NTNU
Norwegian University of
Science and Technology

Bayesian Model Averaging Using Varying Coefficient Regression and Climatology Cumulative Probability Regression

A Case Study of Postprocessing Hydrological Ensembles from Osali

Andreas Kleiven

Master of Science in Physics and Mathematics
Submission date: June 2017
Supervisor: Ingelin Steinsland, IMF

Norwegian University of Science and Technology
Department of Mathematical Sciences

Abstract

In this study, we introduce new Bayesian Model Averaging (BMA) approaches to construct probabilistic discharge forecasts. The approaches are tested for a case study from the Osali catchment in south-western Norway during the period from June, 24 2014 to June 22, 2015, with hydrological deterministic forecasts generated from a HBV-model using ensemble forecasts from European Centre for Medium-Range Weather Forecasts (ECMWF). In the classical BMA approach for postprocessing of ensemble forecasts, a probability density function is associated with each individual ensemble member forecast, and a sliding window training period is used to estimate model parameters, such as the mean and the variance of the individual ensemble member probability density functions. In hydrological forecasting, extreme events caused by snow melting or heavy rainfall, affect the BMA parameters in the following days, resulting in poor predictive performance. We suggest to model the mean in the BMA methodology with a varying coefficient regression (VCR) model to smooth parameter estimates and to better reflect changing weather patterns. Furthermore, we suggest to apply the Climatology Cumulative Probability Regression (CCPR) methodology to construct probabilistic discharge forecasts for each ensemble member, and then combine them with the BMA methodology. The calibration of the probabilistic forecast is assessed using the probability integral transform (PIT) and the predictive performance is assessed according to the continuous rank probability score (CRPS). The results from the case study showed that the predictive performance can be improved by including a varying coefficient regression model and/or the CCPR model in the BMA methodology for postprocessing of hydrological ensemble forecasts.

Sammendrag

I dette studiet introduserer vi nye Bayesian Model Averaging (BMA) metoder for å lage sannsynlighetsvarsel for tilsig. Metodene er testet i et casestudie av Osalifeltet sør-vest i Norge i perioden fra 24. juni 2014 til 22. juni 2015. Hydrologiske prognoser er generert av en HBV-model ved å bruke ensembleprognoser fra European Centre for Medium-Range Weather Forecasts (ECMWF). I den klassiske BMA-metoden til postprosessering av ensembleprognoser blir hvert ensemblemedlem tilegnet en sannsynlighetsfordeling der gjennomsnittet og variansen blir estimert basert på en treningsperiode i form av et glidende vindu. For hydrologiske sannsynlighetvarsel vil ekstreme hendelser, forårsaket av mye regn eller snøsmelting, påvirke BMA-parametrene i dagene som følger. Det vil føre til at sannsynlighetsvarslene vil stemme dårlig overens med observasjonene. Vi foreslår å modellere gjennomsnittet av sannsynlighetsfordelingen assosiert med ensembleprognosene med en Varying Coefficient Regression (VCR) modell, slik at parameterestimatene glattes ut og slik at vi lettere kan tilpasse modellen til varierende værmønster. Videre så foreslår vi å bruke Climatology Cumulative Probability Regression (CCPR) metoden til å lage sannsynlighetsfordelinger for hvert ensemblemedlem, og deretter kombinere dem med BMA-metoden. Kalibrering av det totale sannsynlighetsvarselet blir vurdert ved å bruke Probability Integral Transform (PIT) og kvaliteten på sannsynlighetsvarslene blir vurdert gjennom Continuous Rank Probability Score (CRPS). Resultatene fra casestudiet viste at sannsynlighetsvarslene stemmer bedre overens med observasjonene ved å inkludere en VCR modell og/eller CCPR modellen i BMA-metoden til bruk av postprossesering av hydrologiske ensembleprognoser.

Preface

This thesis is a part of the course TMA4905 Master's Thesis in Statistics at the Norwegian University of Science and Technology (NTNU), Department of Mathematical Sciences. It is the final part of the study program Industrial Mathematics. This work was carried out during the spring of 2017.

The main part of this work has been to develop postprocessing methods for ensembles of streamflow forecasts, and I would like to thank Stian Solvang Johansen from Statkraft for providing the data necessary for the methods developed in this thesis to be tested.

I would especially like to thank my supervisor, Professor Ingelin Steinsland, for our weekly meetings with discussions, and excellent guidance. Her ideas and support have been very valuable for me during the work with this thesis.

Trondheim, June 2017

Andreas Kleiven

Table of Contents

Chapter 1

Introduction

Hydrological forecasting plays an important role in a variety of applications, ranging from flood prevention to water resource management and hydropower production. Therefore, reliable hydrological forecasts are of great importance. Deterministic hydrological models are commonly used for discharge simulations to generate discharge forecasts. In this study, the Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1992) is used. Furthermore, the European Center for Medium-Range Weather Forecasts (ECMWF) provides temperature and precipitation forecasts, which is used as inputs in the HBV model. The forecasts from ECMWF are generated by running the ECMWF numerical weather prediction model which is based on the physical atmosphere. However, the resulting forecasts are not perfect. Forecasts errors originating from errors in the initial state from which the forecast is run, increases as the forecast horizon increases. In addition, the model formulation is only an approximation to the atmosphere, which also is a source of error. In order to determine the uncertainty of a forecast, ECMWF uses an ensemble of forecasts with perturbations made for both the initial state and the model formulation for each member of the ensemble (Woods, 2005). The ensembles are used as inputs in the HBV-model to produce an ensemble of discharge forecasts. From the ensemble forecasts, the uncertainty can be assessed, and the resulting ensemble can be interpreted as a probabilistic forecast.

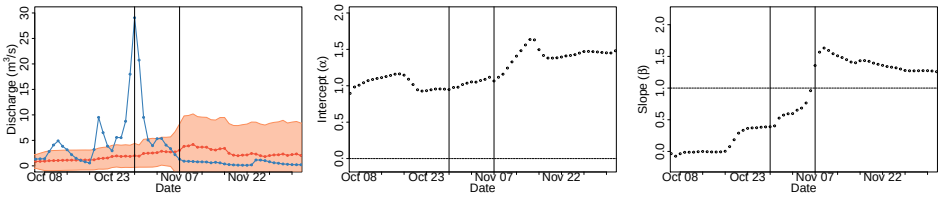
Studies have shown that ensemble forecasts tend to be underdispersive, meaning that the observed value too often lies outside the ensemble range (Raftery et al., 2005). Therefore, the need of statistical postprocessing methods is essential in order to obtain reliable probabilistic forecasts. Different methods have been proposed in scientific papers. Examples are the Hydrological Uncertainty Processor (Krzysztofowicz and Kelly, 2000), the Hydrological Model Output Statistics (Regonda et al., 2013), Quantile Regression (Weerts et al., 2011), Climatology Cumulative Probability Regression (Borhaug, 2014), and Bayesian Model Averaging (Raftery et al., 2005). This study focuses on the two latter methodologies.

The use of Bayesian Model Averaging (BMA) for statistical postprocessing of meteorological forecast ensembles was introduced by Raftery et al. (2005). They generated probabilistic forecasts for temperature by combining deterministic forecasts from different models. In the context of probabilistic weather forecasting, each ensemble member deterministic forecast is considered as a statistical model. Each deterministic forecast is associated with a probability density function (pdf) and the BMA probabilistic forecast, is given by a weighted average of the individual ensemble member pdfs. When modeling temperature, the normal distribution is a natural choice for the individual ensemble member pdfs, but in hydrological forecasting, discharge values are non negative. Hence a normal distribution is not likely to represent the data well.

Extensions of the BMA technique have been developed for cases when the quantity of interest deviates from the normal assumption. Sloughter et al. (2007) modified the method to apply for precipitation forecasts by introducing a discrete-continuous model which combines a logistic regression model and the gamma distribution as individual pdfs for the ensemble member forecasts. Furthermore, Rings et al. (2012) preprocessed the data using the Box-Cox transformation and suggested a more flexible representation of the conditional pdf in the BMA methodology using a joint particle filtering and mixture Gaussian modeling framework.

In the classical BMA methodology, the mean of the individual ensemble member pdfs are the bias-corrected ensemble forecasts, where the bias-correction parameters are estimated from a sliding window training period. In hydrological forecasting, extreme events caused by snow melting or heavy rainfall affects the BMA parameters in the following days, resulting in poor predictive performance. In Figure 1.1a, the BMA predicted mean and a 90% prediction interval for the Osali catchment during the period from October 8, 2014 to December 2, 2014, are plotted in red. The corresponding observations are given in blue. Lead time is the time between the day when the forecast was made, and the day when the forecast is valid, and the lead time of the probabilistic forecasts in Figure 1.1a are 9, but the forecasts are plotted against the day when they are valid. The day when the forecast is made is called issue time. The analysis was carried out in the student project during the autumn 2016. The bias-correction parameters, α and β, for each day are shown in Figure 1.1b and 1.1c, respectively. The horizontal lines indicate the bias-correction values if the ensemble forecasts are perfectly calibrated. The vertical lines show the time delay between the issue time and lead time. It is seen that 90% prediction interval of the BMA probabilistic forecast does not increase much until 10 days after the large forecast-observation error at October 28, since the days between the issue time and lead time is not included in the sliding window training period for the model parameters. Furthermore, the slope parameter β ranges from below 0 to above 1.5. This is not desirable since when β is zero, the mean of the individual ensemble members pdfs are based on α solely, which means that the mean of the probabilistic forecast is just the average observed discharge value in the training period. We further observe that α is large. The model parameters are further described in Chapter 3.2. We suggest to model the mean with a varying coefficient regression model to smooth parameter estimates and to better reflect changing weather patterns. Varying coefficient regression (VCR) is a class of generalized linear regression models where the coefficients are allowed to vary as functions of other

variables (Hastie and Tibshirani, 1993). Examples on applications include nonparametric regression (Cleveland, 1979), dynamic linear models (West, 1996) and spatial and spatio-temporal models (Blangiardo and Cameletti, 2015).



(a) Mean and 90% prediction interval in red, and corresponding observations in blue.

(b) Intercept parameter.

(c) Slope parameter.

Figure 1.1: BMA probabilistic forecast using the method developed by Raftery et al. (2005) with a 60 days sliding window training period, and model parameter estimates for lead time 9 at the Osali catchment in the period from October 8 to December 2, 2014.

We further suggest to use the Climatology Cumulative Probability Regression (CCPR) methodology by Borhaug (2014) to generate the individual ensemble member pdfs and combine them with the BMA methodology to one probabilistic forecast. The CCPR method uses the information available in the climatology, combined with different deterministic forecasts, to fit a beta distribution on the unit scale and then transform the pdf back to original scale.

Before using the probabilistic forecast, it is important to evaluate the quality. The quality can be measured by calibration and sharpness. Calibration is the consistency between the ensemble forecasts and the corresponding observed values. Sharpness is a measure of uncertainty of the probabilistic forecast. When assessing the quality of deterministic forecasts, sufficient measures are scoring rules such as the mean absolute error (MAE) and the root mean square error (RMSE). For probabilistic forecasts, both calibration and sharpness need to be assessed. The continuous rank probability score (CRPS) is a proper scoring rule that is widely used for evaluating probabilistic forecasts, and addresses both calibration and sharpness (Hersbach, 2000; Gneiting and Raftery, 2007). Furthermore the probability integral transform (PIT) (Rosenblatt, 1952), is often used to assess calibration of probabilistic forecasts (Raftery et al., 2005; Gneiting et al., 2005).

In this work three new postprocessing models are suggested: BMA with the beta pdf as the individual ensemble member pdfs (Beta-BMA), BMA with varying coefficient regression (BMA-VCR), and Bayesian model averaging with the beta pdf as the individual ensemble member pdf and varying coefficient regression (Beta-BMA-VCR). These models, together with the classical BMA methodology for forecast ensembles by Raftery et al. (2005), are tested and compared in a case study of the Osali catchment. Discharge observations are provided by Statkraft, which is the largest hydropower producer in Norway.

The thesis consists of seven chapters. In Chapter 2, the data available and a brief ex-

ploratory analysis of the forecast error are presented. In Chapter 3, the background material for the new models is presented, including evaluation methods for probabilistic forecasts, the BMA and CCPR postprocessing methodologies, regression models, and an inference section. The new postprocessing approaches are given in Chapter 4. Chapter 5 gives the models and methods used in the case study, and the results of the case study are presented in Chapter 6. In the last chapter, Chapter 7, the results are summarized and discussed.

Chapter 2

Data and Brief Exploratory Analysis

In this chapter, the data used in the case study is presented. We first give some information about the catchment, and then proceed with introducing the discharge observations and the runoff model used to obtain discharge forecasts. A brief exploratory analysis of the forecast error is presented in the end.

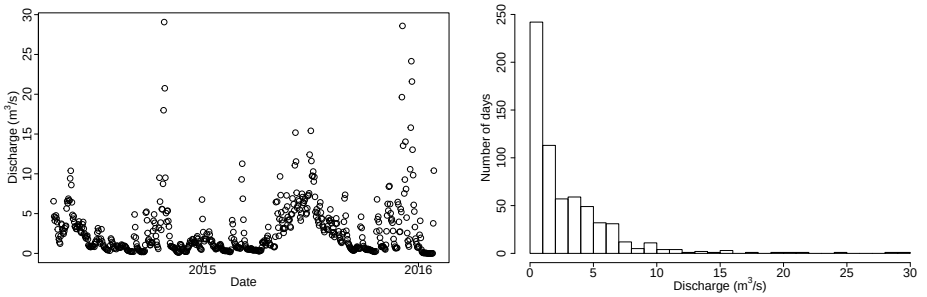
2.1 The Osali Catchment

In this study we consider the Osali catchment which is a part of the Ulla-Førre hydropower complex in south-western Norway. Figure 2.1 shows where Osali is located. The catchment is located at high altitude, and the mean temperature in the area ranges from $-2\ ^\circ\mathrm{C}$ on average in February to $13\ ^\circ\mathrm{C}$ on average in July. The amount of precipitation also varies seasonally, where June has the smallest amount and November the highest with 111 mm and 481 mm, respectively. Furthermore, annual runoff is measured to be 3200 mm on average where the highest average streamflow is in May and November (Engeland and Steinsland, 2014; Engeland et al., 2016).

Figure 2.1: Map over southern Norway. The red dot is the location of the Osali catchment.

2.2 Discharge Observations

Daily discharge, in unit m^3/s, during the period April 4, 2014 to January 27, 2016 is recorded and the data is provided by Statkraft. Figure 2.2a shows the daily observed discharge. We observe that there are two large spikes. One during the fall of 2014, and one late in 2015. This is typical for catchments in the south-western part of Norway. Large streamflows are caused by a major snow melt period during the spring and heavy rain fall during late autumn. We notice that there also are differences from one year to another, e.g. the observed discharge was a lot higher during the summer of 2015 compared to the summer of 2014. Figure 2.2b shows the histogram of discharge values. The majority of days has less than 10 m^3/s observed discharge.



(a) Observed daily discharge.

(b) Histogram of daily discharge.

Figure 2.2: Daily discharge in the period April 4, 2014 to January 27, 2016.

2.3 The HBV Model and Discharge Ensemble Forecasts

The ensemble forecasts used in this study are generated from the Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1992). The HBV model is a deterministic hydrological runoff model that is used for simulations. A simple illustration of the HBV-model is provided in Figure 2.3.

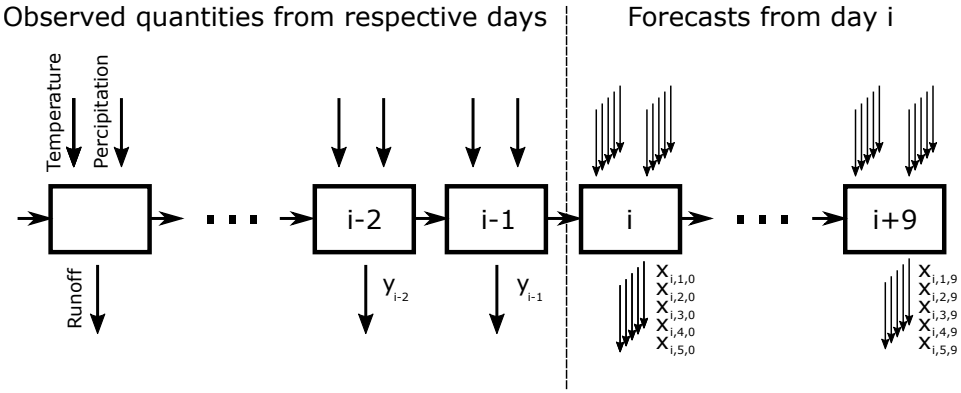


Figure 2.3: Simple illustration of the HBV-model. Inputs are temperature and precipitation, including rain and snow, and runoff is the output. The states before state i use observed quantities from respective days as inputs, and the states after state i use a set of deterministic forecasts from day i as inputs. The straight arrows pointing towards the box at each state represent the inputs and the arrows pointing away from the box represent the output. Temperature- and precipitation forecasts are obtained from ECMWF.

The model takes precipitation and temperature as inputs. Furthermore, the model has a number of free hydrological parameters that are estimated from training data (Engeland et al., 2016). The start state is estimated using observed precipitation and temperature up until yesterday (day $i-1$). Ensembles of temperature and precipitation forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) are used as input in the HBV-model. In Figure 2.3 there are five input arrows representing temperature forecasts and five arrows representing precipitation forecasts from day issue i. The resulting output is an ensemble of five discharge forecasts.

In this study, the ensemble size is $M = 51$, and the ensemble forecasts are treated equally, i.e. they are exchangeable. For each day in the period April 4, 2014 to January 27, 2016, there are initialized forecasts for $L = 10$ days, including the issue day. Throughout this chapter, we let i denote issue time, l lead time, m ensemble member number, and N the total number of days in the period. Lead time is the time between the the day when the forecast was made, and the day when the forecast is valid, i.e. the forecast horizon. The day when the forecast is made is called issue time. Thus, an ensemble forecast is denoted $x_{i,m,l}$ and the corresponding observation is denoted y_{i+l}. Raw ensemble forecasts for discharge are provided by Statkraft.

Figure 2.4 shows observed discharge before May 31, 2014, and the 51 ensemble forecasts for lead time 0 to 9, issued on May 31, together with the corresponding observations. We observe that the uncertainty in the raw ensemble forecasts increases with the lead time. The ensemble forecasts during May 2014 are plotted at lead time 0 and lead time 9 in Figure 2.5a and 2.5b, respectively. In both figures, there are 51 blue points plotted each day, representing the ensemble forecasts. Corresponding observed discharges are plotted as red dots. We observe that the forecasts seem to be biased. Furthermore, we observe that for lead time 0, the ensemble forecasts are concentrated around the same value and do not represent the uncertainty well. The ensemble spread is much larger for lead time 9.

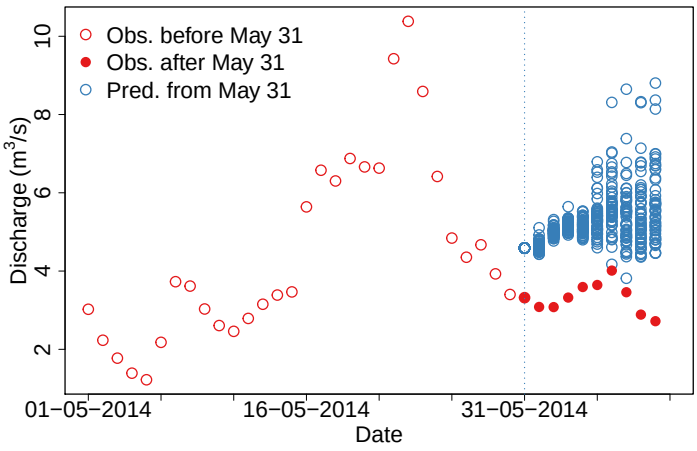
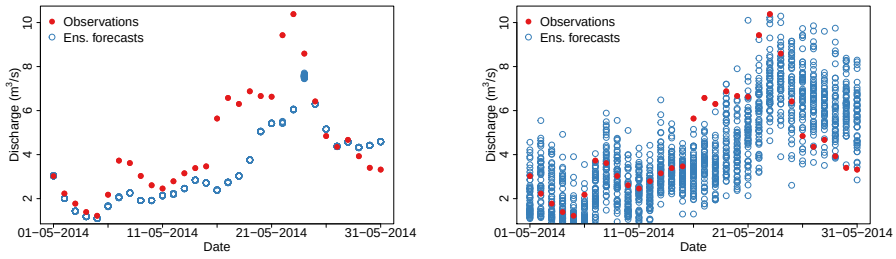


Figure 2.4: Predictions from issue time May 31, 2014 in blue. The solid red points are the corresponding, yet to be observed, observations. The open red points are historically observed discharge.



(a) Lead time 0.

(b) Lead time 9.

Figure 2.5: Ensemble forecasts in blue and verifying observation in red during May, 2014 for respective days.

2.4 Brief Exploratory Analysis of the Forecast Error

In this section, a brief exploratory analysis of the error between forecast and observed values is presented. Since the ensemble members are assumed to be exchangeable we expect the error to behave in a similar manner for all ensemble members. The error, $\epsilon_{i,m,l}$, between a forecast, $x_{i,m,l}$ (at issue time i, ensemble member m, and lead time l), and the verifying observation, y_{i+l} (at time $i + l$), is

$$\epsilon_{i,m,l} = x_{i,m,l} - y_{i+l}. \tag{2.1}$$

The mean error $\epsilon_{i,m,l}$ at each day in the period April 4, 2014 to January 27, 2016, is plotted in Figure 2.6. The line inside the box represents the median, and the boxes show the 25th and 75th percentile. The points are outliers. Negative error indicates that the forecasts more often are smaller than the corresponding observed values.

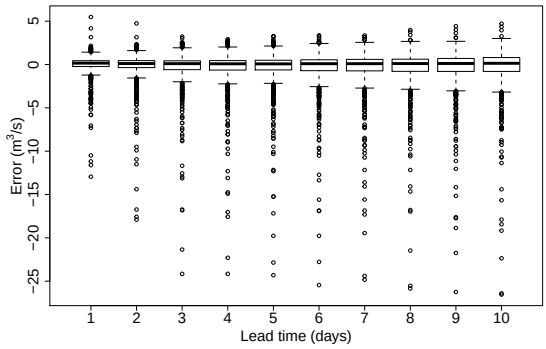


Figure 2.6: The error between forecast and corresponding observation for one ensemble member. The line represents the median, and the boxes show the 25th and 75th percentile. The points are outliers.

