Statistical summer mass-balance forecast model with application to Brúarjökull glacier, South East Iceland

DARRI EYTHORSSON,[1] SIGURDUR M. GARDARSSON,[1] ANDRI GUNNARSSON,[2] BIRGIR HRAFNKELSSON[3]

[1]*Faculty of Civil and Environmental Engineering, University of Iceland, Hjardarhaga 2-6, Reykjavik 107, Iceland*
[2]*Research and Development Division, Landsvirkjun, Haaleitisbraut 68, Reykjavik 103, Iceland*
[3]*Department of Mathematics, Faculty of Physical Sciences, University of Iceland, Reykjavik 107, Iceland*
Correspondence: Darri Eythorsson <dae5@hi.is>

ABSTRACT. Forecasting of glacier mass balance is important for optimal management of hydrological resources, especially where glacial meltwater constitutes a significant portion of stream flow, as is the case for many rivers in Iceland. In this study, a method was developed and applied to forecast the summer mass balance of Brúarjökull glacier in southeast Iceland. In the present study, many variables measured in the basin were evaluated, including glaciological snow accumulation data, various climate indices and meteorological measurements including temperature, humidity and radiation. The most relevant single predictor variables were selected using correlation analysis. The selected variables were used to define a set of potential multivariate linear regression models that were optimized by selecting an ensemble of plausible models showing good fit to calibration data. A mass-balance estimate was calculated as a uniform average across ensemble predictions. The method was evaluated using fivefold cross-validation and the statistical metrics Nash–Sutcliffe efficiency, the ratio of the root mean square error to the std dev. and percent bias. The results showed that the model produces satisfactory predictions when forced with initial condition data available at the beginning of the summer melt season, between 15 June and 1 July, whereas less reliable predictions are produced for longer lead times.

KEYWORDS: glacier mass balance, glacier modelling, ice and climate, melt-surface

1. INTRODUCTION

Water storage in snow and ice is an important factor in the hydrological cycle in many regions of high altitudes and latitudes like Iceland, where 11% of the country is covered by glaciers (Bjornsson and Palsson, 2008). Simulation and prediction of melt behavior provide valuable information for water resources management, e.g. regarding reservoir fill rates, potential power production and load on hydraulic structures. Short-term predictions of a few days improve daily operations and risk analysis, whereas longer term prediction of seasonal melt behavior assists in the systematic optimization of networks of reservoirs and diversions.

Prior work in melt modeling of Icelandic glaciers has focused on either empirical (degree day) and physical (energy balance) modeling approaches. Both have shown good performance for simulating glacial mass balance (e.g. De Ruyter de Wildt and others, 2003b; Marshall and others, 2005; Carenzo and others, 2009; Engelhardt and others, 2014). Empirical approaches to mass-balance modeling have been considered sufficient when the underlying physical processes need not be resolved (e.g. Réveillet and others, 2017). More recently, the vast potential of remote-sensing data has been increasingly applied to snowmelt estimation in basins where little information is available (Kalra and others, 2013; Qiu and others, 2014; Drolon and others, 2016).

Observations have shown that across the globe glaciers are losing mass and retreating. These studies have further concluded that the rapid retreat in the early 21st century is without precedent on a global scale (Barnett and others, 2005; Liu and others, 2015; Zemp and others, 2015; Roe and others, 2017). In line with the trend of glaciers globally, Icelandic glaciers have experienced retreat in recent years and their mass loss since the end of the 19th century has been shown to be responsible for 0.03 mm sea level rise globally (Bjornsson and others, 2013). Studies of the response of Icelandic glaciers to the expected climate change have shown that the country's main ice caps will mostly disappear over the next two centuries, leaving glaciers only at the highest elevations (Aðalgeirsdóttir and others, 2006, 2011).

Studies have predicted that increased glacial ablation will result in increased river runoff in Icelandic rivers throughout the 21st century (Jonsdóttir, 2010; Gudmundsson and others, 2011; Matthews and others, 2015). While little prior work exists on summer mass-balance modeling of Icelandic glaciers, several studies have considered the subject in other regions. These attempts have either employed statistical modeling techniques or used physical models forced with climate simulations (Fujita and Ageta, 2000; Schöner and Böhm, 2007).

The present study considers the prediction of the summer mass balance of Brúarjökull in SE Iceland. The Brúarjökull catchment, which is more than 90% glacierized, was harnessed for hydropower generation by the construction of the Halslon reservoir in 2006 (Gardarsson and Eliasson, 2006). Due to its hydroelectric potential, data have been recorded on hydro-meteorological variables in the catchment, including measurements of glacier mass balance since 1993 (De Ruyter de Wildt and others, 2003a, b; Rasmussen, 2005).

Brúarjökull covers an area of 1550 km², making it the largest outlet glacier of the Vatnajökull ice cap, representing
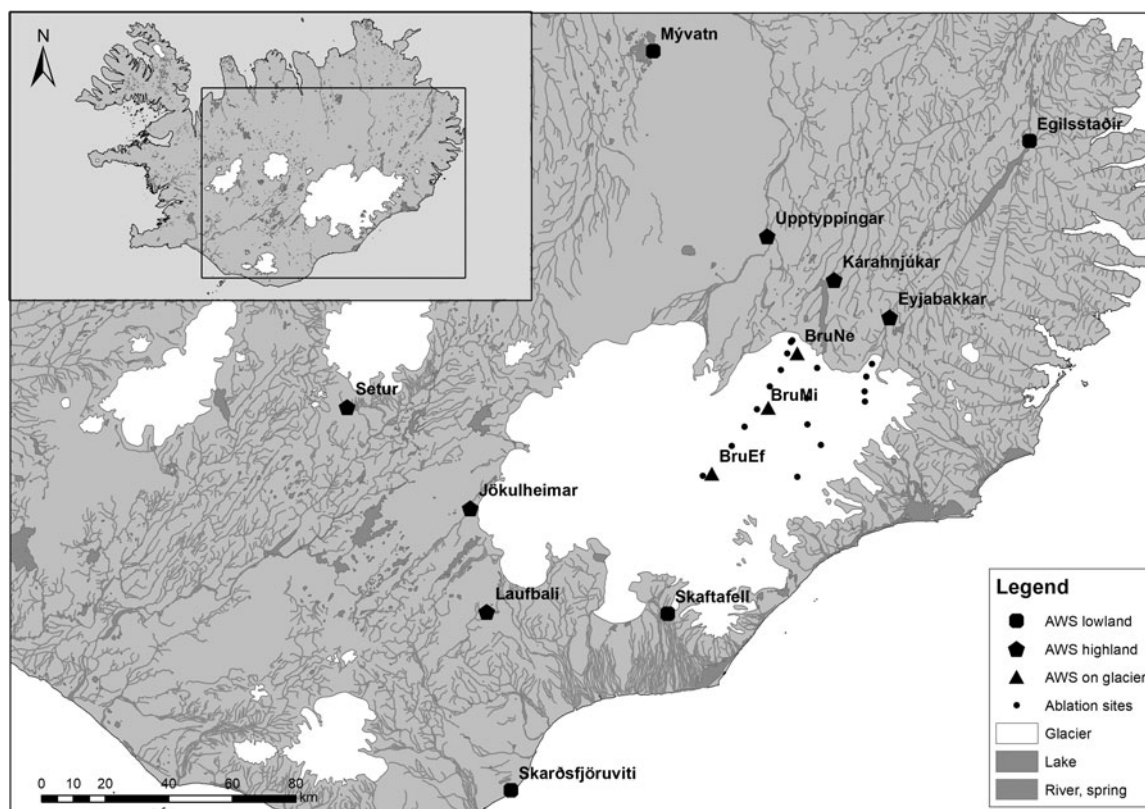
Fig. 1. Location of mass-balance points and automatic weather stations (AWS) which collect the meteorological data that were used in the study (Data on land cover from National Land Survey of Iceland).

