Deltares

Enabling Delta Life

2009.06.01

Uncertainty Framework for Operational Flood and Storm Surge Forecasting

Flood Control 2015

Deltares

2009.06.01

Uncertainty Framework for Operational Flood and Storm Surge Forecasting

Albrecht Weerts
Joost Beckers

Flood Control 2015

1200379-001

Deltares

Title
2009.06.01

Client	**Project**	**Reference**	**Pages**
FC2015	1200379-001	1200379-001-ZWS-0002	64

Keywords
Uncertainty Framework, Data Assimilation, Conditioning, Forecast Calibration

Summary
This research was carried out within the Flood Control 2015 program. The present report the activities on the development of an overarching framework for assessing uncertainties in fluvial and coastal forecasting in a risk-based manner with the aim that it is robust enough to be considered for use in an operational environment. This framework is supplemented by two case studies that show how the framework can be applied.

Versie	Datum	Auteur	Paraaf	Review	Paraaf	Goedkeuring	Paraaf
	dec. 2009	Albrecht Weerts	AW	Jaap Schellekens		Toon Segeren	

State
final

Contents

1. Introduction

Robust forecasts (unbiased and skilful) are vital in providing a comprehensive flood warning service to people and businesses at risk from flooding. For fluvial and storm surge flood forecasting, rainfall–runoff, flow routing, 1D-hydraulic and 2D-hydraulic models are often combined into model cascades and are run automatically in operational flood and storm surge forecasting systems.

Often, the outputs from these models are currently deterministic with one model run delivering the flood forecast which is assumed to be the best representation, although Forecasting Duty Officers assess and advise on the uncertainty in forecasts based on experience and judgement. However, it is widely known that the accuracy of flood forecasts can be influenced by a number of factors, such as the accuracy of input data, and the model structure, parameters and state (initial conditions). Having a sound understanding of these modelling uncertainties is vital to assess and improve the flood forecasting service. This project (2009.06.01) has been carried out under the Flood Control 2015 program. This project has been carried out alongside / in cooperation with a R&D project SC080030 commissioned by the Environment Agency (Environment Agency, 2009). As a result an overarching framework for assessing uncertainties in fluvial and coastal forecasting in a risk-based manner has been developed with the aim that it is robust enough to be considered for use in an operational environment.

This framework is supplemented by a number of practical case studies which will demonstrate how certain uncertainty techniques can add value to the forecasting process (see Appendix B&C).

2. Framework

The main aim of the uncertainty framework is to assist flood forecasting experts (Monitoring and Forecasting Duty Officers) involved in commissioning, maintaining and improving models to decide which uncertainty estimation approaches are suitable in which circumstances and how they should be applied.

Some key goals are (1) to make clear what the sources are of the uncertainties (2) to make clear how these uncertainties propagate through the model cascade used for forecasting, and (3) to determine what methods are available to quantify and reduce the uncertainties.

A particular focus is on cascades of models, such as rainfall-runoff, hydrological and hydrodynamic routing models, and the choices to be made regarding data assimilation and calibration (or conditioning) of forecasts at each model boundary.

For example, when assessing the level of risk, the framework describes which method will be used, and how the resulting decision will influence the choice of method. By contrast, future guidelines will provide more guidance on how to derive these estimates (early examples of how these outputs might appear are shown in the appendices when applying the framework to the case studies).

The key components of the framework are described in the remainder of this chapter under the following section headings:

- Basic Principles – discusses the background to the key uncertainty estimation approaches of forward uncertainty propagation, data assimilation and conditioning that were established in the Phase 1 Report

- Choice of Method - discusses the key factors that influence the choice of uncertainty estimation method

The framework is intended to be generic and, with some further development, could be applied in any organisations responsible for flood and storm surge forecasting. To illustrate this wider application, Appendix B&C provides examples of application of the framework to the coastal forecasting models of the Storm Surge Warning Service (Stormvloed-waarschuwingsdienst/SVSD) and the forecasting system for Rhine and Meuse.

However, a key assumption in applying the framework is that an integrated catchment model is already available for the catchment under consideration, and that it is the uncertainty in model outputs which is of interest. If no model exists, then there are a number of guidelines on model development available from the USA, Europe and elsewhere, including the Environment Agency's Real Time Modelling guidelines (Environment Agency, 2002).
A brief introduction is therefore provided to the issues to consider, with references for further reading.

2.1 Basic Principles

2.1.1 Introduction

There are many sources of error in making flood forecasts. Such errors mean that all forecasts must be considered uncertain, and that there is a real possibility of getting a forecast wrong, both by not issuing a warning and flood damages being incurred or people injured, or by issuing a warning and no flood damages being incurred (i.e. the false alarm or "cry-wolf" problem).

It has long been recognised that both types of error will have an effect on the public perception of and reaction to flood warnings. Flood forecasting is therefore not just a scientific problem, it is a problem of managing and communicating uncertainty to specialist users (e.g. Monitoring and Forecasting Duty Officers), and more widely to the public, local authorities and the emergency services.

One way of dealing with the uncertainty associated with a forecast is simply to be specific and supply a measure of that uncertainty to assist users in decision-making. In order to develop an effective measure of uncertainty we need to address three critical issues:

- the representation of different types of uncertainties in the forecasting system
- the (preferably optimal) constraint of uncertainty in forecasts by means of real-time data assimilation
- the presentation of forecasts and their associated uncertainties to duty officers and possibly to other key decision-makers and the public.

This framework considers methods to address the first two of these problems whilst other Flood Control studies (2009.07 and others) are considering the third more general issue of the communication of uncertainty.

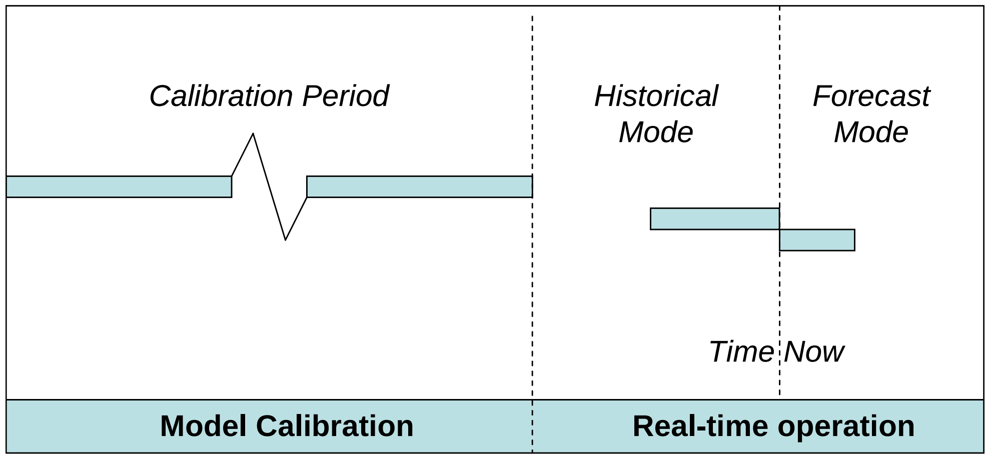
Real-time flood forecasting applications often make use of a cascade of inter-linked hydrological and, in some cases, hydrodynamic models, embedded in a data-management environment such as that of FEWS Rivers Rhine and Meuse. Model cascades (or integrated catchment models) are typically run in two principal modes of operation:

i) a historical mode – in which models are forced by hydrological and meteorological observations over a limited time period prior to the onset of the forecast (e.g. to initialise model stores)

ii) a forecast mode – in which models are run over the required forecast lead time, forced by outputs from other models, with the internal model states at the end of the historic run taken as initial conditions for the forecast run

Increasingly, models are forced using meteorological forecasts of precipitation and sometimes other variables, such as air temperature (e.g. where snowmelt is an issue) and evaporation, in addition to the use of forecasts from river locations further upstream.

Figure 2.1 illustrates these different modes of operation, and how they differ from the much longer period of records which are typically used in model calibration.



Figuur 2.1 Illustration of historical and forecast modes of operation

In each of the steps in the model and data processing chain, uncertainties can be attributed to the model inputs, the model structure, internal model states and model parameterisation, with the total predictive uncertainty accumulating in the forecast outputs (e.g. Beven 2009; Pappenberger et al. 2007).

Depending on the lead-time at which forecasts are issued in comparison to the hydrological response time, the dominant uncertainties will lie in the inputs derived from observations, the rainfall-runoff and routing models, or, if applicable, the hydrodynamic models, and from the uncertainty in rainfall and other meteorological forecasts (if used). The various time delays in the warning process also need to be considered, such as the time taken to collect data, run models, post-process results, take decisions and issue flood warnings, as discussed later.
The process of making a flood forecast can therefore be subdivided into three problems (e.g. Moll, 1986):

- Estimation of the actual state of the basin at the start of the forecast, which consists of interception storage, soil moisture storage, groundwater storage, other possible storages (e.g. snow storage), and the water levels in rivers, reservoirs, wetlands and lakes;

- Modelling of the movement of water during the period covered by the forecast lead-time through the whole cascade of rainfall-runoff, flow routing, and hydrodynamic models;

- Forecasting of the model inputs during the selected lead-time. These can consist of meteorological inputs, but also inflows from locations further upstream or at other model boundaries (for instance tidal influences, abstractions and discharges).

For an individual model component within an integrated catchment model, these three problems result in uncertainties / errors in the forecast consisting of:

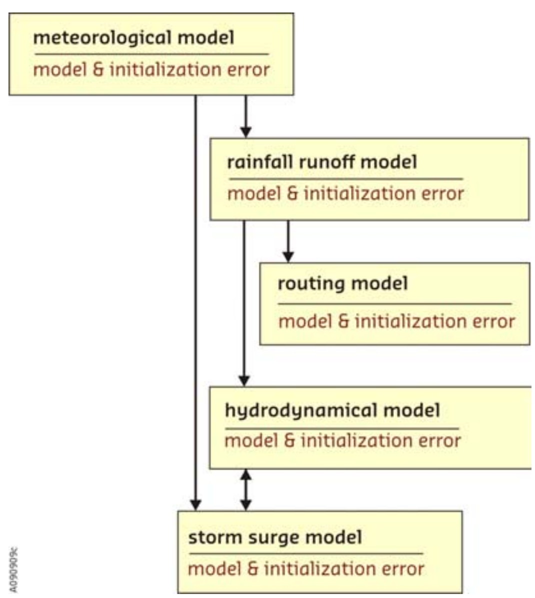
- Initialisation Errors: due to errors in the observations used to estimate precipitation, potential evaporation and temperature, discharge or other boundary conditions in the historical mode of operation;

- Model(ling) Errors: arising from approximating parameterisations/model structures, uncertain model parameters, model resolution limitations, uncertain structure operating/management rules, etc.

- Forcing Errors: errors which occur in the forecast mode of operation when a model component is forced with an input derived from another model with its own Initialisation and Model Errors; for example a Numerical Weather Prediction model or the hydrological or flow routing outputs at a flow forecasting point further upstream.

Note that the initialisation errors are usually not independent from the model errors because normally a model is used to derive the estimate of the actual state of the basins; for instance via data assimilation or just driving the model (cascade) in historical mode until the start of the forecast to estimate soil moisture storage and other variables of interest.

Within an Integrated Catchment Model, these errors combine as illustrated in Figure 2.2. This figure shows the picture for the whole modelling cascade, including the hydrodynamic and coastal components (if relevant). Initialisation and modelling errors occur in each of the different components resulting in forcing errors in the downstream model. The arrows indicate the source of the forcing error, modelling error(s) and initialisation error(s) in the model(s) higher up in the model cascade.



Figuur 2.2 Flood and storm surge forecast model cascade indicating the sources of the errors in the forecasts. The arrows indicate the errors in the forcing (when looking at it from the viewpoint of the receiving model cascade component) or model output (when looking at it from the viewpoint of the producing model cascade component)

2.1.2 Operational Strategy

The errors in the forecast model cascade need to be quantified and/or reduced for the following reasons:

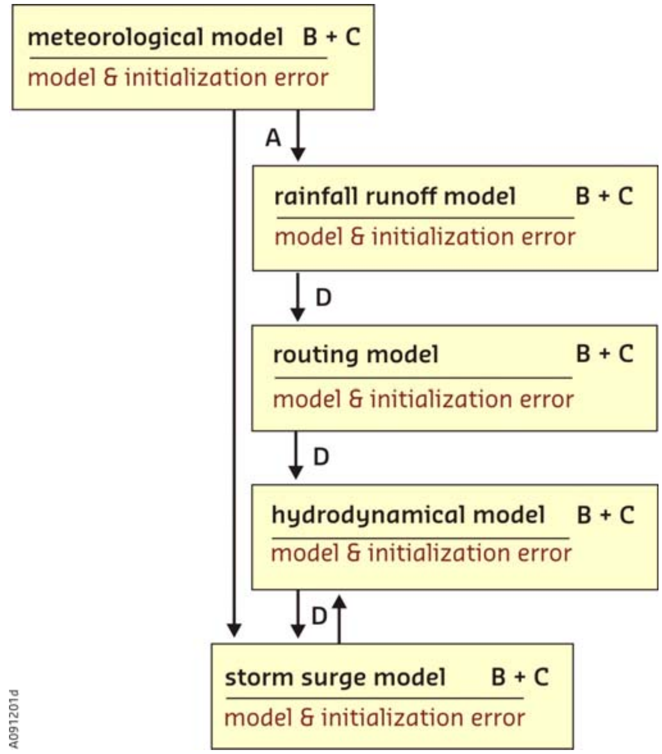
- to provide more accurate forecasts;
- to provide accurate information regarding the uncertainty of the forecast (and, if possible, unbiased and skilful estimates);

Quantification can be seen as providing a description/method to quantify the uncertainties and typically involves the use of forward uncertainty propagation techniques to give an idea of the uncertainty in the forecast. Of course, where suitable information is available, it is preferable to first reduce the uncertainties. This can be achieved by two key approaches (1) making use of recent observations (data assimilation) and (2) applying adjustments based on the historical performance of the forecasts made using the model cascade (forecast calibration).

Data assimilation is a feedback system where the forecast is conditioned on all available information that is available at the time the forecast is made (the forecast origin or ‘time now’). This includes information on the current state of the system, but also entails past performance of the forecast system (possibly further conditioned on secondary information such as the time of year, synoptic situation etc.).

Often, the term data assimilation is used to describe the use of real-time recent data to improve forecasts, whilst the term conditioning (on historical data) or forecast calibration is used to describe methods for improving forecasts based on the historical performance (i.e. not taking account of any real-time data which may be available). Alternative terms in flood forecasting include real-time updating or real-time adaptation.

Figure 2.3 shows schematically where these different approaches interact with the forcing errors, the rainfall-runoff, flow routing and hydrodynamic models, and the forecast produced by the end-to-end model cascade.



Figuur 2.3 Flood and storm surge forecast model cascade indicating the sources of the errors in the forecasts. The arrows indicate the errors in the forcing (when looking at it from the viewpoint of the receiving model cascade component) or model output (when looking at it from the viewpoint of the producing model cascade component), and where and with which methods uncertainty can be reduced and quantified in the model cascade (A) Meteorological Forecast Calibration, (B) State Updating, (C) Parameter Updating, (D) Forecast Calibration/Output Updating

These four operational approaches to updating which are described in the figure are more general forms of the widely used terminology in hydrological forecasting of input updating, state updating, parameter updating and output updating (e.g. Refsgaard, 1997; Serban and Askew 1991). The main operational approaches to quantifying and reducing these sources of uncertainty are as follows:

A: Meteorological Forecast Calibration

Rainfall forecasts are widely used in flood forecasting applications, particularly on fast responding catchments, and to provide long term outlooks on the potential for flooding. The main input is usually forecast rainfall, althoughforecasts of air temperature and other parameters may also be of interest in some applications (e.g. snowmelt forecasting).

