2018-13

Optimal Credible Warnings

Koffi Akakpo
Marie-Amélie Boucher
Vincent Boucher

Juillet / July 2018

Centre de recherche sur les risques
les enjeux économiques et les politiques publiques

www.crrep.ca

Abstract

We consider a decision maker who is responsible for issuing flood warnings for the population. The population is uncertain about the credibility of the warnings and adjusts its beliefs following false alerts or missed events. We show that low credibility leads the decision maker to issue warnings for lower probabilities of flooding. In practice, those probabilities are provided by hydrological forecasts. We therefore use our model to compare welfare under alternative real-world hydrological forecasts. We find that when forecasts include non-realistic extreme scenarios, the economy may remain stuck in a state characterized by many false alerts and poor credibility.

JEL Codes: D81, Q28, C61
Key words: Flood warnings, Renewable resource management, Uncertainty

Koffi Akakpo: Department of Economics, Université Laval, CREATE;
koffi.akakpo.2@ulaval.ca

Marie-Amélie Boucher: Department of Civil and Building Engineering, Université de Sherbrooke;
marie-Amelie.Boucher@usherbrooke.ca

Vincent Boucher: Department of Economics, Université Laval, CRREP and CREATE;
vincent.boucher@ecn.ulaval.ca

We would like to thank Richard Turcotte at the Direction de l'Expertise Hydrique as well as the team responsible for maintaining the TIGGE data portal at the European Center for Medium Range Weather Forecasts in Reading. We would also like to thank Philippe Barla, Markus Herrmann, Patrick Gonzalez, Michel Roland and Rémi Morin Chassé, as well as the many participants to the 2018 Canadian Economic Association Annual Conference, the CREATE-GREDI 2018 Workshop and the Seminaire du jeudi midi for their valuable comments and suggestions. Vincent Boucher gratefully acknowledges financial support from FRQ-SC and SSHRC.

1 Introduction

In most countries, governmental authorities are responsible for issuing warnings in the event of natural disasters (e.g. flood, fire, tornadoes). However, these warnings are only as effective as their credibility to the population; too many false alerts or missed events are likely to affect this credibility.

In this paper, we study the optimal flood warning rule under endogenous credibility. We assume that the population's confidence toward a warning evolves as a function of past warnings and flood events. We show that this leads the decision maker to alert the population for *lower* flood risks. We use a corresponding decision model to rank different types of hydrological forecasts used by the decision maker.

We model the optimal decision of issuing a flood warning, given the probability of a flood (provided by the hydrological forecast) and the population's confidence in the warning. When the decision maker neglects the potential impact of the warning on the population's confidence, the decision rule is given by a simple threshold. However, when the decision maker accounts fully for the impact of the warning on the population's confidence, we find that the optimal rule is decreasing with the population's confidence level, leading the decision maker to alert the population more often when confidence is low.

A dynamic stochastic programming method is used to solve the model. This model has two state variables: the forecasted flood probability and the population's confidence. We formally define the population's confidence as the probability that the population acts if a warning is issued. We assume that the decision maker is risk averse, with preferences represented by a Constant Absolute Risk Aversion (CARA) utility function. We find that the population's confidence has a non-linear impact on the optimal warning rule. Small changes in confidence have a greater impact at low confidence levels.

To discriminate between the different types of hydrological forecasts, we use the optimal decision rule. While the optimal decision rule is a function of the probability of a flood, this probability is usually unknown. The decision maker must therefore use (imperfect) hydrological forecasts.

We use the decision model to compare the performance of different hydrological forecasts. Using a real-world example from the province of Quebec, Canada, we compare a forecasting system that is similar to the operational one with two alternative meteorological ensemble forecasts.[1] We find large welfare differences across forecasts.

This work contributes to the literature focused on regulations when the population has subjective beliefs (e.g. Pollak, 1998). As discussed in Salanié and Treich (2009), the literature usually contrasts populist and paternalistic views of the population's welfare. In our context, we adopt a paternalistic view of the population's welfare in the sense that the decision maker uses the best possible estimate of flood probability. However, much like in a principal-agent problem, the decision maker cannot directly enforce the population's action. The decision maker must therefore use the warning to influence the population's action (itself based on the population's subjective beliefs).

This work also adds to the literature studying the management of uncertain renewable resources (e.g. Pindyck, 1984, 2007; Singh et al., 2006; Weitzman, 2002). We present a binary decision model for a risk-averse decision maker where the growth rate of the resource is random. In our context, the population's confidence level is modeled as the stock of a renewable resource. The effect of the decision maker on the level of the resource is also random: a warning increases the level of confidence if there is a flood. We find that low levels of the resource lead to a greater number of warnings.

Finally, we contribute to the literature dealing with the economic value of hydrological forecasts (e.g. Katz and Murphy, 1997; Murphy, 1977; Roulin, 2007; Verkade and Werner, 2011; Matte et al., 2017). By exploring the welfare effects of alternative hydrological forecasts, we find that the forecasts could be improved by reevaluating the credibility of some forecasted extreme scenarios. Indeed, when the population's confidence is low, a few extreme scenarios will render a warning optimal. If these scenarios are not credible, the decision maker will issue many false alerts, and the economy may remain in a state where both confidence and welfare are low.

The paper is organized as follows. In Section 2, we present the decision maker's problem. In Section 3, we present the data and the hydrological models used for the application, while in Section 4, we present the results and welfare implications. Section 5 presents a robustness analysis, and Section 6 presents the conclusions and future directions.

[1] Details are provided in Section 3.2.

2 Decision Model

In this section, we present the decision model from the point of view of the authorities. We consider a (benevolent) decision maker who must decide whether or not to alert the population. Their decision will affect the population's current action as well as the credibility of future alerts. The formal model is described below.

At any point in time, the population is at risk of being flooded. The population has the possibility of taking a costly action $a \in \{0, 1\}$ to mitigate the effects of the flood (e.g. sandbagging, emptying basements, installing pumps, temporarily relocating, etc.). The instantaneous payoff of the population is given by:

$$\pi(a, y) = bay - ca - dy$$

where $y \in \{0, 1\}$ equals 1 if there is a flood, $d > 0$ is the damage from the flood, $c > 0$ is the cost of the action, and $b > 0$ is the alleviated damage if an action is taken. We assume that $d > b > c$ so that $\pi(0, 0) > \pi(1, 0) > \pi(1, 1) > \pi(0, 1)$.

Crucially, we assume that the population does not know the probability of a flood and must rely on the information provided by the authorities. This assumption is motivated by both conceptual and practical concerns. Indeed, the evaluation of the probability of a flood is obtained from sophisticated hydrological models (see Section 3.2), and it is unrealistic to assume that the population has the necessary knowledge to interpret the models' outputs.

Furthermore, even if the authorities were to communicate the implied probability directly to the population, as the forecasts are made daily, the population would have to constantly follow these flood probabilities. We therefore assume that the population forms subjective beliefs concerning the probability of a flood. In principle, there is no reason for these beliefs to represent the true probability of a flood. For example, in a similar context, Bakkensen and Barrage (2017) recently found a high variability in terms of flood risk beliefs for individuals of a same population (all facing the same risk).[2]

Since we are mostly interested in the decision maker's problem, we model the population's beliefs in a very stylized way. In particular, we assume that if the decision maker does not alert the population, the population takes no action.[3] However, if the decision maker alerts the population, then the population *may* take a costly action.

