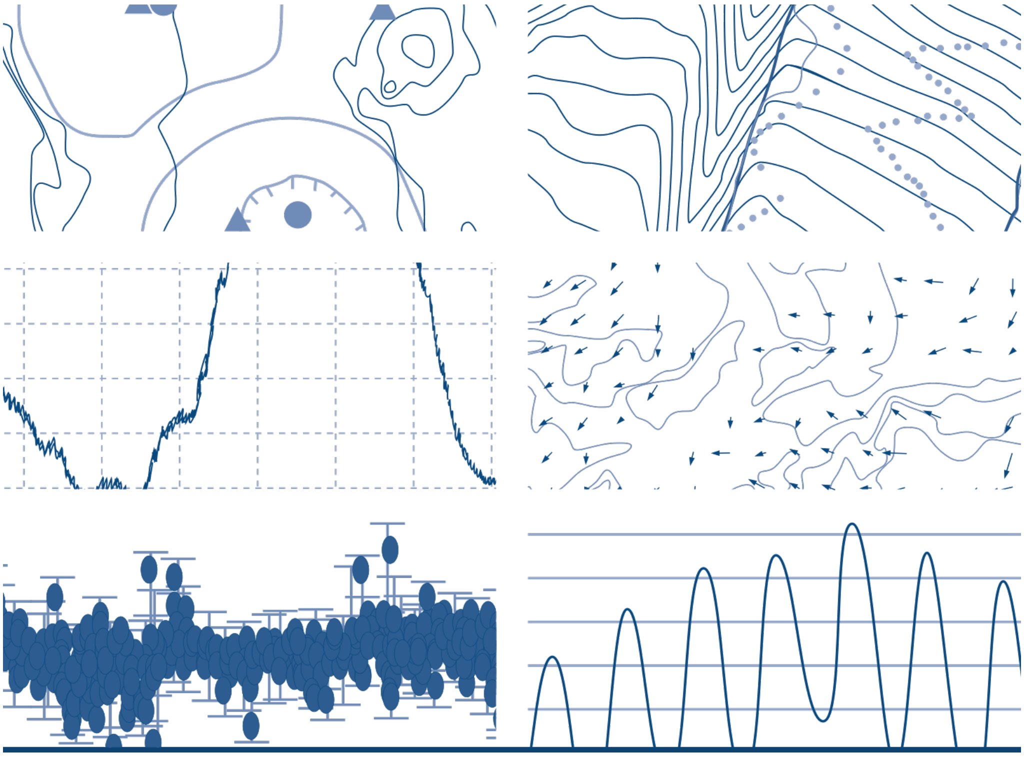
Icelandic
Meteorological Office

Veðurstofa
Íslands



# Probabilistic daily streamflow forecast using an analogue sorting method

Philippe Crochet

**Report**

VÍ 2013-008

Icelandic
Meteorological Office

Veðurstofa
Íslands

# Probabilistic daily streamflow forecast using an analogue sorting method

Philippe Crochet, Icelandic Met Office

Report

Veðurstofa Íslands
Bústaðavegur 7–9
150 Reykjavík

+354 522 60 00
vedur@vedur.is

VÍ 2013-008
ISSN 1670-8261

Veðurstofa Íslands

## Keypage

| Report no.: | Date: | ISSN: | Public ☒ Restricted ☐ |
|---|---|---|---|
| VÍ 2013-008 | November 2013 | 1670-8261 | Provision: |

| Report title / including subtitle | No. of copies: 20 |
|---|---|
| Probabilistic daily streamflow forecast using an analogue sorting method | Pages: 37 |
| | Managing director: Jórunn Harðardóttir |
| Author(s): Philippe Crochet | Project manager: Philippe Crochet |
| | Project number: 4812 |
| Project phase: | Case number: 2013-142 |

**Report contracted for:**
Orkurannsóknasjóður

**Prepared in cooperation with:**

**Summary:**
This report explores the possibility of forecasting daily streamflow using analogues. The principle is based on the search of past situations in a historical archive, most similar to the current situation according to some meteorological and hydrological parameters. Probabilistic and deterministic daily streamflow forecasts are obtained for the next 3 days by selecting observed streamflow in the 3-day period following each selected analogue situation, from an ensemble of analogues. Different algorithms are developed and tested off-line at a number of river catchments, considering various predictors and combination of predictors for the selection of analogues. Results indicate that the method is usually capable of producing reliable prediction intervals and performs better than persistence. The method capitalises on historical information collected on the catchment and can therefore be seen as an objective expert system based on past knowledge and experience.

| Keywords: | Managing director's signature: |
|---|---|
| Iceland, streamflow forecast, analogue, nearest neighbor | Jórunn Harðardóttir |
| | Project manager's signature: |
| | Reviewed by: Bergur Einarsson, SG |

# Contents

# 1 Introduction

The Icelandic Meteorological Office (IMO) is responsible for the hydrological monitoring and issuance of flood warnings. In Iceland, floods are primarily of three different origins (Snorrason et al., 2012): (i) meteorological floods induced by rain and which are often combined with melting of snow and ice, (ii) floods due to ice formation and release within river channels, and (iii) glacier outburst floods which originate in marginal lakes, formed in glacier dammed side valleys, or subglacial lakes, formed as a result of geothermal activity or volcanic eruptions. All types of floods can be problematic. They pose a threat to travellers and populations as well as to infrastructures such as roads and bridges, disrupting traffic. Floods can also be problematic for the hydropower industry with respect to reservoir management and operations and in various other water resources management applications.

Until now, flood warnings at IMO have been based on automatic warning made if water level or other indicating variables such as conductivity at selected flood warning hydrometric stations exceeds a given warning threshold. Before flood warning is issued, the data are manually analyzed and interpreted with other observations such as water temperature, tremor data from the seismic network and weather forecasts from numerical weather prediction (NWP) models, depending on the origin of expected flood. Floods of meteorological origin are the most frequent and can occur in all parts of the country. They are mainly observed in autumn and during wintertime and springtime. Improving the predictability of such floods by increasing the maximum lead time with sufficient accuracy is therefore needed in order to ensure timely warnings.

A classic approach for producing short- and medium-range streamflow forecasts consists of coupling a hydrologic model with a NWP model. The use of hydrologic models at IMO has until now been limited to the simulation of historical and future streamflow series in relation to climate change research projects, water resources management and hydropower planning (Jónsdóttir, 2008; Einarsson & Jónsson, 2010; Þórarinsdóttir, 2012). Work is in progress for the development and evaluation of an operational hydrological forecasting system based on hydrological modeling. Meanwhile, in an effort to rapidly develop the forecasting capabilities of the hydrological service of IMO, an empirical method to short- and medium-range probabilistic streamflow forecast is investigated, based on the use of analogues.

Analogue-based methods have been used with success in streamflow forecasting. They are often referred to as nearest-neighbor methods (Karlsson & Yakowitz, 1987; Galeati, 1990; Akbari et al., 2011). The first step of the technique is to compare a current situation to all past situations collected in an historical archive, according to a set of attributes describing the hydrological and meteorological situations of the catchment, to extract the dates of the $N$ nearest situations according to some analogy criteria and to form a forecast as the sample average of the succeeding discharges of the $N$ nearest neighbors, i.e. the discharge observed some hours or days after each selected past analogue situation.

Analogue-based methods have also been used in weather forecasting (Radinović, 1975; Kruizinga & Murphy, 1983; Van den Dool, 1989; Fraedrich et al., 2003) and climate downscaling (Zorita & Von Storch, 1999; Wetterhall et al., 2005) to extract local weather information which can not be simulated by coarse-resolution meteorological or climate models with sufficient accuracy. In that context, these methods often make use of synoptic predictors describing atmospheric circulation patterns such as mean sea level pressure or geopotential height, assuming that local variables are partly linked to them and partly linked to the local environment. The two main assumptions are that i) similar general atmospheric circulation patterns should lead to similar local effects and that ii) if two atmospheric states are very close initially with respect to certain fields of variables, they will remain close for some time in the future (Radinović, 1975). As in streamflow forecasting, the first step of the technique is to compare a current situation to all past situations collected in an historical archive, according to the selected meteorological parameters, to extract the dates of the closest matches and to form an ensemble forecast with the associated local weather observed on these dates or some hours or days later. In the past decade in particular, this technique has been applied to produce probabilistic quantitative precipitation forecasts (PQPFs) (Obled et al., 2002; Hamill & Whitaker, 2006; Gibergans-Báguena & Llasat, 2007; Diomede et al., 2008; Panziera et al., 2011; Marty et al., 2012) considering that the coarse spatial resolution of NWP models may be a limiting factor for predicting flood-triggering precipitation with reasonable accuracy over small and medium sized mountainous catchments. Analogue-based PQPFs can then be used as input to a hydrological model to produce an ensemble of discharge forecasts (Diomede et al., 2008).

Motivated by the relative success and simplicity of analogue-based methods, the possibility to use these techniques to forecast daily streamflow in Iceland is explored. The method proposed here builds on the different approaches described above and makes use of synoptic meteorological information over a large domain around Iceland and hydrological and meteorological information within the catchments of interest to forecast daily streamflow up to 3 days ahead. In Section 2, the principle of the analogue method is introduced. In Section 3 hydrological and meteorological data used in the analysis are presented. Section 4 describes the different strategies considered for implementing the method and Section 5 presents some results. Some concluding remarks are made in Section 6.

# 2 The analogue method

Let $X(t)$ be a state of a dynamical system at time $t$, known through the observation of $k$ variables $o_j(t)$ (with $j = 1, ..., k$). The general principle of the analogue-based method is to identify, in a historical archive, the most similar state $X(u)$ to $X(t)$, according to some analogy criteria. The future evolution of the current state $X(t)$ after a lead time $T$, $X(t+T)$, can be forecasted by using the evolution of analogous state $X(u)$, after the same lead time $T$, $X(u+T)$. In other words, it is assumed that if $X(u)$ is close to $X(t)$, then $X(u+T)$ should be close to $X(t+T)$ (see also Fraedrich & Rückert, 1998). The vector $X(t)$ defining the state of the system is often referred to as the "feature vector". It is selected to summarize a large data vector by one of smaller dimension which is hoped to contain most of the information relevant to the decision problem at hand (Karlsson & Yakowitz, 1987).

The dynamical system under study here is the hydrologic response of a watershed to atmospheric forcing and physiography, which is governed by complex processes taking place at different spatial and temporal scales. Rather than to attempt to explicitly describe these processes through hydrological modeling, the analogue method proposed here simply attempts to predict future streamflow on the basis of the joint observation of past streamflow and other hydro-meteorological information that can conceivably affect streamflow. The first step of the technique is to compare the current situation to all past situations, according to a set of meteorological and hydrological parameters. Then the dates of the closest matches are extracted and an ensemble forecast is formed with the succeeding discharge of each selected analogue situation, i.e. the discharge observed some hours or days later. The assumption is that given some initial basin conditions, the hydrologic responses associated with two similar meteorological situations should bear a resemblance to each other and remain close for some time in the future. Several practical and methodological aspects need to be considered prior to apply the analogue method:

- The selection of appropriate variables (predictors) defining the state of the system should be physically linked to the variable to be predicted (predictant). The historical archive in which analogue situations are searched for results from a trade-off between i) archive length ii) number of variables defining the system and iii) data homogeneity (Obled et al., 2002). The historical archives need to be as long as possible so that a large variety of situations can be found. This is especially important for events whose return period may be longer than the archive length. Rare situations never observed in the past may prove difficult to predict without some sorts of post-processing. The quality of the archive needs to be homogeneous through the entire period so that no bias is introduced when searching for analogues. Inhomogeneities could result from e.g. instrumental changes, network density and data processing techniques (Obled et al., 2002).

- The size of the analogy domain defining the state of the dynamical system needs to be adapted to the problem to be solved. As the domain size and number of predictors increase, it may prove difficult to find close analogues. According to Van den Dool (1989), simplifying assumptions needs to be made. In particular, if we are only concerned by a limited area, it may be sufficient to restrict the size of the analogy domain around that area to find good analogues.

- The number of analogues to select must result from a compromise between sampling quality and decreasing degree of analogy, which also depends on the analogy criteria (Obled et al., 2002). The larger the number of selected analogues the better the sampling but the lower the analogy with the current situation. In particular, rare events whose return periods are longer than the archive length may prove difficult to sample properly even if the sample size is increased and systematic biases may be expected.

