LUND UNIVERSITY
Faculty of Science

Ensemble Forecasting: A data analysis

Henrik Månsson
THESIS SUBMITTED FOR THE DEGREE OF BACHELOR OF SCIENCE

Department of Physics
HT2015

Supervised by
Elna Heimdal Nilsson and Torbjörn Simann

January 26, 2016

Abstract

This thesis covers the use of ensemble forecasts, how they work and their benefits. It will also give a brief history of weather forecasting followed by an evaluation, where ensemble mean precipitation data from the ECMWF is compared to observation data for 21 different observation stations spread across Sweden. From the evaluation, the mean has difficulty representing amounts of precipitation correctly for high amounts of rainfall but does a good job in describing when the precipitation will fall. The ensemble mean forecast has a generally lower RMSE for the northern observation stations, which would indicate that it does better at accurately forecasting the coming weather for these stations.

Contents

1 Introduction

Everyone has at some point in their life experienced how the weather can affect our society, be it disruption in public transportation, power outages and other similar problems. More devastating storms often cause large monetary damage or even worse, loss of human life. It has therefore become a vital part of today's society to be able to accurately predict how the weather will be within a few hours, to days ahead of time. There are many sectors, e.g. aviation, transportation, agriculture and energy, for which detailed forecasts of the weather development is crucial for their everyday operations.

To make predictions of the weather, advanced numerical models are run based on observed atmospheric conditions, which are measured all over the world. However, with the atmosphere being chaotic in its nature, small errors in measurements for the model input data and imperfections in the models used can cause the forecasts to deviate rapidly from reality and become erroneous. Accurate measurements for the initial values used in the models are therefore important for the accuracy of the forecasts.

This project will look at one of the more recent advances in weather forecasting, ensemble forecasts, and describe its general use with a focus on its use in precipitation and hydrological modelling. An evaluation of ensemble data of precipitation in Sweden provided by the Swedish Meteorological and Hydrological Institute (SMHI) will also be done where the mean of the ensembles for the period 2014-07-31 to 2014-11-30 will be compared to corresponding observations.

The evaluation will be done by comparing the ensemble means to observations for the same period and then calculating the error of the forecast.

1.1 History of weather forecasting

The 20th century brought a significant advances in a number of scientific fields covering computing, physics, mathematical theories , and many other, which would affect weather forecasting greatly [1].

In the beginning of the nineteen-hundreds, weather forecasting was a crude science with forecasts often based on empirical rules for localized areas and they were very dependent on the experience of the forecaster which combined with the very sparse atmospheric measurements available made for rather unreliable forecasts [1].

During the first decades of the 20th century a number of equations using seven basic variables, pressure, temperature, density, humidity and velocity in three coordinates, were developed in an effort to better describe the dynamics of the atmosphere. An early, mentioned in [1], attempt at a forecast using these equations and the most complete set of available data at the time was made by Richardson (a renowned meteorologist of the time) but, unfortunately, it resulted in an unrealistic outcome. The incorrect outcome was most likely caused by the input data Richardson used. However, the ideas of Richardson are the foundation on which today's forecasts are built. The numerical calculations needed for describing the atmosphere are tremendous and in the early 20th century they were too heavy to make actual usable forecasts on an everyday basis. It was not until the late 1940s, when an electronic computer designed for mathematical analysis was built at the Institute for Advanced Studies (IAS) in Princeton that it was possible to use numerical models for atmospheric forecasts. The improved mathematical theories and the use of early computers led the field of weather forecasting into a time of great advances. One of the major advances which came as a result of numerical models was the development of general circulation models (GCM), which improved the forecasts for a number of days ahead of time. The GCMs also led to an increased understanding of different factors that may affect the climate [1].

Nowadays, many of the national weather institutions around the world have their own supercomputers which can handle the enormous amounts of data that is constantly gathered from the global observing system (GOS), which contains around 11,000 weather observation stations spread around the world covering both land and sea [2].

Different institution use these computers and the collected data to run weather models which still are based on many of the same basic principles as the models that were developed during the early 20th century. The models do however still differ in many ways, for example in terms of grid sizes, boundary conditions and on how they handle parameterization. Despite the profound advances made during the 20th century there is a limit to how far into the future forecasts can accurately describe the atmosphere before the inherent unpredictability of the atmosphere cause the models to diverge from reality. This limit is often said to be 14 days but is generally much lower due to imperfections in the models [3].

1.2 Global climate change and energy use

Ever since the industrial revolution mankind has released an increasing amount of carbon dioxide and other greenhouse gasses. At first, the effects of the emissions were ignored and/or unknown, but during the 20th century more knowledge was obtained and an increased understanding of how these emissions actually affect the dynamics of our atmosphere. We are also now starting to see the consequential effects of the emissions around the world with a trend of increasing global surface temperatures and rising sea levels [4]. It is therefore important for the sustainability of our atmosphere that the amount of emissions around the world decrease. A large contributor of greenhouse gasses is the energy production sector, where a large amount of the produced energy still comes from fossil energy sources such as carbon power plants. Reducing the amount of energy generated from fossil sources would be a great step in reducing global emissions. However, for this to be possible, new renewable energy sources must be available to cover the reduction.

Renewable energy sources are in general more reliable on the weather than the fossil counterparts. This then creates a need for more accurate and more long-term weather forecasts that can help to calculate energy demand and how much energy could be generated over a period of time and thereby keeping the energy market more stable.

Wind and hydro-power play a significant role in the Swedish energy production and it is therefore necessary to have good forecasts, which would help with estimating both the energy demand and the energy that potentially could be generated domestically. If produced energy is not thought to be enough to cover the need, energy would have to be imported from foreign sources which might not be as environmentally friendly as the domestically produced energy and could come at a higher cost. Therefore, knowing beforehand approximately how much energy is potentially available could lead to great financial benefits as it would allow for the purchase of energy at different prices when it is needed. Better forecasts could also help with the decisions to further develop the domestic energy production which is generally an expensive affair. Forecasts describing the inflow to water reservoirs or rivers are not only important from an energy production perspective but also for public safety. A sudden influx of water could cause large floodings and in turn enormous damage to populated areas.

