Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

Cost-Loss analysis and probabilistic information for water boards

Laura Janssen

KNMI Internal Report IR-2015-03

Master Thesis
Mathematics

Cost-Loss analysis and probabilistic information for water boards

Author:
Laura Janssen
laura.janssen18@gmail.com

Supervisor Radboud University:
Prof. Dr. Ir. Egbert-Jan Sol
egbert-jan.sol@science.ru.nl

Supervisors KNMI:
Drs. Kees Kok
kees.kok@knmi.nl

Dr. Maurice Schmeits
maurice.schmeits@knmi.nl

Second reader:
Dr. Heleen de Coninck
h.deconinck@science.ru.nl

July 7, 2015

Abstract

The water board Wetterskip Fryslân in the province Friesland is responsible for managing the water level in the Frisian belt canal system. When the expected water level there, calculated by among others the expected precipitation, is too high, the water board intervenes by opening sluices and operating pumps. To know the expected water level in the Frisian belt canal system, the precipitation forecast must be as accurate as possible. Expected precipitation based on probabilistic information takes uncertainty into account, in contrast to the use of deterministic information, which may lead to better decisions also for the short term, forecasts up to two days ahead. To investigate that, a strategy for a simplified model of the water board is constructed to give a suggestion at which hours sluices may need to be opened and a pump may need to operate if the water level reaches a certain threshold. The input of the strategy consists of precipitation data in Friesland from winters of 2012 to 2015. This thesis compares the output based on probabilistic information, which differs for different water level thresholds, with the output based on deterministic information. Then, for the probabilistic information, the thresholds are chosen for which the use of probabilistic information may lead to better decisions. With cost-loss analysis it is shown that in almost all cases the use of probabilistic information leads to better decisions. This investigation has not the intention to mimic the situation of Wetterskip Fryslân, but it can be transferred as a tool supporting the decision making process by other water boards.

Acknowledgements

First of all, I would like to thank KNMI for giving me the opportunity to carry out this research. In particular, I would like to thank Kees Kok and Maurice Schmeits for their guidance and support throughout the process. Also, many thanks to Wetterskip Fryslân, in particular Pier Schaper and Gerben Talsma, for providing the data of the water board and explaining the decision making process of water level management. I would like to thank my supervisor, Egbert-Jan Sol, for his feedback and guidance. Thanks to my second reader Heleen de Coninck. Finally, thanks to my friends and family who supported me and gave me additional feedback.

Contents

Chapter 1

Introduction

1.1 Background information

Since more than half of the Netherlands is near or below sea level, the control of water needs management[1]. In the Netherlands, this is conducted by water boards. The water boards act independently from the national government. Already in 1255, the first official water board was established. Dutch water boards are regional government institutes that manage water barriers, water ways, water levels, water quality and sewage treatment in their respective regions. They are almost the oldest form of local government in the Netherlands, some of them founded in the 13th century. Nowadays the Netherlands has 23 water boards[2] ranging in size from about 400 km^2 to 3000 km^2 [1]. They are also empowered to collect taxes. An average Dutch family pays an average of 315 euros per year to the water boards[3].

Water boards use weather forecasts, which are inherently uncertain, for their water management. Many companies, including water boards, have to make decisions based on weather forecasts. One of them is the water board Wetterskip Fryslân in the province Friesland in the Netherlands. Friesland has been managing its waterways since the 10th century, this in trying to prevent floods such as the All Saints' Flood of 1170[4]. This flood washed away much area, enlarged the Wadden Sea en created the Zuiderzee. Using models, the water board calculates, from among others the expected precipitation, the expected water level in the Frisian rivers and canals to know whether the water level will be too high or too low. They can intervene by opening sluices, operating pumps or letting water in. In this way, they take care that the water levels in the Frisian rivers and canals stay within strict limits to prevent floods and drought.

As mentioned before, forecasts contain inherent uncertainties which have to be taken into account to determine whether sluices need to be opened and pumps need to operate to lower the water level in the Frisian rivers and canals. To calculate

[1] http://www.dutchwaterauthorities.com/about-us/
[2] http://www.dutchwaterauthorities.com/about-us/
[3] http://www.dutchwaterauthorities.com/about-us/
[4] http://www.dutchwatersector.com/our-history/

the expected water level, the expected precipitation must be known. For the short term forecasts, up to the first two days ahead, Wetterskip Fryslân uses deterministic forecasts given by Hirlam (High resolution limited area model) [1]. Here the forecast is computed from a single initial condition and does not contain information about the uncertainty of this forecast. For forecasts more than two days ahead, uncertainty plays a more important role. To this end, an ensemble of (deterministic) forecasts is made by adding small perturbations in the beginning of the forecast, which will be integrated in time. This is done by computing the expected precipitation and others of 50 ensemble members, plus a so-called control run, for a forecast period of 15 days, which gives an indication about the uncertainty in the forecast. This is known as the use of probabilistic information; each ensemble member represents a possible scenario for the future amount of precipitation. The spread of this ensemble serves as a measure of the uncertainty in the forecast. This system is known as EPS (Ensemble Prediction System). Besides, a high resolution deterministic run is performed.

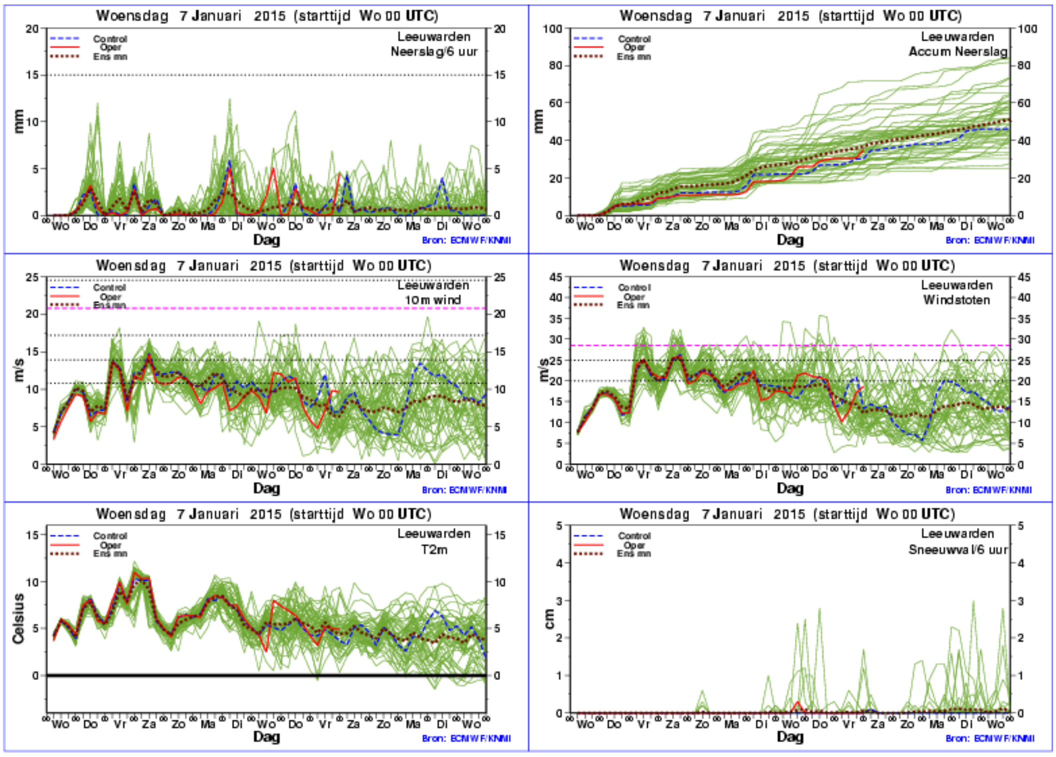

Figure 1.1: *The expected precipitation, accumulated precipitation, wind speed, wind gusts, temperature and snow in Leeuwarden (Friesland) for January 7, 2015. The deterministic run is given in red, the control run given in blue and the 50 ensemble members in green. Together they give an indication of the uncertainty in the forecast (source: ECMWF/KNMI).*

Figure 1.1 is a graphic display of how EPS is used. The expected precipitation from January 7 is shown in the upper left corner. Note for this day that the uncertainty in the first 48 hours is large, as some of the ensemble members show a maximum expected precipitation of about 12 mm per 6 hours where the deterministic run shows a maximum expected precipitation of about 3 mm per 6 hours. Thus, it may happen that even in the short term, the spread of the ensemble is large. Therefore, also taking the uncertainty into account in the short term could lead to better decisions based on these forecasts.

1.2 Problem statement

The Frisian canals and rivers together are referred to as the Frisian belt canal system. How Wetterskip Fryslân determines the expected water level in the Frisian belt canal system will be discussed in Chapter 2. Based on this expected water level, they decide to take precautionary protective action or not. This precautionary action can be sluicing or pumping water away to decrease the water level or letting water in to increase the water level.

A sluice is a barrier between two bodies of water with different water levels[5]. Therefore, it is only possible to drain water through sluices from the belt canal system to the sea, if the water level of the sea is lower than the water level of the corresponding river or canal. Operating pumps is expensive, but can be done at any moment. Sluicing water away and letting water in are relatively cheap. When precautionary actions are needed, the total costs of these actions should be as low as possible. To achieve this, a strategy needs to be developed that suggests when pumps need to operate and sluices need to be opened in the cheapest and most efficient way. Since Wetterskip Fryslân can intervene to increase or decrease the expected water level in many different ways, this thesis presents the strategy for a more simplified construction with only one pump available and at which all sluices can only be opened all at the same time.

Wetterskip Fryslân created boundaries for the water level, indicating when the water board should come into action or not. These actions are opening sluices, operating pumps or letting water in. For the short term forecasts, up to two days ahead, they look at the deterministic expected water level, determined by the deterministic precipitation forecast. When they would also look at the expected water level, based on probabilistic information, uncertainty is taken into account, which may lead to better decisions. This thesis investigates whether the use of short-term probabilistic information instead of deterministic information could lead to better decisions. Therefore this thesis tries to find an answer to the following research questions:

Main question:
How can water boards improve their decision making given deterministic and probabilistic forecast data from KNMI?

Subquestions:

1. Which models does the water board in Friesland use and how does it decide when to undertake protective action?

[5] http://www.ecomare.nl/en/encyclopedia/man-and-the-environment/water-management/management-inland-waters/sluices/

2. What is a cost-loss model, how is it used in determining the economic value of forecasts and which factors could be included to improve the model?
3. What is the uncertainty in economic value using a sensitivity experiment with artificial data?
4. What is the potential added value of the use of probabilistic information for the short term and long term forecasts, using a cost-loss model?

1.3 Approach

With precipitation data in Friesland from 377 days in the winters of 2012 to 2015 this thesis uses the data of Wetterskip Fryslân to calculate the expected water level in the Frisian belt canal system. This expected water level is calculated based on probabilistic precipitation information and based on deterministic precipitation information. We look at a more simplified model for the water board which only has one pump and one sluice possibility. In Chapter 5, a strategy is constructed which suggests the hours at which a pump needs to operate and the sluices need to be opened, based on the use of probabilistic and deterministic information. Then the observed water level is calculated, based on the observed precipitation, with the potential precautionary action chosen by the use of deterministic information and with the potential precautionary action chosen by the use of probabilistic information. In Chapter 6, these observed water levels together with their chosen actions are compared to assess whether the use of probabilistic information results into more efficient sluicing and pumping than by the use of deterministic information.

The cost-loss model states, based on costs of taking precautionary action and potential losses, when a particular user should take precautionary action and when it should not. The use of this model will be discussed in Chapter 3. There is some criticism on this model that it does not take enough factors into account for optimal decision making. These arguments are also discussed. When it is known which actions were chosen given a forecast and whether adverse weather occurred, contingency tables can be constructed which give an indication of the value of that forecast. Then a suggestion can be given to the company which probability threshold it should use to get the most value from the forecast. Given this threshold, the company should take precautionary action if the probability of adverse weather is greater than this threshold.

The question arises how sensitive this optimal threshold is, based on the skill of the forecast. This forecast can have low or high quality and ideally the decisions made by a company using a cost-loss model are based on this weather forecast. Therefore, in Chapter 4, a sensitivity experiment has been carried out (using the statistical programme R [2]) given different skills in forecasts. In this way, the company should not base its decisions on one fixed threshold, but on a range of thresholds, which together result in the highest value of the forecast.

Chapter 2

Task water board

One of the water boards in the Netherlands is Wetterskip Fryslân in the province Friesland. They have provided us with data, giving us the opportunity to look at their decision making process of managing the water level in their belt canal system.

Friesland has many lakes and canals and can drain into to a lake called IJsselmeer and a sea called Wadden Sea. Since this province has so much water, it is an interesting research subject.

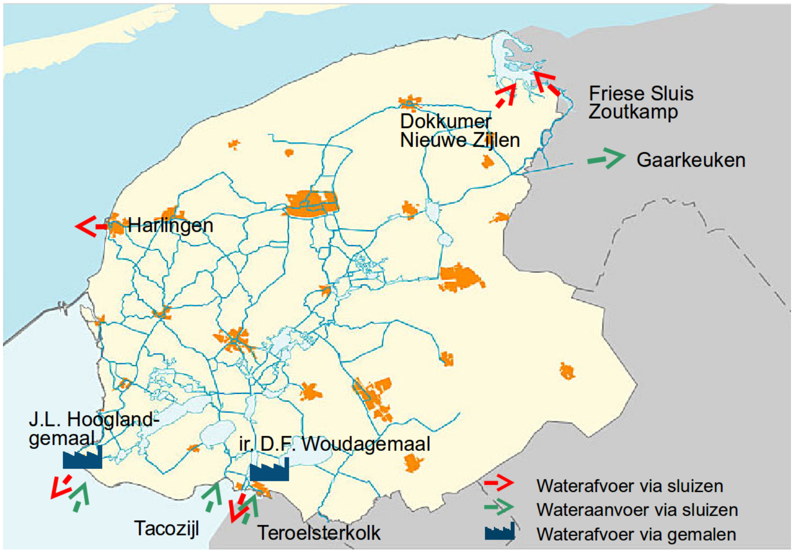

Figure 2.1: *Map of Friesland with location of the pumps and sluices (Courtesy Gerben Talsma).*

2.1 Managing water level

Based on weather forecasts, Wetterskip Fryslân decides whether to take precautionary action or not. If a large amount of precipitation is forecast, pumps may have to operate or sluices may have to be opened to drain water.

The sluices are fast and easily opened. The sluices in Friesland are Dokkumer Nieuwe Zijlen (DNZ), Zoutkamp and Harlingen. Figure 2.1 shows the location of these sluices in Friesland. The red arrows denote the places with drain possibilities and the green arrows denote places with inlet possibilities. The sluices can be opened fully or partly. Table 2.1 shows which options these sluices have. Harlingen consists of four sluice openings, of which two or four can be opened at the same time. Therefore, the capacity of Level 1 and Level 3 can be multiplied with two to get the capacity when all four sluice openings are in use. The costs of opening sluices is cheap, but not exactly known. Therefore, Table 2.1 does not include the variable and fixed costs of opening sluices. Table 2.1 also shows the amount of water the sluices can drain per opening option.