We further consider the mean of the ensemble forecasts at each day. The forecast error between the mean of the ensemble forecasts and the corresponding observation is

$$\epsilon_{i,l} = \frac{1}{M} \sum_{m=1}^{M} x_{i,m,l} - y_{i+l}. \tag{2.2}$$

The mean absolute value (MAE) and the root mean square error (RMSE) between the mean of the ensemble forecasts at each day and the corresponding observation y_{i+l} at lead time l,

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\epsilon_{i,l}| \qquad \text{and} \qquad \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \epsilon_{i,l}^2}$$

are plotted in Figure 2.7. We observe that the MAE and RMSE increase as the lead time increases.

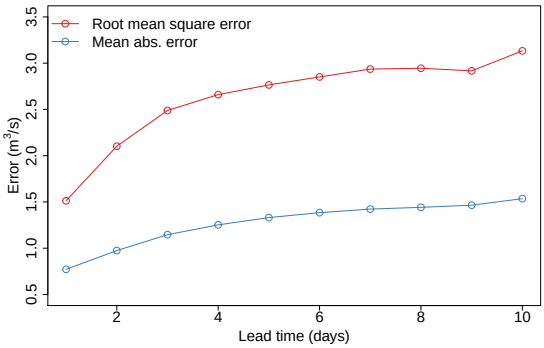


Figure 2.7: The mean absolute error and root mean square error as a function of lead time.

Chapter 3

Background

This chapter describes different approaches for evaluating probabilistic forecasts, and we further describe the original Bayesian Model Averaging (BMA) approach which is a widely used postprocessing technique for ensemble forecasts. Furthermore, some theory from regression models and the Climatology Cumulative Probability Regression (CCPR) approach are presented. In the end of the chapter, the two main inference paradigms, frequentist and Bayesian inference, are described. We use the same notation as introduced in the previous section. The predicted value for model m, issued at time i, and valid for lead time l, is $x_{i,m,l}$, and the corresponding observed value at time $i + l$ is y_{i+l}. However, in this chapter we only consider one issue time and one lead time, and we therefore simplify notation for the ensemble forecasts and corresponding observation to x_m and y, respectively.

3.1 Probabilistic Forecasting

Forecasting future quantities is of great importance in many applications, and forecasts reduce the uncertainty of a future event, but generally not eliminate it, and therefore, forecasts should be probabilistic (Gneiting et al., 2007). Probabilistic forecasts take the form of predictive probability density functions or predictive cumulative distribution functions, and in order for the forecast to be useful, it is important to assess the predictive performance. When evaluating probabilistic forecasts, Gneiting et al. (2005) suggest to maximize sharpness of the probabilistic forecast subject to calibration. Calibration is the statistical consistency between the predictive distributions and the corresponding observed values. Sharpness is a measure of uncertainty of the predictive distributions. Common tools for measuring the calibration of probabilistic forecasts are the verification rank histogram (VRH) (Anderson, 1996), and the probability integral transform (PIT) (Dawid, 1984). Furthermore, the continuous rank probability score (CRPS) is often used when

evaluating the predictive performance of a probabilistic forecasts (Gneiting and Raftery, 2007).

3.1.1 Calibration

Dawid (1982) define a forecast to be well calibrated if the observed empirical frequencies of an event coincides with the forecasted probabilities. The VRH is often used when assessing calibration of ensemble forecasts (Anderson, 1996; Hamill, 2001). The VRH is computed by arranging the ensemble forecasts and the corresponding observation in increasing order. If an ensemble is calibrated, then the observation is equally likely to take any place among the ensemble forecasts.

The probability integral transform (PIT) is common to use for assessing calibration of the probabilistic forecast. Rosenblatt (1952) introduces the probability integral transform, and Dawid (1984) suggested to apply it to assessment of probabilistic forecasts. Let F_i be the predictive cumulative distribution function (cdf) of the observation y_i. The probability integral transform is the value of the predictive cdf at the observation y_i, that is

$$p_i = F_i(y_i). \tag{3.1}$$

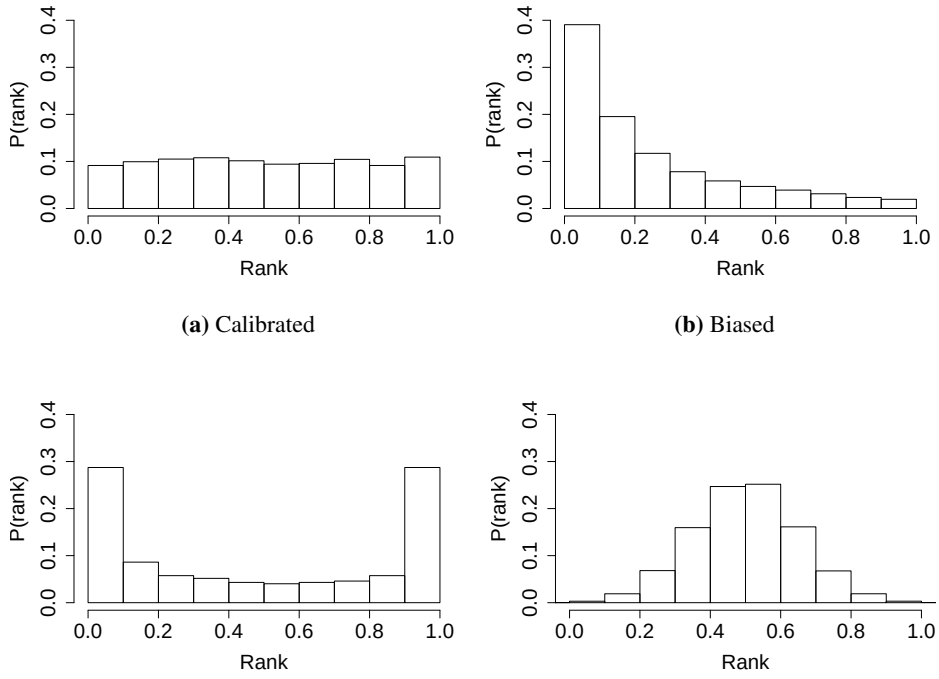
The probabilistic forecast is calibrated if the PIT values, p_i, is uniformly distributed. Uniformity can be assessed by plotting the histogram of PIT values, which we refer to as the PIT histogram.

Considering at the shape of the VRH ans the PIT histogram, gives us an indication whether the probabilistic forecast is calibrated (Gneiting et al., 2007). Hump-shaped histograms indicate that the probabilistic forecast is overdispersed, which means that the prediction intervals on average are too wide. U-shaped histograms indicates underdispersion, meaning that the prediction intervals on average are too narrow. Asymmetrical histograms occur when the probabilistic forecast is biased. Figure 3.1 shows examples of histograms when the probabilistic forecast is calibrated, biased, underdispersed and overdispersed.

3.1.2 The Continuous Rank Probability Score

Proper scoring rules are often used to assess the predictive performance of a probabilistic forecast. A scoring rule is proper if the expected score is minimized when the issued forecast is the true distribution of the quantity to be forecast (Gneiting and Raftery, 2007). The Continuous Rank Probability Score (CPRS) is a proper scoring rule that measures both calibration and the sharpness of a probabilistic forecast. If F is the predictive cdf and y is the corresponding observation, the CRPS is defined as

$$\text{crps}(F, y) = \int_{-\infty}^{\infty} (F(t) - \mathbb{1}_{\{t \geq y\}})^2 dt, \tag{3.2}$$



(a) Calibrated

(b) Biased

(c) Underdispersed

(d) Overdispersed

Figure 3.1: Illustrations of PIT histograms.

where $\mathbb{1}_{\{t \geq y\}}$ is the step function that attains value 1 if $t \geq y$ and 0 otherwise. The CRPS measures the difference between the predicted and occurred cumulative distributions. The value of the CRPS is non-negative and the smaller value the better quality of the probabilistic forecast.

Gneiting and Raftery (2007) show that the CRPS can be written as

$$\text{crps}(F, y) = \mathbb{E}|Y - y| - \frac{1}{2}\mathbb{E}|Y - Y^*|, \tag{3.3}$$

where Y and Y^* are independent random variables with cumulative distribution function F. The CRPS generalizes the absolute error and does therefore provide a way to compare deterministic and probabilistic forecasts. For exchangeable ensemble forecasts of size M with discrete cumulative density function F_{ens}, Equation (3.3) can be written as (Grimit et al., 2006)

$$\text{crps}(F_{ens}, y) = \frac{1}{M}\sum_{m=1}^{M}|x_m - y| - \frac{1}{2M^2}\sum_{m=1}^{M}\sum_{k=1}^{K}|x_m - x_k|, \tag{3.4}$$

where y is the observation and x_m, $m = 1, ..., M$, are the ensemble members. Furthermore, if F is the cdf of a normal distribution, Gneiting et al. (2005) show that there exists an analytic expression for the CRPS.

When evaluating probabilistic forecasts the CRPS is averaged over N days with F_i, $i = 1, ..., N$, being the predictive cdf of the quantity of interest, y_i. The mean CRPS is given as

$$\text{CPRS} = \frac{1}{N} \sum_{i=1}^{N} \text{crps}(F_i, y_i). \tag{3.5}$$

The CRPS reduces to the mean absolute error (MAE) for deterministic forecasts and is defined as

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - x_i|, \tag{3.6}$$

for observation-forecast pairs y_i and x_i.

3.2 Bayesian Model Averaging

The use of Bayesian Model Averaging (BMA) for statistical postprocessing of meteorological forecast ensembles was introduced by Raftery et al. (2005). The technique generates a probabilistic forecast for a quantity of interest by combining forecasts from different models. In the context of probabilistic weather forecasting, each ensemble member forecast, x_m, is considered as a statistical model. Each ensemble member forecast is associated with a probability density function (pdf) $g(y|x_m)$, and the BMA predictive pdf of any future weather quantity, y, is given by a weighted average of the individual ensemble member pdfs,

$$f(y|x_1, ..., x_M) = \sum_{m=1}^{M} w_m g(y|x_m), \tag{3.7}$$

where the weights, w_m, are non-negative and satisfy $\sum_{m=1}^{M} w_m = 1$. The weights are determined from a training period, and they reflect the performance of each ensemble member forecast in the training period relative to the other ensemble members.

The choice of the individual pdfs, $g(\cdot)$, is dependent on the quantity of interest. Raftery et al. (2005) consider temperature where the normal distribution seems to be a reasonable choice. In parts of this study we follow Raftery et al. (2005) and use the normal distribution with mean $\alpha_m + \beta_m x_m$ and standard deviation τ_m. We denote this as

$$Y|x_m \sim \mathcal{N}(\alpha_m + \beta_m x_m, \tau_m^2). \tag{3.8}$$

Dealing with exchangeable ensemble members simplifies the BMA procedure (Fraley et al., 2010). If all the ensemble members are exchangeable, Equation (3.8) can be written as

$$Y|x_m \sim \mathcal{N}(\alpha + \beta x_m, \tau^2), \tag{3.9}$$

and the ensemble BMA probabilistic forecast, Equation (3.7), takes the form

$$f(y|x_1, ..., x_M) = \sum_{m=1}^{M} w g(y|x_m), \tag{3.10}$$

where $w = 1/M$. The mean of the BMA probabilistic forecast is given by

$$\mathrm{E}[y|x_1, ..., x_M] = \sum_{m=1}^{M} w(\alpha + \beta x_m), \tag{3.11}$$

and the variance of the BMA probabilistic forecast can be written as (Raftery et al., 2005)

$$\mathrm{Var}[y|x_1, ..., x_M] = \sum_{m=1}^{M} w \left((\alpha + \beta x_m) - \sum_{i=1}^{M} w(\alpha + \beta x_i) \right)^2 + \tau^2. \tag{3.12}$$

The BMA predictive variance consists of two terms. The first term is a measure of how much spread there is between the ensemble forecasts, and the second term represents the uncertainty within each model.

To illustrate the BMA technique, an example of the BMA probabilistic forecast with five ensemble members is provided in Figure 3.2. Each of the five bias-corrected ensemble forecasts, has an associated normal pdf, and BMA probabilistic forecast is a weighted average over the five individual pdfs. In this example, the weights are $w = 1/5$, and the parameters are $\alpha = 0.50$, $\beta = 1.34$, and $\tau = 1.29$. We observe that the observation lies close to the mean of the BMA probabilistic forecast.

The model parameters in the BMA setting with $g(y|x_m)$ defined in Equation (3.9), is α, β, and τ. The parameters α and β are estimated by linear regression, and τ is estimated by maximum likelihood estimation (MLE) from training data.

Since the ensemble members are exchangeable, the parameters α and β are restricted to be the same for all ensemble members, and they are estimated by linear regression from a sliding window training period. The linear regression estimates for α, β correspond to the estimates obtained from maximizing log-likelihood function of $M \cdot (K-1)$ normally distributed random variables with mean $\mu_m = \alpha + \beta x_m$ and standard deviation τ'

$$(\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\tau}'_{ML}) = \underset{\alpha,\beta,\tau'}{\arg\max} \sum_{k=t-K}^{t-1} \left(\frac{1}{M} \sum_{m=1}^{M} \log(g(y_k|x_{k,m})) \right), \tag{3.13}$$

where M is the total number of ensemble members, K is the length of the training period, and $g(\cdot)$ in this case denotes the normal pdf with mean $\alpha + \beta x_{k,m}$ and variance τ'^2.

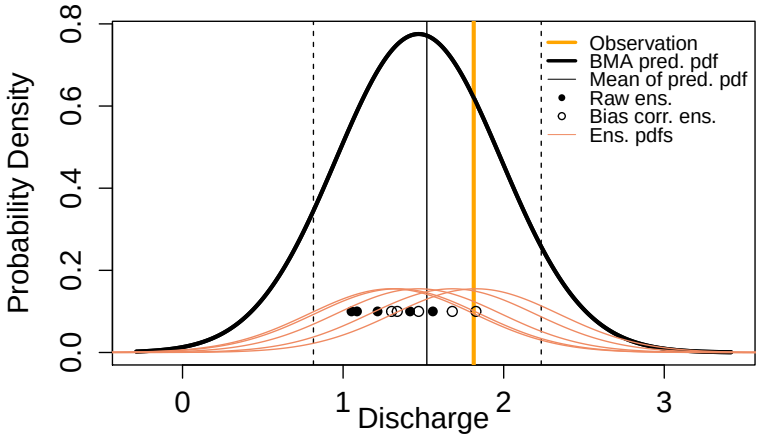


Figure 3.2: An illustration of the BMA probabilistic forecast for January 9, 2015. The filled circles are the raw ensembles, and the empty circles are bias corrected. The yellow vertical line is the verifying observation, and the black vertical line is the mean of the BMA probabilistic forecast. The orange lines are the weighted normal distributed ensemble pdfs, centered at the individual bias-corrected forecast. The vertical stapled lines show the standard deviation of the probabilistic forecast.

The estimate $\hat{\tau}'_{ML}$ is not useful in the BMA context. However, after obtained the bias-correction estimates, the variance parameter τ^2 can be estimated from the log-likelihood function of the BMA probabilistic forecast. The maximum likelihood estimate for τ is

$$\hat{\tau}_{ML} = \arg\max_{\tau} \sum_{k=t-K}^{t-1} \log\left(\frac{1}{M}\sum_{m=1}^{M} g(y_k|x_{k,m})\right), \tag{3.14}$$

where $g(\cdot)$ denotes the normal pdf with mean $\alpha+\beta x_{k,m}$ and variance τ^2, i.e. the individual ensemble member pdfs. This expression can be optimized numerically.

Extensions of the BMA technique have been developed for cases when the quantity of interest deviates from the normal assumption. Sloughter et al. (2007) modified the method to apply for precipitation forecasts. Because the amount of precipitation is highly skewed and has a non-zero probability of being equal to zero, a normal distribution is not appropriate for precipitation. Sloughter et al. (2007) therefore introduce a discrete-continuous model which combines a logistic regression model and the gamma distribution as individual pdfs for the ensemble member forecasts. Furthermore, Ajami et al. (2007) uses BMA to obtain reliable probabilistic hydrological forecasts using the box-cox transformation of streamflow values.

3.3 Generalized Linear Models

Generalized linear models (GLM) was introduced by Nelder and Baker (1972). For a thoroughly review of GLM, we refer to the book by Hardin et al. (2007). GLM is a class of models where the response variable y has a distribution that belongs to the exponential family, but not necessarily the normal distribution. The exponential family is a class of random variables with pdf on the form

$$\pi(y|\boldsymbol{\theta}) = h(y)c(\boldsymbol{\theta})\exp\left(\sum_{k=1}^{K} w_k(\boldsymbol{\theta})t_k(y)\right), \tag{3.15}$$

where $\boldsymbol{\theta}$ is the parameter vector, and y is a scalar or a vector. The functions $h(y)$ and $t_k(y)$ do not depend on $\boldsymbol{\theta}$, and $c(\boldsymbol{\theta}) \geq 0$. Most of the commonly used random variables belongs to the exponential family (Blangiardo and Cameletti, 2015).

In a generalized linear model there is a linear predictor

$$\eta = \alpha + \sum_{k=1}^{K} \beta_k x_k, \tag{3.16}$$

for a set of K explanatory variables x_k. The mean μ of of the random variable Y is connected to the linear predictor η through a link function $g(\cdot)$, i.e.

$$\mathrm{E}(y) = \mu = g^{-1}(\eta). \tag{3.17}$$

The link function $g(\cdot)$ describes the relationship of between the linear predictor η and the mean of the distribution function μ, and is chosen based on problem specific matters, e.g. the domain of the pdf of y should match the range of possible values that μ can take.

3.4 Beta Regression

Regression models, such as the linear regression model, is commonly used for application where data is related to other variables. A regression model describes the relationship between a dependent variable, or response variable, which in our case corresponds to the observation y, and one or more independent variables, or explanatory variables, which in our case correspond to the individual ensemble member forecasts x_m. However, when the response variable y is restricted to the standard unit interval, i.e. the interval from 0 to 1, the linear regression model is not appropriate since it may lead to response values outside the unit interval. Ferrari and Cribari-Neto (2004) proposed a regression model where the response variable is beta distributed, and it is useful for modeling continuous variables that take values in the interval between 0 and 1.

The pdf of a beta distributed random variable y is given by

$$\pi(y; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}, \tag{3.18}$$

where a and b are shape parameters. The expected value is

$$\mathrm{E}(Y) = \frac{a}{a+b}, \tag{3.19}$$

and the variance is

$$\mathrm{Var}(Y) = \frac{ab}{(a+b)^2(a+b+1)}, \tag{3.20}$$

The beta regression model is based on an alternative parametrization of the beta distribution, in terms of a mean parameter μ and a precision parameter ϕ. By setting $\mu = \frac{a}{a+b}$ and $\phi = a + b$ we get

$$\pi(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}. \tag{3.21}$$

Hence, the mean and the variance of y are

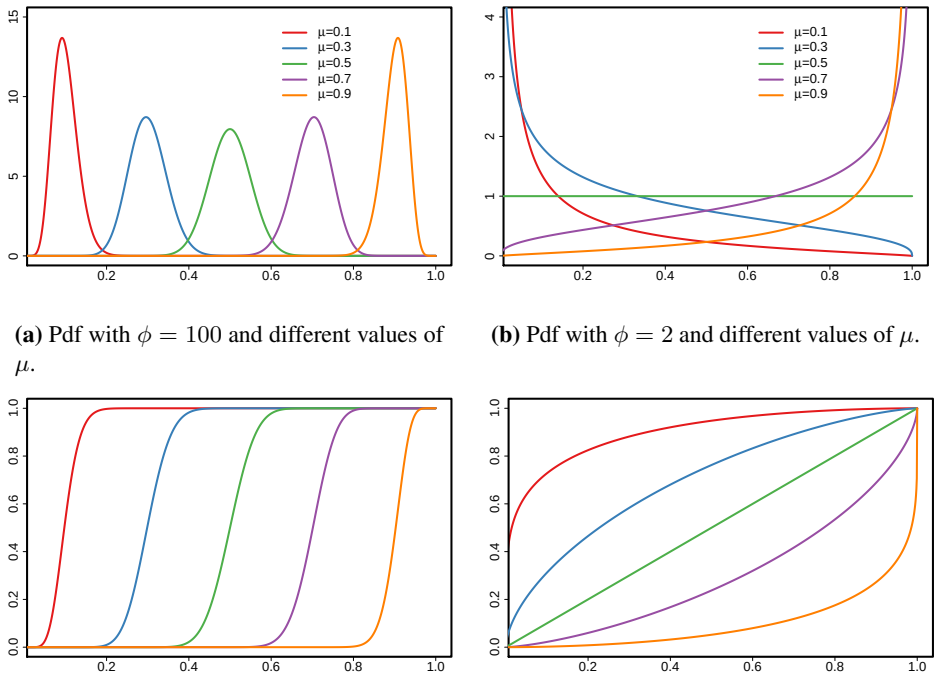
$$\mathrm{E}(Y) = \mu \tag{3.22}$$

$$\mathrm{Var}(Y) = \frac{\mu(1-\mu)}{1+\phi}, \tag{3.23}$$

respectively. The advantage of using μ and ϕ instead of a and b is that it has an easier interpretation which makes the modeling easier. The parameter ϕ is called a precision parameter since for fixed μ, the larger value of ϕ, the smaller variance of Y. The idea behind the beta regression model is to model the mean μ as a function of the linear predictor η defined in Equation (3.16), and the mean μ of the random variable Y is connected to the linear predictor η through a link function $g(\cdot)$. Since $\mu \in (0, 1)$ for the beta distribution, the logit link is a common choice for the beta regression model. Using the logit link function leads to

$$\mu = g^{-1}(\eta) = \frac{1}{1+e^{-\eta}}. \tag{3.24}$$

Figure 3.3 shows examples on the beta pdf and cdf for different values of the mean parameter μ and precision parameter ϕ. We observe that the beta distribution is flexible.



(a) Pdf with $\phi = 100$ and different values of μ.

(b) Pdf with $\phi = 2$ and different values of μ.

(c) Cdf with $\phi = 100$ and different values of μ.

(d) Cdf with $\phi = 2$ and different values of μ.

Figure 3.3: Examples of pdfs and cdfs of a beta distributed random variable with parameters μ and ϕ.

3.5 Climatology Cumulative Probability Regression

The Climatology Cumulative Probability Regression (CCPR) method was developed by Borhaug (2014). The methodology is inspired by Gneiting et al. (2013) who combined predictive pdfs using the beta-transformed linear pool. The CCPR method utilizes the information available in the climatology, combined with different deterministic forecasts. In this section, we consider one ensemble member forecast x and we do therefore suppress subscripts i, m, l.