~19% of the total area covered by the ice cap. The glacier ranges in elevation from 600 to ~1550 m a.s.l. and the mean equilibrium line lies at an altitude ~1200 m a.s.l. (Bjornsson and Palsson, 2008). The glacier slopes gently toward the central Icelandic highland plateau and is classified as a surging outlet glacier with a surge frequency of 80–100 years, the last one occurring in 1964 (Kjær and others, 2008). Unlike other outlets of the ice cap, Brúarjökull is not underlain by geothermal areas. Due to the proximity to surrounding volcanoes, its surface is periodically covered in volcanic tephra, thus decreasing its albedo (Larsen, 1998; Moller and others, 2014). In the three main volcanoes near the basin, Bárðarbunga, Grímsvötn and Kverkfjöll, tephra events occur on average ~15 eruptions per century (Oladottir and others, 2011).

The forcing of physically based melt models with meteorological forecast model output on seasonal time scales inevitably incurs the large uncertainty in the forcing data. In this paper, statistical modeling was investigated to attempt the prediction of summer mass balance directly from the initial conditions of the system on the forecast date, thereby minimizing the uncertainties. The motivation for the study was to investigate whether the mass balance of the Brúarjökull could be predicted at the beginning of the melt season and to develop a simple operational model for reservoir operators. The goal of the study was to assess the predictive power of the information available by employing statistical approaches and the impact of lead times on predictions.

2. DATA

The data used in the present study consisted of glaciological mass-balance measurements, meteorological variables

measured around the Brúarjökull basin, and climate indices which have been shown to correlate with Icelandic weather patterns (e.g. Baldwin and others, 2003; Hanna and others, 2004).

2.1. Glaciological measurements

Winter accumulation and summer ablation of Vatnajökull are measured in biannual measurement surveys at the boundaries of the melt season in spring and autumn. Winter accumulation is estimated by drilling ice cores and the summer ablation is measured from ablation wires or rods that are placed on the glacier in spring, when winter accumulation is measured (Thorsteinsson and others, 2004). The annual net mass balance is calculated as the sum of the winter accumulation and the summer ablation. Figure 1 shows the approximate location of mass-balance sites on the surface of Brúarjökull as small circles.

The annual mass balance within each catchment on the glacier has been estimated based on the ablation stake measurements by extrapolating across the area (Palsson and others, 2014). The summer mass balance within the Halslon reservoir catchment was used as the response variable in the present study, while the winter accumulation at the various accumulation sites was used as an input variable. It should be noted that the estimated mass balance did not include liquid precipitation that fell on the glacier during the summer nor snow that melted outside the survey period. Furthermore, the uncertainty in the mass-balance measurements is not reported. The glaciological summer mass-balance data were selected as response variable based on the overlap of the shorter time series of discharge for the reservoir inflow which started in 2007.

2.2. Meteorological variables

Data were obtained from eight automatic weather stations (AWS) in and around the Brúarjökull basin. Three AWS on the glacier surface were used at elevations of 850, 1200 and 1400 m a.s.l. denoted as BruNe, BruMi and BruEf, respectively. The stations are designed to collect measurements of the components of the snow surface energy balance, shown as triangles in Figure 1. Time series of the measurements were acquired as daily averages of the following parameters: air temperature, relative humidity, net radiation, wind speed and surface albedo.

Data from ten land-based AWS surrounding the Brúarjökull basin were obtained, six based in the central highlands and four in the lowlands. The locations of the land-based AWS are shown as squares and pentagons in Figure 1. Time series were obtained as mean daily values of the following parameters: air temperature, dew point temperature, vapor pressure, relative humidity, atmospheric pressure and wind speed, and additionally precipitation measurements from the AWS at Egilsstadir.

2.3. Climatological variables

Icelandic climate has been shown to be significantly influenced by prevailing ocean conditions surrounding the island as well as changes in the large-scale circulations in the North Atlantic Ocean (Hanna and others, 2001). Large-scale changes in atmospheric circulation have also been shown to correlate significantly with long-term Icelandic climate trends (e.g. Hanna and others, 2004).

To incorporate information on the variability in the ocean conditions surrounding Iceland, the following two datasets were acquired: monthly averages of the Atlantic Multidecadal Oscillation (AMO) index (Enfield and others, 2001) and quarterly averages of the heat content of the Northern Atlantic (60–0°W, 30–65°N) measured in the top 700 m of the ocean by the US National Oceanic Data Center (NODC). The AMO index is defined from the trends in Sea Surface Temperature (SST) in the North Atlantic and has been shown to be correlated with temperature and precipitation patterns in Europe (Ghosh and others, 2017; Zampieri and others, 2017). Furthermore, the heat transport through the North Atlantic by the warm Gulf Stream has been shown to be a key factor in determining the climate of Northern Europe (e.g. Palter, 2015).

To incorporate information about the atmospheric circulations into the model, monthly averages of the North Atlantic Oscillation Index (NAOI) were acquired from the US National Oceanic and Atmospheric Administration (NOAA). The NAOI is a measure of the changes in the difference in atmospheric pressure at sea level between the Icelandic and the Azores. Studies have shown the NAOI to be significantly correlated with temperature and precipitation patterns in Iceland (Hanna and others, 2004).

3. METHODS

3.1. Time series

The data were obtained as hourly or daily averages from the AWS and as point measurements of the winter accumulation data and climatological indices. The AWS data were aggregated to average values to represent the initial conditions of the system at four different dates in spring, specifically for
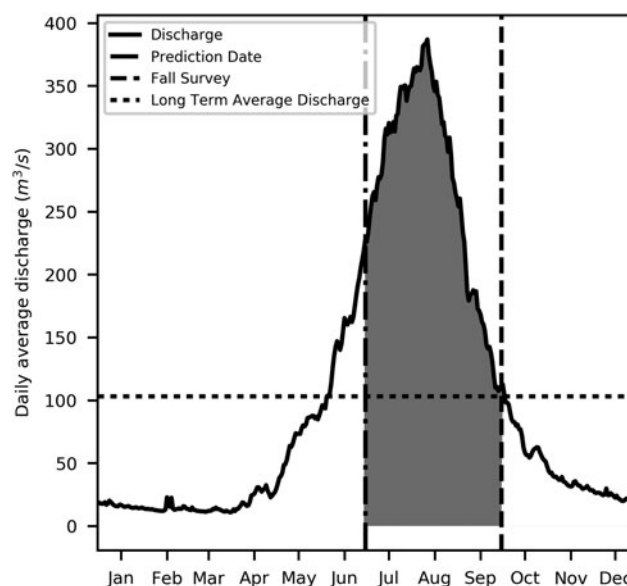

Fig. 2. Average daily discharge into Halslon reservoir for the period 2007–2015. The shaded area represents a proxy for the predicted mass balance.

the periods beginning on 1 April and ending on 15 May, 1 June, 15 June and 1 July.