Sluices

Name	*Condition*	*Capacity (mln m^3 per day)*
DNZ	Level 1	0.3
	M2	1.9
	Level 2	2.9
	Level 3	4.5
	Lock and Drain	5.5
Zoutkamp	0.5 m	0.5
	Full	0.9
Harlingen	Level 1	0.35(×2)
	Level 3	0.55(×2)

Pumps

Name	*Condition*	*Variable Costs (€)*	*Fixed Costs (€)*	*Addition (€)*	*Capacity (mln m^3 per day)*
Hoogland	Level 2L N	185	0	550	1.1
	Level 2L D	810	0	1,100	3.3
	Level 2H N	375	0	550	1.48
	Level 2H D	1,625	0	1,100	4.45
	Level 4L N	370	0	1,100	2.2
	Level 4L D	1,620	0	2,200	6.6
	Level 4H N	750	0	1,100	2.97
	Level 4H D	3,250	0	2,200	8.9
	Level 4HL N&D	1,950	0	2,200	7.4
Wouda	2 Kettles	19,500	15,000	0	5.3
	3 Kettles	25,500	15,000	0	6.2

Table 2.1: *Capacity of sluices and pumps of Wetterskip Fryslân. Variable costs are expressed per day. Fixed costs are expressed per installation. Addition is charged by the Frisian government every first day of the month. The capacity of Harlingen can be multiplied with two, since it has two sluice openings.*

There are two main pumps in Friesland: the *Hoogland pumping station* and the *Wouda pumping station.*

Figure 2.2: *The Hoogland pumping station (Courtesy Gerben Talsma).*

The Hoogland pumping station runs entirely on power. Table 2.1 shows the costs of operating Hoogland. These costs range from €150 to €3,500. This has to do with the many different ways in which Hoogland can operate and the difference between costs in off-peak and peak hours. In the weekends the costs of operating Hoogland are also lower than during the rest of the week. There are four openings through which Hoogland can pump water away which can be operated at a High (H) or Low (L) level. The capitals N and D stand for the time at which Hoogland operates. The N denotes Nights, where the costs of pumping are cheaper caused by the off-peak rates. The D stands for Days, at which operating Hoogland during both peak-rates and off-peak rates have to be paid. The option 'Level 4HL N&D' notes the possibility to operate Hoogland at a low level during peak-hours and at a high level during off-peak hours. The costs of Hoogland can be found in Table 2.1. Because the power needed to operate Hoogland can be large, the Frisian government also imposes an addition on the costs of power. The government charges these costs at the first day of each month. The water board tries to avoid operating the Hoogland pump by sluicing water as much as possible. However, if the expected discharge in the Frisian belt canal system is high due to a large amount of precipitation, even sluicing and operating Hoogland may not be enough.

Figure 2.3: *The Wouda pumping station (Courtesy Gerben Talsma).*

If the expected discharge of water is even too large for the sluices and Hoogland, the Wouda pump must operate. This pump, compared to the Hoogland pump, runs entirely on steam. This is why the government does not charge an addition when this pump operates, as can be seen in Table 2.1. Since this pump does not run on power, it does not matter whether it runs during peak hours or not. The Wouda pumping station is the largest steam-driven pumping station in the world and still in use. It can pump over four million litres of water per minute from the Frisian belt canal system into the IJsselmeer[1]. To heat the water in the kettles, warm up the oil and start up the eight pumps one by one, it takes a team of minimal 11 people. After six hours starting and heating up, the pump works optimally[2]. The start up costs of this pump alone are approximately €15,000 and when it operates, it costs about another €20,000 per day. Two or three kettles can operate at the same time. The costs of operating these kettles can be found in Table 2.1. Because of this high price and the long

[1] http://woudagemaal.nl/7434/information-in-english/

[2] http://www.woudagemaal.nl/7438/stoomgemaal/

start up time, Wetterskip Fryslân tries to prevent that it needs to operate by using the sluices and Hoogland pump. After Hoogland was built, Wouda operated less often. But still Wouda may be necessary to lower the water level. An example is begin 2002, when Wouda operated for 10 days in a row[3].

If even the Wouda pump cannot handle the amount of water that has to be drained away, an emergency organization has to be congregated. In such occasions it is discussed whether there is a probability that Friesland may have floods. It may be the case the whole province needs to be evacuated or just a part of it. This, however, is very expensive and, therefore, needs to be prevented.

2.2 Calculation expected discharge

Before the expected water level in the Frisian belt canal system can be determined, the expected discharge into the belt canal system needs to be calculated. For this Wetterskip Fryslân uses a complicated model called the SAMO-model which has relatively large computing time. This is because this model is not only based on the expected precipitation, but also on factors like the expected evaporation, wind speed, wind direction and temperature. To calculate the expected water level for this thesis, the SAMO-model may be too complicated and may take too much time to compute. Therefore the decision makers of Wetterskip Fryslân suggested a simplified model that gives an indication of the expected discharge in their belt canal system and is mostly based on the expected precipitation. This simplified model is based on the equation given in Equation (2.2.1) where ED stands for the Expected Discharge in million cubic metres per day on the Frisian belt canal system, EP for the Expected Precipitation in millimetres per day and n for the day for which the expected discharge is calculated:

$$\begin{aligned} ED[n] = WF \cdot 3 \cdot (&EP[n] \cdot 0.15 + EP[n-1] \cdot 0.25 + EP[n-2] \cdot 0.1 + \\ &EP[n-3] \cdot 0.08 + EP[n-4] \cdot 0.06 + EP[n-5] \cdot 0.04 + \\ &EP[n-6] \cdot 0.02). \end{aligned} \tag{2.2.1}$$

Equation (2.2.1) is mainly based on the rule of thumb that about 40% of the precipitation that falls in the first two days (including the day looked at) comes into the discharge. This is when the ground does not absorb any precipitation. The three days before that also include some percentage of the precipitation that comes into the discharge. This equation indicates a lump model which is based on experience. Note that Equation (2.2.1) does not require much computation time. For the use in this thesis, it is therefore more convenient to use this equation instead of the SAMO-model.

In Equation (2.2.1), WF represents the Weight Factor which is a number between 0.2 and 1. When the weight factor is 0.2, the ground of the polders is considered dry, which implies that the ground can absorb a great part of the precipitation.

[3] http://www.woudagemaal.nl/7494/waterhuishouding/

When the weight factor is 1.0, the ground of the polders is saturated which implies that the ground does not absorb much of the precipitation.

In Equation (2.2.1), ED is expressed in million cubic metres per day on the belt canal system. Since the surface of Friesland is about 300,000 hectares, one millimetre of precipitation results in three million cubic metres of water in the whole province Friesland. Therefore, to go from millimetre precipitation to million cubic metres per day on the Frisian belt canal system, the precipitation in millimetre has to be multiplied by a factor three. This explains the factor three in Equation (2.2.1).

Next follows an example of how Equation (2.2.1) is used in calculating the expected discharge for a given day and how much the result differs from the expected discharge calculated by the SAMO-model.

Example:
An example is from January 7 2015. Table 2.2 shows the precipitation in millimetres for the previous six days (which are known), January 7, and the following seven days (which are deterministic forecasts).

Date	Jan. 1	Jan. 2	Jan. 3	Jan. 4	Jan. 5	Jan. 6
Observed Precipitation in millimetres	5.43	0.03	1.13	0.27	0.00	1.25

Date	Jan. 7	Jan. 8	Jan. 9	Jan. 10	Jan. 11	Jan. 12
Expected Precipitation in millimetres	2.66	4.14	2.90	1.60	1.50	3.40

Date	Jan. 13	Jan. 14
Expected Precipitation in millimetres	1.60	0.00

Table 2.2: *Observed and expected precipitation in millimetres per day from January 1st to January 14 2015.*

Because January 7 is in the winter, when the ground is more saturated than in the summer, we set the weight factor equal to 0.8. This can be substituted in the simplified equation given in Equation (2.2.1) to calculate the expected discharge from January 7 2015 to January 14 2015. Table 2.3 shows the outcomes.

Table 2.3 shows for example that the expected discharge from January 8 07:00 2015 to January 9 2015 07:00 is equal to 2.19 million cubic metres. Figure 2.4 compares this calculated expected discharge with the expected discharge calculated by the SAMO-model. Here we can see that the expected discharge based on the SAMO-model and the expected discharge based on our simplified model based on Equation (2.2.1) do not differ much. Therefore we will use our simplified model to calculate the expected discharge in the Frisian belt canal system.

Date	Jan. 7	Jan. 8	Jan. 9	Jan. 10	Jan. 11	Jan. 12	Jan. 13	Jan. 14
Expected discharge in million cubic metres per day	1.26	2.19	3.54	4.49	4.01	3.49	3.98	4.23

Table 2.3: *Expected discharge per day in million cubic metres per day from January 7 to January 14 2015.*

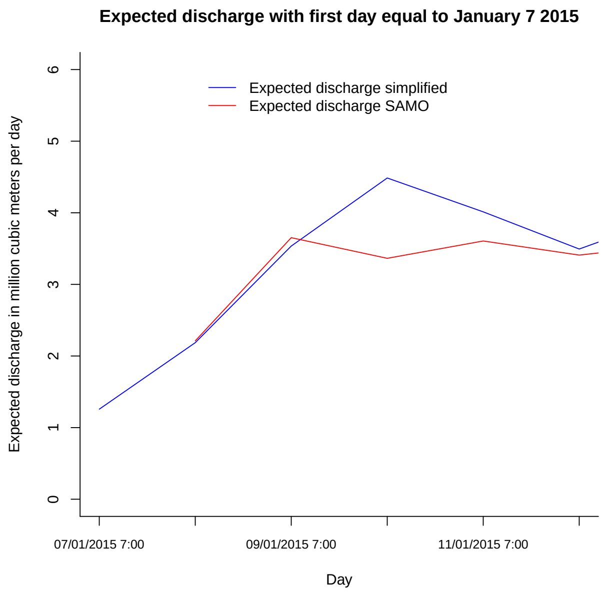

Figure 2.4: *The expected discharge calculated according to a simplified model based on Equation* (2.2.1) *(blue) and according to the SAMO-model used by Wetterskip Fryslân (red).*

2.3 Calculation expected water level

From the expected precipitation, expected discharge and expected evaporation, the expected water level in the Frisian belt canal system can be calculated. The model used by Friesland calculates the expected water level for every quarter of an hour. It is based on the following formula, where i is in quarters of an hour, precipitation and evaporation are in millimetres per quarter of an hour, and discharge is in million cubic metres per quarter of an hour.

```
ExpectedWaterLevelPerQuarter[0] = BeginWaterLevel
ExpectedWaterLevelPerQuarter[i] = ExpectedWaterLevelPerQuarter[i-1] +
                ExpectedPrecipitationPerQuarter[i]/10 -
                ExpectedEvaporationPerQuarter[i]*0.125 +
                ExpectedDischargePerQuarter[i]/(96*1.5)
```

Note that the precipitation in this formula is indirectly added to the expected water level through the expected discharge and is also directly added. Therefore, the precipitation has double influence on the expected water level. This is because the discharge only calculates the precipitation that comes from the polders into the belt canal system. But precipitation will also fall directly into the belt canal system, which will also increase the water level.

Figure 2.5 shows the expected water level from January 7, the day discussed in the example in Section 2.2, based on our simplified formula and based on the SAMO-model.

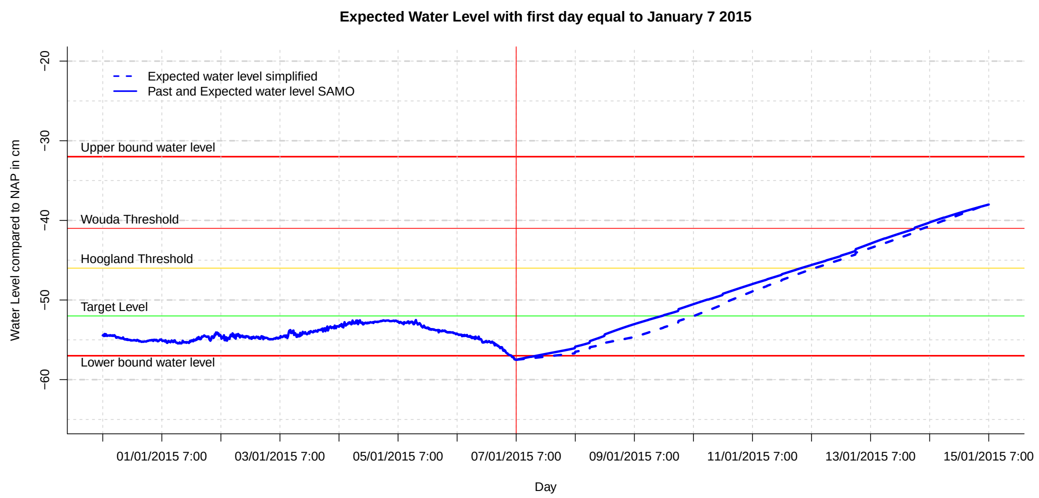

Figure 2.5: *The thick blue line represents the expected water level in Wetterskip Fryslân starting January 7 2015 based on the expected discharge calculated by the SAMO-model. The dotted blue line represents the water level based on the simplified calculation of the expected discharge given by Equation* (2.2.1)*.*

In Figure 2.5 the thick blue line displays the water level as calculated by Wetterskip Fryslân based on expected discharge calculated by the SAMO-model. This water level results when no pumps operate and no sluices are opened. Figure 2.5 also shows the critical water levels established by the water board. These are the upper and lower thick red lines. The upper thick red line is at 32 cm below NAP and the lower thick red line is at 57 cm below NAP. NAP denotes the 'Normaal Amsterdams Peil' and is in the Netherlands the national standard for water levels[4]. The water level should not be above the upper thick red line or below the lower thick red line. When the water level is below the lower thick red line, the water level in the belt canal system is too low, which is bad for nature. When the water level is above

[4]http://www.normaalamsterdamspeil.nl/en/

the upper thick red line, the water level in the belt canal system is too high, which could cause floods in Friesland. Because it is less bad to have the water level in the belt canal system too low than too high, the target level, given in green, is on about one fourth of the difference between the upper red level and the lower red level. The target level is at 52 cm below NAP. This means that Wetterskip Fryslân tries to keep the water level in the belt canal system as close to the target level as possible by opening sluices, operating pumps or letting water in. The Hoogland threshold, which is at 46 cm below NAP, serves as an indication of the need of operating the Hoogland pumping station. This is the case when the expected water level exceeds this Hoogland threshold. In the same way, the Wouda threshold, which is at 41 cm below NAP, serves as an indication of the need the Wouda pumping station.

2.4 Influence of sluices and pumps

The model of Wetterskip Fryslân also shows the influence of operating pumps and opening sluices. Figure 2.6 shows the expected hours of opening sluices and operating pumps from January 7 2015.