Let Y denote the random variable representing the probabilistic discharge and let $G_Y(\cdot)$ denote the climatology cdf. The climatology cdf is based on historical observations and can be constructed in different ways, e.g. by fitting a distribution to the data (Borhaug, 2014). The idea behind the CCPR method is that given a deterministic forecast x and a climatology cdf, we fit a pdf on the unit scale according to a beta regression model

$$Z|x^* \sim \text{Beta}(\mu, \phi) \tag{3.25}$$
$$\mu = \text{logit}^{-1}(\alpha + \beta x^*), \tag{3.26}$$

where x^* is the value of the climatology cdf at the the deterministic forecast x, i.e. $x^* = G_Y(x)$, which means that the climatology is used to transform the original ensemble forecast x to the unit scale. We let $f_{beta}(\cdot)$ and $F_{beta}(\cdot)$ denote the pdf and cdf of $Z|x^*$, respectively. Note that the uniform distribution is a special case of the beta regression model, by setting $\mu = 0.5$ and $\phi = 2$. The interpretation of fitting a pdf on unit scale is that it gives the probability that the observed value y falls into different intervals of the climatology cdf, i.e. it is a pdf for $G_Y(y)$.

The interpretation of the bias-correction parameters α and β in Equation (3.26) can be difficult since the logit link is involved. Note that α determines the mean when $x^* = 0$, i.e. $\mu = \text{logit}^{-1}(\alpha)$. The slope parameter β has an important impact on the shape of μ. This can be seen in Figure 3.4. The mean μ is plotted against the deterministic forecast on unit scale x^* for different values of $\alpha = 0, -3$, and $\beta = 0, 1, 3, 5$. The dashed lines have $\alpha = -3$ and the solid lines have $\alpha = 0$. From the figure it is seen that α determines the intersection with the y-axis, and the lower values of α the closer to zero is the intersection point. The red solid line, which corresponds to $\alpha = 0$ and $\beta = 0$, is an example of poor predictive performance where μ is constant and equal to 0.5. An appealing property of the CCPR model is what we would get for a non-informative forecast: The postprocessed forecast is then the climatology, and the climatology is calibrated. On the other hand, if the predictive performance in the training period is good, β is estimated to be large and α small, which means that x^* has a large impact on μ. The dashed purple line, which corresponds to $\alpha = -3$ and $\beta = 5$ shows this. A drawback with the CCPR method is that it does not include the special case of perfect forecast (Borhaug, 2014).

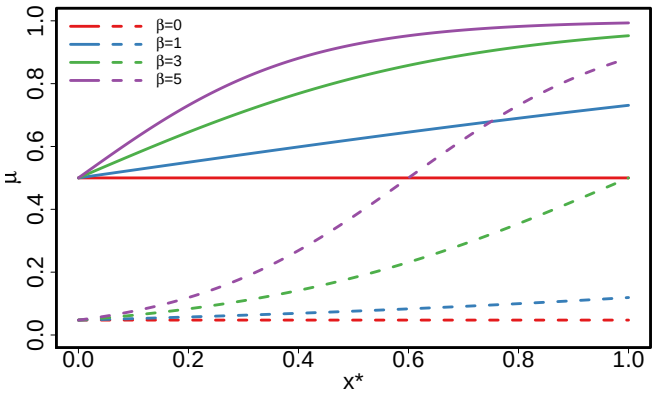


Figure 3.4: The mean μ as a function of the deterministic forecast on unit scale, x^*, for different values of α and β. The dashed lines have $\alpha = -3$ and the solid lines have $\alpha = 0$.

After fitting a pdf on the unit scale, it is transformed back to original scale and the post-processed cdf on original scale, denoted $G_{Y|x}(\cdot)$, is given by

$$G_{Y|x}(y) = G_Y^{-1}(F_{beta}(z|x^*)), \tag{3.27}$$

where $F_{beta}(\cdot)$ is the cdf of $Z|x^*$, and $Z|x^*$ is defined in Equation (3.25) and (3.26). From this expression we find that the special case when $F_{beta}(z|x^*)$ correspond to the uniform cdf, which happens when the prediction in the training period is without any information, then $G_{Y|x}(\cdot)$ and $G_Y(\cdot)$ are equal by the inverse integral transform. However, since the beta distribution is flexible, the postprocessed cdf $G_{Y|x}(\cdot)$ can take a wide range of different shapes.

We illustrate the method with an example where we have one deterministic forecast x and an empirical climatology cdf $G_Y(y)$ based on historical observations $(y_1, ..., y_n)$. This means that the empirical cdf jumps $1/n$ at observation values. The transformation is shown in Figure 3.5. The deterministic forecast x generate a beta cdf $F_{beta}(z)$, which is illustrated by the blue line to the left in the Figure 3.5. Since $F_{beta}(z)$ is the cdf of $G_Y(y)$, it is rotated counterclockwise by 90°. To generate the postprocessed forecast $G_{Y|x}(y)$, each value $G_Y(y_i)$ is assigned the value at the beta cdf, i.e. $F_{beta}(G_Y(y_i))$. The purple line shows the postprocessed cdf $G_{Y|x}(y)$. The transformation procedure is described in details for the point $(7, G_Y(7))$, and the procedure is similar for all points $(y_i, G_Y(y_i))$. We observe in the figure that after the transformation, the probability of observing a streamflow of at most 7 $\mathrm{m/s}^3$ is reduced from 0.9 to 0.5.

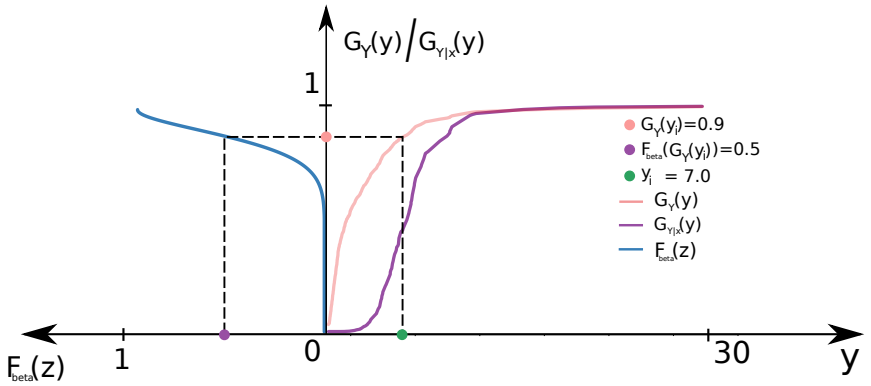


Figure 3.5: Illustration of the transformation procedure from the climatology cdf $G_Y(y)$ in red, via the unit scale cdf $F_{beta}(z)$ rotated counterclockwise 90° in blue, resulting in the postprocessed cdf $G_{Y|x}(y)$ in purple.

The estimation of the model parameters is in the original CCPR method by Borhaug (2014) performed by minimum CRPS estimation which is a procedure of fitting model parameters which minimizes the CRPS over a training period. An alternative procedure is to do maximum likelihood estimation from a training period, i.e

$$(\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\phi}_{ML}) = \underset{\alpha,\beta,\phi}{\arg\max} \left(\sum_{i=1}^{K} \ell_i(\mu_i, \phi) \right), \tag{3.28}$$

where

$$\ell_i(\mu_i, \phi) = \log\Gamma(\phi) + \log\Gamma(\mu_i\phi) - \log\Gamma((1-\mu_i)\phi) \tag{3.29}$$
$$+ (\mu_i\phi - 1)\log(y_i) + ((1-\mu_i)\phi - 1)\log(1-\mu_i), \tag{3.30}$$

and μ is defined in Equation (3.26). Bayesian inference is described later in this chapter, and is a third alternative.

3.6 Varying Coefficient Regression

Varying coefficient regression (VCR) models is a class of generalized linear regression models where the coefficients are allowed to vary as functions of other variables (Hastie and Tibshirani, 1993). Let Y be a random variable, and let the mean be defined as a function of a structured additive predictor η via a link function $g(\cdot)$, such that $g(\mu) = \eta$. A varying coefficient model with explanatory variables $X_0, ..., X_p, R_0, ..., R_p$ takes the form

$$\eta = \alpha(R_0) + \beta_1(R_1)x_1 + \cdots + \beta_p(R_p)x_p, \tag{3.31}$$

where functions $\alpha(\cdot)$ and $\beta_j(\cdot)$, depend on the explanatory variables R_j, which implies that there are some interaction between each R_j and x_j. The varying coefficient model in Equation (3.31) is general and includes a wide range of regression models. Examples include the generalized linear regression model which is obtained if $\beta_j(R_j) = \beta_j$, i.e. the constant function. If $\beta_j(R_j) = \beta_j R_j$ is a linear function, then the model has a product interaction on the form $\beta_j R_j x_j$. Furthermore, if $R_j = X_j$ we end up with a common model for smoothing or nonparametric regression (Cleveland, 1979).

The main focus in this thesis is the model where the regression coefficients are allowed to vary over time, denoted by t. The explanatory variables R_j are then the same variable, which is time t. The model takes the form

$$\eta = \alpha(t) + \beta_1(t)x_1 + \cdots + \beta_p(t)x_p, \tag{3.32}$$

for time steps $t = 1, ..., n$. As an example of a varying coefficient model with time varying parameters, we consider the varying coefficient model

$$Y_t = \boldsymbol{\beta}(t)X_t + \nu_t \qquad \nu_t \sim N(0, V_t), \tag{3.33}$$

where $\boldsymbol{\beta}(t)$ is the vector of regression coefficient functions, and X_t is the vector of explanatory variables at time t.

This model needs to be imposed some structure in order to be useful. An example is the dynamic linear model (DLM) (West, 1996). The DLM is usually formed by an equation describing the relationship between the response variable Y_t and the explanatory variables X_t, and a system equation describing the evolution of the vector of parameters $\boldsymbol{\beta_t}$ through time. One version of the DLM is

$$Y_t = \boldsymbol{\beta}_t X_t + \nu_t, \qquad \nu_t \sim N(0, V_t) \tag{3.34}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \omega_t, \qquad \omega_t \sim N(0, W_t), \tag{3.35}$$

where V_t and W_t are independent of each other.

3.7 Inference

There are two main paradigms for estimation of parameters. In the classical approach, the unknown vector of parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$ is fixed and is estimated by techniques such as maximum likelihood estimation. In the Bayesian approach, $\boldsymbol{\theta}$ is random and assigned a prior $\pi(\boldsymbol{\theta})$. It can then be estimated from the posterior distribution

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y})}. \tag{3.36}$$

Throughout this thesis we let $\pi(\cdot)$ denote a pdf.

3.7.1 ML, MAP and Bayes Estimators

A linear regression model assumes a linear relationship between the predictors and the outcome. As an example we consider the case where the outcome y_i is normally distributed with mean parameter μ_i and variance τ^2. The model setup is

$$y_i \sim \mathcal{N}(\mu_i, \tau^2) \tag{3.37}$$

$$\mu_i = \alpha + \beta x_i. \tag{3.38}$$

Let $\boldsymbol{\theta} = (\alpha, \beta, \tau^2)$ denote the vector of parameters. In the classical approach we are looking for optimal values of parameters by maximizing the likelihood function. The maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{ML}$ is given by

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} \prod_i \pi(y_i|\boldsymbol{\theta}). \tag{3.39}$$

In the Bayesian approach, priors need to be specified for the regression parameters α, β and the variance parameter τ. The maximum a posteriori estimator (MAP) allows us to include our prior beliefs on the parameter vector $\boldsymbol{\theta}$ by returning the maximum of the posteriori distribution, i.e. the mode

$$\hat{\boldsymbol{\theta}}_{MAP} = \arg\max_{\boldsymbol{\theta}} \frac{\pi(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y})} = \arg\max_{\boldsymbol{\theta}} \prod_i \pi(y_i|\boldsymbol{\theta})\pi(\boldsymbol{\theta}). \tag{3.40}$$

By assigning a non-informative prior to the parameter vector, i.e. $\pi(\boldsymbol{\theta}) \propto 1$, the ML estimate and MAP estimate coincide. However, both ML and MAP returns single and specific values for the parameter vector. In contrast, Bayesian estimation computes the full posterior distribution $\pi(\boldsymbol{\theta}|\mathbf{y})$. From the joint posterior distribution one can obtain marginal distribution for each element in the parameter vector, and point estimates can then be found by minimizing the posterior expected loss

$$\hat{\theta}_i = \arg\min_{\hat{\theta}_i} = \mathbb{E}(L(\hat{\theta}_i, \theta_i)), \tag{3.41}$$

where $L(\cdot)$ is the loss function, and θ_i denotes an element in $\boldsymbol{\theta}$. Such an estimator is called a Bayes estimator. A common choice for the loss function is the quadratic error loss function $L(\hat{\theta}_i, \theta_i) = (\hat{\theta}_i - \theta_i)^2$, or the absolute error loss function $L(\hat{\theta}_i, \theta_i) = |\hat{\theta}_i - \theta_i|$. By inserting into equation (3.41), the quadratic error loss function, yields the posterior mean (PM) as parameter estimate

$$\hat{\theta}^i_{PM} = \mathbb{E}(\theta_i|\mathbf{y}), \tag{3.42}$$

while the absolute error loss function yields the median of the posterior density (MPD),

$$\hat{\theta}^i_{MPD} = \text{median}(\pi(\theta_i|\mathbf{y})). \tag{3.43}$$

However, it is important to remember that in the Bayesian framework, the parameters $\boldsymbol{\theta}$ are characterized by a probability distribution, while in the frequentest framework, the parameters are fixed unknown quantities, and only the estimator $\hat{\boldsymbol{\theta}}_{ML}$ is a random variable. The main advantage of using Bayes estimators is that they allow us to express our confidence in any value we choose to use as parameter estimate through the prior.

3.7.2 Integrated Nested Laplace Approximation

The Integrated Nested Laplace Approximation methodology (INLA) was introduced by Rue et al. (2009). It is a method for performing approximate Bayesian inference. The method performs a direct numerical calculation of the posterior distributions for latent Gaussian models. As an alternative to simulation-based Monte Carlo integration, INLA uses the analytic approximation with the Laplace method. For further elaboration of the approximation procedure we refer to Blangiardo and Cameletti (2015), and our introduction is based on this book.

Given some observed data $\mathbf{y} = (y_1, ..., y_n)$, a model for y_i can be characterized by a mean parameter μ_i which is defined as a function of a structured additive predictor η_i via a link

function $g(\cdot)$, such that $g(\mu_i) = \eta_i$. The structured additive predictor η has the following form

$$\eta_i = \alpha + \sum_{k=1}^{K} \beta_k x_{ki} + \sum_{j=1}^{J} f_j(R_{ji}), \tag{3.44}$$

where β_k is the linear effect of the covariates x_k and $f_j(\cdot)$ are unknown functions of the covariates R_j. This structured additive predictor is comparable to the VCR predictor in Equation (3.31), and similar to the VCR-model, it has a wide range of applications due to the numerous forms of the unknown functions $f_j(R_j)$. If the $f_j(R_j)$ terms are omitted, the predictor coincides with the predictor in the generalized linear model in Equation (3.16).

The Latent Gaussian models can be described in a hierarchical structure. At first level, the model of the observations $\mathbf{y}$ is assumed to be conditional independent and identically distributed given the latent field $\boldsymbol{\xi} = (\alpha, \boldsymbol{\beta}, \boldsymbol{f})$ and some hyperparameters $\boldsymbol{\theta}_1$. The hyperparameters $\boldsymbol{\theta}_1$ typically are the measurement error precision. The model of the n observations are given by the likelihood

$$\pi(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\theta}_1) = \prod_{i=1}^{n} \pi(y_i|\xi_i, \boldsymbol{\theta}_1). \tag{3.45}$$

At the second level, the latent field $\boldsymbol{\xi}$ is characterized by a multivariate Normal distribution given the remaining hyperparameters $\boldsymbol{\theta}_2$

$$\boldsymbol{\xi}|\boldsymbol{\theta}_2 \sim \text{MVNormal}(\mathbf{0}, \mathbf{Q}^{-1}(\boldsymbol{\theta}_2)), \tag{3.46}$$

where $\mathbf{Q}^{-1}$ is the covariance matrix reflecting dependence structure of the model. We let $\pi(\boldsymbol{\xi}|\boldsymbol{\theta}_2)$ denote the pdf. Finally, at the third level of the hierarchical model, appropriate priors are assigned to the hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, denoted $\pi(\boldsymbol{\theta})$.

The three-level hierarchical structure is visualized in Figure 3.6. The unobservable quantities characterized by a probability distribution is represented as circles, and the observable quantities, i.e. the data, are represented as squares. The lines represent the conditional dependence structure. Furthermore, the latent Gaussian field in the second level can be assigned a dependence structure from the functions $f_j(\cdot)$.

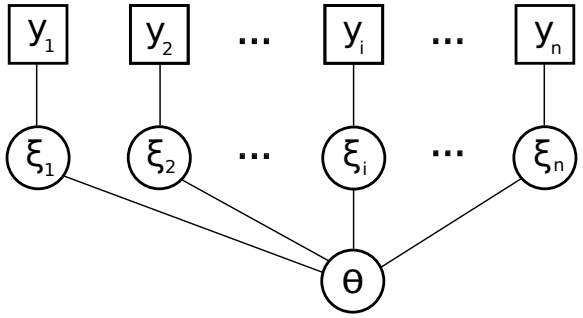


Figure 3.6: The three level hierarchical structure. The upper level is the observable variable y_i, represented by squares. The latent Gaussian field $\boldsymbol{\xi}$ is in the middle. The lower level is the hyper parameters θ. The two lower levels are unobservable quantities characterized by a pdf, and are represented by circles. The lines represent the dependence structure.

The joint posterior for the latent field $\boldsymbol{\xi}$ and the hyperparameters $\boldsymbol{\theta}$ takes the form

$$\pi(\boldsymbol{\xi}, \boldsymbol{\theta}|\mathbf{y}) \propto \prod_i \pi(y_i|\xi_i, \boldsymbol{\theta})\pi(\boldsymbol{\xi}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{3.47}$$

$$\propto \pi(\boldsymbol{\theta})|\mathbf{Q}(\boldsymbol{\theta})|^{n/2}\exp\left(-\frac{1}{2}\boldsymbol{\xi}^T\mathbf{Q}(\boldsymbol{\theta})\boldsymbol{\xi} + \sum_i \log\left(\pi(y_i|\xi_i, \boldsymbol{\theta})\right)\right). \tag{3.48}$$

We are interested in finding the marginal posterior pdfs for each element of the latent Gaussian field $\boldsymbol{\xi}$ and the hyperparameters $\boldsymbol{\theta}$. The marginal distributions can be defined as

$$\pi(\xi_i|\mathbf{y}) = \int \pi(x_i|\boldsymbol{\theta}, \mathbf{y})\pi(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta} \qquad i = 1, ..., n \tag{3.49}$$

$$\pi(\theta_j|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta}_{-j} \qquad j = 1, ..., n. \tag{3.50}$$

The INLA procedure to approximate these marginals is described in details by Blangiardo and Cameletti (2015).

3.7.3 Gaussian Markov Random Fields

Some latent Gaussian models satisfy the conditional independence property. A Gaussian Markov random field (GMRF) is a vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_n)$ with Markov properties, which means that γ_i, γ_j are conditional independent given $\boldsymbol{\gamma}_{-ij}$ where $\boldsymbol{\gamma}_{-ij}$ denotes all the elements in $\boldsymbol{\gamma}$ but γ_i and γ_j (Rue et al., 2009). The conditional independence assumption makes it computational easier. The joint density of the GMRF simplifies to

$$\pi(\boldsymbol{\gamma}) = \pi(\gamma_n|\gamma_1, ..., \gamma_{n-1})\pi(\gamma_{n-1}|\gamma_1, ..., \gamma_{n-2}) \cdots \pi(\gamma_2|\gamma_1)\pi(\gamma_1) \tag{3.51}$$

$$= \pi(\gamma_n|\gamma_{n-1})\pi(\gamma_{n-1}|\gamma_{n-2}) \cdots \pi(\gamma_2|\gamma_1)\pi(\gamma_1). \tag{3.52}$$

As an example of a GMRF we consider the random walk model of order 1. We assume constant time steps $t = 1, 2, ..., T$. The first order random walk model is constructed assuming independent increments

$$\Delta\gamma_t \sim \mathcal{N}(0, \kappa^{-1}), \tag{3.53}$$

where κ^{-1} is a precision parameter. This implies that

$$\gamma_s - \gamma_t \sim \mathcal{N}(0, (s-t)\kappa^{-1}) \qquad s > t. \tag{3.54}$$

The density of $\boldsymbol{\gamma}$ is derived from the $T-1$ increments in Equation (3.53)

$$\pi(\boldsymbol{\gamma}|\kappa) \propto \kappa^{(T-1)/2} \exp\left(-\frac{\kappa}{2}\sum_{t=1}^{T-1}(\Delta\gamma_t)^2\right) \tag{3.55}$$

$$\propto \kappa^{(T-1)/2} \exp\left(-\frac{\kappa}{2}\sum_{t=1}^{T-1}(\gamma_{t+1} - \gamma_t)^2\right), \tag{3.56}$$

which leads to

$$\gamma_t|\boldsymbol{\gamma}_{-t}, \kappa \sim \mathcal{N}\left(\frac{1}{2}(\gamma_{t-1} + \gamma_{t+1}), \frac{1}{2\kappa}\right) \tag{3.57}$$

3.8 Software

For computations, we use existing software in R. Ensemble forecasts and observations are loaded from R-files, and then converted to an *ensembleData* object. The R package *ensembleBMA* (Fraley et al., 2007) is used for simulations and estimation of parameters for the BMA approach for Gaussian ensemble member pdfs with a K days sliding training period. Inference and simulations for the VCR-models and the CCPR-models, is done using the R-package *R-INLA* (Rue et al., 2009) which can be downloaded from `www.r-inla.org`. For optimization of the variance/precision parameter, we use the function *optimize* in *stats* (R Development Core Team, 2008). The function uses Brent's method (Brent, 2013) for numerical optimization.

Chapter 4

New BMA Postprocessing Approaches

In this chapter we suggest new postprocessing models based on the theory described in the previous chapter. Subscript for lead time l and issue time i are suppressed from the notation since the postprocessing models are presented for one specific day. Some of the parameters in the new postprocessing techniques are evolving through time and a subscript t is therefore included on these parameters. Subscript m denotes ensemble member number.