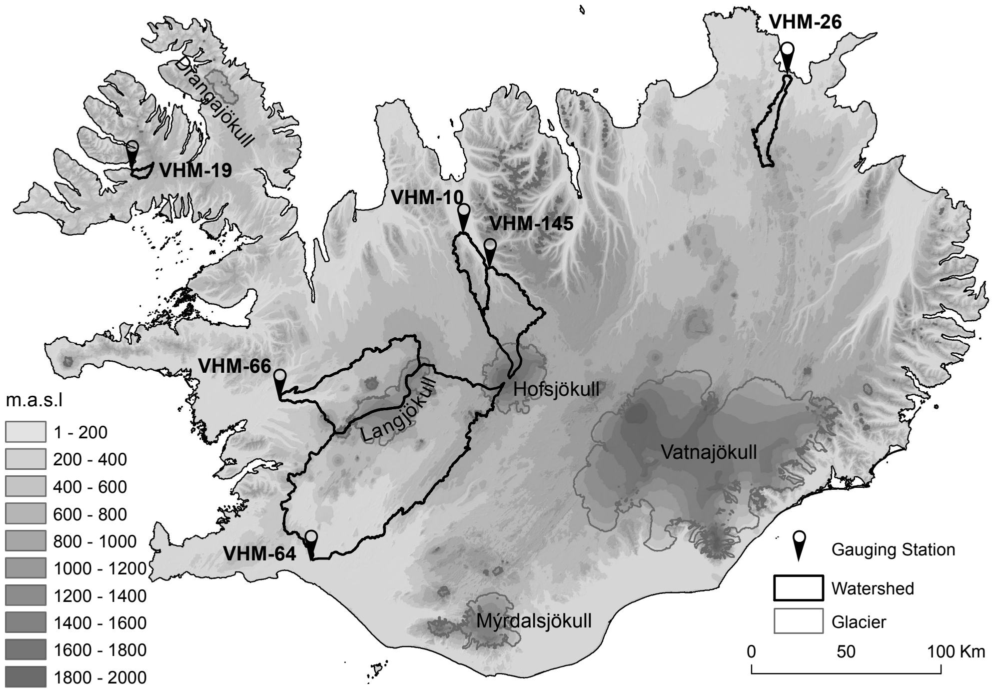



Figure 1. Topography of Iceland and location of watersheds considered here.

Table 1. Characteristics of the considered watersheds. Letter combinations indicate the type of river, with the first letter indicating the primary type. Direct runoff river (D), presence of lakes (S), glacier-fed river (J), groundwater (L).

| Gauging station | vhm 19 | vhm 10 | vhm 26 | vhm 145 | vhm 66 | vhm 64 |
|---|---|---|---|---|---|---|
| Name | Dynjandisá | Svartá | Sandá | Vestari-Jökulsá | Hvítá | Ölfusá |
| Type of river | D+L | D+L | D+L | D+J+L | L+J | L+D+J+S |
| Drainage area ($km^2$) | 42 | 397 | 267 | 850 | 1664 | 5687 |
| Mean altitude (m a.s.l.) | 555 | 535 | 391 | 753 | 664 | 480 |
| % Glacierized area | 0 | 0 | 0 | 11 | 20 | 12 |
| % Grassland | <1 | 43 | 25 | 15 | 10 | 25 |
| % Woodland | <1 | 0 | 0 | 0 | 2 | 2 |
| % Moss | 59 | 25 | 0 | 9 | 14 | 13 |
| % Little or no vegetation | 35 | 27 | 30 | 63 | 47 | 36 |
| % Wetland | 0 | 5 | 46 | <1 | 6 | 7 |
| % Lakes | 6 | <1 | <1 | <1 | <1 | 3 |

# 3 Data

Various meteorological and hydrological parameters were extracted and used to select analogues.

## 3.1 Hydrological data

A set of six river basins of various types and for which daily discharge ($Q$) was available for at least thirty years was selected (Figure 1). Table 1 presents information on the characteristics of each catchment. Rivers in Iceland are usually classified according to their source (Rist, 1990; Jónsdóttir et al., 2008), namely direct runoff (D), groundwater fed (L), glacial rivers (J) and whether they flow through lakes (S). The combination of letters indicates the origin of flow with the first letter indicating the primary origin. Three of these catchments are partly glaciated and two of them have a large groundwater contribution (vhm 64 and vhm 66). A more detailed description of the hydrology of these catchments can be found in Crochet (2013) where a study of their hydrologic response to recent climate variations is presented. In particular, it was observed that winter accumulation of snow cover and its melting in spring play a fundamental role in the hydrology of these watersheds as well as glacier melting in summer for the glacier-covered ones.

## 3.2 Meteorological data

Gridded daily temperature (Crochet & Jóhannesson, 2011) and precipitation (Crochet et al., 2007; Jóhannesson et al., 2007) series with 1-km resolution were used to derive basin-averaged daily temperature ($T$), rain ($R$), accumulated snow water equivalent ($SWE$), snowmelt ($S$) and glacier melt ($G$), following the same procedure as in Crochet (2013). Daily surface input runoff ($W$) was then defined as the sum of rain and snowmelt (and glacier melt). Mean sea level pressure ($MSLP$) fields were used to characterize atmospheric circulation patterns around Iceland. $MSLP$ was extracted twice daily from ERA-40 (Uppala et al., 2005) for the period 1958–2001 and available ECMWF analysis from 2002 to 2006, on a $1^\circ$ X $1^\circ$ latitude-longitude grid. As daily precipitation, temperature and derived variables were averaged from 00 to 00UTC, only the 12UTC $MSLP$ data were used so as to be centered on the same 24-h period.

# 4 Method implementation

Below, details on the method and metrics for evaluating forecast skills are given. The focus is on the probabilistic and deterministic forecast of daily streamflow for the six watersheds described above and for lead times of 1 to 3 days. The data were split into two periods. The validation period (01/09/2001–31/08/2005) for which daily streamflow was forecasted every day up to 3 days ahead and the archive period (01/09/1971–31/08/2001) used to search for analogues.

## 4.1 Predictors

Total streamflow hydrographs can be conceptualized as being composed of direct runoff, produced by rainfall and snowmelt, and baseflow. Snowmelt on a given day will depend on snow storage and temperature. Various methods were tested to identify analogues, in which the number of predictors and combination of predictors defining the feature vectors differed. The following methods are presented:

- Method 1: each day, the $N$ best analogues to current day are selected according to the shape of the *MSLP* fields at 12 UTC over the domain 60–70N, 35–5W (Figure 2). The analogy domain was chosen to be large enough to include areas with noticeable influence on the circulation patterns. The domain size was not strictly speaking optimized but several tests were performed. The idea here was to keep the method as simple and quick as possible, by focusing on the atmospheric circulation patterns only and their influence on precipitation and temperature and therefore runoff, through their associated wind regime. With this method, selected analogues are the same for all catchments.

- Method 2: each day, the $N$ best analogues to current day are selected according to the following catchment-scale predictors: $Q(t)$, $W(t)$, $SWE(t)$, $T(t)$. The idea here was to focus on hydro-meteorological processes taking place at the catchment scale and controlling runoff formation as well as prevailing basin conditions (wetness, snowpack, ...), defined here by $Q$ and $SWE$. Note that various tests were conducted to define initial streamflow conditions such as the average discharge up to current day, over periods of a few days, and the baseflow or mean baseflow estimated with the UKIH baseflow separation method (Piggott et al., 2005). The use of the last observed discharge was observed to give the best results.

- Method 3 combines methods 1 and 2, i.e. synoptic-scale and basin-scale predictors: first the $M$ best analogues to current day are selected with method 1 and then the $N$ best analogues according to method 2 selected among the $M$ best.

- Method 4 combines methods 2 and 1, i.e. basin-scale and synoptic-scale predictors: first the $N$ best analogues to current day are selected with method 2 and then re-ordered with method 1.

Situations characterized by similar atmospheric circulation patterns are expected to be similar with respect to temperature and precipitation but may be quite different in terms of initial catchment conditions, such as baseflow and snow storage, and therefore may lead to different streamflow scenarios in the following days. The use of catchment-scale information (method 3) is therefore a complementary discriminating feature when the synoptic situation alone (method 1) is not sufficient. In the same manner, situations characterized by similar local meteorological features and initial basin conditions may be different in term of atmospheric circulation patterns and therefore may lead to a different hydro-meteorological development and therefore different streamflow scenarios in the following days. The use of synoptic information (method 4) is therefore a complementary discriminating feature when local information alone (method 2) is not sufficient.

In order to take into account seasonal effects, a moving window of +/– 45 days centered on the target day was considered for the selection of analogues, so that in an archive made of $Y$ years, each target day was at most associated to $Y$ x 91 potential analogues. The window size was arbitrarily defined and not optimized. A similar window size was used by Hamill & Whitaker (2006) while Obled et al. (2002) used a 4 months window. By doing so, it is hoped that candidate situations will present similar characteristics in terms of solar energy, surface fluxes and other characteristics such as soil conditions which can partly be frozen in winter, affecting infiltration and therefore baseflow, surface runoff and thus streamflow.

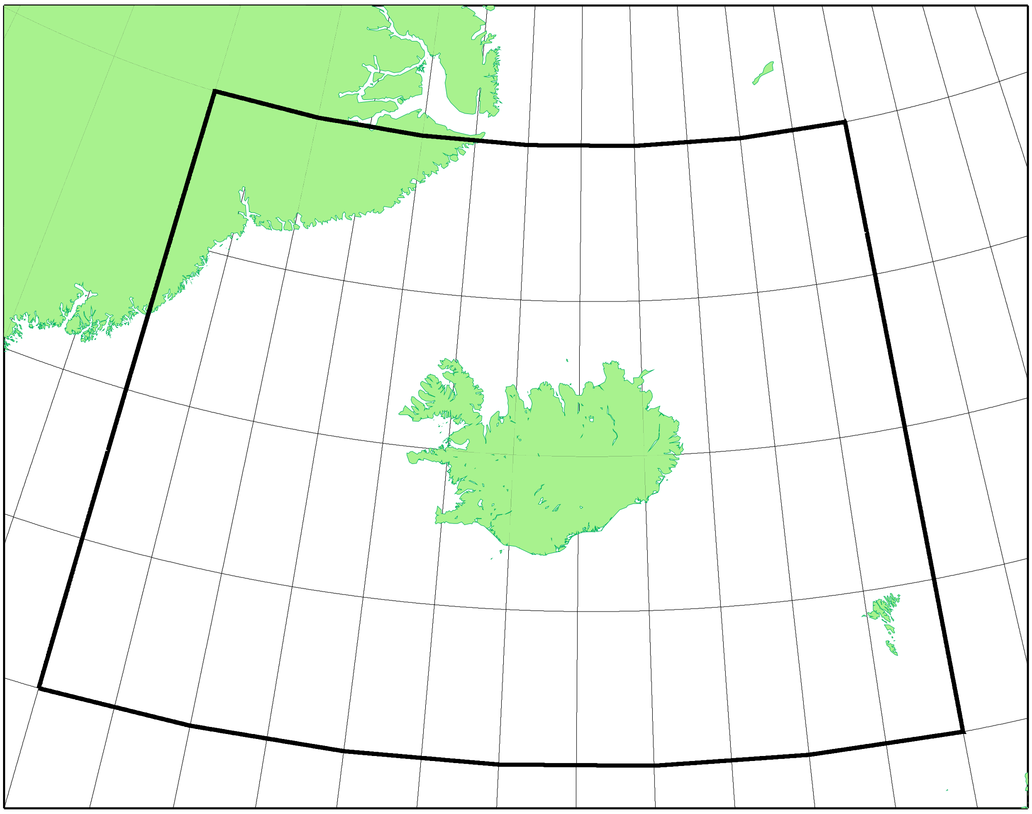


Figure 2. Analogy domain for the *MSLP* fields: 60–70N, 35–5W (thick solid line).

## 4.2 Analogy criteria

The selection of analogues was carried out by mean of three criteria. The similarity of atmospheric circulation patterns was evaluated with the Teweles-Wobus ($S1$) skill score (Wilks, 1995) applied to $MSLP$. This score compares the shape of two $MSLP$ fields by considering their gradient at each grid point of the domain under study:

$$S1(u) = \frac{\sum_{i=1}^{n}|E_i| + \sum_{j=1}^{n}|E_j|}{\sum_{i=1}^{n}G_i + \sum_{j=1}^{n}G_j} \tag{1}$$

with

$$E_l = (MSLP(l-1,u) - MSLP(l+1,u)) - (MSLP(l-1,t) - MSLP(l+1,t)) \tag{2}$$

and

$$G_l = max(|MSLP(l-1,u) - MSLP(l+1,u)|, |MSLP(l-1,t) - MSLP(l+1,t)|) \tag{3}$$

where $E_l$ is the difference around a given grid point $l$, between the $MSLP$ gradient at current time $t$ and time $u$ (i.e. candidate analogue), and $G_l$ is the maximum of these two gradients in the south–north direction ($l = i$) and west–east direction ($l = j$). The best analogue with respect to the shape of the $MSLP$ field at time $t$ is the $MSLP$ field at time $u$ that minimizes $S1$. This score has been used by e.g. Obled et al. (2002), Wetterhall et al. (2005) and Marty et al. (2012) with either $MSLP$ or geopotential height, to forecast precipitation. It was also applied to $MSLP$ fields by Woodcock (1980) for temperature forecasts. According to Woodcock (1980), the $S1$ skill score measures the similarity of the geostrophic winds between two situations and is an eminent suitable tool for selecting analogues for those weather elements largely controlled by wind regimes. Precipitation and temperature in Iceland are partly controlled by wind regimes and therefore surface runoff should be affected as well.