The main aim of the study was to predict the summer inflow into the Hálslón Reservoir. Due to the short time series of inflow (2007–2015) the summer mass balance of Brúarjökull was selected as a proxy. Figure 2 shows the average daily discharge into the Halslon reservoir, where the shaded area shows the period between the forecast date on 1 July to the time of the fall ablation survey when the total summer mass balance of the glacier is calculated. The mass-balance data do not represent hydrological fluxes such as the drainage from the 10% of the basin, which is de-glacierized, baseflow and basal melt due to geothermal fluxes and liquid precipitation that falls on the glacier in summer. Despite this, we consider that the summer mass balance of the 90% glacierized portion of the basin offers a good representation of the inter-annual variability of net summer inflow into the reservoir. Thus, the total inflow, represented by the shaded area under the curve, will be significantly correlated to the summer mass balance, the quantity to be predicted in the present study.

The method was initially applied to 1 July data; then predictions of the summer mass balance were produced for each of the dates to assess the evolution of the predictive performance of the modeling approach in the period leading up to the summer melt season. The availability of the acquired data overlapped for the period 2001–2015, which was selected as the study period for the research. A breakdown of the input variables screened in the study along with their correlation to Brúarjökull summer mass balance is given in Appendix A.

The number of years used in the present study were $N = 15$. The input data were aggregated to a single average value for each year that represented the initial conditions of the system prior to the date of prediction. The data were split into training and test sets using the K-fold cross-validation method. In the present study, K was selected as 5 and the dataset was split into five subsets, each using three observations to test the model and 12 observations for calibration. A

fivefold cross-validation was selected over a leave-one-out approach (where $K = N$) to reduce the variance in the error estimates (James and others, 2013).

3.2. Variable selection

In the present study, many predictor variables were considered, whereas the number of observations of the response variable were few. To reduce the number of predictors and prevent model overfitting, variables were selected that showed a significant correlation with the response variable, the observed mass balance of the glacier. The variables were ranked by their r^2 value, and variables with r^2 values above a certain threshold were selected for further model development. The threshold value for variable selection was determined by sensitivity analysis of model results, as described in the subsequent sections.

3.3. Multivariate model ensemble

The selected variables can be used to create many multivariate regression models, none of which may be obviously superior to any of the others. Rather than selecting any single model, an ensemble of all potential competing models was developed. The selected variables were used to calculate a set of all possible multivariate linear regression models comprising five or fewer input variables. The input variables were limited to five due to computational limitations and the potential risk of overfitting the short time series. The optimal number of input parameters to include in the models of the ensemble was investigated by sensitivity analysis of model results with a range of numbers of input parameters, as described in subsequent sections.

3.4. Multi-model inference

Selection of any single one of the regression models in the set of possible models would recognize the existence of several potential and competing models and introduce additional uncertainty in the estimator due to the model selection. Unless the uncertainty associated with model selection is accounted for, overconfident estimates of model predictions may be inferred (Wang and others, 2009).

An alternative to selecting a single model is to average the prediction over a range of plausible models. This technique, called model averaging, incorporates the uncertainty associated with model selection into predictions of unknown variables (Hjort and Claeskens, 2003). The model averaging approach has in recent years been applied to several hydrological model applications (Diks and Vrugt, 2010; Tsai, 2010).

Methods for model averaging include Bayesian model averaging (BMA) and frequentist model averaging (FMA). In BMA, model uncertainty is evaluated by assigning prior probabilities to all models being considered, whereas in the FMA, no prior probabilities are required and all estimators are determined by the data (Buckland and others, 1997; Raftery and others, 1997; Hoeting and others, 1999). In the present study, the FMA approach to model averaging was chosen as it relies only on the available data.

The response variable was estimated from a model ensemble by calculating the ensemble average. Several weighting functions have been reported in the literature to incorporate the measures of model plausibility into model averaging,

based, for example, on goodness-of-fit metrics Akaike information criterion (Buckland and others, 1997), Bayesian information criterion and focused information criterion (FIC) (Burnham and Anderson, 2002; Zhang and others, 2012). Other strategies for weight function selection include the minimization of Mallow's C_p criterion and weight choice based on the unbiased estimator of risk (Liang and others, 2011). In cases where little prior information is available on the likelihood of each model, or models having similar priors, assigning a uniform weight to each model is a reasonable choice (Raftery and others, 1997). In the present study, a uniform weight was selected.

3.5. Optimal subset of models

Another important consideration of the model averaging methodology is the selection of a set of models over which to average. A complete Bayesian solution to the problem is to average over the entire set of possible models (Madigan and Raftery, 1994). However, as the set of potential models can become large, averaging over the entire set may not be practical. To reduce the number of models to be considered, Madigan and Raftery (1994) suggested excluding models that poorly fit the calibration data.

The quality of each model in the set of possible models was assessed by several evaluation metrics. Moriasi and others (2007) surveyed several model evaluation metrics for watershed simulations and recommended using three metrics: the Nash–Sutcliffe efficiency (NSE), the ratio of the root mean square error to the std dev. of measured data (RSR) and the percent bias ($PBIAS$) for evaluation of hydrological models (Moriasi and others, 2007). These three metrics were selected for model evaluation in the present study; their mathematical formulations are described as:

$$NSE = 1 - \frac{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}} - Y_i^{\mathrm{sim}}\right)^2}{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}} - Y^{\mathrm{mean}}\right)^2}, \tag{1}$$

$$RSR = \frac{RMSE}{STDEV_{\mathrm{obs}}} = \frac{\sqrt{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}} - Y_i^{\mathrm{sim}}\right)^2}}{\sqrt{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}} - Y^{\mathrm{mean}}\right)^2}}, \tag{2}$$

$$PBIAS = \frac{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}} - Y_i^{\mathrm{sim}}\right) \times (100)}{\sum_{i=1}^{n} \left(Y_i^{\mathrm{obs}}\right)}, \tag{3}$$

where n is the number of data points in the dataset, Y_i^{obs} is the observed mass balance in the ith year, Y_i^{sim} is the simulated mass balance in the ith year and Y^{mean} is the mean observed mass balance. Moriasi and others (2007) suggested that a model simulation could be judged as satisfactory if $NSE > 0.5$, $RSR < 0.7$ and $PBIAS < \pm 25\%$.

An ensemble of plausible models was created by evaluating all models in the set of possible multivariate regression models in accordance with the recommended values of NSE, RSR and $PBIAS$. Models with $NSE < 0.5$, $RSR > 0.7$ and $PBIAS > \pm 25\%$ were eliminated from further analysis and the remaining models were stored for multi-model inference.