An important aspect of hydrological ensemble forecasting is whether current atmospheric forecasts account for all of the important meteorological and climatological uncertainties. However, existing raw ensemble weather and climate forecasts meet the above properties only to a limited degree. This is due not only to the number of ensemble members being limited by computing resources (and hence subject to sampling uncertainty), but also currently most ensemble forecasting systems do not account for all significant sources of uncertainty, such as that arising from model parameterisations, and grid resolution. The result is that the forecasts are not necessarily reliable; they can be biased towards the mean and may not display enough variability, leading to an underestimation of the uncertainty (Buizza et al. 2005).

Different approaches have been proposed to derive reliable probabilistic forecasts from raw model ensembles, a process that often involves a combination of bias correction and downscaling in some form (see Weerts et al., 2010 and Environment Agency 2009 for an overview). Most of these methods are based on the idea of adjusting the current forecast using information derived from past forecasts and corresponding observations; an approach which is often called the method of Model Output Statistics (MOS). However, an important conclusion by Wilks and Hamill (2007) is that there appears to be no single best forecast method for all applications, and that extensive work is necessary on ensemble correction methods in the future.

These methods also typically require extensive hindcasting (>30 years) by meteorological offices which is often a problem due to limited money and computing resources, although several organisations such as NOAA/NCEP and ECMWF have performed such exercises. Moreover, the operational models used by meteorological offices are often changed several times per year and each time the model is changed an assessment should be performed because the performance of the meteorological model might have changed in space and time. Due to these practical limitations most flood forecasting agencies presently use the raw ensemble outputs as inputs to their models.

Some flood forecasting offices also use a so-called poor man's ensemble constructed of several deterministic models (see also Werner et al. 2010, Cloke and Pappenberger 2009; Environment Agency 2009), particularly for longer range and seasonal forecasts. Some research studies have also explored approaches to evaluate and combine outputs in real-time using techniques such as Bayesian Model Averaging (see Section 4). Techniques such as weather matching or analogue approaches have also been used with considerable success; these seek to identify features of the general atmospheric circulation in common with previous events, from which an ensemble of likely rainfall fields can be extracted from a historical archive (e.g. Obled et al. 2002).

One consequence of the uncertainties in meteorological inputs is that data assimilation or updating techniques, based on real-time river level or flow observations, are generally required when using rainfall or other meteorological forecasts as inputs to integrated catchment models, and are often combined with forecast calibration or conditioning techniques. This topic is discussed below under the heading of 'Output Updating and Forecast Calibration or Conditioning'.

B: State Updating - Data Assimilation

One role of data assimilation, in both deterministic and ensemble hydrological forecasting, is to produce the best possible estimates of initial hydrological conditions (i.e. to constrain uncertainty at the start of the forecast). Besides a realistic representation of the system state (soil moisture, snow water equivalent, groundwater, etc.) by the ensemble mean, the ensemble members must provide realistic estimates of the uncertainty in the system state. Observed precipitation and temperature data will typically be used in establishing (hydrological) model boundary conditions. Available observations can be effectively used in quantifying and reducing the error in the modelled water levels and discharges from the process models. These observations may include in-situ measurements of water levels, discharge, snow water equivalent, soil moisture and groundwater levels. They may include also remotely sensed observations from radar or satellites.

State updating can be done via manual, sequential or variational data assimilation techniques (Beven, 2009; Seo et al. 2009, Weerts et al., 2009; Weerts and Serafy, 2006; Clark et al., 2008; Heemink et al., 1997; Verlaan et al., 2005). Many flood forecasting offices use manual or deterministic state updating techniques in their flood forecasting system. However, ensemble data assimilation is still very much a research topic in operational hydrology. Moreover the question remains if the ensembles produced are really meaningful and this will probably depend very much on the assumptions made in the data assimilation scheme. Advanced time series analysis techniques, such as the Data-based Mechanistic (DBM) approach, also seek to combine the modelling and data assimilation stages, optimising the model outputs for the forecast lead times of interest.

Quantification of initalisation uncertainty can also be performed via forward uncertainty techniques in a similar way to the use of ensemble forecasting in numerical weather prediction.

C: Parameter Updating

Model errors can also arise from uncertainties in model parameters, and from more fundamental issues with the representation of physical processes i.e. model structural errors. Model parameter updating schemes have been developed for some types of model, but generally other approaches are favoured (such as state updating) for models which have a physical or conceptual basis. However, there can be some merit in sampling parameters within a given physically plausible range; an approach which is widely used in off-line research studies of model performance. One example is the manual modifications (MODS) approach that is used in the operational flood forecasting system used by the National Weather Service in the USA and which is now also available within Delft-FEWS.

However, structural errors in hydrological models are very difficult to correct. Because these errors tend to be strongly correlated in time (for lumped models) or in space and time (for distributed models), addressing them through post-processing requires complex and often heavily parameterised statistical modelling. In principle these errors should also be treated as part of reducing initialisation errors (see above).

One other way of looking at this problem is that using multi-model ensembles offers potential for accounting for structural uncertainty without such data- and parameter-intensive statistical modelling. Also, estimation theory states (e.g. Schweppe 1973) that combining informative forecasts from different models reduces forecast uncertainty (Georgakakos et al., 2004).

An interesting operational example to mention is the flood forecasting system of the River Po in Italy where three different lines of model cascades have been setup to produce forecasts for the whole basin of the Po. However, it is probably fair to say that explicit accounting for parametric uncertainty via parametric uncertainty processors and accounting for model structural uncertainty via multi-model ensemble techniques are only in their infancy in the operational arena.

D: Output Updating and Forecast Calibration or Conditioning

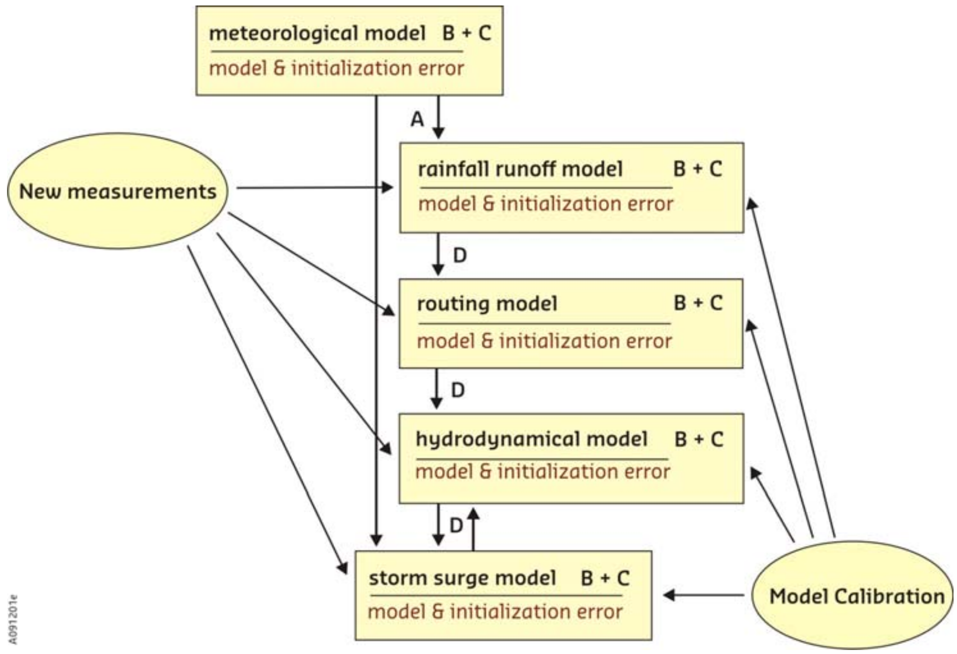
In a similar way to the post-processing of meteorological forecasts (see above), post-processing or output updating techniques have also been developed for hydrological applications. This is often done by applying simple autoregressive (AR) and/or moving average (ARMA) type models (Madsen et al. 2000; Broersen and Weerts, 2005). This type of error correction or reduction of forecast errors is used in many operational flood forecasting systems around the world. Some approaches, such as adaptive gain updating (Lees et al. 1994) also seek to both reduce and quantify the uncertainty in forecast outputs.

The development of post-processing techniques for ensemble forecasts is less far advanced, although many probabilistic techniques are now also being researched and tested in several operational flood forecasting systems around the world. For example, Bayesian uncertainty processors provide a promising approach (Krzyrsztofowicz, 2004). However, implementing such advances requires rather significant upgrades to current forecast systems, so it is likely that operational hydrologic ensemble forecasting will initially employ, a purely statistical “catch-all” ensemble post-processor to reduce and account for the integrated hydrological uncertainty (Seo et al. 2006), such as the Model Output Statistic (MOS) approach discussed earlier for meteorological forecasts

A final observation is that techniques for post-processing of ensemble hydrological forecasts tend to follow or ‘lag’ those developed in the meteorological community. For example, techniques such as quantile regression, Bayesian Model Averaging, and Model Output Statistics were first developed in meteorology, and other technical areas. Whilst some techniques are now established, when considering the best approach to use, it is worth noting that in many cases the methods proposed in the literature are promising but yet need to prove themselves in an operational real-time forecasting setting.

2.1.3 Long-term strategy

Besides the operational strategy to quantify and reduces uncertainties real-time there is also a more long-term strategy necessary, as shown in Figure 2.4, which is focused on (1) improving hydrological, routing, hydrodynamic, storm surge models. This can include improvements to model structure, schematization, model resolution, calibration etc. And (2) the introduction of new and/or better measurements and improved/other numerical weather predictions/forecasts. In the end, this will also help to reduce the uncertainties in the real-time flood and storm surge forecasts. Such improvements will help to reduce both initalisation errors and model errors during operational forecasting.



Figuur 2.4 Flood and storm surge forecast model cascade indicating the sources of the errors in the forecasts. The arrows indicate the errors in the forcing (when looking at it from the viewpoint of the receiving model cascade component) or model output (when looking at it from the viewpoint of the producing model cascade component), and where and with which operational methods uncertainty can be reduced and quantified in the model cascade (A) Meteorological Forecast Calibration, (B) State Updating, (C) Parameter Updating, (D) Forecast Calibration/Output Updating. Together with the long-term options to improve the forecasting system (better models & better/more measurements)

2.2 Choice of Method

The choice for a specific uncertainty estimation method depends on a number of key factors:

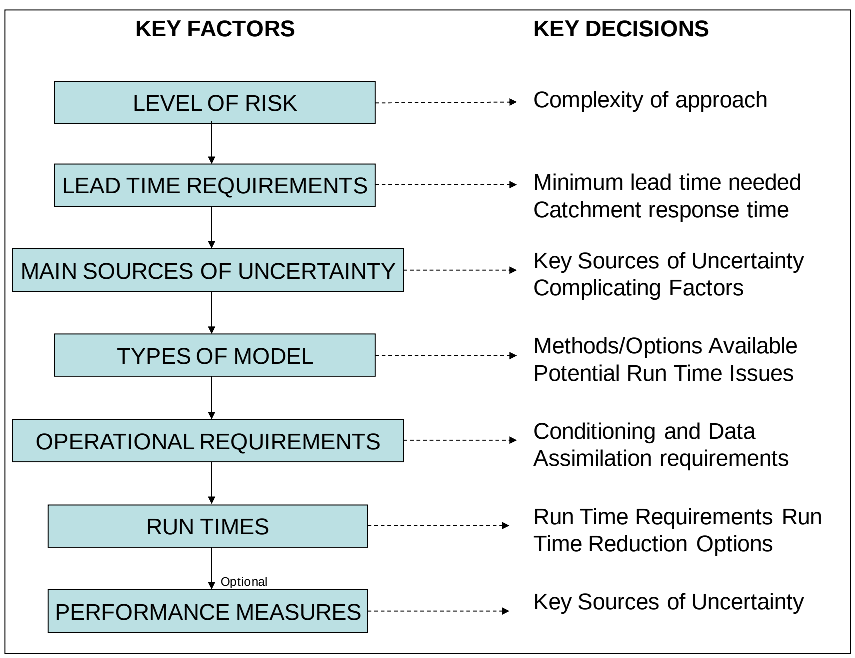
- Level of Risk – what is the risk at the forecasting points of interest, and hence how accurate does the method have to be and/or how much effort is it worth expending?
- Lead Time Requirements – what minimum lead time is required for flood warning or operational reasons, and how does that compare to the catchment response time (assessed in the form of a catchment ‘Type’, on a scale of 1 to 5)?
- Main Sources of Uncertainty – given the catchment type, what the the most likely sources of uncertainty (rainfall forecasts, river flow observations etc.) and what complicating factors may need to be considered (e.g. structures, reservoirs)?
- Types of Models – what types of models does the integrated catchment model use, both in general terms (e.g. rainfall-runoff models, hydrodynamic models), and specific brands (e.g. HBV-96,PDM, ISIS, SOBEK, WAQUA, Delft3D), and how does this influence the choice of uncertainty estimation methods?
- Operational Requirements – how will the forecasts be used operationally and how does this influence the choice of uncertainty estimation approach?
- Runtimes – how much computer processing time is likely to be required for each type of method, and is this achievable with current systems?

- Performance measures – are estimates for model performance already available which would indicate which locations in the catchment/parts of the model should be focussed on?

This section presents the key steps in applying the framework to take account of these factors, including the key decisions to be made, and the main steps towards reaching that decision. Figure 2.5 provides a general overview of the decision making approach. Note that the level of risk is used in several of the key decisions (although not all), and that some steps are optional, depending on what other information may be available.

The background to development of the framework appears in Appendix A, whilst examples of applying the framework appear in Appendices B to D.

The intention of this section is therefore to present a brief outline of the main contents of the framework, which might be used as a checklist, with more detailed information on applying it presented in the appendices. A key decision for the guidelines will be to achieve an appropriate balance between the background material presented in Section 2.2 and Appendix A, and the overall methodology.



Figuur 2.5 Summary of key decisions arising from consideration of key factors

Tabel 2.1 Template for the worksheet used in the case studies

Factor	Key Decisions	Main Findings
Level of Risk	What is the level of risk at individual Forecasting Points or flood risk areas?	
	What is the level of risk at a catchment level ?	
	What complexity of approach is generally to be preferred ?	
Lead Time Requirements	What are the lead time requirements for each Forecasting Point ?	**Flood Warning**
		Outlook Statement
	What are the main forcing inputs for each Forecasting Point at those lead times ?	**Flood Warning**
		Outlook Statement
	What, at a catchment level, are the key forcing inputs to consider for flood warnings and outlook statements ?	**Flood Warning**
		Outlook Statement
Main Sources of Uncertainty	What are the main sources of uncertainty for the catchment for flood warnings ?	**Initialisation Errors** **Modelling Errors** **Forcing Errors**
	What are the main sources of uncertainty for the catchment for outlook statements ?	**Initialisation Errors** **Modelling Errors** **Forcing Errors**
	What additional sources of uncertainty arise from complicating factors ?	
Types of Models	What choices of methods are available for the types of models ?	
	What types of data assimilation routines are an option ?	
	What potential run time issues have been identified ?	
Operational Requirements	Is a purely qualitative approach sufficient for generating ensembles?	
	Is data assimilation desirable or essential ?	
	Is conditioning of forecast outputs required ?	