Both weighted Euclidian distance ($ED$) and Mahalanobis distance ($MD$) were tested to select analogue situations with respect to catchment-scale predictors (methods 2 to 4):

$$ED(u) = \sqrt{\sum_{j=1}^{k} \frac{(o_j(t) - o_j(u))^2}{s_j^2}} \tag{4}$$

where $s_j$ is the standard deviation of $jth$ variable $o_j$ over the sample set.

The Mahalanobis distance is an Euclidian distance which considers the correlation between the different predictors:

$$MD(u) = \sqrt{(X(t) - X(u))^T S^{-1} (X(t) - X(u))} \tag{5}$$

where $X(t)$ $(o_1(t),..,o_j(t),...,o_k(t))$ and $X(u)$ $(o_1(u),..,o_j(u),...,o_k(u))$ denote the multivariate feature vectors at times $t$ and $u$ and $S$ the covariance matrix.

The $N$ best analogues correspond to those $N$ days with the lowest $ED$ or $MD$ score. The $ED$ score has traditionally been used with nearest-neighbor methods (Karlsson & Yakowitz, 1987; Galeati, 1990; Akbari et al., 2011). Both $ED$ and $MD$ scores have been used by Panziera et al. (2011).

## 4.3 Ensemble prediction

For each method, the dates of the $N$ best analogues, i.e. the $N$ closest matches to current situation, at time $t$, were noted and an ensemble of $N$ daily streamflow forecasts, $Q(i,t+T)$ $(i=1,N)$, was obtained for a lead-time of $T$ days, by selecting the discharge observed $T$ days after the day corresponding to each selected analogue:

$$Q(i,t+T) = Q(u_i+T) \tag{6}$$

with $T$ =1, 2 or 3 and $u_i$ is the date of the $i^{th}$ analogue. A $100(1-p)\%$ confidence interval was constructed from this ensemble, from the quantiles corresponding to the non-exceedence probabilities $p/2$ and $1-p/2$: $[Q_{p/2};Q_{(1-p/2)}]$.

## 4.4 Rescaling

A limitation of the method is that it cannot forecast a value lower or larger than recorded in the archive. In order to allow the forecast to exceed these limits, three strategies were tested for rescaling the forecasted streamflow, according to observed streamflow conditions up to current time $t$:

$$Q(i,t+T) = \lambda_i Q(u_i+T) \tag{7}$$

$$Q(i,t+T) = \gamma_i Q(u_i+T) \tag{8}$$

$$Q(i,t+T) = \alpha_i Q_b(u_i+T) + \beta_i\Big(Q(u_i+T) - Q_b(u_i+T)\Big) \tag{9}$$

where $\lambda_i = \frac{Q(t)}{Q(u_i)}$, $\gamma_i = \frac{E[Q(t),Q(t-1),Q(t-2)]}{E[Q(u_i),Q(u_i-1),Q(u_i-2)]}$, $\alpha_i = \frac{Q_b(t)}{Q_b(u_i)}$, $\beta_i = \frac{Q(t)-Q_b(t)}{Q(u_i)-Q_b(u_i)}$,

and $Q_b(t)$ and $Q_b(u_i)$ are baseflows calculated using the UKIH baseflow separation method (Piggott et al., 2005). All rescaling coefficients were limited to a minimum value of 0.25 and a maximum value of 5.

## 4.5 Deterministic prediction

A deterministic forecast was derived from the ensemble by taking the mean of the $N$ analogue forecasts, weighted by the analogy criteria of the method in question:

$$Q(t+T) = \frac{\sum_{i=1}^{N} w_i Q(i,t+T)}{\sum_{i=1}^{N} w_i} \tag{10}$$

where $w_i$ is either $1/S1$, $1/ED$ or $1/MD$.

### 4.6 Forecast evaluation statistics

Several metrics were used to evaluate the forecast skills. Probabilistic forecasts were evaluated in terms of reliability and ensemble spread. A probabilistic forecast is reliable if the observed streamflow discharge lies in the empirical $100(1-p)\%$ prediction interval, $100(1-p)\%$ of the time. The ensemble spread was estimated by the width of the 90% prediction interval, i.e. the difference between the $Q_{95}$ and $Q_5$ quantiles: $Q_{95} - Q_5$. The performance of the deterministic forecast was measured by the mean error (ME) and root-mean squared error (RMSE). The benchmark deterministic prediction was defined as persistence, i.e. $Q(t+T) = Q(t)$, and the analogue-based deterministic forecasts evaluated against it.

## 5 Results

The number of selected analogues was not strictly speaking optimized. A fixed number ($M$=100) of analogues was selected for method 3. For all 4 methods, a comparison between $N$=25 and $N$=50 indicated that $N$=50 was giving better results with respect to both $RMSE$ and $ME$, but the difference was usually small, so only results with $N$=50 are presented. Results also indicated that using a weighted mean (Eq. 10) with either $N$=25 or $N$=50 was usually better than using the best analogue only ($N$=1). It was also observed that the use of the Mahalanobis distance was often giving slightly better results than the Euclidian distance but the improvement was usually not large. Finally, when rescaling was applied, the best results were obtained using the first strategy (Eq. 7), so only this strategy will be presented here.

Tables 2 and 3 present the summary statistics for each analogue technique discussed in Section 4 and the benchmark persistence method, calculated over the validation period (01/09/2001–31/08/2005). Table 2 corresponds to analogue forecasts obtained without rescaling (Eqs. 6 and 10, methods 1–4), selected with the $S1$ score and Euclidian distance. Table 3 corresponds to rescaled forecasts (Eqs. 7 and 10, methods 1–4), selected with the $S1$ score and Mahalanobis distance. Appendix 1 presents the reliability diagrams of the prediction intervals. Figures 3 to 6 illustrate these results and present the scatter plots of observed discharge against deterministic forecast for two river basins, vhm 10 and vhm 64, considering all methods and a forecast range of $T$=2 days. Appendix 2 presents the results for all catchments and all lead times, considering rescaled forecasts only. Finally, Figures 7 and 8 present the probabilistic and deterministic forecasts after rescaling (Eqs. 7 and 10), at vhm 10 and vhm 64, for a forecast range of 2 days and the period 01/09/2004-31/08/2005. Both weighted-mean and nearest analogue forecasts are presented for comparison, together with some statistics.

Results indicate that the probabilistic forecasts are usually reliable for all 4 methods and all 6 river basins, with and without rescaling the forecasts. This means that the observed streamflow discharge lies in the empirical $100(1-p)\%$ prediction interval $100(1-p)\%$ of the time, on average. A comparison between Tables 2 and 3 indicates that the spread of the ensemble forecast is reduced when rescaling is applied. This means that the probabilistic forecast becomes sharper after rescaling, without losing any reliability. In other words, rescaling leads to a better probabilistic forecast. As the forecast range increases, so does the spread of the prediction interval, indicating that the uncertainty increases with $T$. This is also reflected by $RMSE$ which increases with $T$.

Table 2 indicates that persistence is unbiased for all lead times and river basins. The deterministic analogue forecasts obtained with method 1 (Eqs. 6 and 10) perform very poorly and are worse than persistence. This result was somehow expected. As mentioned in Section 4.1, situations characterized by similar atmospheric circulation patterns only, are expected to be similar with respect to rainfall and temperature and therefore surface runoff, but may still be quite different in terms of initial catchment conditions, such as baseflow and snow storage. As a consequence, they may lead to a rather large spread of streamflow scenarios. These results indicate that atmospheric circulation patterns alone are not sufficiently discriminating to be used as streamflow predictors. When basin-scale information alone is used (method 2), results are usually better than persistence with respect to $RMSE$ but not $ME$. The forecasts have the tendency to be negatively biased, especially for the 3 glacial rivers where extreme values in particular are often systematically underestimated. The bias is not as pronounced for the non glacial rivers. Combining synoptic-scale and basin-scale predictors (method 3) leads to a substantial improvement compared to the use of *MSLP* fields alone (method 1). Analogue forecasts become similar or better than persistence, depending on catchment and lead time. Usually, persistence performs better for $T$=1 day and then method 3 becomes better for $T$=2 days and $T$=3 days. As method 2 already gives quite good results, combining basin-scale and synoptic-scale predictors (method 4) does only occasionally improve the forecasts slightly. For some catchments, on the contrary, the results are slightly worse than for method 2, depending on lead time. Overall, methods 2 and 4 are the best and quite equivalent in terms of skills.

When rescaling is applied (Eqs. 7 and 10, Table 3), the deterministic forecasts of all 4 methods improve, both with respect to bias ($ME$) and scatter ($RMSE$), especially method 1. The 4 methods outperform persistence with respect to $RMSE$, for all lead times and all river basins, except method 1 which is slightly worse than persistence for river basin vhm 10. The bias, although practically eliminated, is still slightly larger than for persistence, for all lead times and river basins. The results also indicate that when rescaling is applied, combining local and synoptic information (methods 3 and 4) gives better forecasts than using synoptic-scale predictors alone (method 1) and slightly better forecasts than using basin-scale predictors alone (method 2). Overall, methods 3 and 4 are the best and comparable in terms of skills but method 2 is close behind. It appears that method 4 gives better results than method 3 with respect to $RMSE$ and ensemble spread while methods 3 gives slightly less biased forecasts than method 4. Method 1 performs the poorest of all 4 methods but it is the fastest and simplest method to implement as it only requires discharge series and *MSLP* fields.

Table 2. Forecast skill evaluation over the period 01/09/2001–31/08/2005. No rescaling (Eqs. 6 and 10). All units are in $m^3/s$

| Forecast range | T=1 day | | | T=2 days | | | T=3 days | | |
|---|---|---|---|---|---|---|---|---|---|
| Statistics | ME | RMSE | Q95–Q5 | ME | RMSE | Q95–Q5 | ME | RMSE | Q95–Q5 |
| vhm 64 – method 1 | –20 | 103 | 302 | –19 | 99 | 297 | –20 | 103 | 298 |
| vhm 64 – method 2 | –11 | 44 | 132 | –15 | 64 | 162 | –18 | 87 | 201 |
| vhm 64 – method 3 | –24 | 64 | 206 | –24 | 72 | 221 | –25 | 87 | 241 |
| vhm 64 – method 4 | –16 | 47 | 132 | –19 | 65 | 162 | –22 | 85 | 201 |
| vhm 64 – persistence | 0.027 | 53 | / | 0.02 | 86 | / | 0.04 | 103 | / |
| vhm 66 – method 1 | –10 | 21 | 52.2 | –10 | 22 | 52.8 | –11 | 24 | 52.6 |
| vhm 66 – method 2 | –4.5 | 14 | 31.3 | –5.6 | 20 | 41.1 | –6.2 | 22 | 44.8 |
| vhm 66 – method 3 | –8.2 | 16 | 39.7 | –8.6 | 20 | 44.8 | –9 | 22 | 46.6 |
| vhm 66 – method 4 | –5.6 | 15 | 31.4 | –6.6 | 20 | 41 | –7.2 | 22 | 44.7 |
| vhm 66 – persistence | 0 | 16 | / | 0 | 21 | / | 0 | 24 | / |
| vhm 145 – method 1 | –2.6 | 13 | 30 | –2.5 | 13 | 29.8 | –2.7 | 13 | 29.2 |
| vhm 145 – method 2 | –1.3 | 7.1 | 11.5 | –1.5 | 9.4 | 14.3 | –1.8 | 10 | 16.7 |
| vhm 145 – method 3 | –2.8 | 8.8 | 17.8 | –2.7 | 10 | 19.5 | –2.8 | 11 | 20.5 |
| vhm 145 – method 4 | –1.6 | 7.4 | 11.5 | –1.7 | 9.4 | 14.3 | –2 | 10 | 16.6 |
| vhm 145 – persistence | 0 | 8 | / | 0.015 | 11 | / | 0.02 | 13 | / |
| vhm 10 – method 1 | 1.2 | 4.9 | 15.9 | 1.2 | 5.1 | 16.1 | 1 | 5.3 | 15.8 |
| vhm 10 – method 2 | –0.1 | 3 | 6.2 | –0.1 | 4.2 | 8.28 | –0.1 | 4.6 | 9.4 |
| vhm 10 – method 3 | –0.24 | 3.2 | 9.5 | –0.18 | 4.1 | 10.7 | –0.21 | 4.6 | 11.2 |
| vhm 10 – method 4 | –0.17 | 3 | 6.2 | –0.2 | 4.1 | 8.25 | –0.19 | 4.5 | 9.4 |
| vhm 10 – persistence | 0 | 3.5 | / | 0 | 5 | / | 0 | 5.6 | / |
| vhm 26 – method 1 | 1.3 | 7.4 | 19.8 | 1.3 | 7.4 | 19.9 | 1.3 | 7.5 | 19.8 |
| vhm 26 – method 2 | 0.03 | 3.6 | 7.1 | 0.05 | 5.2 | 9.6 | 0.02 | 5.9 | 10.8 |
| vhm 26 – method 3 | –0.37 | 4.7 | 11.7 | –0.28 | 5.5 | 12.7 | –0.31 | 6 | 13.2 |
| vhm 26 – method 4 | –0.07 | 3.8 | 7.1 | 0.033 | 5.4 | 9.6 | –0.04 | 5.9 | 10.8 |
| vhm 26 – persistence | 0 | 4.2 | / | 0 | 6.1 | / | 0 | 7 | / |
| vhm 19 – method 1 | –0.08 | 2.4 | 5.5 | –0.07 | 2.4 | 5.4 | –0.1 | 2.5 | 5.3 |
| vhm 19 – method 2 | –0.11 | 1.4 | 2.5 | –0.12 | 1.9 | 3.2 | –0.15 | 2.2 | 3.8 |
| vhm 19 – method 3 | –0.31 | 1.7 | 3.6 | –0.29 | 2 | 3.9 | –0.3 | 2.3 | 4.2 |
| vhm 19 – method 4 | –0.16 | 1.4 | 2.5 | –0.18 | 1.9 | 3.2 | –0.21 | 2.2 | 3.8 |
| vhm 19 – persistence | 0 | 1.4 | / | 0 | 2.1 | / | 0 | 2.4 | / |