4.1 BMA with Beta Probability Density Functions

In the BMA methodology, each ensemble member is associated with a pdf $g(y|x_m)$. The form of $g(\cdot)$ is chosen to fit the response variable y. Discharge values are non negative and therefore, the normal distribution is not appropriate. In academic papers several proposals for $g(y|x_m)$ for discharge forecasting have been made. Examples include the normal distribution combined with a box-cox-transformation of the streamflow values (Ajami et al., 2007), the gamma distribution (Vrugt and Robinson, 2007), and particle filtering with Gaussian mixture modeling (Rings et al., 2012). We suggest to apply the CCPR-methodology developed by Borhaug (2014) as the individual ensemble member pdfs.

We transform the ensemble forecasts x_m to the unit scale through the climatology cdf $G_Y(\cdot)$, and we denote the transformed forecast x_m^*. The beta distribution as the individual ensemble member pdfs on unit scale. The same notation as in Section 3.5 is being used, and the model setup is

$$Z_m|x_m^* \sim \text{Beta}(\mu_m, \phi) \tag{4.1}$$

$$\mu_m = \text{logit}^{-1}(\alpha + \beta x_m^*) \qquad m = 1, ..., M. \tag{4.2}$$

The interpretation of the intercept parameter α and slope parameter β discussed Section 3.5. The forecasts are then combined using the BMA methodology with exchangeable ensemble members. On unit scale, the Beta-BMA probabilistic forecast takes the form

$$f(z|x_1^*, ..., x_M^*) = \frac{1}{M} \sum_{m=1}^{M} f_{beta}(z|x_m^*), \tag{4.3}$$

where $f_{beta}(\cdot)$ is the beta distribution with mean μ_m defined in (4.2) and precision ϕ.

Figure 4.1 shows five weighted ensemble pdfs being combined according to Equation (4.3).

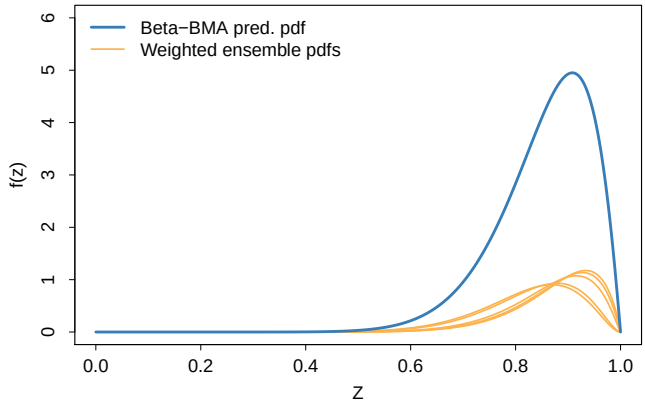


Figure 4.1: Five ensemble member pdfs, the orange lines, being combine to one pdf, the blue line, on unit scale.

The Beta-BMA probabilistic forecast is then transformed back to original scale according to the transformation procedure described in Section 3.5. To the left in Figure 4.2 the ensemble member cdfs, corresponding to the pdfs in Figure 4.1, are combined, and then the combined cdf on unit scale is transformed to original scale.

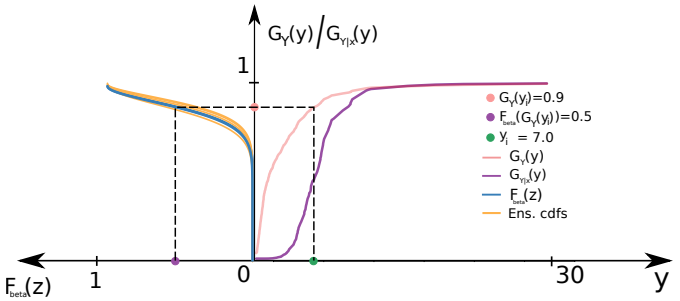


Figure 4.2: Illustration of the transformation procedure from the climatology cdf $G_Y(y)$ in red, via the unit scale cdf $F_{beta}(z)$ rotated counterclockwise 90° in blue, resulting in the postprocessed cdf $G_{Y|x}(y)$ in purple. The unit scale cdf $F_{beta}(z)$ is a combination of five unit scale ensemble member pdfs.

The Beta-BMA cdf on original scale takes the form

$$F(y|x_1^*, ..., x_M^*) = \frac{1}{M} \sum_{m=1}^{M} G_Y^{-1}(F_{beta}(z|x_m^*)), \tag{4.4}$$

where $G_Y(\cdot)$ is the climatology cdf, and $F_{beta}(\cdot)$ denotes the cdf of $Z|x^*$, where $Z|x^*$ is defined in (4.1).

The estimation of the parameters in Beta-BMA can be done the same way as for original BMA. Recall that M denotes the total number of ensemble members and K is the length of the training period. First, the bias-correction parameters α and β are estimated by maximizing the log-likelihood function of $M \cdot (K-1)$ independent Beta-distributed random variables with bias-correction parameters α and β and precision parameter ϕ'

$$(\hat{\alpha}_{ML}, \hat{\beta}_{ML}, \hat{\phi'}_{ML}) = \underset{\alpha,\beta,\phi'}{\arg\max} \left(\sum_{k=1}^{K} \frac{1}{M} \sum_{m=1}^{M} \log\{ f_{beta}\left(z_k|x_{k,m}^*\right) \} \right).$$

An alternative way of getting estimates for the bias-correction parameters is to use Bayesian inference. The priors are

$$\pi(\alpha) \sim \mathcal{N}(0, \sigma_\alpha) \tag{4.5}$$
$$\pi(\beta) \sim \mathcal{N}(0, \sigma_\beta) \tag{4.6}$$
$$\pi(\phi') \sim \text{Gamma}(p, q), \tag{4.7}$$

where $p, q, \sigma_\alpha, \sigma_\beta$ needs to be specified. The joint posterior distribution takes the form

$$\pi(\alpha, \beta, \phi'|\mathbf{z}) \propto \prod_{k=1}^{K} f_{beta}(z_k|\alpha, \beta, \phi')\pi(\alpha)\pi(\beta)\pi(\phi'), \tag{4.8}$$

Posterior marginals can then be obtained. We are interested in point estimates, and if we use the Bayes estimator that minimizes the posterior expected loss with the MSE as loss function, which is the posterior mean, we get

$$\hat{\alpha}_{PM} = \mathbb{E}(\alpha|\mathbf{z}) = \int \alpha\pi(\alpha|\mathbf{z})\mathrm{d}\alpha \tag{4.9}$$

$$\hat{\beta}_{PM} = \mathbb{E}(\beta|\mathbf{z}) = \int \beta\pi(\beta|\mathbf{z})\mathrm{d}\beta. \tag{4.10}$$

$$\tag{4.11}$$

After obtained bias-correction estimates, we fix them and ϕ can be estimated from the same training period based on the likelihood function of the Beta-BMA probabilistic forecast on unit scale, i.e.

$$\hat{\phi}_{ML} = \arg\max_{\phi} \prod_{k=1}^{K} \frac{1}{M} \sum_{m=1}^{M} f_{beta}(z_k|x^*_{k,m}). \tag{4.12}$$

This expression can be optimized numerically.

4.2 BMA with Gaussian Probability Density Functions and Varying Coefficient Regression

In the third method we apply the VCR model to the BMA methodology. Instead of estimating the bias-correcting-parameters from a sliding window training period, we let them evolve according to a VCR model.

When specifying the form of μ_m we use the same notation as in Section 3.7.2 where the INLA methodology is described. The mean μ_m takes the form of a structured additive predictor defined in Equation (3.44) with the identity link $g(\mu_m) = \mu_m$. With one deterministic forecast x_m, μ_m takes the form

$$\mu_m = \alpha + \beta x_m + \sum_{j=1}^{J} f_j(R_j). \tag{4.13}$$

Furthermore, let the modifying variable R_j represent time, denoted by t, and we define $\alpha_t = f_1(t)$ and $\beta_t = f_2(t)$. We now impose the first order random walk structure on α_t and β_t,

$$\alpha_t|\alpha_{t-1} \sim \mathcal{N}(\alpha_{t-1}, A) \tag{4.14}$$
$$\beta_t|\beta_{t-1} \sim \mathcal{N}(\beta_{t-1}, B). \tag{4.15}$$

This corresponds to the DLM model described in Section 3.6 with constant smoothing parameter A and B. We refer to the parameters A and B as smoothing parameters since a smaller value of A and B implies less flexibility for the state variables, or bias-correction parameters, α_t and β_t, respectively. If we gather the parameters for time steps $t = 1, ..., T$, the vectors $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_T)$ and $\boldsymbol{\beta} = (\beta_1, ..., \beta_T)$ form Gaussian Markov random fields.

To summarize, if Y_m denotes the random variable representing the probabilistic discharge for ensemble member m, the model takes the form of a dynamic linear model, i.e.

$$Y_m|x_m \sim \mathcal{N}(\mu_m, \tau^2|x_m) \tag{4.16}$$
$$\mu_m = (\alpha + \alpha_t) + (\beta + \beta_t)x_m \qquad m = 1, ..., M \tag{4.17}$$
$$\alpha_t = \alpha_{t-1} + a_t, \qquad a_t \sim N(0, A) \tag{4.18}$$
$$\beta_t = \beta_{t-1} + b_t, \qquad b_t \sim N(0, B). \tag{4.19}$$

The interpretation of the intercept parameter $\alpha + \alpha_t$ is the same as discussed in the CCPR method in Section 3.5. However, in the Beta-BMA-VCR approach, the intercept parameter consists of two terms, where the first term, α, represent the total bias between the deterministic forecasts and corresponding observations if the smoothing parameter $C = 0$, and α_t are evolving according to a dynamic linear model, and determines how much the total intercept parameter should be allowed to vary based on the smoothing parameter A. The same is valid for the slope parameter $\beta + \beta_t$.

A challenging issue with the original BMA methodology for streamflow forecasting is that the sliding window estimation process leads to poor bias-correction after the ensemble forecasts predicted wrongly. This happens after extreme events caused by heavy rainfall or snow melting. By including the random walk structure, one can easier control the range of values on the intercept and slope parameters by tuning A and B.

The bias-correction parameters are now random, in contrast to original BMA, where the bias-correction estimates are fixed but unknown. The third and last parameter, the variance parameter τ^2, is still considered as a fixed quantity. It is possible to include a random walk structure for the variance parameter as well, but the inference gets a lot more complicated as the variance parameter is on the upper level of the INLA hierarchical structure described in the previous chapter.

In the BMA-VCR model, the probabilistic forecast takes the form

$$f(y|x_1, ..., x_M) = \frac{1}{M} \sum_{m=1}^{M} g(y|x_m), \tag{4.20}$$

where $Y|x_m \sim \mathcal{N}(\mu_m, \tau^2)$ and μ_m is given in Equation (4.17), and τ^2 is the variance. The parameter estimation in BMA-VCR is done in a Bayesian framework. The procedure for obtaining marginal distributions for each element of the latent field $\boldsymbol{\xi}$ is described in Section 3.7.2. As for the original BMA, we first estimate the Gaussian latent field parameters $\boldsymbol{\xi} = (\alpha, \beta, \alpha_t, \beta_t)$ and then the variance parameter. Similar to the BMA model, when modeling the mean, we use a temporary variance parameter τ' instead of the variance parameter τ for the individual ensemble member pdfs. The first step in the estimation procedure is to assign priors to the the hyperparameters τ', A and B. However, to avoid overfitting we let the hyperparameters A and B be fixed and set these by optimizing the predictive performance according to mean CRPS. Furthermore, note that the standard deviation of the individual ensemble member pdfs in the BMA-VCR model, τ is assumed to be fixed, but the temporary variance parameter τ' is random and needs to be assigned a prior. The priors are

$$\pi(\alpha) \sim \mathcal{N}(0, \sigma_\alpha) \tag{4.21}$$
$$\pi(\beta) \sim \mathcal{N}(0, \sigma_\beta) \tag{4.22}$$
$$\pi(\alpha_t|\alpha_{t-1}, A) \sim \mathcal{N}(\alpha_{t-1}, A) \tag{4.23}$$
$$\pi(\beta_t|\beta_{t-1}, B) \sim \mathcal{N}(\beta_{t-1}, B) \tag{4.24}$$
$$\pi(\tau') \sim \text{Gamma}(p, q). \tag{4.25}$$

The joint posterior distribution takes the form

$$\pi(\alpha, \beta, \alpha_T, \beta_T, \tau'|\mathbf{y}) \propto \prod_{t=1}^{T} \pi(y_t|\alpha, \beta, \alpha_t, \beta_t, \tau')\pi(\alpha)\pi(\beta) \tag{4.26}$$
$$\times \pi(\alpha_t|\alpha_{t-1}, A)\pi(\beta_t|\beta_{t-1}, B)\pi(\tau'), \tag{4.27}$$

where A and B are assumed fixed. Posterior marginals can then be obtained. We are interested in point estimates, and we use the Bayes estimator that minimizes the posterior expected loss with the MSE as loss function, which corresponds to the posterior mean,

$$\hat{\xi}_{PM}^{i} = \mathbb{E}(\xi_i|\mathbf{y}) = \int \xi_i \pi(\xi_i|\mathbf{y}) \mathrm{d}\xi_\mathrm{i}, \tag{4.28}$$

where ξ_i denote an element in $\boldsymbol{\xi} = (\alpha, \beta, \alpha_t, \beta_t)$. The variance parameter τ is then estimated by maximum likelihood estimation from a training period of length K. By optimizing we get

$$\hat{\tau}_{ML} = \arg\max_{\tau} \sum_{k=1}^{K} \log\left(\frac{1}{M}\sum_{m=1}^{M} g(y|x_{k,m})\right), \tag{4.29}$$

where $g(y|x_m)$ is normally distributed with mean $\mu_{k,m}$, given in Equation (4.17), and variance τ^2. This expression can be optimized numerically.

4.3 BMA with Beta Probability Density Functions and Varying Coefficient Regression

We have presented two postprocessing approaches, each of them dealing with issues that need to be considered in streamflow forecasting. Beta-BMA accounts for the fact that streamflow values are non negative, and BMA-VCR makes it possible to make better parameter estimates after extreme events. In the third model, Beta-BMA-VCR, we combine the two new postprocessing techniques.

The structured additive predictor is

$$\eta_m = \alpha + \beta x_m + \sum_{j=1}^{J} f_j(R_j), \tag{4.30}$$

and the mean parameter μ_m is connected to the linear predictor through the logit link function. As for BMA-VCR, let R_j represent time, and define the same nonlinear functions $\alpha_t = f_1(t)$ and $\beta_t = f_2(t)$. This results in the following model

$$Z_m|x_m^* \sim \text{Beta}(\mu_m, \phi) \tag{4.31}$$

$$\mu_m = \text{logit}^{-1}\left((\alpha + \alpha_t) + (\beta + \beta_t)x_m^*\right) \qquad m = 1, ..., M \tag{4.32}$$

$$\alpha_t = \alpha_{t-1} + a_t, \qquad a_t \sim N(0, A) \tag{4.33}$$

$$\beta_t = \beta_{t-1} + b_t, \qquad b_t \sim N(0, B), \tag{4.34}$$

and the probabilistic Beta-BMA-VCR cdf on original scale takes the form

$$F(y|x_1^*, ..., x_M^*) = \frac{1}{M}\sum_{m=1}^{M} G_Y^{-1}(F_{beta}(z|x_m^*)), \tag{4.35}$$

The estimation procedure is similar to the procedure in the BMA-VCR model where the posterior mean is used as point estimates for the Gaussian latent field $\boldsymbol{\xi} = (\alpha, \beta, \alpha_t, \beta_t)$, and the precision parameter ϕ is estimated by maximum likelihood estimation. Furthermore, the smoothing parameters A and B are determined by optimizing the predictive performance according to the CRPS over a training period.

Chapter 5

Case Study: Postprocessing Models and Methods for the Osali Catchment

In this chapter, the models and methods used in the case study are presented. Four different models are considered. The period from June 24, 2014 to June 22, 2015 is analyzed. Subscript characters that are specific for the case study are included: Issue time $i = 1, ..., 364$, ensemble member number $m = 1, ..., 51$, lead time $l = 0, 4, 9$, and validation day $t = i + l$.

5.1 Model 1: BMA

The first model is the original BMA methodology used by Raftery et al. (2005) to produce probabilistic forecasts for temperature. The BMA probabilistic forecast is given by

$$f(y|x_{i,1,l}, ..., x_{i,M,l}) = \frac{1}{M} \sum_{m=1}^{M} g(y|x_{i,m,l}) \tag{5.1}$$

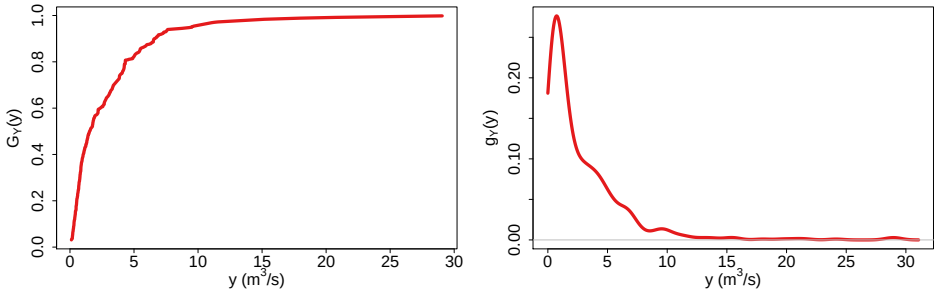
$$Y|x_{i,m,l} \sim \mathcal{N}(\mu_{t,m}, \tau^2) \tag{5.2}$$

$$\mu_{t,m} = \alpha + \beta x_{i,m,l} \qquad m = 1, ..., M, \tag{5.3}$$

and parameter estimates are obtained from a sliding window training period of length K. The estimation procedure is described in Chapter 3.2.

5.2 Model 2: Beta-BMA

The Beta-BMA model is described in Section 4.1. The empirical cumulative cdf $G_Y(y)$ is based on observations in the period from April 24, 2014 to January 27, 2016 from the Osali catchment, and it is shown in Figure 5.1a. The empirical pdf $g_Y(y)$ is seen in Figure 5.1b.



(a) Empirical climatology cdf.

(b) Empirical climatology pdf.

Figure 5.1: The climatology based on observations in the period April 24, 2014 to January 27 2016.

The Beta-BMA cdf at day $t = i + l$ becomes

$$F(y|x^*_{i,m,l}, ..., x^*_{i,M,l}) = \frac{1}{M}\sum_{m=1}^{M} G_Y^{-1}(F_{beta}(z|x^*_{i,m,l})) \tag{5.4}$$

$$Z|x^*_{i,m,l} \sim \text{Beta}(\mu_{t,m}, \phi) \tag{5.5}$$

$$\mu_{t,m} = \text{logit}^{-1}(\alpha + \beta x^*_{i,m,l}) \qquad m = 1, ..., M. \tag{5.6}$$

For the case study, posterior mean estimates are used for the Beta-BMA model. The priors are given in Equation (4.5), (4.6), and (4.7). For ϕ', a vague prior is chosen, the gamma distribution with shape parameters $p = 1$ and $q = 0.00005$. For α and β we choose $\sigma_\alpha = 0$ and $\sigma_\beta = 10^{-5}$, respectively. The estimation procedure is described in Section 4.1.

5.3 Model 3: BMA-VCR

The BMA-VCR model is provided in Section 4.2, and the probabilistic forecast takes the form

$$f(y|x_{i,1,l},...,x_{i,M,l}) = \frac{1}{M}\sum_{m=1}^{M} g(y|x_{i,m,l}) \tag{5.7}$$

$$Y|x_{i,m,l} \sim \mathcal{N}(\mu_{t,m}, \tau^2) \tag{5.8}$$

$$\begin{cases} \mu_{t,m} = (\alpha + \alpha_t) + (\beta + \beta_t)x_{i,m,l} \\ \mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l} \end{cases} \tag{5.9}$$

$$\alpha_t = \alpha_{t-1} + a_t, \qquad a_t \sim N(0, C) \tag{5.10}$$

$$\beta_t = \beta_{t-1} + b_t, \qquad b_t \sim N(0, C), \tag{5.11}$$

where Equation (5.9) describes two different forms for $\mu_{t,m}$. The smoothing parameter C is restricted to be the same for the α_t and the β_t process.

The random walk parameter β_t is plotted for different values of C over the period from October 3 to November 3 in Figure 5.2. The lightest blue line has smoothing parameter $C = 0.14$. The darkest blue line has smoothing parameter $C = 0$, which gives no flexibility, meaning $\beta_t = 0$ for all days in the period. The other lines has smoothing parameter values in between. The lightest blue line is an example of overfitting, where the regression model provides a good fit to the given data, but does generally not provide good predictions.

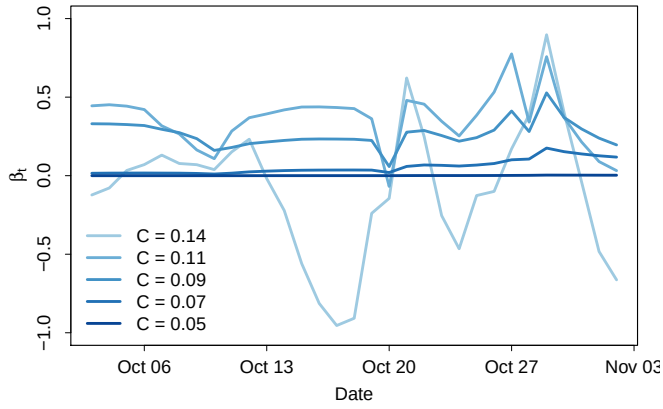


Figure 5.2: The random walk parameter β_t, plotted for smoothing parameter values between $C = 0$ and $C = 0.14$ in the period from October 3 to November 3.

One way to obtain an estimate for the smoothing parameter C is to minimize the CRPS. The special case $C = 0$ correspond to different cases depending on the mean model. The first mean model in Equation (5.9) simplifies to $\mu_{t,m} = \alpha + \beta x_{i,m,l}$, which corresponds to the original BMA technique using all historical observations as training period for bias-correction estimates, instead of a sliding window. The second mean model in Equation

(5.9) simplifies to $\mu_{t,m} = x_{i,m,l}$ which corresponds to the original BMA technique without any bias-correction. The advantage with the second mean model is that for the Osali catchment, there is only a small total forecast-observation bias as seen in Figure 2.6, and by only including the time-varying parameters in the mean model, the model is less affected by events with large forecast-observation errors.

