Indecis

Sectorial Climate Services

Integrated approach for the **d**evelopment across **E**urope of user oriented **c**limate **i**ndicators for GFCS high-priority **s**ectors: Agriculture, disaster risk reduction, energy, health, water and tourism

Work Package 6

Deliverable 6.3

Report on the reliability and uncertainties associated with the (hindcast-type) seasonal forecasts of selected sectorial INDECIS indices

European Research Area for Climate Services

JPI Climate

This report arises from the Project INDECIS which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by FORMAS (SE), DLR (DE), BMWFW (AT), IFD (DK), MINECO (ES), ANR (FR), with co-funding by the European Union's Horizon 2020 research and innovation programme

Applications of seasonal forecasting for agriculture, winter tourism and hydropower

(Contribution from FMI)

Authors:

Andrea Vajda – Finnish Meterological Institute

Otto Hyvärinen – Finnish Meterological Institute

Tiina Ervasti – Finnish Meterological Institute

Alexey Karpechko – Finnish Meterological Institute

Jaakko Ikonen – Finnish Meterological Institute

Table of Contents

Applications of seasonal forecasting in agriculture and winter tourism sector

(Andrea Vajda, Otto Hyvärinen, Tiina Ervasti)

1. Background and motivation

Seasonal climate forecast products are needed for decision making in various sectors, especially in those sectors that are widely affected by climate variability and change, such as agriculture and winter tourism. Due to variability in climate, inter-annual variability of crop yields has increased notably. For example, Finland experienced extreme cold summer in 2017 followed by an extreme hot and dry summer in 2018, both resulting in a decrease by over 25% in crop yield, registering in 2018 the smallest harvest from the last 26 years. Similarly, the ski industry is also facing an increased vulnerability to climate change and variability. The increased uncertainty of snow conditions during the ski season, especially in southern and central regions of Finland; the late start/early end of snowing season and the difficulties in artificial snow production due to high winter temperatures have significant effects on the winter tourism. Seasonal climate forecasts have a large potential to increase preparedness to variability, to manage the business risk and optimize resources. Nevertheless, the uptake of seasonal forecasts for decision making in Europe has been relatively limited (Bruno Soares and Dessai, 2016). This is partly due to the limited skill of forecast systems in this region (Doblas-Reyes et al., 2013; Bruno Soares, 2017) and because users are not aware of the availability of such forecasts. Recent advances in seasonal forecasts resulted in useful predictions that can be beneficial for stakeholders. The applicability of seasonal forecast outputs in agriculture and ski industry has not yet been studied in Finland.

In this project we aimed to assess the applicability of seasonal forecast outputs for the above-mentioned sectoral application in Finland and to develop and pilot a set of seasonal climate indices both with farmers and ski resort managers. The indices developed for agriculture were tested with farmers during two pilot seasons: May-October 2019 and April-October 2020 (still on-going), ski resorts managers tested the seasonal climate indices developed during November 2019-April 2020.

2. Data and methods

2.1 Engaging the users in the seasonal climate service development

One of the core requirements of climate services development is the interaction between providers and users (Buontempo et al. 2014) that implies users' involvement in the co-design and co-development of services. The seasonal forecast products for agriculture and ski resorts was developed in interaction with stakeholders from both sectors through a range of activities, such as meetings during the design and development process, interaction through mails during the testing phase and feedback surveys during the evaluation.

In case of stakeholders from agriculture we have applied a mediate approach, being in direct contact with the Central Union of Agricultural Producers and Forest Owners (MTK) from Finland from the beginning of the process. The Union was engaged through dialogue in the design and development process of the tailored seasonal climate indices in order to harmonize the needs of the farmers in terms of content, visualization and delivery form of forecasts. Furthermore, the Union has involved over 200 farmers into the two pilot stages from various parts of Finland, predominantly from the southern, western and south-western part of the country. The indices were selected together with the Union and the developed seasonal climate outlooks were delivered to the farmers by mail through the Union for testing. To learn about farmers' opinion about the level of usability of the developed seasonal forecast indices and climate outlooks feedback surveys were conducted with them at the end of pilot season 2019 and will be done also at the end of season 2020.

In the case of ski industry pilot, we've been in direct interaction with the managers of ski resorts. We aimed to engage users from various parts of Finland as ski resorts located in Lapland experience different challenges in terms of climate variability than those from the southern or central part of Finland. Six representative ski resorts were identified by the Finnish Ski Association and involved in the pilot: Ruka-Pyhä, Saariselkä and Salla located in Lapland, Hakarinteet and Sappee from Central Finland and Swinghill-Espoo from southern Finland. A workshop was organized with the representatives of the ski resorts in the beginning of the pilot, where the most useful indices were selected, uncertainties related to variables used in the development of indices were presented to the users and the visualization and delivery of climate outlooks was agreed. The selected indices were designed in collaboration with the stakeholders using as input the aspects and criteria applied in the maintenance practices, for ex. conditions for the production of artificial snow or hazard of severe wind speed. The developed seasonal climate outlooks were tested during winter 2019-2020; following the test period a feedback survey was conducted with the users.

2.2 Datasets

The forecast data used in the development of seasonal forecast indices were provided by the SEAS5 seasonal forecast system of ECMWF (Johnson et al. 2019). The variables used were the 2 m temperature, total precipitation, dew temperature, snow depth, snow density and 10 m wind speed. In the first phase of the development and during the pilot from summer 2019, the quality of 2 m temperature and total precipitation forecast was assessed using re-forecast data for the period 1993-2016 accessed from the C3S Copernicus Data Store (CDS), available at 1° spatial resolution for 25 ensemble members. The reason for using the rough spatial resolution but easily accessible CDS re-forecasts was the technical setbacks faced when downloading the reforecast data through the Meteorological Archival and Retrieval System (MARS) (excessively long-lasting and repeatedly interrupted process) during winter-spring 2019. Starting November 2019, the calibration of the parameters, except wind speed, was done using the re-forecast data for the period 1981-2016 spatial resolution accessed through MARS, available at 0.25°. The parameters were evaluated and calibrated using as reference data the ERA5 reanalysis (Hersbach, 2020) for the same period. The reanalysis data was interpolated to the same resolution as reforecast data.

In addition, observational gridded (10x10 km2) growing degree days (GDD) values available from FMI database from 2003 onward were used both in the computation and verification of seasonal GDD indices. Observational wind speed data from three stations located in Lapland, i.e. Inari Saariselkä Kaunispää (68° 26', 27° 26'), Pelkosenniemi Pyhätunturi (67° 1', 27° 13') and Kuusamo Rukatunturi (66° 9', 29° 9') were used in the calibration and computation of wind speed forecasts.

The real time forecasts of seasonal forecast indices were computed using the 0.25° spatial resolution data available for the 51 ensemble members from SEAS5 forecast system, accessed through MARS. Forecasts have been produced for three months ahead, meaning three monthly forecasts, the start month being called lead month (LM) 0, followed by lead month 1 and 2.

2.3 Methods of skill assessment and bias adjustment

To reduce the substantial systematic biases from raw model outputs and produce useful information for sectoral applications, a bias adjustment process was performed for most of the variables. The skill assessment and bias adjustment of model data was performed using the open source R package climate4R (Iturbide et al. 2019). The input data used in the analysis was initially the re-forecasts accessed from the C3S CDS, during autumn all the analyses were updated using the re-forecasts from MARS. The ensemble re-forecasts monthly averages of daily mean temperature, dew temperature, snow depth, maximum wind speed and soil moisture, and monthly sum of daily precipitation were calculated for each grid point and evaluated. In addition, from the seasonal climate indices developed, the skill of growing degree days was also assessed. Ensemble re-forecasts daily mean temperature were used in the calculation of GDD for skill assessment, bias adjustment was also performed on daily data. Several bias adjustment methods were tested for each variable (Table 1), such as variance, scaling, empirical quantile mapping (EQM), parametric quantile mapping (PQM) and power transformation of precipitation (PT) built in the climate4R package, and the ensemble model output statistics (EMOS). In ensemble model output statistics (EMOS) (Gneiting et al 2005), the monthly values of variables was modelled using truncated Gaussian distribution, the mean of this distribution was modelled using the ensemble mean as a predictor and the variance was modelled using a simple constant. For the variance, the ensemble variance of variables as a predictor was also tried, but this turned out to be not a statistically significant predictor and did not improve the results. Other predictors might be tested in future studies. The implementation of Messner et al. 2016 was used.

The raw forecast ensemble and the bias adjusted ensemble data for the three lead months was evaluated against ERA5 reanalysis data using reliability diagrams (e.g. Weisheimer and Palmer 2014) of aggregated grid points of Finnish land areas and maps of verification measures, such as correlation, continuous ranked probability skill score (CRPSS) and mean error. To avoid overfitting we used odd years for fitting and only even years for validation. Verification results are shown for those months (February, May, August and November) that have the same 51 number of reforecast members as the operational forecasts.

Table 1. The bias adjustment methods used in the skill assessment of the variables and seasonal forecast indices. In case of total precipitation the months the bias adjustment method is applied for are listed in brackets, for the other variables the analyses are done for the whole year.

Variables and indices	**Raw model data**	**Bias adjustment methods**						**BA method selected**
		Variance	Scaling	EQM	PQM	PT	EMOS	
2 m temperature	x	x		x			x	EQM / variance
Total precipitation	x	x (V-VIII)	x (I-XII)	x (I-XII)	x (V-VIII)	x (I-XII)	x (I-XII)	No BA applied
Snow depth	x	x		x			x	EMOS
Soil moisture	x	x		x			x	No BA applied
Wind speed	x							Quantile mapping (obs based)
Growing degree days	x	x		x				EQM

3. Forecast skill assessment of seasonal forecast variables for Finland

Verification results for mean temperature, total precipitation, soil moisture, snow depth, wind and the growing degree days (GDD) index are presented hereinafter. The analyses are based on the high resolution re-forecasts data for 1981-2016 accessed from MARS except for the GDD index that is based on re-forecasts accessed from C3S CDS and observations from FMI archives for the period 2003-2016.

3.1 Temperature

The skill of raw model data and bias corrected data with the EQM, scaling, variance and EMOS method was analyzed for 2 m temperature. Reliability diagrams for tercile forecasts indicate that all terciles are useable, varying between perfect and marginally useful for lead month 0 for most of the months. Lower reliabilities are indicated for May (Fig. 1), for which the forecast are useless except the upper tercile. Results for LM1 and 2 diverge: lower and upper terciles are marginally useful for August for LM1 and even February and May for LM 2, for the other months forecasts are unusable.

The continuously ranked probability skill score (CRPSS) for bias corrected data with the variance method for each grid point is shown in Figure 2. For lead month 0 large skill score values that are statistically significant were found over the whole study area for initializations in Feb and Nov. For the other initialization times and lead times CRPSS values are lower. In the case of LM1, skill scores are lower, somewhat better in February and August but near zero in spring. All the bias correction methods tested corrected the raw model biases, improving the raw data in 60-80% of the grid points depending on lead time and months. Results for all the methods for February using CRPSS are shown in Figure 3. Analyses show that all the methods corrected effectively the biases, giving added value to temperature data. Variance method has outperformed the other corrections.

Similar improvement is seen in the variation of mean error (not shown), which was significantly reduced by the bias adjustment methods for each lead time and initialization. The forecast model systematically underestimated the observed temperature values except for Aug, Sep and Oct, when error values were also the lowest.

Based on the skill assessment results we decided to correct the raw temperature data with the variance method starting with autumn 2019. For the first stage of the pilot, i.e. summer 2019 temperature data was corrected using the EQM method. This choice was justified by the results of preliminary skill analyses from spring 2019, based on which EQM method indicated reasonable improvements for temperature. It must be noted that in the preliminary skill assessments fewer methods were tested, also the spatial resolution and length of reanalysis data involved in computation was more reduced. Following the extended skill assessment experience and the outputs presented above the choice of bias correction method was reconsidered.

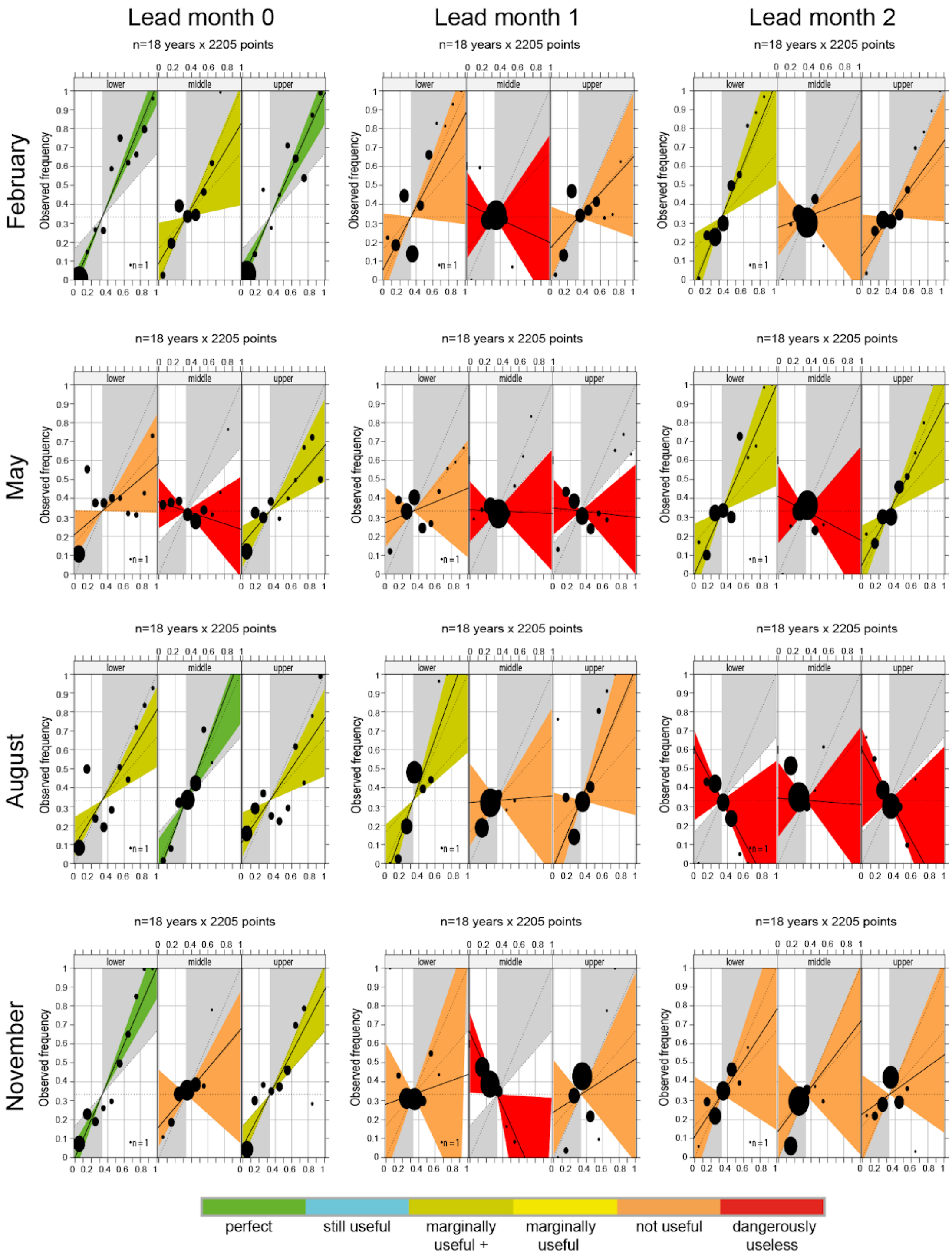

Figure 1. Reliability diagrams of tercile mean temperature forecasts for February, May, August and November with lead month 1, 2 and 3 for the bias corrected data with the variance methods.

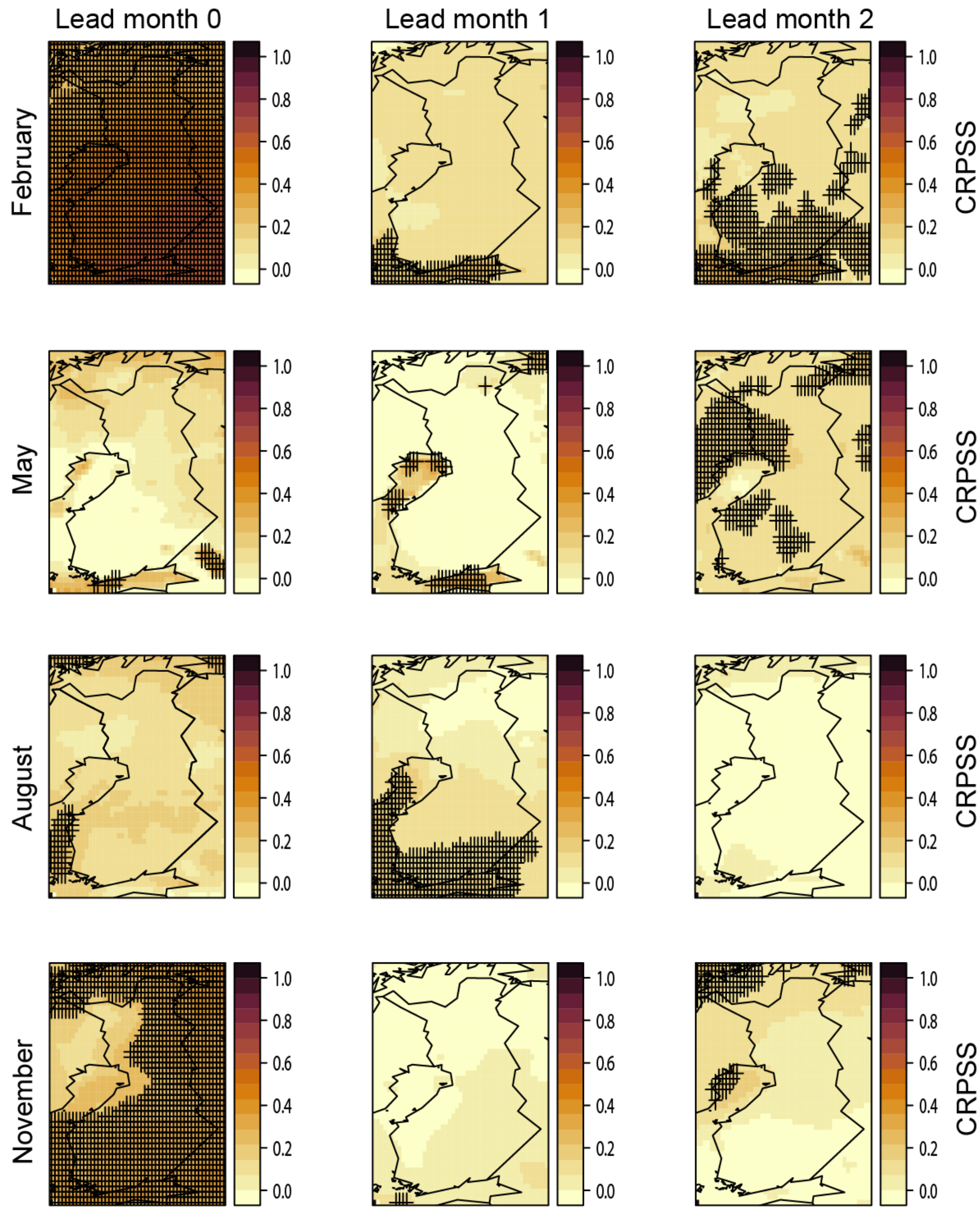

Figure 2. Maps of CRPSS for mean temperature forecasts for February, May, August and November with lead month 1, 2 and 3 obtained for bias corrected data with the variance method. Statistically significant grid points are marked with crosses.

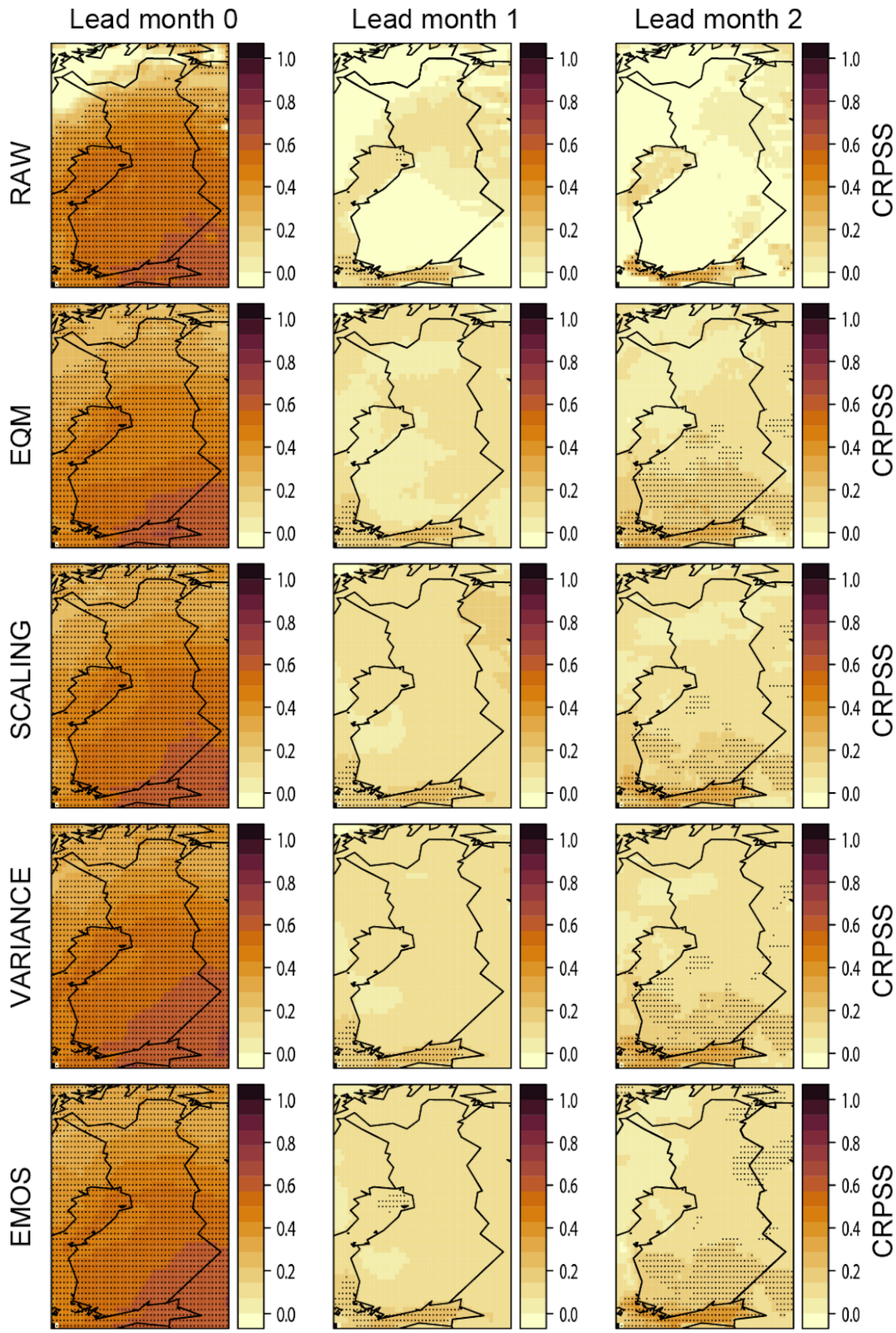

Figure 3. Maps of CRPSS for mean temperature forecasts for February for lead month 1, 2 and 3 shown for raw model data and bias corrected data.