Table 3. Forecast skill evaluation over the period 01/09/2001–31/08/2005. With rescaling (Eqs. 7 and 10). All units are in $m^3/s$

| Forecast range | T=1 day | | | T=2 days | | | T=3 days | | |
|---|---|---|---|---|---|---|---|---|---|
| Statistics | ME | RMSE | Q95–Q5 | ME | RMSE | Q95–Q5 | ME | RMSE | Q95–Q5 |
| vhm 64 – method 1 | 6.8 | 48 | 130 | 13 | 82 | 215 | 14 | 101 | 269 |
| vhm 64 – method 2 | –2 | 35 | 97 | –5.8 | 59 | 154 | –9.4 | 84 | 204 |
| vhm 64 – method 3 | 0.44 | 36 | 107 | 1.2 | 59 | 181 | 0.12 | 82 | 233 |
| vhm 64 – method 4 | –3.6 | 35 | 97 | –6.7 | 57 | 154 | –10 | 79 | 204 |
| vhm 64 – persistence | 0.03 | 53 | / | 0.02 | 86 | / | 0.05 | 103 | / |
| vhm 66 – method 1 | 1.4 | 15 | 33 | 1.9 | 21 | 48 | 1.5 | 23 | 53 |
| vhm 66 – method 2 | –2 | 13 | 29 | –3.1 | 19 | 42 | –3.5 | 21 | 47 |
| vhm 66 – method 3 | –0.46 | 13 | 30 | –0.7 | 19 | 45 | –0.98 | 21 | 50 |
| vhm 66 – method 4 | –2 | 12 | 29 | –3 | 18 | 42 | –3.4 | 21 | 47 |
| vhm 66 – persistence | 0 | 16 | / | 0 | 21 | / | 0 | 24 | / |
| vhm 145 – method 1 | 0.47 | 6.9 | 11.1 | 1 | 10 | 17.9 | 1.2 | 12 | 22.4 |
| vhm 145 – method 2 | –0.37 | 6.8 | 8.9 | –0.57 | 9.6 | 13.9 | –0.8 | 11 | 17.2 |
| vhm 145 – method 3 | –0.03 | 6.6 | 9.2 | 0.24 | 9.9 | 15.1 | 0.23 | 11 | 19.3 |
| vhm 145 – method 4 | –0.38 | 6.9 | 8.9 | –0.46 | 9.7 | 13.8 | -0.7 | 11 | 17.2 |
| vhm 145 – persistence | 0 | 8 | / | 0.015 | 11 | / | 0.02 | 13 | / |
| vhm 10 – method 1 | 0.59 | 3.5 | 7.9 | 0.91 | 5.2 | 12 | 0.94 | 5.7 | 13.6 |
| vhm 10 – method 2 | 0.07 | 2.9 | 5.2 | 0.08 | 4.2 | 8.1 | 0.1 | 4.7 | 9.4 |
| vhm 10 – method 3 | 0.29 | 3 | 6.4 | 0.46 | 4.4 | 10 | 0.46 | 4.9 | 11.4 |
| vhm 10 – method 4 | 0.05 | 2.8 | 5.2 | 0.08 | 4.2 | 8.1 | 0.1 | 4.6 | 9.4 |
| vhm 10 – persistence | 0 | 3.5 | / | 0 | 5 | / | 0 | 5.6 | / |
| vhm 26 – method 1 | 0.29 | 4 | 6 | 0.56 | 5.9 | 9.4 | 0.71 | 6.9 | 11.3 |
| vhm 26 – method 2 | 0.15 | 3.4 | 5.2 | 0.19 | 5.3 | 8.7 | 0.21 | 6 | 10.3 |
| vhm 26 – method 3 | 0.25 | 3.7 | 5.1 | 0.43 | 5.4 | 8.4 | 0.46 | 6.2 | 10.2 |
| vhm 26 – method 4 | 0.1 | 3.4 | 5.2 | 0.23 | 5.3 | 8.7 | 0.23 | 6 | 10.3 |
| vhm 26 – persistence | 0 | 4.2 | / | 0 | 6.1 | / | 0 | 7 | / |
| vhm 19 – method 1 | 0.08 | 1.3 | 2.9 | 0.28 | 2 | 4.7 | 0.35 | 2.3 | 5.6 |
| vhm 19 – method 2 | –0.02 | 1.2 | 2 | –0.025 | 1.8 | 3.3 | –0.05 | 2.2 | 4.2 |
| vhm 19 – method 3 | 0.04 | 1.2 | 2.3 | 0.12 | 1.8 | 3.9 | 0.14 | 2.2 | 4.8 |
| vhm 19 – method 4 | –0.02 | 1.2 | 2 | –0.027 | 1.8 | 3.3 | –0.066 | 2.1 | 4.1 |
| vhm 19 – persistence | 0 | 1.4 | / | 0 | 2.1 | / | 0 | 2.4 | / |

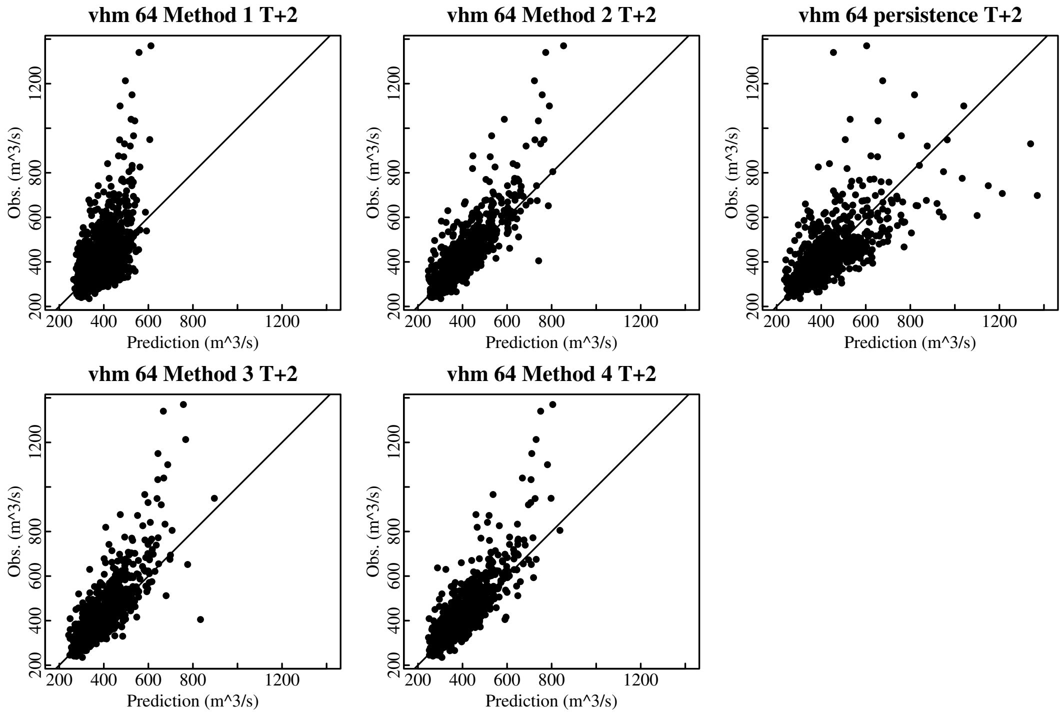



Figure 3. Observed versus predicted daily discharge at vhm 64, for a forecast range (T) of 2 days, over the period 01/09/2001–31/08/2005. The solid line corresponds to a perfect match. No rescaling is applied.

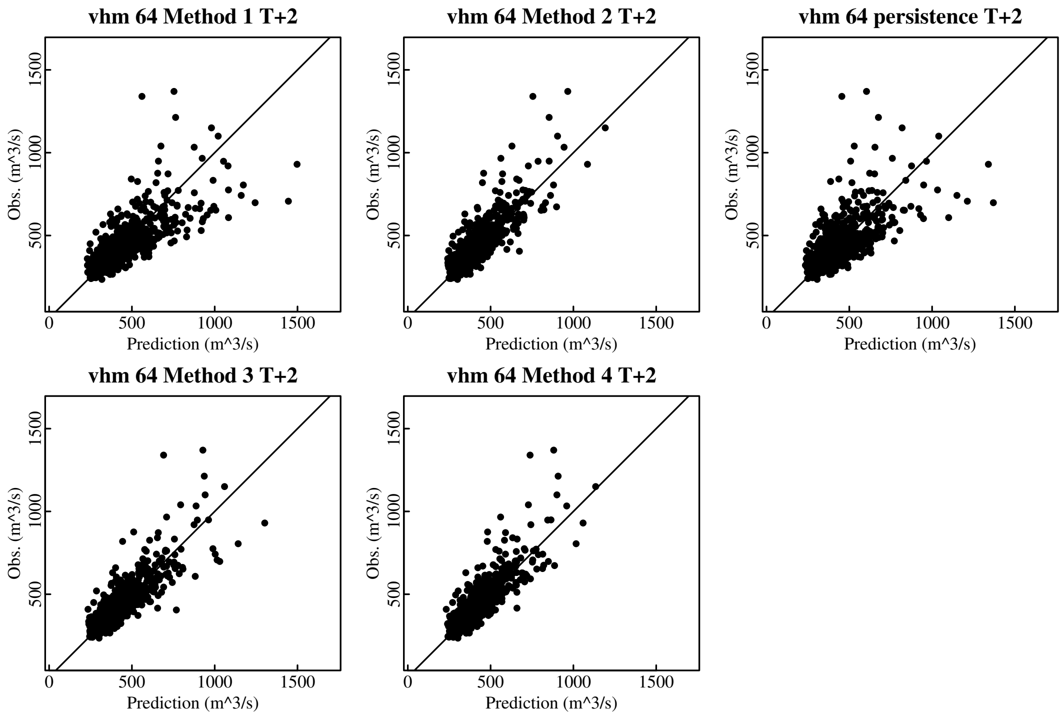



Figure 4. As Fig. 3 but for methods 1 to 4, rescaling is applied (Eq. 7).

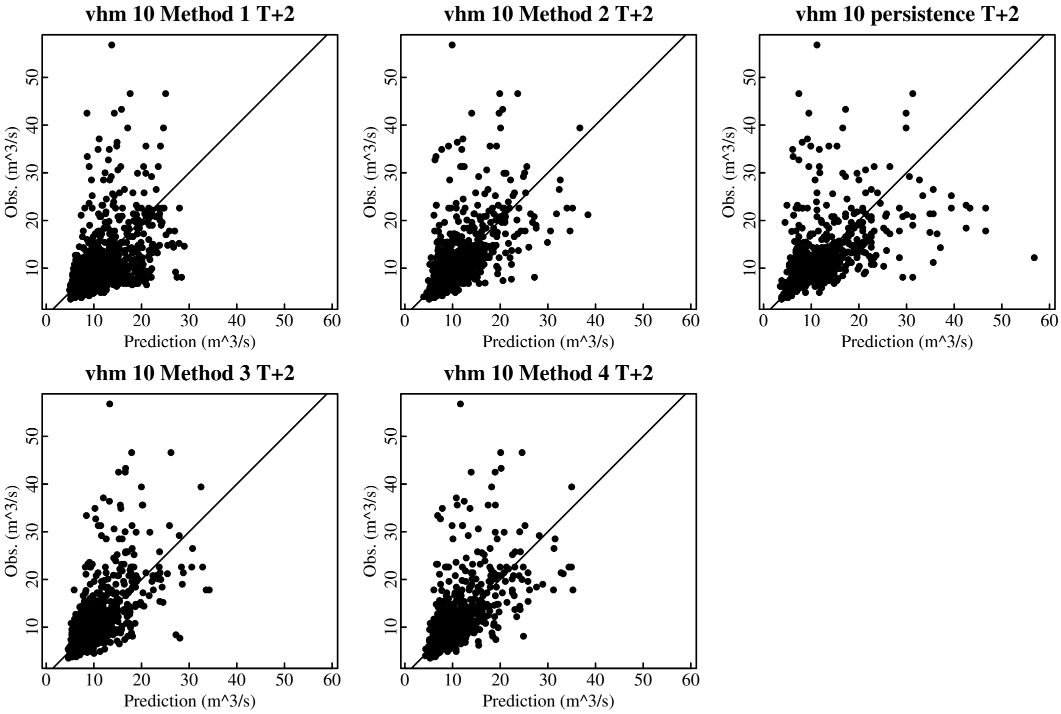



Figure 5. As Fig. 3 but for vhm 10.