[2] See also Salanié and Treich (2009) and the references therein.

[3] That is, the belief that a flood will occur is not high enough to justify the cost of an action.

We formally model the population's beliefs from the point of view of the decision maker. We assume that the population's confidence toward the alert is resumed by $p \in [0,1]$, the probability that the population takes an action if an alert is in place.[4]

Thus, from the point of view of the decision maker, there are two sources of uncertainty: whether or not there will be a flood and the reaction of the population. We assume that, from the decision maker's point of view, these are the only sources of uncertainty. We also assume that the preferences of the decision maker (as well as those of the population) toward risk are given by a CARA function.

$$\mu(a,y) = \frac{-1}{A}\exp\{-A\pi(a,y)\}$$

The choice of the CARA function is attractive as the level of initial wealth can be normalized. This is convenient since it is typically unknown, and it is particularly attractive in a dynamic setting (see below). There is also a normative advantage of assuming that the optimal decision is independent of the population's wealth. Indeed, we want to avoid situations where a decision maker would alert a rich population but not a poor one (or vice versa). Matte et al. (2017) made the same assumption in a similar context.

The decision process is as follows. At any period of time t, the decision maker has the possibility of alerting the population ($\sigma \in \{0,1\}$) in regard to a possible forthcoming flood. This decision is based on the current probability of a flood $x \in [0,1]$ (implied by the received hydrological forecasts) and the probability $p \in [0,1]$ that the population takes the alert seriously (and takes costly action). This probability will be allowed to evolve over time as a function of the previous performance of the alerts. The instantaneous utility for the decision maker is therefore given by the following vNM utility function: $U(x,p,\sigma)$, where:

$$\begin{aligned} U(x,p,1) &= x[p\mu(1,1) + (1-p)\mu(0,1)] + (1-x)[p\mu(1,0) + (1-p)\mu(0,0)] \\ U(x,p,0) &= x\mu(0,1) + (1-x)\mu(0,0) \end{aligned}$$

In a static setting, the decision maker disregards the potential impact of the alert on the population's beliefs p. They therefore choose $\sigma(x,p) = 1$ iff $U(x,p,1) \geq U(x,p,0)$, or equivalently if:

[4] Since we are primarily interested in the decision maker's choice of action, it is not necessary to model explicitly the belief function of the population. From the point of view of the decision maker, it is sufficient to know the probability that the population takes an action, conditional on the alert. Alternatively, p can be thought as the fraction of the population that chooses to act.

$$x \geq \frac{\mu(0,0) - \mu(1,0)}{[\mu(0,0) - \mu(1,0)] + [\mu(1,1) - \mu(0,1)]} \equiv \frac{Type\ I}{Type\ I + Type\ II} \tag{1}$$

where *Type I* is the utility loss of a Type I error (taking an action when there is no flood) and *Type II* is the utility loss of a Type II error (missed flood). Since, Type I errors are typically less costly than Type II errors, we expect this threshold to be smaller than 1/2.

Now, as discussed above, the population does not know the probability x of a flood and only considers the decision maker's recommendation. The population therefore does not know why/if an alert was justified. In this context, false alerts or missed events are likely to affect the population's confidence p.

We therefore study the decision maker's optimal decision in a dynamic setting, where the population's confidence evolves through time. Under standard assumptions (Lucas and Stokey, 1989), the inter-temporal utility of the decision maker can be written recursively:

$$V(x,p) = \max_{\sigma \in \{0,1\}} \{U(x,p,\sigma) + \delta \mathbb{E}_y V(x', p')\}$$

where $x' = g(x,y)$ and $p' = h(p,y,\sigma)$ are the future values for the state variables and $\delta \in (0,1)$ is the discount rate. Importantly, the use of CARA utility functions implies that the population's wealth w is not a relevant state variable. This considerably simplifies the analysis, simulations, and real-world calibrations.

Note that the expectation over the value function is taken with respect to y, which is the only source of randomness for determining future values for the state variables. This is not without loss of generality. In principle, even observing x and y, the decision maker may still be uncertain about the future probability of a flood x'. In the context of this paper, however, we believe this assumption to be reasonable as the value of x and x', which are predicted daily by the hydrological model, are highly correlated.[5] We therefore set $g(x,1) = 1$ (a flood lasts at least one day) and $g(x,0) = x$.

The function $h(p,y,\sigma)$ describes how the population's confidence is affected by missed events $(y = 1, \sigma = 0)$ and by false alarms $(y = 0, \sigma = 1)$. As discussed above, we assume that the population forms subjective beliefs regarding flood risks. Coherently, assuming perfect Bayesian

[5] Specifically, the probability that tomorrow's forecast is exactly in the same ranges from 80% to 90% depending on the selected hydrological forecasting system (see Section 3.2). Moreover, the average difference between two consecutive daily forecasts is of an order smaller than 2×10^{-10} for all hydrological forecasting systems.

updating seems unrealistic. We therefore assume a linear updating as a first-order approximation:

$$h(p, y, \sigma) = [1 - \phi_3 y - \phi_4 \sigma + (\phi_3 + \phi_4) y\sigma]p + [\phi_1(1 - y - \sigma) + (\phi_1 + \phi_2) y\sigma](1 - p)$$

where $\phi_1, \phi_2, \phi_3, \phi_4 \in (0, 1)$.

On the one hand, the population's confidence in a forecast increases when $y = \sigma$; we have $h(p, 0, 0) = p + (1 - p)\phi_1$ and $h(p, 1, 1) = p + (1 - p)\phi_2$. Note that this assumes that the rate at which confidence grows is decreasing in p.[6] On the other hand, the population's confidence decreases when $y \neq \sigma$; we have $h(p, 1, 0) = (1 - \phi_3)p$ (missed event) and $h(p, 0, 1) = (1 - \phi_4)p$ (false alarm). The rate at which confidence is lost is decreasing in p.[7]

The optimal decision rule $\sigma(x, p)$ therefore not only accounts for the current costs and benefits of alerting the population but also on the long-term effect on the population's confidence in regard to flood forecasts. This is analogous to the economic theory on the extraction of renewable resources (e.g. Gordon, 1954). Indeed, the population's confidence p has conditional expected growth rates given by:

$$\dot{p}_{\sigma=0} = -px\phi_3 + (1 - p)(1 - x)\phi_1$$
$$\dot{p}_{\sigma=1} = x(1 - p)\phi_2 - (1 - x)p\phi_4$$

Then, without any intervention (i.e. $\sigma = 0$), the expected steady state value is given by $\mathbb{E}_x p = \frac{(1-x)\phi_1}{x\phi_3+(1-x)\phi_1}$. We clearly see the effect of ϕ_1 and ϕ_3. For $\phi_3 \gtreqless \phi_1$, we have $p \gtreqless (1 - x)$; thus, the difference between ϕ_1 and ϕ_3 represents the population's subjective biases. When $\phi_1 = \phi_3$, the population's confidence is given simply by the probability of the absence of flood: $p = 1 - x$.