Posterior mean estimates are used as bias-correction estimates for the BMA-VCR model. The same priors as for Beta-BMA are used. The estimation procedure for the bias-correction parameters $\alpha, \alpha_t, \beta, \beta_t$ and variance parameter τ^2 is described in Section 4.2.

An alternative to fitting the variance of the probabilistic forecast is to only bias-correct the raw ensemble forecast according to the BMA-VCR model. For the Osali catchment we include an approach where the mean model $\mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l}$ is the bias-corrected ensemble member, without imposing variance. The CRPS of the raw ensemble, and the bias-corrected raw ensemble can be obtained using Equation (3.4) and it can be compared with the CRPS for the probabilistic forecasts of the BMA approaches.

5.4 Model 4: Beta-BMA-VCR

The last model is the Beta-BMA-VCR model which is a combination of the previous two. The Beta-BMA-VCR model is described in Chapter 4.3 and the predictive cdf is

$$F(y|x^*_{i,1,l}, ..., x^*_{i,M,l}) = \frac{1}{M}\sum_{m=1}^{M} G_Y^{-1}(F_{beta}(z|x^*_{i,m,l})) \tag{5.12}$$

$$Z|x^*_{i,m,l} \sim \text{Beta}(\mu_{t,m}, \phi) \tag{5.13}$$

$$\mu_{t,m} = \text{logit}\left((\alpha + \alpha_t) + (\beta + \beta_t)x^*_{i,m,l}\right) \qquad m = 1, ..., M \tag{5.14}$$

$$\alpha_t = \alpha_{t-1} + a_t, \qquad a_t \sim N(0, C) \tag{5.15}$$

$$\beta_t = \beta_{t-1} + b_t, \qquad b_t \sim N(0, C). \tag{5.16}$$

Posterior mean estimators are used as parameter estimates. The estimation procedure is described in Chapter 4.

5.5 Training Period

In BMA and Beta-BMA, the bias-correction parameters, α and β are estimated by maximum likelihood estimation from a K days sliding window training period, i.e. from day $t = i-l-K$ to $t = i-l-1$. In BMA-VCR and Beta-BMA-VCR, the bias-correction parameters, α, α_t, β, and β_t are estimated by the Bayes estimator being the posterior mean of the respective marginal distributions which uses all days available as training period, i.e. from day $t = 1$ to $t = i - l - 1$. After obtained bias-correction estimates, we fix

these and then estimate the variance parameter τ, or the precision parameter ϕ, from the same K days sliding window training period as the bias-correction parameters in BMA and Beta-BMA. This applies to all models in this study.

One way to choose the length of the training period is to set K such that CRPS is minimal. However, since we are exploring several models, and to reduce the computational effort, we choose to consider a short training period with $K = 10$ days and a longer one with $K = 60$ days for each model, and compare the results.

5.6 Predictive Performance

The smoothing parameter C is being optimized based on predictive performance according to the CRPS, given in Equation (3.3). The smoothing parameter C is fixed for all days in the period being analyzed, and we use the same value C for both the evolution processes α_t and β_t. It is being optimized by trying different values for C, keeping in mind that a large C-value leads to overfitting, and by choosing $C = 0$, the mean parameter are not allowed vary locally. The smoothing parameter C influence all other model parameter estimates. First, it affects the bias-correction parameters $\alpha, \beta, \alpha_t, \beta_t$ by determining how much α_t and β_t can vary. This again affect the variance parameter τ, or precision parameter ϕ, because they are estimated based on fixed values of bias-correction parameters. Furthermore, the parameter estimates at given days changes as more data becomes available, meaning that the estimation procedure does not allow a sequential update of parameter estimates. The optimization in the case study for the Osali Catchment is done by trying different values for C, and then model one year from June 24, 2014 to June 22, 2015. The C-value corresponding to the lowest mean CRPS has the best predictive performance, according to mean CRPS. The results are provided in the next chapter.

Chapter 6

Case Study: Results from the Osali Catchment

In this chapter, the results from the case study of the Osali Catchment is presented. The period from June 24, 2014 to June 22, 2015 is analyzed. In order to distinguish between the BMA postprocessing approaches, we use color codes and abbreviations for each approach:

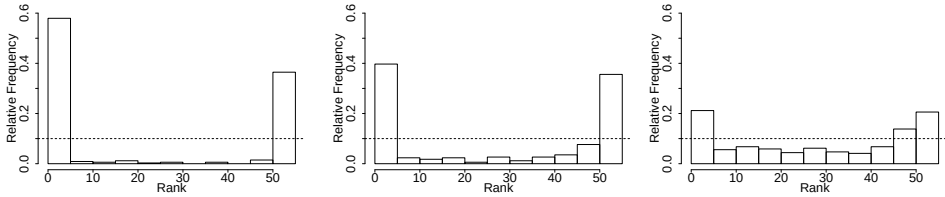
- BMA and Beta-BMA with 60 days training period for all model parameters (BMA-K60, Beta-BMA-K60) are given in orange in the figures.
- BMA and Beta-BMA with 10 days training period for all model parameters (BMA-K10, Beta-BMA-K10) are given in purple in the figures.
- BMA-VCR and Beta-BMA-VCR with mean model $\mu_{t,m} = (\alpha+\alpha_t)+(\beta+\beta_t)x_{i,m,l}$ and 60 days training period for the variance/precision parameter (BMA-VCR-M1-K60, Beta-BMA-VCR-K60) are given in blue in the figures.
- BMA-VCR and Beta-BMA-VCR with mean model $\mu_{t,m} = (\alpha+\alpha_t)+(\beta+\beta_t)x_{i,m,l}$ and 10 days training period for the variance/precision parameter (BMA-VCR-M2-K10, Beta-BMA-VCR-K10) are given in red in the figures.
- BMA-VCR with mean model $\mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l}$ and 60 days training period for the variance parameter (BMA-VCR-M2-K60) is given in brown in the figures.
- BMA-VCR with mean model $\mu_{t,m} = \alpha_t+(1+\beta_t)x_{i,m,l}$ and 10 days training period for the variance parameter (BMA-VCR-M2-K10) is given in green in the figures.

This Chapter is divided into five sections. The first section, presents the results for the deterministic raw ensemble forecasts for lead time 0, 4, and 9. Further, there are three

sections with the results for the BMA postprocessing approaches for each of these lead times separately. The last section contains a comparison between the approaches.

6.1 The Raw Ensemble

The VRH for lead time 0, 4, and 9 for the raw ensemble forecasts is seen in Figure 6.1a, 6.1b, and 6.1c, respectively. It can be seen that the VRH for lead time 0 is strongly U-shaped and slightly biased. If the histogram is U-shaped it means the corresponding observation too often is outside the range of ensemble member forecasts. The VRH for lead time 4 is less U-shaped. For lead time 9 we observe that the observation more often takes place within the ensemble range, compared to lead time 0 and 4, resulting in a VRH that is closer to uniform, but slightly U-shaped. The VRHs indicate that the need of statistical postprocessing approaches is important to obtain calibrated forecasts, specially for for lead time 0 and lead time 4.



(a) Lead time 0. **(b)** Lead time 4. **(c)** Lead time 9.

Figure 6.1: Verification rank histograms.

Mean CRPS for the raw ensembles was calculated to 0.57, 0.89, and 0.98 for lead time 0, 4, and 9, respectively. The CRPS for the raw ensemble for each lead time is given as a black horizontal line in Figure 6.2, 6.6, 6.10, 6.11, 6.12, 6.16, in the following sections.

6.2 BMA Postprocessing Approaches for Lead Time 0

In this section, the results when applying the BMA postprocessing approaches from Section 5 for lead time 0 are presented. We let the optimal value C refer to the value of C with the lowest mean CRPS of the values tested.

6.2.1 Results using Gaussian Probability Density Functions and Varying Coefficient Regression

Figure 6.2 shows the CRPS for the BMA-VCR approaches for different values of the smoothing parameter C. The BMA-K10 and BMA-K60 approaches do not have a dynamic

linear model structure for the bias-correction parameters, the CRPS is therefore given as horizontal lines Figure 6.2. Where lines intersects means that the predictive performance, according to mean CRPS is equally good. It can be seen that a high value of C leads to overfitting and gives poor predictive performance. We further observe that the BMA-VCR-M2-K10 has a minimum around $C = 0.09$, and that the minimum for BMA-VCR-M1-K10 is slightly higher. Recall that the M1-model is $\mu_{t,m} = (\alpha + \alpha_t) + (\beta + \beta_t)x_{i,m,l}$ and the M2-model is $\mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l}$. Smoothing parameter value $C = 0.09$ is used for further analysis. The black dashed varying line is mean CRPS of the bias-corrected ensemble from mean model M2, and the black dashed horizontal line is the raw ensemble.

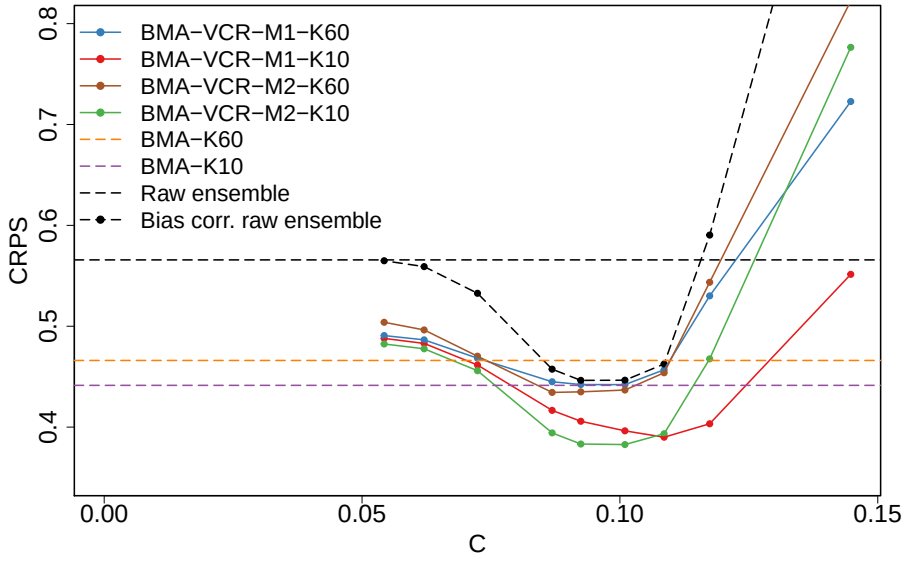
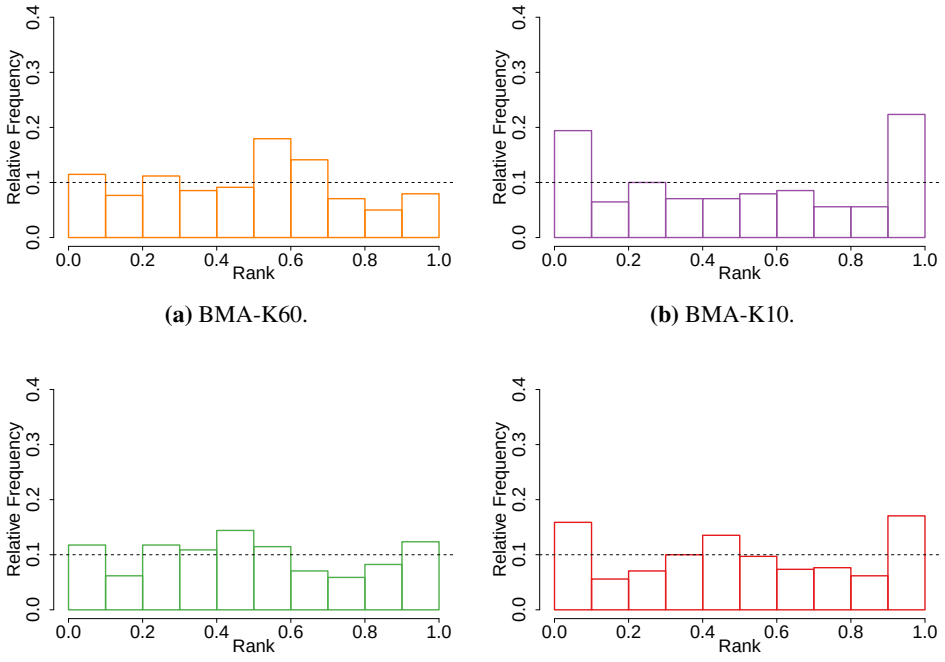


Figure 6.2: The CRPS for the BMA-VCR approaches during the period June 24, 2014 to June 22, 2015 as a function of the precision parameter C in Equation (5.10) and (5.11). The orange and brown dashed lines show the CRPS using the standard BMA methodology, using the normal distribution as the individual ensemble member pdfs, with parameter estimates from a 60 days and 10 days training periods, respectively. The bias-corrected raw ensemble, according to mean model M2, is given by the black dashed varying line, and the raw ensemble is the black dashed horizontal line. The lead time is $l = 0$.

Figure 6.3a and 6.3b show the PIT histogram using a sliding window training period of 60 days and 10 days, respectively. We observe that the histogram of the BMA-K10 approach is U-shaped, indicating underdispersion. The dashed line show the case when the probabilistic forecast is perfectly calibrated. The histogram for the BMA-K60 approach is closer to uniform. Figure 6.3c and 6.3d show the PIT histograms for the BMA-VCR-M2-K10 approach and BMA-VCR-M1-K10 approach, respectively, with smoothing parameter $C = 0.09$, which is a optimal value for the BMA-VCR-M2-K10 approach. We observe

that the forecasts obtained from both approaches are well calibrated.



(a) BMA-K60.

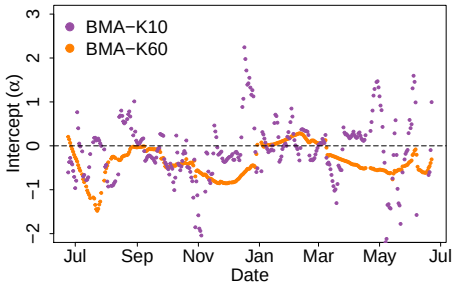
(b) BMA-K10.

(c) BMA-VCR-M2-K10.

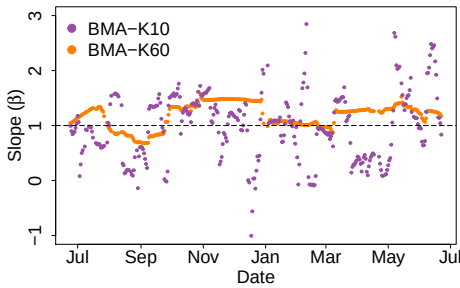
(d) BMA-VCR-M1-K10.

Figure 6.3: PIT histograms for different BMA approaches for lead time 0.

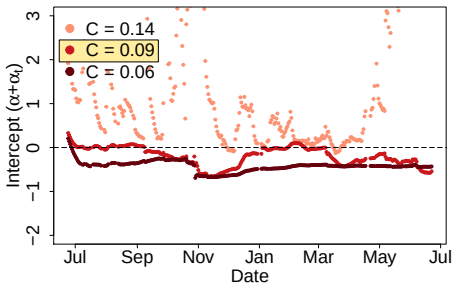
Figure 6.4a and 6.4b show the intercept parameter and slope parameter, respectively, for the original BMA approach, plotted against the day the parameter estimates are valid, i.e. day $t = i + l$. The dashed horizontal lines show the intercept and slope values when the mean parameter equals the ensemble member forecasts. We notice that using 60 days as a training period leads to less variation in the intercept and slope parameter. The intercept and slope parameter for the BMA-VCR-M1-K10 approach are seen in Figure 6.4c and 6.4d for different values of the precision parameter C. We observe that a relative large value gives high flexibility, which leads to overfitting. The less value the less flexibility. The yellow box indicates the optimal value for C for the BMA-VCR-M2-K10 approach from Figure 6.2. The parameters for the BMA-VCR-M2-K10 approach are seen in Figure 6.4e and 6.4f. It can be seen that the intercept parameter varies around 0 for the intercept and around 1 for the slope, as specified in the mean model. If C is small, we observe that the parameters do not deviate from 0 and 1, respectively, meaning that the mean of the M2 model is the individual ensemble member forecasts. However, for an optimal value for C, the parameters vary more.



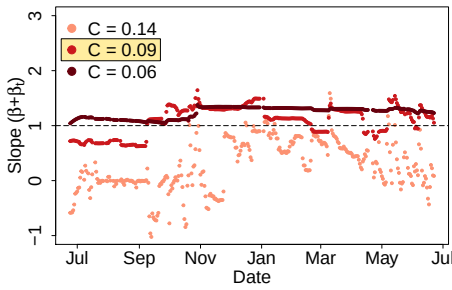
(a) Intercept α for BMA-K60 and BMA-K10.



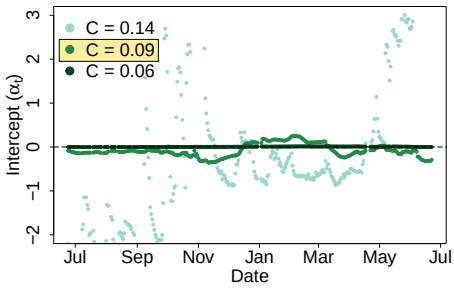
(b) Slope β for BMA-K60 and BMA-K10.



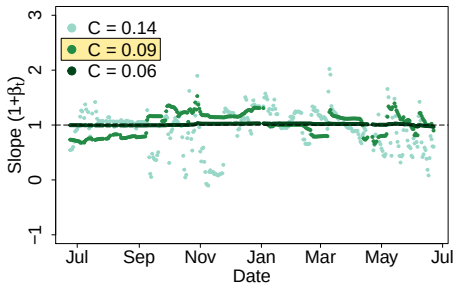
(c) Intercept $\alpha + \alpha_t$ for BMA-VCR-M1.



(d) Slope $\beta + \beta_t$ for BMA-VCR-M1.



(e) Intercept α_t for BMA-VCR-M2.



(f) Slope $1 + \beta_t$ for BMA-VCR-M2.

Figure 6.4: Parameter estimates plotted against validation day $t = i + l$. The yellow box shows an optimal value for the smoothing parameter C based on the BMA-VCR-M2-K10 approach. The lead time is $l = 0$.

Figure 6.5 shows the the variance parameter τ. For the VCR approaches the smoothing parameter C is set to be $C = 0.09$. We refer to τ as a variance parameter, but τ denotes the standard deviation of the individual ensemble member pdfs. We observe that for the $K10$ approaches, the variance varies more. For the $K60$ approaches, the variance remains large for $K = 60$ days after large forecast-observation errors. For the BMA-K60 and

the BMA-VCR-M2-K60 approaches, τ is estimated from the same training period, but the bias-correction parameter are set to be different, since they are estimated differently for the two approaches. This leads to similar trend, but different ML-estimates for τ, as seen in the figure. The same is valid for the BMA-K10 and the BMA-VCR-M2-K10 approaches.

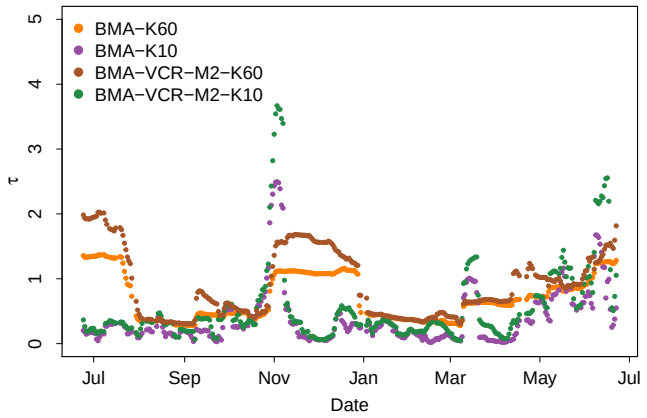


Figure 6.5: The variance parameter τ, which refers to the standard deviation of the individual ensemble member pdfs, plotted against the validation day $t = i + l$, for four different BMA post-processing approaches. The lead time is $l = 0$.

6.2.2 Results using Beta Probability Density Functions and Varying Coefficient Regression

The CRPS for the Beta-BMA-VCR-approaches for different values of the smoothing parameter C is given in Figure 6.6. It is seen that the Beta-BMA-VCR-K10 approach performs better than the Beta-BMA-VCR-K60 approach for smoothing parameter values around the optimal value $C = 0.11$. The purple and orange horizontal line is the CRPS for the Beta-BMA-K10 approach and the Beta-BMA-K60 approach, respectively.

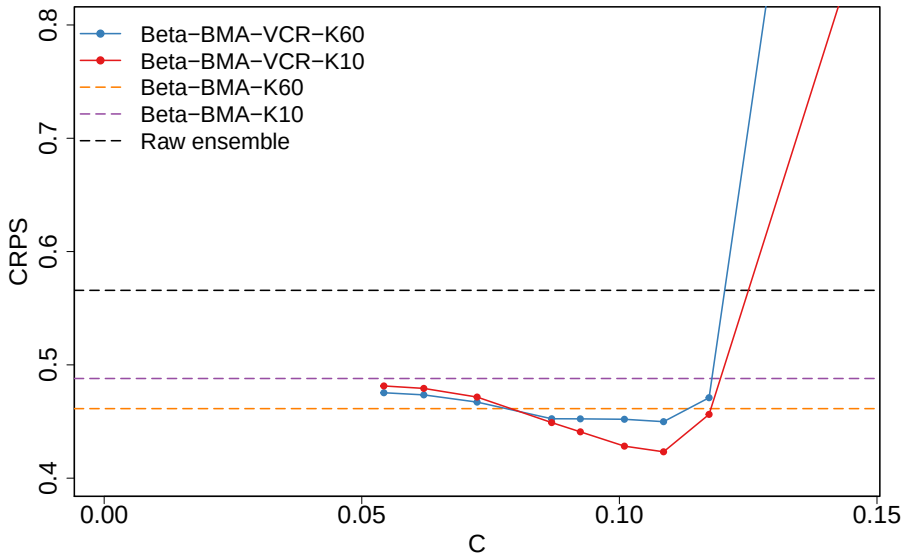
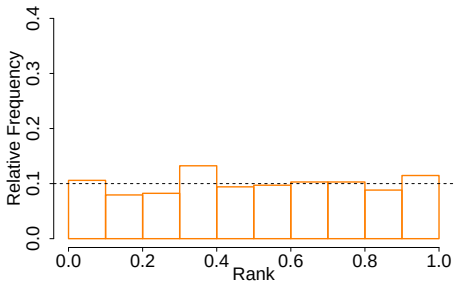
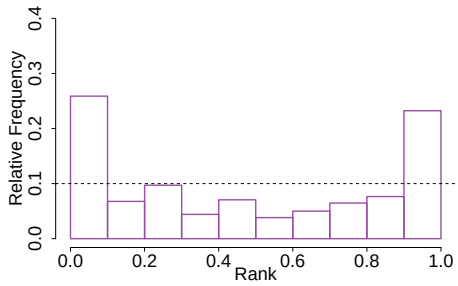


Figure 6.6: The CRPS for the Beta-BMA-VCR approaches from the period June 24, 2014 to June 22, 2015 as a function of the precision parameter C in equation (5.15) and (5.16). The orange and brown dashed lines show the CRPS using the Beta-BMA approach, with parameter estimates from a 60 days and 10 days training periods, respectively. The lead time is $l = 0$. The horizontal black line is the CRPS of the raw ensemble.