1.3 Ensemble Forecasts

As mentioned before, the atmosphere is inherently unpredictable because of its chaotic nature, and small errors in initial values could cause the model to quickly diverge from reality. Unfortunately, there is a limit to how accurately the measurements can be done and even if the error is small it is enough to cause the model to evolve in a different way. A way of compensating for this has therefore been developed: the so called ensemble forecasts. An ensemble forecast consists of a number of model runs, each with different initial values but all equally probable. This method gives a better understanding of how the atmosphere may evolve during the forecast period. Each model run gives a different outcome and by looking at them together one can see within which limits the atmosphere is most likely to evolve. It can also be used as a tool to determine the predictability of the atmosphere by showing how fast the different runs diverge from each other [3].

As mentioned before, this project will focus on ensemble forecasts and their general use together with a data analysis of ensemble forecast data and corresponding observations of precipitation in Sweden.

2 Ensemble Forecasts

2.1 General use

During most of the 20th century the atmosphere was thought of as essentially a deterministic system and the models produced a specific forecast from a specific set of initial values. Deterministic models are still an essential tool for today's weather forecasters and are used on a daily basis.

In 1992 the European Centre for Medium-range Weather Forecasts (ECMWF) started using a new method for describing the evolution of the atmosphere, ensemble forecasts. Ensemble forecasts are probabilistic forecasts, instead of deterministic, meaning that they instead of giving a specific answer of how the atmosphere will develop give a probability distribution of within which limits the atmosphere is likely to evolve. Probabilistic forecasts have shown to have a greater economic benefit for the consumers compared to deterministic forecasts by better representing weather related risks as they give the user multiple probability levels to decide their actions from. Even though a deterministic forecast can be used to create a multiple-value probability forecast using various methods, ensemble forecasts have been found to surpass those forecasts for lead-times greater than 3 days, at 500-hpa height, and give a better picture of the atmospheric development [5].

Ensemble forecasts were created for the use of medium-range forecasts and it is therefore not a surprise that the operational deterministic forecasts, with a higher resolution, outperform the ensemble forecasts for the first days. However, after a few days, when the accuracy of the deterministic forecast starts to drop, the ensemble forecasts generally generate a better picture of the atmosphere [7].

The spread of the ensemble forecasts can be interpreted in terms of predictability of the atmosphere for the forecast period. If the spread of the ensemble members is large, then the atmosphere is in a state where it is harder to forecast and where small changes in the initial conditions may have a large effect on the development of the atmosphere [7].

Ensemble forecasts consist of a number of members, the current medium-range forecast at ECMWF uses 51 individual members, where each member is a separate simulation of the atmosphere, all covering the same period of time but with slightly different initial values, all equally probable, and one control run that uses the same initial values as the operational deterministic models. The different initial values are based on the same atmospheric measurements but perturbed to cover errors that might be caused by the measurements. The resolution of the ensemble forecast members is lowers than the corresponding deterministic forecast due to limitations in computer capacity. There are currently three major methods of creating the initial for the different members of the ensemble, bred-vector perturbation method, singular-vector technique and a Monte-Carlo-like observation approach. Which one of the three methods is better for determining the initial values is uncertain and they are all currently used by different weather institutions around the world [8].

At the ECMWF, the method of singular vectors is currently applied to determine the initial values for their ensemble members. The singular vector model basically consists of identifying regions of the atmosphere where small errors would have the largest effect on the forecast and then creating 50 balanced deviations, one for each member around these unstable points [7].

Since an ensemble forecast consists of a number of model runs it is computationally heavier and each member of the ensemble is therefore run on a lower resolution than the corresponding deterministic model [8].

A common way of visualizing the results of the ensemble forecasts is by using a so called Ensemble Prediction System (EPS) Meteogram, which shows the mean value and spread for different variables during the forecast period, see Figure 1.

Another way of presenting the outcome of an ensemble forecast is by filtering the different members into clusters, where each cluster contains members with similar attributes [7].

Forecast errors can typically be divided into two groups. One contains errors of the initial conditions, which cover measurement errors, incomplete data coverage and data assimilation [9]. The other group consists of different kinds of model errors, for example errors caused by the way the model handles parametrization of physical processes and how it describes boundary conditions. The model errors are, however, also affected by the initial conditions and the two categories are, in reality, therefore heavily connected.

By using an array of perturbed starting values and thereby covering more of the background uncertainty, the use of ensemble forecasts has led to improvements in how data-assimilation is handled and it has reduced the impact of initialization errors. The second group of errors, the model dependent ones, are generally harder to improve on.

The accuracy of weather models has improved drastically over the last 50 years. This is partly due to, as mentioned before, a steady increase in computational power, which allows for models to be run at higher resolutions, currently in the scale of tens of kilometers, and thereby capturing more lower scale physical processes. Improvements on parametrization of the physical processes, which are still too small to be resolved properly by the model, has also contributed to the increase in forecast accuracy. The larger the scale being considered the lower the model errors usually are, as large scale processes are easier to capture in the model [9].

As all weather forecasts, ensemble forecasts are a tool to help the end users in their decision making process. It is therefore of interest to determine how the economic value of ensemble forecasts compare to that of a single high resolution