Run Times	Are there run time issues for the candidate uncertainty estimation methods ?	
	What are the options for reducing run times ?	
Performance Measures	Are suitable performance measures already available to evaluate sources of uncertainty?	
	Which sources of uncertainty (and locations) are identified ?	

2.2.1 Level Of Risk

Key Decisions

What is the level of risk for individual forecasting points or flood risk areas ?

What is the level of risk at a catchment level ?

What complexity of approach is generally to be preferred ?

Key Steps

Produce a catchment map or table of areas at risk, showing the level of risk at individual forecasting points

Assess the risk at a catchment level

Select the complexity of approach based on the level of risk

In recent years, organisations such as the Environment Agency and the Rijkswaterstaat in the Netherlands have been at the forefront of developing risk-based approaches in flood management.

This approach has also been adopted in flood forecasting and warning applications; for example, on deciding on an appropriate level of service in designing flood warning schemes. A similar approach has also been adopted here, and a key output is to define the level of risk at individual forecasting points, and at a catchment scale.

The appendices provide background on the steps required, and on how the level of risk can be used as a deciding factor in the choice of uncertainty estimation method (and in the level of detail to use when applying the framework).

2.2.2 Lead Time Requirements

Key Decisions

What are the lead time requirements for each Forecasting Point ?

What are the main forcing inputs for each Forecasting Point at those lead times ?

What, at a catchment level, are the key forcing inputs to consider for flood warnings and outlook statements ?

Key Steps

Define the catchment response time at each forecasting point
Define the lead time requirements, including an allowance for decision-making and warning times, and hence the catchment type
Aggregate the findings to catchment level for flood warnings and outlook statements

The forecasting lead-time required depends primarily on the lead-time requirement for flood warning, and may extend from as little as 1-2 hours to several days ahead.

In the latter case, the shorter lead-time forecasts may be used in issuing the actual operational warning, while forecasts at the longer lead time are used mainly as guidance in moving to a flood alert status, rather than to guide the issuing of a flood warning. A distinction is therefore made between the lead-time required for issuing flood warnings, and the time required for a lower level of alert (called an Outlook Statement here).

The appendices describe methods for estimating the lead time requirement and a catchment classification scheme for flood forecasting in which five types of catchment are defined (Types 1 to 5).

2.2.3 Main Sources of Uncertainty

Key Decisions
What are the main sources of uncertainty for the catchment for flood warnings ?
What are the main sources of uncertainty for the catchment for outlook statements ?
What additional sources of uncertainty arise from complicating factors ?
Key Steps
Use the catchment types defined earlier to assess sources of uncertainty
Assess the impact of complicating factors on uncertainty

If information is not already available from previous studies (e.g. performance measures) an initial assessment of the main sources of uncertainty can be performed using the catchment classification scheme mentioned above. This identifies which inputs and other factors affect the uncertainty in flood forecasting model outputs.

The appendices describe the methodology which is used to link catchment type to uncertainty, and the influence of any complicating factors, such as from reservoirs and control structures.

2.2.4 Types of Models

Key Decisions
What methods are available for the types of models being used ?
What types of data assimilation routines are an option ?
What potential run time issues have been identified ?
Key Steps

Assess the uncertainties for the types of model in the integrated catchment model
Consider complicating factors such as reservoirs and snowmelt models
Assess options based on the model types in the integrated catchment model

When considering which sources of uncertainty to consider, the individual types of models used in the existing integrated catchment model are an important factor to consider, and can include some or all of the following general types of model:

- Rainfall-runoff models (gauged and ungauged catchments)
- Flow routing models (hydrological and/or hydrodynamic)

The appendices provide an overview of which data assimilation and conditioning methods are available for both generic types of model (e.g. rainfall-runoff) and specific model types (e.g. HBV-96, PDM, SAC-SMA etc).

2.2.5 Operational Requirements

Key Decisions
Is a purely qualitative approach sufficient for generating ensembles ?
Is data assimilation desirable or essential ?
Is conditioning of forecast outputs required ?
Key Steps
For each Forecasting Point, assess the operational requirement
Decide, at a catchment level, which is the most important Forecasting Point to consider

The operational requirement describes the intended operational use of probabilistic forecasts in decision-making for flood warning and operational and emergency response.

One key question to consider is whether a quantitative estimate of probability is required for input to the decision-making process, or whether a more visual, qualitative approach is envisaged.

The degree of data assimilation and conditioning depends on this decision, as described in the appendices.

2.2.6 Runtimes

Key Decisions
Are there run time issues for the candidate uncertainty estimation methods ?
What are the options for reducing run times ?
Key Steps
Assess run times for the existing deterministic model
Estimate run times required for ensemble forecasts
Evaluate run time reduction options

Some probabilistic forecasting techniques can significantly increase the run times required for each model run, so that forecasts cannot be obtained within an operationally useful time.

This is typically an issue for models which include a hydrodynamic component, and may be an issue for other types of models if large numbers of ensemble runs are envisaged.

The appendices describe the main issues to consider for different types of models, and the main options for reducing run times.

2.2.7 Performance Measures

Key Decisions

Are suitable performance measures already available to evaluate sources of uncertainty ?

Which sources of uncertainty (and locations) are identified ?

Key Steps

Review previous studies, and ongoing operational assessments, of model performance

Assess what this shows about the performance of individual models, data sources, forcing inputs etc.

One of the most important steps in the forecasting process is forecast verification, since this determines how much a forecaster can trust/rely on the forecasting system when issuing a forecast.

For an integrated catchment model, the performance measures at each forecasting point should provide a good indication of the locations in the catchment where the performance of individual components of the model is poor (and hence uncertainty is high), and also of some of the underlying causes of that uncertainty.

However, it should be noted that usually only an estimate of overall uncertainty is provided, not of individual sources unless, for example, systematic testing has been performed of the effects of adding or removing model components or data streams, and of the sensitivity of outputs to key model parameters and data inputs.

If this information is already available, this can supplement or replace some of the previous steps in the analysis. More information on performance measures is provided in Appendix A.

2.3 Summary of Finding

The uncertainty framework has been developed from a combination of a review of available techniques, findings from previous or related studies, and the expertise of the project team. Trials were performed for the coastal forecasting system in the Netherlands and the river forecasting system for the Rhine and Meuse (FEWS Rivieren).

3. FC2015 2009.06 & the Framework

3.1 Introduction

The project 2009.06 under the Flood Control program contains 11 sub-projects carried out by and together with the different partners. These projects are

- 2009.06.01 Development Uncertainty Framework (Ontwikkeling van raamwerk voor kwantificeren en reduceren van onzekerheden;
- 2009.06.02 Grid-based estimation precipitation amounts from combining Radar and Raingauges (Optimale schatting gevallen hoeveelheid neerslag per roostercel van 1x1 km, inclusief betrouwbaarheidsmarges);
- 2009.06.02 Operationalising spatial dependencies and uncertainty in precipitation estimates (Operationaliseren van ruimtelijke afhankelijkheid en onzekerheid in neerslagmetingen);
- 2009.06.04 Bayesian Model Calibration of Rainfall-Runoff models (Bayesiaans afregelmechanisme voor het identificeren en kwantificeren van onzekerheden in neerslag-afvoermodellen);
- 2009.06.05 Uncertainty in Onzekere afvoerverdeling bij extreem hoogwater;
- 2009.06.06 Uncertainty in flood scenarios and consequences (Onzekerheden in overstromingsscenario’s en gevolgen) ;
- 2009.06.07 Uncertainty in Evacuation Modelling (Onzekerheid bij evacuaties);
- 2009.06.08 Postprocessing of Hydrological Ensemble Forecasts (Postprocessing van probabilistische informatie over overstromingsrisico’s);
- 2009.06.09 Operational River and Coastal Water Level Forecast using Bayesian Model Averaging (Bepalen conditionele kansen van stormvloedwaarschuwingen);
- 2009.06.10 Operational uncertainty analysis of dike failure – piping and macro-instability (Operationele faalkansanalyse dijkringgebied tijdens hoogwatergolf);
- 2009.06.11 Real-time data assimilation using spatially distributed hydrological models (Real-time data assimilatie in ruimtelijk gedistribueerde hydrologische modellen).

These projects may be categorized under the framework to see where possible omission in the framework are and to determine to which strategy (operational or long-term) the different projects contribute. This is shown below:

Development Framework	2009.06.01
Operational Strategy	
A-Condition Meteorological Forcing:	-
B-Conditioning Initial State:	2009.06.11
C-Condition Model Parameters:	-
D-Conditioning Forecasts :	2009.06.08 & 2009.06.09 (in DFCR)
Long-term Strategy	
Model:	2009.06.04, 2009.06.05, 2009.06.10
Measurements:	2009.06.02&2009.06.03
Flood Scenarios and Consequences and Uncertainty	
Flood Scenarios:	2009.06.06
Evacuations:	2009.06.07

As can be seen above, uncertainty in Flood Scenarios and Flood Consequences has not been considered when constructing the framework, although operational Flood Scenarios and Flood Consequences forecasts might fall under C or D. This has been done because of time and budget constraints and may be considered in future Flood Control 2015 projects.

3.2 Conclusions

This report presents a first version of an overarching framework for assessing uncertainties in fluvial and coastal forecasting in a risk-based manner with the aim that it is robust enough to be considered for use in an operational environment. Two strategies to reduce and quantify uncertainties have been indentified: an operational strategy and a a long-term strategy. The operational strategy is focused an reduction and quantification of uncertainties in real-time. The long-term strategy is focused on structural improvements (models and measurements) that help to reduce (and quantify) uncertainties.

This framework is supplemented by two case studies which will demonstrate how the framework can be used. The framework needs further development and application (for instance for/to local water boards). In future versions uncertainty in prediction of flood scenarios and consequences should be considered.

4. Literature

BEVEN, K J (2009) Environmental Modelling: An Uncertain Future?, Routledge: London

BROERSEN, P.M.T, WEERTS, A.H. (2005) Automatic error correction of rainfall–runoff models in flood forecasting systems. In: Proceedings IEEE/IMTC Conference, Ottawa, Canada, 1531–1536.

BUIZZA, R., HOUTEKAMER, P.L., TOTH, Z., PELLERIN, G., WEI, M. AND ZHU, Y. (2005) A comparison of ECWMF, MSC, and NCEP global ensembles prediction systems, Monthly Weather Review, 133, 1076-1097.

CHEN, Y, SENE, K, HEARN, K. (2005) Converting 'Section 105' or SFRM Hydrodynamic River Models For Real Time Flood Forecasting Applications, 40th DEFRA Flood and Coastal Management Conference, The University of York, 5th-7th July 2005, Paper 6b-3.

CLARK M.P., D. E. RUPP, R.A. WOODS, X. ZHENG, R. P. IBBITT, A. G. SLATER, J. SCHMIDT, M.J. UDDSTROM (2008), Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Advances in Water Resources 31,1309–1324.

CLOKE H L, PAPPENBERGER F (2009) Ensemble flood forecasting: A review. J. Hydrology, 375(3-4), 613-626

ENVIRONMENT AGENCY (2002) Real Time Modelling Guidelines. R&D project WSC013/5

ENVIRONMENT AGENCY (2009) Risk-based Probabilistic Fluvial Flood Forecasting for Integrated Catchment Models. Phase 1 report. Science Report SR-SC080030

GEORGAKAKOS, K. P., SEO, D.-J., GUPTA, H., SCHAAKE J., AND BUTTS, M. B. (2004) Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, 298, 222-241.

HEEMINK, A. W, K. BOLDING and M. VERLAAN, Storm Surge Forecasting using Kalman Filtering, Journal of Meteorological Society of Japan, Special issue,1997.

JOLLIFFE I T, STEPHENSON D B (2003) Forecast Verification: A Practitioner's Guide in Atmospheric Science, John Wiley and Sons, Chichester

JONES, A.E., JONES, D.A., MOORE, R.J. (2003) Development of Rainfall Forecast Performance Monitoring Criteria. Phase 1: Development of Methodology and Algorithms. Report to the Environment Agency and the Met Office, CEH Wallingford, 291pp.

JONES, A.E., JONES, D.A., MOORE, R.J. AND ROBSON, A.J., (2004) Heavy Rainfall Warning Assessment Tool User Guide. Report to the Environment Agency and the Met Office, CEH Wallingford, 65pp.

KRZYSZTOFOWICZ, R. (2004), Bayesian Processor of Output: A new technique for probabilistic weather forecasting, Paper 4.2, 17th Conference on Probability and Statistics in the Atmospheric Sciences, 84th AMS Annual meeting, seatlle, 11-15 Jnauary 2004, AMS. URL: http://ams.confex.com/ams/84Annual/17PROBSTA/abstracts/69608.html.

LAIO, F., TAMEA, S (2007) Verification tools for probabilistic forecasts of continuous hydrological variables, Hydrol. Earth Syst. Sci., 11, 1267-1277

LEES, M. J., YOUNG, P. C., FERGUSON, S., BEVEN, K. J., BURNS, J. (1994). An adaptive flood warning scheme for the River Nith at Dumfries. In W.R., W. & Watts, J. (Eds.) 2nd International Conference on River Flood Hydraulics. J. Wiley & Sons.

LETTENMAIER, D P, WOOD, E F (1993) Hydrologic forecasting. In Handbook of Hydrology, Maidment R.D. (Ed.) 26.1-26.30, McGraw-Hill

MADSEN H., BUTTS M. B., KHU S. T., LIONG S.Y. (2000) Data assimilation in rainfall–runoff forecasting. 4th International Conference on Hydroinformatics, July 2000, Cedar Rapids, Iowa, USA.

MOLL J.R. (1986) Short range Flood Forecasting on the River Rhine. In: River Flow Modelling and Forecasting, Water Science and Technology Library, D. Reidel Publishing Company,Kluwer, Dordrecht, Eds Kraijenhoff and Moll

OBLED C, BONTRONA G, GARCON R (2002) Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach. Atmospheric Research, 63(3-4), 303-324

PAPPENBERGER, F., BEVEN, K.J., FRODSHAM, K., ROMANOVICZ, R. AND MATGEN, P. (2007). Grasping the unavoidable subjectivity in calibration of flood inundation models: a vulnerability weighted approach. Journal of Hydrology, 333, 275–287.

REFSGAARD, J.C (1997) Validation and intercomparison of different updating procedures for real-time forecasting. Nordic Hydrology, 28, 65–84.

RENNER M, WERNER M G F, RADEMACHER S, SPROKKEREEF E (2009) Verification of ensemble flow forecasts for the River Rhine. J. Hydrology (in press)

ROULIN E, VANNITSEM S (2005) Skill of Medium-Range Hydrological Ensemble Predictions. Journal of Hydrometeorology 6(5): 729

SCHWEPPE (1973) Uncertain dynamic systems, Prentice-Hall, 563pp.

SEO, D.J., HERR, H.D. AND SCHAAKE, J.C., 2006. A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. Hydrology and Earth System Sciences Discussions, 3, 1987–2035.

SEO, DJ., L. CALINA, R. CORBY, T. HOWIESON (2009) Automatic state updating for operational streamflow forecasting via variational data assimilation, Journal of Hydrology, Volume: 367 Issue: 3-4 Pages: 255-275.

SERBAN P, ASKEW A J (1991) Hydrological forecasting and updating procedures In: Hydrology for the Water Management of Large River Basins, IAHS Publication No. 201

UPPALA et al. (2006) The ERA-40 re-analysis. Q.J.R.Meteorol., 131(612), 2961-3012.