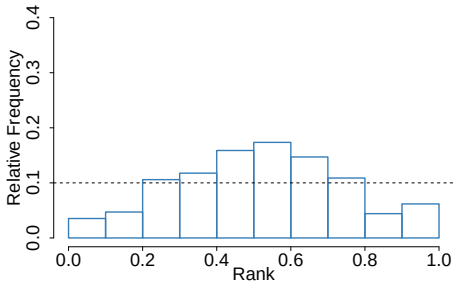
Figure 6.7a and 6.7b show the PIT histograms for the Beta-BMA approaches with 60 days training period and 10 days training period, respectively. Using 10 days sliding window as training period results in a U-shaped histogram, indicating underdispersion. A 60 days sliding window yields a calibrated forecast. Figure 6.7c and 6.7d give the PIT histograms for the Beta-BMA-VCR-K60 approach and the Beta-BMA-VCR-K10 approach, respectively. We observe that the histogram in Figure 6.7c is hump-shaped, meaning that the variance of the probabilistic forecast on average is too large. For the Beta-BMA-VCR-K10 approach, the forecast is well calibrated.



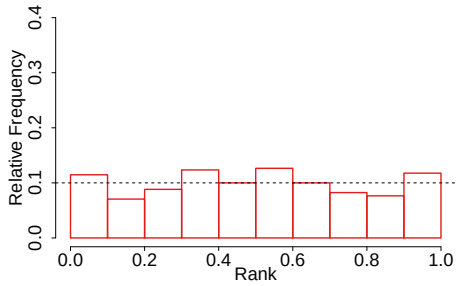
(a) Beta-BMA-K60.



(b) Beta-BMA-K10.



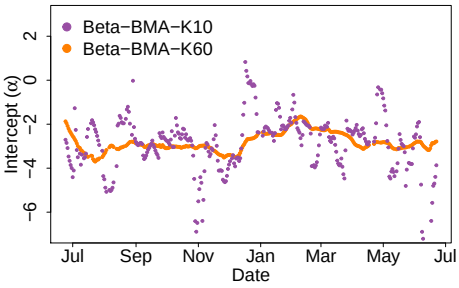
(c) Beta-BMA-VCR-K60.



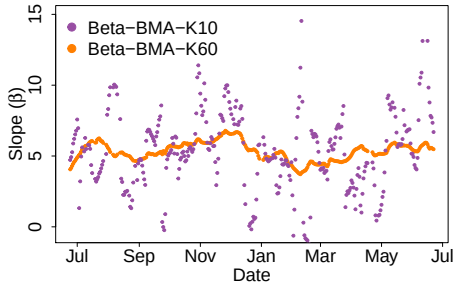
(d) Beta-BMA-VCR-K10.

Figure 6.7: PIT histograms for Beta-BMA approaches for lead time 0.

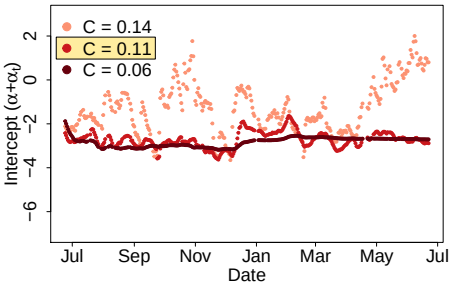
The intercept parameter α and the slope parameter β for the Beta-BMA approaches are seen in Figure 6.8a and 6.8b, respectively. The intercept parameter $\alpha + \alpha_t$ and slope parameter $\beta + \beta_t$ for the Beta-BMA-VCR-K10 approach are shown in Figure 6.8c and 6.8d, respectively. For the optimal value $C = 0.11$ we observe that the the intercept parameter varies around -3 and the slope parameter around 5, which means that the individual ensemble member forecasts has a large impact on the mean parameter, as discussed in Section 3.5.



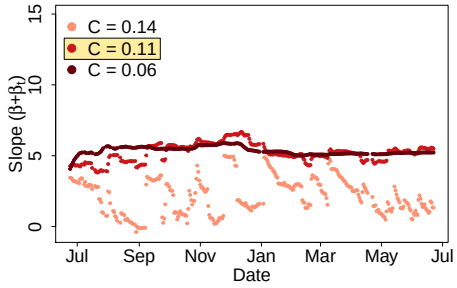
(a) Intercept α for Beta-BMA-K60 and Beta-BMA-K10.



(b) Slope $\beta$for Beta-BMA-K60 and Beta-BMA-K10.



(c) Intercept $\alpha + \alpha_t$ for Beta-BMA-VCR.



(d) Slope $\beta + \beta_t$ for Beta-BMA-VCR.

Figure 6.8: Parameter estimates plotted against the validation day $t = i+l$, for different Beta-BMA postprocessing approaches. The lead time is $l = 0$. The yellow box shows the optimal value for the smoothing parameter C.

The inverse of the precision parameter ϕ is plotted in Figure 6.9. The inverse of the precision parameter, can be thought of as a variance parameter, i.e. the larger value, the larger variance. However, the interpretation of ϕ, or ϕ^{-1}, is difficult since it is a precision, or variance, parameter for a pdf on unit scale. On original scale, the variance of the probabilistic forecast will in addition depend on the shape of the climatology cdf. For the Beta-BMA-VCR approaches the smoothing parameter C is set to be $C = 0.11$. We observe that the Beta-BMA-VCR-K60 approach has large variance parameter compared to the other approaches.

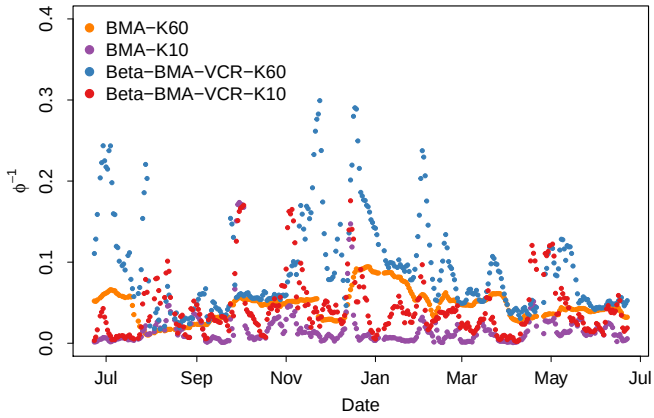


Figure 6.9: The variance parameter ϕ^{-1}, plotted against the validation day $t = i + l$, for four different Beta-BMA postprocessing approaches. The lead time is $l = 0$.

6.3 BMA Postprocessing Approaches for Lead Time 4

In this section, the results for lead time 4 is presented. We only present the CRPS for each approach. Figures related to parameter estimates and PIT histograms for lead time 4 can be seen in Appendix A.

6.3.1 Results using Gaussian Probability Density Functions and Varying Coefficient Regression

The CRPS for the BMA-VCR approaches for different smoothing parameter values is shown in Figure 6.10. We observe that the BMA-VCR-M2-K10 and the BMA-VCR-M2-K60 approaches have the lowest CRPS, with an optimal smoothing parameter value close to $C = 0.09$. It is further seen that the CRPS for the raw ensemble forecasts is slightly larger than the CRPS for the forecast obtained from the best performing method, the BMA-VCR-M2 approach. Furthermore, we observe that the bias-corrected ensemble forecasts, without being assigned an associated pdf, perform better than all the BMA-VCR approaches for $C = 0.09$. The bias-corrected raw ensemble, according to mean model M2, is given by the black dashed varying line, and the raw ensemble is the black dashed horizontal line.

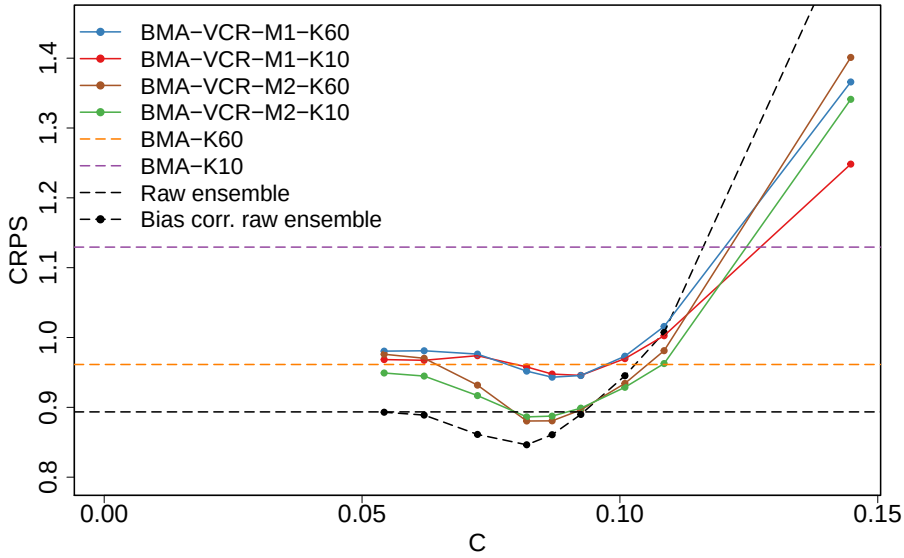


Figure 6.10: The CRPS for the BMA-VCR approaches during the period June 24, 2014 to June 22, 2015 as a function of the precision parameter C in Equation (5.10) and (5.11). The orange and brown dashed lines show the CRPS using the standard BMA methodology, using the normal distribution as the individual ensemble member pdfs, with parameter estimates from a 60 days and 10 days training periods, respectively. The bias-corrected raw ensemble, according to mean model M2, is given by the black dashed varying line, and the raw ensemble is the black dashed horizontal line. The lead time is $l = 4$.

6.3.2 Results using Beta Probability Density Functions and Varying Coefficient Regression

The CRPS for different values of the smoothing parameter C for the Beta-BMA-VCR approaches is shown in Figure 6.11. It can be seen that the Beta-BMA-VCR-K10 approach performs better than the Beta-BMA-VCR-K60 approach for $C < 0.11$, indicating that a long training period for the precision parameter is better than a short one, in contrast to the results for lead time 0. We further observe that mean CRPS for the Beta-BMA-VCR approaches are lower than the CRPS for the raw ensemble forecasts.

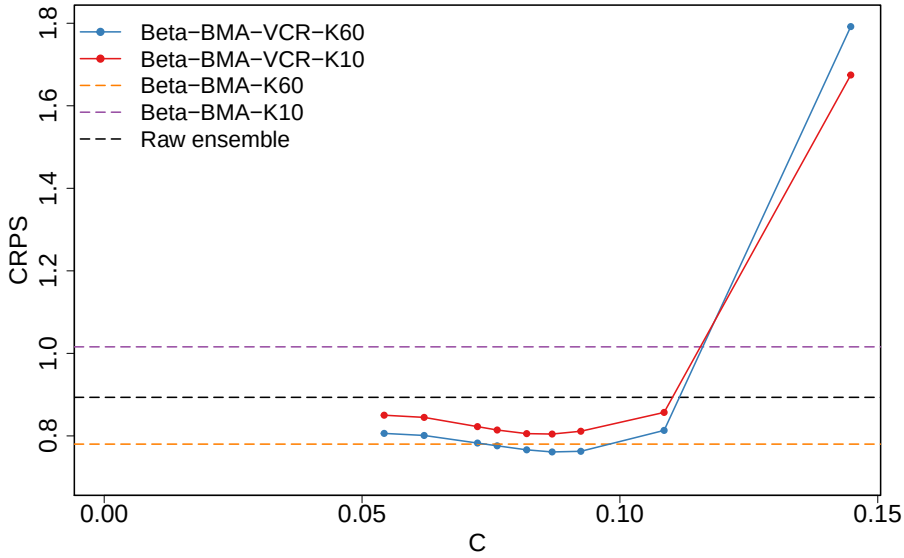


Figure 6.11: The CRPS for the Beta-BMA approaches during the period June 24, 2014 to June 22, 2015 as a function of the precision parameter C in equation (5.15) and (5.16). The orange and brown dashed lines show the CRPS using the standard Beta-BMA approach, with parameter estimates from a 60 days and 10 days training periods, respectively. The lead time is $l = 4$. The horizontal black line is the CRPS of the raw ensemble.

6.4 BMA Postprocessing Approaches for Lead Time 9

The results for lead time 9 are presented in this section.

6.4.1 Results using Gaussian Probability Density Functions and Varying Coefficient Regression

The CRPS for the BMA-VCR approaches for different smoothing parameter values are given in Figure 6.12. We observe that BMA-VCR-M2-K10 and BMA-VCR-M2-K60 approaches have the lowest CRPS among the BMA postprocessing approaches, similarly to lead time 4. However, the raw ensemble has better predictive performance. We further observe that it is possible to obtain a lower mean CRPS by bias-correct the raw ensemble with mean model M2, i.e. $\mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l}$, and smoothing parameter $C = 0.09$.

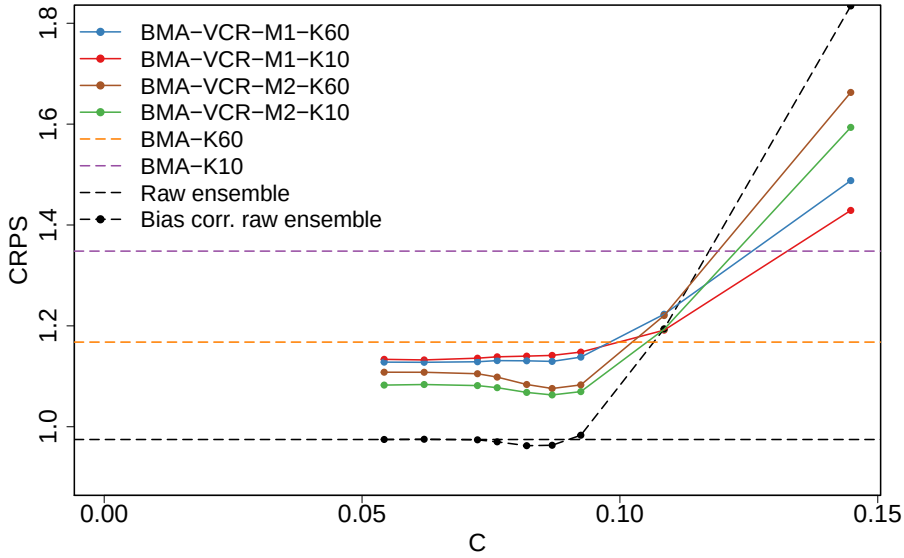
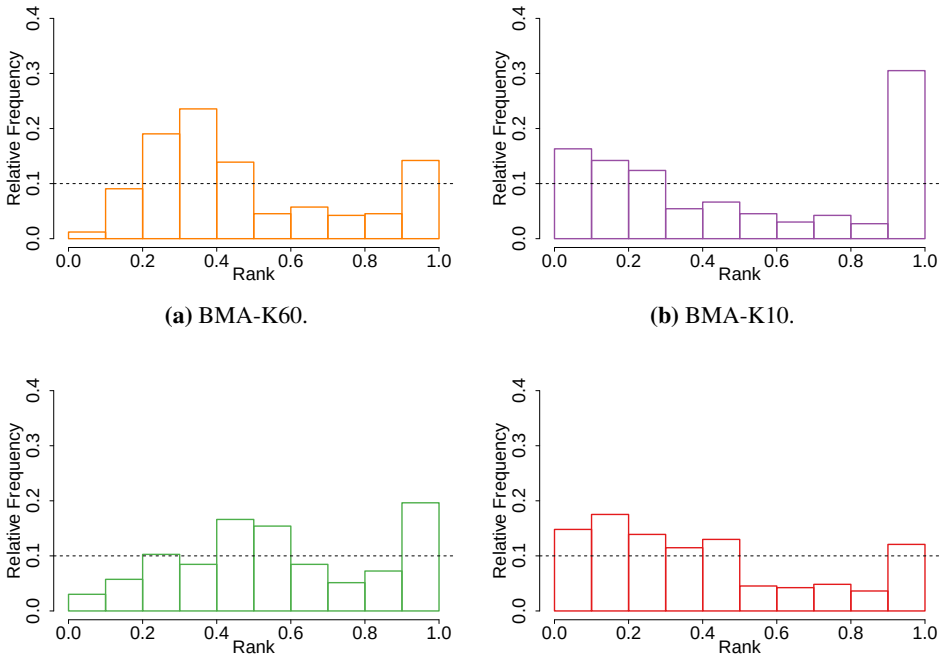


Figure 6.12: The CRPS for the BMA-VCR approaches during the period June 24, 2014 to June 22, 2015 as a function of the precision parameter C in Equation (5.10) and (5.11). The orange and brown dashed lines show the CRPS using the standard BMA approach, using the normal distribution as the individual ensemble member pdfs, with parameter estimates from a 60 days and 10 days training periods, respectively. The bias-corrected raw ensemble, according to mean model M2, is given by the black dashed varying line, and the raw ensemble is the black dashed horizontal line. The lead time is $l = 9$.

Figure 6.13a and 6.13b show the PIT histograms for the BMA-K60 and BMA-K10 approaches, respectively. The PIT histogram for the BMA-VCR-M2-K10 approach and the BMA-VCR-M1-K10 approach are seen in Figure 6.13c and 6.13d, respectively, with smoothing parameter $C = 0.09$. We observe that the probabilistic forecast for all four approaches are worse calibrated than the raw ensemble forecasts, which can be seen by the VRH in Figure 6.1c.



(a) BMA-K60.

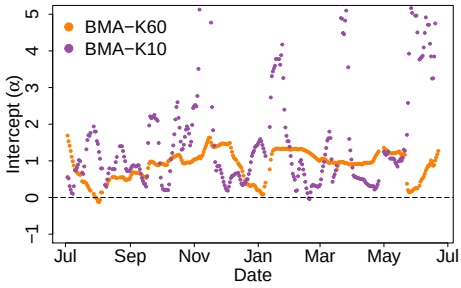
(b) BMA-K10.

(c) BMA-VCR-M1-K10.

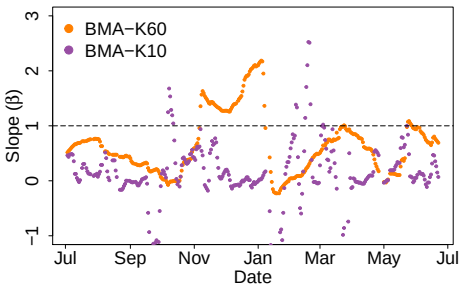
(d) BMA-VCR-M2-K10.

Figure 6.13: PIT histograms for BMA approaches for lead time 9.

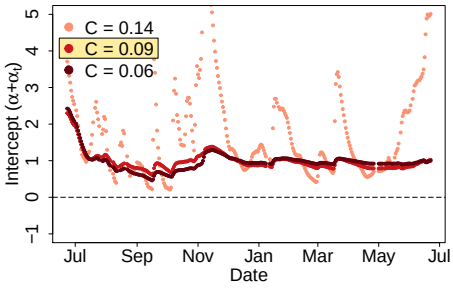
Figure 6.14a and 6.14b show the intercept parameter and slope parameter, respectively, plotted against the day they are valid, i.e. day $t = i + l$. We observe the sudden increase in the slope parameter after the large streamflow in October, as discussed in Chapter 1. The intercept and slope for the BMA-VCR-M1-K10 approach are given in Figure 6.14c and 6.14d for different values of the smoothing parameter C. The parameters for the BMA-VCR-M2-K10 approach are seen in Figure 6.14e and 6.14f. We observe that the parameter estimates, for the optimal value $C = 0.09$ are smoother compared the parameter estimates for the BMA-K60 approach. This results in a slightly better calibrated forecast, which is found by comparing the PIT histograms from Figure 6.13a and Figure 6.13c. However, both approaches are worse calibrated and have a higher CRPS than the raw ensemble forecasts.



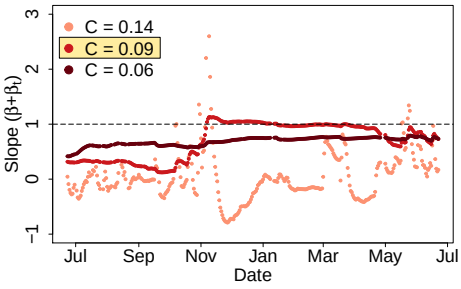
(a) Intercept α for BMA-K60 and BMA-K10.



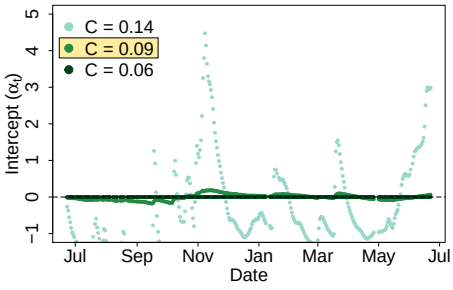
(b) Slope β for BMA-K60 and BMA-K10.



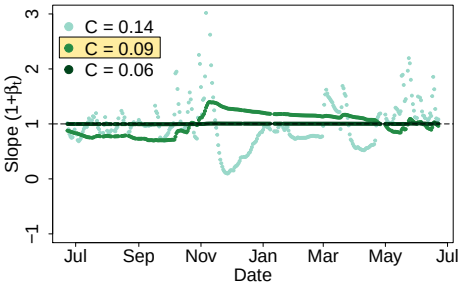
(c) Intercept $\alpha + \alpha_t$ for BMA-VCR-M1.



(d) Slope $\beta + \beta_t$ for BMA-VCR-M1.



(e) Intercept α_t for BMA-VCR-M2.



(f) Slope $1 + \beta_t$ for BMA-VCR-M2.

Figure 6.14: Parameter estimates plotted against validation day $t = i + l$. The yellow box shows a optimal value for the smoothing parameter C based on the BMA-VCR-M2-K10 approach. The lead time is $l = 9$.

Figure 6.15 shows the variance parameter τ plotted against the validation day $t = i + l$. For the BMA-VCR approaches the smoothing parameter C is set to be 0.09. We observe that the standard deviation of the probabilistic forecast for the $K10$ approaches is varying more than for the $K60$ approaches.