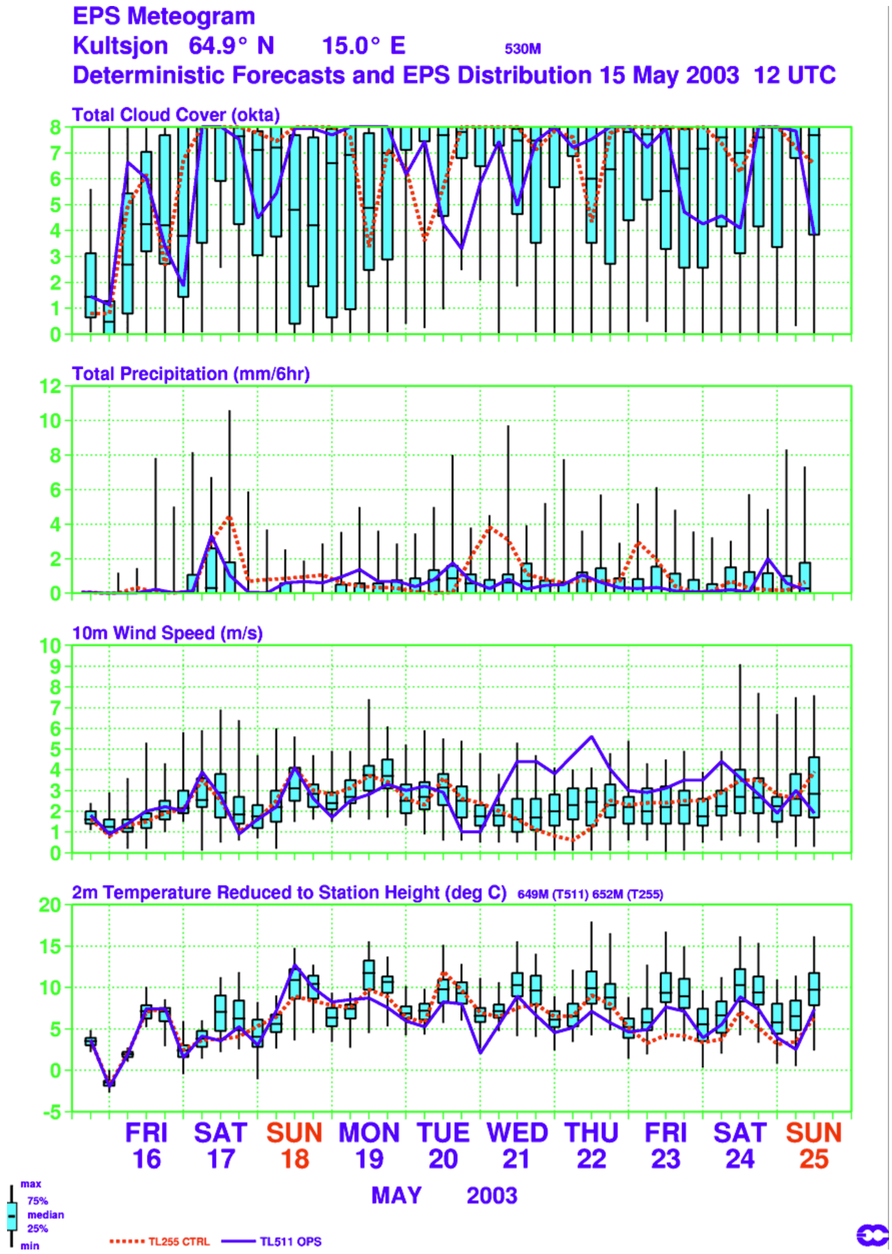


Figure 1: Shows an example meteogram of an ensemble forecast as used in [7], covering an area around Kultsjön in Sweden. The meteogram shows cloud cover, precipitation, wind speed and 2 meter tempereature. The blue line, TL511 OPS, is a high resolution forecast for comparison and the red line, TL255, is a control forecast which uses best, unperturbed, initial values.

deterministic one. A common way of calculating the economic value of forecasts is to use what is called a decision-analytical model where the cost benefit can be evaluated depending on which action is taken from the forecast [10]. From an analysis done by [10], ensemble forecasts provide the greatest value all through the medium-range, usually up to 10 days. During the first days of a forecast, a deterministic model with higher resolution performs better than the corresponding ensemble forecast. However, ensemble forecasts, which handle the impact of errors in initial values better, will be more accurate at longer lead-times than a deterministic forecast. Although, at which times the ensemble forecast starts to outperform the deterministic is dependent on the variable in question [9]. There may be some cases where the deterministic forecast has the upper hand over the ensemble and that is when an important event is too small to be captured by the lower resolution ensemble forecast [10].

The benefits of a probability forecast are that instead of a straight yes or no answer the user will get is a set of probabilities for different outcomes allowing a wider range of decisions to choose from. However, when dealing with probabilities it is important to present the data well otherwise it could cause confusion for those users who are inexperienced in terms of thinking in probabilities [5].

2.2 Ensembles in hydrology

For regulation of dams and power production planning the main forecasts are the long-term, seasonal, forecasts. Through the use of these forecasts it is possible to optimize the power production while at the same time keep enough capacity to handle sudden large inflows. The most important seasonal forecasts for Sweden is the so called spring flood, which culminate around May-June in the northern parts of the country. It is during this period when the greatest amount of snow melting occurs and most water is added to the system [11].

The method used for forecasting of the spring flood in Sweden is based on the Hydrologiska Byråns Vattenavdelnings model (HBV-model), which was first developed in the 1970s but has since then been continually improved [11].The HBV-model is now used by several countries around the world and its primary applications are: to produce short-term inflow predictions for hydro-power, to assist with dam safety and volume control, and to help with assessing the risk of flooding

For the seasonal forecast, the first step of the HBV-model is to do calibrated runs of the model using data from observations, covering the same catchment of interest. These runs are done for a certain period back in time up to the current date from which historical years are selected, which are thought to be a good representation of the coming period. The model runs for the historical years create an ensemble of the possible development of the atmosphere, from which possibilities for different scenarios can be extracted. A downside of this method is that it relies on the normal climate, which has the effect that if the spring flood deviates from the climatological mean, the errors will be significantly larger.

In a report from the SMHI, [11], three different methods of improving on the current spring flood forecasts are tested and evaluated. One of these three methods is the reduced historical ensemble, in which historical years are used to create an ensemble with the hope that several of them will be a good representation of the coming season.

The idea of using historical years to give a representation of what is to come is not new and in the early nineteen-hundreds there was an idea of using a catalog of historical years to describe the atmosphere [1]. This got heavily criticized at the time, because the method assumes that what the atmosphere has done before will repeat itself in a similar way. However, by using several historical years, which show similar trends as the current one, the outcome will hopefully be closer to actuality.

A study, [12], of two different catchment areas in Belgium found that the use of ensembles for flow prediction provided important data of extremes that were missed when using historical years.

At the SMHI, a version of the HBV-model is used for hydrological purposes. A calibrated model is run with initial values based on different ensemble members and thereby giving an estimate of how water flow may evolve and the resulting hydrological ensemble is then statistically post-processed to give the final results. The meteorological forecasts used for the initial values are based on forecasts for ten days ahead of time [7].

