MVR applied to Statistical Downscaling for Prediction of Monthly Mean Land Surface Temperatures: model documentation.

R.E. Benestad

DNMI, December 15, 1998

Contents

1 Introduction

This report is a part of the DNMI KLIMA *Linear Statistical Downscaling Methods* series, documenting the construction and testing of linear statistical downscaling models. The focus of this report will be on multivariate regression (MVR) models. We have already discussed the Canonical Correlation Analysis (CCA) and Singular Vector Decomposition (SVD) techniques applied to the construction of linear models in the KLIMA technical reports numbers 28 and 30 respectively (*Benestad*, 1998a,c). One important part of the statistical models is that they are based on historical data, and it is therefore crucial that the data sets on which these models are based are free from serious errors. An evaluation of several predictor data sets used in the model calibration described here is discussed in *Benestad* (1998b).

The same notations as in the previous reports in the *Linear Statistical Downscaling Methods* series will be adopted here. The predictands will be referred to as Y and the predictors will be called X. The locations of the 24 stations analysed in this report are marked on the map in figure 1. In some of the sections, however, the stations with observations not up to date have been excluded in order to extend the analysis to the 1990s.

The first section derives the mathematical framework for the regression model, and is followed by the description of different regression models employing different predictor data sets. In this report, we define *best skill* as the prediction with the highest correlation score. An evaluation of the MVR models will be based on comparisons with the CCA models, as the SVD models discussed in *Benestad* (1998c) had systematically lower skill scores than the corresponding CCA models. However, we will also discuss some MVR model results based on the same EOF predictor combination as the optimal SVD models (referred to as SVDMVR models). The last section concludes the report with a discussion of the main results.

2 Multivariate Regression models

While the CCA and SVD models were based on maximizing correlations and covariance, the regression models aim to minimize the root mean square (RMS) error. The regression models are based on the least squares solutions to an inconsistent system:

$$Y = X\Psi - \zeta, \tag{1}$$

of n equations of q unknowns, where ζ represents a noise term. If the noise term is insignificant then the linear expression is satisfied by the normal