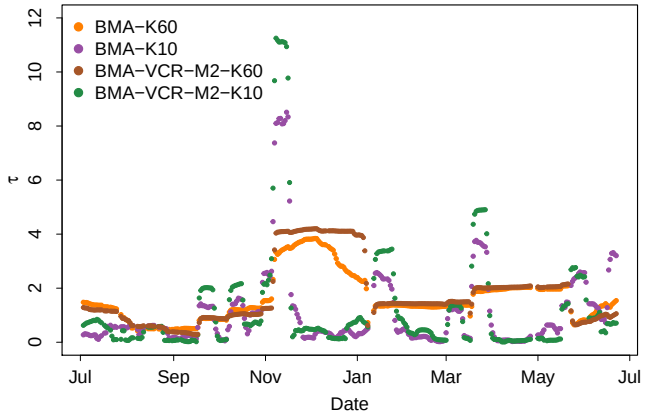


Figure 6.15: The variance parameter τ, which refers to the standard deviation of the probabilistic forecast, plotted against the validation day $t = i + l$, for four different BMA postprocessing approaches. The lead time is $l = 9$.

6.4.2 Results using Beta Probability Density Functions and Varying Coefficient Regression

The CRPS for the Beta-BMA approaches for different values of the smoothing parameter C is shown in Figure 6.16. It is seen that the Beta-BMA-VCR-K60 approach performs better than the Beta-BMA-VCR-K60 approach for $C < 0.11$. However, there is no clear optimal value for C. For the following analysis, we use $C = 0.06$. We further observe that the Beta-BMA-VCR-K60 approach obtains the same CRPS value as the raw ensemble forecasts, while all other approaches has higher mean CRPS.

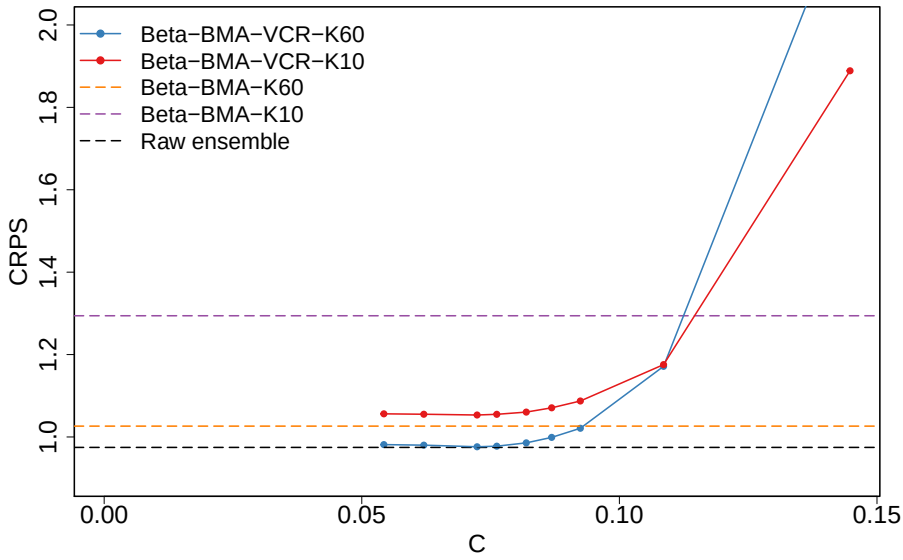
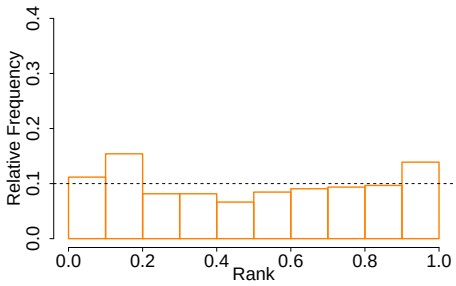
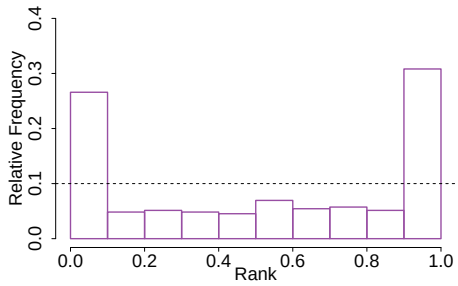


Figure 6.16: The CRPS from the period June 24, 2014 to June 22, 2015 as a function of the smoothing parameter C in equation (5.15) and (5.16). The orange and brown dashed lines show the CRPS using the Beta-BMA, with parameter estimates from a 60 days and 10 days training periods, respectively. The lead time is $l = 9$. The black horizontal line is the CRPS for the raw ensemble.

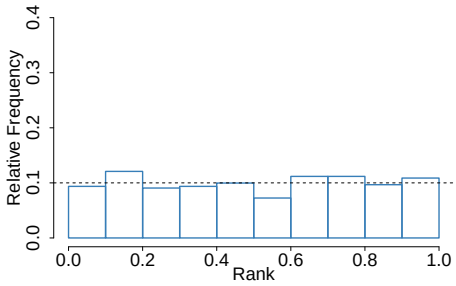
The PIT histograms for the Beta-BMA-K60 and Beta-BMA-K10 approaches are seen in Figure 6.17a and 6.17b, respectively. We observe that the probabilistic forecast for the Beta-BMA-K10 approach is underdispersed, meaning that the variance on average is too small. A 60 days training period gives a better calibrated forecast. Figure 6.17c and 6.17d show the PIT histogram for the Beta-BMA-VCR-K60 and the Beta-BMA-VCR-K10 approaches, respectively. Both forecasts are well calibrated. Comparing Figure 6.17c with the VRH for the raw ensemble forecasts in Figure 6.1c we find that the forecast from the Beta-BMA-VCR-K60 approach is better calibrated.



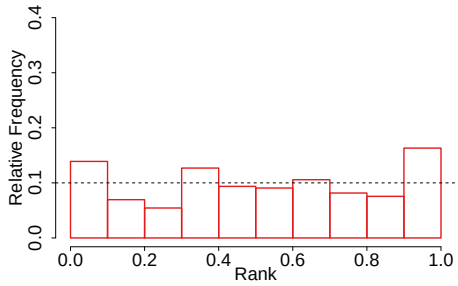
(a) Beta-BMA-K60.



(b) Beta-BMA-K10.



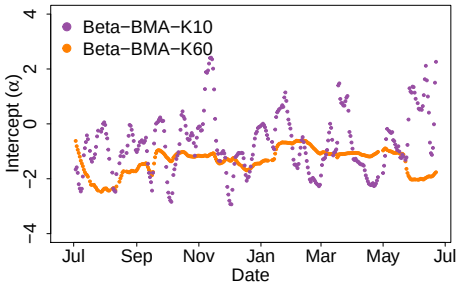
(c) Beta-BMA-VCR-K60.



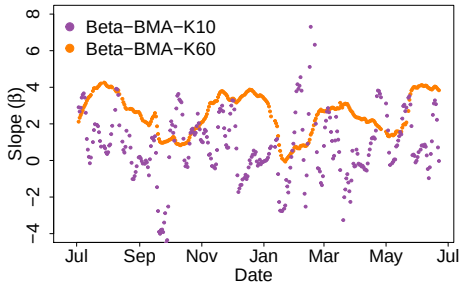
(d) Beta-BMA-VCR-K10.

Figure 6.17: PIT histograms for Beta-BMA approaches for lead time 9.

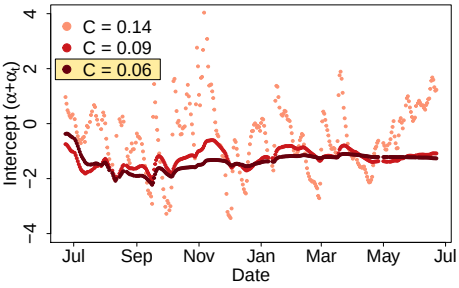
The intercept parameter α and slope parameter β for the Beta-BMA-K10 and Beta-BMA-K60 approach are seen in Figure 6.18a and 6.18b, respectively. The intercept parameter $\alpha + \alpha_t$ and slope parameter $\beta + \beta_t$ for the Beta-BMA-VCR-K10 approach are seen in Figure 6.18c and 6.18d, respectively for the optimal value $C = 0.06$. We observe that the intercept parameter varies around -1 and the slope parameter around 3, which means that the individual ensemble member forecasts has less impact on the mean parameter than for lead time 0. Hence, the climatology is more important.



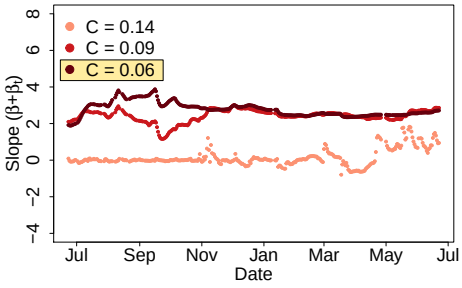
(a) Intercept α for Beta-BMA-K60 and Beta-BMA-K10.



(b) Slope $\beta$for Beta-BMA-K60 and Beta-BMA-K10.



(c) Intercept $\alpha + \alpha_t$ for Beta-BMA-VCR.



(d) Slope $\beta + \beta_t$ for Beta-BMA-VCR.

Figure 6.18: Parameter estimates plotted against the validation day $t = i + l$, for different Beta-BMA postprocessing approaches. The lead time is $l = 9$. The yellow box shows the optimal value for the smoothing parameter C.

The variance parameter ϕ^{-1} is plotted in Figure 6.19. For the VCR approaches the smoothing parameter C is set to be $C = 0.06$. As explained earlier, the interpretation is difficult. We observe that the $K10$ approaches vary much. We further observe that the estimates from the BMA-K60 and Beta-BMA-VCR-K60 approaches are close to each other because the smoothing parameter value C is small.

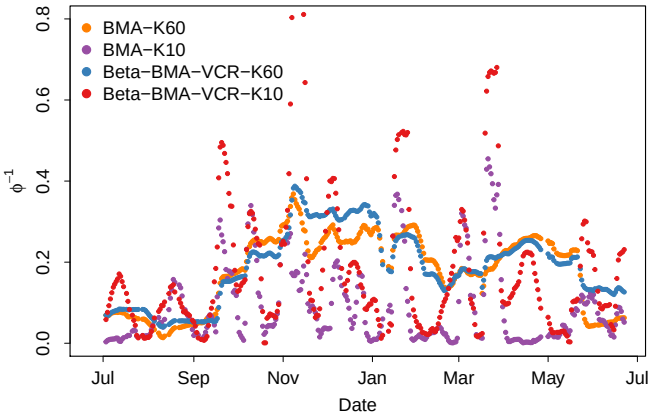


Figure 6.19: The variance parameter ϕ^{-1}, plotted against the validation day $t = i + l$, for four different Beta-BMA postprocessing approaches. The lead time is $l = 9$.

6.5 Comparison Between Approaches

The CRPS for each approach and for the raw ensemble are given in Table 6.1. For each lead time, the lowest CRPS is highlighted. We find that the BMA-VCR-M2-K10 approach has the best predictive performance for lead time 0, and Beta-BMA-VCR-K60 has the best predictive performance for lead time 4. For lead time 9, the Beta-BMA-VCR-K60 approach and the raw ensembles perform equally good, but the bias-corrected raw ensemble preforms slightly better. Furthermore, in the analysis we found that general trend is that for higher lead times, the less optimal smoothing parameter value. The optimal value seemed to range from $C = 0.11$ for lead time 0 to $C = 0.06$ for lead time 9.

CRPS	L0	L4	L9
BMA-K60	0.47	0.96	1.17
BMA-K10	0.44	1.13	1.35
BMA-VCR-M2-K60	0.43	0.88	1.08
BMA-VCR-M2-K10	**0.38**	0.89	1.06
Beta-BMA-K60	0.46	0.78	1.03
Beta-BMA-K10	0.49	1.02	1.29
Beta-BMA-VCR-K60	0.45	**0.76**	0.98
Beta-BMA-VCR-K10	0.42	0.80	1.05
Raw ensemble	0.57	0.89	0.98
Bias corr. raw ensemble	0.45	0.85	**0.96**

Table 6.1: CRPS for eight different BMA approaches and for the raw ensemble in the period from June 24, 2014 to June 22, 2015.

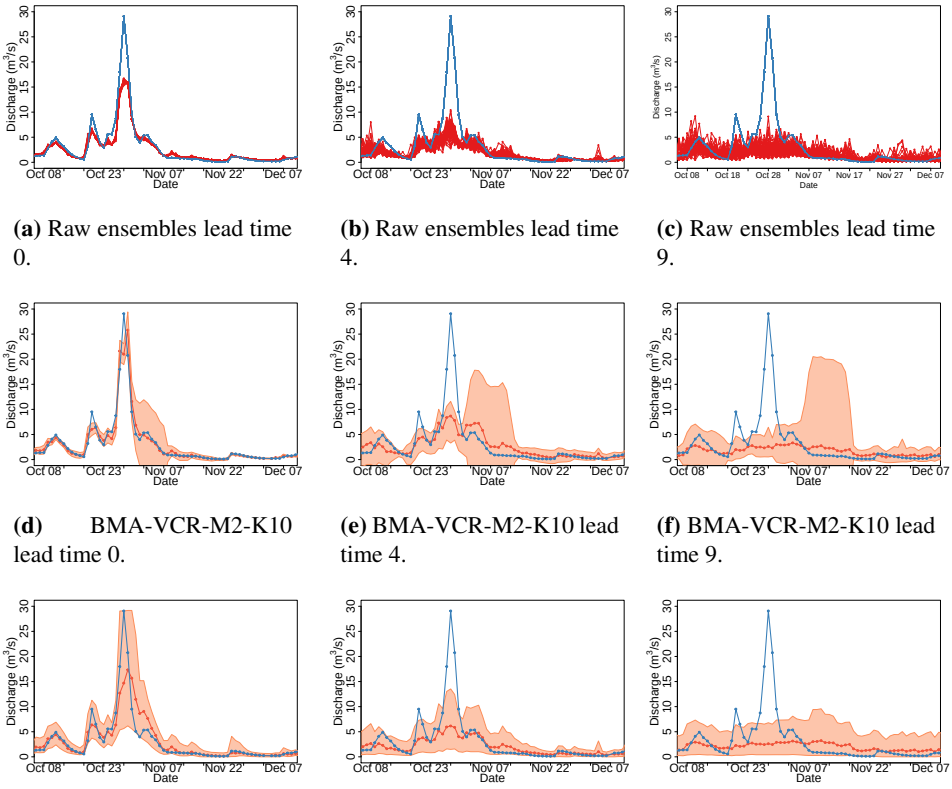
We compare the ensemble forecasts with the probabilistic forecast obtained from the two approaches having the lowest CRPS in Table 6.1, i.e. BMA-VCR-M2-K10 and Beta-BMA-VCR-K60.

The raw ensemble forecasts for lead time 0, 4 and 9 are shown in Figure 6.20a, 6.20b and 6.20c, respectively. The blue line in the figures are the corresponding observations, and the red lines are the ensemble forecasts. We observe that for lead time 0, the ensemble forecasts are following the observation curve closely, but the ensemble forecasts are not predicting the magnitude of the large streamflows late in late October. For lead time 4, the ensemble forecasts are only partly able to detect the late October streamflows, and for lead time 9, there are few signs indicating the large streamflow late in October.

The postprocessed forecasts for lead time 0, 4 and 9 for the BMA-VCR-M2-K10 are seen in Figure 6.20d, 6.20e and 6.20f, respectively. The blue line is the observation curve, and the red line is the mean of the probabilistic forecast. The red area around the red line is a 90% prediction interval. The width of the 90% prediction interval is a measure of sharpness of the postprocessed probabilistic forecast. The BMA-VCR-M2-K10 approach at lead time 0 provides good predictions with some errors at days at the days with large streamflow late in October. When the forecast is wrong we observe that the prediction interval for the next $K = 10$ days is wider than usual. For lead time 4, the forecast-observation error late in October is larger than for lead time 0, and since the days between the issue time and lead time are not included in the training period for the variance parameter, we observe that the variance of the probabilistic forecast does not increase until 4 days after the large forecast-observation error. For lead time 9, we observe that this late effect increases.

The postprocessed forecasts using the Beta-BMA-VCR-K60 for lead time 4 and 9 are seen in Figure 6.20g, 6.20h and 6.20i, respectively. For lead time 0, we observe that the mean of the probabilistic forecast is corresponding well with low streamflow values. We further observe that there is a large variance for the days with large streamflows. For lead time 4 we observe that the mean of the probabilistic forecast is following the trend of the corresponding observation curve, and as for lead time 0, we observe that when the mean of the probabilistic forecast is large, the uncertainty is large. For lead time 9, we observe that the ensemble forecasts are not detecting the large streamflow at October 28, resulting in a large forecast-observation errors around October 28. However, the probabilistic forecasts the following days are not affected the same way as for the BMA and BMA-VCR approaches.

The postprocessed probabilistic forecasts for the rest of the year, from June 24, 2014 to June 22, 2015, are provided in Appendix B.



(a) Raw ensembles lead time 0.

(b) Raw ensembles lead time 4.

(c) Raw ensembles lead time 9.

(d) BMA-VCR-M2-K10 lead time 0.

(e) BMA-VCR-M2-K10 lead time 4.

(f) BMA-VCR-M2-K10 lead time 9.

(g) Beta-BMA-VCR-K60 lead time 0.

(h) Beta-BMA-VCR-K60 lead time 4.

(i) Beta-BMA-VCR-K60 lead time 9.

Figure 6.20: Raw ensembles and postprocessed probabilistic forecasts for the period from October 6, 2014 to December 10, 2014. The blue line is the observation curve, and the red line is the mean of the probabilistic forecast. The red area around the red line is a 90 % prediction interval.

Chapter 7

Discussion and Conclusion

In this study, three new BMA postprocessing approaches have been developed. We have applied the BMA methodology combined with the CCPR methodology and VCR models to generate probabilistic forecasts for discharge. The approaches were tested in a case study for the Osali catchment for lead times 0, 4 and 9.

VCR models have been studied when modeling the mean parameter instead of a sliding window training period. The sliding window training period in the original BMA methodology makes the parameter estimates sensitive to large forecast-observation errors, which for the Osali catchment often happens during the fall due to heavy rainfall, and during the spring because of snow melting. The results showed that it is possible to obtain better calibrated and sharper probabilistic forecasts by imposing a dynamic linear structure for the bias-correction parameters in the BMA methodology.

To assess calibration of the probabilistic forecast we considered the PIT histogram. We found that the BMA-VCR approaches had close to uniform PIT histograms for lead time 0. By considering the PIT histograms for lead time 4 and 9 we found for the BMA-VCR approaches that the probabilistic forecast lack calibration. The Beta-VCR approaches were well calibrated for all lead times. Mean CRPS were used to assess the predictive performance. The smoothing parameter C was estimated based on predictive performance, and a value close to the optimal value for each approach were used for evaluation of the probabilistic forecast. For lead time 0 it was the BMA-VCR-M2-K10 approach with a relative high smoothing parameter $C = 0.09$ that gave the lowest CRPS. This means that a flexible model which adapts easily to recent weather patterns is favorable for lead time 0. For lead time 4 and 9, the Beta-BMA-VCR-K60 approach obtained the lowest CRPS among the BMA approaches, with smoothing parameter $C = 0.09$ and $C = 0.06$, respectively. This means that the Beta-BMA model with a long training period for the model parameters and a strict smoothing parameter, which is a model that are less influenced by local variations, provides better results for lead time 4 and 9.

Two mean models were suggested for the BMA-VCR approaches. The first model (M1) included the total bias between forecasts and observations based on all data available, as well as bias-correction parameters with a dynamic linear structure. This means that the smoothing parameter C determines to what extent the model should learn from much data, i.e. small C, or adapt locally with high flexibility for the bias-correction parameters. The second mean model (M2) only included the dynamic linear model structure parameters, meaning that the mean parameter varied around the individual ensemble member forecasts. It was seen that mean model M2 performed better than mean model M1 for all lead times. One reason is that model M2 recovers faster after huge forecast-observation errors compared to mean model M1.

The variance of the probabilistic forecast for lead time 4 and 9 using the normal distribution as the individual ensemble member pdf with a sliding window training period, is difficult to estimate since the days between the issue time and lead time are not included in the training period. Two lengths of the training period were tested. For the BMA-VCR models, it was only the variance parameter τ that was estimated from a sliding window training period. The results showed that a short training period provides better results for lead time 0, and a long training period is better for lead time 4 and lead time 9. A short training period makes the variance parameter more adaptable to changing weather patterns, or local adjustments, and for lead time 0, this is desirable. A long training period generally estimate the model parameters better because of more data, and since the days between the issue time and lead time is not included in the training period, local adjustments are less important. However, for lead time 9, the raw ensemble forecasts have a lower CRPS than the forecasts obtained from the BMA-VCR approaches. This is despite the fact that the BMA-VCR-M2 approach includes the raw ensemble forecasts as a special case, by letting $C = 0$. Hence, we need a better way to estimate the variance of the probabilistic forecast in order to improve calibration and sharpness. By only model the mean, i.e. by bias-correct the raw ensemble forecast according to mean model $\mu_{t,m} = \alpha_t + (1 + \beta_t)x_{i,m,l}$, we found that for a smoothing parameter value $C = 0.09$, the predictive performance for lead time 9 improved slightly compared to the raw ensemble, indicating that the ensemble forecasts have some benefits of local adjustments.

For the Beta-BMA approaches, the CCPR methodology is applied to each individual ensemble member and combine them with the BMA methodology. The postprocessed probabilistic forecast in the CCPR methodology is based on the climatology and deterministic forecasts, and provides a flexible representation of the individual ensemble member pdfs, because of the many forms of the beta distribution. However, the interpretation of the parameters is difficult because the pdfs are fitted on the unit scale. The beta distribution was parametrized by a mean parameter μ and a precision parameter ϕ, and the shape of the beta distribution depends on both parameters. In addition, the shape of the postprocessed probabilistic forecast on original scale is affected by the shape of the climatology.

The precision parameter ϕ was estimated from a sliding window training period. On the unit scale, the larger ϕ the sharper forecast. On original scale, however, the mean of the beta distribution μ and the shape of the climatology cdf have a large impact on the total variance. This makes the interpretation of ϕ difficult. If the mean μ is low, then the certainty of the discharge value to be forecasted increases since the climatology for the

Osali catchment mainly consists of low discharge values. On the other hand, if μ is large, the uncertainty is large. This effect is beneficial since high streamflows often are more difficult to predict than low streamflows.