The main variables used from the meteorological forecasts for hydrological forecasts are precipitation and temperature. This is no surprise as precipitation describes how much extra water is added to the system and the temperature affects melting of snow caps and evaporation rates.

In [7] data from 50 different catchment areas was analyzed to see how well the model values compared to the observed values. Overall, the model functions well but in some of the studied areas the results from the model deviated heavily from reality. This was believed to be caused by the geographic location of said areas as the local hydrology may be more or less complicated depending on where you are. Unfortunately, the studied areas lacked data during some shorter periods of the evaluation period, approximately 9% of the time.

2.3 Forecast evaluation

Evaluating forecasts is important, to see whether the model manages to produce a good prediction and to help with the choice of models for different situations.

For ensemble forecasts there are two main ways of doing this; one, a deterministic evaluation where a single forecast, a mean forecast, is extracted from the ensemble which then gets evaluated and can be compared to both the operational deterministic forecast and observations. Two, a probability evaluation, where the spread of the ensemble forecast and the probability distribution are evaluated [7].

For the deterministic evaluation, the ensemble first has to be reduced to a single forecast. For instance, a common way of doing this is by taking the median of all the members of the ensemble forecast and thereby creating one single forecast.

Errors and quality of the forecast are typically described in statistical terms such as the mean error (ME), mean absolute error (MAE) and root mean square error (RMSE).

In the probability evaluation the total spread of the ensemble is examined and compared to reality. The ensemble forecast is often divided into quartiles to which then the observations are compared. As an example, if the maximum quartile is 75%, then 25% of the observations will be higher than the value predicted by the ensemble forecast.

3 Data analysis

Ensemble data from the ECMWF ensemble forecasts, provided by the SMHI, was used for an analysis where the forecast data for a number of stations in Sweden was compared to their corresponding observations. The forecast and observation data covers the time period of 2014-07-31 to 2014-11-30. During the autumn of 2014 Sweden received high amounts of rainfall and the water reservoirs around the country were filled to high capacity.

The data from the ECMWF that is used is from the forecast issued at 12pm each day and stretches 15 days forward in time. The forecasts contains both daily and six hourly forecast values of temperatures and precipitation. However, only the daily values of precipitation will be used for this analysis.

3.1 Method

The analysis will be done in the form of a deterministic evaluation of the ensemble forecasts, where all the ensemble members have been reduced to produce a mean forecast. This mean forecast will then be compared to observations for the same stations and 15 day periods. From these two data-sets the root-mean-squared-error (RMSE) for the two week forecasts was calculated.The error for forecast day 1-15 for each station was also calculated. The data used was provided by the SMHI and covers the time period of 2014-07-31 to 2015-09-29. The data contains daily observations for the whole time period and 15 day forecasts issued every day.

In the observation data stations, at occasion, lacked observation due to technical reasons. The analyzed periods were selected so that both the forecast and observation data had full coverage during the whole 15 days.

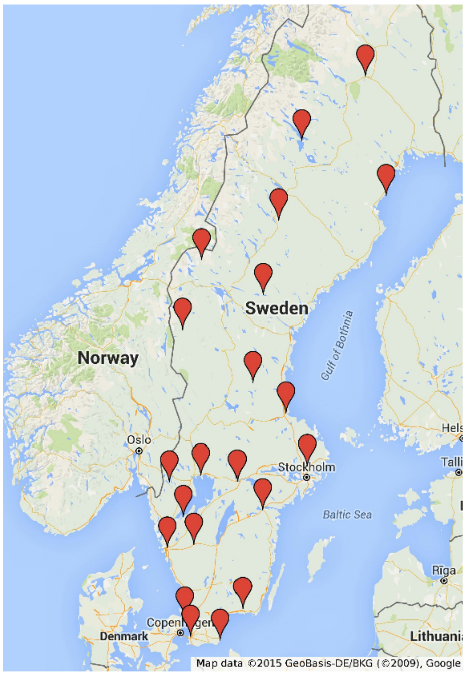


Figure 2: Shows the location of the examined stations in Sweden, generated with Google maps [13]

Due to time constraints, the data

for the whole year could not be treated and instead had to be limited to a shorter time span. The examined period was reduced to be from late summer, 31st of July, to the 30th of November. Reason for this was that the late autumn of 2014 received high amounts of rainfall and during the winter of 2014-2015 the water reservoirs around Sweden had reached high levels. Record numbers of rainfall for the month of October were reached in Heden, in Gothenburg, with 330 mm of rain.

When selecting which stations to examine, a list of Swedish WMO stations was compared to the two data-sets to see which stations in both sets. This resulted in 21 stations spread across the country with a higher concentration of stations in the central parts of Sweden.

The examined period was divided into sets of 15 days, i.e day 1 to 15, 16 to 30 and so on, and for each period the fifteen-day forecast that was made on the first day of the period for each station was used for the comparison.

At some occasions, a few of the stations were missing observation data and the error for these days was then set to -99 to denote a lack of data.

The examined stations with their respective coordinates can be found in Table 1 and their locations can be seen in Figure 2.

3.2 Root-Mean-Squared Error

Root-mean-squared error is a statistical measure of the difference between two data-sets and is calculated by:

$$RMSE = \sqrt{\frac{\sum(Y_{Obs} - Y_{Fcst})^2}{n}} \tag{1}$$

where n is the number of data-points in the series, Y_{Obs} is observed data and Y_{Fcst} is the forecasted data. With RMSE, high errors have a greater effect on the total error due to the fact that the errors are squared before they are averaged [3].

3.3 Program

To work with the data a self written program in Matlab was used. The first step of program was to read the data files from both the observations and forecasts for the studied period. Then the forecast data, which contained more than the needed data, had to be filtered to only provide the data of interest in organized matrices. After that the data sets were compared to find stations which were contained in both sets and this resulted in the 21 stations that were used.

The studied period was then divided into eight 15 day periods where the forecasts issued on the first day of each of the 15 day were used to calculate the RMSE for each 15 day period (see table 2).