There is a similar intuition for $\sigma = 1$. The expected steady state value is $\mathbb{E}_x p = \frac{x\phi_2}{x\phi_2+(1-x)\phi_4}$; thus for $\phi_2 \gtreqless \phi_4$, we have $p \gtreqless x$. Finally, one can easily verify that the expected confidence at the steady state is higher for $\sigma = 1$ if:

$$\frac{1 - x}{x} \leq \sqrt{\frac{\phi_3 \phi_2}{\phi_4 \phi_1}} \tag{2}$$

which reduces to $x > 1/2$ when $\phi_1 = \phi_3$ and $\phi_2 = \phi_4$.

The fact that the evolution of p is stochastic, even for fixed decision rules, adds to the model's complexity. Accordingly, as is often the case with reasonably detailed dynamic optimization problems, few results hold for all parameters' values, and most of the analysis must be done numerically.

[6] $h(p, 1, 1) - p$ and $h(p, 0, 0) - p$ are positive and decreasing in p.

[7] $h(p, 1, 0) - p$ and $h(p, 0, 1) - p$ are negative and decreasing in p.

2.1 Numerical Resolution

In this section, we solve the model numerically for a credible set of parameters. Specifically, we solve the decision maker's problem using value function iteration (e.g. Rust, 1996). Table 1 presents the value of the calibrated parameters used to produce Figures 1 to 3. Values for b and d are calibrated based on a previous cost analysis of the Montmorency River watershed (Leclerc et al., 2001, in French), where we applied the hydrological forecasting systems described in Section 3. Numbers are in millions of 2001 Canadian dollars. A robustness analysis is presented in Section 5.

Table 1: Calibrated Parameters

Parameters	**Values**	**Parameters**	**Values**
δ	0.90	ϕ_1	0.01
A	0.01	ϕ_2	0.1
c	0.364	ϕ_3	0.2
b	1.285	ϕ_4	0.1
d	15.460	-	-

Some generic properties emerge from our analysis.[8] We first present these properties and then discuss their intuition and robustness. In Section 7.1 of the Appendix, we also provide additional technical details as to why, for all the parameters' values, these results do not hold analytically.

Numerical Result 1 *For reasonable values of the parameters, we have the following:*

1. *For any p, $\sigma(x,p) = 1$, whenever inequality (1) holds. See Figure 1.*

2. *The set of (x,p), such that $\sigma(x,p) = 0$, is convex. See Figure 1.*

3. *For any x, $V(x,p)$ is increasing and convex in p. See Figure 2.*

4. *For any p, $V(x,p)$ is decreasing and convex in x. See Figure 3.*

Numerical result 1 implies that a decision maker who is aware of their impact on the population's confidence would alert the population *more often* than a decision maker who is unaware. Indeed,

[8] Those results are also supported by additional exploration of the parameter space. Those simulations are not reported here but are available from the authors. The source codes are also available online at `https://www.vincentbouchereconomist.com/codes-and-links`.

as missed floods are likely to hurt the population's confidence more than false alerts, this pushes the decision maker to alert the population for lower probabilities of flooding. This is especially true for low values of p (Numerical result 2). Indeed, when the population's confidence is low, there is little short-term benefit or cost to alerting the population, as the population is likely to ignore the alert.

The choice of issuing an alert or not is therefore governed by its impact on the population's confidence (see Equation (2)). Since false alerts are less costly than missed events (in terms of current utility as well as for the long-term impact on confidence), and that predicted floods increase the population's confidence, this pushes the decision maker to alert the population to lower probability thresholds.[9]

Numerical result 3 implies that a high level of confidence is good for welfare. Indeed, the only way a high level of confidence may hurt welfare would be if the impact of missed floods or false alerts was extreme; this only occurs for unrealistic values of ϕ_3 and ϕ_4.

Numerical result 4 is intuitive since increasing the flood probability has a first-order negative impact on welfare. Indeed, even if the population is alerted, not all damage is avoided. A situation where $V(x,p)$ is increasing in x can only occur at very large values of x *and* at very small values of p. In such a case, even if the optimal decision is to alert the population, the impact on the current utility is limited as the population will likely disregard the alert. However, the alert may have a large impact on the population's confidence. Then, when x increases, it decreases the probability of a false alarm and increases the probability that the predicted flood will materialize. This may have an overall positive impact on welfare.

The next section describes the real-life context and data that we use to experiment with the decision model that was just described.

3 Data

3.1 Montmorency Watershed

The Montmorency River, Quebec, Canada, is a North to South-flowing river ending at a spectacular 84-m-high waterfall that is one of the major tourist attractions of the area. Its watershed covers

[9]Note that this result holds even if we do not assume that missed events hurt the population's confidence more than false alerts. Again, see Equation (2).

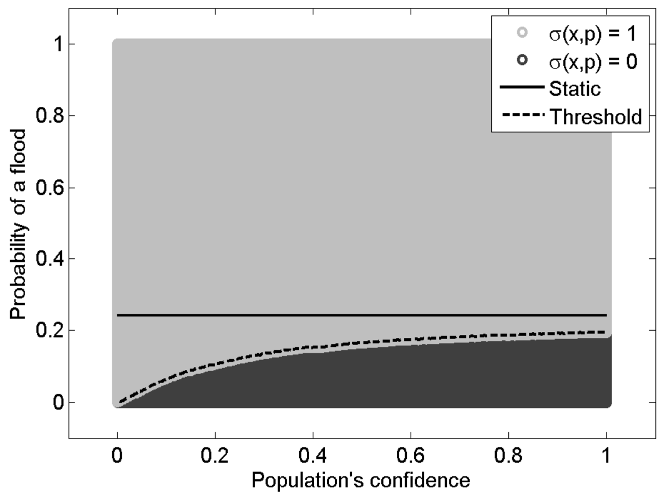

Figure 1: Decision rule.

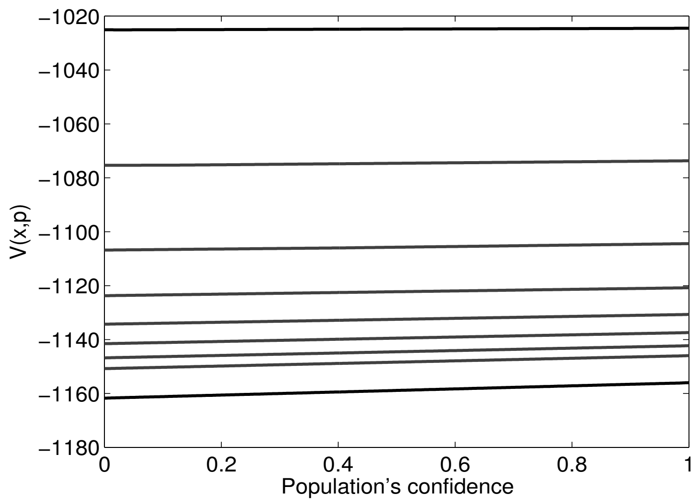

Figure 2: Value function $V(x,p)$, as a function of the population's confidence p, for fixed values of x.

1150 km^2. The upper part of the watershed is densely forested, but the southern portion is home to approximately 30 000 people.[10]

[10] See Matte et al. (2017) for more information and maps.