VERLAAN, M., A. ZIJDERVELD, H. DE VRIES and J. KROOS (2005), Operational storm surge forecasting in the Netherlands: developments in the last decade, Phil. Trans. R. Soc. A.,363 no 1831, 1441-1453.

WEERTS, A.H., 2008. Hindcast of water levels for the Rhine branches and the Meuse for 2006 & 2007, Deltares, Research Report, Q4234.

WEERTS, A.H., D MEIßNER, S RADEMACHER, 2008. Input Data Rainfall-Runoff Model Operational Systems FEWS-NL & FEWS-DE, Deltares & BfG, Research Report, Q4234.

WEERTS, A.H. AND EL SERAFY, G.Y.H., 2006. Particle filtering and Ensemble Kalman filtering for state updating with conceptual rainfall–runoff models. Water Resources Research, 42(9), W09403, doi:10.1029/2005WR004093

WEERTS, A.H., EL SERAFY, G.Y.H., HUMMEL, S., DHONDIA, J. AND GERRITSEN, H., 2009. Application of generic data assimilation tools (DATools) for flood forecasting purposes. Computers and Geoscience (accepted).

WEERTS A H, SEO D J, WERNER M, SCHAAKE J (2010). Operational hydrological ensemble forecasting. To appear in Applied uncertainty analysis for flood risk management edited by K Beven and J Hall

WERNER M, SELF K (2005) Measuring Performance in Flood Forecasting: A question of model reliability or of forecast reliability. ACTIF International conference on innovation advances and implementation of flood forecasting technology, 17 to 19 October 2005, Tromsø, Norway

WERNER M, REGGIANI P, WEERTS A H (2010) Reducing and quantifying uncertainty in operational forecasting: Examples from DELFT-FEWS forecasting platform/system. To appear in Applied uncertainty analysis for flood risk management edited by K Beven and J Hall

WILKS D.S. (2006) Statistical Methods in the Atmospheric Sciences. 2d ed. International Geophysics Series, Vol. 91, Academic Press, 627 pp

WILKS D. S., HAMILL T. (2007) Comparison of Ensemble-MOS methods using GFS reforecasts, Monthly Weather Review, 135, 2379-2390.

YOUNG, P.C., 2009. Real-time updating in flood forecasting and warning. In: I. Cluckie, G. Pender, C. Thorne and H. Faulkner (eds), The Flood Management Handbook (in press). Wiley-Blackwell (available from author).

A. Uncertainty Framework

This appendix presents the technical background to development of the framework which is described in Chapter 2 of the main report, and to the methods which are recommended for selecting an appropriate uncertainty estimation approach.

The main section headings follow the key factors introduced in Chapter 2 as follows:

- Appendix A.1 - Level of Risk
- Appendix A.2 - Lead-time Requirements
- Appendix A.3 - Main Sources of Uncertainty
- Appendix A.4 - Types of Models
- Appendix A.5 - Operational Requirements
- Appendix A.6 – Run Times
- Appendix A.7 - Performance measures

A.1 Level of Risk

Many organisations are increasingly adopting a risk-based approach to decision-making, both for long-term planning and investment, and for shorter term operational decisions. Risk is often defined as the combination of probability and consequence, sometimes also taking account of factors such as vulnerability and exposure. One potential benefit of probabilistic forecasts is the ability to tailor flood warnings to the risk profile of the recipient (risk adverse, risk neutral etc), and the level of risk at which decisions are taken (e.g. to evacuate a town, raise a temporary barrier).

For the purposes of this framework, a risk-based approach offers the possibility of using the level of risk as a guide to the choice of an appropriate uncertainty estimation technique. For example, if the risk is low, less data hungry and computationally intensive approaches might be used, such as visualisation of outputs. However, if the risk is high, the decision might be made to use a state-of-the-art approach, investing in additional computer processing power and a hindcasting exercise for meteorological ensembles.

Many different methods are used for calculating risk, ranging from look-up tables or charts to numerical assessments. Methods need to be tailored to each situation; in particular taking account of the purpose of issuing a flood forecast, and the likely impacts if that forecast is incorrect. For example, in flood warning applications, the term ‘consequences’ can include the number of properties flooded, the economic damages, the number of lives lost, business disruption, and a range of intangible factors, such as the impacts on health and stress, and the loss of memorabilia and other personal items.

When using an integrated catchment model, there is also the issue of whether risk should be assessed individually for each forecasting point in a catchment, or aggregated at a catchment level. If the model does not cover all areas at risk in a catchment, then that factor also needs to be considered.

Since methods for addressing risk are so organisation-dependent, it is probably best to continue the discussion through the use of examples, as illustrated in Box A.1.

Box A.1 Risk-based Approaches used in the Netherlands

The Dutch flood protection standards for coastal defences are among the most stringent in the world. These standards were set after the 1953 flood, which cost the lives of some 1800 people in the Netherlands and 300 in the UK. The disaster put flood hazard on the political agenda and gave cause for extensive flood protection works in Zeeland and Zuid-Holland. The flood protection standards were developed in the late 1950's and eventually led to the Flood Defences Act ("Wet op de Waterkering", WWk (1995)). This Act states that the primary sea defences in the Netherlands should be able to withstand the maximum hydraulic load that is expected to occur at a probability of 1/1250 to 10,000 per year, depending on the area at risk (see Figure A.1).

Figure A.1 Safety standards for primary flood defences in the Netherlands.

The Flood Defences Act also demands that the flood are monitored every five years (2001, 2006, 2011, etc.) for the required level of protection. This assessment is based on the Hydraulic Boundary Conditions (HBC) and the Safety Assessment Regulation (VTV: Voorschrift op Toetsen op Veiligheid). The HBC are derived every five years and approved by the Minister of Transport, Public Works and Water Management.

The Safety Assessment Regulation (VTV) describes three levels of assessment:

1. The basic assessment uses approximate models and rules of thumb. These assessment rules take conservative margins for simplifications or assumptions that were made to make the assessment feasible for non-experts. If a dike fails the basic assessment, a detailed assessment is usually the next step.

2. The detailed assessment generally uses the same models as in the basic assessment, but without making simplifications. This implies using more data and advanced numerical models. The assessment result can be positive, negative or undecided, in which case an advanced assessment is required.
3. The advanced assessment requires the involvement of experts on dike safety, who will perform specific research on the dike at stake. They will employ state-of-the-art models and all available knowledge to come to a final judgement.

Operational flood forecasting in the Netherlands is done separately for the coastal and fluvial systems. The Storm Surge Warning Service is responsible for sea level forecasting. If the forecast exceeds a critical level within the next 6 hours, the SVSD issues a warning to the dike and dam authorities and other bodies responsible for public safety. High tide levels are forecast by hydrodynamic models that are coupled to meteorological forecasts from KNMI. A Kalman filter is used to assimilate water level measurements along the British and Dutch coast. This reduces the uncertainty of the forecasts up to a 12-hour lead time (see Figure A.2).

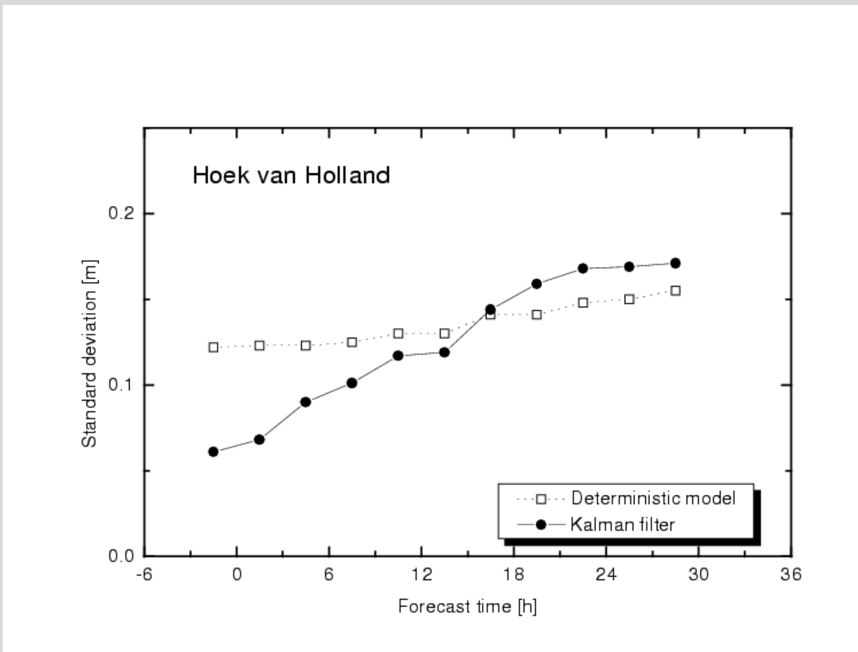


Figure A.2 Uncertainty (RMSE from observations) of the sea level forecasts at Hoek van Holland as a function of lead time.

Water-level forecasts for the rivers Rhine and Meuse in the Netherlands are the responsibility of the Centre for Water Management of Rijkswaterstaat. The Centre for Water Management and Deltares have developed in the past decade, a so-called Flood Early Warning System for the Rhine and Meuse Rivers, called FEWS NL. FEWS NL is an advanced combination of hydrological and hydraulic models with software for import, validation, interpolation and presentation of data. Every 30 minutes the system receives observed water levels from about 60 gauging stations in the Rhine basin. Every hour meteorological observations are downloaded from servers at the national Dutch (KNMI) and German (DWD) weather services of more than 600 stations in the basin of Rhine and Meuse. The system uses output from four numerical weather models at KNMI, DWD and the European Centre for Medium Range Weather Forecasts (ECMWF).

The large investments made in the coastal and fluvial forecasting systems are justified by the required accuracy for adequate flood warnings in both regions. Although the probability of flooding is very low, the consequences of a flood in a densely populated area such as the Netherlands are dramatic. Any measures that can be taken to reduce the losses depend on an accurate forecasting system. At the same time a false alarm is very costly. The cost for the 1995 evacuation of 240,000 people from the threatened areas along the River Rhine and Meuse was estimated at 1 billion euros.

A.2 Lead-time requirements

The forecasting lead-time depends primarily on the lead-time requirement for flood warning, and may extend from as little as 1-2 hours to several days ahead. In the latter case, the shorter lead-time forecasts may be used in issuing the actual operational warning, while forecasts at the longer lead-time are used mainly as guidance in moving to a flood alert status, rather than to guide the issuing of a flood warning. For example, for a large-scale flood event in a UK situation, given a week's lead-time, the sequence of information actions might be (Golding 2009):

- 3-5 days ahead: issue 'advisory' or 'period of heightened risk'; engage in awareness raising activities through the media, mobilize support organisations for the vulnerable; initiate 'participatory' information sharing by local flood response organisations
- 1-2 days ahead: issue 'early warning' or 'watch'; activate mitigation measures for flood minimization and protection of critical infrastructure; provide active support to vulnerable groups; move to a consultative engagement with those in the most vulnerable areas
- Hours ahead: issue 'flood warning'; activate emergency response; evacuate most vulnerable groups if appropriate; provide 'prescriptive' advice to individuals

Although the lead-times may differ between organisations, it is convenient in this framework to distinguish between two types of lead-time requirement:

- Flood Warnings – the typical lead-time at which emergency response actions need to be taken
- Outlook Statements – the lead-times at which forecasts are used to provide a general 'outlook', 'advisory', 'early warning' or 'watch'

In the absence of existing information on the key sources of uncertainty, such as from performance measures (see later), the lead-time requirement is a key factor in deciding on the most appropriate uncertainty estimation technique(s) to use. A comparison with the catchment response time then indicates whether sufficient lead-time can be obtained using catchment observations (e.g. river flows, raingauges) or rainfall forecasts are required as inputs.

This balance between lead-time requirements and catchment response time can be formalised by using a development of a simple classification scheme for flood forecasting originally developed by Lettenmaier and Wood (1993).

Considering first a single forecasting point in a catchment, the adapted classification scheme compares the desired warning time ($T_{warning}$) to the total response time (T_{total}) at the location for which the forecast is to be provided. This response time is further subdivided into the hydraulic response time (travel time through main river, T_{river}) and the hydrological response time (which is less than the response time of the catchment, $T_{catchment}$).

An additional lead-time (T_{surge}) is also applicable for coastal forecasting situations (although coastal forecasting is outside the scope of the present study). This division is somewhat arbitrary but generally the river channel is considered to be the main river (system), whilst the hydrological response is the response of sub-catchments before water flows into the main river system.

The situations in Table A.1 are defined, and these general categories are illustrated in Figure A.3, and indicate the types of forcing inputs which may be required at each forecasting point in the catchment. For example, for Type 1 situations, rainfall forecasts are essential and, if conditioning of outputs is used, this would require an archive of forecast values (perhaps obtained using a hindcasting exercise). For catchments with multiple forecasting points, then each point needs to be considered in turn and the forcing inputs assessed.

Table A.1 Links between lead-time requirements and catchment response (adapted from Lettenmaier and Wood 1993)

Type	Catchment	Criterion	Description and key forcing inputs for flood warnings
1	very fast responding basins	$T_{warning}$>>T_{total}	The desired lead-time is such that the warning or outlook must be issued on the basis of water that has not yet fallen as rain. In this case a rainfall forecast is the only means to provide a timely warning when using a flood forecasting model
2	small to medium basins	$T_{warning}$ <T_{total} & $T_{catchment}$>>T_{river}.	The warning or outlook will be issued on the basis of water that is already in the catchment and is mainly determined by the hydrological travel time. This may be the case for point I in Figure A.2.
3	medium size basins	$T_{warning}$<T_{total} & $T_{catchment}$~T_{river}	The warning or outlook will be issued on the basis of water that is already in the catchment and river and the response time is determined by the hydrological response time and the hydraulic response time. This may be the case for forecast point IV in Figure A.2.
4	large river basin	$T_{warning}$<T_{river} or $T_{catchment}$<<T_{river}	The warning or outlook will be issued on the basis of water that is already in the main channel; or the hydrological response time is insignificant compared to the hydraulic response time. This may be the case for the forecast point VII in Figure A.2, assuming catchments E and F have only minor contributions.
5	coastal / tidal zone	$T_{warning}$>>T_{surge}	The desired lead-time is such that the warning or outlook may be issued on the basis of wind conditions that have not yet occurred. In this case wind and pressure forecasts are necessary for a timely warning.

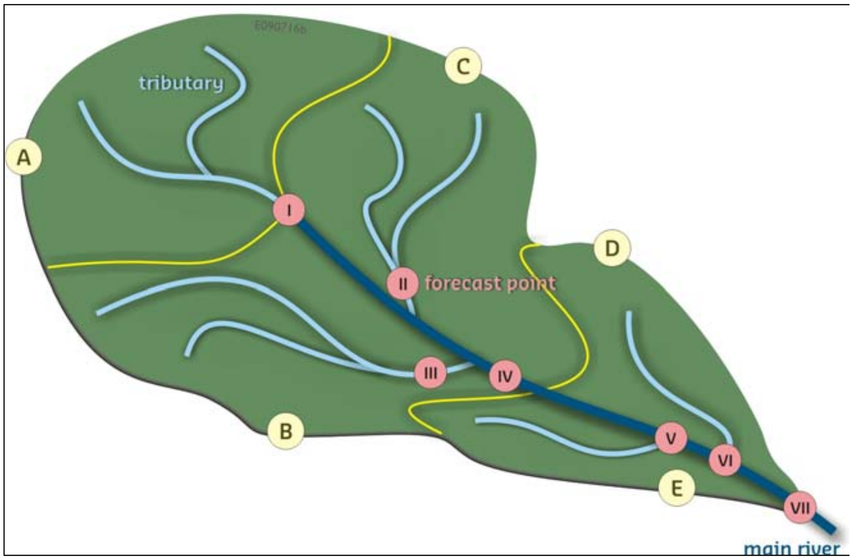


Figure A.3 Schematic layout of a catchment, including the main river, tributaries and catchments (adapted from Lettenmaier and Wood 1993)

In estimating the actual time available, an allowance also needs to be made for the various time delays in the decision-making and warning process, which can include (Environment Agency 2002):

- The time taken for information to be received by telemetry
- The time taken for a routine real-time model run
- The time taken for flood forecasting and warning staff to decide to act upon a forecast of levels exceeding a Flood Warning trigger level (e.g. whilst performing 'what if' runs)
- The time taken for all properties to be warned (e.g. via an automated dialling system.