Madigan and Raftery (1994) suggested that, in the case of models that fit the calibration data equally well, the less complicated model should be selected as it receives more support from the data. In the present study, a sensitivity analysis was

Table 1. *NSE* of different model configurations with varying $r^2_{threshold}$ and number input variables, optimal value of $NSE = 1$

	Number of variables				
	1	2	3	4	5
$r^2_{threshold}$					
0.2	−0.09	0.48	0.53	0.51	0.44
0.25	−0.03	0.55	0.56	0.54	0.48
0.3	0.23	0.69	0.71	0.71	0.66
0.35	0.24	0.61	0.64	0.64	0.44

Table 3. *PBIAS* of different model configurations, optimal value of $PBIAS = 0$

	Number of variables				
	1	2	3	4	5
$r^2_{threshold}$					
0.2	0.76	2.55	2.45	3.15	4.29
0.25	0.85	2.46	2.39	3.02	3.66
0.3	−2.65	−0.08	0.41	0.75	1.23
0.35	−3.05	−0.75	0.38	0.47	1.45

performed on model predictions by varying the number of allowed input variables in the models and thus identifying the optimal number of variables to include.

The model averaging estimate of glacier ablation, $\hat{A}$, is then given by

$$\hat{A} = \frac{1}{M}\sum_{k=1}^{M}\hat{A}_k, \qquad (4)$$

where the index k denotes the kth model considered, M is the total number of models and $\hat{A}_k$ is the estimated ablation based on the kth model. The uncertainty in the estimate is taken as the spread in predicted values of the ensemble of models.

4. RESULTS AND DISCUSSION

4.1. Multimodel inference

The selection of a threshold value of r^2 for variable selection and the number of input variables used in the model were optimized by performing a sensitivity analysis of the model results. The results were evaluated using the metrics *NSE*, *RSR* and *PBIAS* and were calculated using four threshold values of r^2 [0.2, 0.25, 0.3, 0.35] and five options for the number of model input variables [1, 2, 3, 4, 5]. Model ensembles were calculated for each combination of model options and the ensemble mean was used to calculate the evaluation metrics. The use of the median ensemble response was also investigated and yielded almost identical results. Tables 1–3 show the results for the evaluation metrics: *NSE*, *RSR* and *PBIAS*, respectively, while Table 4 shows the total number of models in each ensemble.

The results of the sensitivity analysis showed that the optimal values of *NSE* and *RSR* were obtained using a threshold value of $r^2 = 0.3$ and constraining the number of input parameters in the models to four (optimal results are

highlighted in Tables 2–4). In terms of *PBIAS*, the optimal configuration was found with a threshold $r^2 = 0.3$ and two input variables in the models. However, the *PBIAS* of several configurations showed very low bias including the optimal configuration in terms of *NSE* and *RSR*. Hence it was concluded that the optimal model ensemble was achieved by selecting potential input variables with $r^2 > 0.3$ and restricting the number of inputs into each model in the ensemble to three. As shown in Table 4, this model ensemble contains 35 plausible models.

4.2. Variable selection

The time series of all the acquired potential input variables were assessed based on their correlation with the observed summer mass balance of Brúarjökull. Variables with a correlation coefficient below a set threshold value of 0.3 as determined in Section 4.1 were eliminated from further analysis. The variables selected for model development and their corresponding r^2 values are presented in Table 5.

4.3. Model evaluation

The model was evaluated according to its ability to predict observed values of mass balance of the glacier in terms of the evaluation metrics *NSE*, *RSR* and *PBIAS* described in Section 3.5. The models were evaluated using fivefold cross-validation; thus, the data were split five ways providing 12 observations for calibration, leaving three observations for model evaluation for each fold. Table 5 shows the evaluation metrics obtained in the present study for each of the five folds used for cross-validation.

The results in Table 6 show that for four out of the five folds, all evaluation metrics indicated a satisfactory prediction in accordance with the specifications of Moriasi and others (2007). However, for the third fold, evaluated with

Table 2. *RSR* of different model configurations, optimal value of $RSR = 0$

	Number of variables				
	1	2	3	4	5
$r^2_{threshold}$					
0.2	1.04	0.72	0.69	0.69	0.75
0.25	1.01	0.67	0.66	0.68	0.72
0.3	0.88	0.56	0.53	0.53	0.58
0.35	0.87	0.62	0.60	0.59	0.75

Table 4. Number of models in the ensemble of plausible models with different configurations of number of input variables and threshold r^2 value

	Number of variables				
	1	2	3	4	5
$r^2_{threshold}$					
0.2	15	105	455	1365	3003
0.25	11	55	165	330	462
0.3	7	21	35	35	21
0.35	5	10	10	5	1

Table 5. Final variables selected for model development and their correlation to the observed summer mass balance of Brúarjökull, given as r^2 values

Variable	Location	r^2
Net radiation	BruNe (850 m a.s.l)	0.65
AMO index	Atlantic Ocean	0.48
Albedo	BruNe (850 m a.s.l)	0.47
Albedo	BruMi (1200 m a.s.l)	0.36
Atmospheric pressure	Karahnjukar	0.35
Precipitation	Egilsstadir	0.33
Ocean heat content	North Atlantic	0.32

Table 6. Evaluation metrics for model averaged predictions using fivefold cross-validation

Evaluation period	NSE	RSR	PBIAS (%)
2001–2003	0.71	0.53	4.9
2004–2006	0.66	0.57	10.7
2007–2009	−0.72	1.3	−3.2
2010–2012	0.75	0.49	−6.1
2013–2015	0.87	0.35	−8.7

Models show satisfactory performance with NSE > 0.5, RSR < 0.7 and PBIAS ⩽ ±25%.

observations from the period 2007–2009, low *NSE* and high *RSR* values were observed, whereas PBIAS was within acceptable range.

Figure 3 shows the observed mass balance of the Brúarjökull for the study period with predicted values from each fold in a box and whiskers plot. The observed summer mass balance is shown as black stars; the notch in the box represents the median of the ensemble predictions, while the ends of the box represent the upper and lower quartiles; the whiskers encompass the range of all ensemble predictions. Considering the time series of simulated and observed values shown in Figure 3 during the period 2007–2009, both were very close to the long-term average mass balance of the glacier. When the observed values were close to the mean value, the denominator in Eqns (1) and (2) took on a small value, inflating the evaluation metrics, *NSE* and *RSR*. Thus, during periods where the mass balance is consistently close to the mean, these metrics may provide a poor indication of the quality of the model outputs.

The results in Figure 3 show that the range of ensemble predictions encompasses the observed values for all observations in the study period except 2004 and 2012. Furthermore, the ensemble mean provides a reasonable estimate of the observed mass balance for the range of observations considered. The evaluation metrics described in Eqns (1)–(3) were calculated for all predictions yielding values of *NSE* = 0.71, *RSR* = 0.54 and PBIAS = 0.27%. The results suggest that the model has satisfactory performance with low bias for under- or overpredicting the mass balance.

4.4. Model performance in time

The results described in the previous sections were obtained using data available on 1 July. To assess how the predictive performance of the method evolved over time in spring, as the melt season approached predictions were calculated for data available for 15 May, 1 June and 15 June. Table 6 shows the evaluation metrics observed for predictions at each of these dates compared with 1 July.

The evaluation metrics presented in Table 7 show that the predictive information in the data diminished quickly as the lead time increased to dates prior to 1 July. For predictions made on 15 May, negative values of *NSE* were observed, indicating that the model predictions performed worse than simply reporting the mean mass balance of the glacier.