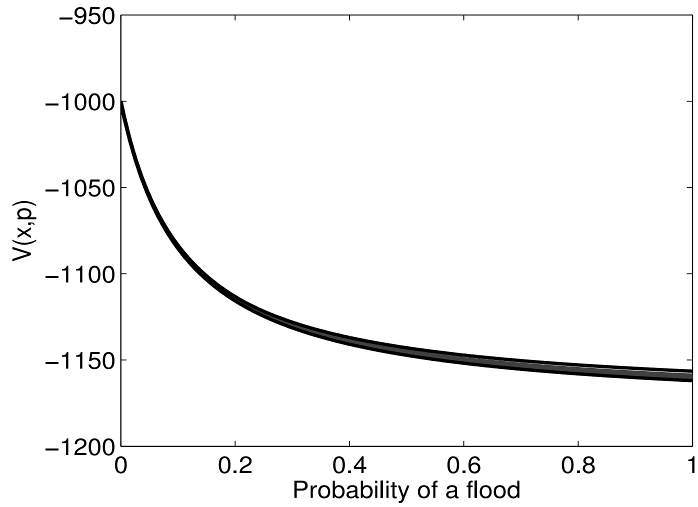

Figure 3: Value function $V(x, p)$, as a function of the flood probability x, for fixed values of p.

The upper part of the Montmorency River watershed receives quite important amounts of precipitation and, in particular, abundant snowfall (Matte et al., 2017). In addition, the response time of the watershed—defined as the delay between the onset of a precipitation event and the following rise in streamflow—is fairly short (12 h). Consequently, flooding is frequent during spring melt. The worst recorded flooding took place in the late fall of 1996 after an event of intense rainfall onto accumulated snow. Some downstream sections of the river are also subject to ice jams in the spring, which can create intense flooding when coupled with high streamflow. For instance, in April 2014, the breakup of an ice jam coupled with important runoff from snowmelt increased water levels to a semi-centennial level.

The topography of the downstream portion of the watershed causes certain communities to become entirely isolated when water levels rise. Consequently, the greatest concern for public authorities is related to people refusing to evacuate following a flood warning and evacuation order. Previous studies have argued against the implementation of structural flood control measures within this watershed. Thus, emergency evacuations, pumping, and sandbagging become the most common flood mitigation efforts to protect people and property. Alerts and evacuations are fairly common.

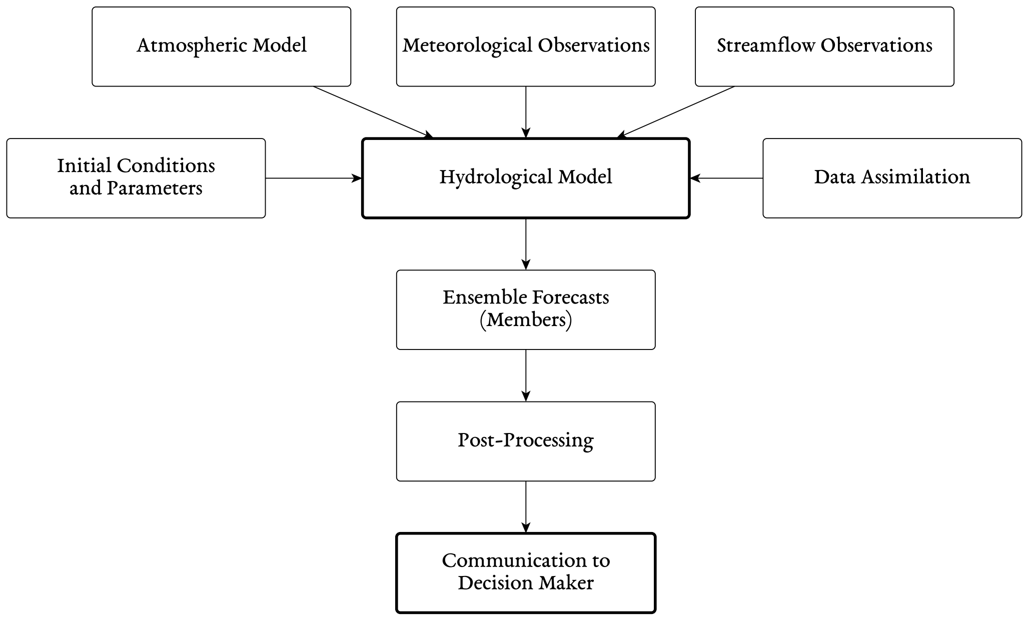

Figure 4: Schematic representation of the hydrological forecasting system.

3.2 Competing Hydrological Forecasting Systems

In this paper, we use the same three competing hydrological ensemble forecasting systems as in Matte et al. (2017). These three systems differ in how they represent uncertainties. Figure 4 presents a schematic description of the general forecasting framework.

At the core of all three forecasting systems is HYDROTEL (Fortin et al., 1995), a spatially distributed, physics-based model developed and maintained by the Institut National de Recherche Scientifique (INRS). It runs at a daily time step.

A hydrological model is conceptually analogous to a computable general equilibrium model in economics (e.g. Dixon and Jorgenson, 2013). There exists a plethora of models, and each can be conceptualized as a particular set of hypotheses in regard to the main hydrological processes (infiltration, groundwater and surface water flow, evapotranspiration, etc.). In practice, the choice of a specific model for operational purposes is often driven by the availability and cost of the model, and by the ease of implementation, its computational burden, and the data requirements.

HYDROTEL is also the model that is used operationally by the *Direction de l'Expertise Hydrique* (DEH), the group responsible for hydrological forecasting for the province of Quebec; HYDROTEL has been applied to the Montmorency River watershed since 2008 (Rousseau et al., 2008). Its implementation requires the calibration of unknown parameters. The calibration process consists of optimizing the value of those parameters so that the simulated streamflow values match the observations as closely as possible. For calibration purposes, HYDROTEL is fed with past observed (not forecasted) precipitation and temperature data. The operational implementation of HYDROTEL involves 27 free parameters that were calibrated using the Shuffle Complex Evolution algorithm of the University of Arizona (SCE-UA, Duan et al., 1994).

Once calibrated, HYDROTEL is driven by meteorological forecasts (daily precipitation and temperature, Figure 4). The three concurrent forecasting systems that are compared in this paper differ in this regard. Within the current operational set-up at the DEH, deterministic meteorological forecasts are used. This corresponds to the single most probable scenario per time step for future precipitation and temperature. To consider the uncertainty in future streamflow, the deterministic forecasts obtained from deterministic meteorological forecasts are then "dressed" using the statistics of past errors, as explained by Huard (2013). This allows for considering not only several possible future scenarios for streamflow, but also to compute the probability of exceeding a certain threshold related to flooding.

It is also possible to explicitly consider the uncertainty of future precipitation and temperatures by using meteorological ensemble forecasts. They consist of an ensemble of equiprobable scenarios for future meteorological variables (inputs). Each scenario can be fed, in turn, to HYDROTEL to obtain an ensemble of streamflow forecasts. Each scenario (called a member) in meteorological ensemble forecasts corresponds to a specific initial state of the atmosphere. The initial condition of the atmosphere cannot be observed exactly nor everywhere, and because of the chaotic nature of the atmosphere, this initial state at the onset of a forecast has a large influence on forecasts.