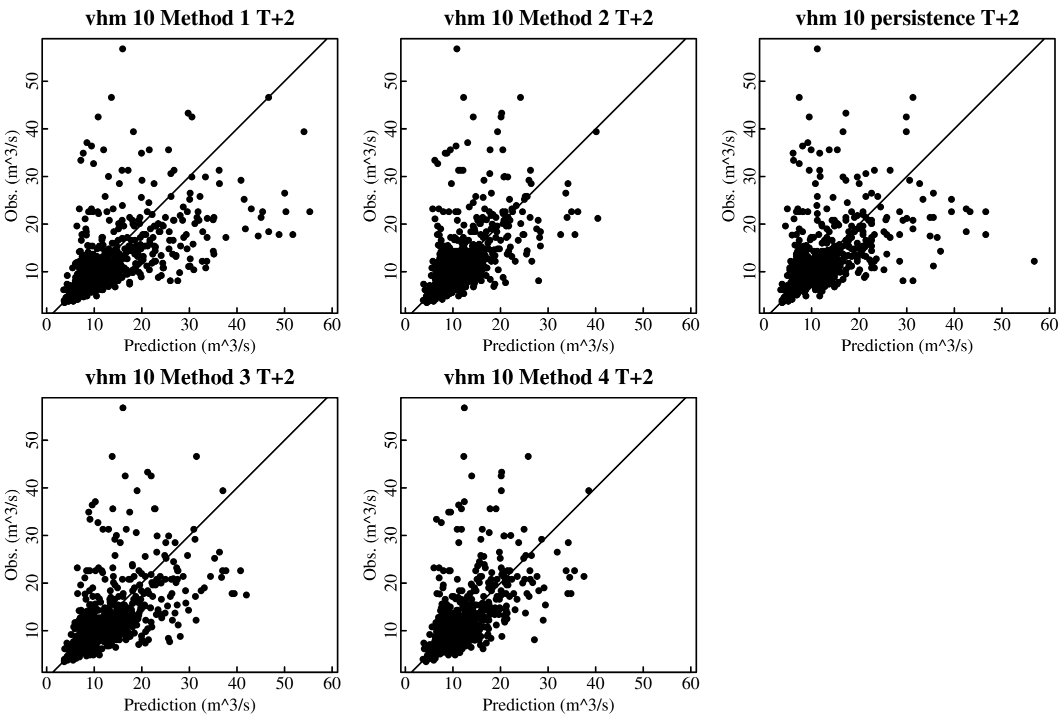



Figure 6. As Fig. 4 but for vhm 10.

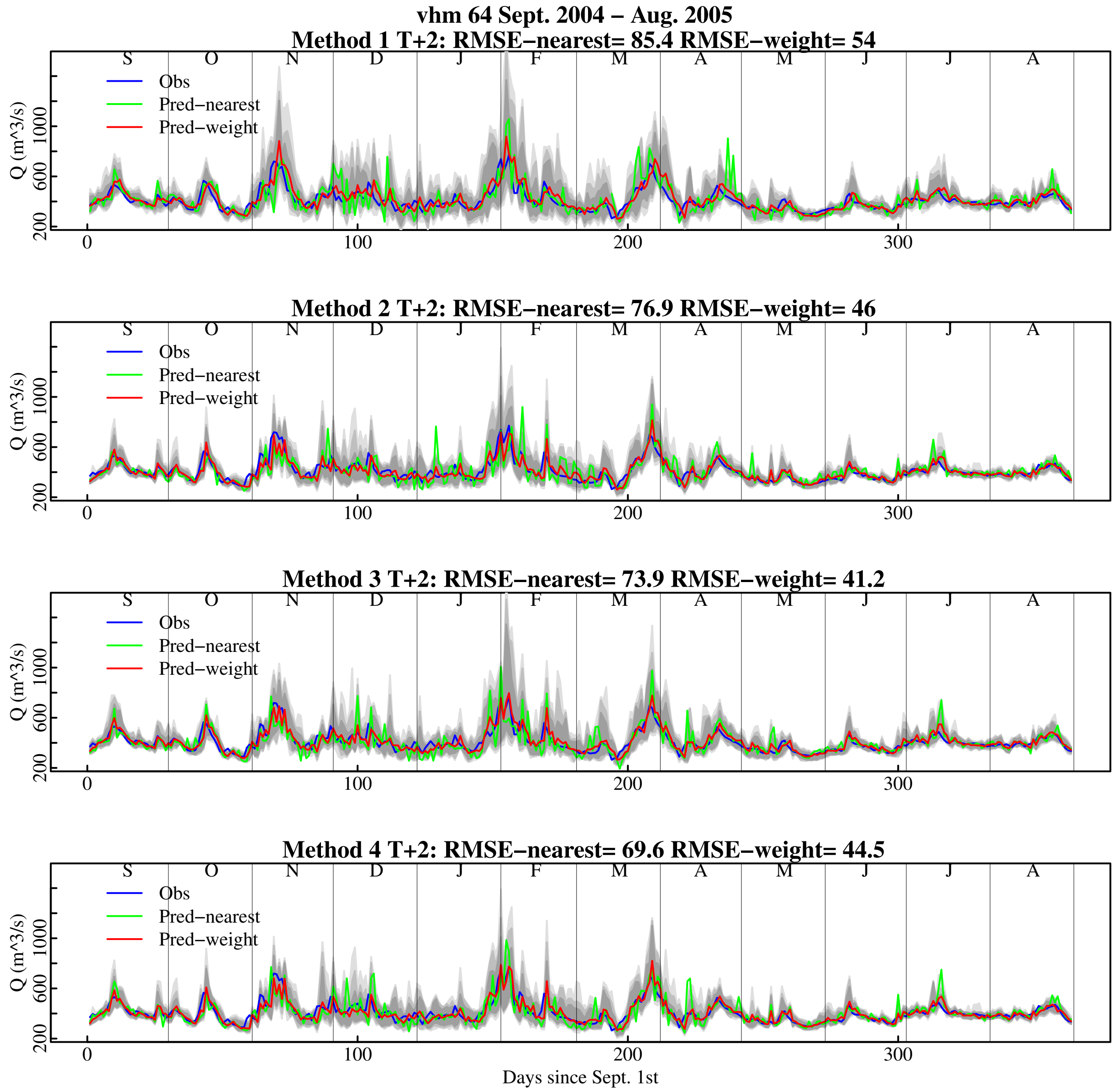



Figure 7. Observed and predicted daily discharges at vhm 64 for a forecast range (T) of 2 days and water-year 2004–2005, using methods 1 to 4 with rescaling. The 80%, 90% and 95% prediction intervals are represented by grey shades of decreasing densities. The observed discharge is in blue, the nearest analogue forecast in green and the weighted-mean of the 50 best analogues (Eq. 10) in red.

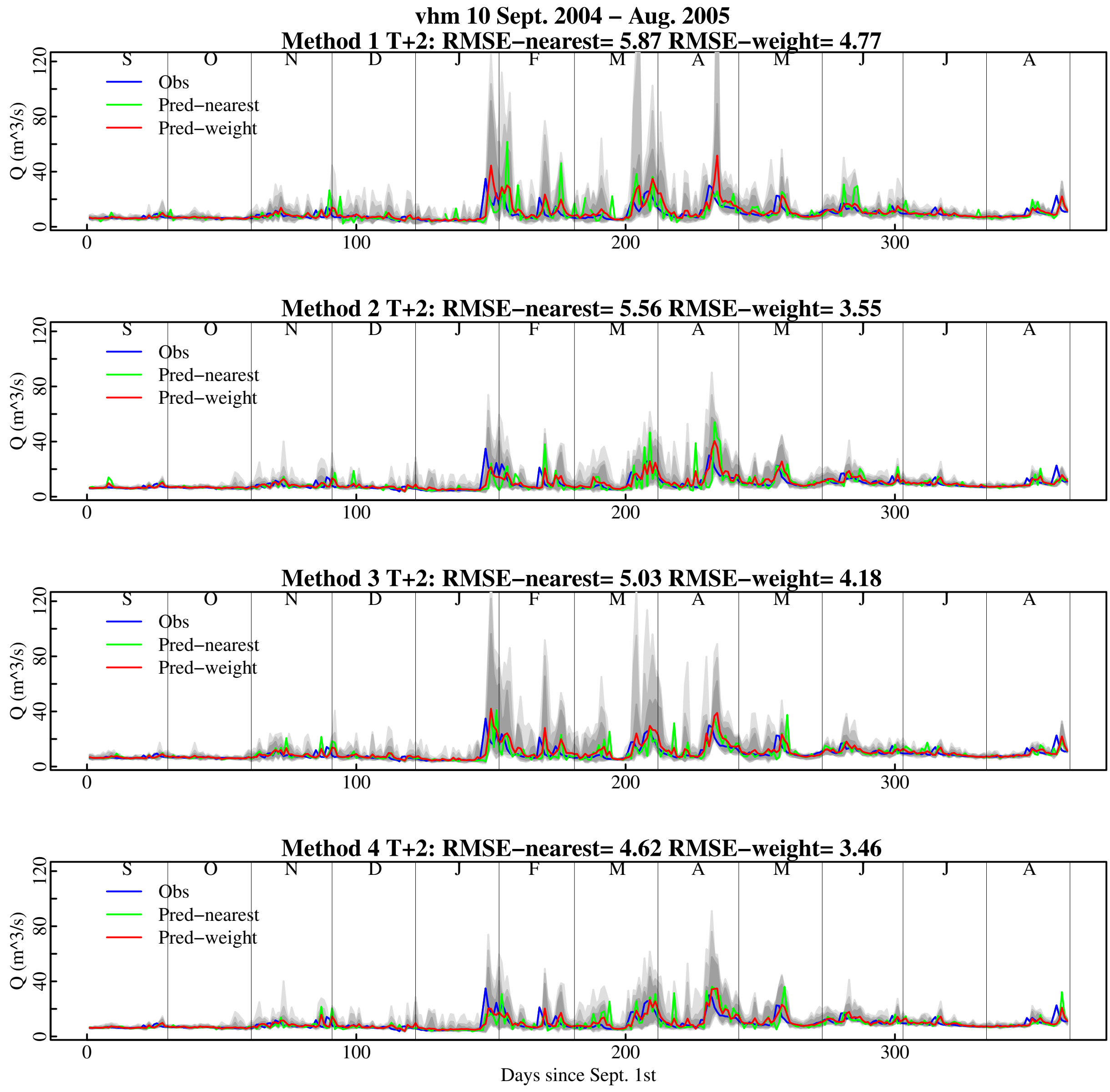



Figure 8. As Fig. 7 but for vhm 10.

# 6 Summary

A probabilistic streamflow forecasting system based on analogues was proposed and tested on 6 river basins of various types. The principle of the method is based on the search of past days in a historical archive, most similar to the current situation, according to some meteorological and hydrological parameters relevant to streamflow formation. Different algorithms were developed and tested off-line, considering various predictors and combination of predictors for the selection of analogues. For each proposed method, probabilistic and deterministic daily streamflow forecasts were obtained each day for the next 3 days, by selecting observed streamflow in the 3-day period following each selected analogue, from the ensemble of 50 best analogues.

Results indicated that all tested methods provided reliable prediction intervals. Rescaling the analogues with the last observed discharge gave the best results and substantially outperformed persistence. The two methods combining synoptic-scale and basin-scale predictors performed the best and were relatively comparable in terms of skills, although not much better than the method using basin-scale predictors only. Using *MSLP* fields only gave the poorest forecasts of the 4 methods but the advantage of this technique is that it is quick and easy to implement. This particular method delivers a forecast for all catchments in once, while methods using basin-scale information are specific to each catchment and require more data for their implementation. In particular, if accurate precipitation and temperature estimates are not available for the catchment of interest, these methods cannot be applied and the method using *MSLP* fields only offers a good alternative.