Figure 4 shows the predictions from the model suite using data from the four dates in spring. The results show that for longer lead times, the spread in ensemble predictions increased. Especially, ensembles for the years 2008 and 2009 showed a large variability in model outputs. The model also had poorer performance in predicting the extremes of the observations with longer lead times. The results show that predictions made prior to 1 July were less reliable and the earliest time when satisfactory predictions could be made was between 15 June and 1 July.

Figure 5 shows the correlation of the selected input variables shown in Table 5 to the Brúarjökull summer mass balance at the four forecast dates in spring. The table shows that the key predictor on 1 July forecast, net radiation at BruNe, has much less predictive power earlier in the spring. The same applies to the albedo at both BruNe and BruMi, which show
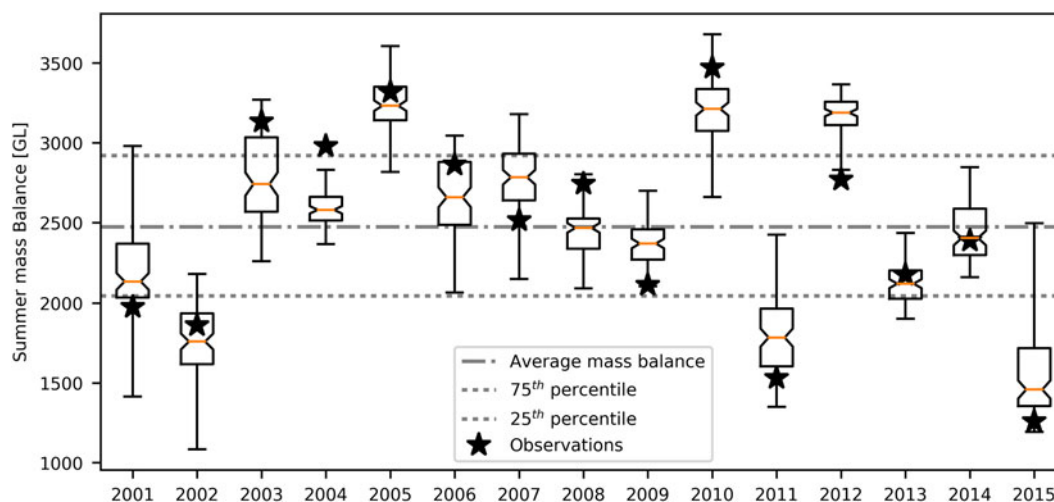

Fig. 3. Model averaged predictions of Brúarjökull summer mass balance for all fivefold cross-validations. Observed glacier mass balance is shown as black stars.

Table 7. Evaluation metrics for predictions with longer lead times with models showing satisfactory performance on 1 July

Forecast date	NSE	RSR	PBIAS (%)
15 May	−0.95	1.39	9.2
1 June	0.13	0.93	3.2
15 June	0.45	0.75	2.7
1 July	0.71	0.54	0.27

Satisfactory performance is defined as models having $NSE > 0.5$, $RSR < 0.7$ and $PBIAS \leqslant \pm25\%$.

some predictive power on 15 June but none earlier. This suggests that spring snow conditions on the glacier are not indicative of the summer melt. By the end of June, as the melt season is beginning on the glacier, these variables start showing the power to predict the summer melt patterns.

The precipitation at Egilsstadir and atmospheric pressure at Karahnjukar similarly show little to no correlation to the summer mass balance on the earliest forecast dates. This suggests that the precipitation patterns in late spring and early summer play an important role in determining the summer melt, whereas precipitation patterns during the winter are less important in determining the ultimate summer melt. This can also be deduced from the fact that none of the winter accumulation measurements showed correlation to the observed summer mass balance.

The AMO index and the North Atlantic Ocean Heat content, however, show a persistent correlation to the summer mass balance throughout all the forecast dates. This suggests that a large portion of the inter-annual variability in Icelandic glacier mass balance is affected by the large-scale oceanic circulations and heat transport in the North Atlantic Ocean. The trends in these variables persist in much longer time frames than local climate conditions and contain significant predictive power for glacial mass-balance forecasts at least as early as the end of the first annual quarter. These results suggest that a significant portion of the variability in summer mass balance of Icelandic glaciers can be forecast well in advance of the melt season with lead times up to the length of the autocorrelative time scales of the AMO index.
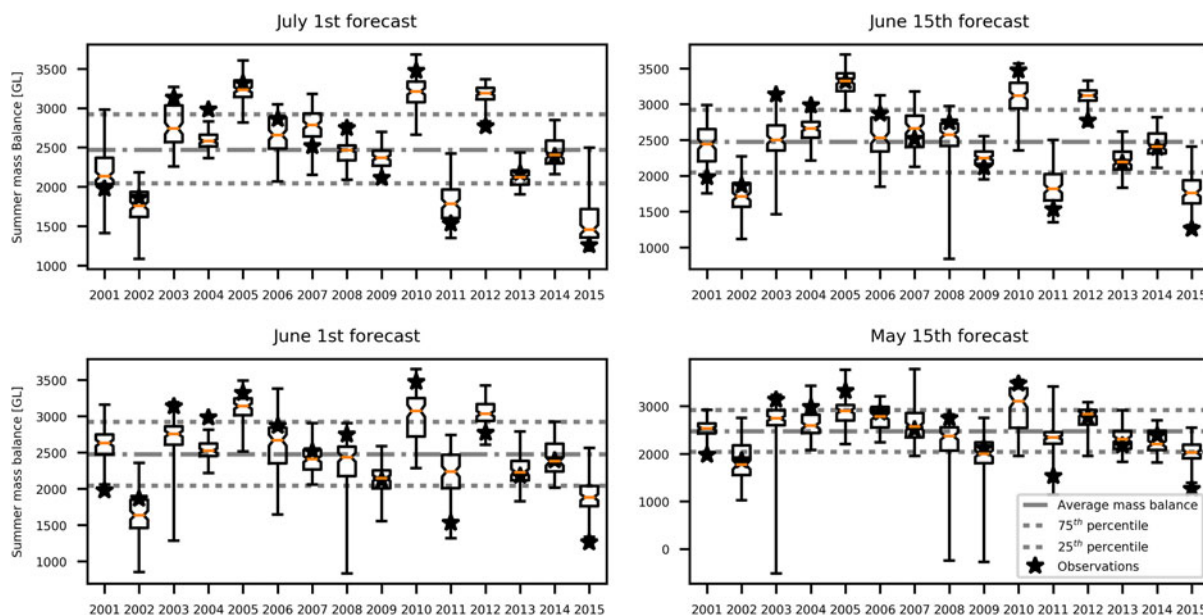

