Acknowledgment

This work was carried out in the framework of the Labex MS2T, which was funded by the French Government, through the program "Investments for the future" managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02).

References

[1] M. Aickin, "Connecting Dempster-Shafer belief functions with likelihood-based inference," *Synthese*, vol. 12, pp. 347-364, 2000.

[2] H. Apel, A. H. Thikeken, B. Merz and G. Blosch, "Flood risk assessment and associated uncertainty," *Natural Hazards and Earth System Sciences*, vol. 4, pp. 295-308, 2004.

[3] J. L. Bamber and W.P . Aspinall, "An expert judgement assessment of future sea level rise from ice sheets," *Nature Climate Change* DOI: 10.1038/nclimate1778, 2013.

[4] G. A. Barnard, G. M. Jenkins and C. B. Winsten, "Likelihood inference and time series," *Journal of the Royal Statistical Society,* vol. 125, pp. 321-372, 1962.

[5] N. Ben Abdallah, N. Mouhous-Voyneau and T. Denoeux, "Combining statistical and expert evidence using belief functions: Application to centennial sea level estimation taking into account climate change," *International Journal of Approximate Reasoning*, in press, 2013. doi: http://dx.doi.org/10.1016/j.ijar.2013.03.008.

[6] A. Birnbaum, "On the foundations of statistical inference", *Journal of the American Statistical Association*, vol. 57, pp. 269-306, 1962.

[7] R. J. Dawson, J. W. Hall, P. D. Bates and R. J. Nichols, "Quantified analysis of the probability of flooding in the Thames Estuary under imaginable worst case sea-level rise scenarios," *International Journal of Water Resources Development*, vol. 21, pp. 577591, 2005.

[8] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals of Mathematical Statistics*, vol. 38, pp. 325-339, 1967.

[9] A. P. Dempster, "Upper and lower probabilities generated by random closed interval," *Annals of Mathematical Statistics*, vol. 39, pp. 957-966, 1968.

[10] A. P. Dempster, "The Dempster-Shafer calculus for statisticians," *International Journal of Approximate Reasoning*, vol. 48, pp. 365-377, 2008.

[11] T. Denoeux. Likelihood-based belief function: justification and some extensions to low-quality data. International Journal of Approximate Reasoning (accepted for publication), 2013. Available at http://hal.archives-ouvertes.fr/hal-00813021.

[12] D. Dubois, "Possibility theory and statistical reasoning," *Computational Statistics and Data Analysis*, vol. 57, pp. 47-69, 2006.

[13] A. W. F. Edwards, *Likelihood*, expanded edition, Baltimore, USA: The John Hopkins University Press, 1992.

[14] EurOtop. *Wave Overtopping of Sea Defences and Related Structures: Assessment Manual*, 2007. http://www.overtopping-manual.com/eurotop.pdf.

[15] S. Ferson, V. Kreinovitch, L. Ginzburg, D. S. Meyers and K. Sentz, "Constructing probability boxes and Dempster-Shafer structures", Albuquerque, USA: Sandia National Laboratories, SAND2002-4015, 2003.

[16] Ferson S, Nelsen R, Hajagos J, Berleant D, Zhang J, Tucker W, Ginzburg and L, Oberkampf W, " Dependence in probabilistic modelling, Dempster-Shafer theory and probability bounds analysis", Technical report, Sandia National Laboratories, 2004.

[17] R. A. Fisher, *Statistical Methods and Scientific Inference*, Edinburgh: Oliver and Boyd, 1956.

[18] Y. Goda, "Irregular wave deformation in the surf zone Coastal Engineering in Japan," *Coastal Engineering in Japan, JSCE*, vol. 18, pp. 13-26, 1975.

[19] E. J. Gumbel, *The statistics of extremes*, New York: Colombia University Press, 1958.

[20] P. J. Hawkes, B. P. Gouldby, T. A. Tawn and M. W. Owen, "The joint probability of waves and water levels in coastal engineering design," *Journal of Hydraulic Research*, vol. 40, pp. 241-251, 2002.