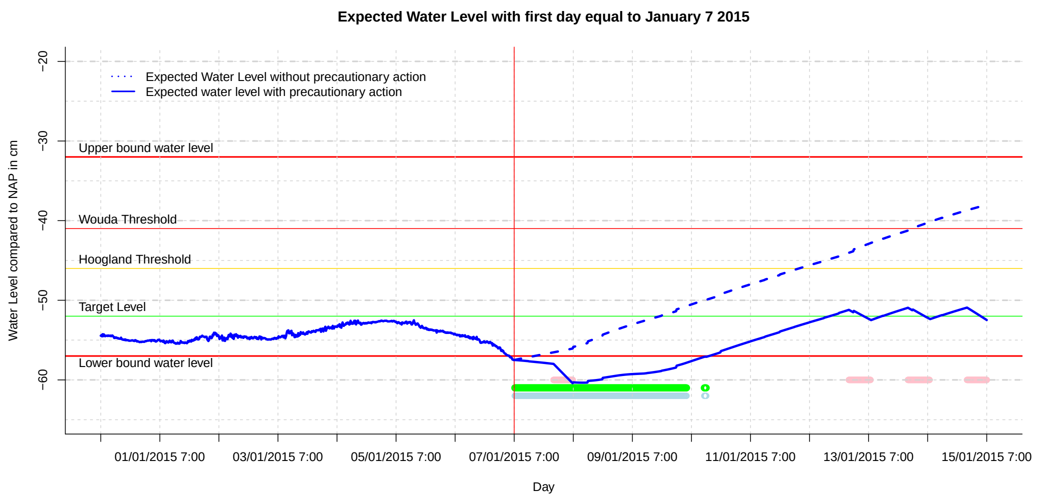

Figure 2.6: *The expected water level calculated by the water board without opening sluices or operating pumps (blue dotted line) and the expected water level after opening sluices and operating pumps (blue thick line).*

It shows that the water board decided to let the water level drop below the lower thick red level to get the water level at the target level for the subsequent days. The thick green line at 61 cm below NAP represents the hours at which the sluice Zoutkamp was opened. The thick light blue line at 62 cm below NAP represents the hours at which the sluice Dokkumer Nieuwe Zijlen was opened. The pink thick line at 60 cm below NAP represents the hours at which the pumping station Hoogland operated. Since the decision maker of Wetterskip Fryslân decided to let the water level come below the lower thick red line, it was even prevented that the Wouda pumping station had to operate in the upcoming days. This may have led to much less expenses than would occur when the Wouda pumping station had to operate. Therefore deliberately decreasing the water level below the critical lower level may have led to less expenses.

Chapter 3

Introducing cost-loss

Based on a weather forecast, a company may have to decide whether to take protective action in the face of uncertainty as to whether or not adverse weather will occur. This is known as a *cost-loss situation* [3]. When a company does not take protective action and adverse weather occurs, this will cost an amount L. If the company chooses to take protective action and adverse weather occurs, then part of the losses, L_1, are prevented. If all potential losses can be prevented, then $L_1 = L$. In this way this event will cost $C + (L - L_1)$. If the company takes protective action and no adverse weather occurs, the costs are C. If the company does not take protective action and no adverse weather occurs, then the costs will be 0. Altogether the following concluding Table 3.1 is constructed showing the costs for each of the four cases:

		Observed	
		Adverse weather	No adverse weather
Forecast	Action	$C + (L - L_1)$	C
	No action	L	0

Table 3.1: *Costs and Losses based on adverse weather and no adverse weather stated by observations and potential precautionary actions.*

As an example, suppose a farmer has a fruit crop. When frost is predicted, the farmer may decide to sprinkle the blossoms with water so that the fruit will not be lost. The costs of sprinkling this water over the fruit is C. When he does not do that, part of the fruit may be lost. These losses are L_1. In this case, L_1 consists of the losses that can be prevented by sprinkling the blossom. When the frost is so severe that parts of the losses cannot be prevented, then these losses are in L, but not in L_1.

Because C, L_1 and L are expressed in amounts of money, it holds that $C, L_1, L \geq 0$. If $C = 0$, the company would always take protective action and if $L_1 = 0$, the company would never take protective action. Since the decisions in these cases are fixed, cost-loss analysis is not necessary. Therefore, these cases will not be further discussed. Note that if $C \geq L_1$ it would never pay off for the company to take protective action, thus these cases are also not further discussed. Altogether, only the cases are discussed where the *cost-loss ratio* C/L_1 lies between 0 and 1. Assessing its cost-loss ratio may help a company in making the most optimal decision which has minimal costs.

For most companies the potential losses are relatively high compared to the costs. Therefore, for many companies the cost-loss ratio C/L_1 is small.

According to an article of Thornes (2001) [4] the simple cost-loss model to judge the value of weather forecasts can be applied to the situations where

(a) the effects of adverse weather and the cost of taking action to avoid weather damage are known;

(b) the potential dissatisfaction of the decision maker is a linear function of the value of the loss;

(c) the probability of occurrence of adverse weather is known precisely.

Note that (a) and (b) should be known in advance and (c) should be known by experience. Consequently, cost-loss analysis can only be done with known data.

Also note that (b) is a quite strong assumption that does not always hold. Many economic studies about behaviour in finance note that decision makers are not always rational and risk neutral.

One of these studies is from J.K. Lazo (2010) where he states that the cost-loss analysis cannot be used in real life, because it is too simplified [5]. He refers to Millner who showed in one of his articles (2009) that incorporating a beneficial feature in the cost-loss model induced a decrease in the economic value compared to the basic cost-loss model and therefore, behavioural aspects limit the effectiveness of the cost-loss model [6]. The limitations of cost-loss are discussed in Section 3.6.

3.1 Applications of cost-loss

Cost-loss analysis is used in different fields. In an article of Thornes about road maintenance (2001) [4] cost-loss analysis is used to decide whether the government should salt the roads, given some weather forecast indicating the probability of frost. If the road is not salted and frost takes place, the frost can cause slippery roads, potentially causing accidents. The avoidable losses L_1 are, in this case, the costs for the insurance policies when an accident has taken place. The precautionary action that can be taken is salting the roads. The costs of salting the roads are expressed as C. If the roads are salted but no frost takes place, then the roads are unnecessarily salted. The simplest cost-loss analysis states that the government should salt the roads if the probability of frost is greater than the cost-loss ratio C/L_1.

Another example where cost-loss can be used is volcanology which is described by Woo in his article ‘Probabilistic criteria for volcano evacuation decisions’ (2008)

[7]. Here, C expresses the costs of evacuation. He defines N as the size of the population which may be evacuated in the event of a volcanic eruption, E as the proportion of the N people at risk who would owe their lives to the evacuation call and V as the notational value of a life to save. Then, the losses L can be expressed as $L = E \times N \times V$. The same calculation holds in case of a hurricane as the hurricane Katrina in New Orleans in 2005. Here, the behaviour of people played a role in whether they were willing to evacuate when it was necessary. Some did not trust meteorologists, did not want to leave their pets behind, were confident after survival of earlier hurricanes etcetera. In determining the losses that may exist, the evacuation time distribution has a long tail, which implies that there is no finite time by which all endangered people will be evacuated. Therefore, there is no safe time window for an alert, during which a mass evacuation can be fully completed.

3.2 Construction of contingency tables

With data from a previous forecast and the knowledge of (a), (b) and (c) given by Thornes [4] the company decides what the value of the forecast is using a simple cost-loss model [8]. First a *contingency table* is constructed. Table 3.2 is such a contingency table. Here, $\bar{o}$ is the fraction of observed events.

		Observed		
		Adverse weather	No adverse weather	Total
Forecast	Adverse weather	a	b	$a+b$
	No adverse weather	c	d	$c+d$
	Total	$a+c=\bar{o}$	$b+d=1-\bar{o}$	$a+b+c+d=1$

Table 3.2: *Contingency table given forecasts of adverse weather and whether it eventually occurred. Given the forecast, a company decides to take precautionary action or not, which places the event in the contingency table.*

In Table 3.2, a denotes the fraction of events at which adverse weather was predicted and occurred, b denotes the fraction of events at which adverse weather was predicted but did not occur, c the fraction at which adverse weather was not predicted but did occur and d the fraction at which adverse weather was not predicted and did also not occur. In this way $a+d$ denotes the fraction of correct forecasts, b denotes the fraction of Type I errors and c denotes the fraction of Type II errors[1]. Note from Table 3.1 that the costs of a Type I error is C and the costs of a Type II error is L with $C < L$, thus according to the cost-loss model it is more attractive to make a Type I error than a Type II error. This is most commonly known as *overforecasting* a particular event so that the probability of a Type I error becomes larger.

To determine how correct the forecast was, the Hit rate H and the False alarm rate

[1] In Hypothesis Testing (statistics), a Type I error is known as a 'false positive' and a Type II error is known as a 'false negative'.

F can be calculated. The Hit rate H is calculated as the forecasts at which adverse weather was predicted and did occur divided by the adverse weather observations. In terms of the variables given in Table 3.2, it holds that $H = a/(a + c)$.
The False alarm rate F is calculated as the forecasts at which no adverse weather was predicted but did occur divided by the total number of no adverse weather observations. In terms of the variables given in Table 3.2, it holds that $F = b/(b+d)$.

From now on, the cost-loss ratio C/L_1 will be denoted as α, the Hit rate as H, the False alarm rate as F and the fraction of observed adverse weather as $\bar{o}$:

$$\bar{o} = a + c, \qquad \alpha = C/L_1, \qquad H = \frac{a}{a+c} = \frac{a}{\bar{o}}, \qquad F = \frac{b}{b+d} = \frac{b}{1-\bar{o}}. \tag{3.2.1}$$

3.3 Cost-Loss model given only climatological information

Note from Table 3.1 that always taking action to prevent a loss L_1 leads to an average expense of $C + \bar{o}(L - L_1)$. Never taking action leads an average expense of $\bar{o}L$. Therefore, given only climatological information, a company should take precautionary action if $C + \bar{o}(L - L_1) < \bar{o}L$, thus when $C < \bar{o}L_1$. This leads to the following summary:

if $\bar{o} > C/L_1$ it will pay to take precautionary protective action,
if $\bar{o} < C/L_1$ it will not pay to take precautionary protective action,
if $\bar{o} = C/L_1$ it does not matter either way.

Recall that $0 < C/L_1 < 1$ and $0 \leq \bar{o} \leq 1$ which entails that the previous three statements are correctly defined.

3.4 Determining forecast value

There are different scores trying to give some indication about the quality of a forecast. Examples are the Percent Correct score PC, the Miss rate M and the Kuipers score [4]. The Percent Correct score (PC) is the fraction of correct forecasts. With the notation given in Table 3.2 it holds that $PC = (a+d) \times 100\%$. This score, however, does not always give the right suggestion, as shown by Murphy in his article called 'The Finley Affair' (1996) [9]. This article refers to another article by Finley (1884) in which he summarized results of an experimental tornado forecasting programme [10]. He used the percent correct score as a verification measure and came to a PC-score of 96.6%. The article of Murphy pointed out that always forecasting no tornadoes would have led to a better PC-score of 98.2%. This is caused by the low occurrence of tornadoes. This phenomenon is known as the *Finley affair*.

For a company that has to make decisions based on weather forecasts, it is important to have an indication of the economic value of a forecast. To compare the quality and economic value of forecast providers, an index is constructed that takes into account the number of Type I and Type II errors as well as the cost-loss ratio. Therefore, the relative economic value V [4] of a forecast system is defined as:

$$V = \frac{E(climate) - E(forecast)}{E(climate) - E(perfect)}. \tag{3.4.1}$$

In Equation (3.4.1) E stands for expected Expense. When the forecast system is perfect, V will have a value of 1; if V is not better than decision making without any forecast system then it will have value 0. $E(climate)$ denotes the expected expenses when only climatological information is available. In this case, the optimal course of action is always to act if $C < \bar{o}L_1$ and to never act otherwise. Therefore, with Table 3.1, the expected expense $E(climate)$ is

$$\begin{aligned} E(climate) &= \min\{C + \bar{o}(L - L_1), \bar{o}L\} \\ &= \min\{C, \bar{o}L_1\} + \bar{o}(L - L_1). \end{aligned} \tag{3.4.2}$$

If the weather is perfectly forecast, then the decision maker only takes precautionary action if the event occurs. Consequently, the expected expense $E(perfect)$ becomes

$$E(perfect) = \bar{o}(C + L - L_1). \tag{3.4.3}$$

$E(forecast)$ denotes the expected expense of the forecast system, which by Table 3.1 and Table 3.2 equals:

$$E(forecast) = a(C + (L - L_1)) + bC + cL$$

Note from Equation (3.2.1) that $H\bar{o} = a$, $(1 - H)\bar{o} = c$ and $F(1 - \bar{o}) = b$ which implies that

$$\begin{aligned} E(forecast) &= H\bar{o}(C + (L - L_1)) + F(1 - \bar{o})C + (1 - H)\bar{o}L \\ &= H\bar{o}(C - L_1) + F(1 - \bar{o})C + \bar{o}L. \end{aligned} \tag{3.4.4}$$

Substituting (3.4.2), (3.4.3) and (3.4.4) in (3.4.1) gives

$$\begin{aligned} V &= \frac{\min\{C, \bar{o}L_1\} + \bar{o}(L - L_1) - (H\bar{o}(C - L_1) + F(1 - \bar{o})C + \bar{o}L)}{\min\{C, \bar{o}L_1\} + \bar{o}(L - L_1) - \bar{o}(C + L - L_1)} \\ &= \frac{\min(\alpha, \bar{o}) - F\alpha(1 - \bar{o}) + H\bar{o}(1 - \alpha) - \bar{o}}{\min(\alpha, \bar{o}) - \bar{o}\alpha}. \end{aligned} \tag{3.4.5}$$

Thus, the relative value of a particular forecast system depends on the cost-loss ratio α and fraction of occurrence $\bar{o}$, which are external to the system, and hit rate H and false alarm rate F which are model dependent [8].

Figure 3.1 displays V for $H = 0.88$, $F = 0.14$ and $\bar{o} = 0.43$.

Figure 3.1 suggests that the highest value for V is reached when $\alpha = \bar{o} = 0.43$. This is, in fact, true. This is proven in Theorem 3.4.1.

Theorem 3.4.1. *The maximum value of V given in Equation* (3.4.5) *is reached when $\alpha = \bar{o}$.*

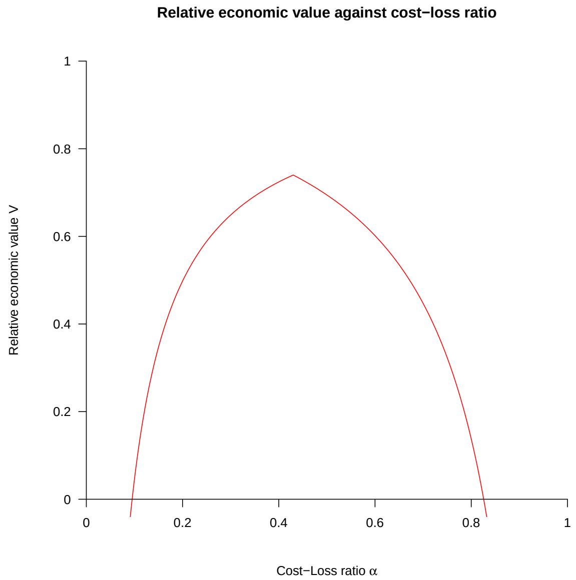

Figure 3.1: *Relative economic value V expressed in terms of cost-loss ratio α with $H = 0.88, F = 0.14$ and $\bar{o} = 0.43$.*

Proof: For $0 \neq \alpha \leq \bar{o}$ Equation (3.4.5) becomes

$$\begin{aligned} V &= \frac{\alpha - F\alpha(1-\bar{o}) + H\bar{o}(1-\alpha) - \bar{o}}{\alpha - \bar{o}\alpha} = \frac{1 - F(1-\bar{o})}{1-\bar{o}} + \frac{H\bar{o} - H\bar{o}\alpha - \bar{o}}{\alpha - \bar{o}\alpha} \\ &= \frac{1 - F(1-\bar{o})}{1-\bar{o}} + \frac{(H-1)\bar{o}}{\alpha(1-\bar{o})} - \frac{H\bar{o}}{1-\bar{o}}. \end{aligned} \tag{3.4.6}$$

Since $H - 1 \leq 0$, Equation (3.4.6) is maximum when α is maximal. As it is assumed that $\alpha \leq \bar{o}$, this happens when $\alpha = \bar{o}$.