As described in Figure 4, the model also includes "data assimilation". As for an atmospheric model, the initial state of a watershed (as represented in the model) at the onset of a forecast cannot be measured precisely. Variables that define the hydrological state of a watershed are typically soil moisture in the top layers, ground water level, snow water equivalent, etc. These variables fluctuate substantially across space and time. If a new forecast is to be produced at time t, data assimilation consists of using available observations (for instance, streamflow) from previous periods to update and improve the estimation of state variables in the model.

Several methods exist for performing automated data assimilation (Montzka et al., 2012). However, in the operational setting at the DEH, data assimilation is performed manually. The forecaster modifies past observed precipitation and/or temperatures until the model's simulation run agrees with the observed streamflow at time t.

This manual data assimilation process was performed for all three forecasting systems. However, one system (System C) also includes an estimate of state variable uncertainty. Instead of considering the single best estimation of state variables, plausible scenarios are considered using the framework already described in Matte et al. (2017).

In the next section, we describe the properties of the forecasts under each system.

3.3 Characteristics of the Forecasting Systems

For each system, forecasts have been produced for the 2011–2014 period. We first look at the distribution of the errors. Specifically, we first compare the predicted streamflow value[11] with the observed streamflow values. Figure 5 displays the errors for the three systems over time. Figure 6 presents the kernel density estimates of the error distribution, aggregated over time.[12]

All systems exhibit some sort of bias. System A (the one used operationally) underestimates the streamflow value, while Systems B and C (the meteorological ensemble forecasts) typically overestimate the streamflow value. As demonstrated in Figure 10 in Appendix 7.2, this occurs mostly because extreme scenarios are more often included in the meteorological ensemble forecasts.

Of course, this is not necessarily representative of the usefulness of the forecasts. Since the systems provide an estimate of the distribution of streamflow values at each period, we compute the implied probability of a flood.[13] Table 2 and Figure 7 illustrate that ensemble forecasts (Systems B and C) predict a positive flood probability more often than the operational forecast (System A). However, most of the probability mass is on very small probabilities of flooding. By itself, this is neither positive nor negative, as it depends on the observations and on the decision maker's reaction to those forecasts.

In the next section, we use our decision model (Section 2) to discriminate between the three

[11]That is, the expected value, averaged across ensemble members.

[12]Formal statistical tests are presented in Table 6 in the Appendix. Note that Figures 5 and 6 hide considerable heterogeneity. Indeed, the distribution of errors is highly heteroskedastic, larger streamflow values being much harder to predict.

[13]This is simply given by the proportion of ensemble members that predict the streamflow value above the flood threshold.

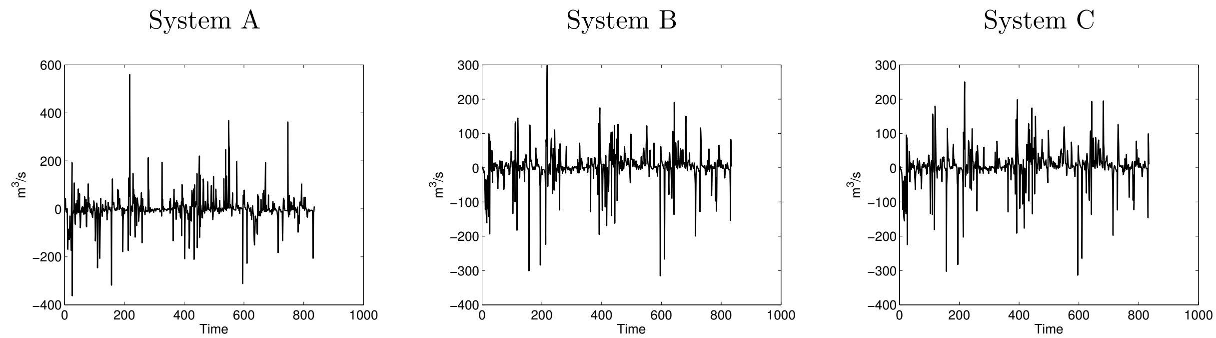

Figure 5: Average of prediction error for each system.

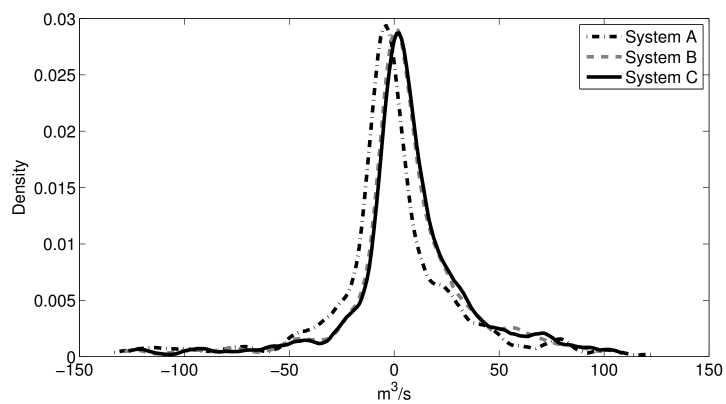

Figure 6: Kernel density of prediction error for each system.

systems.

Table 2: Probabilities of flooding.

	System A	System B	System C
Freq. $x = 0$	93.30%	85.17%	85.17%
Average prob.	2.09%	2.24%	2.37%

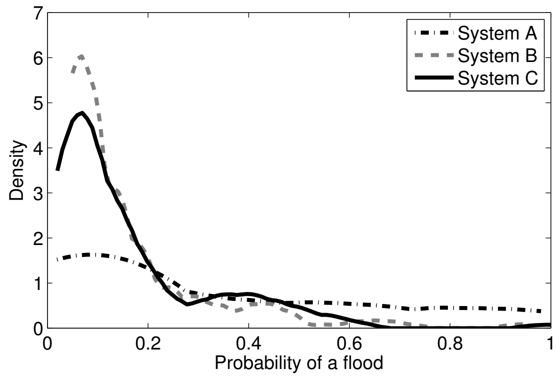

Figure 7: Probabilities of a flood. $(x|x > 0)$

4 Results

We fed the decision model with streamflow forecasts from the three concurrent systems described above. Figure 8 presents the evolution of the two state variables as a function of time. Figure 9 also displays the density estimates for the population's confidence, over time, for the three models.