Fig. 4. Model averaged predictions of Brúarjökull summer mass balance for all fivefold cross-validations. The optimal model found with 1 July data was forced with earlier data at three different dates between 15 May and 15 June.
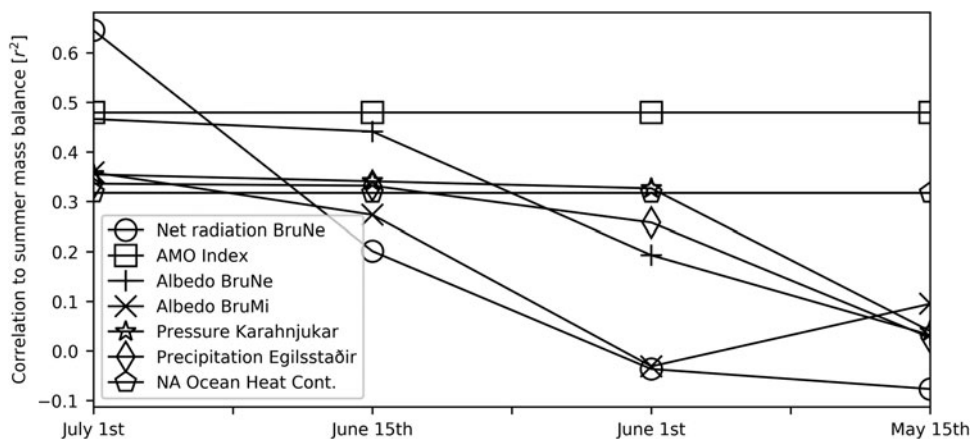

Fig. 5. Correlation of the selected predictor variables to Brúarjökull summer mass balance on four forecast dates in Spring.

Lastly, we note that the NAOI did not show any correlation to the summer mass balance. This suggests that large-scale atmospheric circulation in the North Atlantic, while indicative of Icelandic climate, is not an important factor in summer melt patterns of Icelandic glaciers. Hence, in terms of glacial mass balance, the ocean circulation in the North Atlantic is a much more important variable than atmospheric circulation.

5. CONCLUSION

The study showed that, of all the potential input variables available in the basin, seven showed a significant correlation with the summer mass balance. The variables deemed to contain predictive information at the beginning of the melt season were associated with average net radiation, glacier albedo, precipitation, atmospheric pressure and heat flux in the North Atlantic. It was observed that out of all potential multivariate regression models incorporating these variables, only a few adequately predicted summer mass balance. As selection of any single model would cause additional uncertainty in the estimation of the response variable due to model selection, an ensemble of plausible multivariate regression models was calculated and the average of the model results was used to predict the glacier mass balance.

The selection of a subset of plausible models over which to average was investigated. The results suggest that the optimal subset was found by eliminating models with poor fit to calibration data. Sensitivity analysis of model predictions suggested that the optimal number of input variables to include in the models was three and with variables exhibiting significant correlation included as inputs. The results showed that, in terms of the model evaluation measures *NSE*, *RSR* and PBIAS, satisfactory predictions of summer mass balance could be made by calculating a uniform average of model forecasts over the set of plausible models.

Investigation of the lead time with which predictions are calculated showed that model performance becomes less reliable as simulations are performed earlier in spring. Satisfactory predictions can be produced between 15 June and 1 July, at which time the melt season is beginning and predictions of summer melt volumes are valuable to water resources managers.

ACKNOWLEDGEMENTS

REFERENCES

Aðalgeirsdóttir G, Johannesson T, Bjornsson H, Palsson F and Sigurdsson O (2006) Response of Hofsjokull and southern Vatnajokull, Iceland, to climate change. *J. Geophys. Res. Earth Surf.*, **111**, F03001 (doi: 10.1029/2005jf000388)

Aðalgeirsdóttir G and 6 others (2011) Modelling the 20th and 21st century evolution of Hoffellsjokull glacier, SE-Vatnajokull, Iceland. *Cryosphere* **5**, 961–975 (doi: 10.5194/tc-5–961–2011)

Baldwin MP and 5 others (2003) Stratospheric memory and skill of extended-range weather forecasts. Science *(New York, N.Y.)*, **301**(5633), 636–640 (doi: 10.1126/science.1087143)

Barnett TP, Adam JC and Lettenmaier DP (2005) Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438**(7066), 303–309 (doi: 10.1038/nature04141)

Bjornsson H and Palsson F (2008) Icelandic glaciers. *Jökull*, **58**(58), 365–386.

Bjornsson H and 6 others (2013) Contribution of Icelandic ice caps to sea level rise: trends and variability since the Little Ice Age. *Geophys. Res. Lett.*, **40**(8), 1546–1550 (doi: 10.1002/grl.50278)

Buckland ST, Burnham KP and Augustin NH (1997) Model selection: an integral part of inference. *Biometrics*, **53**(2), 603 (doi: 10.2307/2533961)

Burnham KP and Anderson DR (2002) *Model selection and multimodel inference*, Springer, New York (doi: 10.1007/978-3-319-02868-2_3)

Carenzo M, Pellicciotti F, Rimkus S and Burlando P (2009) Assessing the transferability and robustness of an enhanced temperature-index glacier-melt model. *J. Glaciol.*, **55**(190), 258–274 (doi: 10.3189/002214309788608804)

De Ruyter de Wildt M, Klok E and Oerlemans J (2003a) Reconstruction of the mean specific mass-balance of Vatnajokull (Iceland) with a seasonal sensitivity characteristic. *Geogr. Ann. Ser. A Phys. Geogr.*, **85**(1), 57–72. (doi: 10.1111/1468-0459.00189)

De Ruyter de Wildt M, Oerlemans J and Bjornsson H (2003b) A calibrated mass-balance model for Vatnajökull, Iceland. *Jökull*, **52**, 1–20.

Diks CGH and Vrugt JA (2010) Comparison of point forecast accuracy of model averaging methods in hydrologic applications. *Stoch. Environ. Res. Risk Assess.*, **24**(6), 809–820 (doi: 10.1007/s00477-010-0378-z)

Drolon V, Maisongrande P, Berthier E, Swinnen E and Huss M (2016) Monitoring of seasonal glacier mass-balance over the European Alps using low-resolution optical satellite images. *J. Glaciol.*, **62**(235), 912–927 (doi: 10.1017/jog.2016.78)

Enfield DB, Mestas-Nuñez AM and Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.*, **28**(10), 2077–2080 (doi: 10.1029/2000GL012745)

Engelhardt M, Schuler TV and Andreassen LM (2014) Contribution of snow and glacier melt to discharge for highly glacierised catchments in Norway. *Hydrol. Earth Syst. Sci.*, **18**(2), 511–523 (doi: 10.5194/hess-18-511-2014)

Fujita K and Ageta Y (2000) Effect of summer accumulation on glacier mass-balance on the Tibetan Plateau revealed by mass-balance model. *J. Glaciol.*, **46**(153), 244–252 (https://doi.org/10.3189/172756500781832945)

Gardarsson SM and Eliasson J (2006) Influence of climate warming on Halslon reservoir sediment filling. *Nord. Hydrol.*, **26**, 553–569 (doi: 10.2166/nh.2006.014)

Ghosh R, Müller WA, Baehr J and Bader J (2017) Impact of observed North Atlantic multidecadal variations to European summer climate: a linear baroclinic response to surface heating. *Clim. Dyn.*, **48**(11–12), 3547–3563 (doi: 10.1007/s00382-016-3283-4)

Gudmundsson S and 6 others (2011) Response of Eyjafjallajokull, Torfajokull and Tindfjallajokull ice caps in Iceland to regional warming, deduced by remote sensing. *Polar Res.*, **30** (doi: 10.3402/Polar.V30i0.7282)