The analogue method proposed here has proven to be a useful, yet simple tool for probabilistic and deterministic streamflow forecasting. The method capitalises on historical information collected on the catchment and can therefore be seen as an objective expert system based on past knowledge and experience. This method could be used in complement to a traditional hydrological forecasting system based on the coupling of a NWP model and hydrologic model. Once such a system will be in place at IMO, the performance of the two approaches should be compared. More work is still needed to refine the analogue method, such as the optimization of the synoptic domain size, the size of the moving temporal window and the number of selected analogues. More work is also needed to investigate the use of other parameters or combination of parameters, in the analogue search.

In practise, some rivers may be affected by icing conditions in winter, disrupting the water-level measurements and therefore the validity of the rating curves and the conversion into discharge. This problem could have an effect on the operational application of the proposed method and should be investigated. An alternative method to probabilistic streamflow forecasting could be to apply analogue methods for the probabilistic forecast of meteorological parameters used as input to a hydrological model. This method should be investigated as well.

# 7 Acknowledgements

This study was supported by the Landsvirkjun Research Fund. The author is grateful to Bolli Pálmason, Tinna Þórarinsdóttir and Bogi Brynjar Björnsson for their assistance. Comments and suggestions by Bergur Einarsson were greatly appreciated.

# 8 References


Akbari, M., Van Overloop, P. J. & Afshar, A. (2011). Clustered K nearest neighbor algorithm for daily inflow forecasting. Water Resour. Manage. 25, 1341–1357.

Crochet, P., Jóhannesson, T., Jónsson, T., Sigurðsson, O., Björnsson, H., Pálsson, F. & Barstad, I. (2007). Estimating the spatial distribution of precipitation in Iceland using a linear model of orographic precipitation. J. Hydrometeorol., 8, 1285–1306.

Crochet, P. & Jóhannesson, T. (2011). A dataset of daily temperature in Iceland for the period 1949–2010. Jökull, 61, 1–17.

Crochet, P. (2013). Sensitivity of Icelandic river basins to recent climate variations. Jökull, 63, (in press).

Diomede T., Nerozzi, F., Paccagnella, T. & Todini, E. (2008). The use of meteorological analogues to account for LAM QPF uncertainty. Hydrol. earth Syst. Sci., 12, 141–157.

Einarsson, B. & Jónsson, S. (2010). *The effect of climate change on runoff from two watersheds in Iceland.* Icelandic Meteorol. Office, Rep. 2010-016. Available at http://www.vedur.is/-media/ces/2010_016.pdf

Fraedrich K. & Rückert, B. (1998). Metric adaption for analog forecasting. Physica A, 253, 379–393.

Fraedrich K., Raible, C. C. & Sielmann, F. (2003). Analog ensemble forecasts of tropical cyclone tracks in the Australian region. Wea. Forecasting, 18, 3–11.

Galeati, G. (1990). A comparison of parametric and non-parametric methods for runoff forecasting. Hydrol. Sci. J., 35, 1, 79–94.

Gibergans-Báguena, & Llasat, M.C. (2007). Improvement of the analog forecasting method by using local thermodynamic data. Application to autumn precipitation in Catalonia. Atmos. Res., 86, 173–193.

Hamill, T. M. & Whitaker, J. S. (2006). Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. Mon. Wea. Rev., 134, 3209–3229.

Jóhannesson, T., Aðalgeirsdóttir, G., Björnsson, H., Crochet, P., Elíasson, E. B., Guðmundsson, S., Jónsdóttir, J. F., Ólafsson, H., Pálsson, F., Rögnvaldsson, Ó., Sigurðsson, O., Snorrason, Á., Blöndal Sveinsson, Ó. G. & Thorsteinsson Th. (2007). *Effect of climate change on hydrology and hydro-resources in Iceland.* National Energy Authority, Rep. OS-2007/011, 91 pp.

Jónsdóttir, J. F. (2008). A runoff map based on numerically simulated precipitation and a projection of future runoff in Iceland. Hydrol. Sci. J. 53(1), 100–111.

Jónsdóttir, J. F., Uvo, C. B. & Clarke, R. T. (2008). Trend analysis in Icelandic discharge, temperature and precipitation series by parametric methods. *Hydrology Res.* 39(5–6), 425–436.

Karlsson, M. & Yakowitz, S. (1987). Nearest-neighbor methods for non parametric rainfall-runoff forecasting. Water Resour. Res., 23 (7), 1308–1330.

Kruizinga, S. & Murphy, A.H. (1983). Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands. Mon. Wea. Rev., 111, 2244–2254.

Marty, R., Zin, I., Obled, C., Bontron, G. & Djerboua, A. (2012). Toward real-time daily PQPF by an analog sorting approach: application to flash-flood catchments. Journal of Applied Meteorology and Climatology, 51, 505–520.

Obled, C., Bontron, G. & Garcon, R. (2002). Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach. Atmos. Res., 63, 303–324.

Panziera, L., Germann, U., Gabella, M. & Mandapaka, P. V. (2011). NORA–Nowcasting of orographic rainfall by means of analogues. Quart. J. Roy. Meteor. Soc., 137, 2106–2123.

Piggott, A. R., Moin, S. & Southam, C. (2005). A revised approach to the UKIH method for the calculation of baseflow. Hydrol. Sci. J. 50(5), 911–920.

Radinović, D. (1975). An analogue method for weather forecasting using the 500/1000 mb relative topography. Mon. Wea. Rev., 103, 639–649.

Rist, S. (1990). *Vatns er þörf.* Bókaútgáfa Menningarsjóðs, 248 pp. (In Icelandic).

Snorrason, Á., Einarsson, B., Pagneux, E., Harðardóttir, J., Roberts, M. J., Sigurðsson, O., Þórarinsson, Ó., Crochet, P., Jóhannesson, T. & Þorsteinsson, T. (2012). Floods in Iceland (Ch. 13). In: "Changes in flood risk in Europe". IAHS Special publication 10. Edited by Z.W. Kundzewicz.

Uppala, S.M. & co-authors (2005). The ERA-40 re-analysis. Quart. J. Roy. Meteor. Soc., 131, 2961–3012.

Van den Dool, H. M. (1989). A new look at weather forecasting through analogues. Mon. Wea. Rev., 117, 2230–2247.

Wetterhall, F., Halldin, S. & Xu, C.Y. (2005). Statistical precipitation downscaling in central Sweden with the analogue method. Journal of Hydrology, 306, 174–190.

Wilks, D. S. (1995). Statistical methods in the atmospheric sciences. Academic Press, San Diego, California, 467 pp.

Woodcock, F. (1980). On the use of analogues to improve regression forecasts. Mon. Wea. Rev., 108, 292–297.

Zorita, E. & von Storch, H. (1999). The analog method as a simple downscaling technique: Comparison with more complicated methods. J. Climate, 12, 2474–2489.

Þórarinsdóttir, T. (2012). Development of a methodology for estimation of technical hydropower potential in Iceland using high resolution hydrological modeling. University of Iceland. 106 pp.