3.2 Total precipitation

Total precipitation results for the months for which the full 51 reforecast members were available (February, May, August and November) are shown below. For the comparison between the methods, the results for February using CRPSS are shown (Fig. 4). All bias-adjustment methods improve on the raw forecasts. The results are statistically significant for the first lead month, but after that difference between the climatology is rather small, or negligible.

The EMOS is arguably the most promising method for further study, as even with the very simple implementation (the Gaussian approximation without any covariates), it gives competitive results and could be easily expanded with additional information sources. However, the results in November were disappointing (CRPSS no better than the climatology in any grid point, for reasons to be studied later), and the results for all four months are shown using the scaling method (Figs. 5 and 6).

All in all, the "winter" months (February and November) got higher scores than the "summer" months (May and August). Both CRPSS and reliability diagrams agree that after the first lead time, the probabilities are, mostly, not really useful or even somewhat misleading. Since according to the preliminary verification results from spring 2019 no improvement was seen in the quality of bias corrected precipitation data, the raw forecast was used in the production of climate indices during the pilot seasons. However, based on the new skill assessment results the biases can be effectively reduced with EMOS, this method will be considered in future applications.

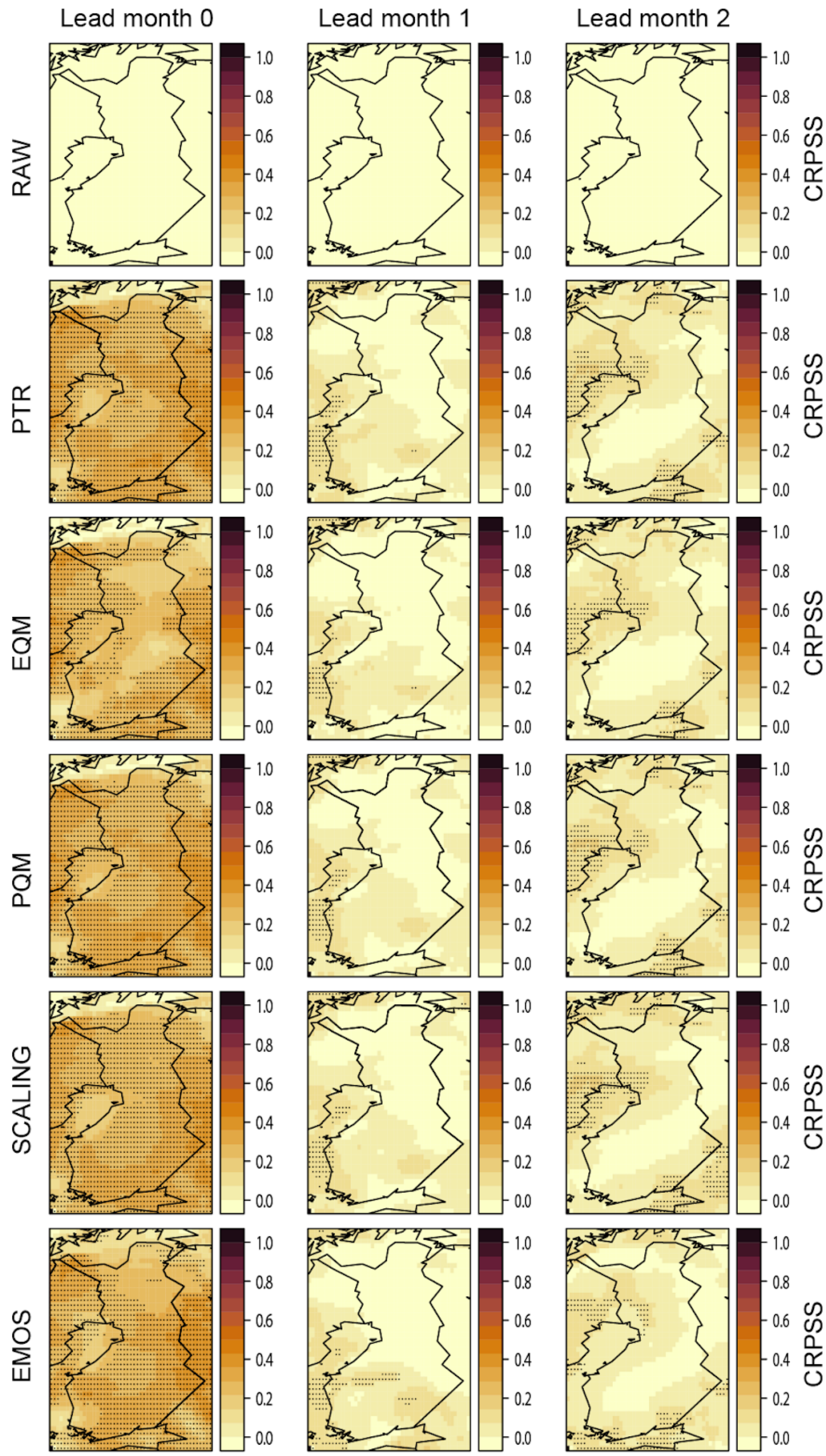

Figure 4. Maps of CRPSS for total precipitation forecasts for February with lead month 1, 2 and 3. Both not corrected or "raw" forecasts and forecasts bias-corrected with different methods (see Table 1) are shown. Statistically significant grid points are marked with small crosses.

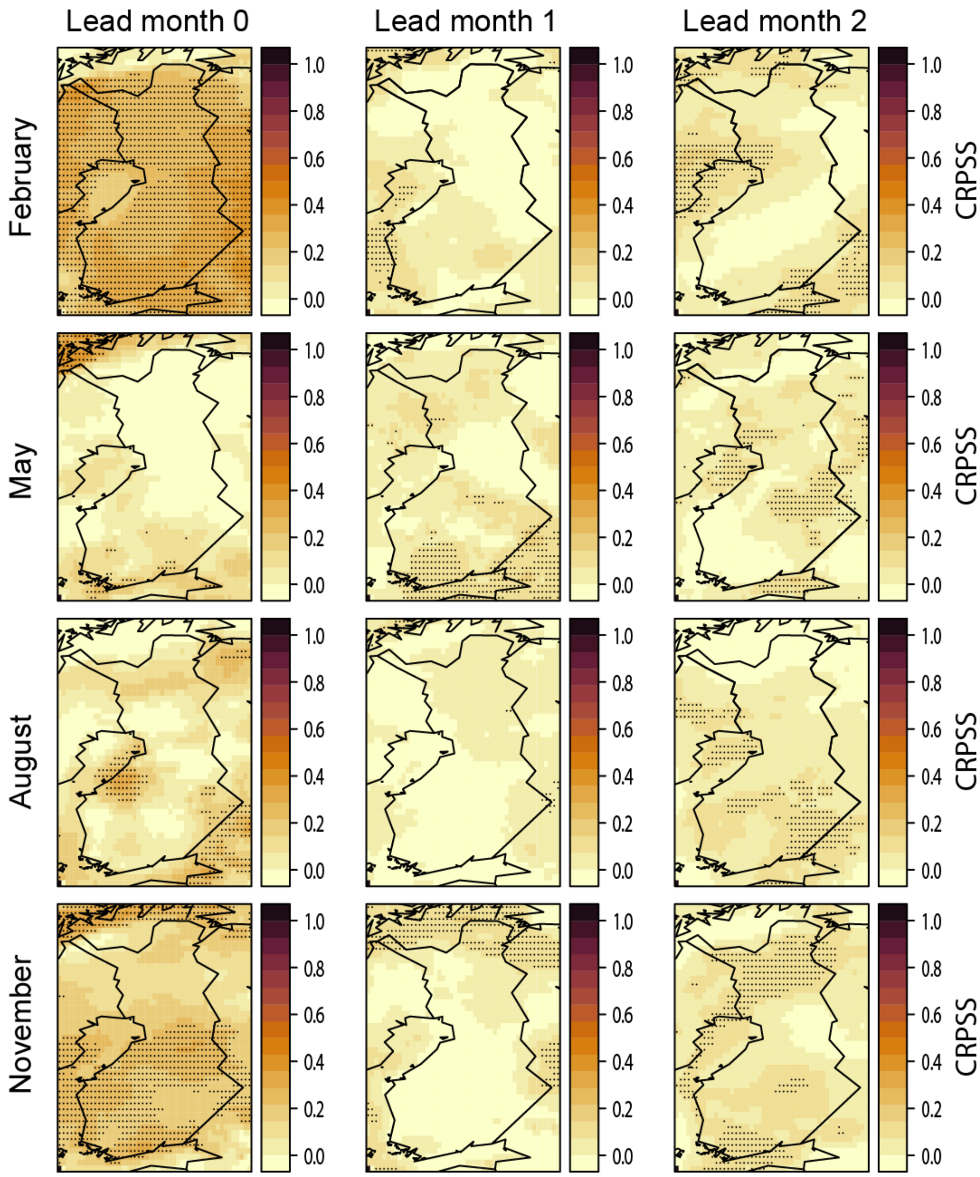

Figure 5. Maps of CRPSS for mean total precipitation forecasts for February, May, August and November with lead month 1, 2 and 3 obtained for bias-corrected data with the scaling method. Statistically significant grid points are marked with small crosses.

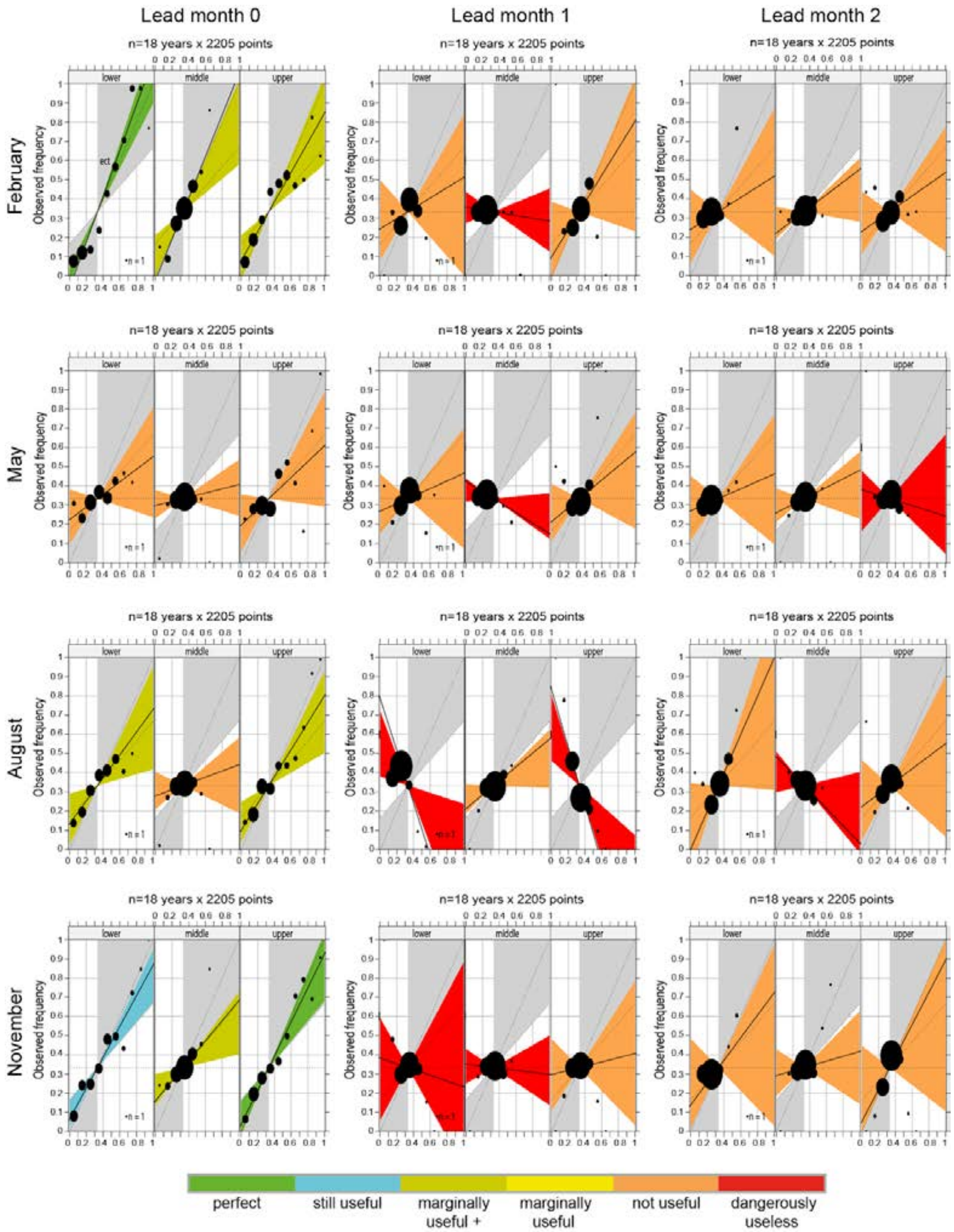

Figure 6. Reliability diagrams of tercile total precipitation forecasts with lead month 1, 2 and 3 for February, May and November obtained with the scaling method.

3.3 Soil moisture

Verification results of soil moisture are shown for those months (February, May, August, and November) that have the same 51 number of reforecast members as the operational forecasts.

Forecasts were bias-adjusted using the variance method. These results are based on Hyvärinen et al. 2020.

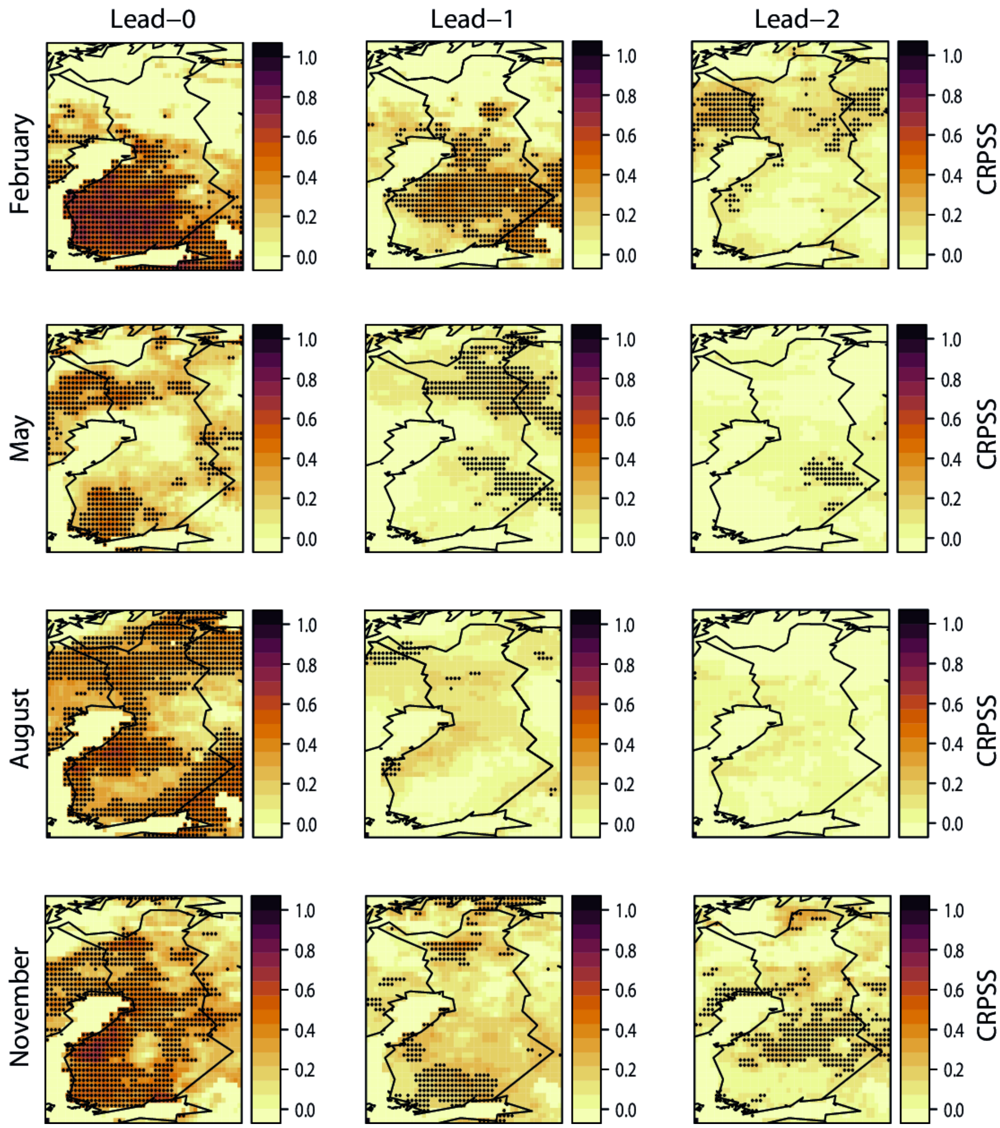

Figure 7. The maps of CRPSS for soil moisture forecasts in February, May, August, and November with lead times of 0, 1, and 2 months. Statistically significant grid points marked with points. Forecasts are bias-adjusted using the variance method. From Hyvärinen et al. 2020.

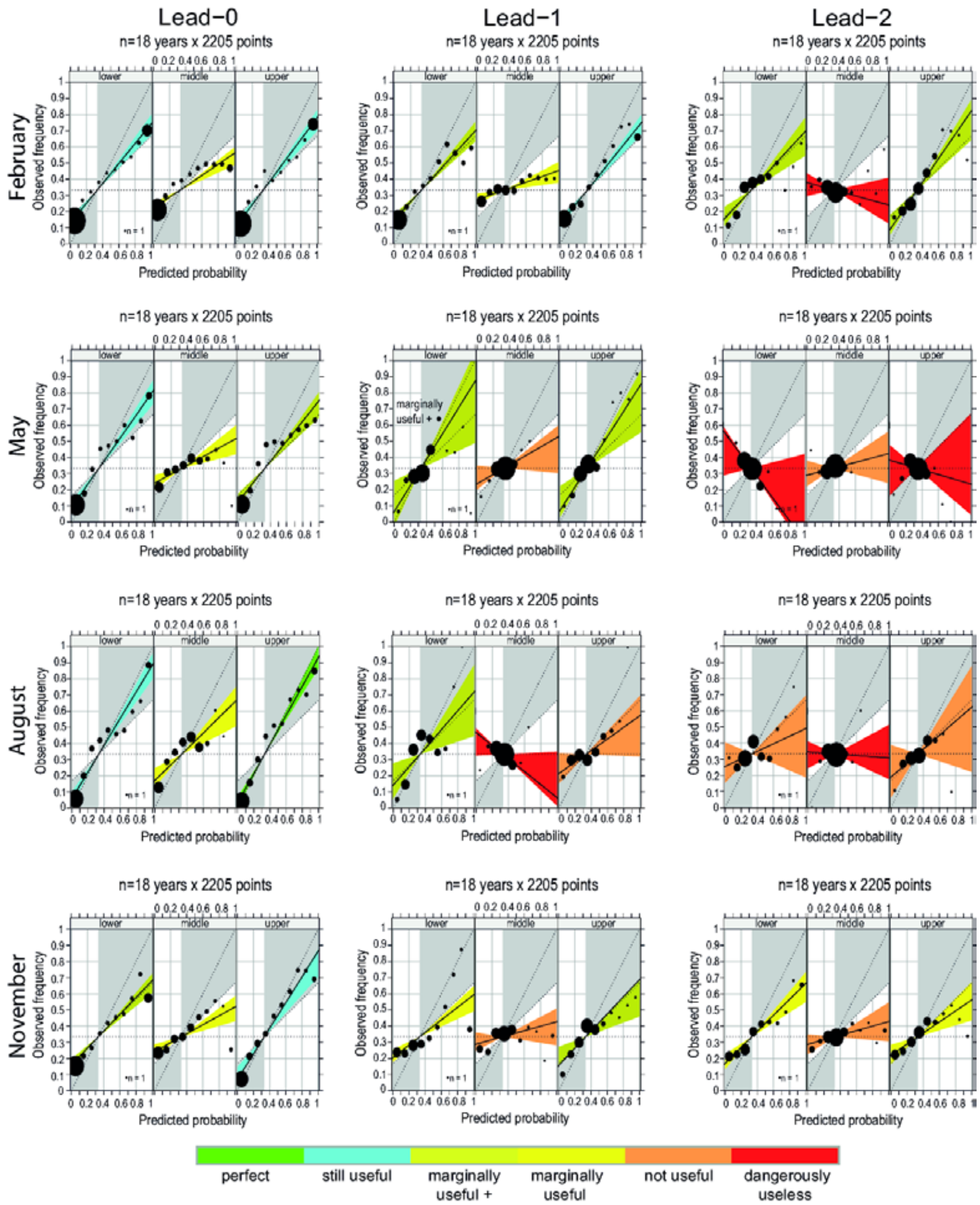

Figure 8. The reliability diagrams of tercile forecasts for soil moisture for February, May, August, and November for Finnish land areas with lead times of 0, 1, and 2 months. Forecasts are bias-adjusted using the variance method. From Hyvärinen et al. 2020.