Hanna E, Jónsson T and Box JE (2001) Recent changes in Icelandic climate. *Spring*, **61**, 3–9 (doi: 10.1256/wea.80.04)

Hanna E, Jónsson T and Box JE (2004) An analysis of Icelandic climate since the nineteenth century. *Int. J. Climatol.*, **24**(10), 1193–1210 (doi: 10.1002/joc.1051)

Hjort NL and Claeskens G (2003) Frequentist model average estimators. *J. Am. Stat. Assoc.*, **98**(464), 879–899 (doi: 10.1198/016214503000000828)

Hoeting JA, Madigan D, Raftery AE and Volinsky CT (1999) Bayesian model averaging: a tutorial. *Stat. Sci.*, **14**(4), 382–417 (doi: 10.2307/2676803)

James G, Witten D, Tibshirani R and Hastie T (2013) *An introduction to statistical learning with applications in R*, Springer, New York (doi: 10.1007/978-1-4614-7138-7)

Jonsdóttir JF (2010) A runoff map based on numerically simulated precipitation and a projection of future runoff in Iceland. *Hydrol. Sci. J.*, **53**, 100–111 (doi: 10.1623/hysj.53.1.100)

Kalra A, Ahmad S and Nayak A (2013) Increasing streamflow forecast lead time for snowmelt-driven catchment based on large-scale climate patterns. *Adv. Water Resour.*, **53**, 150–162 (doi: 10.1016/j.advwatres.2012.11.003)

Kjær KH, Korsgaard NJ and Schomacker A (2008) Impact of multiple glacier surges – a geomorphological map from Bruarjokull, East Iceland. *J. Maps*, **4**(1), 5–20 (doi: 10.4113/jom.2008.91)

Larsen G (1998) Eight centuries of periodic volcanism at the center of the Iceland hotspot revealed by glacier tephrostratigraphy. *Geology*, **26**(10), 943–946 (doi: 10.1130/0091-7613(1998)026<0943:ECOPVA>2.3.CO;2)

Liang H, Zou G, Wan ATK and Zhang X (2011) Optimal weight choice for frequentist model average estimators. *J. Am. Stat. Assoc.*, **106**(495), 1053–1066 (doi: 10.1198/jasa.2011.tm09478)

Liu W and 5 others (2015) Impacts of climate change on hydrological processes in the Tibetan Plateau: a case study in the Lhasa River basin. *Stoch. Environ. Res. Risk Assess.*, **29**(7), 1809–1822 (doi: 10.1007/s00477-015-1066-9)

Madigan D and Raftery AE (1994) Model selection and accounting for model uncertainty in graphical models using Occam's Window. *J. Am. Stat. Assoc.*, **89**(428), 1535–1546 (doi: 10.1080/01621459.1994.10476894)

Marshall SJ, Bjornsson H, Flowers GE and Clarke GKC (2005) Simulation of Vatnajokull ice cap dynamics. *J. Geophys. Res. Earth Surf.*, **110**, F03009 (doi: 10.1029/2004JF000262)

Matthews T, Hodgkins R, Gudmundsson S, Palsson F and Bjornsson H (2015) Inter-decadal variability in potential glacier surface melt energy at Vestari Hagafellsjokull (Langjokull, Iceland) and the role of synoptic circulation. *Int. J. Climatol.*, **35**(10), 3041–3057 (doi: https://doi.org/10.1002/joc.4191)

Moller R and 7 others (2014) MODIS-derived albedo changes of Vatnajokull (Iceland) due to tephradeposition from the 2004 Grimsvotn eruption. *Int. J. Appl. Earth Obs. Geoinf.*, **26**(1), 256–269 (doi: 10.1016/j.jag.2013.08.005)

Moriasi DN and 5 others (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, **50**(3), 885–900 (doi: 10.13031/2013.23153)

Oladottir BA, Larsen G and Sigmarsson O (2011) Holocene volcanic activity at Grimsvotn, Bardarbunga and Kverkfjoll subglacial centres beneath Vatnajokull, Iceland. *Bull. Volcanol.*, **73**(9), 1187–1208 (doi: 10.1007/S00445-011-0461-4)

Palsson F and 5 others (2014) *Vatnajokull: mass-balance, melt water drainage and surface velocity of the glacial year 2013–14*. Landsvirkjun. Report No LV-2014-138 (Retrieved 01/05/2017, http://gogn.lv.is/files/2014/2014-138.pdf)

Palter JB (2015) The role of the Gulf Stream in European climate. *Ann. Rev. Mar. Sci.*, **7**(1), 113–137 (doi: 10.1146/annurev-marine-010814-015656)

Qiu L, You J, Qiao F and Peng D (2014) Simulation of snowmelt runoff in ungauged basins based on MODIS: a case study in the Lhasa River basin. *Stoch. Environ. Res. Risk Assess.*, **28**(6), 1577–1585 (doi: 10.1007/s00477-013-0837-4)

Raftery AE, Madigan D and Hoeting JA (1997) Bayesian model averaging for linear regression models. *J. Am. Stat. Assoc.*, **92**(437), 179–191 (doi: 10.1080/01621459.1997.10473615)

Rasmussen LA (2005) Mass-balance of Vatnajokull reconstructed back to 1958. *Jokull*, **55**, 139–146

Réveillet M, Vincent C, Six D and Rabatel A (2017) Which empirical model is best suited to simulate glacier mass-balances? *J. Glaciol.*, **63**(237), 39–54 (doi: 10.1017/jog.2016.110)

Roe GH, Baker MB and Herla F (2017) Centennial glacier retreat as categorical evidence of regional climate change. *Nat. Geosci.*, **10**(2), 95–99 (doi: 10.1038/ngeo2863)

Schöner W and Böhm R (2007) A statistical mass-balance model for reconstruction of LIA ice mass for glaciers in the European Alps. *Ann. Glaciol.*, **44**, 161–169 (doi: 10.3189/172756407782871639)

Thorsteinsson T, Einarsson B and Kjartansson V (2004) *Afkoma Hofsjokuls 1997–2004*. Icelandic National Energy Authority, Reykjavik, Report NO OS-2004/029

Tsai FTC (2010) Bayesian model averaging assessment on ground-water management under model structure uncertainty. *Stoch. Environ. Res. Risk Assess.*, **24**(6), 845–861 (doi: 10.1007/s00477-010-0382-3)

Wang H, Zhang X and Zou G (2009) Frequentist model averaging estimation: a review. *J. Syst. Sci. Complex.*, **22**, 732–748 (doi: 10.1007/s11424-009-9198-y)

Zampieri M, Toreti A, Schindler A, Scoccimarro E and Gualdi S (2017) Atlantic multi-decadal oscillation influence on weather regimes over Europe and the Mediterranean in spring and summer. *Glob. Planet. Change*, **151**, 92–100 (doi: 10.1016/j.gloplacha.2016.08.014)

Zemp M and 38 others (2015) Historically unprecedented global glacier decline in the early 21st century. *J. Glaciol.*, **61**(228), 745–762 (doi: 10.3189/2015JoG15J017)

Zhang X, Wan ATK and Zhou SZ (2012) Focused information criteria, model selection, and model averaging in a tobit model with a nonzero threshold. *J. Bus. Econ. Stat.*, **30**(1), 132–142 (doi: 10.1198/jbes.2011.10075)

APPENDIX A

See Table A1.