Table 1: Shows the geographical information, latitude and longitude, the station numbers and the elevations of the examined stations.

WMO Nr	Station name	Latitude	Longitude	Meters above mean sea level
02049	Gallivare	67°09'N	020°39'E	359
02124	Arjeplog	66°02'N	07°52'E	430
02176	Ronnskar	65°02'N	021°34'E	3
02221	Korsvattnet	63°50'N	013°30'E	717
02245	Villhelmina	64°35'N	016°51'E	348
02247	Krangede	63°09'N	016°10'E	183
02308	Tannas	62°27'N	012°40'E	723
02338	Edsbyn	61°22'N	015°43'E	184
02408	Blomskog	59°13'N	012°05'E	170
02418	Karlstad Flygplats	59°22'N	013°28'E	46
02432	Orebro	59°14'N	015°03'E	53
02453	Gavle	60°43'N	017°10'E	16
02485	Stockholm	59°34'N	018°06'E	44
02513	Goteborg	57°42'N	012°00'E	5
02520	Satenas	58°26'N	012°42'E	54
02536	Rangedala	57°47'N	013°10'E	297
02574	Smhi	58°36'N	016°09'E	32
02611	Helsingborg	56°02'N	012°46'E	43
02625	Skillinge	55°26'N	014°19'E	4
02635	Malmo	55°35'N	013°01'E	13
02664	Ronneby	56°16'N	015°17'E	58

Thereafter, all of the forecasts were used to calculate the RMSE for each separate forecast day for each station. The results can be seen in table 3.

The forecasts for each station were also plotted together with the corresponding observations. Examples of this can be seen in figure 4 and 3, which shows the forecasts issued on day 15 and 76 for all of the stations.

3.4 Results

In this section, results of error calculations are presented in Table 2 and 3. Plots of two of the eight 15-day periods are shown in Figure 3 and 4. The two periods are the first 15 days of the data and day 76-90, where day one is the 31st of July 2014.

Table 2: Shows the root-mean-squared error (in mm) for each of the 15-day precipitation forecasts for all stations. -99 indicates that the station in question had one or more days where no observations were available.

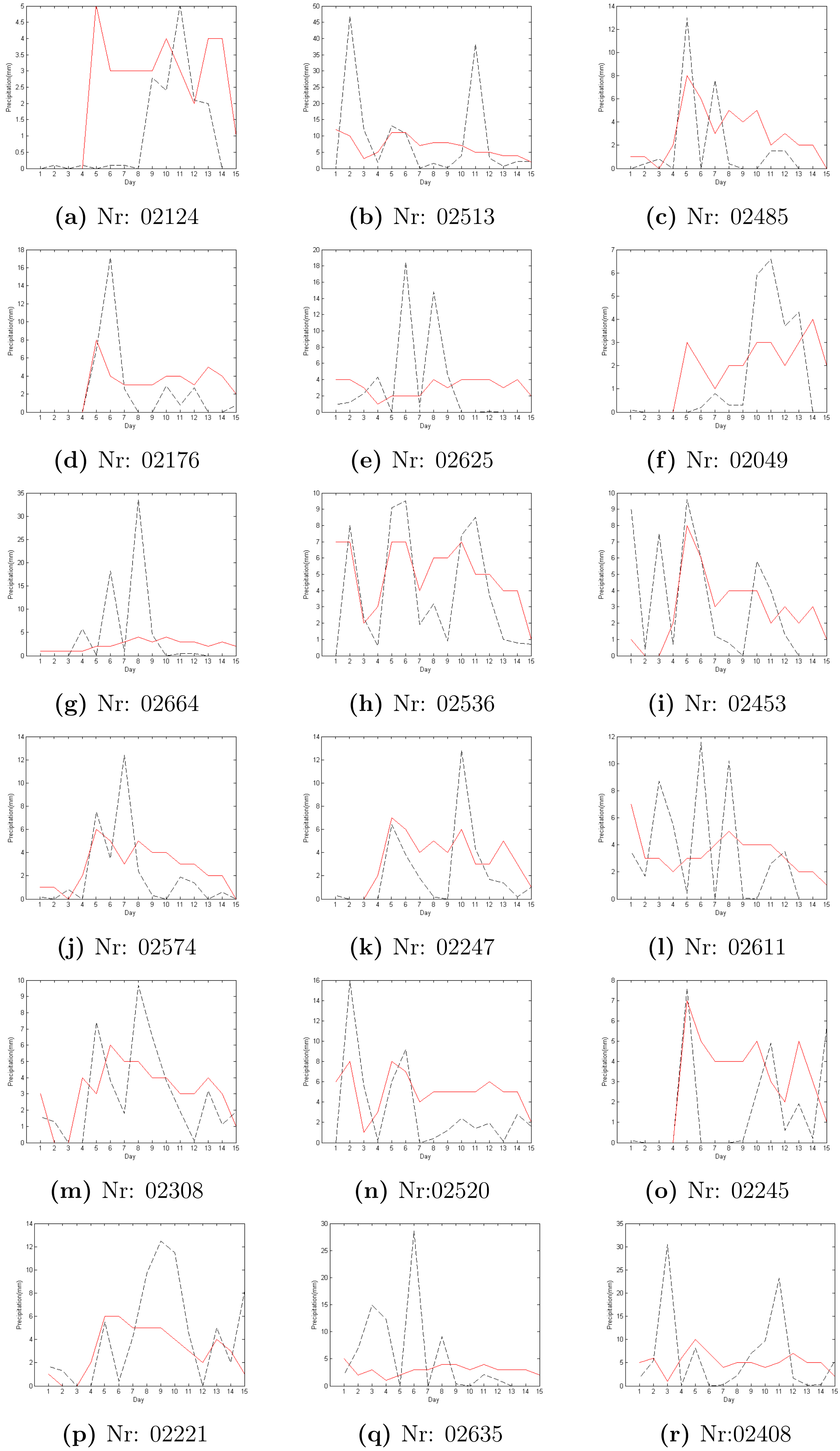
Station	1-15	16-30	31-45	46-60	61-75	76-90	91-105	105-120
02124	3.66	5.22	1.31	1.46	1.74	2.29	3.60	2.18
02513	7.16	6.07	4.20	1.84	3.83	13.80	4.69	3.22
02485	7.60	5.16	2.32	11.85	8.20	3.27	3.05	1.40
02176	1.25	5.05	1.14	1.30	4.07	4.04	2.61	3.78
02418	-99	-99	-99	-99	-99	-99	-99	-99
02625	5.66	11.98	7.15	1.70	3.15	5.79	1.65	3.12
02049	8.63	4.11	1.26	2.30	1.41	2.07	3.50	2.81
02432	-99	-99	-99	-99	-99	-99	-99	-99
02664	12.67	3.23	2.66	1.89	7.48	9.02	3.61	4.64
02536	4.00	6.97	1.88	5.24	3.01	3.04	3.61	-99
02453	2.31	2.59	1.27	12.99	4.42	3.46	2.88	1.17
02574	2.21	2.10	2.08	5.11	5.51	3.11	3.89	1.98
02247	3.88	4.46	1.00	1.54	5.43	2.87	3.97	2.90
02611	4.76	3.26	6.78	1.96	4.46	3.87	2.12	2.50
02308	5.21	3.66	1.40	1.80	6.27	2.54	1.86	2.24
02520	-99	5.68	1.88	4.05	2.22	4.11	4.19	2.07
02245	5.67	5.74	1.23	1.27	2.49	2.87	2.94	2.22
02338	2.77	3.77	1.13	1.36	6.11	5.05	4.38	4.31
02221	7.46	4.37	1.42	8.43	2.59	3.94	2.93	3.18
02635	2.73	5.10	15.39	2.30	6.06	8.36	1.87	2.63
02408	4.76	6.10	2.41	3.55	4.35	9.81	3.67	3.41