A.3 Main sources of Uncertainty

If information is not already available from previous studies (e.g. performance measures) an initial assessment of the main sources of uncertainty can be performed using the catchments types (Types 1 to 5) defined in Table A.1. Table A.2 illustrates some key sources of uncertainty which are typical for flood warnings and outlook statements.

In the table, the focus is on the primary sources of errors, rather than derived variables such as antecedent conditions (which are typically obtained from an initial state, and observed areal mean rainfall, potential evaporation and flows in the historical mode of operation).

Note that, for very fast responding catchments, with current meteorological forecasting performance, rainfall forecasts are likely to be the most significant source of uncertainty of those listed. Also, as noted earlier, the key sources of uncertainty are likely to be different for flood warnings and outlook statements.

Table A.2 *Dominating uncertainties for each forecasting situation*

Type	Catchment	Source
1	very fast responding basins	**Initialisation Errors**: -Past Areal Mean Rainfall -Potential Evaporation **Modelling Errors**: -Rainfall-Runoff Model Parameters -Rainfall-Runoff Model Structure **Forcing Errors:** -Rainfall Forecasts
2	small size basins	**Initialization Errors**: - Past Area Mean Rainfall -Potential Evaporation **Modelling Errors**: -Rainfall-Runoff Model Parameters -Rainfall-Runoff Model Structure **Forcing Errors:** -Rainfall Forecasts (for longer lead-times/outlook statements)
3	medium size basins	A mixture/combination of 2&4
4	large river basins	**Initialisation Errors**: -High Flow Ratings -Ungauged Lateral Inflows -Tidal Boundary **Modelling Errors**: -Hydraulic/Routing Model Parameters -River Channel/Floodplain Survey -River Control Structures **Forcing Errors:** **-**High Flow Ratings (forecast inflows) -Forecast Tidal Boundary -Forecast Lateral and other Inflows (for longer lead-times/outlook statements) -Rainfall Forecast (for even longer lead-times/outlook statements)
5	coastal/tidal zone	**Initialisation Errors**: -Water levels in the coastal zone **Modelling Errors**: -Bathymetry -Model domain/resolution **Forcing Errors:** -High Flow Ratings (forecast inflows/levels) -Boundary Conditions -Wind and Pressure Forecast

A number of catchment-related complicating factors can also influence the choice of an uncertainty estimation approach, and can include some or all of the following items:

- Permeable catchments – runoff response may normally be minor, but major flows generated once saturated conditions are reached
- Groundwater influences – runoff response may vary with groundwater levels, and groundwater flooding may need to be considered
- Urban influences – a number of factors may influence response (e.g. drainage networks, flood detention basins)
- Snowmelt – additional runoff may occur from melting snow, requiring a snowmelt modelling component
- Embanked floodplains – water may be lost to the river network permanently, or return some time later via drainage paths
- Reservoirs, lakes and wetlands – flow modifications/attenuation may occur due to storage effects and/or gate operations
- Flow Control Structures – flows may be influenced by structure operations at gates, barriers etc.
- Off-line storage, abstractions, discharges and diversions – influences may occur from washland operations, pumps, flood relief channels etc.
- Event specific problems – issues may occur with channel blockage due to debris, defence breaches, dam breaches etc.

If any of these components are present, the key sources of uncertainty may also need to be assessed for these model components as illustrated in Table A.3. Again, the focus is on primary sources of error, rather than observed or calculated values such as reservoir levels:

Table A.3 Some typical additional sources of uncertainty from complicating factors

Complicating Factor	Initialisation Errors	Modelling Errors	Forcing Errors
Permeable catchments		Groundwater/soil moisture store parameters Model Structure	
Groundwater influences		Groundwater store parameters Model Structure	
Urban influences		Surface Runoff Parameters	
Snowmelt	Past Precipitation	Snow Store Parameters; Model Structure	Precipitation Forecasts
Embanked floodplains	Past Floodplain Drainage	Flood Defence and Embankment Survey/Flow Paths; Model Structure	Forecast Floodplain Drainage
Reservoirs, lakes and wetlands	Past Ungauged Inflows; Past Open Water Evaporation	Stage-Volume Survey; Release Rules; Model Structure	Forecast Ungauged Inflows
Flow Control Structures	Past Structure Settings (gates etc)	Control/Logical Rules; Model Structure	Forecast Structure Settings (gates etc)
Off-line storage, abstractions, discharges and diversions	Past Abstractions, Discharges and Diversions	Outflow Relationships; Operating Rules; Model Structure	

Event specific problems		Representation of Blockage, Breach etc; Model Structure	

A.4 Types of Model

When considering which sources of uncertainty to consider, the types of models used in the existing integrated catchment model are an important factor to consider, and can include some or all of the following general types of model:

- Rainfall-runoff models (gauged and ungauged catchments)
- Flow routing models (hydrological and/or hydrodynamic

The sources of uncertainty are generally specific to each type of model used by an organisation.

A.5 Operational Requirements

The operational requirement describes the intended operational use of probabilistic forecasts in decision-making for flood warning and operational and emergency response.

One key question to consider is whether a quantitative estimate of probability is required for input to the decision-making process, or whether a more visual, qualitative approach is envisaged.

For example, the following techniques have been used in meteorology and application areas for interpretation of the outputs from ensemble forecasts:

A. Visualisation – 'eyeball' assessments of the spread of ensemble members with forecast lead-time, and relative to threshold values, and of other factors such as the clustering of ensembles

B. Persistence-based approaches – which compare the number of threshold exceedances between successive model runs

C. Threshold-frequency approaches – calibration of thresholds based on the historical model performance, obtained over a calibration period (e.g. based on flow return periods)

D. Physical-threshold approaches – as for threshold-frequency approaches but using actual threshold values in decision-making (e.g. flood defence levels)

E. Cost-loss approaches – assessment of appropriate actions based on consideration of the economic value or utility of forecasts, and the optimum probability thresholds for taking action

F. Bayesian Uncertainty Processors (whole-system versions) – similar to cost-loss approaches, considering the predictive uncertainty taking account of all information available up to the time of the forecast, and including economic and subjective views of flood warning decision criteria

The probabilistic content for Types A and B might be classed as qualitative, Type C as indicative/qualitative, and Types D to F as quantitative.

Although it is difficult to generalise, the requirement for the probabilistic content of forecasts increases moving from Type 1 to Type 6. Also, for any given type, the data requirements for low probability-high impact events are higher, requiring either long runs of historical data to calibrate methods, and/or stochastic or other simulation of data to extend record lengths.

A.6 Run Times

Some probabilistic forecasting techniques can significantly increase the run times required for each model run, so that forecasts cannot be obtained within an operationally useful time. This is typically an issue for models which include a hydrodynamic component, and may be an issue for other types of models if large numbers of ensemble runs are envisaged. Some options for reducing run times include:

- Computational Improvements – e.g. parallel processing, faster processors
- Model Configuration changes – e.g. nested models, model simplification or rationalisation
- Statistical Approaches – e.g. sampling or grouping of ensembles
- Model Emulators – e.g. simpler models to emulate the behaviour of more complex models

For any given integrated catchment model, the options which can be used depend on a number of factors, including the type of model, existing run times, and the catchment and local response (e.g. influence of structures). Also, some options, such as faster processors, typically require an investment at organisational level, so are not always an option when considering an individual catchment model.

For hydrodynamic models, the main options for reducing run times include rationalisation and simplification, parallel processing, and the use of emulators. Some studies have shown (e.g. Chen et al. 2005) that even for complex models run time reductions of one or two orders of magnitude (10-100) are often possible without sacrificing model performance at forecasting points.

Emulators also provide an attractive option, particularly if real-time inundation mapping is envisaged (e.g. Young et al. 2009). The run time requirement for this approach is minimal.

A.7 Performance Measures

Performance measures may be also be used as a guide to the main sources of uncertainty and may be available for both deterministic and probabilistic forecasts. For example, for probabilistic forecasts, the methods which are used to quantify the uncertainties include the reliability, skill and resolution of forecasts. Many of these techniques have been developed within the atmospheric sciences (e.g. Wilks 2006; Jolliffe and Stephenson 2003), but are often equally applicable to other disciplines, such as the hydrological sciences.

A distinction can also be made between real-time and historical verification, where historical verification is used to assess the performance of the forecasting system; for example, by looking at different attributes of the forecasts, such as reliability, skill, resolution, discrimination, etc., to diagnose the performance of the forecasting system so that cost-effective improvements may be made. Of course it is also possible to perform a trend analysis to assess improvement in forecast quality over time.

Methods for verification of deterministic forecasts are well established (e.g. Jones et al. 2003, 2004; Werner and Self 2005; Wilks 2006), and provide clear insights into value and skill of predictions at different lead-times, giving valuable information to the forecaster in interpreting forecast products.

Recently, more attention has been paid to the verification of ensemble forecasts (Roulin and Vannitsem, 2005; Laio and Tamea 2008, Renner et al., 2009; CEH Wallingford 2009). However note that the records of ensemble forecasts are often limited and that therefore verification statistics are difficult to determine without recourse to a hindcasting or reforecasting exercise. This is in principle straightforward for a river forecasting model, but can be a considerable undertaking for a meteorological forecasting model (e.g. Uppala et al. 2006).

To assess the quality of probabilistic hydrological forecasts, several verification techniques can be applied. Methods typically assess the reliability of forecasts in representing historical probability distributions (median, spread, moments etc.), the long-term forecasting success relative to thresholds, such as skill scores, and the information content of forecasts, such as the proximity to a yes/no response (e.g. the sharpness). As with deterministic forecasts, measures can be presented for different forecast lead-times, with and without the use of data assimilation.

For example, reliability can be assessed through reliability diagrams or attribute diagrams. These diagrams measure the agreement between predicted probabilities and observed frequencies. If the forecast is reliable then, over the long-term, whenever the forecast probability of an event occurring is P, that event should occur a fraction P of the time.

Skill measures for assessing ensemble forecasts include the Brier Score, which measures the mean squared error in the probability space, and the Brier Skill Score (BSS) measures skill relative to a reference forecast (usually the climatological or naïve forecast).

The Ranked Probability Score (RPS) is another way of determining the accuracy of the probabilistic forecast. RPS measures the squared difference in probability space when there are multiple categories (when there are only two categories, the RPS is equal to the BS). As with the Brier Skill Score, the Ranked Probability Skill score measures skill relative to a reference forecast, and applies when there is a discrete number of categories, but can be extended to continuous categories as the Continuous Ranked Probability Score (CRPS). CRPS is particularly attractive in that it does not depend on the particular choice of thresholds and that it allows comparative verification with single-value forecasts, for which CRPS reduces to the absolute mean error.

The Relative Operation Characteristic (ROC) is a measure to assess the ability of the forecast to discriminate between events and non-events. The ROC curve plots the hit rate (POD) against the false alarm rate (POFD), which differs from the False Alarm Ratio (FAR). The curve is created using increasing probability thresholds to make the yes/no decision.

As with performance measures for deterministic forecasts, probabilistic measures provide information on different aspects of forecast performance, and are often used in combination to assess the performance of a forecasting model.

B Applying the Framework for the SVSD

This appendix presents the main results from application of the uncertainty framework to the forecasting models of the Storm Surge Warning Service (Dutch: Stormvloed-waarschuwingsdienst, SVSD). The descriptions are presented as follows:

- Appendix B.1 – Introduction
- Appendix B.2 – Application of the Uncertainty Framework

B.1 Introduction

B1.1 SVSD

The Storm Surge Warning Service is responsible for timely notification of the dike and dam authorities and other bodies responsible for public safety in case of a storm surge hazard. During the storm season the SVSD keeps a meticulous watch on meteorological developments and coastal tide conditions, particularly if the wind direction is between south-westerly and northerly. High tide levels are forecast by hydrodynamic models that are coupled to meteorological forecasts from the Royal Netherlands Meteorological Institute (KNMI). If the forecast exceeds a critical level, the SVSD issues advance warnings to the relevant authorities.

There are three critical water levels. The 'pre- warning' level will call for dam authorities to take very limited precautions. The second ‘warning’ level will trigger some further measures. The third and highest ‘alarm’ level will call for drastic precautions, such as continuous monitoring of the dike status. The pre-warning, warning or alarm is issued 6 hours in advance of the expected time of exceedance. Currently, the SVSD is investigating if this can be increased to 12 hours in advance.

Because the timing of high tides varies from one place to another and because a gale will seldom affect the whole coastline with equal force, the coastal region is divided into sectors. In each sector, there is a reference monitoring station and each sector has different critical water levels as shown in Table B.1.

Table B.1 Sectors and critical water levels

Sector	Schelde	West-Holland	Dordrecht	Den Helder	Harlingen	Delfzijl
Reference station	Vlissingen	Hook of Holland	Dordrecht	Den Helder	Harlingen	Delfzijl
Pre-warning level	310	200				260
Warning level	330	220		190	270	300
Alarm level	370	280	250	260	330	380

The procedure for issuing warnings and alarms to the various sectors is as follows:

Every day the Hydro Meteo Centre in Hook of Holland (an auxiliary office of the KNMI) produces tidal rise forecasts. The SVSD is notified when the high tide at any reference station is expected to exceed the "information level" (which is as much as 40 to 50 cm below the warning level). This message is generally issued about ten hours before the water is likely actually to reach that level.

Based on the information supplied and experience, the SVSD officer on duty (a tidal hydrologist) - will decide whether or not it is expedient to fully staff the SVSD command centre. If the warning level is expected to be exceeded in one or more sectors, the SVSD will issue warnings and/ or alarms. Wherever possible, this will be done at least 6 hours in advance of high tide, so that the dike and dam authorities have time to prepare. These warnings will be issued to a number of bodies concerned with the safety of the coastal provinces, including:

- water boards and dike and dam authorities
- Rijkswaterstaat field services
- the provincial public works authorities
- the Ministry of the Interior (Fire Service and Disaster Response Department).

As soon as an alarm is issued, announcements are also broadcast on radio and TV news bulletins.