Table A1. Breakdown of the potential input variables surveyed along with their correlation to Brúarjökull summer mass balance

Variable	Location	Period	R^2_{adjusted}	Variable	Location	Period	R^2_{adjusted}
Net radiation	BruNe	Q2	0.645	Relative humidity	Hvalsnes	Q2	0.050
AMO	North Atlantic	Past month	0.479	Dew point temperature	Eyjabakkar	Q1	0.045
Albedo	BruNe	Q2	0.466	Relative humidity	Myvatn	Q1	0.026
Albedo	BruMi	Q2	0.359	Outgoing longwave radiation	BruNe	Q2	0.014
Atmospheric pressure	Karahnjukar	Q2	0.355	Winter accumulation	BB0	Winter	0.013
Precipitation	Egilsstaðir	Q2	0.336	Winter accumulation	B10	Winter	0.008

Table A1. (*Cont.*)

Variable	Location	Period	$R^2_{adjusted}$	Variable	Location	Period	$R^2_{adjusted}$
Heat content	North Atlantic Ocean	Past month	0.318	Relative humidity	BruMi	Q1	0.007
Reflected solar radiation	BruNe	Q2	0.298	Wind speed	Myvatn	Q1	0.004
Wind speed	Myvatn	Q2	0.291	Wind speed	Upptyppingar	Q1	0.000
Winter accumulation	B12	Winter	0.261	Temperature	Skaftafell	Q2	−0.003
Atmospheric pressure	Eyjabakkar	Q2	0.259	Outgoing longwave radiation	BruMi	Q2	−0.004
Atmospheric pressure	Egilsstaðir	Q2	0.256	Wind speed	Karahnjukar	Q1	−0.004
Wind speed	Egilsstaðir	Q2	0.232	Vapor pressure	Karahnjukar	Q2	−0.008
Temperature	Eyjabakkar	Q2	0.227	Temperature	Egilsstaðir	Q2	−0.017
Winter accumulation	B19	Winter	0.218	Positive degree days	BruNe	Q1	−0.021
Wind direction	BruNe	Q2	0.200	Atmospheric pressure	Laufbali	Q2	−0.023
Atmospheric pressure	Egilsstaðir	Q1	0.195	Negative degree days	BruNe	Q1	−0.024
Wind speed	Hvalsnes	Q2	0.192	Temperature	BruNe	Q1	−0.027
Winter accumulation	B16	Winter	0.191	Temperature	Egilsstaðir	Q1	−0.030
Relative humidity	Jökulheimar	Q2	0.191	Atmospheric pressure	Teighorn	Q2	−0.030
Incoming longwave radiation	BruNe	Q2	0.187	Wind speed	BruNe	Q1	−0.030
Atmospheric pressure	Eyjabakkar	Q1	0.182	Dew point temperature	Karahnjukar	Q2	−0.033
Temperature	Karahnjukar	Q2	0.182	Relative humidity	Eyjabakkar	Q1	−0.033
Sea temperature anomaly	North Atlantic	Past month	0.180	Wind direction	Upptyppingar	Q1	−0.034
Albedo	Brune	Q1	0.177	Wind speed	Eyjabakkar	Q1	−0.035
Relative humidity	Jökulheimar	Q2	0.167	Wind speed	Setur	Q2	−0.037
Wind speed	Laufbali	Q2	0.158	Temperature	Teighorn	Q2	−0.044
Wind speed	Skarðsfjöruviti	Q2	0.157	Relative humidity	Eyjabakkar	Q2	−0.046
Winter accumulation	B13	Winter	0.157	Temperature	Laufbali	Q2	−0.046
Relative humidity	BruNe	Q1	0.155	Winter accumulation	B09	Winter	−0.047
Reflected solar radiation	BruMi	Q2	0.152	Temperature	BruMi	Q2	−0.049
Winter accumulation	Bru	Winter	0.151	Wind direction	BruNe	Q1	−0.051
Relative humidity	Karahnjukar	Q2	0.146	Incoming solar radiation	BruMi	Q2	−0.051
Wind speed	Skaftafell	Q2	0.143	Wind direction	Karahnjukar	Q2	−0.055
Temperature	Eyjabakar	Q1	0.139	Precipitation	Egilsstaðir	Q1	−0.055
Wind speed	BruMi	Q2	0.139	Atmospheric pressure	Jökulheimar	Q2	−0.056
Temperature	Setur	Q2	0.135	Temperature	Skarðsfjöruviti	Q2	−0.056
Incoming longwave radiation	BruNe	Q1	0.133	Wind direction	Upptyppingar	Q2	−0.058
Wind speed	Egilsstaðir	Q1	0.126	Incoming longwave radiation	BruMi	Q2	−0.060
Winter accumulation	BUD	Winter	0.124	Wind speed	BruNe	Q2	−0.060
Positive degree days	BruMi	Q2	0.117	Wind direction	Karahnjukar	Q1	−0.064
Reflected solar radiation	BruNe	Q1	0.116	Atmospheric pressure	Setur	Q2	−0.065
Wind speed	Upptyppingar	Q2	0.116	Incoming solar radiation	BruNe	Q2	−0.066
Wind speed	Eyjabakkar	Q2	0.110	Wind direction	Eyjabakkar	Q2	−0.068
Incoming solar radiation	BruNe	Q1	0.107	Relative humidity	Myvatn	Q2	−0.068
Temperature	Jökulheimar	Q2	0.107	Wind direction	Eyjabakkar	Q1	−0.068
Winter accumulation	B18	Winter	0.106	Negative degree days	BruNe	Q2	−0.069
Net radiation	BruMi	Q2	0.103	NAOI	North Atlantic	Past month	−0.069
Wind speed	Jökulheimar	Q2	0.102	Negative degree days	BruMi	Q2	−0.071
Relative humidity	BruNe	Q2	0.098	Wind Speed	Teighorn	Q2	−0.072
Atmospheric pressure	Karahnjukar	Q1	0.097	Wind direction	BruMi	Q2	−0.072
Wind speed	Karahnjukar	Q2	0.097	Relative humidity	Skaftafell	Q2	−0.074
Dew point temperature	Karahnjukar	Q1	0.088	Temperature	Hvalsnes	Q2	−0.075
Temperature	Myvatn	Q2	0.080	Temperature	BruNe	Q2	−0.075
Temperature	Myvatn	Q1	0.078	Vapor pressure	Eyjabakkar	Q2	−0.075
Temperature	Karahnjukar	Q1	0.070	Relative humidity	Karahnjukar	Q1	−0.077
Outgoing longwave radiation	BruNe	Q1	0.069	Relative humidity	Egilsstaðir	Q1	−0.077
Winter accumulation	B14	Winter	0.067	Dew point temperature	Eyjabakkar	Q2	−0.077
Vapor pressure	Eyjabakkar	Q1	0.067	Positive degree days	BruNe	Q2	−0.082
Relative humidity	Egilsstaðir	Q2	0.054	Net radiation	BruNe	Q1	−0.090