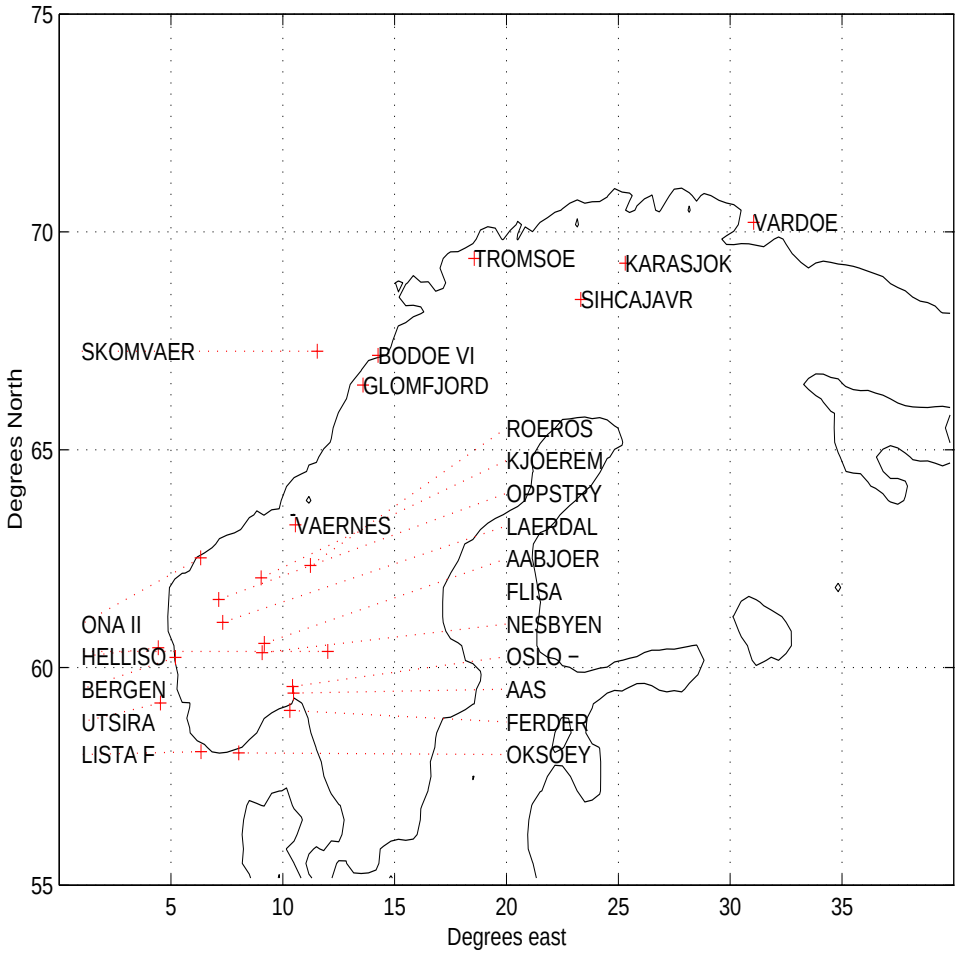

Figure 1: Map showing the location of the the stations (predictand locations) referred to in this report.

equations:

$$X^T X \Psi = X^T Y, \tag{2}$$

The matrix product $X^T X$ is invertible if the columns of X are independent, and we can express Ψ in terms of X and Y only (*Strang*, 1995, pp.156):

$$\Psi = (X^T X)^{-1} X^T Y. \tag{3}$$

Equation 1 may, however, involve significant noise levels and a "true" estimate of Ψ is then:

$$\Psi = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} X, \tag{4}$$

where $\Omega = E(\zeta \zeta^T)$ is the error-covariance matrix. One problem is that we only have an estimate of Ω if ζ is known ($\hat{\Omega} = \zeta \zeta^T$). Ω may furthermore be non-invertible. We can get around these problems by excluding the noise term from the analysis, and only attempt to predict the signal in Y, which we refer to as $\hat{Y}$.

$$\hat{Y} = \Psi X. \tag{5}$$

By applying PCA to the data and truncating to the kth leading EOF, we also remove noise in X and ensure that $(X^T X)^{-1}$ is invertible by writing the matrix in terms of its PCA products,

$$(X^T X)^{-1} = [(E_{(k)} \Sigma_{(k)} A_{(k)}^T)^T E_{(k)} \Sigma_{(k)} A_{(k)}^T)^{-1} = (A_{(k)} \Sigma_{(k)}^2 A_{(k)}^T]^{-1}.$$

Hence equation 4 can be expressed as:

$$\Psi = (A_{(k)} \Sigma_{(k)}^2 A_{(k)}^T)^{-1} X^T Y, \tag{6}$$

A Matlab script, *mvr.m*, estimates the regression model, Ψ according to equation 6 and then predicts the values of $\hat{Y}$.

3 SST Models

Two types of SST models will be discussed here. The first kind is a model calibrated with regional SSTs covering the area 10°W to 40°E and 55°N to 75°N, which will be referred to as the 'Nordic Seas model', although SSTs from the North Sea, the Barents Sea, the Norwegian Sea, Skagerrak, Kattegat, and the Baltic Sea also are included. The second model type covers a larger area, 90°W to 40°E and 15°N to 80°N, and is called the 'North Atlantic model', although the North Sea, the Mediterranean, the Black Sea,

the Labrador Sea, the Barents Sea, the Norwegian Sea, Skagerrak, Kattegat, the Baltic Sea, and Hudson Bay also are included.

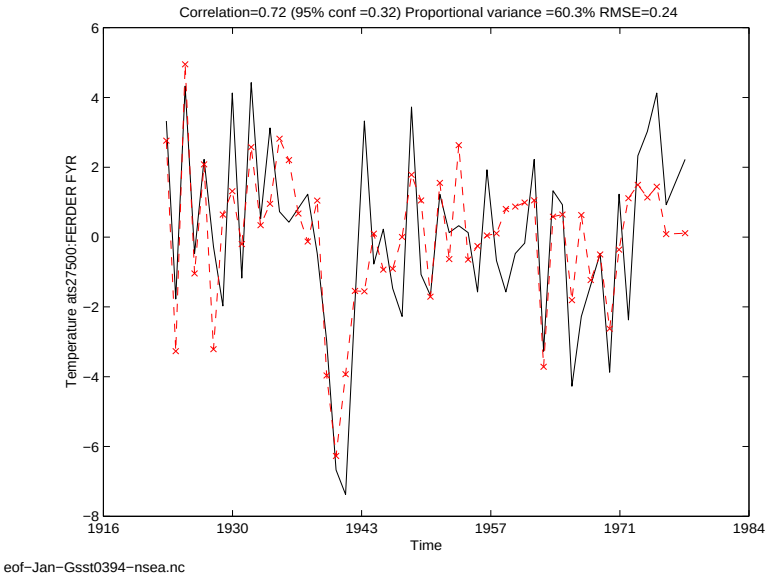

Figure 2: Time series of predicted January temperatures (dashed) for Ferder, employing the cross-validation method with GISST2.2 Nordic Seas SSTs, shown with the observations (black solid line).