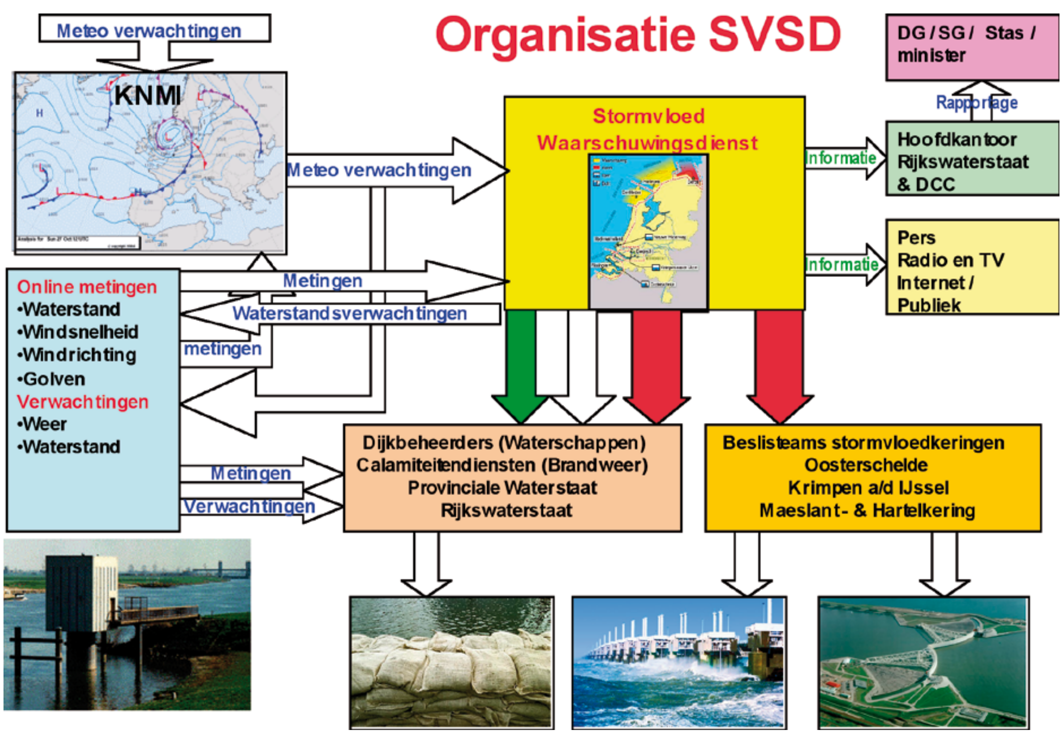


Figure B.1 *Information flow of the SVSD (courtesy of SVSD)*

When a dike watch is advised (alarm level), the dike authorities will call up personnel and recruit local volunteers to form dike teams. The central security station is staffed and sandbags are filled and loaded onto lorries. The sea dikes are patrolled by people carrying mobile phones. Fire brigades and police services are put on full alert. Army and navy commanders confine their personnel to barracks or to their ships. Dike cuttings - places where the dikes are crossed by roads and railway lines - are sealed off. Mayors and chief police officers are notified in the municipalities concerned. They take measures to ensure law and order. The public is informed by the national press agency (ANP), which spreads the information via radio and TV.

B.1.2 DCSM model

The WAQUA - DCSM98 storm surge model is a numerical model of the shallow water equations applied to the Northwest European Continental Shelf (see Figure G.2). DCSM was developed by Rijkswaterstaat, Deltares, and KNMI. The model is run several times a day, to calculate sea levels and the depth averaged currents on a grid with cells of approximately 8 by 8 km, using as input the wind and pressure forecasts from the KNMI limited area model HIRLAM.

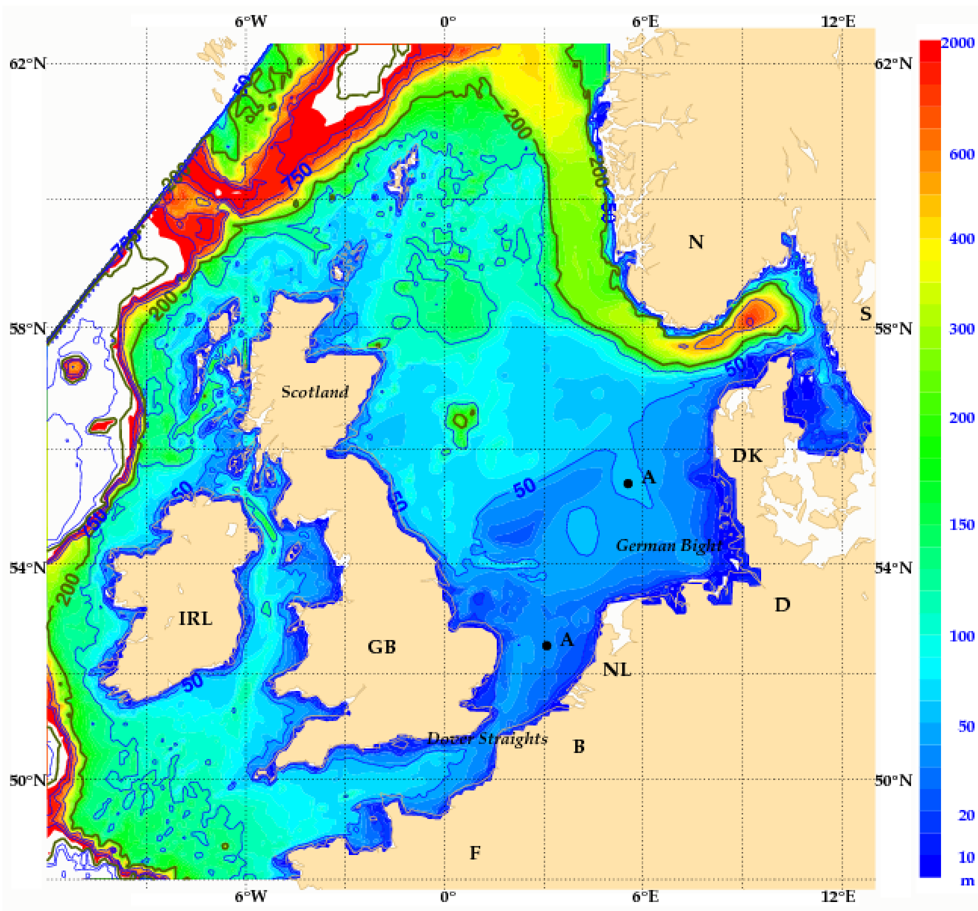


Figure B.2 Bathymetry and outline of the WAQUA-DCSM model (courtesy of KNMI)

Deterministic sea level forecasts for up to 48 hours ahead are produced four times per day. A Kalman filter is used for data assimilation of recent sea level observations. In particular for sea levels along the British and Dutch coasts this gives a significant improvement of the accuracy at short lead times. This benefits the decision support for the dynamic storm surge barriers in the Oosterschelde and the Rotterdam Waterway (Maeslantkering).

The standard deviation of the sea level forecasts for up to 48 hours is generally less than 15 cm along the Dutch and British coasts. For lead times of less than 12 hours, the Kalman filter typically brings this down to less than 10 cm. For extreme surges, the forecast accuracy is hard to evaluate, because these events are rare.

B.2 Applying the Uncertainty Framework

Following the uncertainty framework, the following key factors will be considered (see Chapter 2 of the main report):

- Level of Risk
- Lead Time Requirements
- Types of Model
- Main Sources of Uncertainty
- Operational Requirements
- Run Times
- Performance Measures

A worksheet has been developed to assist in this process and a completed version for the SVSD is attached to the end of this description. The following sections describe the analysis and decision making process which contributed to completing this worksheet. The coastal system is of Type 5.

B.2.1 Level of Risk

The protection standards for sea defences vary along the Dutch coast, between 1/10,000 per year for the coast of West-Holland to 1/2000 per year for Dordrecht. Table A.2 displays the protection standards and corresponding water levels for the reference locations to the SVSD sectors. The standards are well below the SVSD alarm levels, except for location Dordrecht (alarm level 2.5 m+NAP).

Table B.2 Protection standards for dike rings corresponding to the SVSD sectors

Sector	Protection standard	Water level
Schelde	1/4000	5.3 m+NAP
West-Holland	1/10,000	5.1 m+NAP
Dordrecht	1/2000	3.0 m+NAP
Den Helder	1/10,000	5.7 m+NAP
Harlingen	1/4000	4.9 m+NAP
Delfzijl	1/4000	6.0 m+NAP

The protection standards are relatively high, at least compared to fluvial standards and other countries. This is justified by the scale of the consequences should a flooding by a storm surge occur. The potential for enormous damage and a high number of casualties demands an extremely low flooding probability in order to keep the risk (probability x consequences) at a reasonable level.

B.2.2 Lead Time Requirements

The current lead time of the SVSD warnings and alarms before the expected critical level exceedance is 6 hours. The SVSD is currently investigating how this can be increased to, if possible, 12 hours. This lead time is sufficient for the tasks of the SVSD service, namely calling for operation of the storm surge barriers, organising a dike watch and sealing off dike cuttings. Fire brigades, police services and the army can also be alerted in time.

However, a warning lead time of 12 hours is by a long way not sufficient to organize an evacuation of the densely populated coastal regions. An organized evacuation of a coastal dike ring at risk requires multiple days. An accurate surge forecast several days ahead is currently not feasible, mainly due to the uncertainty of the meteorological forecast (wind and pressure) for longer lead times.

B.2.3 Main Sources of Uncertainty

For a Type 5 system, the main sources of uncertainty are listed in Table B.3, together with a summary of the likely importance for the Dutch coastal region.

The meteorological input for the storm surge model is the main source of uncertainty for the storm surge forecasts. The uncertainty originates from errors in the position and intensity of depressions. Small-scale intense phenomena also play an important role. Local wind gusts can affect the water level considerably at specific locations.

Predictability of such phenomena is intrinsically limited. Near the coast, the influence of topography on the wind over the sea should be taken into account.

Table B.3 Initial analysis of likely sources of uncertainty

Type	Source	Importance for DCSM
Initialisation Errors	Water levels	The initial water levels (and currents) over the entire model domain are an important source of uncertainty, although this is partly corrected for by the Kalman filter.
Model Errors	Bathymetry	The bottom topography and roughness coefficient are sources of uncertainty in the DCSM model. The roughness coefficient is a constant whose value is usually determined in the calibration, because it cannot be measured directly.
	Grid resolution	The accuracy of storm surge forecasts is also limited by the accuracy of the storm surge model itself. Although the model was calibrated extensively, the limited resolution is still a source of uncertainty, especially near the coast. The model grid resolution of DCSM8 is approximately 8 km in each dimension. A former version of the model (DCSM16) used a 16 km grid.
	Surface roughness	The surface roughness or wind drag coefficient is a source of uncertainty that affects the wind setup. A Charnock-type wind-drag relation is used, but there are indications that this formulation is no longer correct for extreme wind speeds.
Forcing Errors	Wind and pressure forecast	The most important source of uncertainty for the WAQUA/DCSM model is the wind forcing by meteorological forecasts. The increasing uncertainty of the wind forcing for longer lead times is the main reason for the limitations to the accuracy of the hydrodynamic model.
	Astronomical tide	The simulation of the astronomical tide suffers from uncertainty. This is corrected for by taking the difference between a simulation that includes wind forcing and a second run for only the astronomical tide. This way, most of the error in the astronomical tide cancels out. The effect on forecasts of high surges is therefore limited, but it does enter into the assimilation process of observed water levels.
	Boundary conditions	The boundary conditions for the forecasting model are located at distant deep water locations. Nevertheless, these boundary conditions are based on purely astronomical tide. The influence of wind outside the model domain is thus neglected.

B.2.4 Types of Models

The WAQUA/DCSM is a semi-implicit numerical model of the 2D shallow water equations, which are solved on a regular grid (spherical coordinates). A steady state Kalman filter is applied to assimilate recent water level measurements. Filtering locations are obtained form Wick, North Shields, Lowestoft, Southend, Vlissingen, Hoek van Holland, IJmuiden and Den Helder. The meteorological forcing is derived from the KNMI limited area model HIRLAM7.2.

The uncertainty framework shows that the following techniques are potentially available to provide uncertainty information for this type of system:

Table B.4 Summary of approaches potentially available to provide uncertainties

Name	**Data Assimilation (internal)**	**Data Assimilation (external options)**	**Parameter/state exchange**
Kalman Filter	WAQUA-KF		Native text files
Bayesian Model Averaging		*NOOS*	
Ensemble Prediction System (EPS)		*EMCWF-EPS*	

A solution to deal with the uncertainty in meteorological input is the use of ensemble forecasts. These provide information about the range of possible realizations, and, if applied to the storm surge model, also for the water levels. For small scale phenomena ensemble runs can be useful for assessing different possible scenarios.

Other countries surrounding the North Sea use similar although not identical models for operational water level forecasting. These models use different meteorological inputs. The forecasts are exchanged on an online FTP server and made available using an application called MATROOS. The different model forecasts can be employed as a poor man’s ensemble in a multi-model approach.

B.2.5 Operational Requirements

The SVSD storm surge warning service is currently based on the deterministic WAQUA/DCSM-8 model forecast and on the practical experience of the SVSD staff.

Most research efforts in recent years have focused on improving the accuracy of the deterministic forecast as much as possible. The current uncertainty for the 48 hour forecast of 10 to 15 cm is considered acceptable.

No probabilistic or uncertainty information is included in the decision whether or not to issue a warning. However, there is an interest in such information, particularly for longer lead times.

B.2.6 Run Times

The computing time for a 48 hour forecast of WAQUA/CSM-8 model is about 2 hours on a standard PC. The operational forecasting for the SVSD is done at KNMI using non-standard hardware. Uncertainty information based on an ensemble or multi-model approach can only be done using parallel computing, in order to produce practically useable results for the warning service.

B.2.7 Performance Measures

Commonly used performance measures are the RMSE and bias. An example of an RMSE output for the high tide at Hoek van Holland is shown below.

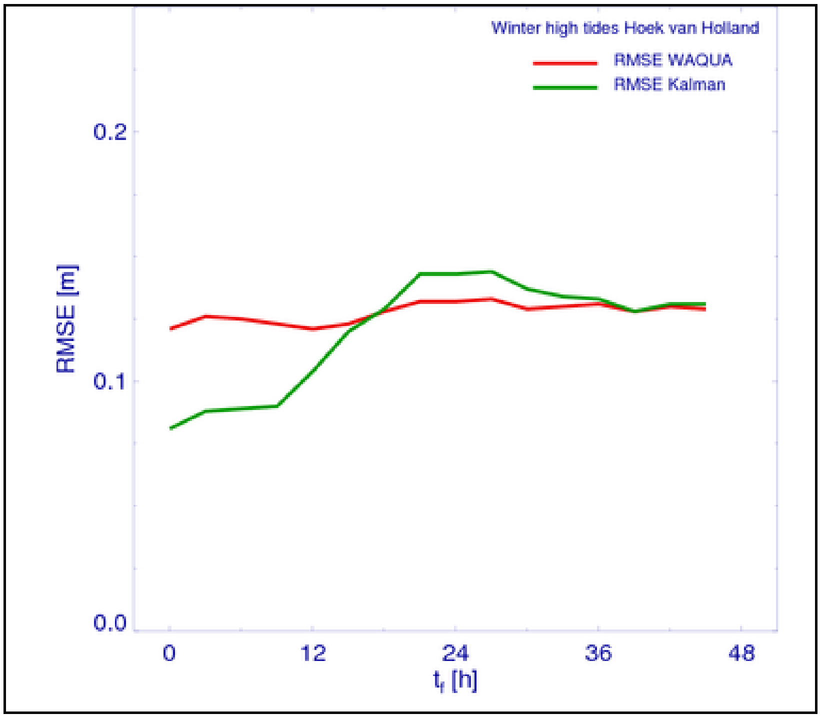


Figure B.3 RMSE of the WAQUA/DCSM-8 model forecast for high tide at Hoek van Holland with and without Kalman filtering. Different lead times up to 48 hours

B.2.8 Choice of Methods

Having considered the various key factors, the final stage in the analysis is to select the most appropriate uncertainty reduction and uncertainty estimation techniques for the SVSD forecasting system. Table B.5 provides an overall summary of the results from this analysis. Table B.6 presents the completed worksheet for this case study.