The CRPSS maps are shown in Figure 7. For the first lead month, there are rather large areas with reasonably large skill score values that are statistically significant. For the second lead month, values are lower, somewhat better in winter months than in summer months. For third lead time, values are very near zero everywhere and statistical significant areas are scarce, so it is mostly hard to argue that the forecasts would be better than the reference.

Reliability diagrams of tercile forecasts for Finnish land areas are shown in Figure 8. All terciles of all months for the first lead month forecasts are usable, but results for other lead times diverge. In winter months (November and February), lower and upper terciles are somewhat usable even with the lead-2 forecasts, but for summer forecasts these terciles are no longer very usable, and can be downright misleading. For all months, the forecasts of middle tercile are mostly unusable for longer than the lead-0 forecasts. The probable reason for this is that it is difficult to discriminate between no signal and the average conditions.

3.4 Snow depth

Verification results shows good skill both for raw model snow depths and for snow depths corrected using EQM, variance and EMOS methods. Reliability diagrams of snow depth forecasts, shown for Feb, May and Nov (Fig. 9), indicates that all the terciles for lead time 0 varies between useful and marginally useful for all the initializations. In terms of lead month 1 and 2 forecasts all the terciles are usable for February, varying between perfect and marginally useful categories but with varying between marginally useful and useless for May and November. This can be related to challenges in forecasting the onset and melting of snow cover in the beginning and end of the season.

For lead month 0, the CRPSS values are large, with values close to 1 and statistically significant over most of Finland in February and over large areas in November (Fig. 10). High skill score values are indicated also for LM 1 and 2 for Feb, lower values but still showing skill are given for the other months. All the bias correction methods reduced the mean error from the row data and improved the skill. According to the CRPPS metrics, EMOS reduced the most effectively the biases compared to the EQM and variance methods. Thus, the EMOS adjusted snow depth data was used in the production of snow related seasonal forecast indices and climate outlooks.

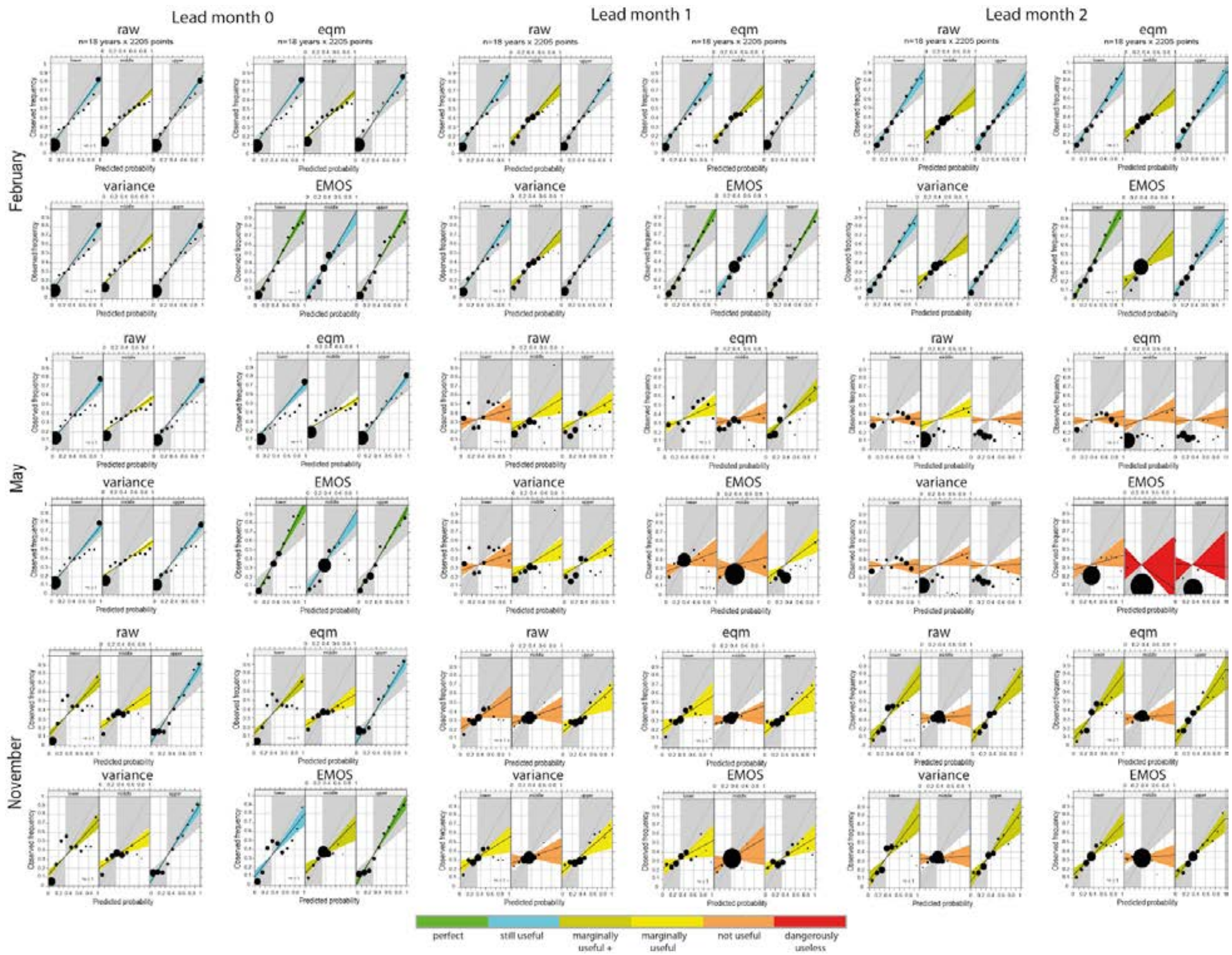

Figure 9. Reliability diagrams of tercile snow depth forecasts with lead month 1, 2 and 3 for February, May and November obtained for the raw model data and bias corrected data with the EQM, variance and EMOS methods.

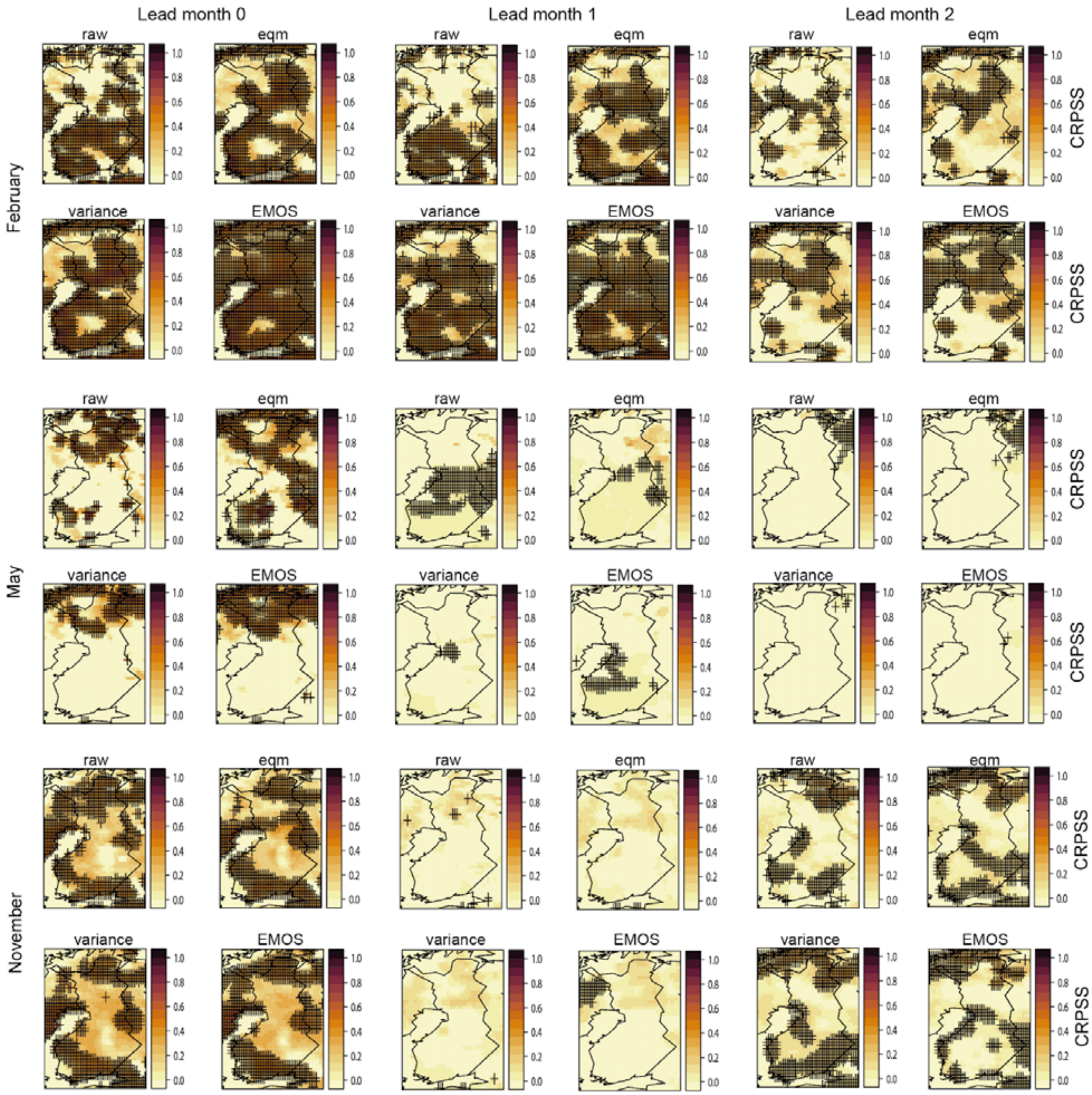

Figure 10. Maps of CRPSS for snow depth forecasts with lead month 1, 2 and 3 for February, May and November obtained for the raw model data and bias corrected data with the EQM, variance and EMOS methods. Statistically significant grid points are marked with crosses.

5.6 Growing degree days index

Verification results for the growing degree days (GDD) index calculated from the raw model data and post-processed data with EQM and variance method were analyzed. Reliability diagrams for tercile forecasts of GDD index indicate that for LM0 the lower and upper terciles fall between marginally useful and perfect category for July-Aug-Sep initialization but for LM1 and 2 the forecasts are useless both for raw model data and bias adjusted data. In June the forecasts are not useful for any lead time (Fig. 11).

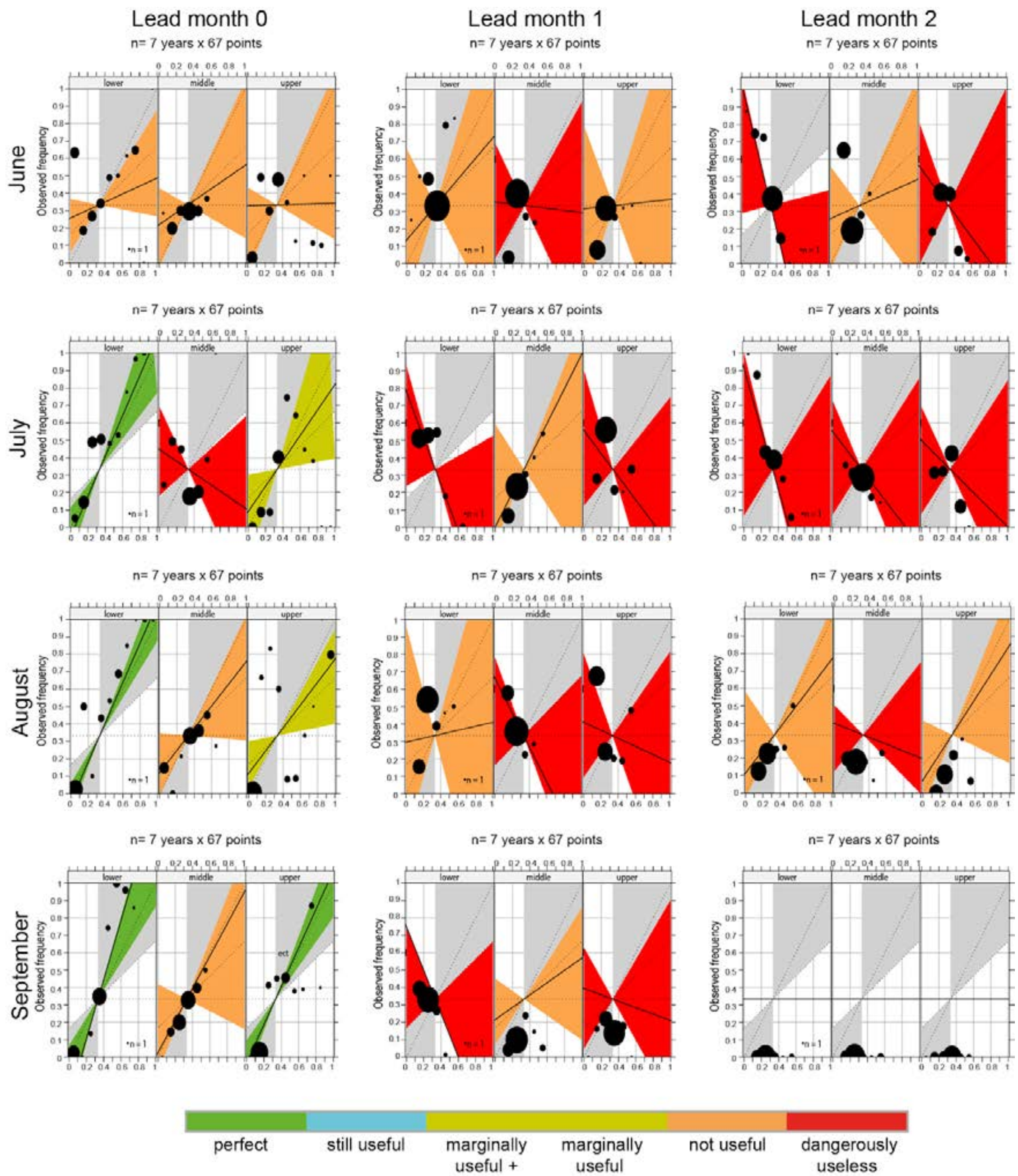

Figure 11. Reliability diagrams of tercile growing degree days index forecasts with lead month 1, 2 and 3 for June-September obtained for the bias corrected data with the EQM method.

Considerably large CRPSS values that are statistically significant were found over large areas for LM0 in July and September (Fig. 12) somewhat lower but still statistically significant skill score values were in June for LM1. For the other initializations and lead times the skill scores are lower, in some cases near-zero. Both variance and EQM method improved slightly the skill of GDD index for each LM and initialization except for August; the improvements shown by the two bias correction methods were of the same level (Vajda and Hyvärinen, submitted). Also, using the observed GDD as a starting point in the computation of forecasted GDD values improved the forecast to some degree. It has been decided to apply the EQM in the bias adjustment process during the pilot.

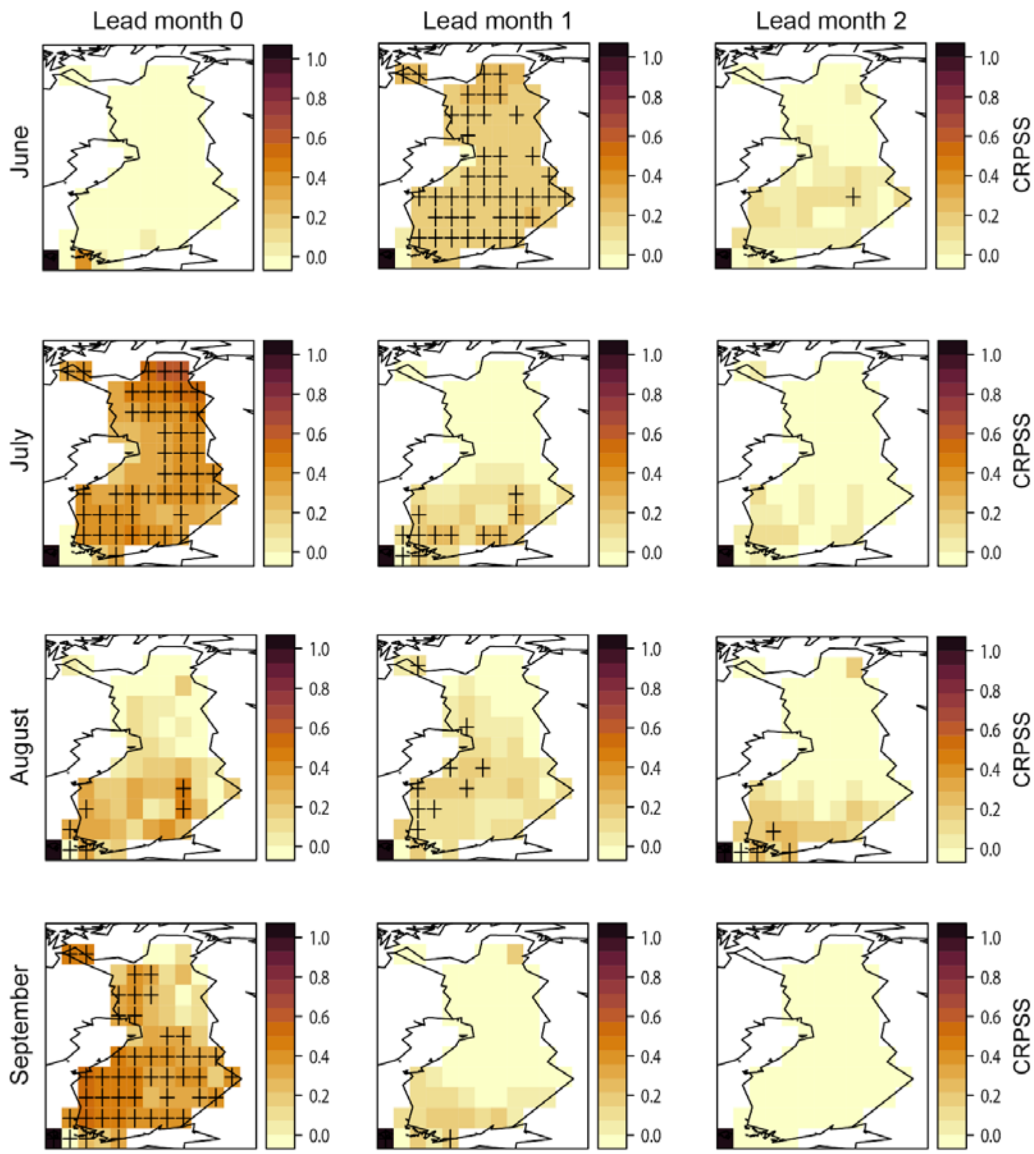

Figure 12. Maps of CRPSS for bias corrected growing degree days index (applying the EQM method) for June-September with lead months 0, 1 and 2. Statistically significant grid points are marked with crosses.

6. The seasonal forecast indices developed and pilot phases

6.1 Seasonal forecast indices developed

A set of seven seasonal forecast indices were tailored for agriculture and five seasonal forecast indices for ski resorts (Table 2). The indices developed for the farmers were: mean temperature, development of growing season, growing degree days, cold spell, total precipitation and dry/rainy conditions, soil moisture. The development of growing season, cold spell and dry/rainy condition indices described the likelihood of expected weather conditions through probability forecast, while the mean temperature, growing degree days and total precipitation indices through absolute values (Vajda and Hyvärinen, submitted). All the indices were computed for three months ahead and updated every month.

The mean temperature index was calculated for each grid from the bias corrected (using EQM method for May-Oct 2019 and variance starting from Nov 2019) monthly ensemble mean values. Total precipitation index was calculated in a similar way but using the raw model monthly ensemble precipitation sum.

The growing degree day index provided the forecasted GDD sum accumulated by the end of each month. GDD was defined as the degree sum above the base temperature (5 °C) since the start of the growing season and was calculated from the bias corrected (using EQM method) daily values. The forecasted GDD consists of the observed GDD by the initialization month and the additional forecasted sum of temperatures above the base temperature for each ensemble member. The development of growing season index indicates the likelihood of growing season being behind, normal or ahead of climatology (Fig. 13). The development phase was defined by the proportion of forecast members indicating below normal, normal and above normal category compared to climatology for the normal period 1981-2010. In addition, probabilities within these categories describe the proportion of members falling in the tercile categories (below 33%, 33-66% and above 66%) of the forecast. For instance, if most of the ensemble members lied within the below normal category the growing season was predicted to be behind the climatology and the uncertainty is relatively small (Vajda and Hyvärinen, submitted).

Table 2. The seasonal climate indices developed including the variables, thresholds and methods used in the calculations.