For $\bar{o} \leq \alpha$ Equation (3.4.5) becomes

$$\begin{aligned} V &= \frac{\bar{o} - F\alpha(1-\bar{o}) + H\bar{o}(1-\alpha) - \bar{o}}{\bar{o} - \bar{o}\alpha} = \frac{-F\alpha(1-\bar{o})}{\bar{o}(1-\alpha)} + \frac{H\bar{o}(1-\alpha)}{\bar{o}(1-\alpha)} \\ &= H - \frac{F\alpha(1-\bar{o})}{\bar{o}(1-\alpha)}. \end{aligned} \tag{3.4.7}$$

Note that Equation (3.4.7) is maximum when α is minimal. As it is assumed that $\bar{o} \leq \alpha$, this happens when $\alpha = \bar{o}$. □

With Equation (3.4.7) (or Equation (3.4.6)) the maximum value of V given H and F is equal to:

$$V_{\max} = H - \frac{F\bar{o}(1-\bar{o})}{\bar{o}(1-\bar{o})} = H - F.$$

This is equal to the Kuipers Score (KS) which has the desirable characteristics of equitability, in that random or constant forecasts will score 0 and perfect forecasts will have a score of 1 [8]. The Kuipers score is defined as follows:

$$KS = \frac{ad - bc}{(a+c)(b+d)}.$$

Equation (3.4.5) expresses the economic value V of a forecasting system given a cost-loss ratio α. In the next section, it is discussed how a company with a specific situation can use this equation to determine in which case it has to take precautionary protective action.

3.5 More advanced cost-loss models

Until now only a one-staged cost-loss model is discussed, in which one decision is made about whether or not to take protective action. It is, however, possible and maybe even more accurate, to look at cost-loss models consisting of more than two states. This model is described in an article by Roulin (2007) [11]. Figure 3.2 shows an example of a two-staged cost-loss model.

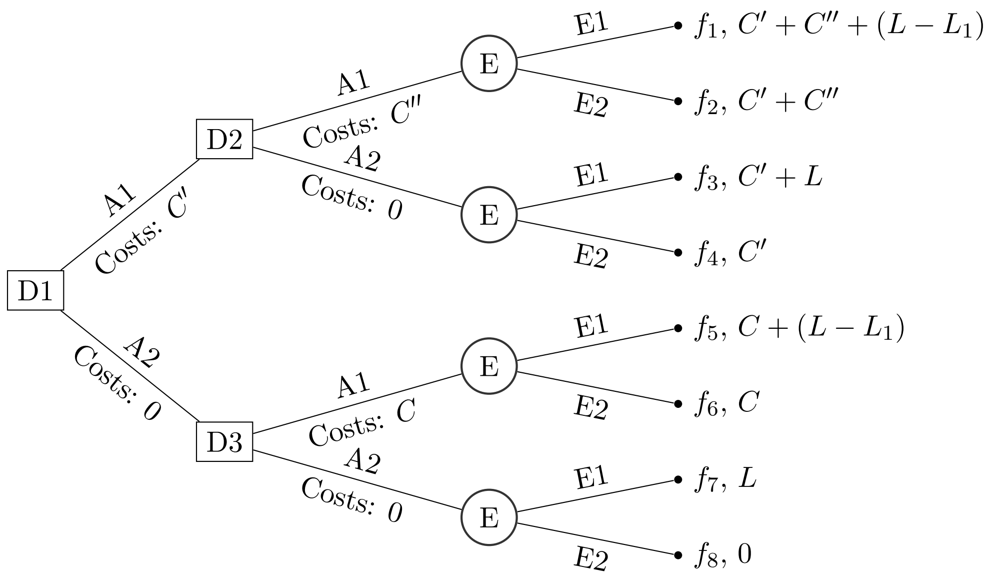

Figure 3.2: Two stage cost-loss model.

In the first stage the first decision (D1) has to be made about whether to take precautionary action or not. This decision is to take precautionary action (A1) or not to take precautionary action (A2) which may or may not induce costs C'. After that, another decision should be taken. This can be A1 or A2, which may entail costs C''. When there is no action taken in the first stage, the next decisions may or may not cause costs C; it represents the one-staged cost-loss model discussed in the beginning of this chapter. When these decisions are taken, the event will take place (E1) or not (E2) which may cause losses L or $L-L_1$. The two-staged cost-loss model ends with eight expenses that are represented by the decisions made and the events occurred and with the fractions that these event occur (f_i for $i \in \{1, \ldots, 8\}$). When this is known, the branch is selected that gives the least expense multiplied with the probability of occurrence [11].

3.6 Comments on cost-loss

As stated before, an article by Andrew Millner claimed that incorporating a beneficial feature in the simple cost-loss model would decrease the economic value of the forecast.

An example of an beneficial feature is the *false alarm intolerance* stated in the article 'The boy who cried wolf revisited: The impact of false alarm intolerance on cost-loss scenarios' by Roulston and Smith (2004) [12]. In this article they refer to Aesop's fable 'The boy who cried wolf' in which a boy cries every time he thinks to hear a wolf. However, this happens so often that the inhabitants of the village doubt his warnings and do not respond correctly when he cries again. This is stated as a false alarm intolerance.

This article also looks at the behaviour of people. It notes that whether an individual will leave his or her home based on some extreme weather forecasts, like for example the forecast of a hurricane, may depend on what his or her neighbours are doing. Because the false alarm intolerance is time dependent, Roulston and Smith even state that the optimum warning threshold will be time dependent.

In a paper by Murphy and Katz (1985) they summarized some extensions of the cost-loss model [13]. They suggest that the cost-loss model should be rebuilt with a different structure. Since part of the losses may be recoverable over time, this should also be implemented in the cost-loss model. They also note that focussing on minimizing the expected expenses not always is what the decision maker wants. It may be more important to the decision maker to postpone a potential loss. Also the attitude of the decision maker is important, as is also stated by Lazo [5] in the beginning of Chapter 3. This can be done by maximizing the expected utility instead of minimizing the expected expenses.

These articles all give disadvantages of the use of the simple cost-loss model. However, the cost-loss model can be considered as a support tool in making the optimal decision. The decision maker himself may decide how he uses this information. In this way he can still involve for example behavioural characteristics.

Chapter 4

Sensitivity experiment

Equation (3.4.5) expresses the economic value V in terms of the cost-loss ratio $\alpha = C/L_1$. Suppose a company gives us its cost-loss ratio α. It wants to know which probability threshold to select to obtain the most value from the forecast. The simplest cost-loss model states that when the probability of adverse weather is greater than this threshold, the company should take protective action. In selecting the right probability threshold the question rises how sensitive a threshold is given the skill of a weather forecast. This chapter will look at weather forecasts with three different artificial skills. For a given cost-loss ratio these forecasts give a threshold which is supposed to be optimal for a company. The calculations are derived from an article by D.S. Richardson [8].

4.1 Data experiment

The set `forecast` consists of n precipitation forecasts representing the precipitation probabilities for n days. According to the cost-loss model, a company should use a probability threshold p_t to take protective action when the precipitation probability is greater than this chosen threshold p_t. Which days the company takes protective action is stated in the set `forecastpt`. Consider those days having a higher precipitation probability than p_t as days at which the company takes protective action. These days are in `forecastpt` represented by 1. Consider those days having a lower precipitation probability than p_t as days at which the company does not take protective action. These days are represented by 0. This construction gives the set `forecastpt` consisting of zeros and ones dependent on p_t and on the data in `forecast`.

Also the real precipitation data is given. This set is called `observed` and contains the observed precipitation for the same n days. Per day it denotes 1 if precipitation occurred and 0 if not. The set `forecastpt` gives an impression of the economic value of the forecast for the company. This economic value is calculated as presented in Chapter 3. For a given p_t, first the corresponding contingency table is constructed. With this contingency table, the hit rate and false alarm rate can be calculated whereby the economic value of the forecast V can be calculated. This reveals the economic value for a forecast with a particular skill and a given threshold p_t.

To conclude how sensitive the choice of the threshold is, this threshold p_t will be varied from 0 to 1 where the forecast probabilities and observations are fixed. Each of these thresholds will yield their own contingency table, hit rate H, false alarm rate F and economic value V. From all these thresholds p_t, the company can select the threshold which results in the highest economic value V. Since this economic value V depends on the fraction of occasions $\bar{o}$ and the cost-loss ratio α, this optimal threshold p_t may differ for different companies, different weather events and differences in skill of the forecast. When the company decides to completely rely on the cost-loss model, the company will take precautionary protective action when the probability of precipitation for that day is greater then the chosen threshold p_t. This method converts the probabilistic forecasts, given the probabilistic thresholds p_t, to a decision model.

4.2 Example construction contingency table

In the following example the number of days n is equal to five and the data available is stated in Table 4.1.

Day	1	2	3	4	5
`forecast`	0.73	0.07	0.23	0.88	0.63
Prob. categorized `forecastpt`	1	0	1	1	1
`observed`	0	0	1	1	1

Table 4.1: *Precipitation data for five days. The data in* `forecast` *is the probability of precipitation for each day. The probability threshold* p_t *is chosen to be* 0.2.

The data in `forecastpt` consists of ones on the days that the probability of precipitation is larger than 0.2 and zeros elsewhere. Therefore, it may be interpreted as the days at which the company takes precautionary action given the weather forecast, if the probability threshold for this company equals $p_t = 0.2$.

For probability threshold $p_t = 0.2$, Table 4.1 shows that on day one precipitation was forecast, but did not occur. On day two precipitation was not forecast and did not occur. On day three, four and five precipitation was forecast and did occur. This information places the five days in the contingency table stated in Section 3.2. Table 4.2 shows the contingency table for this example where the probability threshold p_t equals 0.2.

The contingency table given in Table 4.2 implies $\bar{o} = 3/5$, $H = \frac{3}{5}/\frac{3}{5} = 1$ and $F = \frac{1}{5}/(\frac{1}{5} + \frac{1}{5}) = 0.5$. With this information and Equation (3.4.5) the economic value V can be calculated as a function of cost-loss ratio α. For this example with

		Observed: Precipitation	Observed: No Precipitation	Total
Forecast	Action	3/5	1/5	4/5
	No Action	0	1/5	1/5
	Total	$\bar{o} = 3/5$	$1 - \bar{o} = 2/5$	1

Table 4.2: *Contingency table given five precipitation forecasts and precipitation occurrences given in Table 4.1. The probability threshold p_t equals* 0.2.

$n = 5$ and $p_t = 0.2$ this implies:

$$V = \frac{\min(\alpha, \bar{o}) - F\alpha(1-\bar{o}) + H\bar{o}(1-\alpha) - \bar{o}}{\min(\alpha, \bar{o}) - \bar{o}\alpha}$$
$$= \frac{\min(\alpha, 3/5) - 1/5\alpha + 3/5(1-\alpha) - 3/5}{\min(\alpha, 3/5) - 3/5\alpha}.$$

This calculation for V can be done for many different thresholds p_t. For each threshold, first the contingency table is constructed, then H, F and $\bar{o}$ are calculated and then the economic value V is determined as a function of cost-loss ratio α. This economic value V can be plotted for all thresholds $p_t \in \{0, 0.1, \ldots, 0.9, 1\}$. This visually advises which threshold p_t a company should select to gain the highest economic value V.

4.3 Artificial forecast skills

A sensitivity experiment is executed based on the example mentioned in the previous section. The elements in the dataset `observed` are as described in the previous example, but now for $n = 100$ days. The elements are fixed for all sensitivity experiments. To investigate how sensitive the thresholds are given the skill of a weather forecast, forecasts with three different skills will be presented. These are forecasts with no skill, little skill and high skill.

No Skill The data in `forecast` for forecasts with no skill contains 100 random numbers between zero and one. Since this is totally independent of de data in `observed`, this forecast data has completely no skill. For clarification the forecast data in `forecast` is constructed as follows:

$$\text{for } i \in \{1, \ldots, n\} : \texttt{forecast}[i] = x,\ x \in [0,1],\ x \text{ random}.$$

Little skill The data in `forecast` for forecasts with little skill is partly depending on the observed data. If on day $i \in \{1, \ldots, 100\}$ the `observed` data states that there is no precipitation, then on day i in `forecast` a random number between zero and one squared is assumed, rounded to two decimals. When the `observed` data states there is precipitation, then, identical to the forecasts with no skill, on day i in `forecast` a random number between zero and one is assumed. This gives little skill to the forecast on the days there is no precipitation observed. For clarification the forecast data in `forecast` is constructed as follows:

$$\text{for } i \in \{1,\ldots,n\} : \text{if } \texttt{observed}[i] = 0 : \texttt{forecast}[i] = x^2,\ x \in [0,1],\ x \text{ random.}$$
$$\text{if } \texttt{observed}[i] = 1 : \texttt{forecast}[i] = x,\ \ x \in [0,1],\ x \text{ random.}$$

High skill The data in `forecast` for forecasts with high skill is most linked to the observed data. If, on day $i \in \{1, \ldots, n\}$ the `observed` data states there is no precipitation, then the same is done as in the forecast data with little skill. When, on day i, the `observed` data states that there is precipitation, then on day i in `forecast` there is one minus a random number between zero and one squared assumed. This also gives skill in the forecast on the days the observed data states that there is precipitation. Therefore, this forecast dataset has high skill. For clarification the forecast data is constructed as follows:

$$\text{for } i \in \{1,\ldots,n\} : \text{if } \texttt{observed}[i] = 0 : \texttt{forecast}[i] = x^2,\qquad x \in [0,1],\ x \text{ random.}$$
$$\text{if } \texttt{observed}[i] = 1 : \texttt{forecast}[i] = 1 - x^2,\ x \in [0,1],\ x \text{ random.}$$

For these three types of skill, three different forecast data sets are investigated. This is done to also provide uncertainty for a fixed skill. These are constructed as described in the skill descriptions. For each skill the economic value V for $p_t \in \{0, 0.1, 0.2, 0.3, \ldots, 0.9, 1.0\}$ will be calculated as described in the example in the previous section. For each skill and for each threshold p_t first the corresponding contingency table is constructed, then H, F and $\bar{o}$ are calculated and then the economic value V. After we plot V for each threshold p_t for all different skills, it can visually show to a company with cost-loss ratio α which threshold gives the highest economic value V.

The Hit rate H and the False alarm rate F give some indication about the skill of the forecast. The higher the hit rate and the lower the false alarm rate, the better the forecast. Therefore the Relative Operating Characteristics (ROC) is a visual indication of the skill of the forecast. An example can be found in the upper left figure in Figure 4.1. Here, the false alarm rate is set against the hit rate for all thresholds p_t. The surface A below the ROC plot is a numeric indication of the skill of the forecast. If A is close to 1 the forecast is considered to be good and when A equals 0.5, the forecast is not better than using climatological information.

4.4 Calculation perfectly reliable forecast

The forecast has the most value when the points in the ROC plot are close to the upper left corner. Then, the forecast system has the highest hit rate and lowest false alarm rate. However, weather forecasts are not perfect and therefore the ROC will never reach this upper left corner. Therefore, we also plot the most reliable ROC based on the forecast data. The calculation of this most reliable ROC is done as follows.