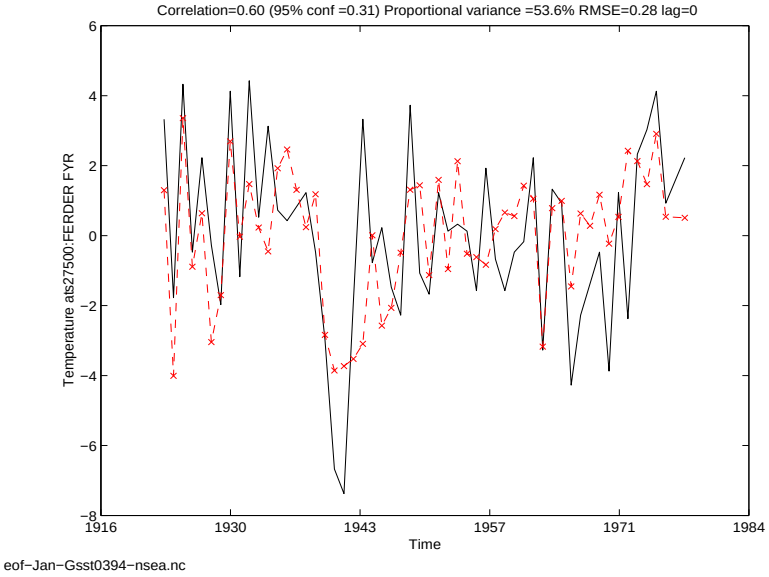

Figure 3: Same as figure 2, but using the same EOF combination as the corresponding optimal SVD model. The time series represents the January temperatures from Ferder fyr.

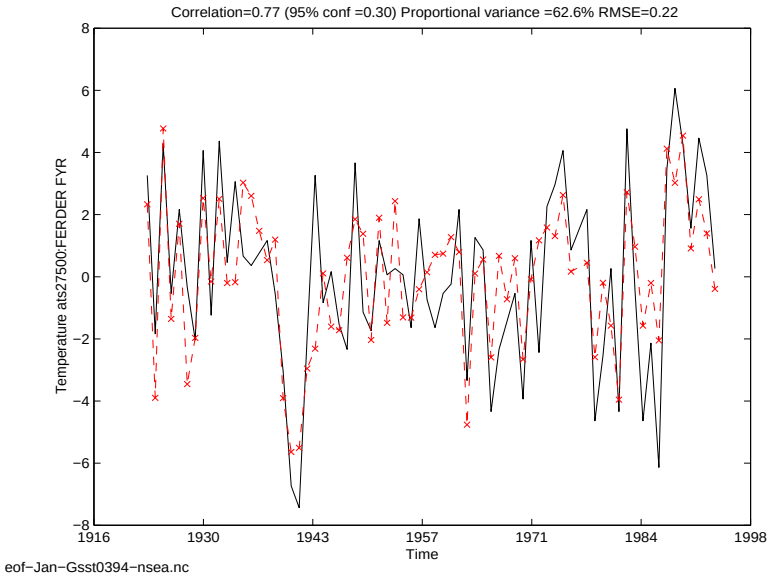

Figure 4: Same as figure 2, but using a longer time series than the corresponding optimal CCA model. The time series represents the January temperatures from Ferder fyr.