Seasonal climate index	Variables used in product development	Methodology and thresholds applied
Mean temperature	Forecasted t2m	Monthly ensemble mean t2m value
Development of growing season	Forecasted daily mean temperature (T_{daily}), observed GDD, GDD climatology (1981-2010)	t_{ref} = 5 °C GDD= ΣT_{daily}- t_{ref} $GDD_{forecast}$=GDD_{obs}+ $\Sigma T_{daily\ forec}$ - T_{ref} $GDD_{forecast}$ compared to GDD_{obs}

Growing degree days	Forecasted daily mean temperature (T_{daily}), observed GDD	$t_{ref} = 5$ °C $GDD= \Sigma T_{daily} - t_{ref}$ $GDD_{forecast}=GDD_{obs}+ \Sigma T_{daily\ forec} - T_{ref}$
Cold spell	Forecasted daily T_{mean}, re-forecasted daily mean temperature	Forecasted 6-day T_{mean} falls below 0.1 quantile of re-forecast T_{mean} climatology
Total precipitation	Total precipitation (TP)	Forecasted monthly ensemble TP value
Dry/rainy conditions	Forecasted total prec, re-forecasted total prec	Forecasted monthly ensemble mean TP amount compared to the monthly 0.5 quantiles of re-forecasted TP
Soil moisture	Forecasted volumetric soil water layer 1 and 2, re-forecasted volumetric soil water layer 1 and 2	SM= swlv1*0.25+swlv2*0.75 Forecasted monthly SM of the ensemble mean compared to the monthly 0.5 quantiles of re-forecasted SM
Probability of snow cover	Forecasted monthly mean snow depth	Percentage of forecast ensemble members indicating ≥1 cm of monthly mean snow depth
Snow depth	Forecasted monthly mean snow depth	Forecasted monthly ensemble mean snow depth
Conditions for artificial snow production	Forecasted daily dew point temperature, forecasted daily wind speed	$T_{dew\ point} < -5$ °C WS < 5 m/s Daily minimum dew point temperature combined with daily maximum wind speed
Occurrence of maximum wind speed	Forecasted daily wind speed, observed daily wind speed	Forecasted number of days with maximum wind speeds between given thresholds related to the station observations

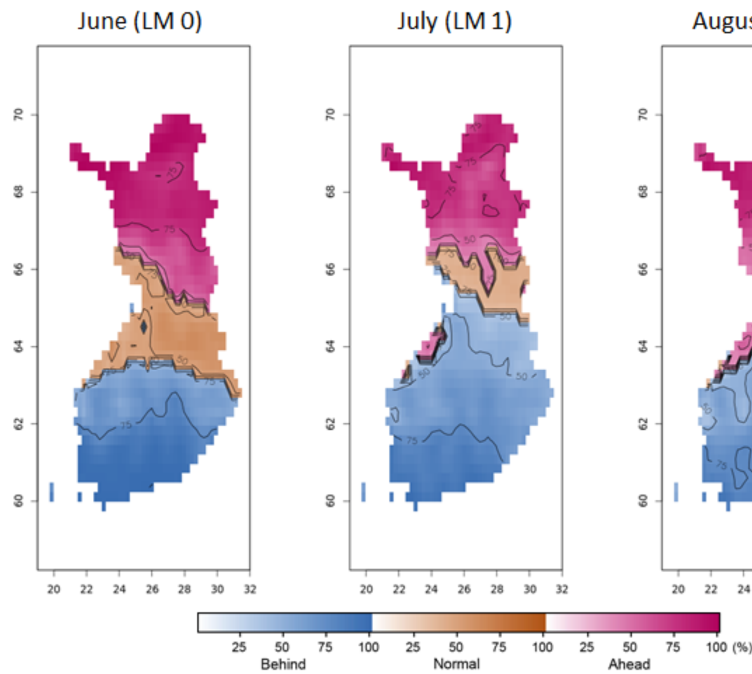

Figure 13. Seasonal climate outlook describing the development of growing season issued on June 6, 2020.

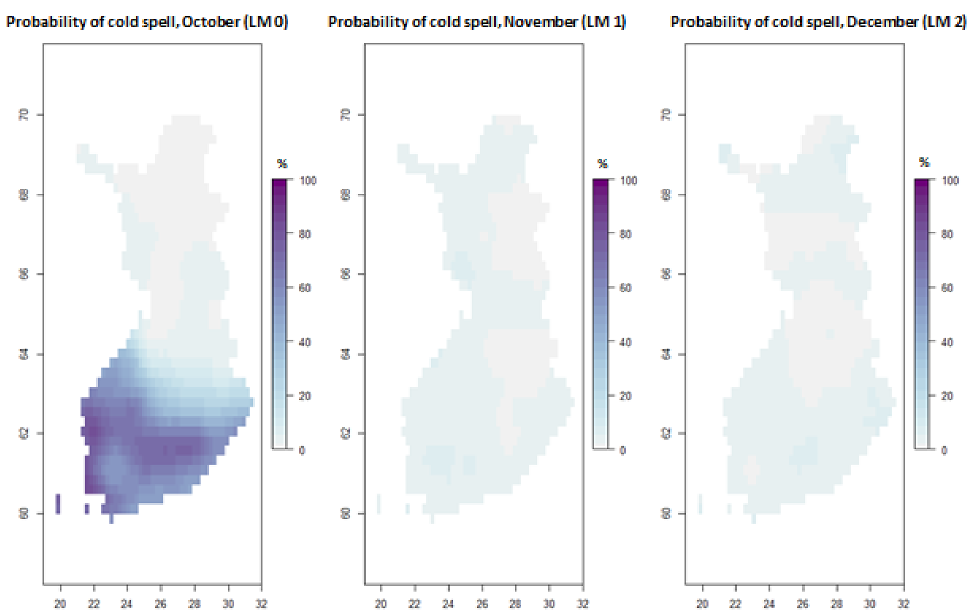

Figure 14. Seasonal climate outlook describing the probability of cold spell issued on October 6, 2019.

The cold spell index (Klein Tank et al. 2009) was implemented by combining the forecasted mean temperature with the reforecast mean temperature for the period 1993-2016. A cold spell is considered to occur when the mean temperature of a 6-day period falls below the 0.1 quantile of the re-forecast mean temperature climatology. The proportion of ensemble members satisfying this criterion gives the probability of occurrence (Fig. 14).

The dry/rainy conditions index defines the drier than average and rainier than average conditions. It was determined by relating the forecasted monthly total precipitation amount of the ensemble mean to the monthly 0.5 quantiles of re-forecasted total precipitation for the respective month from the 1993-2016 period. Values below 0.5 quantile marked the drier, those above 0.5 quantile the rainier than usual conditions (Vajda and Hyvärinen, submitted).

The soil moisture index described the conditions of soil, i.e. being either drier or wetter than usual. First, the soil moisture was calculated from the volumetric soil water layer 1 (soil layer from 0 to 7 cm, swlv1) and volumetric soil water layer 2 (soil layer from 7 to 28 cm, swlv2) as swlv1*0.25+swlv2*0.75 (Hyvärinen and Vajda, 2020). The index was computed by relating the forecasted monthly soil moisture of the ensemble mean to the monthly 0.5 quantiles of re-forecasted soil moisture for the respective month from the 1981-2016 period. Values below 0.5 quantile marked the drier, those above 0.5 quantile the wetter than usual soil conditions (Fig. 15).

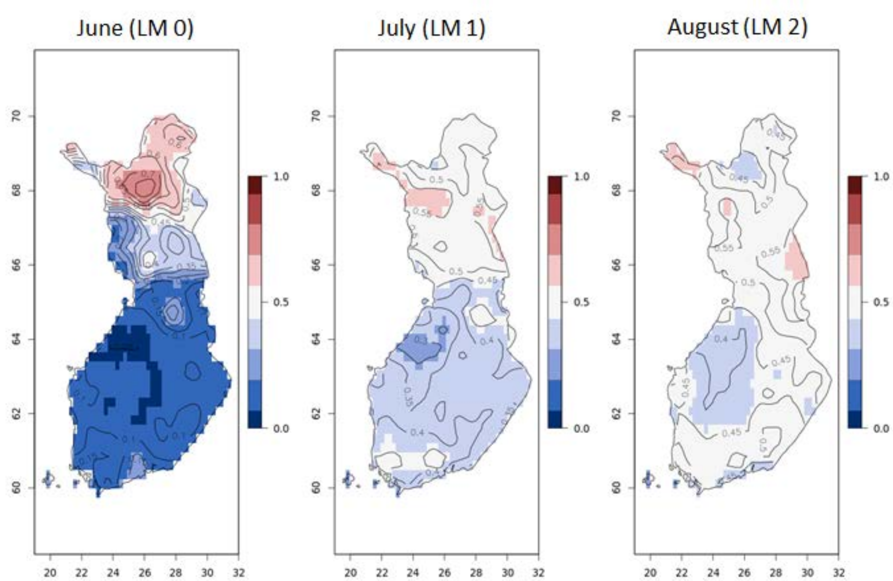

Figure 15. Seasonal climate outlook describing the development of soil moisture for lead month 0, 1 and 2 issued on June 6, 2020.

The following indices were developed and tested with the ski resort stakeholders: mean temperature (as provided for the farmers), probability of snow cover, snow depth, conditions for artificial snow production and occurrence of maximum wind speed.

Both the probability of snow cover and snow depth were calculated from the bias corrected monthly mean snow depth values. The probability of snow cover was defined as the percentage of forecast ensemble members indicating at least 1 cm of monthly mean snow depth. The snow depth index describes the monthly mean snow depth for each grid box.

Conditions for artificial snow production index described the probability of days with favorable weather likely to produce snow. The suitable weather conditions for snow production was defined together with the end-users, effective snowmaking depends on low air temperature, humidity and relatively calm wind conditions. Suitable conditions for artificial snow production were considered to occur when the dew point temperature fell below -5 °C and wind speed is below 5 m/s. Bias corrected (using variance method) daily minimum dew point temperature and raw forecast data for daily maximum wind speed was used in the calculations. The probability of days/month fulfilling the above mentioned criteria is given in three categories: less than 10, 10-20, more than 20 days (Fig. 16). The proportion of ensemble members that satisfied the criterion expressed the probability of occurrence for the three categories.

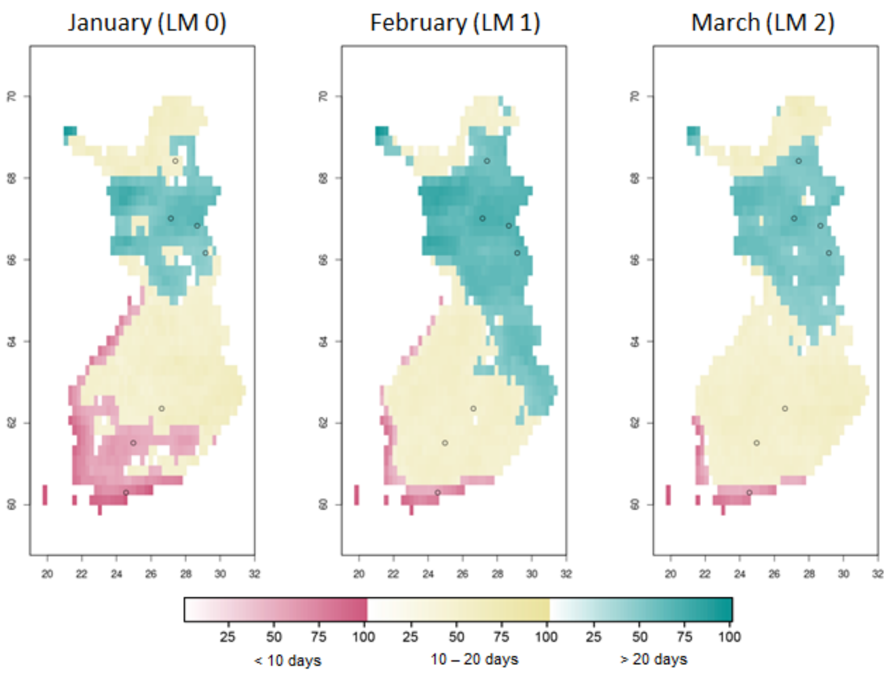

Figure 16. Seasonal climate outlook describing the probability of days suitable for artificial snow production for lead month 0, 1 and 2 issued on January 6, 2020.

The occurrence of maximum wind speed index predicts the number of days with maximum wind speed exceeding certain thresholds in the given location (Fig. 17). Raw wind speed model data was bias corrected using observations for the 1996-2016 period from the nearby weather stations located on mountaintop and using the quantile mapping method with gamma distribution. The forecasted number of days with certain maximum wind speeds was related to the station observations for each location. Uncertainty information is given through the 0.1 and 0.9 quantiles (shown with whiskers) and the median of the observed number of days (in dashed lines).

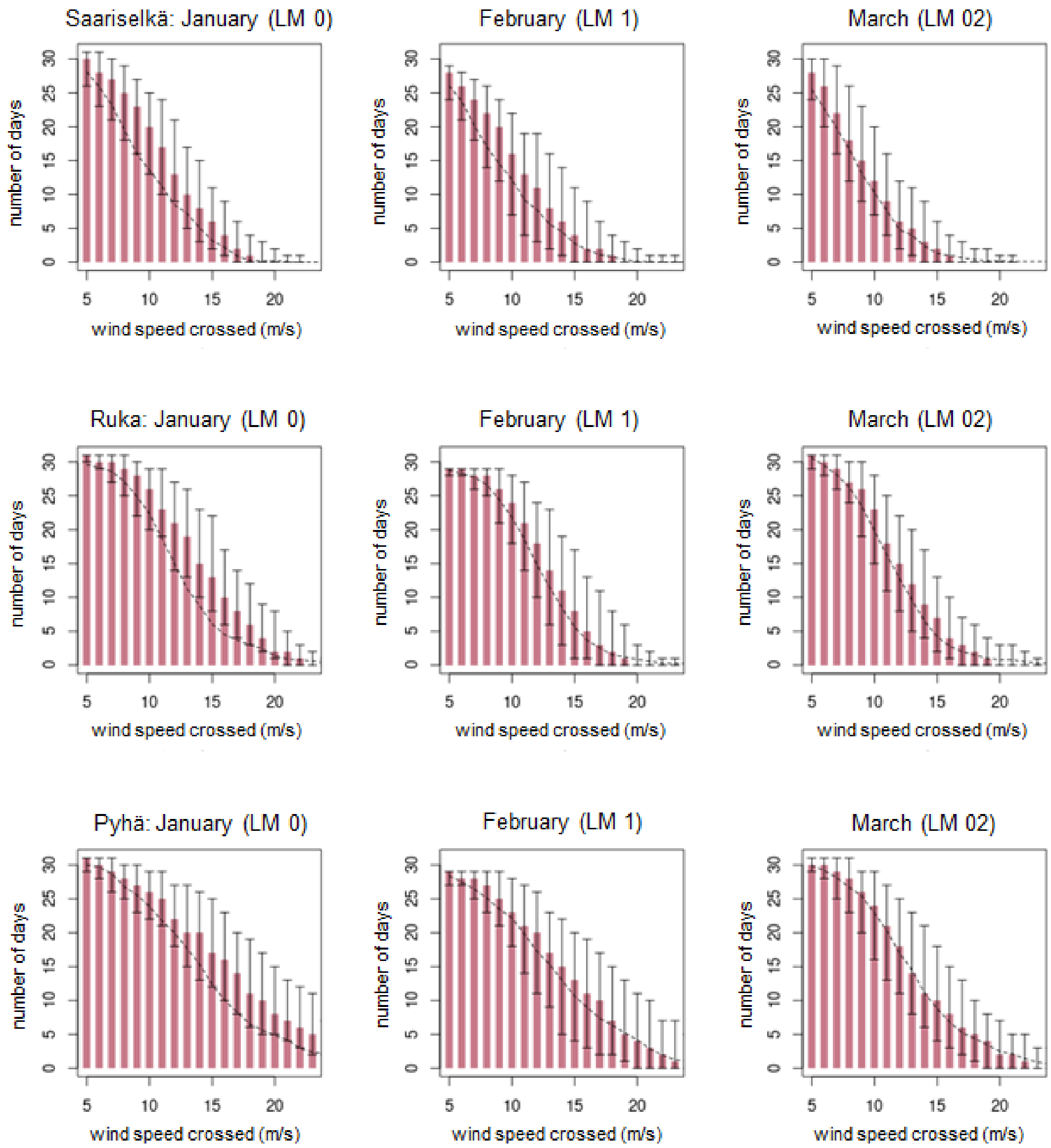

Figure 17. Seasonal climate outlook describing the occurrence of maximum wind speed for LM1, 2 and 3 initialized on January 6, 2020.

6.2 Climate outlooks and pilot phases

The forecasts of tailored seasonal indices were produced manually on the 5th of every month, when updates from ECMWF were released. The indices visualized through monthly maps were combined into seasonal climate outlooks and sent to the users. When designing the outlooks we aimed for an easily understandable and interpretable format. The outlooks also contained information on the indices, such as a short description of the index, advice on how to interpret the forecast and information on the computation of the product.

Two testing stages were run with the farmers, the first pilot during May-October 2019 followed by a second pilot during April–October 2020 (still on-going). The first seasonal climate outlooks were issued to the farmers in June 2019 and included four indices: mean temperature, development of growing season, growing degree days and total precipitation. The indices describing dry conditions and cold spells were additionally added in August. However, all the indices were subsequently computed for the time period May-October 2019 for evaluation purposes. Following the farmers' wishes regarding new indices from the feedback given at the end of the first pilot, the soil moisture index was developed and included to the climate outlook starting from April 2020.

The pilot of seasonal climate outlooks generated for the ski resorts took place during November 2019-April 2020. The first outlook issued included the mean temperature, probability of snow cover, snow depth and conditions for artificial snow production forecasts. The occurrence of maximum wind speed index was added in December 2019 and it was developed and issued exclusively for the resorts located in Lapland where hazardous gusty winds may impact the operation of ski lifts.

7. Evaluation of seasonal forecast products from the first pilot phase

7.1 Performance of seasonal climate outlooks during the pilot seasons

Following the pilot season from summer 2019 and winter 2019/2020, the accuracy of forecasts were tested by comparing the climate outlooks provided with observations for temperature, growing degree days and precipitation for the summer testing phase, and snow depth and maximum wind speed for winter season. Precipitation and GDD forecasts were tested for three stations located in various regions: Jokioinen (60°48', 23°29') located in southern part of Finland, Seinäjoki (62°56', 22°29') from central Finland and Sodankylä (67°21', 26°37') from Northern Finland. An analysis of snow conditions and maximum wind speed outlooks were run for the ski resorts involved in the pilot. The performance of the climate outlooks during one pilot phase clearly does not assess the skill of the forecasts and developed indices, the intention is rather to estimate if the produced forecasts were of any use for the farmers.

The skill of **mean temperature** was in general good during the whole pilot (May 2019-Aug 2020). The forecasts performed better for LM0 (not shown), marking biases of maximum 1-2 °C range, especially in the southern and southwestern coastal area, where the mean temperature was systematically underestimated both during summer 2019. Similar analogy were found for LM1 and LM2 forecasts for summer months 2019, except for July, for which mean temperature was overestimated by both May and June runs. During the winter season the model forecasted lower monthly mean temperature for Nov-Jan, with larger errors (up to 3-4 °C) for longer lead times.

This underestimation of winter temperature was more pronounced in Southern and Central Finland. Winter was exceptionally warm in Finland. Although the forecast indicated a warmer than average winter already in November initialization, the anomaly was predicted to be lower. During late winter and early spring 2020 the forecasts overestimate the monthly mean temperature. The quality of summer forecast was in general good, although the warm June was systematically underestimated.

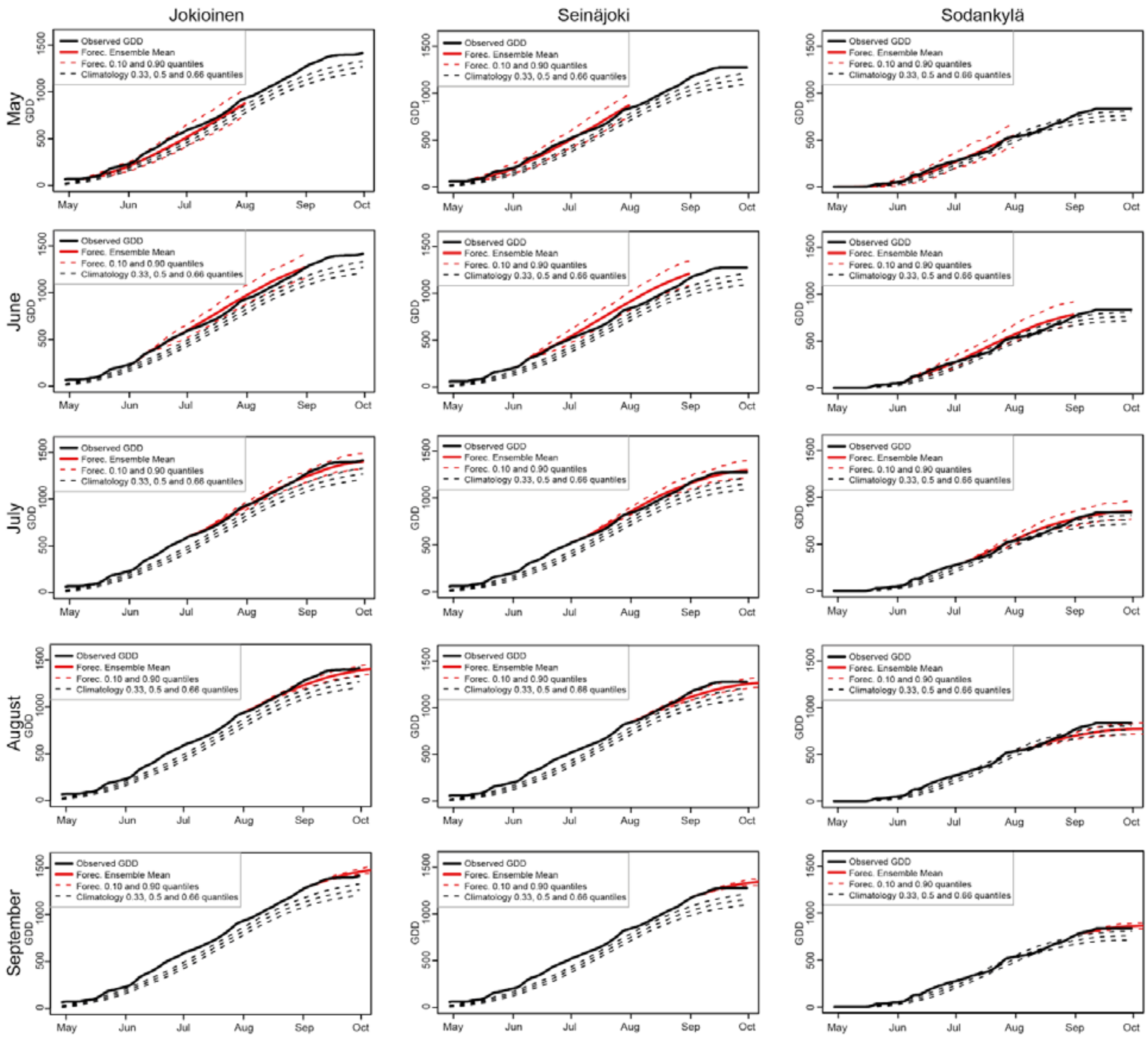

Figure 18. Forecasted and observed GDD during May-September 2019 for the selected locations: Jokioinen, Seinäjoki and Sodankylä (from Vajda and Hyvärinen submitted to ASR).