Let the function $g(p)$ denote the frequency in which our forecast data given in `forecast` forecasts precipitation with probability p. In our example in Section 4.2

it holds that $g(p) = 1$ for all p in `forecast`. Let $p'(p)$ denote the frequency at which the event occurs when the forecast probability is p. This can be found in the dataset `observed`. In our example in Section 4.2 we have $p'(0.73) = p'(0.23) = p'(0.88) = p'(0.63) = 1$ and $p'(0.07) = 0$. In a perfectly reliable forecast it would hold that $p'(p) = p$. Let S be the set of possible thresholds, thus $S = \{0.0, 0.1, \ldots, 0.9, 1.0\}$. Then, the hit rate and false alarm rate for the perfectly reliable forecast, H_{p_t} and F_{p_t} [8], would be

$$
\begin{aligned}
H_{pt} &= \left[\sum_{p\in S, p\geq p_t} p'(p)g(p)\right] \Bigg/ \left[\sum_S p'(p)g(p)\right] \\
&= \left[\sum_{p\in S, p\geq p_t} p\cdot g(p)\right] \Bigg/ \left[\sum_S p\cdot g(p)\right], \\
F_{pt} &= \left[\sum_{p\in S, p\geq p_t} (1-p'(p))g(p)\right] \Bigg/ \left[\sum_S (1-p'(p))g(p)\right] \\
&= \left[\sum_{p\in S, p\geq p_t} (1-p)g(p)\right] \Bigg/ \left[\sum_S (1-p)g(p)\right],
\end{aligned}
$$

This hit rate H_{pt} and false alarm rate F_{p_t} based on a perfectly reliable forecast give the economic value V_{p_t}. This is the perfectly reliable economic value. They also give the perfectly reliable ROC.

Figures 4.1, 4.2 and 4.3 show the output of the experiment. Each figure represents a skill and expresses three different forecast data based on this skill. The observations are equal for all skills. The left side of each figure represents the Relative Operating Characteristics (ROC) with the ROC based on the forecast as the black line and based on the perfectly reliable forecast given in red. This perfectly reliable ROC is denoted in the legend as 'perfect' ROC. The right side of each figure represents the economic value V. The economic value V given each probability threshold p_t is given in different colours, as stated in the legend of each figure. The maximum economic value V reached, given all the probability thresholds p_t, is displayed as the thick black line. The thick black dashed line represents the maximal economic V reached by the perfectly reliable economic value.

4.5 Analysis of sensitivity experiment

Figure 4.1, 4.2 and 4.3 show that the more skill the forecasts have, the greater the area A under the ROC curve is. This also holds to how close the black line representing the ROC of the forecast lies to the most reliable ROC curve.

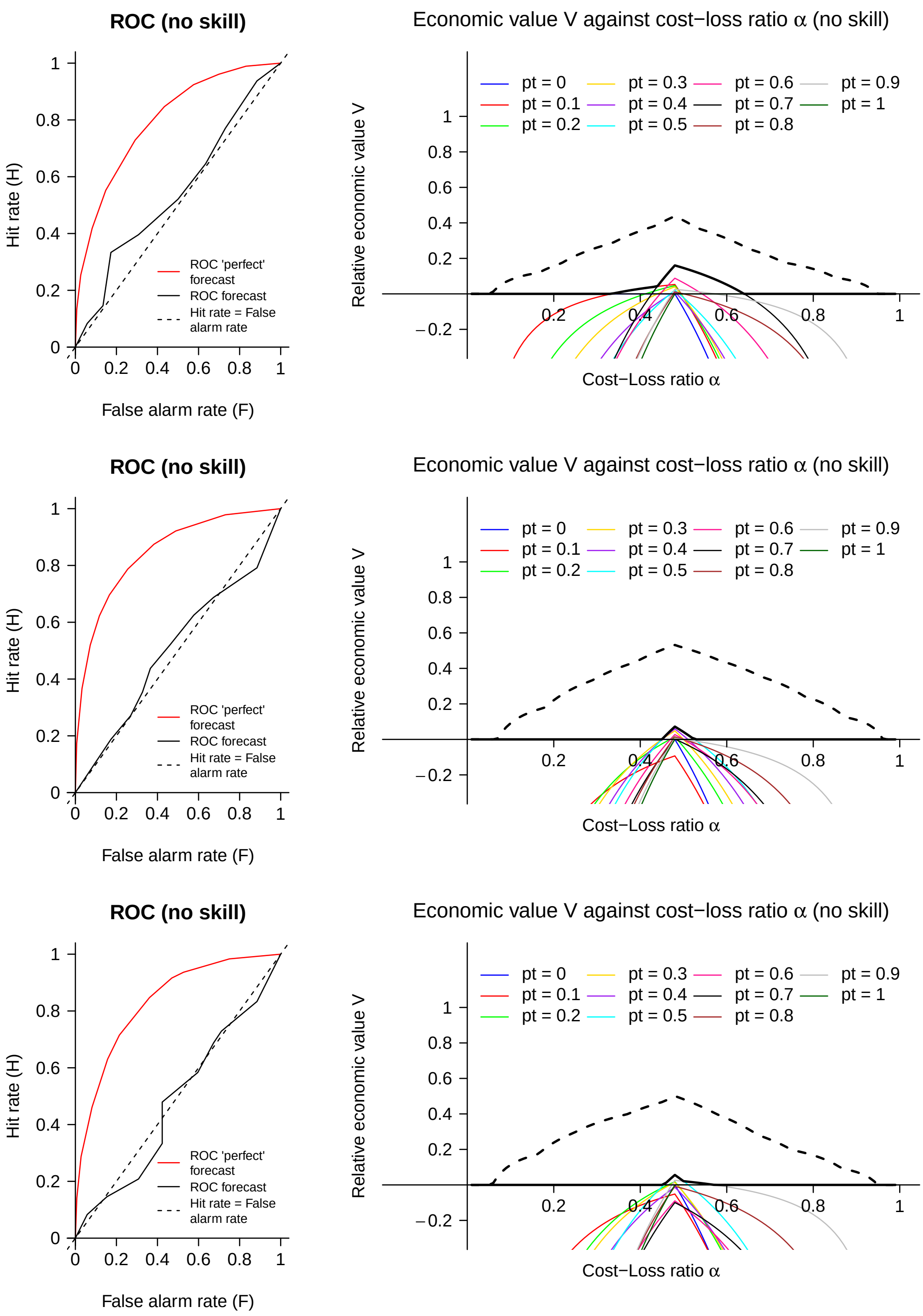

Figure 4.1: *Three different data sets with no skill with corresponding ROC (left) and corresponding economic value V (right). The corresponding areas under the ROC's A are respectively:* $A = 0.5475$, $A = 0.5066$ *and* $A = 0.4810$.

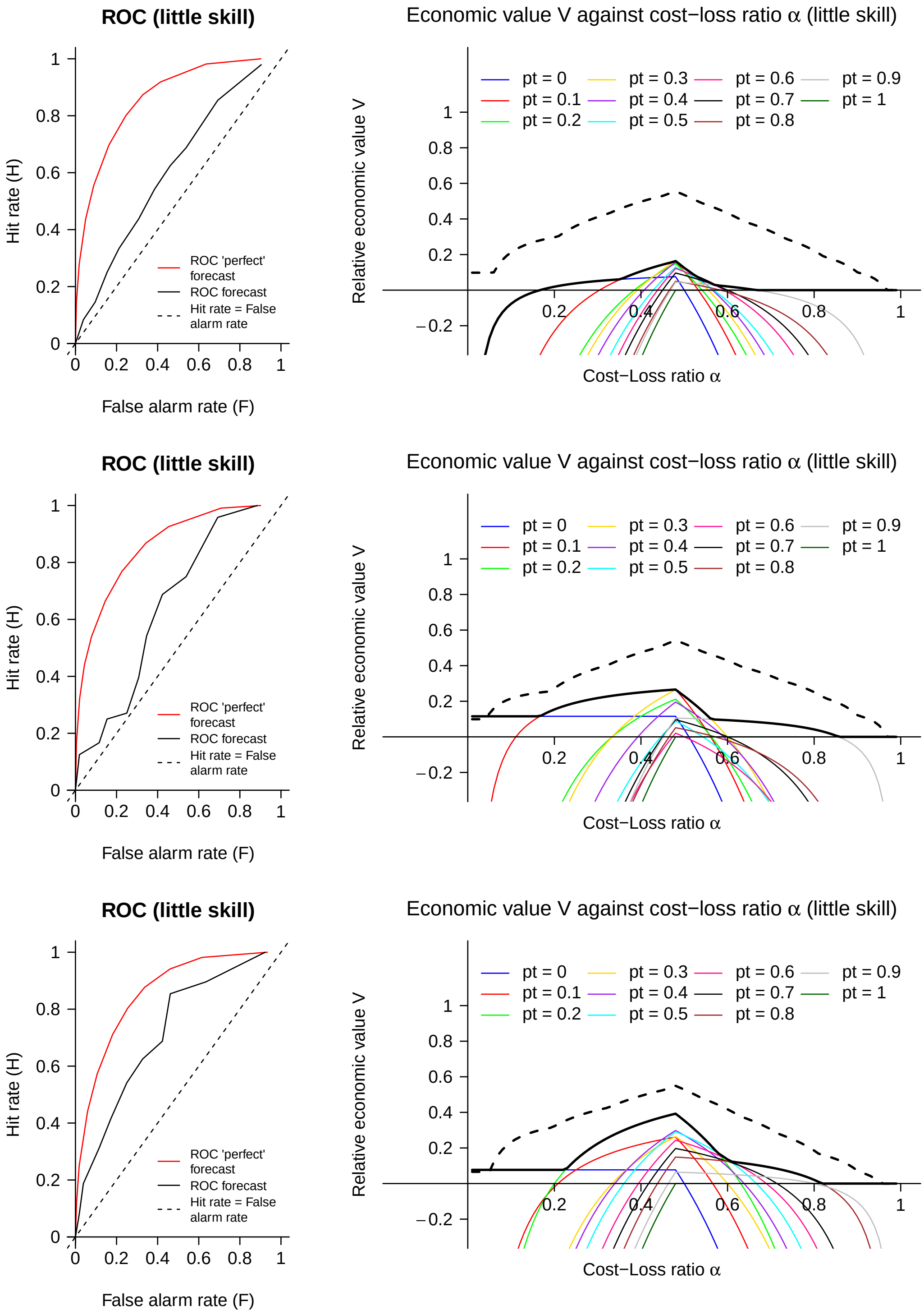

Figure 4.2: *Three different data sets with little skill with corresponding ROC (left) and corresponding economic value V (right). The corresponding areas under the ROC's A are respectively:* $A = 0.5190, A = 0.5355$ *and* $A = 0.6430$.

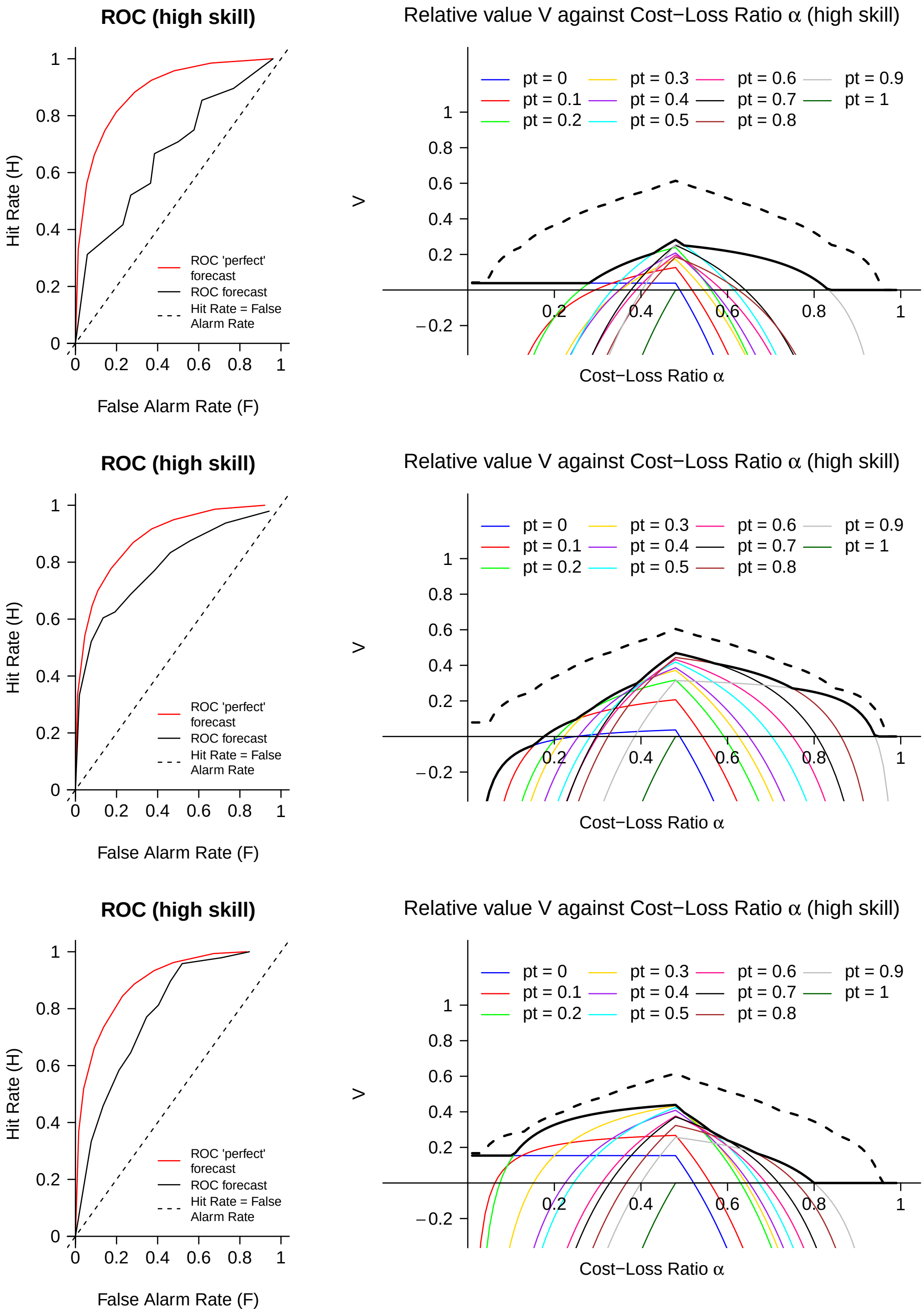

Figure 4.3: *Three different data sets with high skill with corresponding ROC (left) and corresponding economic value V (right). The corresponding areas under the ROC's A are respectively:* $A = 0.6372$, $A = 0.7338$ *and* $A = 0.6336$.

However, these differences are not significant as can be seen in the captions of the figures: the third dataset in Figure 4.2 has an area under the ROC of $A = 0.6430$, because the first dataset in Figure 4.3 has an area under the ROC of $A = 0.6372$. Therefore, even though the forecasts are constructed to differ in skill, the forecasts with these skills do not have to differ significantly.

Suppose a company has a cost-loss ratio α equal to 0.6. When the forecast data has no skill, Figure 4.1 shows that the optimal threshold for this company is $p_t = 0.7$ or $p_t = 1.0$ where the maximum economic value V ranges between 0 and 0.055. Figure 4.2 shows that the optimal threshold for this company is $p_t = 0.9$ or $p_t = 0.5$ where the maximum economic value V ranges between 0.021 and 0.136. Figure 4.3 shows that the optimal threshold for this company is $p_t = 0.9$, $p_t = 0.8$ or $p_t = 0.7$ where the maximum economic value V ranges between 0.218 and 0.396. So the optimal probability threshold for this company ranges between $p_t = 0.5$ and $p_t = 1.0$ and the maximum value of V ranges between 0 and 0.396. Therefore, we should advise the company to select a threshold p_t between 0.5 and 1, since we do not know the skill of the forecast beforehand. In reality, most of the time, the cost-loss ratio of the company are also not exactly known, since the costs and losses of taking precautionary actions are often not exactly known. This also favors a range of probability thresholds, given a range of cost-loss ratios, which together should gain the most economic value V. With the threshold chosen, the company can gain an economic value between 0 and 0.396, which is also quite some range. Therefore, we have a lot of sensitivity in the thresholds given the forecast which makes it harder to advise a company given its cost-loss ratio.