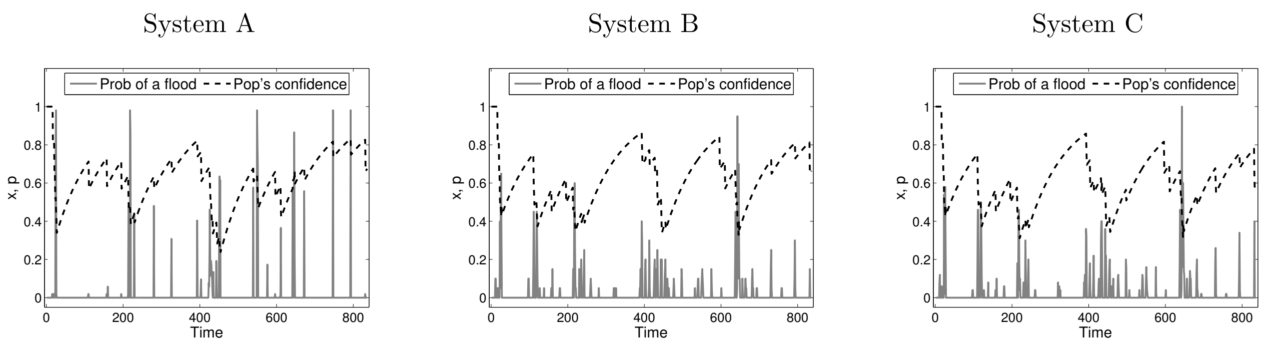

Figure 8: Evolution of the population's confidence and the forecast probability of floods for each forecast.

The population's confidence varies more for the two ensemble models (i.e. Systems B and C). Coherently with the results presented in Section 3.3, the fact that Systems B and C predict higher flood probabilities leads the decision maker to alert the population more often. Table 3 shows the number of alerts, false alerts, and missed events for the three systems. System B shows a lower rate

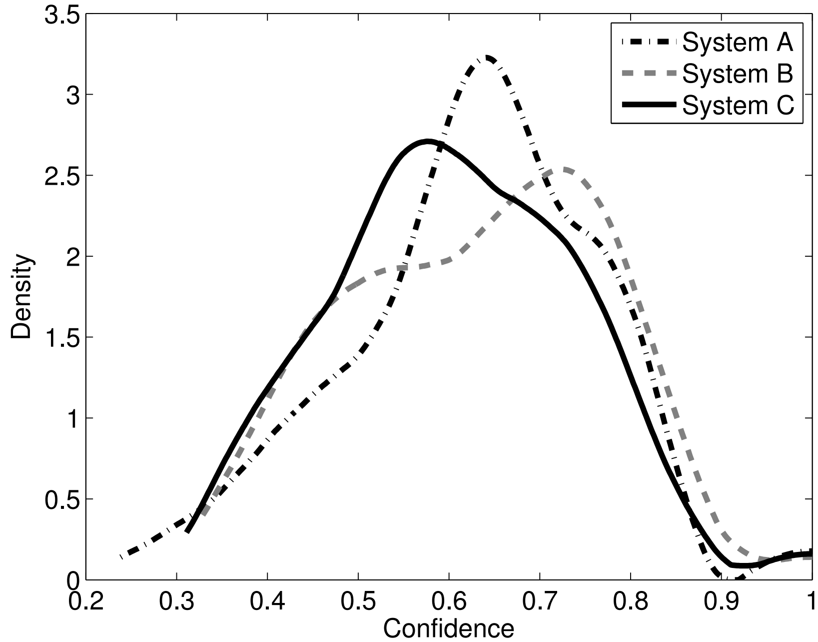

Figure 9: Kernel density of the population's confidence for each system.

of false alerts and fewer missed events.

That being said, the implications of Table 3 for welfare are not necessarily obvious. Indeed, the population's welfare depends on the monetary payoffs associated with each state, as well as on the curvature of the utility function. In the next section, we present a formal welfare analysis.

Table 3: Number of alerts, false alerts, and missed events for the three forecasts.

	System A	System B	System C
Alerts	29	32	38
False alerts	27	29	35
Rate of false alerts	93.10%	90.63%	92.11%
Missed events	14	13	13

4.1 Population's Welfare

To assess the population's welfare, we chose to evaluate the expected instantaneous utility under each forecasting system and under the optimal dynamic decision rule $\sigma(x, p)$. An estimator of that quantity is given by the following "ex-post" utility:

$$\mu_{post} = \frac{1}{T}\sum_t \sigma_t[p_t\mu(1, y_t) + (1 - p_t)\mu(0, y_t)] + (1 - \sigma_t)\mu(0, y_t),$$

where $\sigma_t = \sigma(x_t, p_t)$ is the optimal rule given x_t and p_t. Note that although the flood dynamic y_t is the same across forecasting systems, the population's confidence p_t and the optimal decision σ_t differ depending on which hydrological forecasting system is used.

Note also that using the value function $V(x, p)$ from the decision model would not provide a meaningful evaluation of the population's welfare. The main reason is that the value function includes the decision maker's beliefs (expectations) about the future values of state variables and utility (even if those are not realized). Then, while the value function $V(x, p)$ is a good tool to analyze the decision maker's optimal decision, welfare must be defined for the population, given these decisions.

Table 4 presents the values for each forecast. Note that this comparison allows to find the "best" forecast (i.e. System B), but it does not provide a measure of the value of the forecasts.

To gain more insight, we borrowed from the theory of decision under risk and defined the quantity Q as follows:

$$\begin{aligned}
\mu_{post}(Q) &= \frac{1}{T}\sum_t \sigma_t[p_t\frac{-1}{A}\exp\{-A[(b-d)y_t - c + Q]\} + (1-p_t)\frac{-1}{A}\exp\{-A[-dy_t + Q]\}] \\
&\quad +(1-\sigma_t)\frac{-1}{A}\exp\{-A[-dy_t + Q]\} \\
&= \mu^*_{post}
\end{aligned}$$

where μ^*_{post} is the ex-post utility for the best forecast. Importantly, as the preferences are represented by a CARA utility function, adding Q does not change the optimal decision rule.

The value Q is therefore interpreted as the daily willingness to pay for the best forecast (given the actual forecast used). Table 4 presents the values for Q. Estimating the long-term gains from changing the system is, however, less obvious. The reason is that even if wealth does not change the optimal decision for a CARA utility function, it does affect the marginal utility of money. Then, inter-temporal utility functions, such as $U_{post} = \sum_{t=0}^{T} \delta^t U_{w_t}(y_t, p_t, \sigma_t)$, are highly path dependent.

For this reason, we prefer to simply focus on the long-term monetary benefits (discounted or not) implied by Q. Those are also displayed in Table 4.

Looking at Table 4, we see that the willingness to change System A is consequent ($36 500/year). However, this number must be compared to the cost of implementing a new operational system, which includes infrastructure spending, but also to training costs. Also, we see that the cost of using System C (conceptually similar to System B) would be expensive.

	System A	System B	System C
Ex-Post utility	-100,32463	**-100,32455**	-100,3259
Q	100$	-	-13 000$
(Undiscounted) Yearly value	36 500$	-	-4 745 000$

Table 4: Welfare

The reasoning can be observed in Figures 1 and 7. System C, in particular, features a number of extreme scenarios where the predicted streamflow value is very high. These scenarios never occur in the data, and they usually imply a positive (albeit small) flood probability. If, for some reason, the economy is in a state where the population's confidence is low, then the optimal rule is guided by the very small probability of severe flooding (see Figure 1). If those small probabilities are biased, the decision maker will issue more false alerts, and the economy could remain in a state where both the confidence and welfare are low.