In terms of growing season outlooks, the **development of growing season** was predicted well for LM0 in case of each run, however the LM1 and LM2 forecasts slightly overestimated the development of growing season for the whole Finland except for the southern part. The accumulation of forecasted and observed **growing degree days sum** during the pilot season is shown in Figure 18 for the three selected locations. The predicted GDD sum is in strong resemblance with the observed accumulation, the skill increasing for shorter lead time. Although the LM2 forecast has slightly reduced skill, the observed values fall within the 10 and 90% probability distribution of ensemble members. Reliable predictions of growing season development and growing degree days sum provide valuable information on crops stages, agricultural productivity permitting better management and planning the timing of activities.

The quality of **precipitation** forecast was poor. The amount of precipitation was in general overestimated during summer 2019 for each lead time (Fig. 19), with the largest errors in the extremely dry month of July which was predicted to be rainier than usual. The quality of precipitation forecast for the summer season 2019 was lower than that of the reforecast climatology for the same months for the selected location. Similar features were seen during the summer season, 2020 with overestimation of precipitation during April, May and August. Precipitation forecasts for June and July were of better quality, except for more local outlier rainfall episodes.

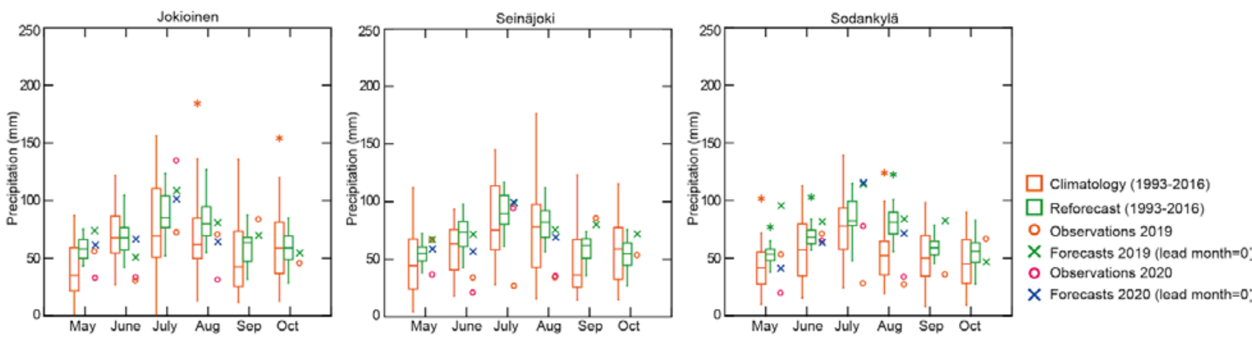

Figure 19. The forecasted and observed monthly precipitation sums for summer 2019 and 2020 and the climatology based on reforecast precipitation values and observations from 1993-2016 for the selected locations. Observed and forecasted precipitation values are missing for Sep-Oct 2020.

As winter 2019/2020 was exceptionally warm in Finland, snow conditions were also exceptional: the southern part was snow free (in Helsinki 3 cm of maximum snow depth during the winter, the lowest registered from the last 100 year) while exceptionally large snow amounts were registered in Lapland (with the largest snow depth registered form the last 100 year in Sodankylä, 127 cm). The forecasted and observed **snow depth** during Nov-Apr for the ski resorts involved in the pilot is shown in Figure 20. Observations are obtained from the closest weather station.

The quality of snow depths forecast was in general good, with best performance for lead month 0. However, the snow amount was slightly underestimated in the northern located resorts, especially towards the end of the season. In case of the ski resorts located in southern Finland the snow amount was underestimated, the errors ranging between 3 and 15 cm.

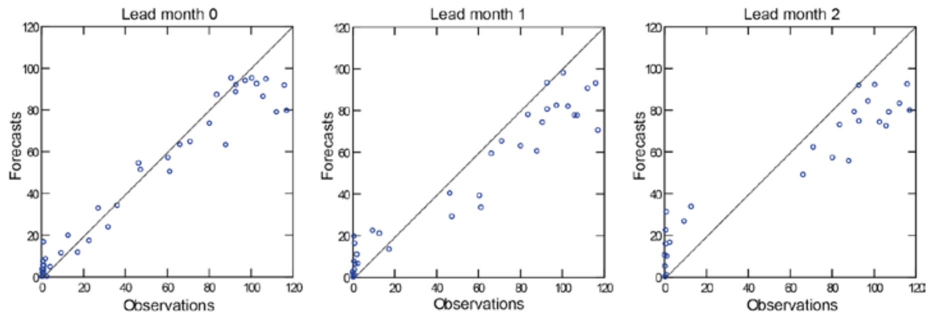

Figure 20. The forecasted and observed snow depths for November 2019-April 2020 for the ski resorts involved in the pilot.

For the performance of **occurrence of maximum wind speed outlook**, the forecasts made in February are shown in Figure 21 as an example. The dark red bars show the mean forecasted number of days when the wind speed threshold is exceeded with the 80% confidence intervals. The black dotted line is the climatology. The red dashed line is the actual number of days based on observations. When the observations are not very extreme, the forecasts seem to be reasonable. For example, in Saariselkä and Ruka and with the lead month 0, the forecasts suggest somewhat higher than normal winds, and observations agree. On the other hand, the rather unusually strong winds, as in Pyhä, were not forecasted very well.

All in all, the mean forecast of the number of days might not work very well, but the observations are most of the time inside the confidence intervals, as they should be. Wider confidence intervals would have covered more observations. This suggests that a probabilistic forecast would be more informative, but would be more complicated and the user might need more training before they can gain from such forecasts. Of course, a more extensive validation using reforecasts would be needed before we can say if the forecasts are better than climatology.

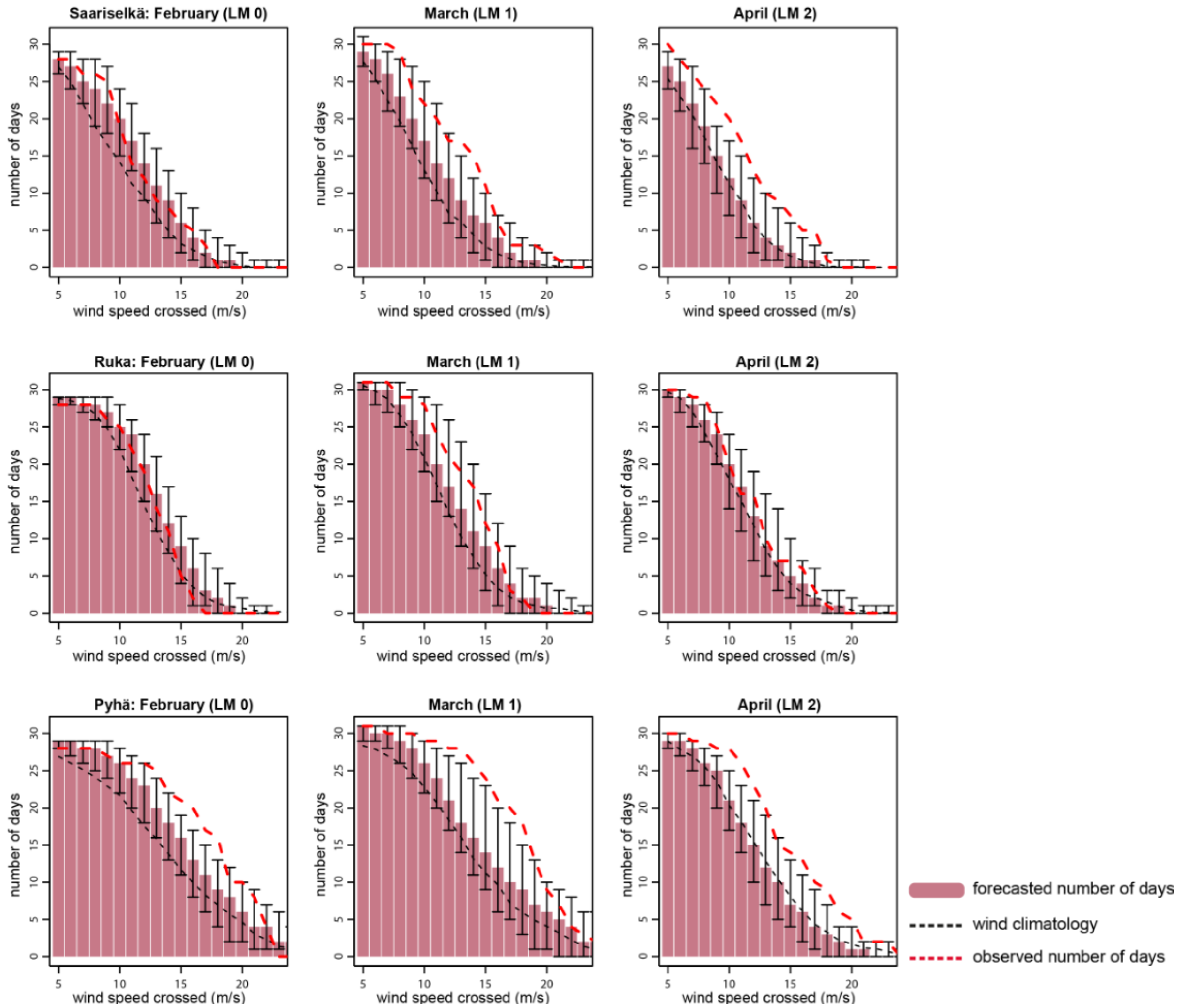

Figure 21. Forecasted and observed occurrence of maximum wind speed for February, March and April 2020 initialized in February 2020.

7.2 Evaluation by the users

Following both the pilot from summer 2019 and winter 2019/2020 a feedback survey was conducted with the users to gather their opinion about the seasonal climate outlooks and the usefulness of the developed indices. A summary of the feedback forms and their results are presented below.

The information asked from the farmers referred to (1) how useful and (2) how understandable the climate outlooks were, (3) how the forecasts were used, (4) have the users changed their planes based on the forecasts and how, (5) what kind of improvements they wished for and (6) what other indices would be useful for farmers. The farmers were also asked about their willingness to test the seasonal outlooks during spring-autumn 2020. Background information of the respondents, such as the province they operate in, experience as a farmer and weather and climate information they usually use in their work was also asked.

Of the involved farmers, 43 answered the survey, of which 83% found the climate outlooks very easily or easily understandable. All the tested seasonal indices were found useful, the total precipitation index was marked as very useful and useful by 78%, and mean temperature and development of growing season by 74% of the respondents (Fig. 22). The evaluation regarding precipitation is not in line with our verification results for the pilot season. Nevertheless it might indicate users' urge for seasonal precipitation forecast but it also shows that users need more information on the skill of forecast to better understand its usability.

When asking about how the outlooks were used by the farmers, they listed the following activities: planning the time of harvest, the threshing, in feed preparation and growth regulation of cereals. A few respondents indicated that they were rather following the forecasts than using those in everyday work. 29% of repliers have changed their operations based on the seasonal outlooks, especially the time of the harvests while 45% haven't changed their normal course of actions although they followed the forecasts.

Most of the respondents (93%) indicated their willingness to continue the piloting during the year 2020. From the listed indices to be included in the pilot 2020, the soil moisture and wind speed were the most wished. Respondents also expressed their wishes for the following improvements: detail the forecasted conditions in the text description, add the variation of forecasted indices during the past years and the 30-year climatology, and include risk analysis for various regions. Although it was not possible to fulfil all the requirements within the running project, improvements on the included forecast text have been done and the soil moisture has been included for the 2020 pilot. Further improvements are considered for the future pilots.

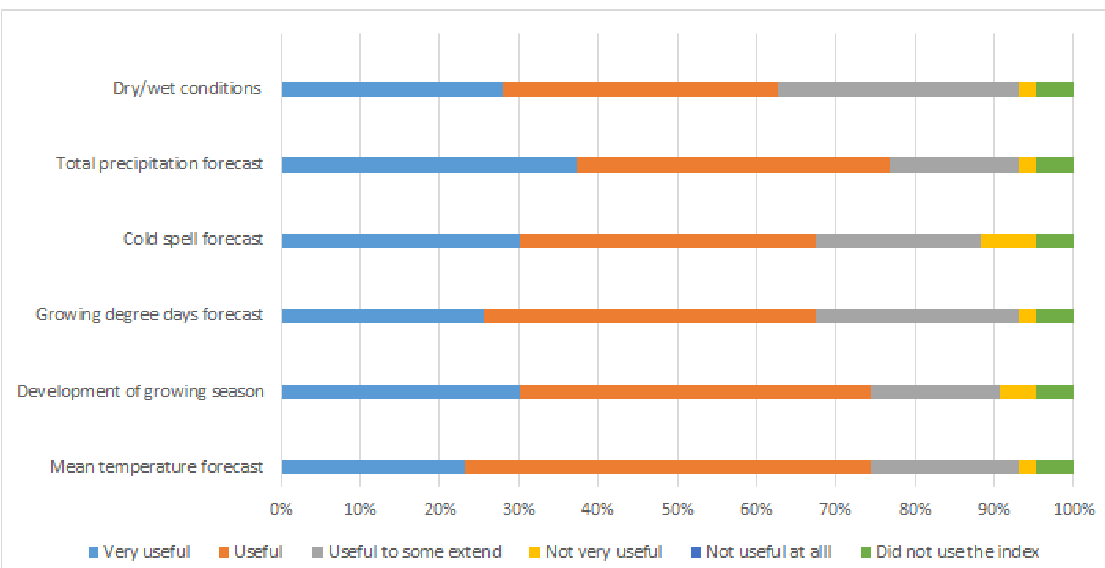

Figure 22. Rated values of level of benefit for the tested seasonal climate indices by the farmers based on the survey following the summer 2019 pilots.

To the feedback survey conducted with the managers of the ski resorts 4 replies were received. Users were asked for feedback on the usefulness, easiness of interpretation of the outlooks and how the forecasts were used. We also asked about their willingness of being involved in further pilots and the need for summer climate outlooks. According to respondents the climate outlooks were easily understandable. The most useful indices were considered the mean temperature and conditions for artificial snow production, these being used for assuring snow on slopes and producing extra amounts of snow. However, only 25% of the respondents (1 user) changed their plans based on the provided outlooks, 50% of the respondents couldn't say if they changed their activities in any way. It must be noted that half of the respondents run their business in Lapland and half in central and southern Finland. The snow conditions during the pilot season were exceptional and very different in southern and central Finland compared to Lapland. While an exceptionally large amount of snow was registered in Lapland, the winter was mild and thus snow free in the rest of the country, thus ski resort managers from various regions encountered different issues during the season. All of the respondents were willing to continue the testing of seasonal outlooks in the future; they also indicated their need for summer outlooks including temperature, precipitation and wind related forecasts. Also, 75% of the users showed an interest in sub-seasonal forecasts for the winter season.

8. Summary

In this project, we have concluded the first tests of applicability of seasonal forecasts for agriculture and the ski industry in Finland. Forecasts, a set of seasonal climate indices, were co-developed with farmers and ski resort managers. The indices developed for agriculture were tested during two pilot seasons: May-October 2019 and April-October 2020 (still on-going), those developed for ski resorts were tested during November 2019-April 2020.

Based on the reforecasts and for all variables, the forecasts for winter months were more skillful than for the summer months. In winter at least the forecasts of the first month were skillful for the whole country, sometimes also for the second month, at least in some parts of the country. The probabilities of lower and upper terciles were often more reliable than forecasts of middle tercile. The probable reason for this is that it is difficult to discriminate between no signal and the average conditions.

During the pilot season from summer 2019 and winter 2019/2020, the temperature-based forecasts fared well, while the precipitation forecasts fared less well. Snow forecasts were also of good quality.

The users were surveyed about the usefulness of the forecasts. Farmers found the climate outlooks very easily or easily understandable. All the tested seasonal indices were found useful, the total precipitation index, mean temperature and development of growing season being the most useful. For the ski resort managers, the most useful indices were the mean temperature and conditions for artificial snow production. How the perceived usefulness of forecasts transfers to the decisions made by the users is not so straight-forward. For farmers, almost one third of repliers have changed their operations based on the seasonal outlooks, especially the time of the harvests, while the half of repliers haven’t changed their normal course of actions although they followed the forecasts. For the ski resort managers, only one user of the four repliers changed their plans based on the provided outlooks, and half of the respondents couldn't say if they changed their activities in any way.

References

Bruno Soares, M.: Assessing the usability and potential value of seasonal climate forecasts in land management decisions in the southwest UK: challenges and reflections. Adv. Sci. Res. 14:175-180, https://doi.org/10.5194/asr-14-175-2017, 2017.

Bruno Soares, M. and Dessai, S.:Barriers and enablers to the use of seasonal climate forecasts amongst organisations in Europe, Clim. Change, 137, 89-103, 2016.

Buontempo, C., Hewitt, C. D., Doblas-Reyes, F. J., and Dessai, S.: Climate service development, delivery and use in Europe at monthly to inter-annual timescales, Climate Risk Management, 6, 1–5, https://doi.org/10.1016/j.crm.2014.10.002, 2014.

Doblas-Reyes, F. J., Garcia-Serrano, J., Lienert, F, Biescas, A. P., Rodrigues, L. R. L.: Seasonal climate predictability and forecasting: status and prospects. Wiley Interdiscip Rev Clim Change 4(4):245–268, https://doi.org/10.1002/wcc.217, 2013.

Gneiting, T., A. E. Raftery, A. H. Westveld, T. Goldman: Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. Mon. Weather Rev., 133, 1098–1118, doi:10.1175/MWR2904.1., 2005 http://journals.ametsoc.org/doi/abs/10.1175/MWR2904.1

Hersbach, H., Bell, B.; Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 Global Reanalysis. Q. J. R. Meteorol. Soc., 1-51, https://doi.org/10.1002/qj.3803, 2020.

Hyvärinen, O., Venäläinen, A., and Vajda, A.: Bias-adjusted seasonal forecasts of soil moisture for forestry applications in Finland, Adv. Sci. Res., 17, 23–27, https://doi.org/10.5194/asr-17-23-2020, 2020.

Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M. D., Manzanas, R., San-Martín, D., Cimadevilla-Alvarez, E., Cofiño, A. S., Gutiérrez, J. M.: The R-based climate4R open framework for reproducible climate data access and post-245 processing, Environmental Modelling and Software 111: 42-52, https://doi.org/10.1016/j.envsoft.2018.09.009, 2019.

Johnson, S. J. et al.: SEAS5: The new ECMWF seasonal forecast system, Geosci. Model Dev. 12, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019, 2019.

Klein Tank, A.M.G., Zwiers, F.W., Zhang, X.: Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation, climate data and monitoring WCDMP-No 72, WMO-TD No 1500, p 5, 2009.

Messner, J. W., G. J. Mayr, and A. Zeileis, 2016: Heteroscedastic censored and truncated regression with crch. The R Journal, 8, 173–181, doi:10.32614/rj-2016-012.

Vajda A. and Hyvärinen O.: Development of seasonal climate outlooks for agriculture in Finland, submitted to Adv. Sci. Res.

Weisheimer, A. and Palmer, T. N.: On the reliability of seasonal climate forecasts, J. Roy. Soc. Interface 11, https://doi.org/10.1098/rsif.2013.1162, 2014.

Applications of seasonal forecasts for hydropower

(Alexey Karpechko, Jaakko Ikonen)

1. Snowmelt driven streamflow forecasts and hydropower operations

Streamflow prediction consists of simulating complex hydrological processes that lead to the transformation of precipitation into runoff through various types of hydrological models, ranging from simple conceptual models to complex physically-based models. In snowmelt-dominated watersheds, the majority of streamflow occurs during the spring snowmelt period. Uncertainties and errors in simulated snow accumulation, melt timing and melt rate can consequently cause large errors in streamflow response simulations and predictions. It is widely recognized that consideration of uncertainty in snowmelt timing and melt rate is critical for both research and operational modelling. Errors in snowmelt timing and melt rate simulations are one of the most common and largest sources of streamflow predictions errors in snowmelt-dominated watersheds.

Kemijoki Oy is the most important producer of hydropower and regulating power in Finland. Kemijoki Oy operates 16 hydropower plants and 4 regulated reservoirs in the Kemijoki watercourse area. Currently hydropower production accounts for approx. 19% of total energy production in Finland, which whilst on-par with European averages is quite low when compared to the neighbouring Scandinavian countries. For example, in Sweden hydropower accounts for nearly half of the country's electrical energy production. As with other hydropower operations elsewhere, the river's natural conditions during different seasons of the year are the most important factor affecting the planning of hydropower production. In addition to fluctuations in natural conditions, hydropower plant operations planning involves other holistic considerations such as energy consumption demand and the many users of the rivers, for example bobbers, fishermen and boaters.

Spring snowmelt driven inflows to hydropower reservoirs (typically from mid-April/early-May to the end of June) accounts for between 55 and 70% of the total annual inflows to reservoirs in the Kemijoki watershed. This means that most of the annual water resources available for hydropower production is only available to producers during this period. It is therefore essential to have carefully planning and reservoir management schemes in place well before the onset of the peak inflow volumes. Reservoir management is important as the energy demand is out of phase with the natural availability of the water resources; typically, demand is higher during the colder months when the inflows are lower and vice versa. Therefore, hydropower producers need to redistribute the availability of these resources from the spring inflow periods to other times of the year when electricity demand is higher i.e. during the six months of colder winter, while maintaining the balance between a sufficiently large volume of water for optimal production and enough remaining capacity for safe flood risk management. The typical strategy for hydropower operators is to have reservoirs at around 90% capacity at the end of the spring inflow period which is then ideally maintained until the beginning of winter. To achieve this, operators require reliable seasonal forecast information to help them in planning the operations both leading up to and during the spring high inflow period.