4.6 Conclusions sensitivity experiment

When a company with a cost-loss ratio α needs advise about which probabilistic threshold it should choose to gain the most economic value of the forecast, it is necessary to give a range of possible thresholds. This range can be large when the skill of the forecast is unknown. Since forecasts nowadays have quite some skill, the range of thresholds that give the most economic value for a company may not be as large as the sensitivity experiment suggests. However, it is still more valuable to give the company a range of possible thresholds that could give the most economic value of the forecast. In this way, more uncertainty in the skill of the forecasts is taken into account, so that the company may choose the right precautionary actions given the forecast.

Chapter 5

Determining strategy

The precipitation data of Friesland for the winters in 2012 to 2015 will be analysed to come up with a strategy for our simplified model indicating when Friesland should operate pumps and open sluices. However, since the water board in Friesland has many pumps and sluices, we will look at a simplified model for the water board. This model has only one pump, Hoogland, available. It has as many sluices as the original water board in Friesland, but in this case the sluices can only be opened all at the same time. For this simplified model of the water board we will construct a strategy which determines at which hours sluices need to be opened and Hoogland needs to operate. First, it is discussed what data is available and what the assumptions are for the strategy and for the water board.

5.1 Data available

First, there is data of Friesland for the winters in 2012 to 2015. This dataset consists of precipitation forecasts from EPS and from the deterministic ECMWF model, observed precipitation and surges. The EPS data denotes of days in the winters of 2012 to 2015 the forecast precipitation for ten days ahead. This EPS consists of 51 ensemble members, which together provide insight into the uncertainty in forecasting. All the members denote the precipitation per six hours and are expressed every tenth of a millimetre. Besides, also the high-resolution deterministic ECMWF model precipitation output is used.

The observed precipitation data denotes for days in the winters of 2012 to 2015 the observed precipitation in tenth millimetres. Also here the data displays precipitation per six hours. This dataset is based on calibrated radar data [14].

The surge data denotes the water level of the Wadden Sea, the sea connected to Friesland. This surge data also consists of 51 ensemble members and a deterministic run, which are calculated from EPS and the deterministic ECMWF model. When the water level of the Wadden Sea is below the water level of the Frisian belt canal system, it is possible to sluice water away. Since sluicing is cheapest, this is the most attractive way to drain water.

5.2 Assumptions simplified model

It is assumed that the sluice can lower the water level of the Frisian belt canal system by 0.21 cm per hour. This is approximately the total amount that the sluices in Friesland can sluice per hour. This can be deduced from Table 2.1. The maximal capacity of the sluices is equal to $5.5 + 0.9 + 0.55 \cdot 2 = 7.5$ million cubic metres per day. Since the Frisian belt canal system has a surface of 15,000 hectare, this should be divided by 1.5 and by 24 to get the influence in centimetres per hour of the sluices on the belt canal system. This results in approximately 0.21 cm per hour. Only one sluice possibility is assumed; when the sluices are opened, they can sluice 0.21 cm per hour to the IJsselmeer and Wadden Sea.

It is also assumed that the Hoogland pumping station can pump away 0.25 cm from the Frisian belt canal system per hour. This is approximately the total amount of centimetres that the Hoogland pump in Friesland can pump away per hour. This can also be deduced from Table 2.1. The maximum capacity of Hoogland is equal to 8.9 million cubic metres per day. Since the belt canal system has a surface of 1.5 million cubic metres, this should again be divided by 1.5 and by 24 to get the influence in centimetres of the pumps on the belt canal system. This results in approximately 0.25 cm per hour. There presence of only one pump is assumed, Hoogland, which can pump away 0.25 cm per hour.

It is also assumed that the sluice level is equal to -40 cm, which implies that sluices can drain water when the water level of the Wadden Sea is lower than -40 cm. Which hours this happens depends on the surge level of the Wadden Sea which is given by our data. In Friesland the sluice level is dependent on the amount of water that is drained by sluicing or pumping, thus in their case this sluice level is variable. For our simplified model for the water board we set the sluice level as a constant.

Friesland can also decide to let water into the belt canal system if the water level is too low. Therefore, in this experiment, it is assumed for our simplified model for the water board, that the begin water level for each day will be fixed and greater or equal to -52 cm, which equals the target level (the green line in Figure 2.5). Since letting in water is cheap, the hours at which water is let in are not specified. For our simplified model for the water board, it is more important at which hours sluices need to be opened and Hoogland needs to operate.

Executing this strategy, we want to investigate whether there is a difference between the probabilistic chosen actions and the deterministic chosen actions. Taking action is in this case defined as operating Hoogland. For each day in the winters of 2012 to 2015, the expected water level is calculated based on either the deterministic run or the EPS run. The experiment is executed for different fixed begin water levels, namely -52, -50, -48, -46 and -44. Because we are most interested in the differences between the use of probabilistic and deterministic information in the short term, we fill first execute this experiment for a forecast period of 48 hours. To also look at these differences in the long term, we will repeat the experiment for a forecast period of 120 hours.

Figure 5.1 shows the simplifications in our model compared to the Frisian decision model. One difference between the model of Wetterskip Fryslân and our model is the way in which the expected discharge is calculated, as has been discussed in Section 2.3. With this expected discharge, the expected water level in the Frisian belt canal system can be calculated. Based on this expected water level, the water board may decide to open sluices or operate pumps. During our simplified model, we will only use one pump (Hoogland) and one sluice possibility, while Friesland has many more options to drain water. When there is known which hours the sluices are opened and which hours pumps are operated, the observed water level can be calculated. This observed water level is also based on the observed precipitation and the observed discharge, which is calculated by Equation (2.2.1).

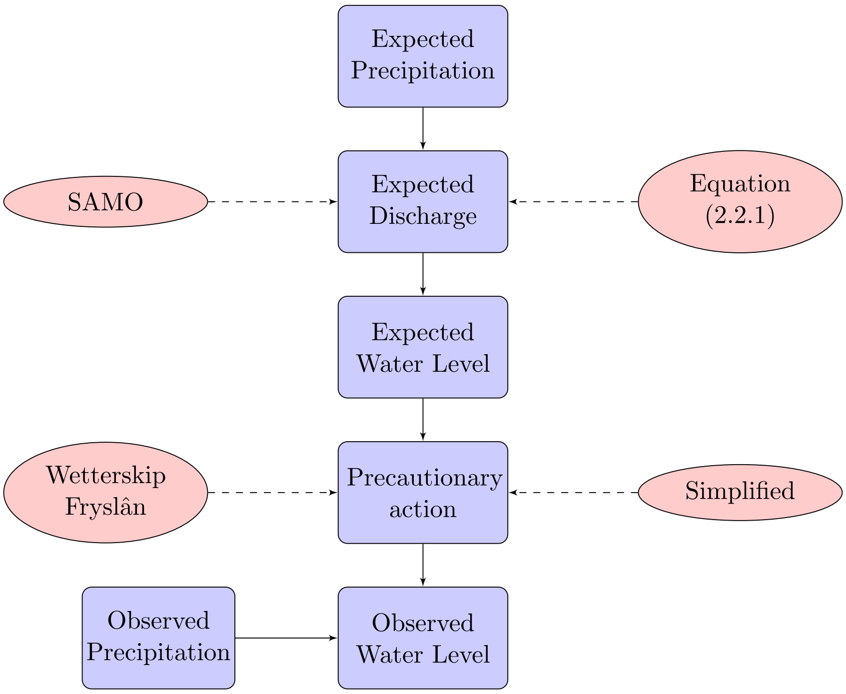

Figure 5.1: Model used by Wetterskip Fryslân compared to the simplified model for our experiment.

5.3 Construction strategy

From the EPS precipitation forecast data, the expected discharge to the Frisian belt canal system is calculated. This is done with the simplified equation given by Equation (2.2.1). Since the ground during the winter is considered saturated, which implies that much of the precipitation runs off directly, the weight factor in Equation (2.2.1) is set equal to 0.8. Note that since the ECMWF data consists of one deterministic run and 51 ensemble runs, this results in one deterministic expected water level and 51 probabilistic expected water levels. With these expected water levels the simplified model implies when to operate Hoogland and open sluices. Recall the calculated expected water level of Friesland on January 7 2015 with the corresponding target levels which can be found in Figure 2.5.

The goal is to investigate whether, and if so, under what condition, the use of probabilistic information can lead to better decisions than the use of deterministic information in the short term and long term. Therefore we will compare the action chosen by the deterministic run (action denotes here whether Hoogland needs to operate) with the action chosen by the probabilistic runs during a forecast period of 48 and 120 hours. First a strategy for the simplified model is constructed that tells when the water board should operate Hoogland and open sluices.

The strategy is as follows:

1. For a day in the winters of 2012 to 2015 we can calculated the expected water level. Since opening sluices is cheap, if necessary, the sluices are first opened to come as close to the target level (-52 cm) after the first 24 hours as possible. This with the exception that the water level must never become lower than the lower critical level (bottom thick red line in Figure 2.5). After that, there results an expected water level with possibly opening sluices. Given this expected water level, there is a criterion given which decides whether Hoogland needs to operate or not.

2. Now we can calculate the observed water level for that day given by the observed precipitation data. If it is decided that Hoogland needs to be turned on, the hours at which Hoogland operates, is determined backwards in time. At the time where the Hoogland threshold is exceeded, Hoogland is turned on and it is checked backwards in time per hour whether Hoogland needs to operate to get the observed water level below the Hoogland threshold. Hoogland can operate from six hours, since that is the first hour at which the data of the deterministic run is available. The advantage of determining, the hours at which Hoogland operates, backwards in time is that costs of pumping are postponed as long as possible. If the observed water level drops below the Hoogland threshold, we will pump another four hours backwards in time before we stop Hoogland. This is a buffer for discharge that may come into the Frisian belt canal system for the upcoming days. If it is decided that Hoogland not needs to be turned on, then Hoogland will not be operated. After these steps, it is known for the observed water level whether Hoogland is turned on (action) or not (no action) and we know the number of hours the sluices were opened and the number of hours Hoogland operated. In this way, we get the observed water level with the action chosen by the expected water level. It is considered a loss if the observed water level exceeds the Hoogland threshold. The hours that the sluices were opened and the number of hours at which Hoogland was operated were already known, and now it is also known whether a loss occurred.

Part one of this strategy is executed on the deterministic precipitation runs from 377 days in the winters of 2012 to 2015. The criterion for the deterministic precipitation run to operate Hoogland, is that the expected water level must exceed the Hoogland threshold. Then part two of the strategy can be executed, which results in 377 deterministically chosen actions of whether Hoogland is operated not. For each day, there are 51 observed water levels and one deterministic observed water level, coming from the surge data. We will choose the second of the 51 observed water levels. While the perturbations in the beginning of the ensemble members are random, the second water level is a random choice. Now we know whether the observed water levels after opening sluices and operating pumps contained Hoogland exceedings or not. Therefore we can fill in each day of the winters of 2012 to 2015 in the contingency table as stated in Section 3.2.

The criterion for the probabilistic precipitation runs to operate Hoogland is different, since we can calculate 51 expected water level given by 51 ensemble members. For each of these ensemble members, part one of the strategy is executed. This leads to 51 expected water levels. We will calculate the percentage of ensemble members for which the expected water level exceeds the Hoogland threshold. For probability targets $p_t \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ we will operate Hoogland if this percentage of exceedings is above p_t and not otherwise. For each probability target p_t, part two of the strategy can be executed which now results in 377 probabilistically chosen actions. For each probability target p_t, we can again fill in these days of the winters in 2012 to 2015 in the contingency table as stated in Section 3.2.

Now we have obtained for each day in the winters of 2012 to 2015 probabilistically chosen actions and a deterministically chosen action. With this information, we can construct contingency tables corresponding to the deterministic run and probabilistic runs. For the probabilistic runs, we can construct contingency tables for each probability threshold $p_t \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. For each day in the winters of 2012 to 2015, we know whether the probabilistic run with a fixed probability threshold choose to take protective action or not and whether the Hoogland threshold was still exceeded after the chosen action. Therefore we can fill in each day of the winters of 2012 to 2015 in the contingency table. This can also be done for the deterministic run.

However, these contingency tables are not the same as we know them from Chapter 3. This will be explained and discussed in the next section.

5.4 Reinterpretation contingency tables

The contingency tables resulting from our experiment are not the same as those discussed in Chapter 3. This is because in our experiment, we have influence on whether the observed water level will exceed the Hoogland threshold. In the contingency tables in Section 3.2, the observations given in the columns are observations in adverse weather. Because we cannot influence weather, we can place each event from the data in the contingency table, independently of whether to take precautionary action. Because this is not the case for our experiment, the quantities in the contingency tables also have different meanings. Table 5.1 shows the contingency table as it should be interpreted from our experiment.

In Table 5.1 a denotes days for which Hoogland operates, but still in the observations the Hoogland threshold was exceeded. There are two possibilities. Hoogland operates less than it should be, caused by the observed precipitation being higher

		Observed: Exceed Hoogland threshold	Observed: Not exceed Hoogland threshold	Total
Action	Operate Hoogland	a	b	$a+b$
	Not operate Hoogland	c	d	$c+d$
	Total	$a+c$	$b+d$	$a+b+c+d$

Table 5.1: *Contingency table given dates in the winters of 2012 to 2015 based on the strategy described before. a, b, c and d are not expressed as fractions, but as integers.*

than the expected precipitation, or Hoogland operated as long as it could operate, but it could still not prevent the exceeding of the Hoogland threshold. This last statement states unavoidable losses, while the first statement states an underestimation of the expected precipitation.

In Table 5.1 b denotes the days for which Hoogland operates, and in the observations the Hoogland threshold was not exceeded. In this case, operating Hoogland prevented a loss in the future, or Hoogland was operated unnecessary, which happened if the Hoogland threshold would not have been exceeded, even when Hoogland was not operated. Thus a part in b states unnecessarily operating Hoogland and a part states the right moments of operating Hoogland, where it prevented Hoogland exceedings in the future.

In Table 5.1 c denotes the days for which Hoogland was not operated, but in the observations the Hoogland threshold was exceeded. This means, on the days in this stage, Hoogland was not operated while it should be. This has the same meaning as the fraction c in the original contingency table, since not operating Hoogland has no influence on the observed water level.

In Table 5.1 d denotes the days for which Hoogland was not operated and in the observations the Hoogland threshold was not exceeded. This means that these days the right choices were made of not operating Hoogland. This has the same meaning as the fraction d in the original contingency table, since not operating Hoogland has no influence on the observed water level.

As the contingency table given in Table 5.1 has a different interpretation compared to the contingency tables given in Section 3.2, it is not of any use to calculate the hit rate and false alarm rate. Therefore our conclusions will be made based on the complete contingency tables that appear from the given data.

Chapter 6

Results

For 377 days in the winters of 2012 to 2015 we obtained results from the observed water levels. We filled in each day in the contingency table presented in Section 5.4. We also determined the total hours of the winters of 2012 to 2015 on which the sluices were opened and Hoogland was operated.

Appendix A.1 gives an example of the contingency tables that result from our computations. For these contingency tables the begin water level equals -50 cm and the forecast period equals 48 hours.

For begin water levels equal to -52, -50, -48, -46 and -44, we will compare the probabilistic contingency tables for each threshold $p_t \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ with the deterministic contingency table. Recall that the target level equals -52 cm and the Hoogland threshold equals -46 cm.

For each begin water level, the total hours the sluices were opened given by the probabilistic contingency tables and the deterministic contingency table were equal. Therefore, only the hours at which Hoogland operated are interesting in comparing the hours that precautionary action was taken.