Instead of using the forecast produced each day, the whole period was divided into eight 15-day forecasts. The error for each of the eight periods can be seen for all of the stations in Table 2. Two of the stations continuously lacked measurements for one or more days in all of the eight periods and errors are therefore set to -99.

In Table 2, occurrences where a station has a very high error can be seen for a few of the stations during several periods. This is a result of using RMSE. As mentioned before, large errors will have a greater significance for the RMSE and it is therefore enough that one event of high rainfall is missed to raise the error. This can be seen when comparing the error of station 02513 (Table 2 in the period 76-90 to its plotted forecast, see sub-figure (b) in Figure 3. There the mean forecast greatly missed two events which in turn resulted in a high error for the station during this period.

Figure 3 and 4 are presented as examples of how well the mean predicts the precipitation for the different stations. The two periods were chosen at random but show similar features compared to the remaining six periods, with some differences.

Table 3 shows the calculated RMSE value for each forecast day for every station. For the forecast day error calculation, the data from each daily forecast day was used and not only the eight forecasts seen in Table 2.



(a) Nr: 02124 **(b)** Nr: 02513 **(c)** Nr: 02485

(d) Nr: 02176 **(e)** Nr: 02625 **(f)** Nr: 02049

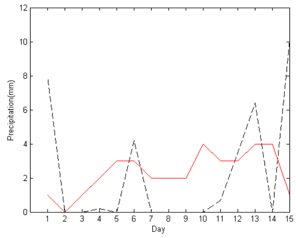
(g) Nr: 02664 **(h)** Nr: 02536 **(i)** Nr: 02453

(j) Nr: 02574 **(k)** Nr: 02247 **(l)** Nr: 02611

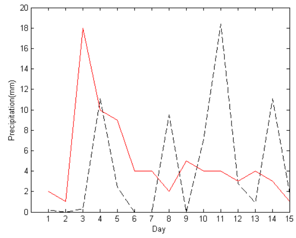
(m) Nr: 02308 **(n)** Nr:02520 **(o)** Nr: 02245

(p) Nr: 02221 **(q)** Nr: 02635 **(r)** Nr:02408

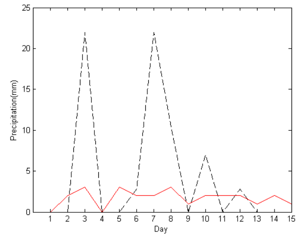
Figure 3: Plots of the forecast for day 76-90 for each station except Nr: 02338. The Y-axis shows the precipitation in millimeters and the X-axis is the day of the forecast. The red line shows the forecasted values and the dashed line is observations.