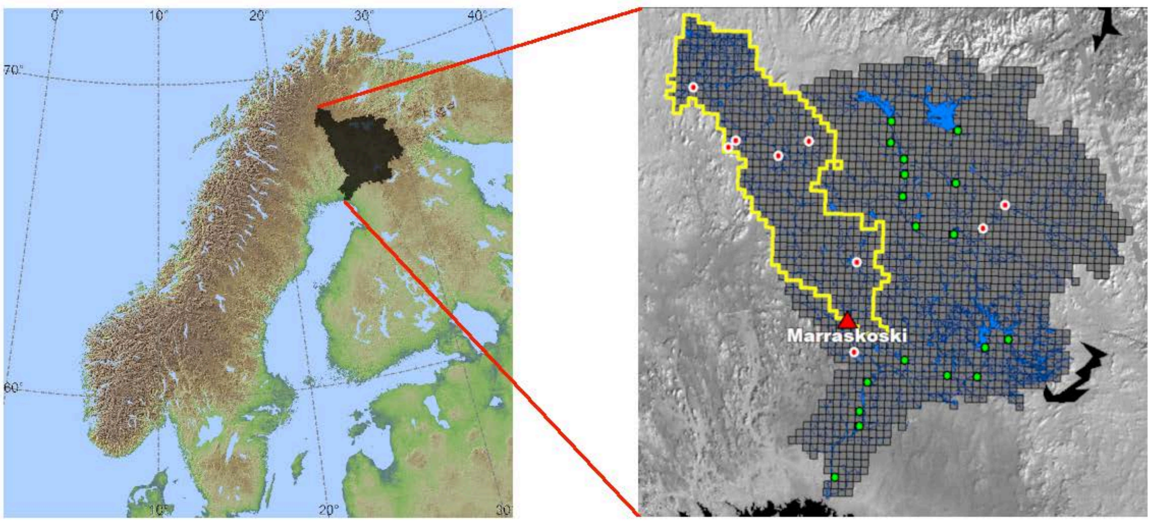

Figure 1 Location of the Kemijoki watershed and the study basin Ounasjoki basin. The green circles indicate the locations Kemijoki Oy hydropower plants and streamflow regulation structures. Red/white circles indicate the location of natural unregulated water level / streamflow observation stations operated by SYKE (the Finnish Environment Institute). Streamflow observations from the Marraskoski observation station are used assess hydrological model performance.

This report focuses on case studies assessing the accuracy of spring season, snowmelt driven streamflow ensemble hindcasts in Northern Finland. The accuracy of streamflow ensemble hindcasts are assessed in an unregulated tributary sub-basin (Ounasjoki Basin) of the Kemijoki watershed (Fig.1). The main water course of the Kemjoki watershed is 550 km long and is the longest river in Finland. It runs through municipalities of Kemijärvi and Rovaniemi before reaching the Gulf of Bothnia at Kemi. The Kemjoki watershed has a subarctic climate with short, mild summers and very cold and snowy winters. The average (30 years) annual temperature is just above freezing and snow cover generally lasts from mid-October to mid-May, with an average of 213 days with snow cover per annum. The coldest months are January and February with average temperatures ranging from -10 to -14°C. Although precipitation is relatively light during April and May, annual high runoff peaks are observed during this period due to snow melt.

Three recent and hydrologically unique spring snowmelt seasons (2018, 2019 and 2020) are assessed by driving a hydrological model with seasonal meteorological forecasts performed by an operational seasonal forecast system SEAS5 by the European Centre for Medium-range Weather Forecasts (ECMWF).

The hydrological model used in this study, FMI (Finnish Meteorological Institute) Hydrological Operations and Predictions System (HOPS), requires daily precipitation and evapotranspiration sums as well as daily average 2-metre air temperature fields as forcing input data.

2. Meteorological forcing and bias correction

Meteorological fields needed for driving hydrological model are daily means of 2-metre temperature, and daily sums of total precipitation and evaporation. These fields are obtained from seasonal forecasts performed by operational seasonal forecast system SEAS5 by European Centre for Medium-range Weather Forecasts (ECMWF). Forecasts are initialized at first day of each month and run for seven months. For this application, forecasts initialized on 1 April are employed. The fields are interpolated on regular 0.5°×0.5° grid in latitude and longitude.

In order to be applicable for impact studies, forecast outputs usually require bias correction to remove the difference between modelled and observed climates. This difference, if not removed, may make outcomes of impact calculations unrealistic. For 2m temperature bias correction we first calculated the mean difference between SEAS5 and ERA-5 over the control period at each grid point and each forecast day. Thereafter the difference at each grid point was smoothed with a 15-day moving average. Finally, the difference was removed from the forecasts. The control period is 1981-2017. The bias corrected procedure was applied to the forecasts for years 2018, 2019 and 2020 which were used for the hydrological modelling.

For daily total precipitation and evaporation sums, the mean ratios between SEAS5 and ERA-5 were calculate instead of the difference; otherwise the bias correction procedure is the same as that for the temperature.

A more elaborated quantile mapping method was also tried to remove the bias. In quantile mapping method the forecast values are mapped in the observed space, i.e. the bias correction depends on the forecasted value. However, we did not find a noticeable difference between the mean bias correction and quantile mapping and therefore the mean bias correction was used in the study.

Before moving to the results of hydrological modelling, we first quantify the role of bias correction for improving the forecasts. For this purpose, the forecasts from the control period 1981-2017 were used. Each forecast from this period was bias corrected using the procedure outlined above; however, this particular year was not used for calculating the biases to avoid artificial skill improvement. The mean RMS forecast errors for this period are shown in Figure 2, Figure 3 and Figure 4 for the first three forecast month.

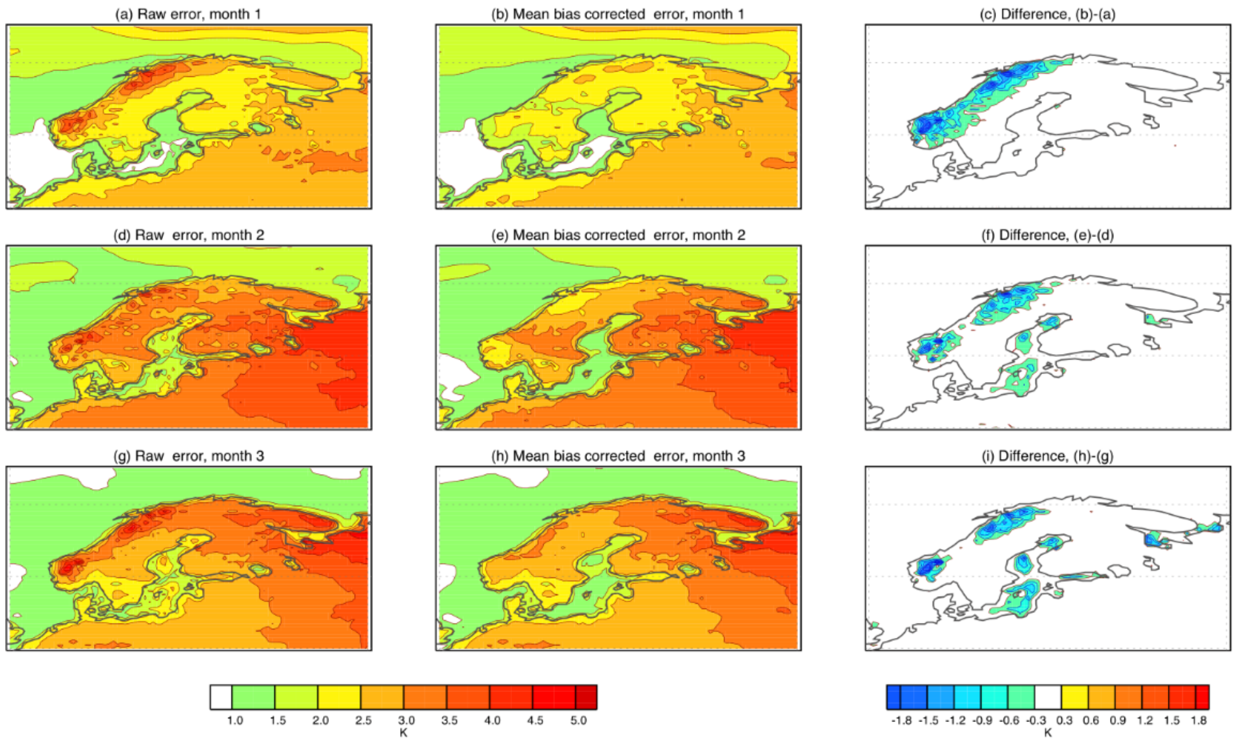

Figure 2. Mean root mean square errors of 2-metre temperature for the forecasts initialized on 1 April during 1981-2017. Shown are errors for (a,b) April, (d,e) May and (g,h) June for (a,d,g) raw and (b,e,h) bias corrected forecasts. Differences between bias corrected and raw forecasts are shown in (c,f,i) for April, May and June correspondingly.

As expected, RMS error grows, in general, with forecast time. For 2-metre temperature (Fig.2) the largest absolute forecast error is over northwestern Russia and Scandinavian maintains. The bias correction reduces forecast error most effectively over Scandinavian mountains and also over Baltic and White Seas during 2nd and 3rd forecast months. RMS error for total precipitation (Fig. 3) is largest over Scandinavian mountains. Unlike in 2-metre temperature, the forecast error over the mountains is not reduced by the bias correction but is somewhat increased. The positive impact of the bias correction is mostly seen over Russia. Finally, in case of evaporation (Fig. 4), the RMS error is more spatially homogenous. The largest errors are found over ocean and also over Russia. The impact of bias correction on evaporation forecast error is mostly negligible although some reduction of forecast error can be found in limited areas over Baltic Sea.

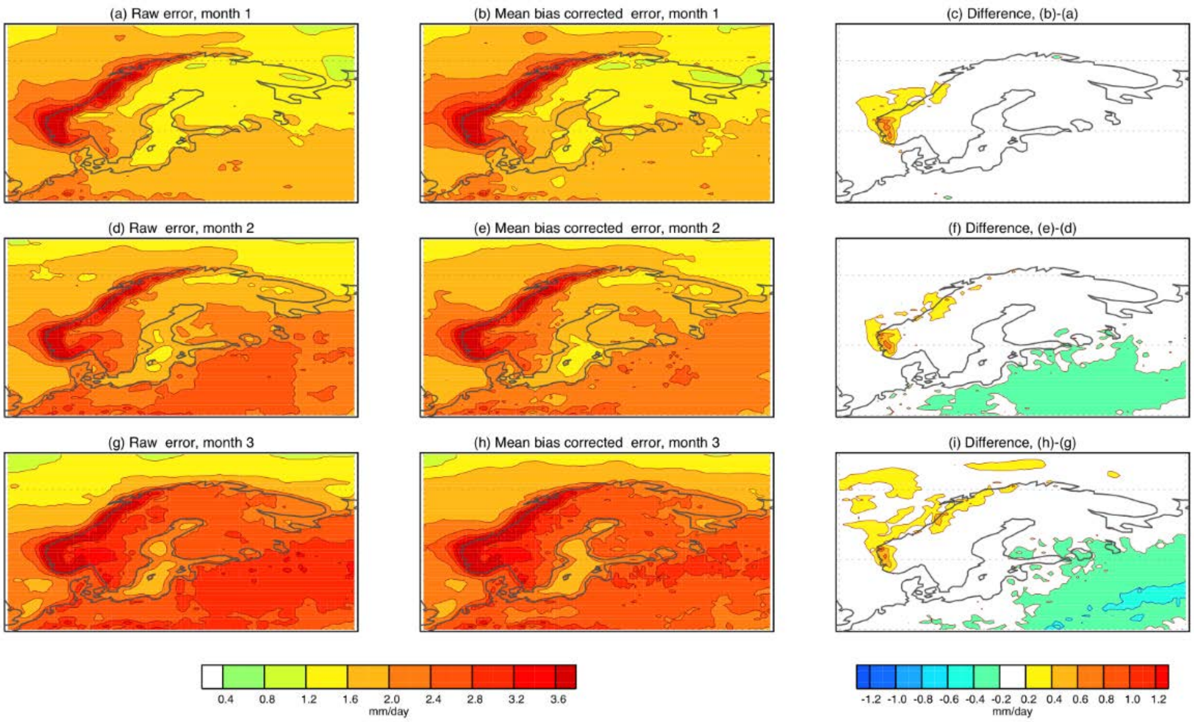

Figure 3. The same as Figure 1 but for total precipitation.

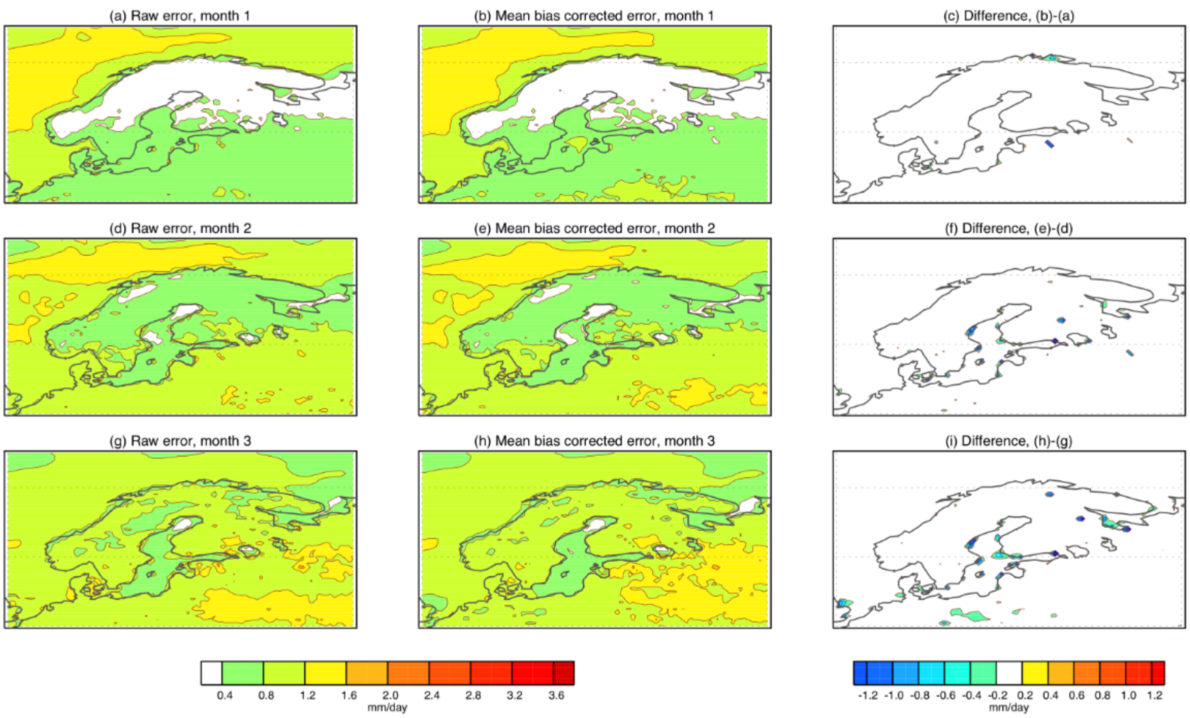

Figure 4. The same as Figure 1 but for evaporation.

3. Hydrological modelling and streamflow simulation

FMI's Hydrological Operations and Predictions System (HOPS) consists of distributed and modified version of the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash, 1995) coupled with a modified version of the SNOW-17 temperature index snow accumulation and ablation model, a soil temperature model based on Rankinen et al. 2004 and an in-house developed distributed routing model accounting for overland, and channel flow retention and attenuation, based on hydrographic properties. The parameters for the SAC-SMA model are derived from a semi-physical a-priori parametrization scheme introduced by Koren et al. 2000. Satisfactory simulation results have been achieved by running the modelling system without any calibration (apart from snowmelt base temperature and parameters governing snowmelt rates) or deviation from the a-priori parametrization. This allows rapid model implementation even in areas where hydrological reference data is either sparse or not available.

The HOPS model is used operationally by FMI to provide daily hydrological nowcasts and 10-day deterministic forecasts for Kemijoki Oy over the entire Kemijoki watershed. The first seasonal ensemble forecasts tests were conducted during the spring of 2020 with the aim of providing as much lead time as possible to hydropower operators due to unusually large volume of snow observed over the watershed during the winter of 2019-2020. The results of the preliminary seasonal ensemble forecast for the snowmelt period were however rather inconclusive and it was difficult to quickly assess the usefulness of such forecasts. Therefore, it was decided to conduct a larger snowmelt season ensemble forecasting accuracy assessment with the latest three snowmelt seasons (2018, 2019 and 2020) with the aim of attempting to conclude an approach that could prove useful in interpreting future snowmelt season forecasting results to the end user, Kemijoki Oy.

For the purposes of this report, the performance of the modelling system was evaluated by running the model with historical meteorological forcing (daily precipitation sum, average daily temperature and daily evapotranspiration sum) between the years 2017-2020, after an initial cold start period of 2016-2017. Specifically, the model's performance was evaluated by ingesting ERA-5 Land meteorological data instead of in situ observations in order to maintain consistency, especially with the snow melt model's parametrization, to the seasonal meteorological dataset used to drive the seasonal ensemble forecasting as closely as possible. Simulated streamflow is compared to observed streamflow at the outlet point of the Ounasjoki basin, Marraskoski

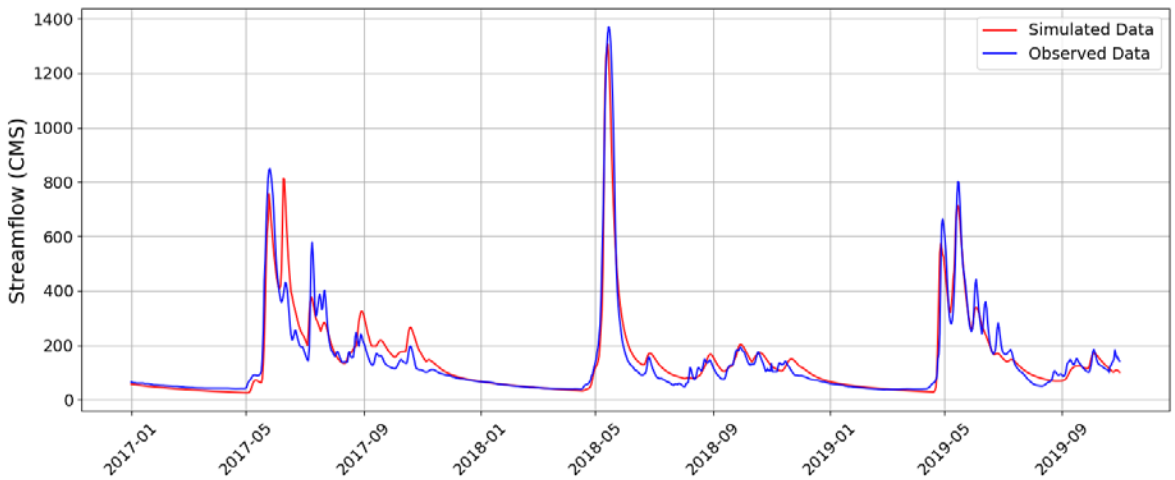

Figure 2. HOPS model validation runs hydrograph at the Marraskoski observation station.

Overall, the HOPS model validation period simulation for the Ounasjoki basin can be considered as satisfactory, if not good. The Nash-Sutcliffe coefficient of efficiency was 0.902, the correlation between observed and simulated discharge was 0.901 while the RMSE was 53 m^3/s and the Mean Average Error was 30 m^3/s. The shape of the simulated hydrograph matches the observed hydrograph very well (Fig. 5). The model tends however to somewhat overestimate late autumn streamflow whilst underestimating post spring peak streamflow. Both model tendencies are however very likely due to deficiencies in forcing evapotranspiration data; i.e. overestimation during early summer and underestimation during late summer and autumn. These issues however have very little bearing on snowmelt peak flow and timing simulations and therefore have a negligible effect on this study's focus. Winter period low flow months are generally simulated well by the model. Validation results for the year 2020 were omitted due to very clear problems in winter period low flow observations data. Low flow observations for the winter period of 2019-2020 were nearly quadruple of those observed since the beginning of observations (1971) at the Marraskoski station. Including these in model validation would introduce a clear bias and skew validation results. The suspected errors in streamflow observations during this period will be revisited later in this report.

4. Snowmelt period hydrological ensemble hindcasts

For this report, two types of ensemble seasonal streamflow hindcasts are run for the spring snowmelt periods of 2018, 2019 and 2020; 1) Un-adjusted ensemble hindcasts and 2) Bias-adjusted ensemble hindcasts. Both types of hindcasts use ECMWF SEAS5 meteorological forecast data (as explained in previous chapters) as a basis to drive the HOPS model streamflow simulations. Each hindcast is run with a 90-day lead time from April 1st to June 29th covering the snowmelt driven stream peak flow period.

The setup procedure for each hindcast consists of the following steps:

1. Initial HOPS model state parameter conditions are extracted from HOPS validation runs' conditions corresponding to those of the ensemble hindcast forecasts' first valid date -1 day. This is repeated for each year and the same initial conditions are used for both un-adjusted and bias-adjusted ensemble hindcast runs.

2. Each ensemble hindcast is run for a period resulting in a "lead-time" of 90-days for each year and type of ensemble hindcast type run, mimicking a 90-day forecast issued on April 1.