We will first look at the results of the short term forecasts, where the forecast period equals 48 hours, and conclude whether the use of probabilistic information would lead to better decisions. Then we will look at the results of the long term forecasts, where the forecast period equals 120 hours, and also draw conclusions.

The maximum number of hours that Hoogland can operate, are for the short term (48 hours) equal to $(48 - 6) \cdot 377 = 15.834$. This is since we have 377 days in our dataset and operating Hoogland the first six hours is not possible. For the long term (120 hours) this amount equals $(120 - 6) \cdot 377 = 42.978$. This is important to note when we look at how many hours Hoogland had to operate more or less in the probabilistic run compared to the deterministic run, as it gives an idea whether the differences between the deterministic and probabilistic contingency tables are significant.

6.1 Results for short term forecasts

Table 6.1 compares the output of the experiment using probabilistic information and deterministic information. It shows for each begin water level and each threshold how many hours Hoogland was operated more or less by the use of probabilistic information compared to the use of deterministic information and how many Hoogland exceedings were more or less by the use of probabilistic information compared to the use of deterministic information.

Begin water level	p_t	*Difference hours*	*Difference exceedings*	*cost-loss range*
−52	0.0	+22	−9	$0 < C/L < 9/22 \approx 0.409$
	0.2	+16	−7	$0 < C/L < 7/16 \approx 0.438$
	0.4	+3	−3	$0 < C/L < 3/3 = 1$
	0.6	−21	+1	$1/21 \approx 0.048 < C/L < 1$
	0.8	−82	+8	$8/82 \approx 0.098 < C/L < 1$
	1.0	−1235	+65	$65/1235 \approx 0.053 < C/L < 1$
−50	0.0	+26	−5	$0 < C/L < 5/26 \approx 0.192$
	0.2	+12	−3	$0 < C/L < 3/12 = 0.25$
	0.4	−3	0	Probabilistic better
	0.6	−18	+4	$4/18 \approx 0.222 < C/L < 1$
	0.8	−49	+7	$7/49 \approx 0.143 < C/L < 1$
	1.0	−1971	+98	$98/1971 \approx 0.050 < C/L < 1$
−48	0.0	+51	−10	$0 < C/L < 10/51 \approx 0.196$
	0.2	+23	−6	$0 < C/L < 6/23 \approx 0.261$
	0.4	+23	−6	$0 < C/L < 6/23 \approx 0.261$
	0.6	+2	−1	$0 < C/L < 1/2 = 0.5$
	0.8	−25	+3	$3/12 = 0.25 < C/L < 1$
	1.0	−2979	+128	$128/2979 \approx 0.043 < C/L < 1$
−46	0.0	+3	0	Deterministic better
	0.2	+3	0	Deterministic better
	0.4	0	0	Indifferent
	0.6	−1	0	Probabilistic better
	0.8	−6	0	Probabilistic better
	1.0	−4552	0	Probabilistic better
−44	0.0	0	0	Indifferent
	0.2	0	0	Indifferent
	0.4	0	0	Indifferent
	0.6	0	0	Indifferent
	0.8	0	0	Indifferent
	1.0	−7048	0	Probabilistic better

Table 6.1: *Calculation of cost-loss range at which the use of probabilistic information leads to less expenses than the use of deterministic information. The forecast period is equal to 48 hours.*

Table 6.1 can be used to answer the question whether using probabilistic information may help the decision making process of a water board. Whether it also leads to less expenses depends on the costs of operating Hoogland (C) and the fixed losses (L) that occur when the Hoogland threshold is exceeded. Since we had to reinterpret our contingency tables, it is not that clear when the water board had unavoidable losses and when not. Therefore we will look at the total potential losses L and we will compare these to the costs of operating Hoogland C.

For example (Appendix A.1) for the begin water level of -50 cm, choosing the probability threshold $p_t = 0.0$ and operating Hoogland for 26 hours more, prevented 5 exceedings of the Hoogland threshold. If the costs of operating Hoogland are much less relative to the potential losses that occur when the Hoogland threshold is exceeded (the cost-loss ratio is low), then operating Hoogland, trying to prevent losses, will mostly lead to less expenses. Let C be the costs of operating Hoogland one hour and L be the fixed losses when the Hoogland threshold is exceeded. Then, for our begin water level of -50 cm, if $26 \cdot C < 5 \cdot L$, thus if $0 < C/L < 5/26 \approx 0.192$, the use of probabilistic information leads to less expenses. However, if the costs of operating Hoogland do not differ much from the losses that occur when the Hoogland threshold is exceeded (the cost-loss ratio is high) then this mostly will lead to more expenses. For the begin water level equal to -50 cm, this is the case when $5/26 \approx 0.192 < C/L < 1$. Therefore it depends on the cost-loss ratio whether the use of probabilistic information instead of deterministic information and the action chosen based on this type of information leads to less expenses.

Table 6.1 also shows for the other begin water levels whether the use of probabilistic information instead of deterministic information will lead to less expenses, dependent on the cost-loss ratio C/L of the water board.

Note from 6.1 for the threshold $p_t = 1.0$ that for each begin water level the difference in hours that Hoogland is operated are large, compared to the total number of hours at which Hoogland can be operated of 15.834 hours. This is because this threshold causes that Hoogland never operates. It can be seen for this threshold that the difference in hours and difference in exceedings increases with increasing begin water levels. However, at a begin water level of -46 cm, the Hoogland exceedings for the deterministic and probabilistic contingency tables are equal. This is since a begin water level of -46 cm equals the Hoogland threshold and therefore the Hoogland threshold is easily exceeded, which results in no difference between the exceedings by the use of deterministic information and probabilistic information.

From Table 6.1 it seems that the lowest probability threshold $p_t = 0.0$ is an attractive probability threshold for many begin water levels. However, this threshold implies that Hoogland has to operate even when only one of the 51 ensemble water levels exceeds the Hoogland threshold. Therefore, for many days, Hoogland would operate, which results in a large hour difference compared to the hours Hoogland was operated based on the deterministic contingency table and a lower exceeding difference based on deterministic contingency table. Because Hoogland has zero installation costs, this is not a problem. However, when the installation costs of Hoogland would not be zero, then this would be included in the costs. In this way, the probability target $p_t = 0.0$ may not be that attractive any more.

6.2 Conclusion use probabilistic information for short term forecasts

Figure 6.1 visualizes the cost-loss ranges given in Table 6.1. In this way the water board with a particular cost-loss ratio, or a range of cost-loss ratios, can see for which probability thresholds the use of probabilistic information leads to less expenses than the use of deterministic information.

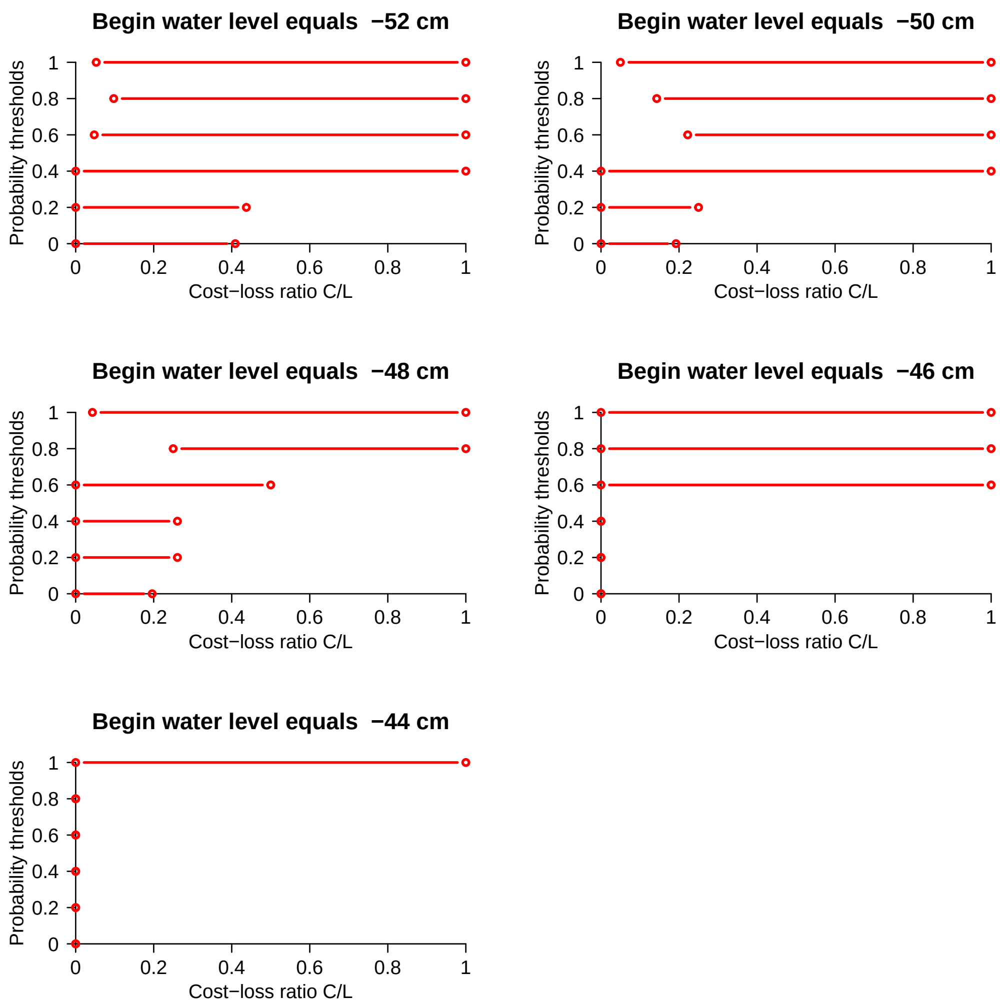

Figure 6.1: *Cost-loss range for which the use of probabilistic information leads to better decisions. The forecast period equals 48 hours.*

Note from Figure 6.1 that for each cost-loss ratio, there is a probability target for which the use of probabilistic information leads to less expenses than the use of deterministic information. This implies that the total cost-loss ratio is covered by the probability thresholds. For a fixed cost-loss ratio, Figure 6.1 often suggests more than one probability threshold. Therefore the water board can choose which probability threshold from this range it finds the most suitable. In this way, the behavioural aspects of the decision makers of the water board can be included; when the water board prefers taking more precautionary action than is necessary, then it should choose the lowest probability threshold from the suggested thresholds. If the water board wants to keep the costs of operating Hoogland as low as possible, even when this will lead to more Hoogland exceedings, then it should choose the highest probability threshold from the suggested thresholds.

Table 6.1 concludes for a begin water level of -50 cm, that with a probability threshold of $p_t = 0.4$, actions based on probabilistic information result in less expenses than based on deterministic information. This can also be seen in Figure 6.1. For the begin water level of -50 cm, the probability threshold $p_t = 0.4$ is a suggested threshold for all cost-loss ratios $0 < C/L < 1$.

Table 6.1 concludes for a begin water level of -46 cm that a probability threshold larger than 0.6 always results in less expenses when the actions are based on probabilistic information. This is since the Hoogland threshold equals -46 cm which implies that the Hoogland threshold is easily exceeded.

Table 6.1 shows that for a begin water level of -44 cm, the expenses based on probabilistic and deterministic information are equal, except for the probability threshold $p_t = 1.0$. Then there is the same amount of Hoogland exceedings for the probabilistic information and deterministic information, but Hoogland was operated 7048 hours less. This however is not an attractive threshold for the water board, because the choice of this threshold will probably increase the water level for the upcoming day, which implies that the probability of large losses in the upcoming days is greater. This depends however on the behaviour of the water board; is the water board risk seeking or risk averting. If it is risk seeking, then it will hope that less precipitation will fall the upcoming days so that the water level will decrease without taking precautionary action. If the water board is risk averse, it wants to take precautionary action, potentially more often than is necessary, and will therefore choose a lower probability threshold than $p_t = 1.0$.

Next the results of the long term forecasts are discussed, where the forecast period equals 120 hours.

6.3 Results for long term forecasts

As for the short term, whether actions determined by probabilistic information would lead to less expenses, depends on the cost-loss ratio. Also here the costs C are the costs of operating Hoogland for one hour and L are the fixed losses when the Hoogland threshold is exceeded. Table 6.2 compares the output of the experiment using probabilistic information and deterministic information for a long term forecast, up to 120 hours ahead.

Appendix A.2 shows the contingency tables for the begin water level of -50 cm. Table 6.2 shows for every begin water level and every threshold the cost-loss ratio range for which the use of probabilistic information leads to less expenses than the

Begin water level	p_t	*Difference hours*	*Difference exceedings*	*cost-loss range*
−52	0.0	+113	−6	$0 < C/L < 6/113 \approx 0.053$
	0.2	+105	−4	$0 < C/L < 4/105 \approx 0.038$
	0.4	+87	−3	$0 < C/L < 3/87 \approx 0.034$
	0.6	+40	+2	Deterministic better
	0.8	−39	+8	$8/39 \approx 0.205 < C/L < 1$
	1.0	−5788	+130	$130/5788 \approx 0.022 < C/L < 1$
−50	0.0	+46	−8	$0 < C/L < 8/46 \approx 0.174$
	0.2	+13	−2	$0 < C/L < 2/13 \approx 0.154$
	0.4	+4	−1	$0 < C/L < 1/4 = 0.25$
	0.6	−49	+3	$3/49 \approx 0.061 < C/L < 1$
	0.8	−113	+7	$7/113 \approx 0.062 < C/L < 1$
	1.0	−7075	+143	$143/7075 \approx 0.020 < C/L < 1$
−48	0.0	+86	−9	$0 < C/L < 9/86 \approx 0.105$
	0.2	+50	−7	$0 < C/L < 7/50 = 0.14$
	0.4	+10	−2	$0 < C/L < 2/10 = 0.2$
	0.6	−7	+1	$1/7 \approx 0.143 < C/L < 1$
	0.8	−95	+8	$8/95 \approx 0.084 < C/L < 1$
	1.0	−8458	+161	$161/8458 \approx 0.019 < C/L < 1$
−46	0.0	+48	−2	$0 < C/L < 2/48 \approx 0.042$
	0.2	+37	−1	$0 < C/L < 1/37 \approx 0.027$
	0.4	+34	−1	$0 < C/L < 1/34 \approx 0.029$
	0.6	+33	−1	$1/33 \approx 0.030 < C/L < 1$
	0.8	−22	+1	$1/22 \approx 0.045 < C/L < 1$
	1.0	−10325	+4	$4/10325 \approx 0.000 < C/L < 1$
−44	0.0	0	0	Indifferent
	0.2	0	0	Indifferent
	0.4	0	0	Indifferent
	0.6	0	0	Indifferent
	0.8	0	0	Indifferent
	1.0	−13039	0	Probabilistic better

Table 6.2: *Calculation of cost-loss range at which the use of probabilistic information leads to less expenses than the use of deterministic information. The forecast period is equal to 120 hours.*

use of deterministic information. The calculation is executed the same way as for short term.

6.4 Conclusion use probabilistic information for long term forecasts

Figure 6.2 visualizes the cost-loss ranges given in Table 6.2. In this way the water board with a particular cost-loss ratio, or a range of cost-loss ratios, can see for which probability thresholds the use of probabilistic information leads to less expenses than the use of deterministic information.