5 Alternative Specifications

In this section, we present some alternative specifications from the decision model to reach more general conclusions. Table 5 presents changes in the calibration along two dimensions: the value of avoidable damages b and the "elasticity" of the belief function. We gain the following insights:

1. Increasing the value of avoidable damages favors System A over Systems B and C. When the value of avoidable damages increases, the value of the population's confidence increases. As displayed in Figure 9, Systems B and C lead to more risky distributions of confidence. In particular, the lower tail of the distribution of beliefs (of paramount importance for a risk-averse decision maker) is thicker. When avoidable damages increase, the cost of low confidence

also increases.

This is especially true as increasing the value of avoidable damages increases the incentive to alert the population (and for lower probability thresholds). As displayed in Table 7 in the Appendix, this leads the decision maker (under Systems B and C) to issue many false alerts.

2. Higher variability of the population's confidence favors System A over B and C.[14] The reasoning is similar. When confidence reacts more to the sequence of alerts and floods, the cost of missed floods and false alerts is higher. Since Systems B and C exhibit a greater variability, this translates to an even higher variability in confidence levels and in particular, to thicker lower tails.

   In particular, the reduced variability of the population's confidence decreases the probability of achieving a very low confidence. This, in turns, prevents Systems B and C from justifying frequent alerts, as opposed to the baseline scenario (see Table 8 of the Appendix).

3. Higher variability of the population's confidence can be beneficial for welfare. Everything else being equal, increasing the variability of beliefs may improve welfare (see Table 5). Indeed, the cost of previous mistakes vanishes more rapidly since the decision maker may "escape" periods of low confidence more easily.

As also discussed previously, Systems B and C feature extreme (although not necessarily realistic) scenarios (see Figure 10 of the Appendix). A risk-averse decision maker tends to focus more heavily on these scenarios and, therefore, this is detrimental for welfare.

In the context of this paper, the effect of risk aversion can be seen by the large weighting on low flood probabilities in Figure 7. If the population's confidence remains high enough that the alert threshold is above this mass, then Systems B and C are preferred to System A. However, the opposite is true if the parameters of the model require alerting the population to a very low flood probability. In this latter case, Models B and C lead to a vicious cycle of false alerts and low confidence.

[14] We consider the second parametrization of ϕ as being less elastic or less variable. From Equation (2), we see that decreasing the values of ϕ reduces the expected growth rate of the belief function. Note, however, that since ϕ_1 is left untouched, the "natural growth rate" (no flood, no alert) is relatively more important. This helps to avoid states of low confidence.

Table 5: Ex-post utility for each system (maximal values are in bold).

$(\phi_1, \phi_2, \phi_3, \phi_4)$	(c, b, d)	System A	System B	System C
$(0.01, 0.1, 0.2, 0.1)$	$(0.364, 1.285, 15.46)$	-100.3246	**-100.32455**	-100.3259
	$(0.346, \mathbf{1.546}, 15.46)$	**-100.3246**	-100.3267	-100.3248
	$(0.346, \mathbf{2.570}, 15.46)$	**-100.3213**	-100.3262	-100.3245
$(0.01, \mathbf{0.08}, \mathbf{0.08}, \mathbf{0.05})$	$(0.346, 1.285, 15.46)$	-100.3250	**-100.3249**	-100.3252
(less elastic beliefs)	$(0.346, \mathbf{1.546}, 15.46)$	-100.3255	**-100.3250**	-100.3253
	$(0.346, \mathbf{2.570}, 15.46)$	-100.3246	-100.3257	**-100.3244**

6 Next Steps

We developed a dynamic model for the design of optimal warnings when the population's confidence is imperfect. We found that the optimal decision rule implies that, for a warning to be issued, higher flooding probabilities are required when the population's confidence is high.

We used the decision model to compare welfare under alternative hydrological forecasts. We found that the forecasting system that mimics the operational one (in the province of Quebec, Canada) could be improved by using meteorological ensemble forecasts rather than deterministic forecasts. However, since ensemble meteorological forecasts often lead to unrealistic extreme scenarios, there is a risk that the economy becomes stuck in a state where confidence is low, and the decision maker issues many false alerts.

The mechanisms identified in this paper may therefore provide additional guidelines to decision makers and help design better hydrological forecasting systems. Moreover, the decision model developed in this paper could be extended in a number of ways.

We assumed, coherent with our application, that a flood event is binary. However, in other contexts, the population may face different flood magnitudes. This in turn may lead the decision maker to issue different types of alerts (e.g. yellow, orange, red). Similarly, the question of timing may be important. Some adaptation or mitigation measures take time to implement, and the decision maker may want to alert the population sooner rather than later, while taking into account that uncertainty is reduced over time.

The study of these important issues involves solving non-trivial technical challenges and is left for future research.

References

Bakkensen, L. A. and L. Barrage (2017): "Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water?" Tech. rep., National Bureau of Economic Research.

Dixon, P. B. and D. W. Jorgenson (2013): *Handbook of computable general equilibrium modeling: volume 1A*, Elsevier.

Duan, Q., S. Sorroshian, and V. Gupta (1994): "Optimal use of the SCE-UA global optimization method for calibrating watershed models," *Journal of Hydrology*, 158, 265–284.

Fortin, J.-P., R. Moussa, C. Bocquillon, and J.-P. Villeneuve (1995): "HYDROTEL, un modèle hydrologique distribué pouvant bénéficier des données fournies par la télédétection et les systèmes d'information géographique," *Revue des Sciences de l'Eau / Journal of Water Science*, 8(1), 97–124.

Gordon, H. S. (1954): "The economic theory of a common-property resource: the fishery'," *The Journal of Political Economy*, 62, 124–142.

Huard, D. (2013): "Analyse et intégration d'un degré de confiance aux prévisions de débits en rivière," Tech. rep., David Huard Solution, Quebec.

Katz, R. and A. Murphy (1997): *Economic value of weather and climate forecasts*, New York: Cambridge University Press.

Leclerc, M., M. Morse, J. Francoeur, M. Heniche, P. Boudreau, and Y. Secretan (2001): "Analyse de risques d'inondations par embâcles de la rivière Montmorency et identification de solutions techniques innovatrices - Rapport de la Phase I - Préfaisabilité," Tech. Rep. R577, INRS-Eau and Laval University, Quebec.

Lucas, R. and N. Stokey (1989): *Recursive methods in dynamic economics*, Cambridge MA: Harvard University Press.

Matte, S., M.-A. Boucher, V. Boucher, and F. F. Thomas-Charles (2017): "Moving beyond the cost–loss ratio: economic assessment of streamflow forecasts for a risk-averse decision maker," *Hydrology and Earth System Sciences*, 21, 2967–2986.

Montzka, C., V. Pauwels, H. Franssen, X. Han, and H. Vereecken (2012): "Multivariate and multiscale data assimilation in terrestrial systems: A review," *Sensors*, 12, 16,296–16,316.

Murphy, A. (1977): "The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation," *Monthly Weather Review*, 105, 803–816.

Pindyck, R. S. (1984): "Uncertainty in the theory of renewable resource markets," *The Review of Economic Studies*, 51, 289–303.

——— (2007): "Uncertainty in environmental economics," *Review of Environmental Economics and Policy*, 1, 45–65.

Pollak, R. A. (1998): "Imagined risks and cost-benefit analysis," *The American Economic Review*, 88, 376–380.

Roulin, E. (2007): "Skill and relative economic value of medium-range hydrological ensemble predictions," *Hydrology and Earth System Sciences*, 11, 725–737.