3. Each ensemble hindcasts' streamflow predictions are compared to observed streamflow at the outlet of the Ounasjoki basin (Marraskoski). For each hindcast 5 critical aspects of the ensemble streamflow forecast are assessed; 1) Difference in onset of snowmelt streamflow in days, 2) Difference in peak streamflow timing in days, 3) Difference in peak streamflow volume as a percentage to observed, 4) Difference in accumulated streamflow volume as a percentage to observed and 5) Difference in number of days considered "high streamflow" days. In the case of the Ounasjoki basin high streamflow days are considered as days where simulated streamflow exceeds 300 m^3/s. This is direct measure of the steepness of the peak streamflow hydrograph; a higher number indicating a shallower hydrograph shape and vice versa.

A range of observed spring snowmelt season streamflow derived from 50-years of observation is also included for each ensemble hindcast simulation comparison to visually inspect the skill of the hindcasts ensembles to distinguish season specific conditions from climatological averages.

4.1 Snowmelt period 2018

The winter of 2017-2018 saw exceptionally high snowfall during the early months of the snow accumulation period. In November 2017 there was an estimated 28% more snowfall compared to the previous ten-years. The following two months (December 2017 and January 2018) were on par with the previous ten-year average. The later part of the winter/spring, February through June, was considerably drier than the ten-year average. February 2018 and May 2018 in particular, were significantly drier and sunnier, with February experiencing colder than average temperatures while May significantly higher temperatures than on average.

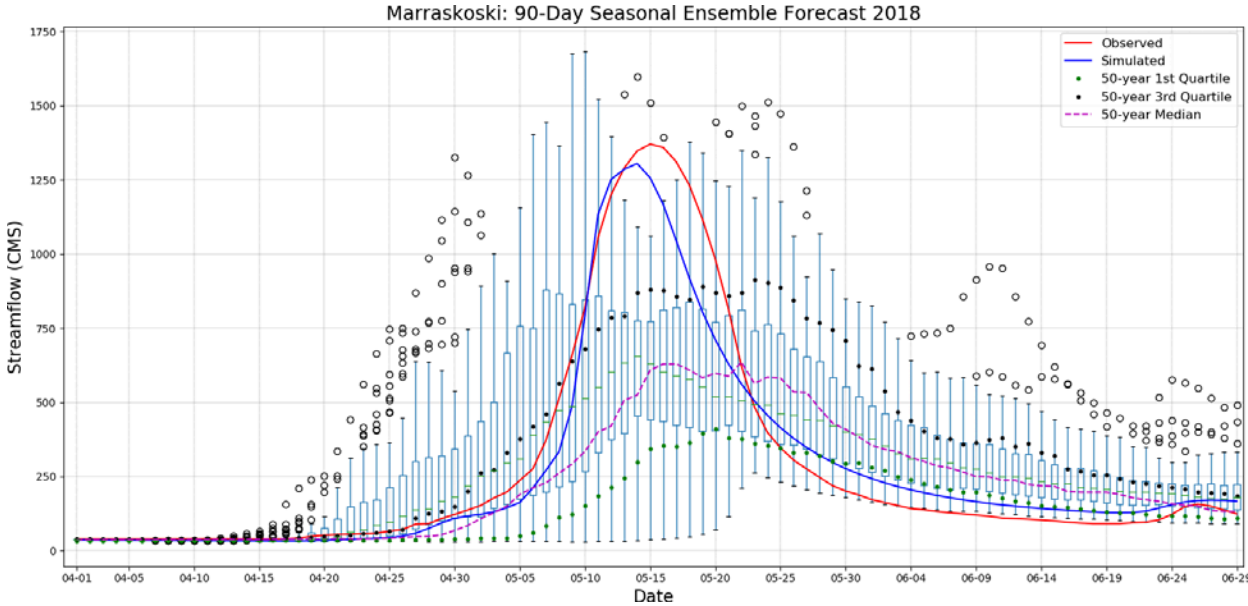

Figure 3. Seasonal (un-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2018, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

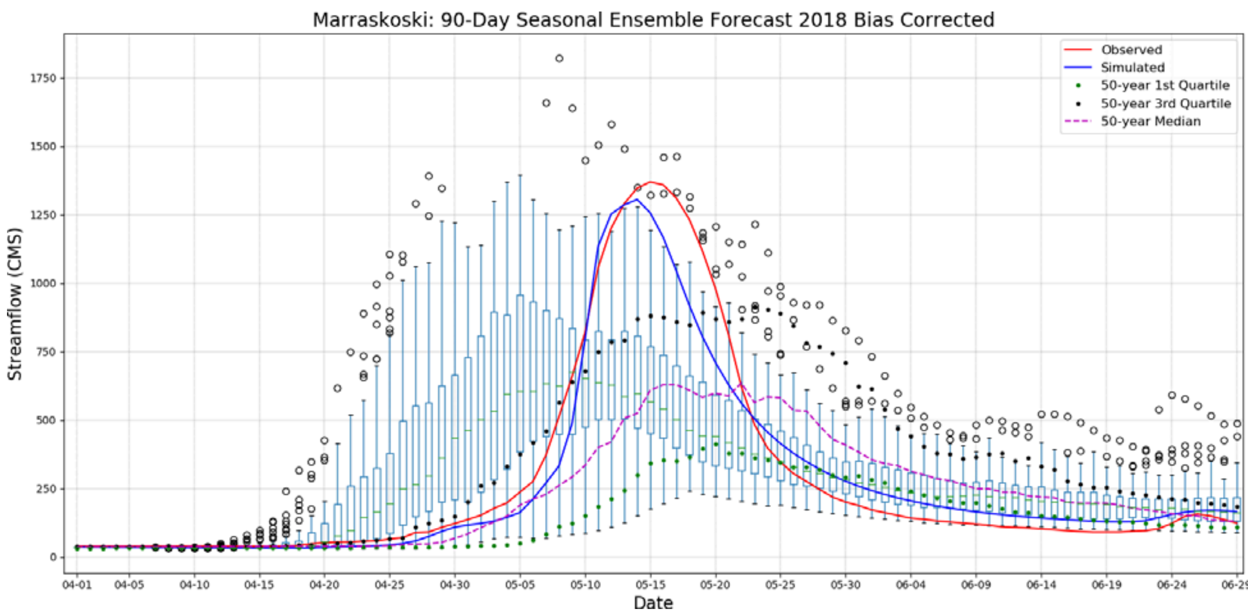

Figure 4. Seasonal (bias-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2018, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

Overall, the temperature during the snow accumulation period, between November 2017 and end of March 2018 was clearly cooler than the previous ten-years. The average temperature during the primary snowmelt period between mid-April and mid-May was however significantly higher than the previous ten-years. In particular, the Ounasjoki basin experienced significantly higher than average temperatures from the beginning of May until mid-May resulting in very rapid snowmelt. The onset of snowmelt began on par with the ten-year average, however the high rate of snowmelt from the beginning of May to mid-May resulted in a very sharp single runoff peak that resulted in significant flooding of the Ounasjoki river.

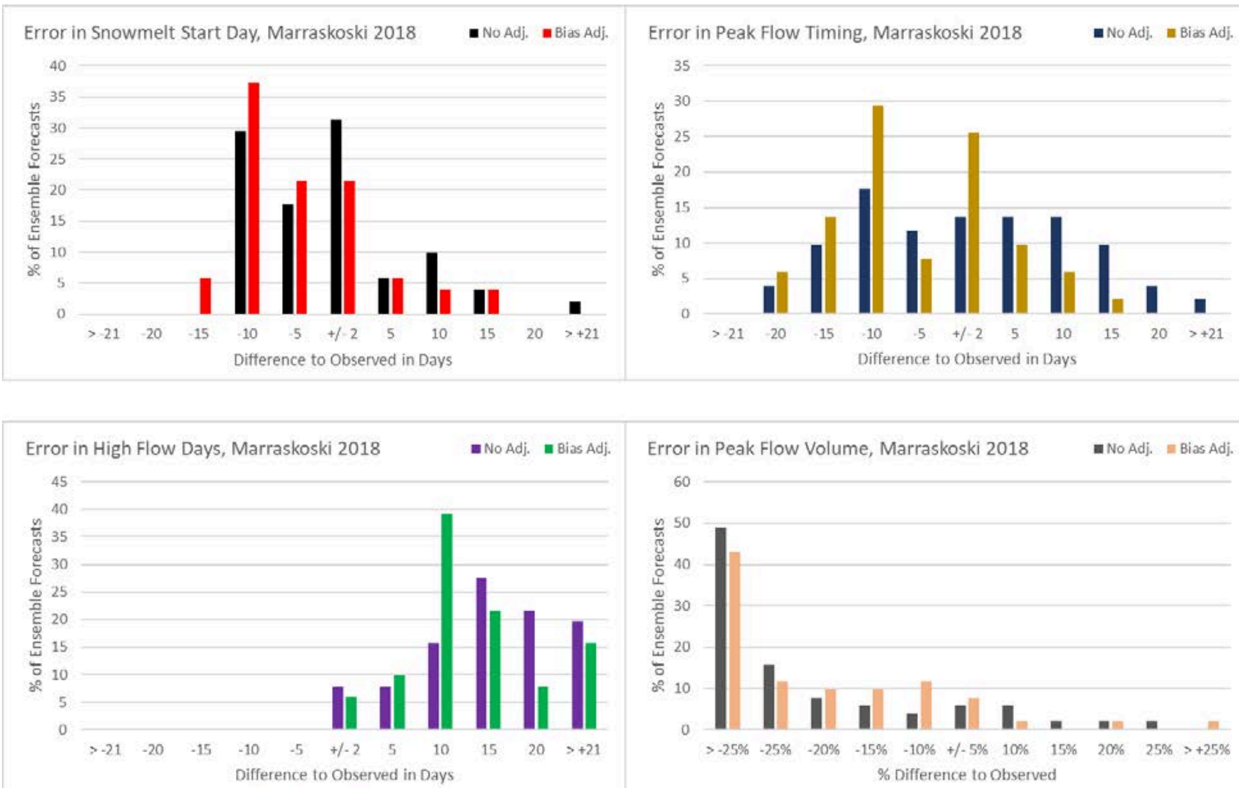

Figure 5. Percentage breakdown of key evaluation metrics of seasonal ensemble hindcast member errors, spring 2018.

Table 1. Categorical un-adjusted and bias-adjusted streamflow timing metrics compered to observed, spring 2018

	Diff. in Days:	> -21	-20	-15	-10	-5	+/- 2	5	10	15	20	> +21
No Adj.	Snowmelt Start	0	0	0	29.4	17.6	31.4	5.9	9.8	3.9	0	2
	Peak Flow	0	3.9	9.8	17.6	11.8	13.7	13.7	13.7	9.8	3.9	2
	High Flow	0	0	0	0	0	7.8	7.8	15.7	27.5	21.6	19.6
Bias Adj.	Snowmelt Start	0	0	5.9	37.3	21.6	21.6	5.9	3.9	3.9	0	0
	Peak Flow	0	5.9	13.7	29.4	7.8	25.5	9.8	5.9	2	0	0
	High Flow	0	0	0	0	0	5.9	9.8	39.2	21.6	7.8	15.7

Table 2. Categorical un-adjusted and bias-adjusted streamflow volume metrics compered to observed, spring 2018.

	Diff in %:	> -25%	-25%	-20%	-15%	-10%	+/-5%	10%	15%	20%	25%	> +25%
No Adj.	Peak Flow	49	15.7	7.8	5.9	3.9	5.9	5.9	2	2	2	0
	Accum. Flow	0	0	2	5.9	7.8	29.4	15.7	15.7	3.9	9.8	9.8
Bias Adj.	Peak Flow	43.1	11.8	9.8	9.8	11.8	7.8	2	0	2	0	2
	Accum. Flow	0	0	0	3.9	11.8	31.4	13.7	11.8	9.8	9.8	7.8

In general, both the un-adjusted and bias-adjusted seasonal ensemble streamflow hindcasts for 2018 predict a very wide range of individual streamflow volumes and peak times as was expected (Fig. 6 and Fig. 7). The un-adjusted ensemble median as well as the daily 1st and 3rd streamflow quartiles bear a striking resemblance to the corresponding 50-year observations, with only a slight deviation towards an earlier than usual snowmelt onset and streamflow peak time. This visible tendency to predict an earlier than observed snowmelt period is somewhat increased with the bias-adjusted ensemble runs. Since the snowmelt season was in terms of onset and peak flow timing quite close to the long-term average, the 50-year observational median and range of streamflow fits well with both the HOPS simulated validation run and observed streamflow. Approximately 65% of bias-adjusted hindcast ensembles predict an earlier than observed snowmelt onset time, whereas only 47% of un-adjusted predict an earlier than observed snowmelt onset time. Further, the majority (31%) of un-adjusted ensemble members predict a snowmelt onset difference to observed of only +/-2 days (Fig. 8 and Table 1).

A clear benefit of meteorological bias-adjustment can however be observed in that the individual predicted peak flow timing difference to observed peak flow is reduced and the spread of peak flow occurrences are less varied. Approximately 78% of the ensemble members of the bias-adjusted hindcast exhibit a +/-10 day difference to observed peak flow timing, whereas the in the un-adjusted hindcast the percentage of peak flow occurrences for the same period is 70%. Another, but less obvious and slightly less meaningful in terms of hydropower operations, benefit from bias-adjustment is the decrease in high flow (> 300 m^3/s) days difference to observed. Nearly 40% of the ensemble members have a difference to observed of only +5-10 days, indicating a that a healthy majority of ensemble members produce a rather sharp hydrograph, with a fast rising and declining limb as was observed. The un-adjusted ensemble hindcast noticeably spreads the hydrograph base over many more days, with approx. 69% of ensemble members having a more than 10 days error to observed high flow days total (Fig. 8 and Table 1).

The error in peak flow volume however for both hindcast runs is quite significant, with both hindcast runs producing a majority ensemble member error of >-25 % to observed peak flow volume. The bias-adjusted hindcast run does however decrease this error, but not very significantly (Fig. 8 and Table 2). The reason for the majority of both hindcasts runs in underestimating the peak flow volume is due to both hindcast types generally overestimating overall high flow days, thereby depleting their snowmelt reservoir fields too soon and too slowly to produce the observed sharp hydrograph, albeit the bias-adjusted hindcast runs tends to provide slightly better results. In terms of accumulated streamflow, there is little difference between the two hindcast run types. This is as be expected since most of the excess moisture in the watershed had already accumulated before the initiation of the ensemble hindcast runs and forthcoming precipitation during the simulation period has very limited impact on total accumulated runoff for the analysis period.

4.2 Snowmelt Period 2019

With regard to snowfall and accumulation the winter of 2018-2019 was fairly typical, however February and March 2019 saw significantly more snowfall than an average. April 2019 was on the other hand however abnormally dry. Overall the temperature during the snow accumulation period, between November 2018 and end of March 2019 was very typical and comparable to those observed during the past ten-years. Similarly, the average temperature during the primary snowmelt period between mid-April and mid-May was also typical and on-par with the previous ten-years. The estimated snowpack water content was only slightly lower than during the past ten-years. Since the average temperature during the primary snowmelt period between mid-April and mid-May was also typical, the snowmelt peak time also occurred within the "normal" timeframe of snowmelt. Average observed streamflow volumes for the snowmelt period (April to June 2019) were very typical and on-par with those observed during the previous ten-years. Regarding the shape of the runoff hydrograph the most dominating features are the relatively high double peak snow melt driven runoff peaks, and the rapid rate at which runoff volumes increase from very low volumes to relatively high volumes.

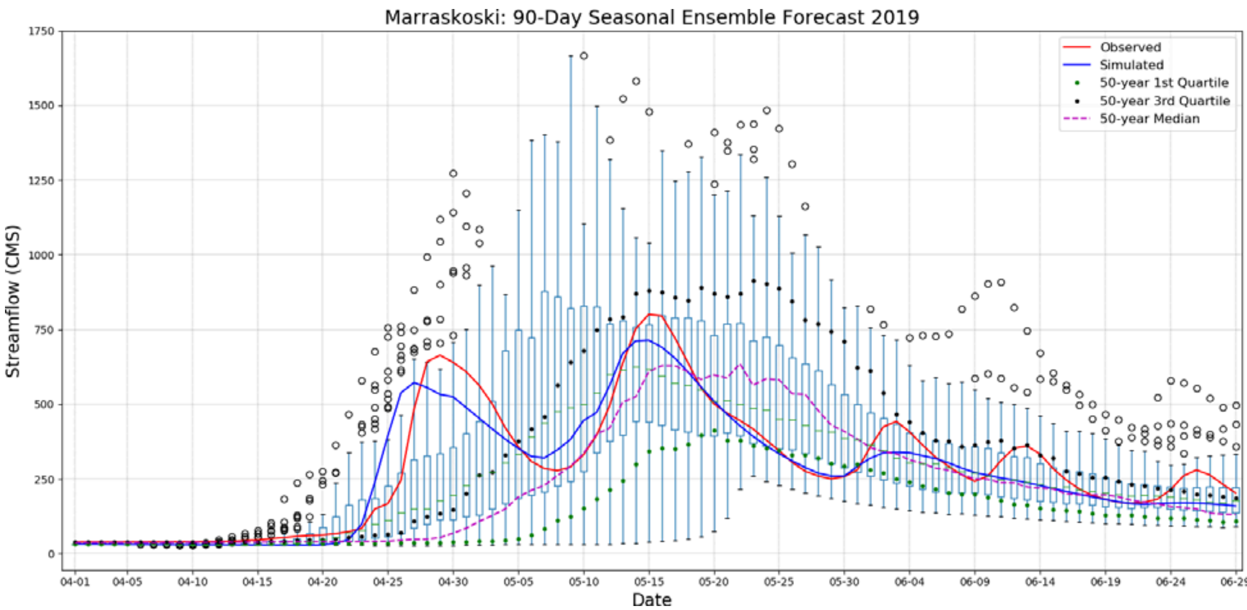

Figure 6. Seasonal (un-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2019, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

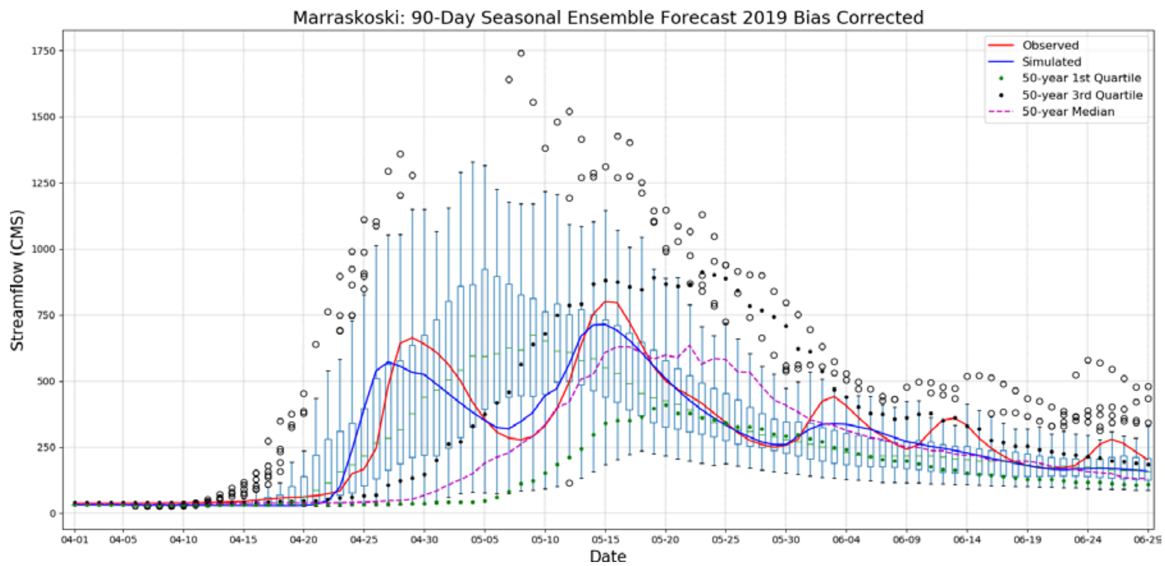

Figure 7. Seasonal (bias-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2019, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

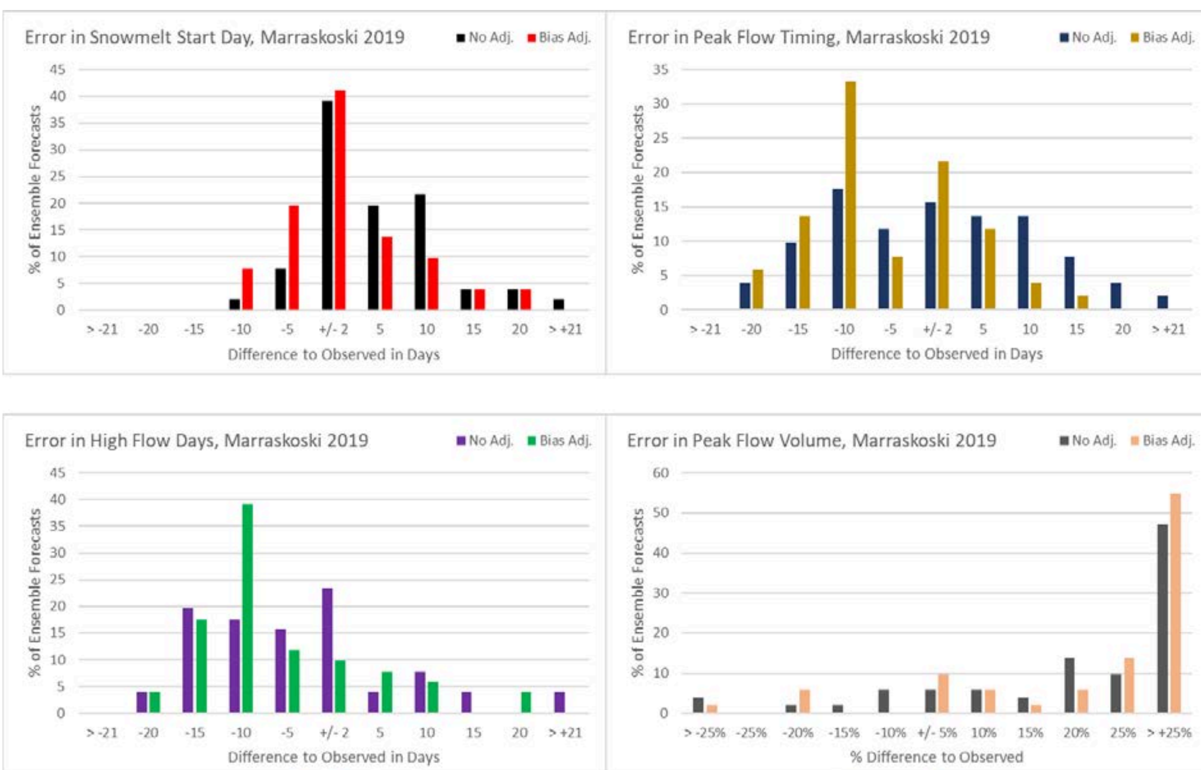

Figure 8. Percentage breakdown of key evaluation metrics of seasonal ensemble hindcast member errors, spring 2019.