Data assimilation (Kalman Filter) is already employed and has proven to yield a valuable uncertainty reduction for short lead times (0-12 hours). For lead times around 24 hours, however, the Kalman Filter actually increases the uncertainty. This is due to a delayed response of the model to the adjustments made by the Kalman Filter.

Two approaches are recommended to produce uncertainty information for different lead times:

- For short lead times, the ensemble of forecasts from the NOOS community (hydrodynamic models similar to WAQUA/DCSM-8, but largely fed by different meteorological forcing), can be employed in a multi-model approach called Bayesian Model Averaging (BMA). This technique has been implemented on the forecast-exchange server MATROOS. Below is a screenshot from that system, showing the confidence interval for the probabilistic forecast. Since most of the NOOS forecasts are issued for relatively short lead times (2 days), this approach will work best for these short lead times.

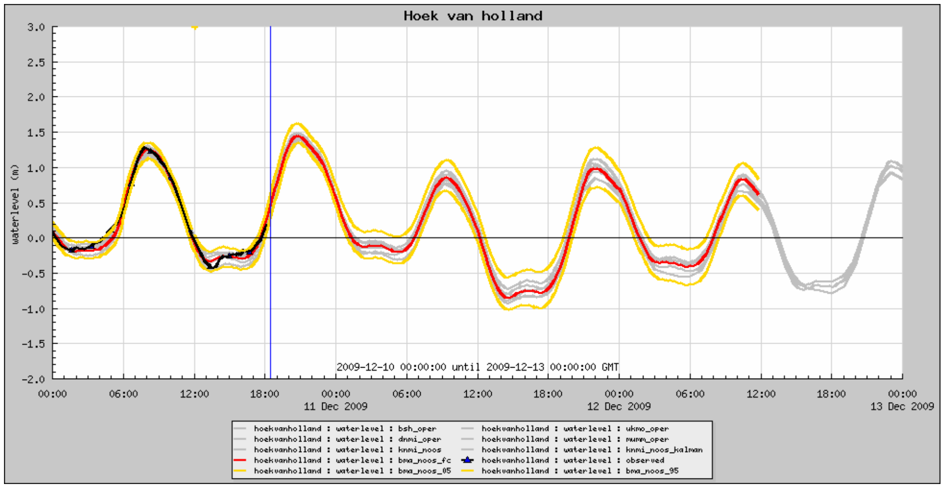


Figure B.4 Probabilistic water level forecasts, based on the BMA-NOOS

- For longer lead times, the meteorological forcing is by far the dominant source of uncertainty. This uncertainty is captured by ensemble runs of the ECMWF meteorological model. The 51 members of the ECNWF-EPS can be used as input to 51 hydrological model runs. The figure below shows an example of the results. The EPS requires calibration, because it is over-dispersed. The calibration of the probabilistic forecast can be done using historical forecasts and observations. Furthermore, the ECMWF-EPS is known to underestimate the uncertainty of the meteorology for short lead times (up to 2 days). Therefore this approach is most appropriate for longer lead times.

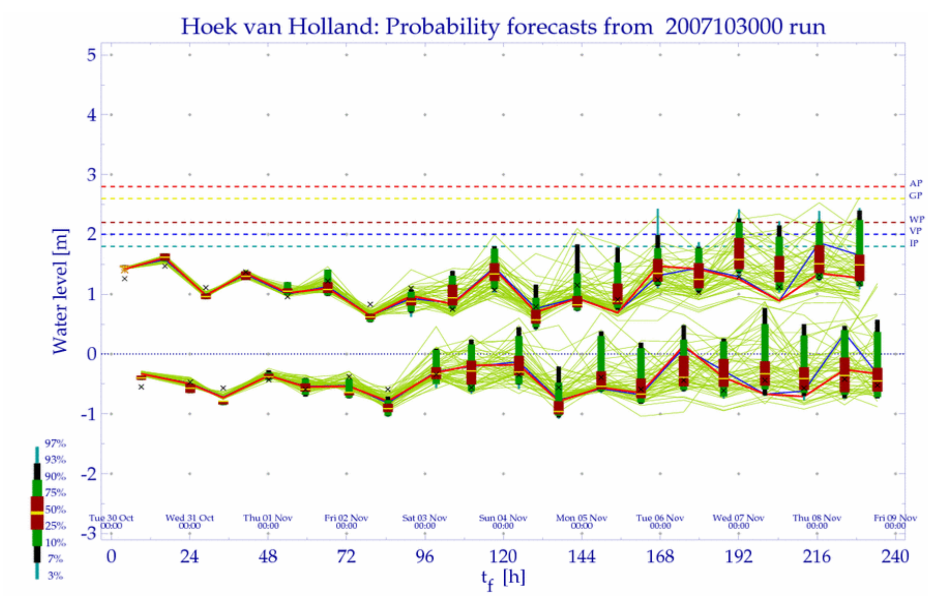


Figure B.5 Probabilistic water level forecasts, based on the ECMWF-EPS

Table B.5 Summary of Application of the Uncertainty Framework to the SVSD

Catchment	**North Sea**	
Forecasting Points	Hoek van Holland	
Model Type	Type 5	
Model(s)	WAQUA/DCSM-8	
	Flood Warning	**Outlook Statements**
Level of Service Requirement	6 - 12 hours	None
Main Uncertainties	**Initialisation Errors**: Water levels	**Initialisation Errors:** -
	Modelling Errors: Grid resolution	**Modelling Errors:** -
	Forcing Errors: Meteorological forecasts (wind and pressure)	**Forcing Errors:** Meteorological forecasts (wind and pressure)
Quantification of uncertainties	Multi-model approach, such as BMA, on the NOOS models.	ECMWF Ensemble Prediction System.
Reduction of uncertainties	Kalman Filter	-

C Applying the Framework for FEWS Rivieren Rhine & Meuse

This appendix presents the results from the uncertainty framework case study to the operational forecasting system for the Lobith Rhine location on the Dutch-German border. The case study is presented in two parts:

- Part D.1 – Description of the FEWS RIVIEREN system
- Part D.2 – Application of the Uncertainty Framework

C.1 Lobith and FEWS Rivieren

C.1.1 River forecasting in the Netherlands

Reliable water-level forecasts for the Dutch rivers are of great importance for operational water management. About one quarter of the Netherlands lies below sea level and more than 60% of the country is potentially threatened by high water levels at sea and floods from the rivers. The endangered area along the Rhine river is extremely densely populated and is of significant economic and historic value. More than half of the Dutch population lives and works in this part of the country; the harbors of Rotterdam and the national airport Schiphol are located here. The potential economic damage of a flood in this area is roughly estimated at 1,200 billion Euro.

The accepted risk of an inundation in such an area is low. At the same time extreme events are expected to occur more frequently and in a more severe extent as a result of soil subsidence as well as climate-change induced extreme weather conditions and sea-level rise. One approach to reducing the risks and to limit the consequences of these increasing threats is through the development of improved operational warning systems.

Lobith is the location where the Rhine enters the Netherlands. Although in reality, this happens about 4 km further upstream, near Spijk, Lobith has always been a reference location for hydrologic studies of the lower Rhine branches in the Netherlands. For example, the dikes of the lower Rhine branches (see Figure C.1) are designed to withstand a design discharge at Lobith having an exceedance probability of 1/1250 per year.

Lobith discharges are forecast using hydrologic (HBV) and hydraulic models (SOBEK) and a statistical model, called LobithW. Until the late 1990's, the statistical model was the only operational forecasting model for the lower Rhine branches that was used by the Dutch water authorities. LobithW consists of several sub-models and uses statistical correlations between water levels, discharges and rainfall at upstream gauging stations. To take into account some of the nonlinearity of the system, the sub-models are split into high and low water level regimes.

Figure C.1 Rhine branches in The Netherlands

After the 1993 and 1995 floods of the Meuse and Rhine Rivers in the Netherlands the need for forecasts with a longer lead time became clear. The limited time that was available for the large scale evacuations during the 1995 flood was seen as highly undesirable. In order to increase the lead time for Lobith use was made of a hydraulic model and rainfall-runoff models for the tributaries.

Water-level forecasts for the rivers Rhine and Meuse in the Netherlands are the responsibility of the Centre for Water Management (formerly RIZA) of Rijkswaterstaat. Under normal circumstances these forecasts are made every morning on a daily basis (365 days a year), mainly for navigation on the Rhine. During floods the frequency of forecasts is increased to at least twice a day.

Figure C.2 1995 Flooding of the lower Rhine branches

C.1.2 FEWS Rivieren Rhine & Meuse

Until the late nineties only a relatively simple computer model, called LobithW, was used to forecast the water level of the Rhine. This model is based on knowledge and arithmetic principles from the fifties and calculates the water level at the gauging station Lobith near the German-Dutch border on the basis of statistical relations with a number of reference points. The model produces a forecast of the water level at the German-Dutch border with a lead time of four days. Experience shows that only the first two days are reliable. During the last big floods in 1993 and 1995 it was shown that the preparation time for the evacuation of a larger area should be at least 2 ½ to 3 days. The existing forecasting system was not able to produce a reliable forecast for this lead time.

In the period after 1998 river-stage forecasting went through a spectacular development, not only regarding computer models, but especially in the field of available data. Because of the immense increase in available data and the developments in the field of IT (internet, data transmission, faster computers) it became possible to use more advanced physical models in operational mode. The Centre for Water Management and Deltares have developed in the past decade, in cooperation with sister organizations in Switzerland and Germany, a so-called Flood Early Warning System (FEWS). An operational version of this system for the Rhine and Meuse Rivers, called FEWS Rivieren, has been installed in 2008 (see Figure C.3).

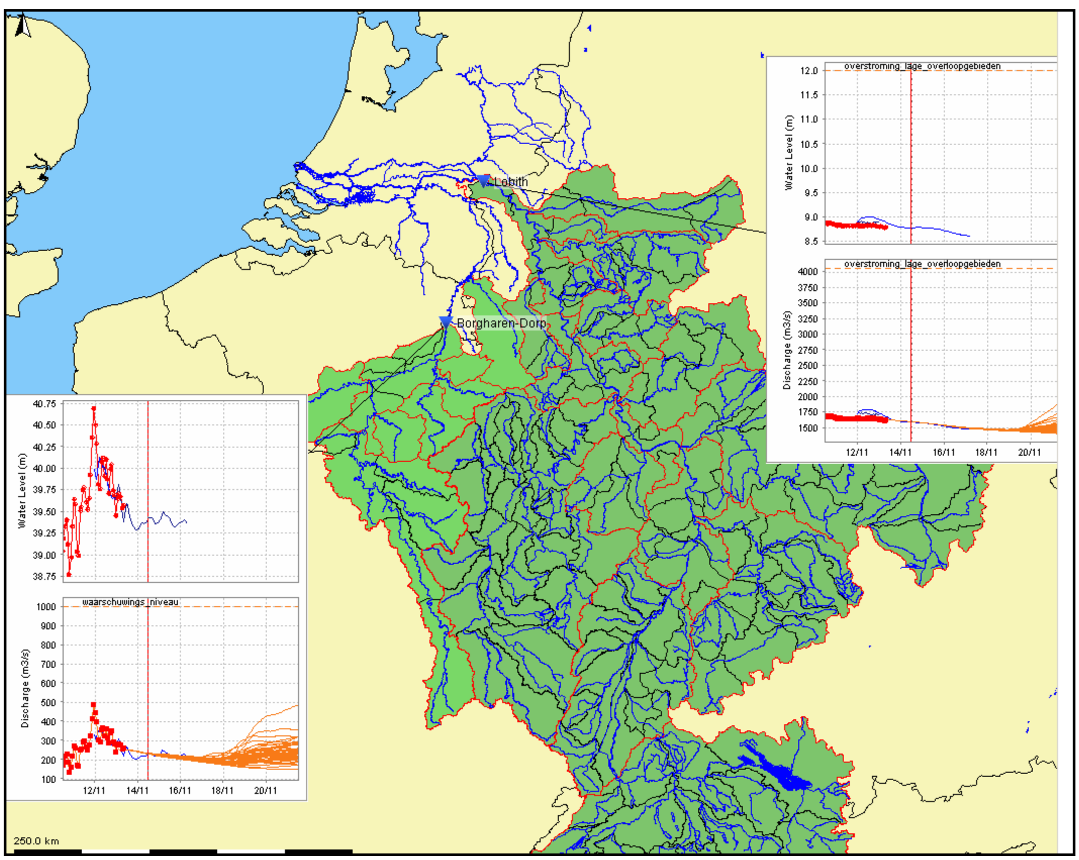


Figure C.3 FEWS RIVIEREN screendump

FEWS Rivieren is an advanced combination of hydrological and hydraulic models with software for import, validation, interpolation and presentation of data. In comparison to the former statistical model FEWS Rivieren uses significantly more data as input. Every 30 minutes the system receives observed water levels from about 60 gauging stations in the Rhine basin. Every hour meteorological observations are downloaded from servers at the national Dutch (KNMI) and German (DWD) weather services of more than 600 stations in the basin of Rhine and Meuse. The system uses output from four numerical weather models at KNMI, DWD and the European Centre for Medium Range Weather Forecasts (ECMWF).

This extreme increase of available data has great advantages but also creates new challenges. In addition to weather forecasts from four deterministic models, the Centre for Water Management also receives ensemble weather predictions from the ECMWF. The ECMWF global ensemble produces 51 ensemble members at 40 km resolution with a lead time up to 14 days. The limited area ensemble (COSMO LEPS) is composed of 16 members on a 10 km grid with a lead time of 120 hours. These ensemble scenarios are all fed into hydrological models, so that in principle more than 70 discharge predictions are now available (see Figure C.4 and Figure C.5).

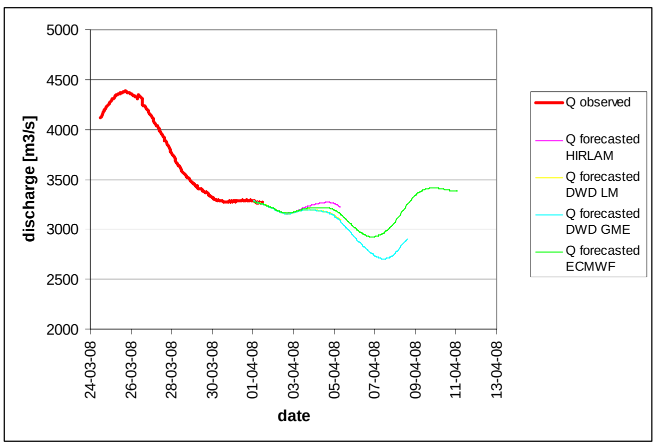


Figure C.4 Deterministic discharge forecasts for the Rhine at Lobith with FEWS RIVIEREN

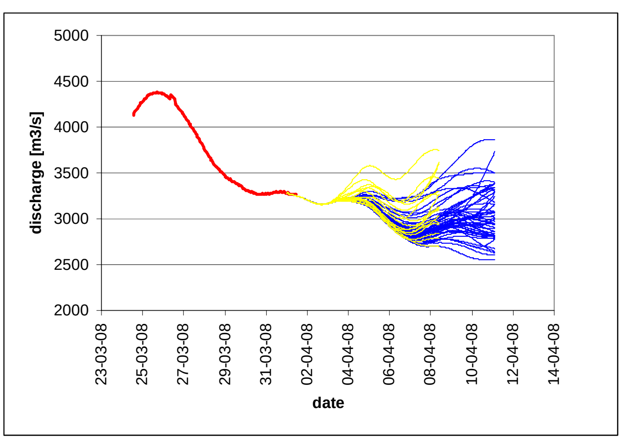


Figure C.5 Probabilistic discharge forecasts for the Rhine at Lobith with FEWS RIVIEREN (red = observed discharge, yellow = forecasts from COSMO LEPS ensembles, blue = forecasts from ECMWF ensembles)

This abundance of forecasts complicates the forecasting process, but at the same time also provides valuable information about the (un)certainty of the forecast. This requires, however, a translation of the spread of the ensemble forecasts to a probability, which is not straightforward. Another problem is how to communicate uncertainties to those that have to make management decisions, such as whether to evacuate an area or not. A crisis manager requires an unambiguous forecast and is not used to taking decisions based on probabilities.