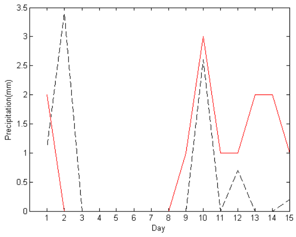
(a) Nr: 02124



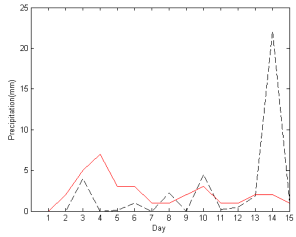
(b) Nr: 02513



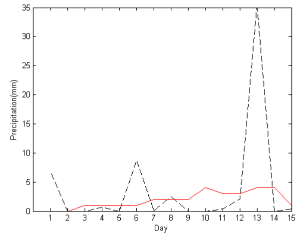
(c) Nr: 02485



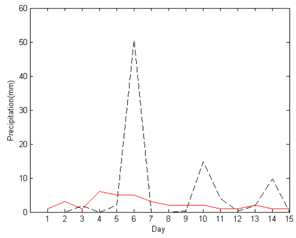
(d) Nr: 02176



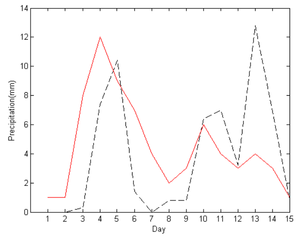
(e) Nr: 02625



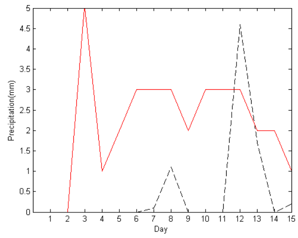
(f) Nr: 02049



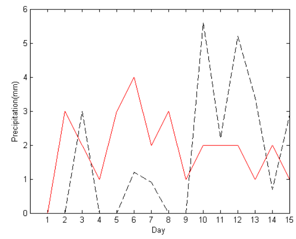
(g) Nr: 02664



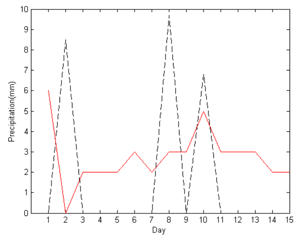
(h) Nr: 02536



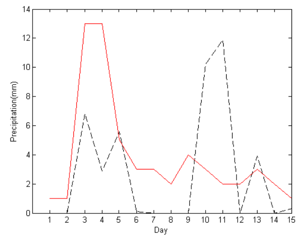
(i) Nr: 02453



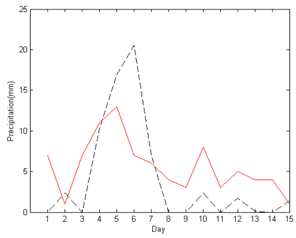
(j) Nr: 02574



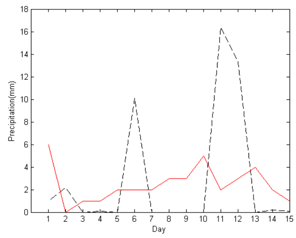
(k) Nr: 02247



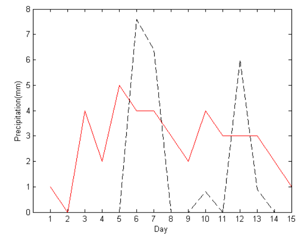
(l) Nr: 02611



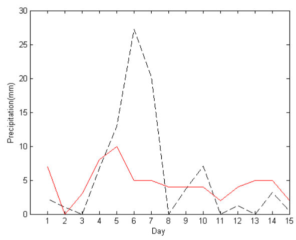
(m) Nr: 02308



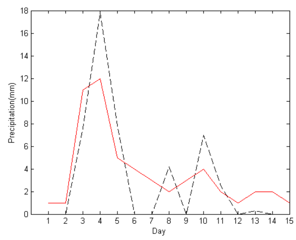
(n) Nr:02520



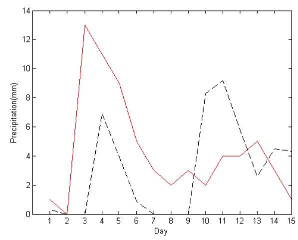
(o) Nr: 02245



(p) Nr: 02221



(q) Nr: 02635



(r) Nr:02408

Figure 4: Plots of the forecast for day 1-15 for each station except Nr: 02338.The Y-axis shows the precipitation in millimeters and the X-axis is the day of the forecast. The red line shows the forecasted values and the dashed line is observations.

Table 3: Shows the daily forecast error, in mm, for each station for all 15 day forecasts during the examined period. Zeros indicate a lack of observation data in the calculation.

Station Nr:	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14	Day 15
02124'	4,09	3,02	2,89	2,43	2,58	2,62	2,73	2,57	2,47	2,97	2,50	2,56	2,72	2,68	2,43
02513'	7,07	7,09	6,98	6,53	6,43	6,82	6,65	6,79	6,85	6,89	6,90	6,68	6,91	6,73	7,12
02485'	6,50	6,30	6,56	6,29	6,26	6,40	6,57	6,30	6,28	6,32	6,36	6,26	6,29	6,24	6,48
02176'	3,19	3,18	2,98	3,06	2,95	3,11	3,24	3,26	3,21	3,38	3,05	3,36	3,55	3,57	3,48
02418'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02625'	6,90	6,95	6,23	5,89	5,71	5,85	6,00	6,17	6,42	6,47	6,35	6,40	6,40	6,28	6,10
02049'	4,50	4,13	4,05	4,06	4,09	3,87	4,10	4,12	4,06	4,25	4,35	4,21	4,25	3,04	2,87
02432'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02664'	7,66	7,68	6,98	6,50	6,45	6,64	5,18	5,16	5,28	5,27	5,08	5,14	5,48	5,49	5,52
02536'	4,49	3,87	0	0	0	0	0	0	0	0	0	0	0	0	0
02453'	5,35	5,42	5,47	5,36	5,35	5,38	5,64	5,7	5,68	5,63	5,58	5,67	5,71	5,61	5,76
02574'	4,50	3,68	3,42	3,48	3,30	3,29	3,61	3,74	3,78	3,71	3,77	3,67	3,71	3,66	3,82
02247'	4,17	3,98	3,33	3,22	3,38	3,36	3,60	3,63	3,58	4,00	3,71	3,90	3,90	3,84	4,05
02611'	5,07	4,80	4,41	4,18	3,66	3,90	4,11	4,15	4,26	4,18	4,01	3,95	4,08	3,95	4,15
02308'	3,51	3,45	3,14	3,18	3,32	3,52	3,22	3,27	3,24	3,70	3,32	3,38	3,35	3,32	3,49
02520'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02245'	4,09	3,81	3,47	3,18	2,82	3,08	3,25	3,20	3,06	3,44	3,33	3,04	2,97	2,86	2,64
02338'	4,22	3,95	3,90	3,81	3,84	3,75	3,72	3,87	3,94	4,29	4,16	4,25	4,37	4,29	4,66
02221'	3,96	4,01	4,06	4,07	4,51	4,94	4,45	4,31	4,29	4,28	4,44	4,68	4,56	4,61	4,83
02635'	8,04	7,95	7,94	8,17	8,03	8,22	8,20	8,15	8,48	8,52	8,40	8,54	8,58	8,52	8,65
02408'	5,22	5,22	5,34	4,98	4,94	4,82	4,73	5,06	5,02	5,09	5,18	5,45	5,43	5,41	5,68

4 Discussion and conclusions

With the growing trends of renewable energy sources, improved weather forecasts are important to keep production and demand stable. As an example, if the prediction for the energy output, which is dependent on the weather forecasts, deviates by the magnitude of 1 GW over a number of hours, the following potential economic cost could be in the area of one million Euros [14]. Also, having good indications of water levels in reservoirs or how much water inflow is expected could have great effects on both public safety and potential production from water power plants. Ensemble forecasts are a good improvement for the forecasting industry as it gives a wider picture of how the atmosphere may develop, and as a result the forecasts can be utilized more efficiently.