## Appendix 1

## Test of reliability for the prediction intervals. Observed versus predicted prediction intervals.

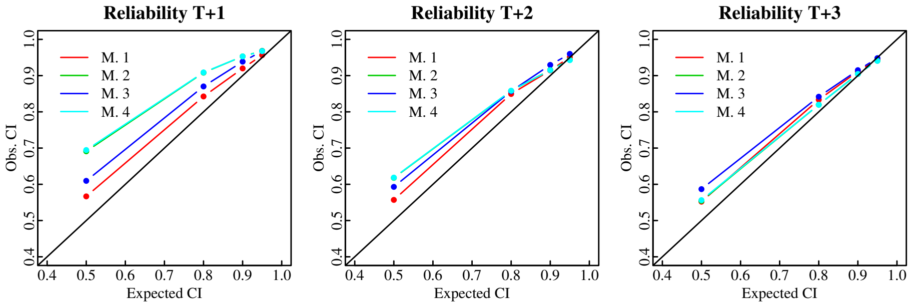



*Figure I.1. vhm 64*

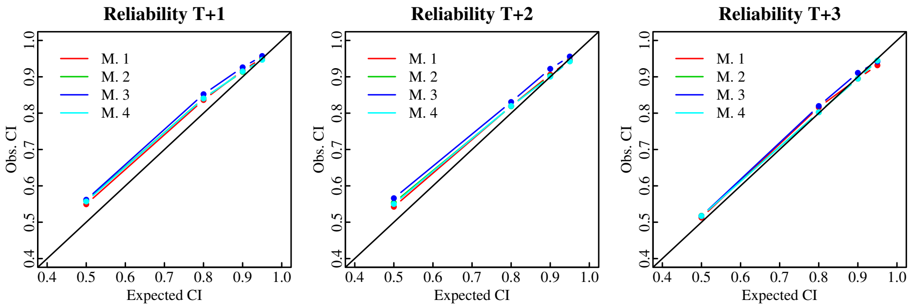



*Figure I.2. vhm 64 with rescaling*

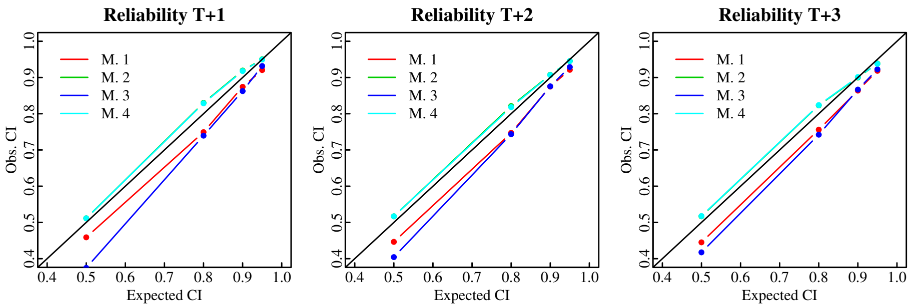



*Figure I.3. vhm 66*

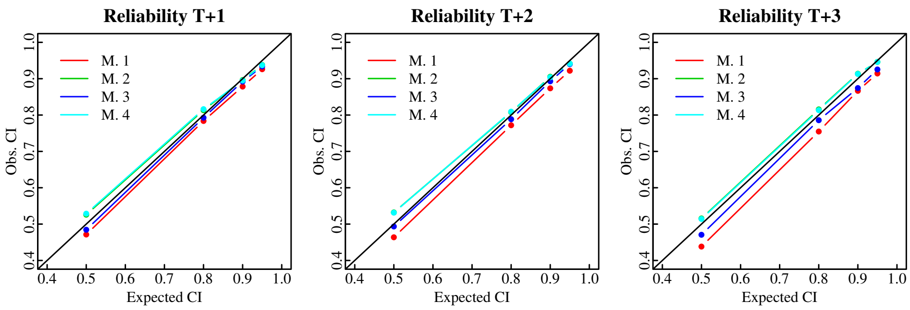



*Figure I.4. vhm 66 with rescaling*

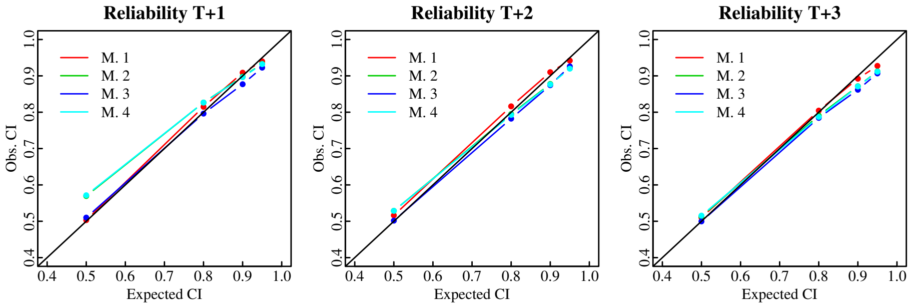



*Figure I.5. vhm 145*

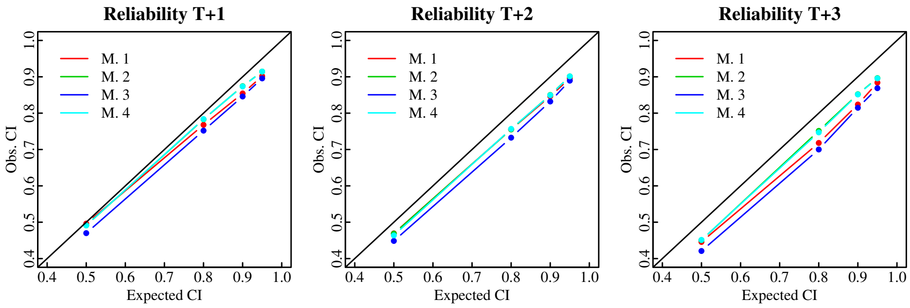



*Figure I.6. vhm 145 with rescaling*

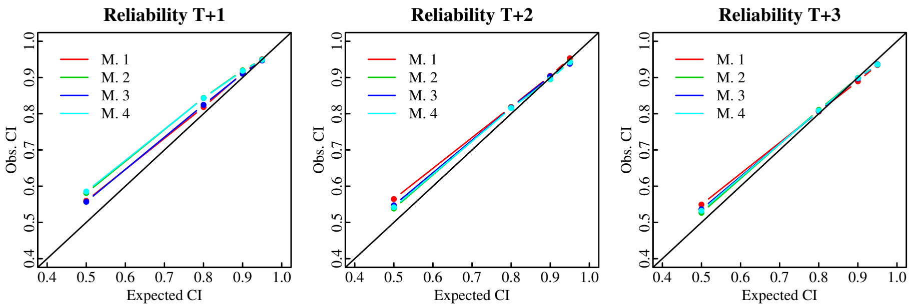



*Figure I.7. vhm 10*

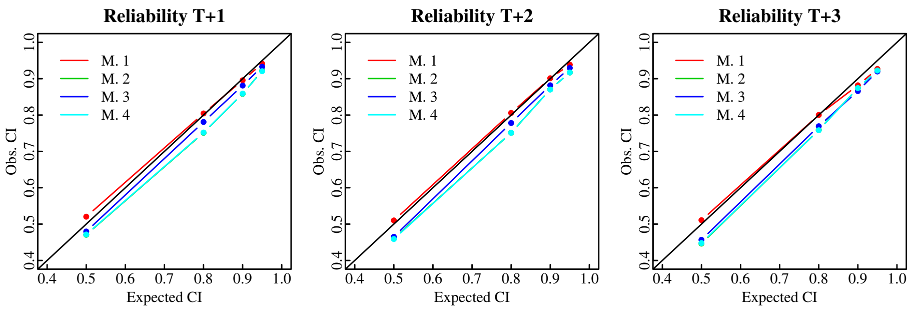



*Figure I.8. vhm 10 with rescaling*

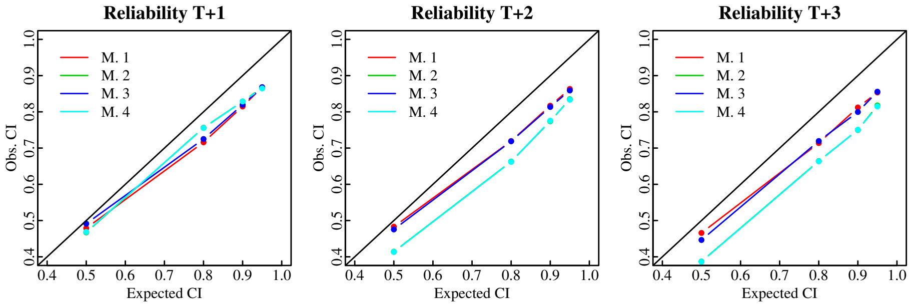



*Figure I.9. vhm 19*

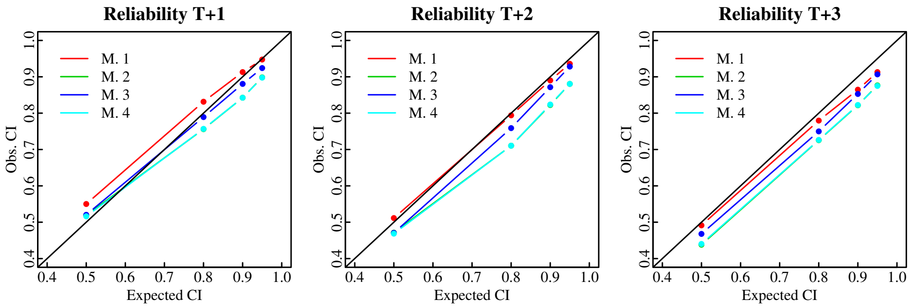



*Figure I.10. vhm 19 with rescaling*

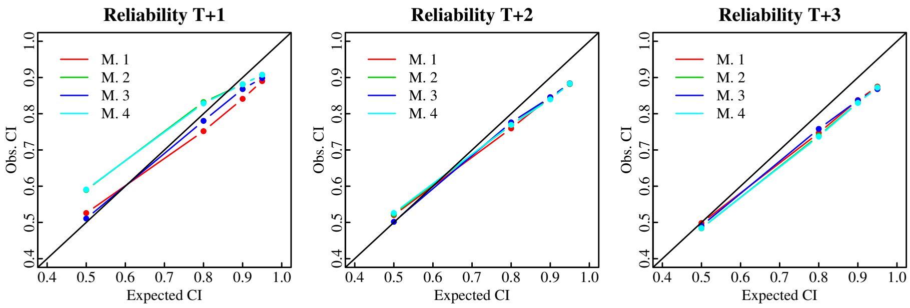



*Figure I.11. vhm 26*

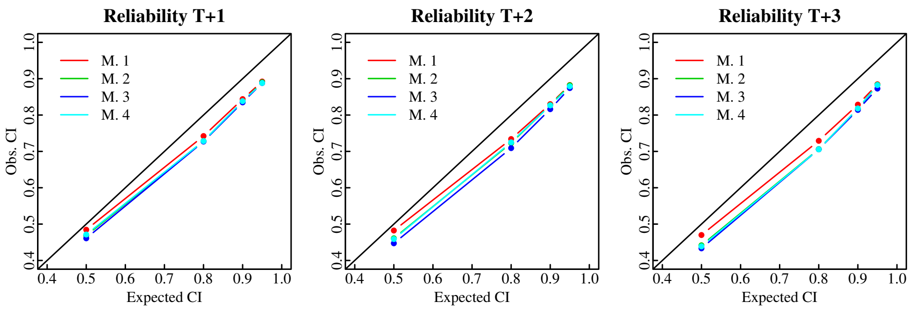



*Figure I.12. vhm 26 with rescaling*

## Appendix 2

**Observed daily discharge versus deterministic forecast over the period 01/09/2001–31/08/2005. The solid line corresponds to a perfect match. For methods 1 to 4, rescaling is applied.**