Table 3. Categorical un-adjusted and bias-adjusted streamflow timing metrics compered to observed, spring 2019.

	Diff. in Days:	> -21	-20	-15	-10	-5	+/- 2	5	10	15	20	> +21
No Adj.	Snowmelt Start	0	0	0	2	7.8	39.2	19.6	21.6	3.9	3.9	2
	Peak Flow	0	3.9	9.8	17.6	11.8	15.7	13.7	13.7	7.8	3.9	2
	High Flow	0	3.9	19.6	17.6	15.7	23.5	3.9	7.8	3.9	0	3.9
Bias Adj.	Snowmelt Start	0	0	0	7.8	19.6	41.2	13.7	9.8	3.9	3.9	0
	Peak Flow	0	5.9	13.7	33.3	7.8	21.6	11.8	3.9	2	0	0
	High Flow	0	3.9	17.6	39.2	11.8	9.8	7.8	5.9	0	3.9	0

Table 4. Categorical un-adjusted and bias-adjusted streamflow volume metrics compered to observed, spring 2019.

	Diff in %:	> -25%	-25%	-20%	-15%	-10%	+/-5%	10%	15%	20%	25%	> +25%
No Adj.	Peak Flow	3.9	0	2	2	5.9	5.9	5.9	3.9	13.7	9.8	47.1
	Accum. Flow	0	2	5.9	7.8	21.6	25.5	13.7	7.8	5.9	3.9	5.9
Bias Adj.	Peak Flow	2	0	5.9	0	0	9.8	5.9	2	5.9	13.7	54.9
	Accum. Flow	0	0	3.9	13.7	17.6	29.4	13.7	5.9	7.8	3.9	3.9

As with the hindcast runs for 2018, in general, both the un-adjusted and bias-adjusted seasonal ensemble streamflow hindcasts for 2019 predict a very wide range of individual streamflow volumes and peak times as was expected (Fig. 9 and Fig. 10). However due to the observed multi-peaked and very variable nature streamflow during the snowmelt period of 2019, the ensemble hindcasts runs appear to predict the streamflow conditions quite well. Again, the un-adjusted ensemble median as well as the daily 1st and 3rd streamflow quartiles bear a striking resemblance to the corresponding 50-year observations, with only a slight deviation towards an earlier than usual snowmelt onset and streamflow peak time. As with the simulation runs for 2018, this visible tendency to predict an earlier than observed snowmelt period is somewhat increased with the bias-adjusted ensemble runs.

Contrary to 2018, and since the onset of snowmelt occurred somewhat earlier than usual in 2019, the tendency of the hindcast model runs to produce snowmelt prematurely provides a better fit with observations. This tendency is yet again exacerbated by the bias-adjusted hindcast ensemble runs. Approximately 27% of bias-adjusted hindcast ensembles predict an earlier than observed snowmelt onset time, whereas only 10% of un-adjusted predict an earlier than observed snowmelt onset time. Further, the majority (40%) of un-adjusted ensemble and the majority (41%) of bias-adjusted ensemble member forecasts predict a snowmelt onset difference to observed of only +/-2 days (Fig. 11 and Table 3).

As with 2018, the bias-adjusted peak flow times are more concentrated around the actual observed peak flow day. Approximately 41% of both the bias-adjusted and un-adjusted hindcasts only exhibit a +/-5 day difference to observed peak flow timing. Of the 51 bias-adjusted hindcast ensemble members 21% have a peak flow timing difference to observed peak flow time of merely +/-2 days, while 15% of the un-adjusted hindcast ensemble members have their peak flow within the same timeframe. Significantly, the bias-adjusted hindcast correctly predicts an early streamflow peak with 33% of forecasts predicting peak flow to occur up to 10 days before the observed peak. Although, this could be interpreted as an error, the majority of forecasts correctly predict both an early peak and a later peak flow event, the peak flow volumes are simply reversed, i.e. the lower peak actually occurred before the higher peak (Fig. 11 and Table 3).

Whilst the peak flow timing predicted by the bias-adjusted hindcasts is more accurate than with the un-adjusted hindcasts, bias-adjustment again (as with 2018) decrease the amount of high flow (> 300 m^3/s) days. A decrease in the number of high flow days of more than -10 days to observed is predicted by approx. 61% of ensembles. This results in shorter and sharper flow peaks. The same situation can be observed for the un-adjusted hindcast runs albeit to a lesser degree with 41% of forecast members having more than -10 high flow days when compared to observed streamflow (Fig. 11 and Table 3).

The error in peak flow volume however for both hindcast runs is, as with 2018 significant. A clear majority of both hindcast runs produce a difference of >+25 % to observed peak flow volume. Since the bias-adjusted runs tend to concentrate streamflow peaks, there are more ensemble members with higher than observed peak flow volumes than with the un-adjusted hindcast runs (Fig. 11 and Table 4). In contrast to 2018, the reason for the majority of both hindcasts runs in overestimating the peak flow volumes is due to both hindcast types generally underestimating overall high flow days, thereby depleting their snowmelt reservoir fields too rapidly. In terms of accumulated streamflow, there is little difference between the two hindcast run types, as was the case with 2018.

4.3 Snowmelt Period 2020

As with the winter of 2017-2018, the winter of 2019-2020 saw exceptionally high snowfall during the early months of the snow accumulation period. October 2019 saw extraordinarily large amounts of snowfall that due to the somewhat exceptionally cool temperatures remained until the onset of snowmelt in the spring of 2020. Although typically, the first snowfall events are generally observed in October, usually these accumulations tend to melt away before the onset of enduring snow cover. An enduring snowpack usually does not form until November or December. The winter of 2019-2020 was exceptional also in that temperature in Northern Finland remained constantly above freezing through-out the winter, while in southern and central Finland the low pressure from the south kept the weather warm and wet. Although many parts of the Northern Finland were 2-5° C milder than usual, these cannot be considered exceptional. Meanwhile through-out the winter of 2019-2020 precipitation levels in north Finland were 1.5 – 2 times higher than usual resulting in much higher snow water content than is usually observed. Although there was grave concern for record flooding in the spring of 2020 due to unusually large snow reservoir volumes, for the most part, although there was considerable flooding, record breaking floods did not occur. This is largely due to the considerably cooler than average temperatures between April and May 2020 coupled with relatively dry conditions. As with the winter of 2017-2018, snowmelt driven streamflow occurred in a single peak. This streamflow peak however occurred considerably later than during the previous ten-year average and the rate of snowmelt was also considerably slower than expected resulting an elongated peak streamflow curve.

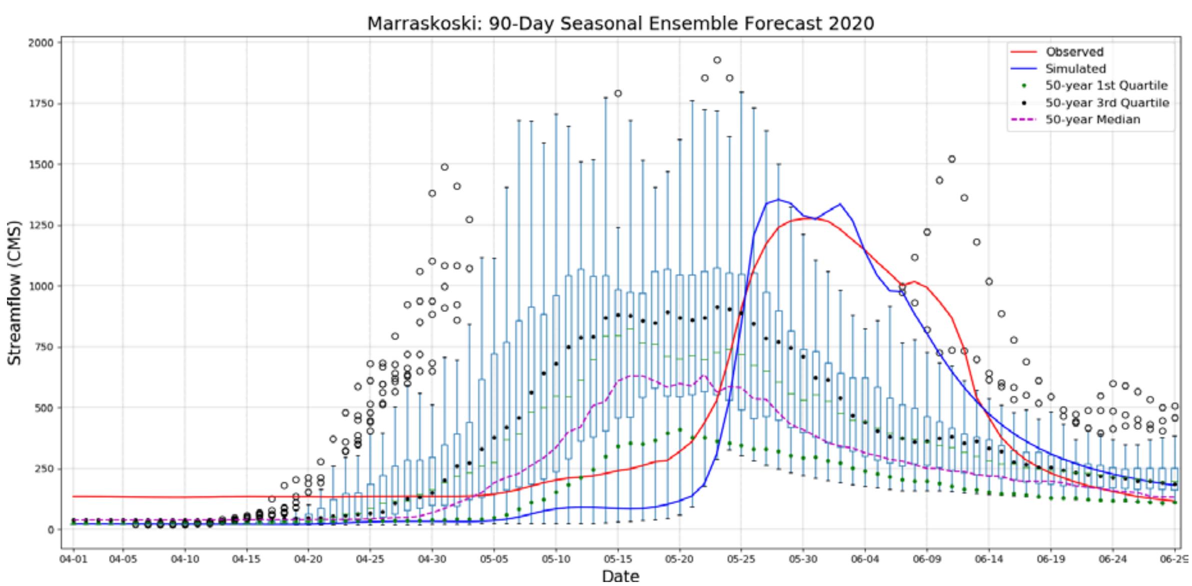

Figure 9. Seasonal (un-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2020, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

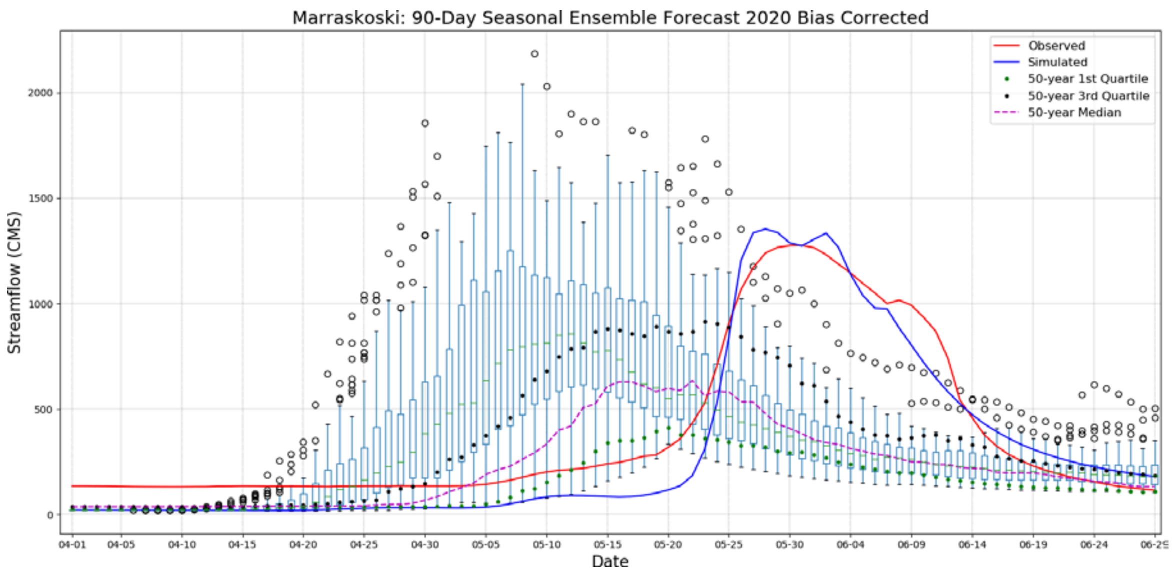

Figure 10. Seasonal (bias-adjusted) snowmelt period ensemble forecast member streamflow range, spring, 2020, HOPS validation run streamflow, SYKE observed streamflow and 50-year 1st quartile, median and 3rd quartlie streamflow observations.

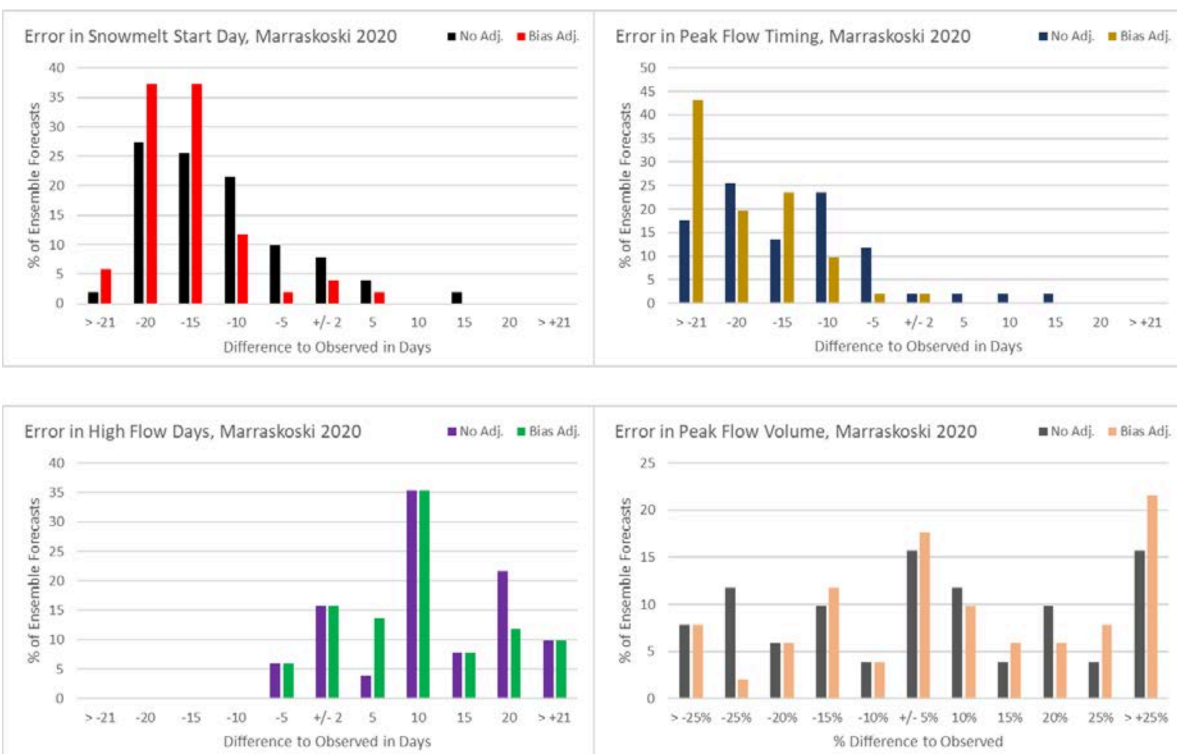

Figure 11. Percentage breakdown of key evaluation metrics of seasonal ensemble hindcast member errors, spring 2020.

Table 5. Categorical un-adjusted and bias-adjusted streamflow timing metrics compered to observed, spring 2020.

	Diff. in Days:	> -21	-20	-15	-10	-5	+/- 2	5	10	15	20	> +21
No Adj.	Snowmelt Start	2	27.5	25.5	21.6	9.8	7.8	3.9	0	2	0	0
	Peak Flow	17.6	25.5	13.7	23.5	11.8	2	2	2	2	0	0
	High Flow	0	0	0	0	5.9	15.7	3.9	35.3	7.8	21.6	9.8
Bias Adj.	Snowmelt Start	5.9	37.3	37.3	11.8	2	3.9	2	0	0	0	0
	Peak Flow	43.1	19.6	23.5	9.8	2	2	0	0	0	0	0
	High Flow	0	0	0	0	5.9	15.7	13.7	35.3	7.8	11.8	9.8

Table 6. Categorical un-adjusted and bias-adjusted streamflow volume metrics compered to observed, spring 2020.

	Diff in %:	> -25%	-25%	-20%	-15%	-10%	+/-5%	10%	15%	20%	25%	> +25%
No Adj.	Peak Flow	7.8	11.8	5.9	9.8	3.9	15.7	11.8	3.9	9.8	3.9	15.7
	Accum. Flow	2	11.8	17.6	13.7	21.6	25.5	3.9	3.9	0	0	0
Bias Adj.	Peak Flow	7.8	2	5.9	11.8	3.9	17.6	9.8	5.9	5.9	7.8	21.6
	Accum. Flow	2	7.8	21.6	17.6	21.6	21.6	3.9	2	2	0	0

The spring of 2020 saw both high peak streamflow volumes and a much later than average peak flow time in the Kemijoki watershed as a whole and in the study basin of Ounasjoki as well. The abnormally late spring snowmelt season causes large errors for both the un-adjusted and bias-adjusted seasonal ensemble streamflow hindcasts for 2020 (Fig. 12 and Fig. 13). This may partly be due to a longer lead-time to peak flow and/or due to the fact the neither hindcast type is able to accurately predict the meteorological conditions that occurred in general. Ignoring the obviously large errors, it can be observed that as with the other hindcasts runs for 2018 and 2019 the bias-adjusted ensemble runs tend to predict earlier snowmelt onset timing than the un-adjusted hindcast runs do. Further, and again, the un-adjusted ensemble median as well as the daily 1st and 3rd streamflow quartiles bear a striking resemblance to the corresponding 50-year observations, with only a slight deviation towards an earlier than usual snowmelt onset and streamflow peak time. Approximately 94% of bias-adjusted hindcast ensembles predict an earlier than observed snowmelt onset time, and 86% of un-adjusted predict an earlier than observed snowmelt onset time. Further, the majority (55%) of un-adjusted ensemble and the majority (37%) of bias-adjusted ensemble member forecasts predict a snowmelt onset difference to observed of equal to or over -15 days (Fig. 14 and Table 5). As can be seen from Figure 12 and Figure 13 there appears to be a clear base level discrepancy with streamflow observations for the low flow period, before the onset of snowmelt, as mentioned previously in this report. This however does not directly affect the analysis of the hindcast runs since the focus is on snowmelt onset and peak flows.

Nearly all the bias-adjusted hindcast run peak flow times occur significantly earlier than observed. The un-adjusted hindcast run only has 16% of forecasts with a +/-5 days error to observed peak flow time (Fig. 14 and Table 5). The difference between the number of high flow days (>300 m^3/s) for the bias-adjusted and un-adjusted hindcasts runs is negligible, however the difference of both to observed is considerable. A large majority of ensemble members for both hindcasts overestimate the number high flow days significantly (Fig. 14 and Table 5).

The error in peak flow volume for both hindcast runs is very variable. It is not possible to distinguish any clear trend or majority in peak flow volumes for either hindcast run (Fig. 14 and Table 6). Interestingly, in contrast to the other simulation years, both hindcast runs have a clear tendency to underestimate accumulated streamflow for the spring snowmelt period. This however is clearly due to the fact that the low streamflow observations appear to have a base level bias, and such can be ignored (Table 6).

5. Conclusions

Daily long-term ensemble streamflow forecasts covering the snowmelt period have been requested by the end-user of FMI's observations and forecast data, Kemijoki Oy, for hydropower optimization purposes and planning. With this in mind, a preliminary analysis of the potential of such forecasts to provide meaningful information to the end user as conducted. The main objective of this study was to assess the forecasting skill of seasonal spring snowmelt ensemble hydrological simulations in an unregulated basin of the Kemijoki watershed, namely in the Ounasjoki basin. Seasonal, ensemble hindcasts initiated on April 1st with a 90-day lead time were configured for the years 2018, 2019 and 2020.

The main findings are summarised as follows:

1) Seasonal hydrological forecast for snowmelt periods driven by meteorological ensemble forecasts tend to only slightly improve forecasts based purely on analysis of long-term

historical streamflow observations. This is likely particularly true if a reasonable means of interpreting the results of the ensemble forecast are not available.

2) Meteorological bias-adjustment tends to reduce the spread of ensemble streamflow forecasts thus reducing overall uncertainty. This however does not mean that the overall skill of the hydrological forecast is improved.

3) Meteorological bias-adjustment tends to (at least in this study) result in earlier snowmelt onset as well as sharper rising hydrograph limbs and recession curves.

4) Interpretation of seasonal hydrological forecasts for snowmelt periods should not be based on ensemble mean or other statistical metrics involving all ensemble members, but rather a means of categorisation is required. For example, mixing ensemble members with snowmelt driven peak flow occurring early in the analysis period with those having a late peak flow to calculate ensemble means as well as quartiles can lead to dilution of overall streamflow peaks as well as recession curves, since for each forecast snowpack water equivalent is generally not a replenishable reservoir field.

5) Ensemble spread whilst indicating larger forecast uncertainty, tends to also indicate larger forecast errors. This is potentially useful to users as the ensemble spread could be used as a measure of the forecast quality.

Looking forward, future studies need to address some form of ensemble member categorisation and assignment of probabilities to those in order to produce a more meaningful and easily interpretable seasonal snowmelt period forecast. Regarding the bias adjustment of meteorological seasonal forecast data, results from this study show that while the seasonal forecasts where bias adjustment was conducted were disappointing, the performance of the seasonal hydrological forecasts were in some respects improved. The methodology behind bias adjustment of meteorological data should possibly be revisited.

Further, perhaps a better understanding of how the performance of the different modelling components are affected by the initial conditions and lead time would shed more light on how to best approach the shortcomings of this study. Further development and testing along these lines are planned for the near future.

References

Burnash, R.J.C.,"The NWS river forecast system—catchment modeling", In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, Colorado, pp. 311–366, 1995.

Koren, V.I., Smith, M., Wang, D., Zhang, Z., "Use of Soil Property Data in the Derivation of Conceptual Rainfall-Runoff Model Parameters Proceedings of the 15th Conference on Hydrology", AMS, Long Beach, CA, pp. 103–106, 2000.

Rankinen, K., Karvonen, T., and Butterfield, D.: A simple model for predicting soil temperature in snow-covered and seasonally frozen soil: model description and testing, Hydrol. Earth Syst. Sci., 8, 706–716, https://doi.org/10.5194/hess-8-706-2004, 2004.