As all forecasts, ensemble forecasts lose accuracy with time due to the inherent unpredictability of the atmosphere. However, this does not guarantee that the forecasts for the first days will be perfect, only that they are more likely to provide a better description of the atmosphere than the last days of the forecast.

Certain physical processes are harder to forecast well than others. These processes are usually on the small scales, which due to model resolution can not be captured properly and it will therefore be hard to tell exactly where the precipitation may fall. However, synoptic scale phenomena, such as frontal zones, are usually well depicted by the model. As ensemble forecasts, due to computational limitations, are run with a lower resolution than the operational deterministic models there is a greater chance that more of the small scale events are missed.

From the root-mean-squared errors presented in Table 2 the overall error for all of the stations are fairly high. This is partly due to, as mentioned above, that the further ahead in time you look the harder it is to make a good forecast but also the fact that the amount of precipitation, especially during the more intense rainfalls, seem to have been underestimated by the model mean at times. Looking at the forecasts plots in Figure 4 and 3 one can see that the mean of model data does a good job, with a few exceptions, of describing when the precipitation will fall, but fails to accurately describe the amount in many cases.

When comparing the errors in Table 2 to the location of the stations, see Table 1, the mean of the ensemble shows better performance for the northern stations during most of the periods, with few exceptions. The exceptions are likely caused by events which the model mean had a hard time capturing properly. This trend is also seen in Table 3 where the northern stations have a generally lower error for all of the days of the forecasts.

It would be expected that the error would increase with the later days in the forecast but from Table 3 no clear trend of increasing errors with time is noticeable and all stations seem to preform with the even root-mean-squared errors for all of the forecast days.

By looking at the daily errors for the stations, the ensemble mean gives a good inclination of the daily precipitation with relatively low errors. Although, due to a few days of high precipitation, either missed or highly underestimated by the mean of the models, the total error is brought up drastically.

However, when expecting an extreme, be it high precipitation or another variable, the mean value of a forecast is not the best value to go by. One of the ensemble members might have indicated the risk of this extreme but by using the mean value this data will not show. If one of the ensemble members had indicated an extreme event, such as high amount of rainfall, it would have been evaluated with the other possibilities to determine the risk. By looking at only the mean of the ensemble, a great deal of the benefits of this method, such as the multi-level decision making, are lost. The multi-level decision making has shown great economic value for ensemble forecasts as it allows for greater usage by the customer. With the probability-levels provided by the ensemble forecast, customers can make more informed decisions than if they only had a single yes/no answer for a specific situation. The mean of the ensemble is in principle a single-deterministic forecast generated with lower resolutions than the operational deterministic forecasts and will therefore likely be outperformed by those. With the mean of the ensemble the large-scale atmospheric features are still retained but smaller features might be lost.

The stations which had the highest errors for all the forecast days were all situated close to the sea. Weather at the sea-land transition is a bit more variable and harder with different interactions, such as sea-breeze and convective effects. This would explain why the ensemble mean performs worse for the coastal stations.

To improve on the evaluation it would have been good to also look at the data from the operational deterministic model alongside the ensemble forecasts to see how they would perform comparatively.

Overall, ensemble forecasts are a great tool for today's forecasters for all different variables. However, it is important to use alongside other methods to get the best outcome.

In the evaluation of the precipitation ensemble mean for the period of 2014-07-31 to 2014-11-30, when Sweden received high amounts of precipitation, it showed that the mean is not a reliable value to look at as it does not capture extreme values well, which could be crucial for the forecast.

For hydrological modeling, it has been shown, [12], that models using data from ensemble forecasts give better results when evaluating the risks of flooding.

A possible continuation of this project would be to include data from both observations, ensemble models and deterministic models to provide a better picture of their respective benefits.

References

[1] Lynch, Peter. The origins of Computer weather prediction and climate modeling, *Journal of Computational Physics* 227, 2008, 3431-3444

[2] World Meteorological Organization, 2015, Observation components of the Global Observing System; https://www.wmo.int/pages/prog/www/OSY/Gos-components.html, Accessed 19-11-2015

[3] Inness, P., Dorling, S.,*Operational Weather Forecasting*, JoHn Wiley & Sons Ltd, 2013

[4] IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

[5] Zhu, Y., Wobus, R., Richardson, D., Mylne, K., The Economic Value of Ensemble-Based Weather Forecasts, *Bulletin of American Meteorological Society* January 2002, 73-83

[6] Gneiting, T., Raftery, E., Weather Forecasting with Ensemble Methods,*Science*, 14 October 2005: 310, 248-249

[7] Olsson, J., Lindström, G., Johnell, A., Jacobsson, K., Hydrological ensemble forecasts, RH No. 21, SMHI, 2006

[8] Stensrud, D., Brooks, E., Du, J., Tracton, S., Rogers, E., Using Ensembles for Short-Range Forecasting, *Monthly Weather Review*, Vol 127, 1999, 433-446

[9] Leutbecher, M., Palmer, T.N., Ensemble Forecasting, *Journal of Computational Physics*, Vol 227, 2008, 3515-3539

[10] Richardson, D.S., Skill and relative economic value of the ECMWF ensemble prediction system, *Q.J.R. Meteorol. Soc*, vol 126, 200, 649-667.

[11] Olsson, J., Uvo, C., Foster, K., Yang, W., Södlking, J., German, J., Johansson, B., A multi-model system for spring flood forecasts, Elforsk, 11:72, 2011.

[12] Gouwleeuw, B.T., Thielen, J., Franchello, G., De Roo, A.P.J., Buizza, R., Flood forecasting using medium-range probabilistic weather prediction, *Hydrology and Earth System Sciences* 9, 2005, 365-380.

[13] Google, 2015,
https://www.google.com/maps/d/u/1/edit?mid=zzLO5f9mAZ5E.klqhjvFQCS5w,
Accessed 08-12-2015

[14] Lange, M.; Focken, U., "New developments in wind energy forecasting," in Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE , vol., no., pp.1-8, 20-24 July 2008 doi: 10.1109/PES.2008.4596135