C.2 Applying the Uncertainty Framework

Following the uncertainty framework, the following key factors will be considered:

- Level of Risk
- Lead Time Requirements
- Types of Model
- Main Sources of Uncertainty
- Operational Requirements
- Run Times
- Performance Measures

A worksheet has been developed to assist in this process. A completed worksheet for the Lobith Rhine case study is attached to the end of this description. The following sections describe the analysis and decision making process, which contributed to completing the worksheet.

C.2.1 Level of Risk

The accepted risk of an inundation is low. The protection standard for river dikes along the Dutch Rhine branches is the 1/1250 per year discharge at Lobith. The corresponding water levels at each location are derived from a numerical model calculation, using the design discharge at Lobith as a boundary condition. The alarm and warning levels are well below these critical levels.

C.2.2 Lead Time Requirements

The Rhine River has a rather stable discharge regime near Lobith. Peak flows in the downstream part of the Rhine basin generally occur during the winter months. These can develop if large amounts of rainfall occur over a large part of the basin, mostly in combination with snowmelt. It takes some time for a peak flow to reach the Netherlands. The flow from the main tributaries in Germany (Mosel and Main) takes 2 days or more to reach Lobith. Consequently, the lead time for flood forecasting is long compared to the individual tributaries of the Rhine or the River Meuse. This makes this case study fall into the category 4 type catchment (large river basins).

Model Lobith computes water levels up to four days ahead. However, only the first two days are reliable. During the last big floods in 1993 and 1995 it was shown that the preparation time for the evacuation of a larger area should be at least 2 ½ to 3 days. The forecasts based on hydrologic and hydraulic models are reliable up to four days ahead.

C.2.3 Main Sources of Uncertainty

The main sources of uncertainty for a Type 4 system such as the FEWS Rivieren depend on the lead time. This is visualized in Figure C.6 . For short lead times, the uncertainty of the stage discharge relation of upstream locations is most important. For longer lead times, typically for 5 days or more ahead, the rainfall forecasts dominate the uncertainty.

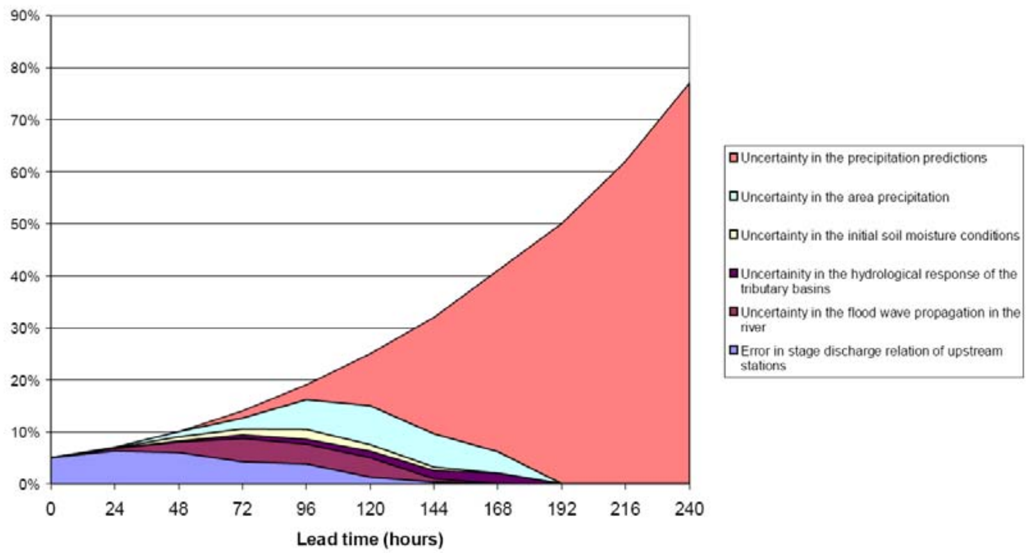


Figure C.6 Contributions to the uncertainty of the forecast discharge at Lobith

C.2.4 Types of Models

Discharges at Lobith are forecast using an integrated catchment model cascade consisting of a rainfall runoff models (HBV-96) and a hydraulic model (SOBEK-RE) using various forms of data assimilation besides this model cascade a statistical model, called LobithW, is used for daily forecasting. Meteorological inputs to these models is generated by the national Dutch (KNMI) and German (DWD) weather services and the European Centre for Medium Range Weather Forecasts (ECMWF).

C.2.5 Operational Requirements

For the Rhine at Lobith with a 1-day leadtime an accuracy of 10 cm in water level is considered acceptable for decision taking during a crisis situation. This corresponds to an accuracy of the discharge of around 250 m^3/s. For a 2-day leadtime an accuracy of 20 cm is considered acceptable, for a 3-day leadtime 30 cm and for a 4-day leadtime 40 cm (Rhine Action plan, 1999).

For the Meuse, the leadtime is currently 1-day with the goal to extend it in the near future to two 2-days. No clear accuracy limits are set for the Meuse.

C.2.6 Run Times

Runtimes are currently not an issue for both Rhine and Meuse. However, Ensemble Kalman Filter (EnKF) is only applied in the historical window (1day) because use in forecast window (4-10 days) would cost too much runtime.

C.2.7 Performance Measures

For the Rhine and Meuse each year a hindcast over the last two years is performed with the newly updated system to test its performance. Below some results of the hindcast over 2006 & 2007 from Weerts (2008) are shown which make clear that for the first few days initial states in the hydraulic model are the dominate source of uncertainty. Based on past performance of the HBV-96 model, the rainfall interpolation method was also evaluated and improved in 2008 (Weerts et al., 2008).

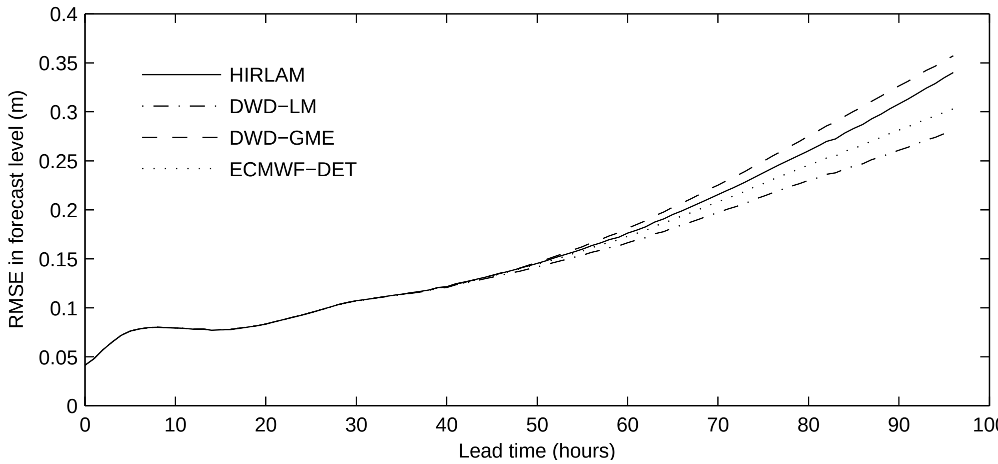


Figure C.7 *Root mean squared error of the water level forecast at the gauge of Lobith on the Rhine determined over a two year hindcast (2006&2007) with EnKF using all NWP forecasts (HIRLAM, DWD-LM, DWD-GME, ECMWF-DET)*

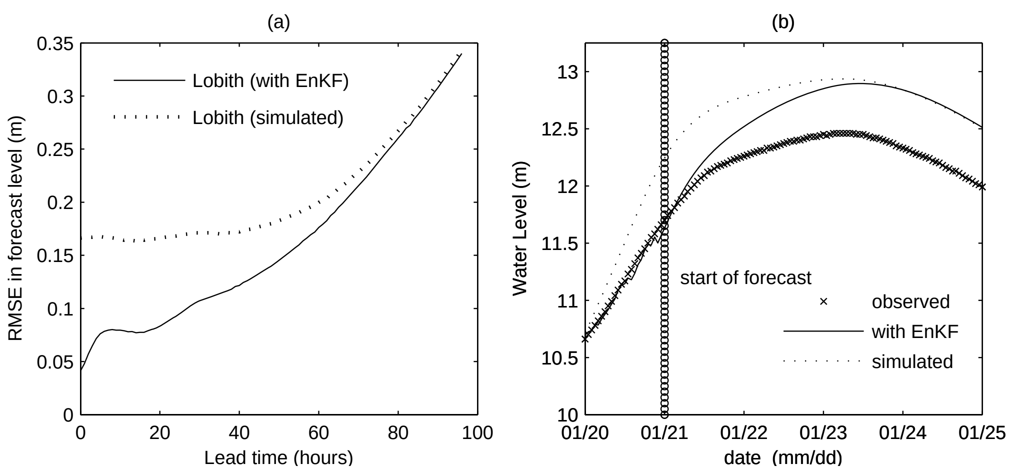


Figure C.8 *Root mean squared error of the water level forecast at the gauge of Lobith on the Rhine with EnKF and without assimilation as a function of lead time determined over a two year hindcast (2006&2007). (b) Observed water level together with the mean of the EnKF water level forecast and the water level forecast without assimilation at Lobith for an event in January 2007. The HBV-96 - SOBEK-RE model cascade is forced using HIRLAM NWP*

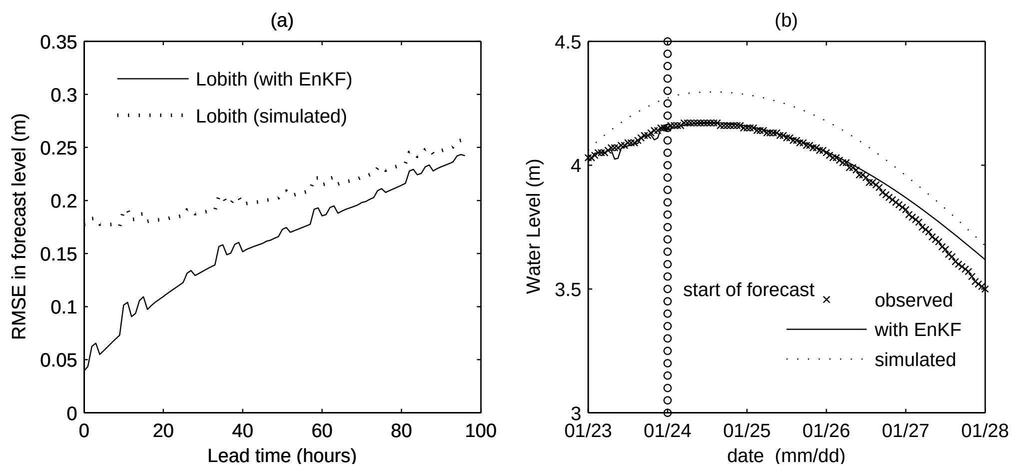


Figure C.9 Root mean squared error of the water level forecast at the gauge of Olst on the Rhine with EnKF and without assimilation as a function of lead time determined over a two year hindcast (2006&2007). (b) Observed water level together with the mean of the EnKF water level forecast and the water level forecast without assimilation at Olst for an event in January 2007. The HBV-96 - SOBEK-RE model cascade is forced using HIRLAM NWP

C.2.8 Choice of Methods

Table C.1 provides an overall summary of the results from this analysis. Table C.2 presents the completed worksheet for this case study.

Table C.1 Summary of Application of the Uncertainty Framework

Catchment	**Lower Rhine River branches**	
Forecasting Points	Lobith	
Model Type	Type 4	
Model(s)	LobithW, HBV, SOBEK	
	Flood Warning	**Outlook Statements**
Level of Service Requirement	6 -12 hours for immediate response actions.	3 days required for evacuation operations
Main Uncertainties	**Initialisation Errors**: Actual water levels	**Initialisation Errors:** Upstream water levels.
	Modelling Errors: Rating curves, hydraulic modelling (eg roughness)	**Modelling Errors:** Hydrological modelling
	Forcing Errors: -	**Forcing Errors:** Upstream inflows, precipitation
Quantification of uncertainties	HUP, multimodel, BMA, Quantile regression	Meteorological ensemble methods .
Reduction of uncertainties	ARMA, bias correction, EnKF	-

Table C.2 *Worksheet for selection of uncertainty estimation method*

Factor	Key Decisions	Main Findings
Level of Risk	What is the level of risk at individual Forecasting Points or flood risk areas ?	The flood probability standards are relatively low.
	What is the level of risk at a catchment level ?	
	What complexity of approach is generally to be preferred ?	
Lead Time Requirements	What are the lead time requirements for each Forecasting Point ?	**Flood Warning** 6 -12 hours
		Outlook Statement For evacuation several days ahead are required.
	What are the main forcing inputs for each Forecasting Point at those lead times ?	**Flood Warning** Upstream water levels and inflows
		Outlook Statement Forecasted precipitation
	What, at a catchment level, are the key forcing inputs to consider for flood warnings and outlook statements ?	**Flood Warning** Upstream water levels and inflows
		Outlook Statement Forecasted precipitation
Main Sources of Uncertainty	What are the main sources of uncertainty for the catchment for flood warnings ?	**Initialisation Errors:** Actual water levels **Modelling Errors:** Rating curves, Hydraulic modelling **Forcing Errors:** -
	What are the main sources of uncertainty for the catchment for Outlook Statements ?	**Initialisation Errors:** Upstream water levels **Modelling Errors:** Hydrological modelling **Forcing Errors:** Upstream inflows, precipitation
	What additional sources of uncertainty arise from complicating factors ?	Weir regulations in the Netherlands, poorly modelled inflows Netherlands,
Types of Models	What choices of methods are available for the types of models ?	HBV-96 SOBEK-RE LOBITHW
	What types of data assimilation routines are an option ?	ARMA EnKF
	What potential run time issues have been identified ?	EnKF takes long should be optimized Interpolation takes long
Operational Requirements	Is a purely qualitative approach sufficient for generating ensembles?	No
	Is data assimilation desirable or essential ?	Essential

	Is conditioning of forecast outputs required ?	Yes
Run Times	Are there run time issues for the candidate uncertainty estimation methods ?	Yes, EnKF can current only be used in historical mode.
	What are the options for reducing run times ?	Use other methods like HUP, BMA, Quantile Regression.
Performance Measures	Are suitable performance measures already available to evaluate sources of uncertainty?	Yes, RMSE but others are also used (see Renner et al., 2009)
	Which sources of uncertainty (and locations) are identified ?	Hydrological Initial state & forecasted precipitation