[21] "Intergovernmental Panel on Climate Change (IPCC) climate change 2007: The Physical Scientific Basis: Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge, UK: Cambridge University Press, 2007.

[22] H. Joe and J. J. Xu, "The Estimation Method of Inference Functions for Margins for Multivariate Models," University of British Columbia, Department of Statistics, Technical Report no. 166, 1996.

[23] E. Kriegler and H. Held, "Utilizing belief function for the estimation of future climate change," *International Journal of Approximate Reasoning*, vol. 39, pp. 185-209, 2005.

[24] G. Kim and P. Silvapulle, "Comparison of semiparametric and parametric methods for estimating copulas," *Computational Statistics and Data Analysis*, vol. 51, pp. 28362850, 2007.

[25] B. Merz and A. H. Thieken, "Flood risk curves and uncertainty bounds," *Natural hazards*, vol. 51, pp. 437-458, 2009.

[26] W. T. Pfeffer, J. T. Harper and S. O'Neel, "Kinematic constraints on glacier contribution to 21st century sea level rise," *Science*, vol. 321, pp. 1340-1343, 2008.

[27] M. Purvis, P. D. Bates and C. M. Hayes, "Probabilistic methodology to estimate future coastal flood risk due to sea level rise," *Coastal Engineering*, vol. 55, pp. 1062-1073, 2008.

[28] S. Rhamstorf, "A semi empirical approach to projecting future sea level rise," *Science*, vol. 315, pp. 368-37, 2007.

[29] S. Rahmstorf, G. Foster and A. Cazenave, "Comparing climate projections to observations up to 2011," *Environmental Research Letters*, vol. 7, 2012.

[30] T. A. Severini, *Likelihood Methods in Statistics*, Oxford, UK: Oxford University Press, 2000.

[31] G. Shafer, *A mathematical Theory of Evidence*, Princeton, New Jersey, USA: Princeton University Press, 1976.

[32] G. Shafer, "Belief Functions and Parametric Model," *Journal of the Royal Statistical Society, Series B*, vol 44, pp. 322-352, 1982.

[33] C. Shouyu and G. Yu, "Variable Fuzzy Sets and its Application in Comprehensive Risk Evaluation for Flood-control Engineering System," *Fuzzy Optimization and Decision Making*, vol. 5, pp. 153-162, 2006.

[34] A. Sklar, "Fonction de répartition à n dimensions et leurs marges," *Publications de Institut de Statistique Université de Paris,* vol. 8, pp. 229-231, 1959.

[35] D. A. Sprott and J. D. Kalbfleisch, "Examples of Likelihoods and Comparison with Point Estimates and Large Sample Approximations," *Journal of the American Statistical Association*, vol. 64, pp. 468-484,1969.

[36] M. Vermeer, S. Rhamstorf, "Global sea level linked to global temperature," in *Proc Natl Acad Sci, USA, 2009*, p. 21527.

[37] Y. P. Xu YP, M. J. Booji and Y. B. Tong, " Uncertainty analysis in statistical modeling of extreme hydrological events," *Stochastic environmental research and risk assessment,* vol. 24, pp. 567-578, 2010.

[38] P. Walley, *Statistical Reasoning with Imprecise Probabilities,* London, UK: Chapman and Hall, 1991.

[39] L. Wasserman, "Belief functions and statistical inference," *The Canadian Journal of Statistics,* vol. 18, pp. 183-196, 1990.

[40] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, pp. 3-28, 1978.

[41] G. Kim and P. Silvapulle, "Comparison of semiparametric and parametric methods for estimating copulas," *Computational Statistics and Data Analysis,* vol. 51, pp. 28362850, 2007.

[42] J. A. Church and N. J. White, "A 20th century acceleration in global sea-level rise," *Geophys Research Letter,* 33(1):L01602. doi:10.1029/2005GL024826, 2006.