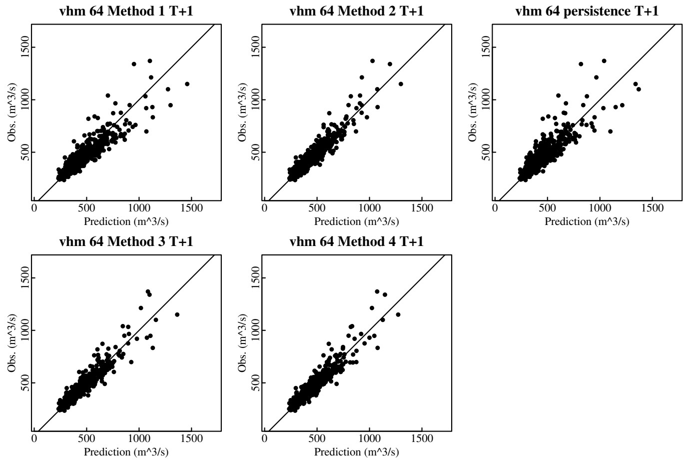



*Figure II.1. vhm 64, T=1 day*

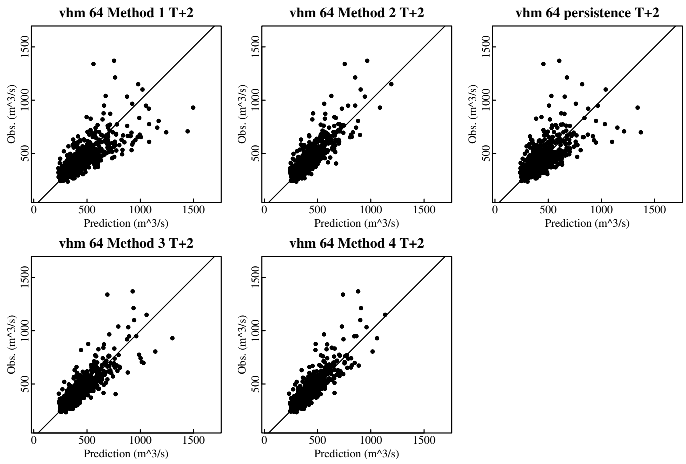



*Figure II.2. vhm 64, T=2 days*

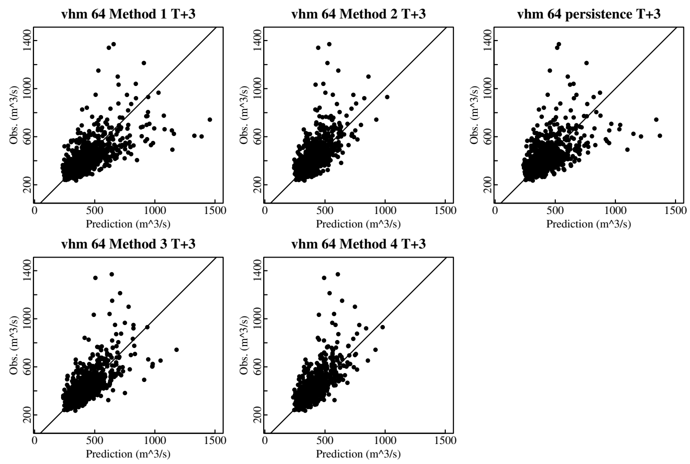



*Figure II.3. vhm 64, T=3 days*

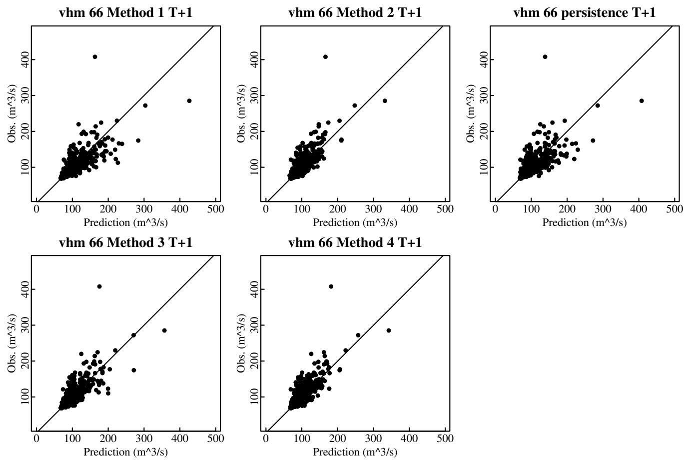



*Figure II.4. vhm 66, T=1 day*

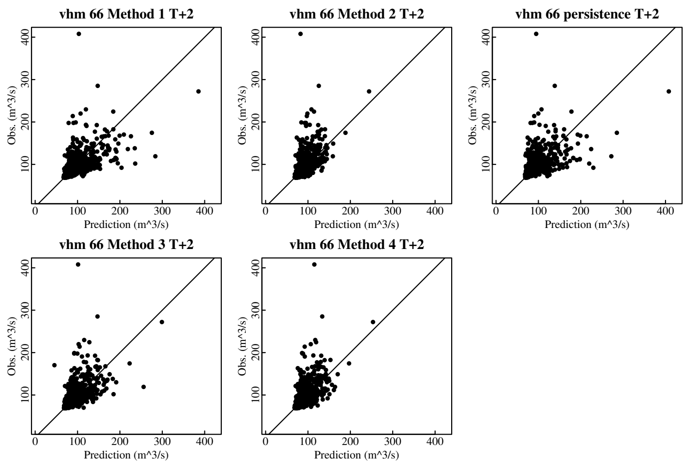



*Figure II.5. vhm 66, T=2 days*

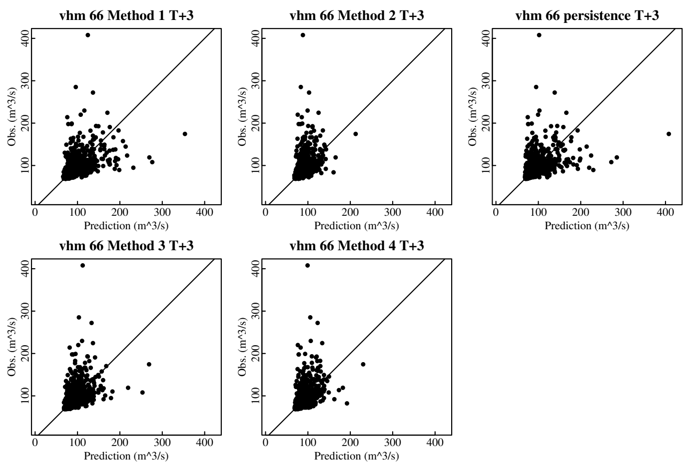



*Figure II.6. vhm 66, T=3 days*

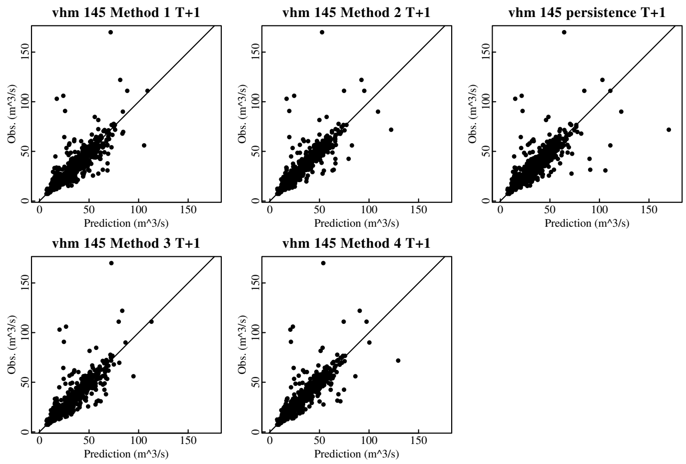



*Figure II.7. vhm 145, T=1 day*

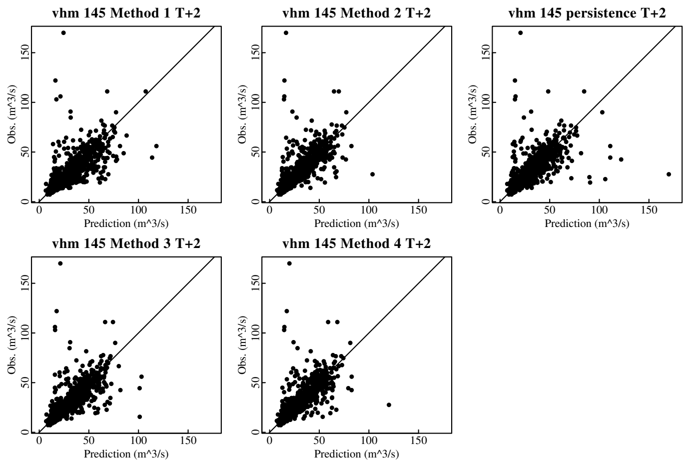



*Figure II.8. vhm 145, T=2 days*

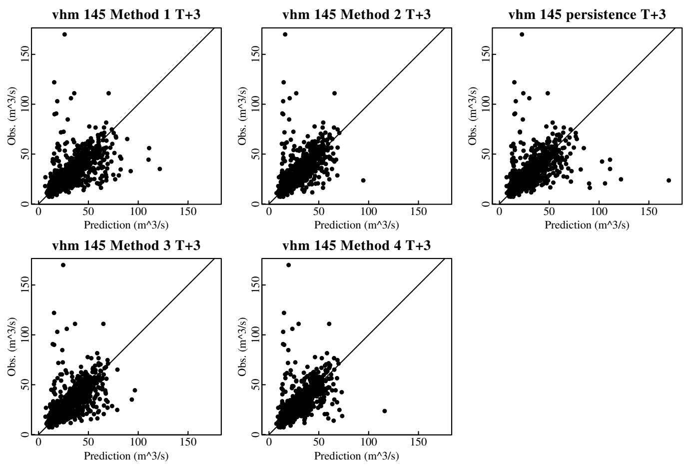



*Figure II.9. vhm 145, T=3 days*

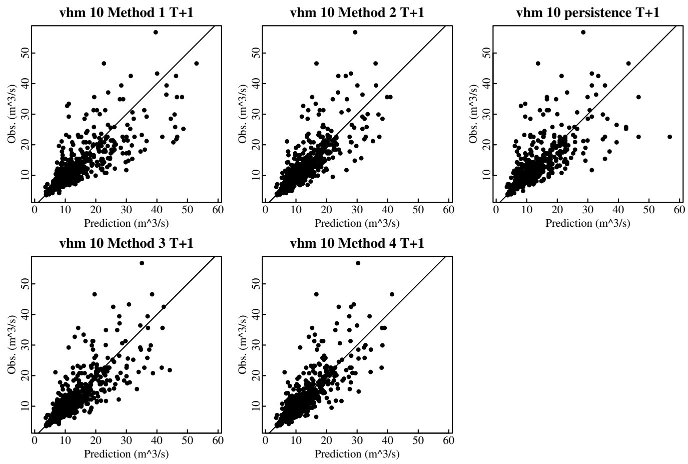



*Figure II.10. vhm 10, T=1 day*

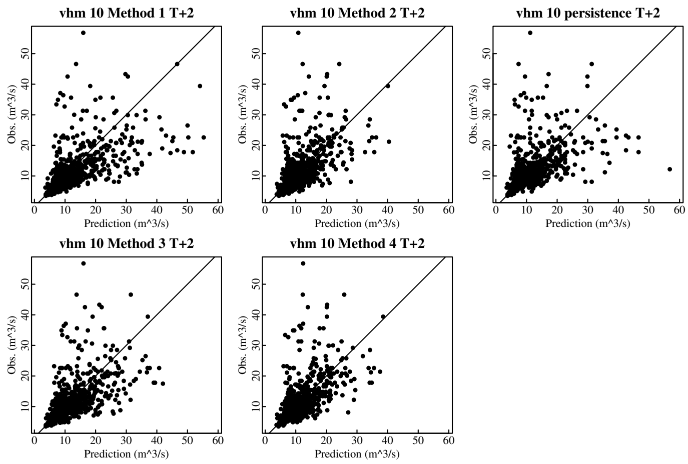



*Figure II.11. vhm 10, T=2 days*

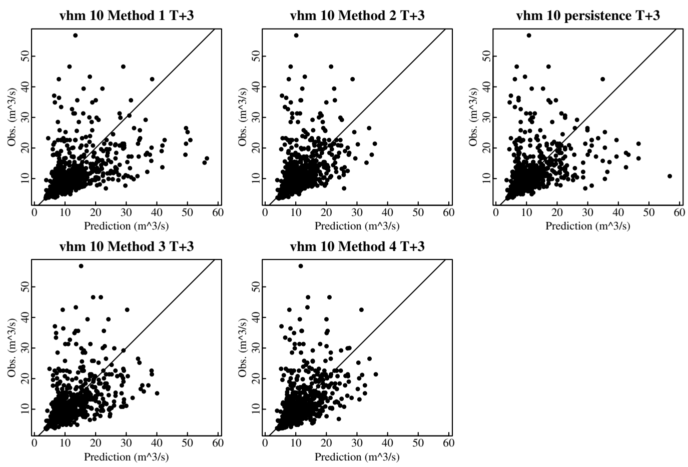



*Figure II.12. vhm 10, T=3 days*

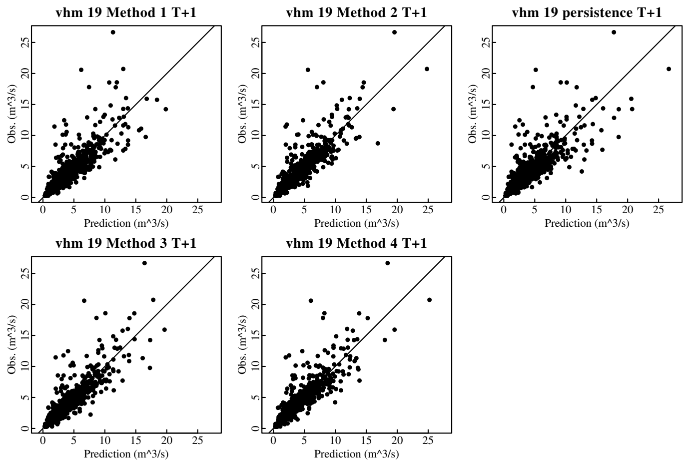



*Figure II.13. vhm 19, T=1 day*

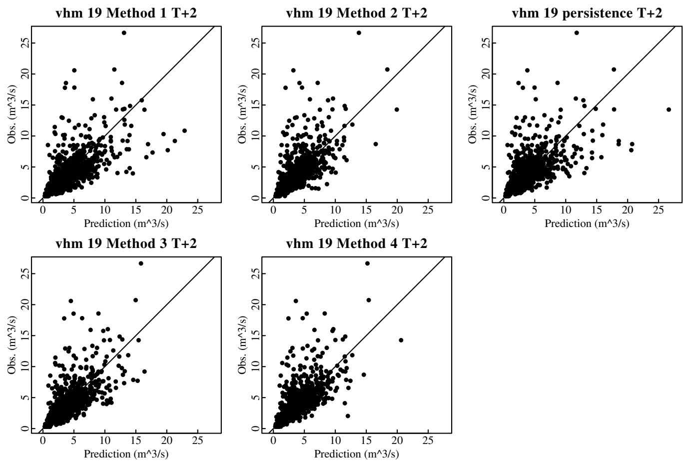



*Figure II.14. vhm 19, T=2 days*

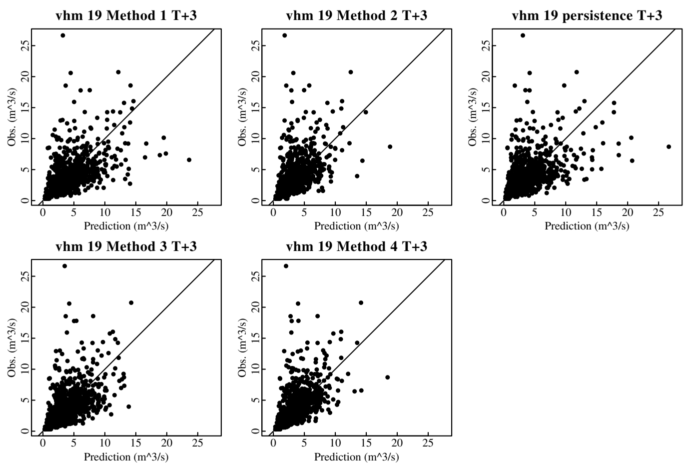



*Figure II.15. vhm 19, T=3 days*

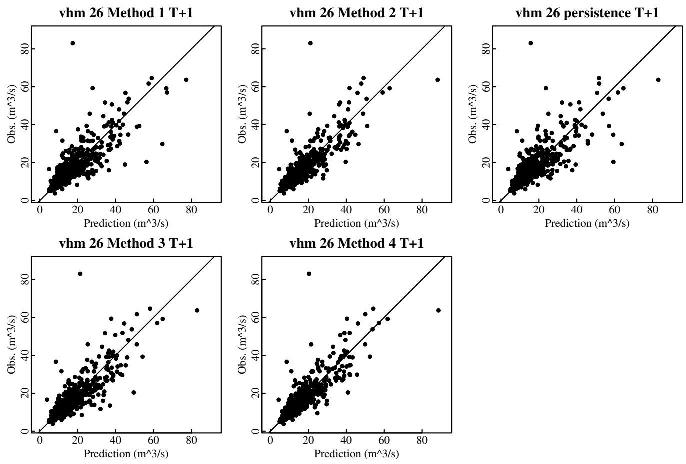



*Figure II.16. vhm 26, T=1 day*

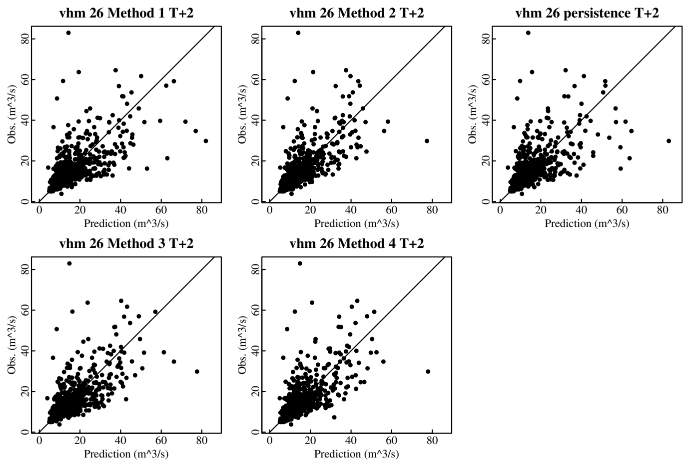



*Figure II.17. vhm 26, T=2 days*

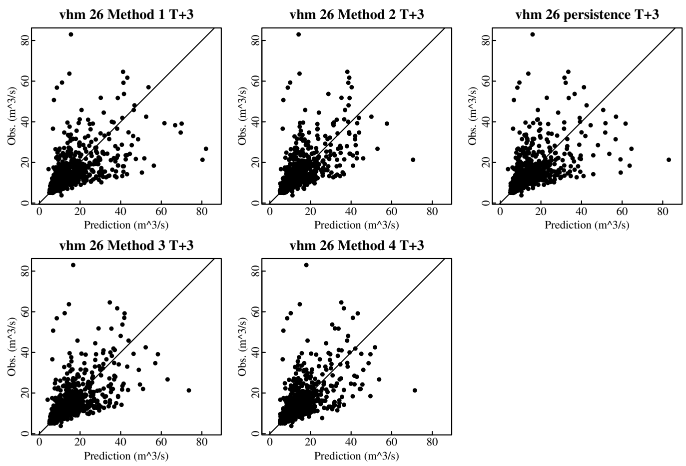



*Figure II.18. vhm 26, T=3 days*