Rousseau, A., S. Savary, and B. Konan (2008): "Implantation du modèle HYDROTEL sur le bassin de la rivière Montmorency afin de simuler les débits observés et de produire des scénarios de crues du printemps pour l'année 2008," Tech. Rep. R921, INRS-Eau, Quebec.

Rust, J. (1996): "Numerical dynamic programming in economics," *Handbook of Computational Economics*, 1, 619–729.

Sah, R. and J. Zhao (1998): "Some envelope theorems for integer and discrete choice variables," *International Economic Review*, 39, 623–634.

Salanié, F. and N. Treich (2009): "Regulation in happyville," *The Economic Journal*, 119, 665–679.

Singh, R., Q. Weninger, and M. Doyle (2006): "Fisheries management with stock growth uncertainty and costly capital adjustment," *Journal of Environmental Economics and Management*, 52, 582–599.

Verkade, J. and G. Werner (2011): "Estimating the benefits of single value and probability forecasting for flood warning," *Hydrology and Earth System Sciences*, 15, 3751–3765.

Weitzman, M. L. (2002): "Landing fees vs harvest quotas with uncertain fish stocks," *Journal of Environmental Economics and Management*, 43, 325–338.

7 Appendix

7.1 Analytical Properties

The standard argument used to establish the properties of V is as follows (e.g. Lucas and Stokey, 1989). We suppose that V belongs to the set $\mathcal{V}$ (e.g. the set of functions that are increasing in p and decreasing in x).

Then, we define the following mapping:

$$T(x,p) = \max_{\sigma}\{U(x,p,\sigma) + \delta \mathbb{E}_y V(g(x,y), h(p,y,\sigma))\}$$

If T also belongs to $\mathcal{V}$, the contraction mapping theorem (e.g. Lucas and Stokey (1989)) implies that the fixed point V effectively belongs to $\mathcal{V}$.

Here, this could be done by invoking a discrete version of the envelope theorem (Sah and Zhao, 1998). However, note that for $\sigma = 1$, $U(x,p,\sigma)$ is not necessarily increasing in p (only if inequality (1) holds). This is intuitive. If the optimal decision is not to alert the population, then higher confidence levels translate into higher costs of being wrong if a flood actually happens.

A similar argument holds for x. Although $U(x,p,\sigma)$ is decreasing in x, $T(x,p)$ is not necessarily decreasing as well. For example, if ϕ_4 is close to 1 (false alarms strongly affect confidence), then higher values of x (for x close to 1) does not change $U(x,p,\sigma)$ much, but reduces the probability of a false alert, which may have a positive effect on the decision maker's utility. Of course, these effects are unlikely to occur in practice, but they do prevent any analytical results on the monotonicity of $V(x,p)$.

Similar effects impact the analytical study of $\sigma(x,p)$.

We have $\sigma(x,p) = 1$ if:

$$U(x,p,1) - U(x,p,0) + \delta \mathbb{E}_y[V(g(x,y), h(p,y,1)) - V(g(x,y), h(p,y,0))] \geq 0$$

(while (1) holds if $U(x,p,1) - U(x,p,0) \geq 0$). Even with strong conditions for ϕ_1, ϕ_2, ϕ_3 and ϕ_4 and for the convexity/monotonicity of V, the inequality $\mathbb{E}_y[V(g(x,y), h(p,y,1)) - V(g(x,y), h(p,y,0))] \geq 0$ is not guaranteed, when the exact shape of $V(x,p)$ is unknown.

Consider, for example, Numerical result 1. It holds for reasonable parameter values but does not hold in general. To see why, let us assume that V is increasing in p, and consider $\phi_1 = 1$, and $\phi_2 = \phi_3 = \phi_4 = 0$, as well as $p = 0$ and x high enough so that (1) holds. Then, the optimal solution for the static problem is to alert the population. However, the optimal solution for the dynamic problem is $\sigma(x,0) = 1$ if $V(x,0) \geq V(x,1)$, which is not true as $V(x,p)$ is increasing in p. By continuity, this argument holds for small values of ϕ_2, ϕ_3, ϕ_4 and p and large values for ϕ_1.

Finally, since the exact shape of V is unknown, it is therefore not possible to establish bounds on those parameters (as a function of the model's primitives).

7.2 Additional Tables and Figures

Table 6: Diagnostics of prediction errors.

	A vs B	A vs C	B vs C	A†	B‡	C‡
Student test	*H0: difference in means is equal to 0*			*H0: distribution's mean is zero*		
statistic	-1.8623	-2.4822	-0.7065	-0.5823	2.2273	3.1739
$p-value$	0.0627	0.0132	0.4800	0.2803	0.0131	0.0008
Wilcoxon test	*H0: Both distributions have equal medians*			*H0: median of distribution is equal to 0*		
statistic	-6.5168	-7.5151	-1.1558	-2.3809	6.7228	8.1422
$p-value$	0.0000	0.0000	0.2478	0.0086	0.0000	0.0000

† left side, ‡ right side

Table 7: Number of alerts, false alerts, and missed events for the three forecasting systems. Baseline belief updating, $b = 1.546$.

	Baseline belief updating, $b = 1.546$		
	System A	System B	System C
Alerts	31	46	49
False alerts	29	43	44
Rate of false alerts	93.5%	93.5%	89.8%
Missed events	14	13	11

Table 8: Number of alerts, false alerts, and missed events for the three forecasting systems. Less elastic belief updating, $b = 1.546$.

	Less elastic belief updating, $b = 1.546$		
	System A	System B	System C
Alerts	28	31	37
False alerts	26	28	33
Rate of false alerts	92.9%	90.3%	89.2%
Missed events	14	13	12

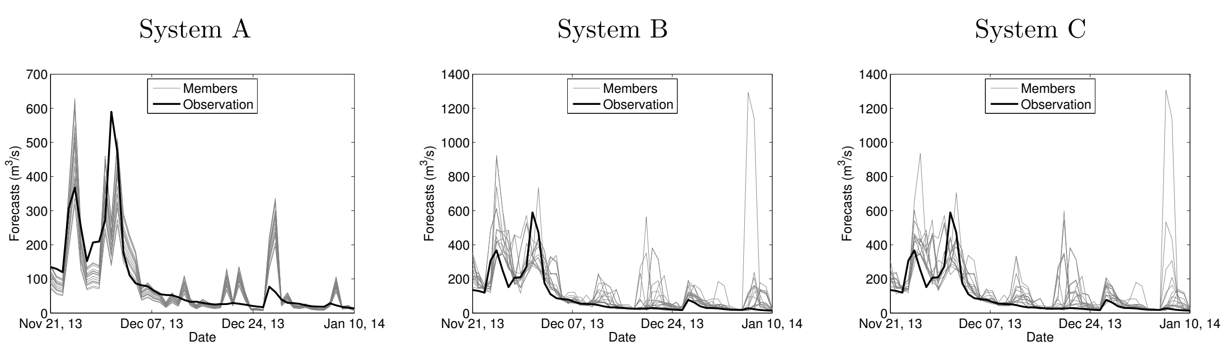

Figure 10: Forecast members for each system.