For the Beta-BMA approaches, a small intercept parameter α combined with a large slope parameter β, gave a wide range for μ, which means that the deterministic forecast has a large impact on the mean parameter. If β is zero μ is constant and independent of the deterministic forecast. For the Beta-BMA-K60 approach for lead time 0, α was estimated to be close to -3 and β close to 5. For lead time 9, the estimates were around -1 and 2, respectively, indicating less influence of the deterministic forecast for increasing lead time. By including the VCR model, the probabilistic forecast was sharper and better calibrated for certain values of the smoothing parameter C.

There are several ways to further develop the BMA-VCR and Beta-BMA-VCR models. For the BMA-VCR model, a dynamic linear structure could also be imposed to the variance parameter, instead of being estimated from a sliding window training period with fixed size. We did not consider this case since the inference gets complicated and is not supported by the INLA framework. Furthermore, as seen in the case study, the variance of the probabilistic forecast often is too small in periods with high discharge forecasts. This could be solved by modeling the variance parameter τ as a function of the deterministic ensemble forecasts (Sloughter et al., 2007) or the ensemble spread. The precision parameter for the Beta-BMA-VCR model could also be modeled the same way. In addition, it is possible to construct the climatology pdf differently. In our case study, the domain of the climatology pdf only ranges from 0 to the largest observed value in the period from April 4, 2014 to January 31, 2016. The climatology pdf should have a larger domain, e.g. by letting the tail of the climatology decay exponentially (Borhaug, 2014). Further, if there are data from several years available, it could be interesting to have a more flexible representation of the climatology pdf by including seasonal variations.

Bibliography

Ajami, N. K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resources Research 43 (1).

Anderson, J. L., 1996. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. Journal of Climate 9 (7), 1518–1530.

Bergström, S., 1992. The HBV model: Its structure and applications. Swedish Meteorological and Hydrological Institute.

Blangiardo, M., Cameletti, M., 2015. Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons.

Borhaug, J., 2014. Climatology cumulative probability regression: A postprocessing methodology based on climatology and deterministic forecasts, with a case study of streamflow forecasts at osali.

Brent, R. P., 2013. Algorithms for minimization without derivatives. Courier Corporation.

Cleveland, W. S., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American statistical association 74 (368), 829–836.

Dawid, A. P., 1982. The well-calibrated bayesian. Journal of the American Statistical Association 77 (379), 605–610.

Dawid, A. P., 1984. Present position and potential developments: Some personal views: Statistical theory: The prequential approach. Journal of the Royal Statistical Society. Series A (General), 278–292.

Engeland, K., Steinsland, I., 2014. Probabilistic postprocessing models for flow forecasts for a system of catchments and several lead times. Water resources research 50 (1), 182–197.

Engeland, K., Steinsland, I., Johansen, S. S., Petersen-Øverleir, A., Kolberg, S., 2016. Effects of uncertainties in hydrological modelling. a case study of a mountainous catchment in southern norway. Journal of Hydrology 536, 147–160.

Ferrari, S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. Journal of Applied Statistics 31 (7), 799–815.

Fraley, C., Raftery, A. E., Gneiting, T., 2010. Calibrating multimodel forecast ensembles with exchangeable and missing members using bayesian model averaging. Monthly Weather Review 138 (1), 190–202.

Fraley, C., Raftery, A. E., Gneiting, T., Sloughter, J. M., 2007. Ensemblebma: An r package for probabilistic forecasting using ensembles and bayesian model averaging. Tech. rep., DTIC Document.

Gneiting, T., Balabdaoui, F., Raftery, A. E., 2007. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69 (2), 243–268.

Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102 (477), 359–378.

Gneiting, T., Raftery, A. E., Westveld III, A. H., Goldman, T., 2005. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. Monthly Weather Review 133 (5), 1098–1118.

Gneiting, T., Ranjan, R., et al., 2013. Combining predictive distributions. Electronic Journal of Statistics 7, 1747–1782.

Grimit, E. P., Gneiting, T., Berrocal, V., Johnson, N. A., 2006. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. Quarterly Journal of the Royal Meteorological Society 132 (621C), 2925–2942.

Hamill, T. M., 2001. Interpretation of rank histograms for verifying ensemble forecasts. Monthly Weather Review 129 (3), 550–560.

Hardin, J. W., Hilbe, J. M., Hilbe, J., 2007. Generalized linear models and extensions. Stata press.

Hastie, T., Tibshirani, R., 1993. Varying-coefficient models. Journal of the Royal Statistical Society. Series B (Methodological), 757–796.

Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting 15 (5), 559–570.

Krzysztofowicz, R., Kelly, K. S., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. Water Resources Research 36 (11), 3265–3277.

Nelder, J. A., Baker, R. J., 1972. Generalized linear models. Encyclopedia of statistical sciences.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL http://www.R-project.org

Raftery, A. E., Gneiting, T., Balabdaoui, F., Polakowski, M., 2005. Using bayesian model averaging to calibrate forecast ensembles. Monthly Weather Review 133 (5), 1155–1174.

Regonda, S. K., Seo, D.-J., Lawrence, B., Brown, J. D., Demargne, J., 2013. Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts–a hydrologic model output statistics (hmos) approach. Journal of Hydrology 497, 80–96.

Rings, J., Vrugt, J. A., Schoups, G., Huisman, J. A., Vereecken, H., 2012. Bayesian model averaging using particle filtering and gaussian mixture modeling: Theory, concepts, and simulation experiments. Water Resources Research 48 (5).

Rosenblatt, M., 1952. Remarks on a multivariate transformation. The annals of mathematical statistics 23 (3), 470–472.

Rue, H., Martino, S., Chopin, N., 2009. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. Journal of the royal statistical society: Series b (statistical methodology) 71 (2), 319–392.

Sloughter, J. M. L., Raftery, A. E., Gneiting, T., Fraley, C., 2007. Probabilistic quantitative precipitation forecasting using bayesian model averaging. Monthly Weather Review 135 (9), 3209–3220.

Vrugt, J. A., Robinson, B. A., 2007. Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and bayesian model averaging. Water Resources Research 43 (1).

Weerts, A., Winsemius, H., Verkade, J., 2011. Estimation of predictive hydrological uncertainty using quantile regression: examples from the national flood forecasting system (england and wales). Hydrology and Earth System Sciences, 15,(1).

West, M., 1996. Bayesian forecasting. Wiley Online Library.

Woods, A., 2005. Medium-range weather prediction: The European approach. Springer Science & Business Media.

Appendix A

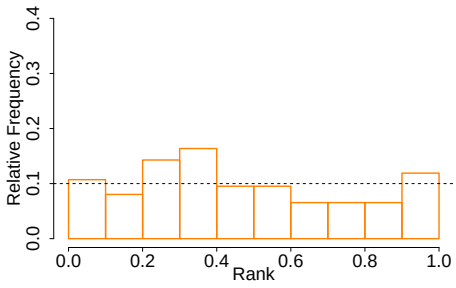
Results for Lead Time 4

Figure A.1a and A.1b show the PIT histograms for the BMA-K60 and BMA-K10 approaches using a sliding with a sliding window training period of 60 days and 10 days, respectively. Figure A.1c and A.1d show the histogram of pit values for the BMA-VCR-M2-K10 approach and BMA-VCR-M1-K10 approach, respectively, with smoothing parameter $C = 0.09$.

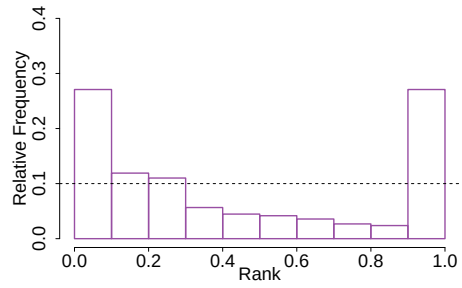
Figure A.2a and A.2b show the intercept parameter and slope parameter, respectively. The intercept and slope for the BMA-VCR-M1-K10 approach is seen in Figure A.2c and A.2d for different values of the precision parameter C. The yellow box indicates the optimal value for C or the BMA-VCR-M2-K10 approach from Figure 6.10. The parameters for the BMA-VCR-M2-K10 approach are seen in Figure A.2e and A.2f

Figure A.3a and A.3b show the PIT histograms for the Beta-BMA approach using the using a sliding with a sliding window training period of 60 days and 10 days, respectively. Figure A.3c and A.3d are the pit-histograms for the Beta-BMA-VCR-K60 approach and the Beta-BMA-VCR-K10 approach, respectively.

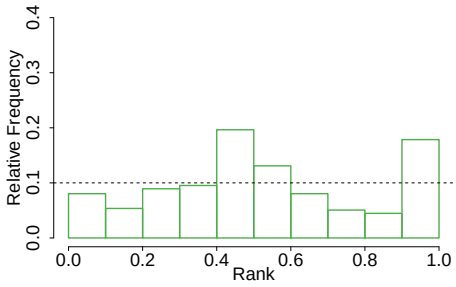
The intercept parameter α and slope parameter β are seen in Figure A.4a and A.4b, respectively, for the approaches Beta-BMA-K10 and Beta-BMA-K60. Figure A.4c and A.4d show the intercept parameter $\alpha + \alpha_t$ and slope parameter $\beta + \beta_t$ for the approach Beta-BMA-VCR-K10, respectively for the optimal value $C = 0.09$. We observe that the intercept parameter varies around -2 and the slope parameter around 4, which means that the individual ensemble member forecasts has less impact on the mean parameter than for lead time 0, but more than for lead time 9.



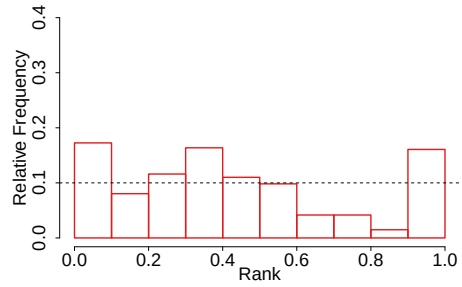
(a) BMA-K60.



(b) BMA-K10.

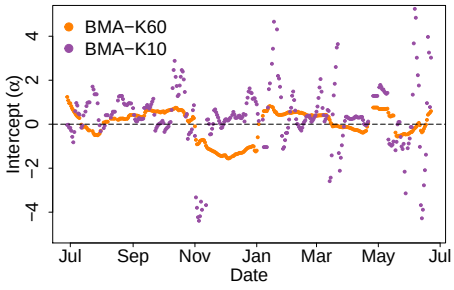


(c) BMA-VCR-M2-K10.

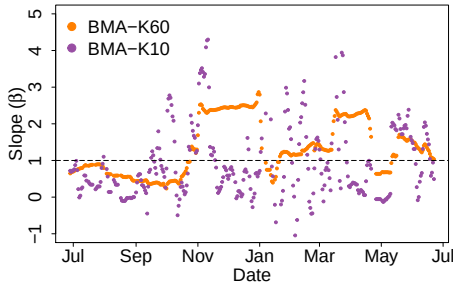


(d) BMA-VCR-M1-K10.

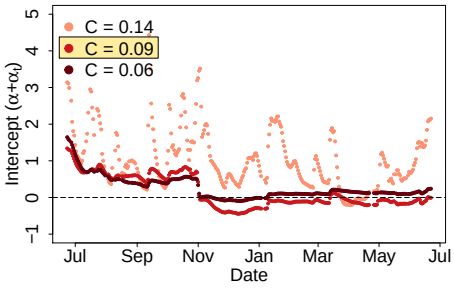
Figure A.1: PIT histograms for lead time 4



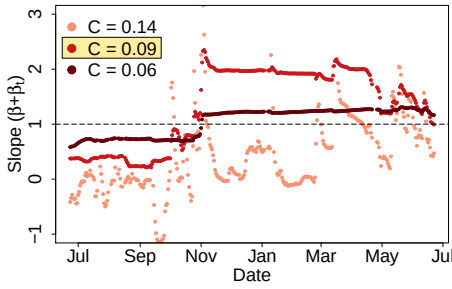
(a) Intercept α for BMA-K60 and BMA-K10.



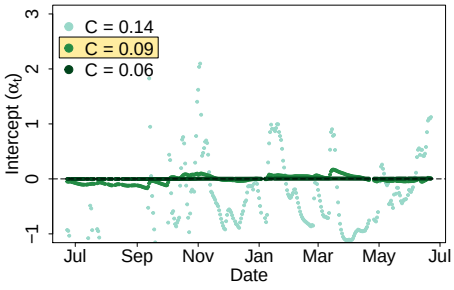
(b) Slope β for BMA-K60 and BMA-K10.



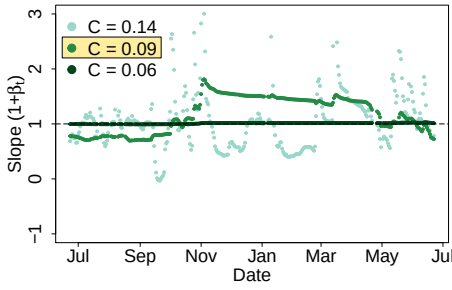
(c) Intercept $\alpha + \alpha_t$ for BMA-VCR-M1.



(d) Slope $\beta + \beta_t$ for BMA-VCR-M1.

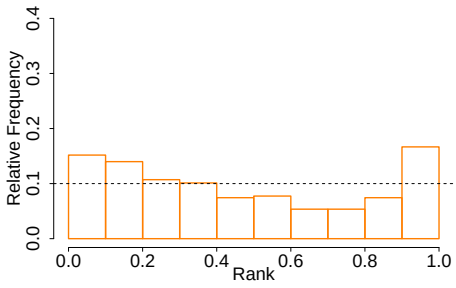


(e) Intercept α_t for BMA-VCR-M2.

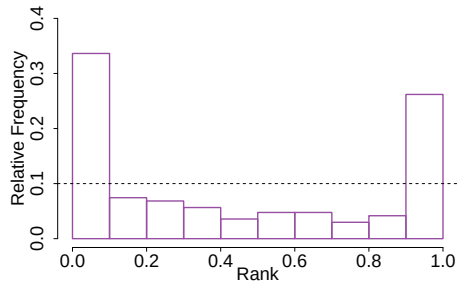


(f) Slope $1 + \beta_t$ for BMA-VCR-M2.

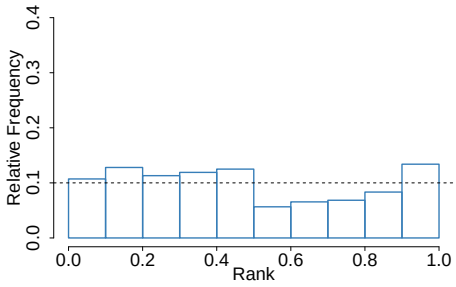
Figure A.2: Parameter estimates for lead time 4.



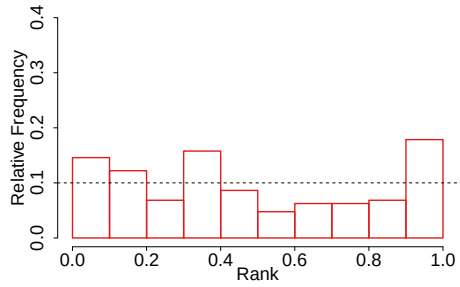
(a) Beta-BMA-K60.



(b) Beta-BMA-K10

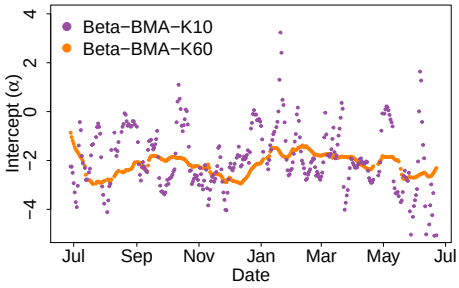


(c) Beta-BMA-VCR-K60.

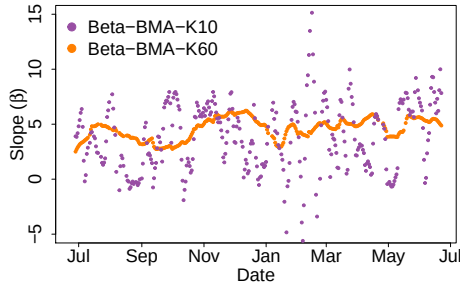


(d) Beta-BMA-VCR-K10.

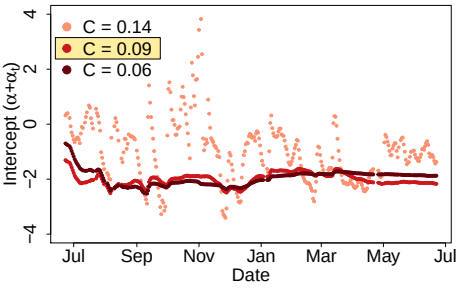
Figure A.3: PIT histograms for lead time 4.



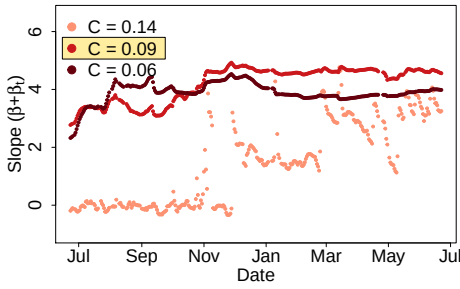
(a) First subfigure



(b) Intercept α for Beta-BMA-K60 and Beta-BMA-K10.



(c) Third subfigure



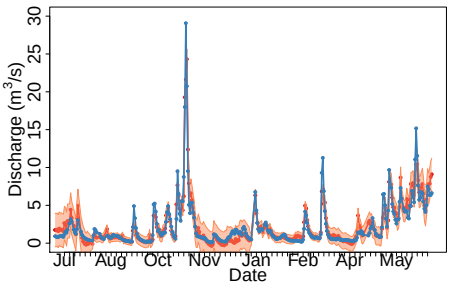
(d) Slope $\beta + \beta_t$ for Beta-BMA-VCR.

Figure A.4: Parameter estimates for lead time 4.

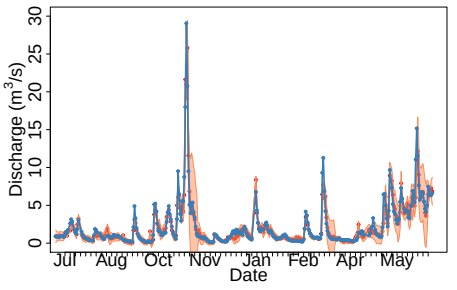
Appendix B

Postprocessed Probabilistic Forecasts

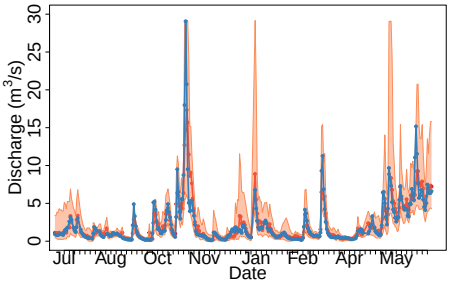
The postprocessed probabilistic forecast for the period from June 24, 2014 to June 22, 2015, for the BMA-K60, BMA-VCR-M2-K10, the Beta-BMA-VCR-K60 approaches for lead time 0, 4 and 9, are provided in the following figures. The red line is the mean of the postprocessed probabilistic forecast, and the red area around is a 90% prediction interval. The blue line is the corresponding observations.



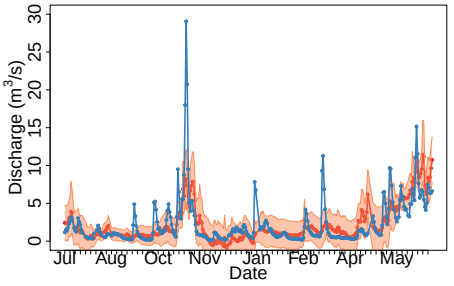
(a) BMA-K60 lead time 0



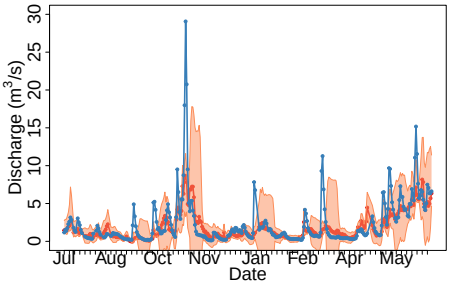
(b) BMA-VCR-M2-K10 lead time 0



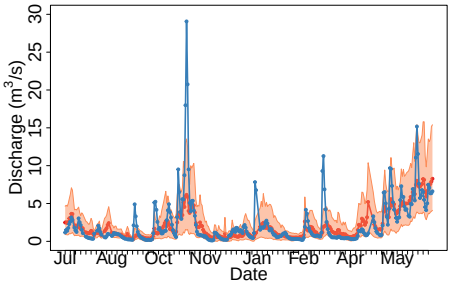
(a) Beta-BMA-VCR-K60 lead time 0



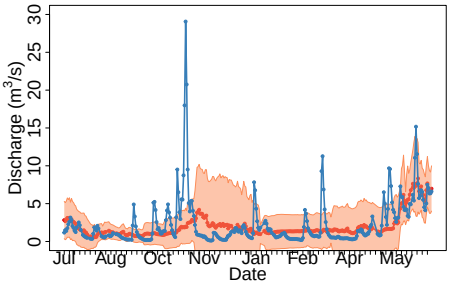
(b) BMA-K60 lead time 4



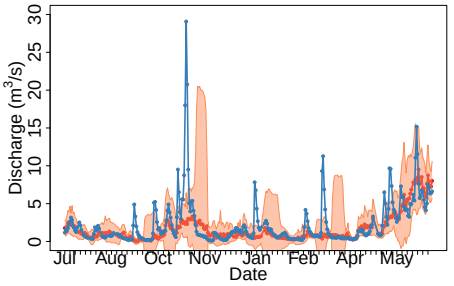
(a) BMA-VCR-M2-K10 lead time 4



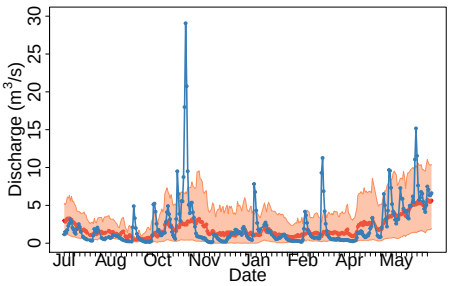
(b) Beta-BMA-VCR-K60 lead time 4



(a) BMA-K60 lead time 9



(b) BMA-VCR-M2-K10 lead time 9



(a) Beta-BMA-VCR-K60 lead time 9