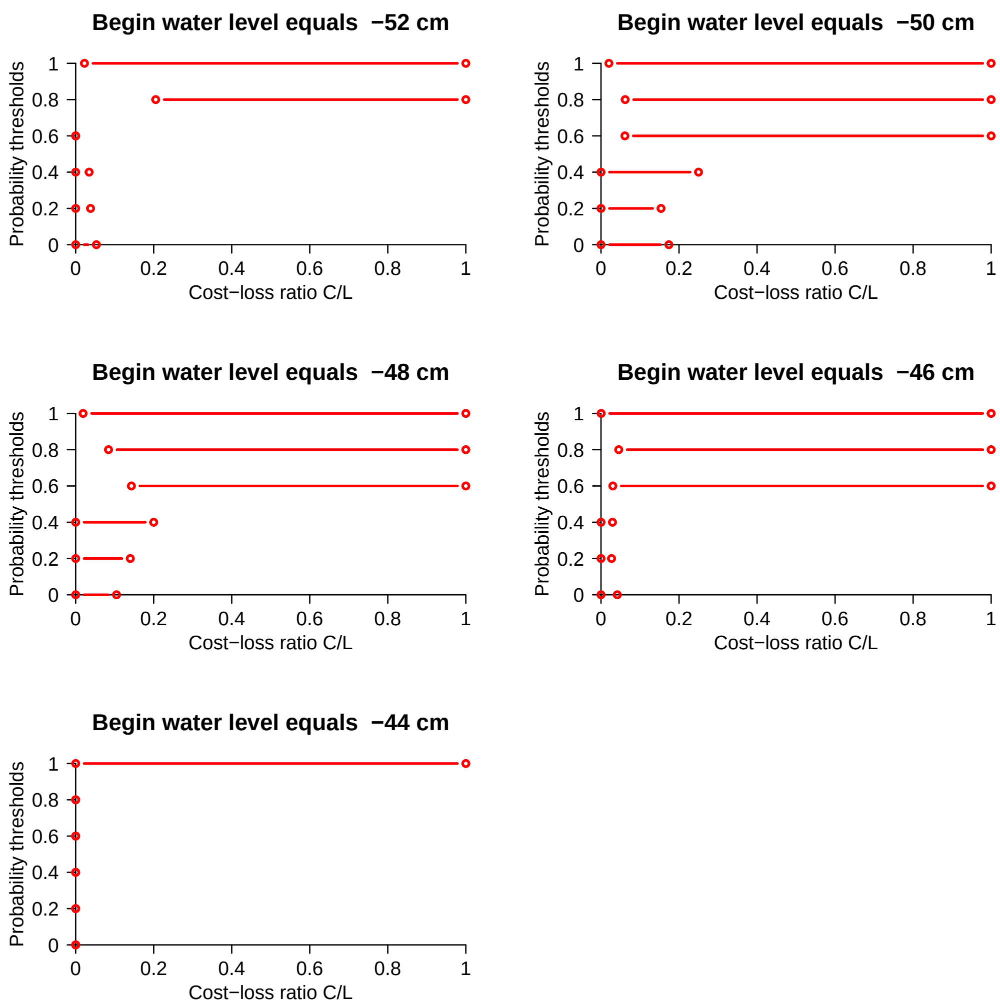

Figure 6.2: *Cost-loss range for which the use of probabilistic information leads to better decisions. The forecast period equals 120 hours.*

Note that Table 6.2 implies, as for the short term, that for each cost-loss ratio there is a probability threshold for which the use of probabilistic information leads to less expenses than the use of deterministic information. Therefore, for the long term, the use of probabilistic information is, for this simplified model of the water board, always more efficient than the use of deterministic information.

We can note differences between Figure 6.1 and Figure 6.2. For the begin water level of −52 cm, for each cost-loss ratio there are less suggested probability thresholds in Figure 6.2 than in Figure 6.1. Especially the target level $p_t = 0.6$ denotes a difference between the short term and the long term forecasts. This is since during the long term forecasts, the probability of exceeding the Hoogland threshold is higher than during the short term forecasts. It can be seen in Figure 6.1 and Figure

6.2 that for the long term, the high probability thresholds are more in favour than for the short term.

Chapter 7

Conclusion

In this thesis it is investigated whether the use of probabilistic information may lead to better decisions in water level management than the use of deterministic information. First the decision making process of Wetterskip Fryslân was investigated in how they make their decisions in water level management. If the expected water level in the Frisian belt canal system is too high, they may intervene by opening sluices or operating pumps. A simplified model of the water board is created in Chapter 5 which only has the Hoogland pumping station and sluices available, of which the sluices may be opened all at the same time. For this simplified model, the use of probabilistic information instead of deterministic information in the short term, i.e. forecasts up to 48 hours ahead, leads to better decisions for all investigated begin water levels. This also holds for the use of probabilistic information instead of deterministic information for the long term, i.e. forecasts up to 120 hours ahead. Whether the decisions made by the water board also leads to less expenses, depends on the cost-loss ratio of the water board.

Since the skill of the forecasts might differ and since the costs and losses of the water board may not be accurately determined, we also looked at a sensitivity experiment for different artificial skills. Here we have seen that the probability threshold is sensitive to the cost-loss ratio and to the skill of the forecast. Therefore it is more accurate for a company to present a range of cost-loss ratios, so that the economic value that the company can get from its cost-loss ratio and the chosen probability threshold is the highest.

We have seen that the use of cost-loss and the use of probabilistic information may lead to better decisions. The cost-loss model is used as a tool in choosing the most optimal probability threshold. Because for most begin water levels, more than one probability threshold can be suggested, the decision maker of the water board can choose the most optimal probability threshold. In this way, a behavioural feature can be included after the use of the cost-loss model.

7.1 Discussion

First we tried to create a model that was as close to the decision making process of Wetterskip Fryslân as possible. It was not intented to see whether the decisions of Wetterskip Fryslân were right, but to create a model that can be transferred as a tool supporting the decision making process by other water boards as well. This was, however, hard because there are many factors through which the water board can influence the water level in the Frisian belt canal system. We also wanted to compare the use of probabilistic and deterministic information in the most direct way, so that less external factors could influence the differences. This has resulted in a simplified model for the water board (Chapter 5). This model can be extended to be used by more water boards and the use of more sluices and pumps.

In Chapter 6 we have compared the use of probabilistic and deterministic information by the output given in contingency tables. Since the contingency table had to be reinterpreted, as has been discussed in Section 5.4, it was not possible to use the scores of the original contingency tables given in Section 3.2. We have not designed a score that would show the benefit of the use of probabilistic information compared to deterministic information. A possible way to do that, is to calculate the relative increase (or decrease) of the total expense.

In Chapter 4 we have looked at the value calculation from the contingency tables. Since the contingency tables that resulted from our experiment had to be reinterpreted, we were not able to calculate ratios as the hit rate and the false alarm rate. Therefore, we were not able to calculate the value of the forecast in the way it has been done in Chapter 4. It would be interesting to come up with an expression for the value of the forecast formulated in terms of quantities that are directly linked to the skill of the forecast that can be derived from our contingency tables.

Because the installation costs of Hoogland were equal to zero, the results of the experiment, given in Chapter 6, showed in many cases that selecting a low probability threshold often leads to less expenses. However, if we had defined the loss of exceeding the Hoogland threshold as the hours at which the Wouda pumping station had to be operated, then the losses would entail a variable part (the number of hours Wouda had to be installed) and a fixed part equal to the installation costs (of the Wouda pumping station). In this way, it is less attractive to operate pumps and therefore the lower probability thresholds would probably be less attractive than stated in Chapter 6.

7.2 Related work

Related work on cost-loss is of A.H. Murphy in a number of his articles ([3], [13], [9]). He looked at the value of cost-loss in decision making processes. Thornes [4] looked at the quality and value of weather forecasts, where the use of cost-loss is also discussed. This is also done in an article by Roulin [11], where he suggested the use of a more-staged cost-loss model. A. Millner [6] also discussed the inclusion of a behavioural feature to cost-loss analysis.

Related work on water management by water boards is from Schalk Jan van Andel and Roland K. Price [15]. Here the water board Rijnland in the Netherlands was discussed and different weather forecast products were examined. They look, however, at more extreme weather events and do not include cost-loss analysis.

A paper of Kees Kok [1] also focused on Wetterskip Fryslân and their decision making process. In this article a warning system for extreme precipitation is discussed. This warning system advises when the water board should take precautionary actions, but not explicitly on which hours pumps need to operate or sluices need to be opened.

7.3 Future work

As stated in the discussion before, the model defined in Chapter 5 can be extended to more water boards and/or to the use of more pumps and sluices. In this way, the use of probabilistic information and deterministic information can also be compared in a more realistic setting. Then the strategy that is constructed in Chapter 5 could be rewritten in a way that water boards can use it in general for their decision making process.

Also stated in the discussion before is that we were not able to define a meaningful score from the contingency tables, as found in Appendix A.1 and Appendix A.2. This could be done by taking a closer look at the reinterpreted contingency tables stated in Section 5.4 and distinguish the moments at which unavoidable losses and correctly chosen actions were taken place. In this case, the hours that Hoogland was operated for each day in the winters of 2012 to 2015 needs to be evaluated to decide whether Hoogland could operate more hours to prevent the loss that first did occur.

Finally, it is interesting not to take exactly the same begin water level each day in the winters of 2012 to 2015, but to take the observed water level of the next day as begin water level of the upcoming day. This resembles the real water levels more and in this way it can be seen on which days in the winters of 2012 to 2015 the water level was high. In that case, it can also be evaluated whether operating pumps and opening sluices created a buffer so that losses in the future were prevented.

Bibliography

[1] C. J. Kok *et al.*, "Meteorological support for anticipatory water management.", *Atmospheric Research*, August 2010. doi:10.1016/j.atmosres.2010.08.013.

[2] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. `http://www.R-project.org/`.

[3] A. H. Murphy, "The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation.", *Monthly Weather Review*, vol. 105, no. 7, July 1977.

[4] J. E. Thornes and D. B. Stephenson, "How to judge the quality and value of weather forecast products.", *Meteorological Applications*, vol. 8, pp. 307–314, 2001.

[5] J. K. Lazo, "The Costs and Losses of Integrating Social Sciences and Meteorology", *Weather, Climate and Society*, vol. 2, pp. 171–173, July 2010.

[6] A. Millner, "What Is the True Value of Forecasts?", *Weather, Climate and Society*, vol. 1, pp. 22–37, July 2009.

[7] Gordon Woo, "Probabilistic criteria for volcano evacuation decisions", *Natural hazards*, vol. 45, pp. 87–97, 2008.

[8] D. S. Richardson, "Skill and relative economic value of the ECMWF ensemble prediction system.", *Q.J.R. Meteorological Society*, vol. 126, pp. 649–667, 2000.

[9] A. H. Murphy, "The Finley Affair: A Signal Event in the History of Forecast Verification.", *Weather and Forecasting*, vol. 11, no. 1, March 1996.

[10] J.P. Finley, "Tornado Predictions", *American Meteorological Journal*, vol. 1, pp. 85–88, 1884.

[11] E. Roulin, "Skill and value of hydrological ensemble predictions", *Hydrology and Eath System Sciences Discussions, Copernicus Publications*, vol. 11(2), pp. 725–737, 2007. hal-00305048.

[12] M. Roulston and L. Smith, "The Boy Who Cried Wolf Revisited: The Impact of False Alarm Intolerance on Cost-Loss Scenarios", *Weather and Forecasting*, vol. 19, pp. 391–397, April 2004.

[13] Allan H. Murphy, Richard W. Katz, Robert L. Winkler, Wu-Ron Hsu, "Repetitive Decision Making and the Value of Forecasts in the Cost-Loss Ratio Situation: A Dynamic Model", *Monthly Weather Review*, vol. 113, pp. 801–813, May 1985.

[14] Iwan Holleman, "Bias adjustment and long-term verification of radar-based precipitation estimates", *Meteorological Applications*, vol. 14, pp. 195–203, 2007.

[15] Schalk Jan van Andel, Roland K. Price, "Ensemble Prediction and Water-Level Forecasts for Anticipatory Water-System Control", *Journal of Hydrometeorology*, vol. 9, pp. 776–788, 2007.

Appendix A

Output example of experiment

A.1 Contingency tables for data winters 2012 to 2015 with begin water level equal to -50 cm, short term

Deterministic run Forecast period = 48h.		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	98	110
	Hoogland turned off	6	261	267
Total		18	359	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1971

Probabilistic run Forecast period = 48h. Probability target = 0.0		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	103	115
	Hoogland turned off	1	261	262
Total		13	364	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1997

Probabilistic run Forecast period = 48h. Probability target = 0.2		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	101	113
	Hoogland turned off	3	261	264
Total		15	362	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1983

Probabilistic run Forecast period = 48h. Probability target = 0.4		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	98	110
	Hoogland turned off	6	261	267
	Total	18	359	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1968

Probabilistic run Forecast period = 48h. Probability target = 0.6		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	94	106
	Hoogland turned off	10	261	271
	Total	22	355	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1953

Probabilistic run Forecast period = 48h. Probability target = 0.8		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	12	91	103
	Hoogland turned off	13	261	274
	Total	25	352	377

Total hours sluices open: 3102
Total hours Hoogland operates: 1922

Probabilistic run Forecast period = 48h. Probability target = 1.0		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	0	0	0
	Hoogland turned off	116	261	377
	Total	116	261	377

Total hours sluices open: 3102
Total hours Hoogland operates: 0

Table A.1: *Contingency tables based on a simplified model for the water board and the constructed strategy discussed in Chapter 5 that show the differences in using deterministic and probabilistic information given 377 days in the winters of 2012 to 2015. The begin water level is equals -50 cm and the forecast period equals 48 hours.*

A.2 Contingency tables for data winters 2012 to 2015 with begin water level equal to -50 cm, long term

Deterministic run Forecast period = 120h.		Observed Exceed Hoogland threshold	Observed Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	143	162
	Hoogland turned off	8	207	215
Total		27	350	377

Total hours sluices open: 8062
Total hours Hoogland operates: 7075

Probabilistic run Forecast period = 120h. Probability target = 0.0		Observed Exceed Hoogland threshold	Observed Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	251	170
	Hoogland turned off	0	207	207
Total		19	258	377

Total hours sluices open: 8062
Total hours Hoogland operates: 7121

Probabilistic run Forecast period = 120h. Probability target = 0.2		Observed Exceed Hoogland threshold	Observed Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	145	164
	Hoogland turned off	6	207	213
Total		25	352	377

Total hours sluices open: 8062
Total hours Hoogland operates: 7088

Probabilistic run Forecast period = 120h. Probability target = 0.4		Observed Exceed Hoogland threshold	Observed Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	144	163
	Hoogland turned off	7	207	214
Total		26	351	377

Total hours sluices open: 8062
Total hours Hoogland operates: 7069

Probabilistic run Forecast period = 120h. Probability target = 0.6		Observed Exceed Hoogland threshold	Observed Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	140	159
	Hoogland turned off	11	207	218
Total		30	347	377

Total hours sluices open: 8062
Total hours Hoogland operates: 7026

Probabilistic run Forecast period = 120h. Probability target = 0.8		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	19	136	155
	Hoogland turned off	15	207	222
Total		34	343	377

Total hours sluices open: 8062
Total hours Hoogland operates: 6962

Probabilistic run Forecast period = 120h. Probability target = 1.0		Observed Exceed Hoogland threshold	 Not exceed Hoogland threshold	Total
Action	Hoogland turned on	0	0	0
	Hoogland turned off	170	207	377
Total		170	207	377

Total hours sluices open: 8062
Total hours Hoogland operates: 0

Table A.2: *Contingency tables based on a simplified model for the water board and the constructed strategy discussed in Chapter 5 that show the differences in using deterministic and probabilistic information given 377 days in the winters of 2012 to 2015. The begin water level equals* -50 *cm and the forecast period equals 120 hours.*