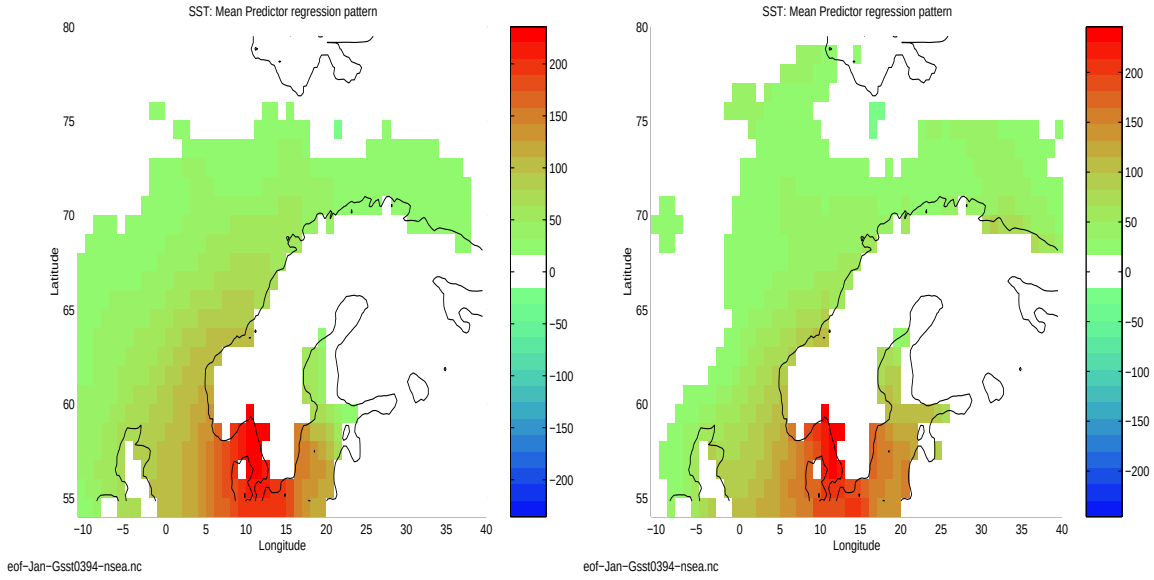

Figure 5: The mean January MVR GISST2.2SST pattern associated with Flisa land surface temperature. Left panel shows the optimal MVR model pattern while the right panel shows the SVDMVR model pattern.

- Similar predictor pattern and skill scores to the CCA results.

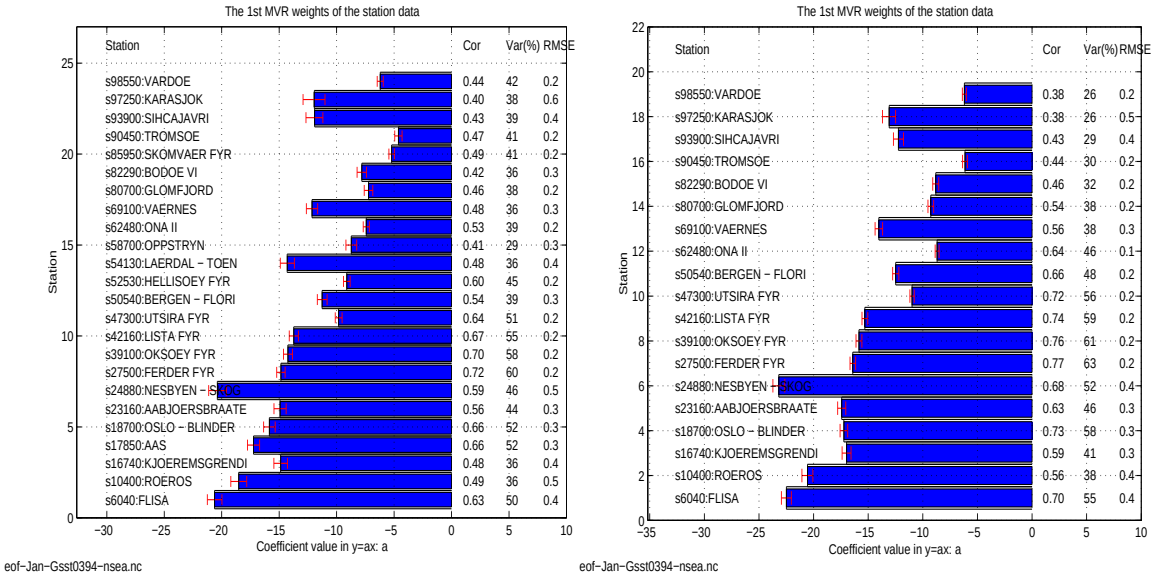

Figure 6: The mean January predictand weights of EOF1 shown in filled bars for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation results from the cross-validation analysis are given on the right hand side.

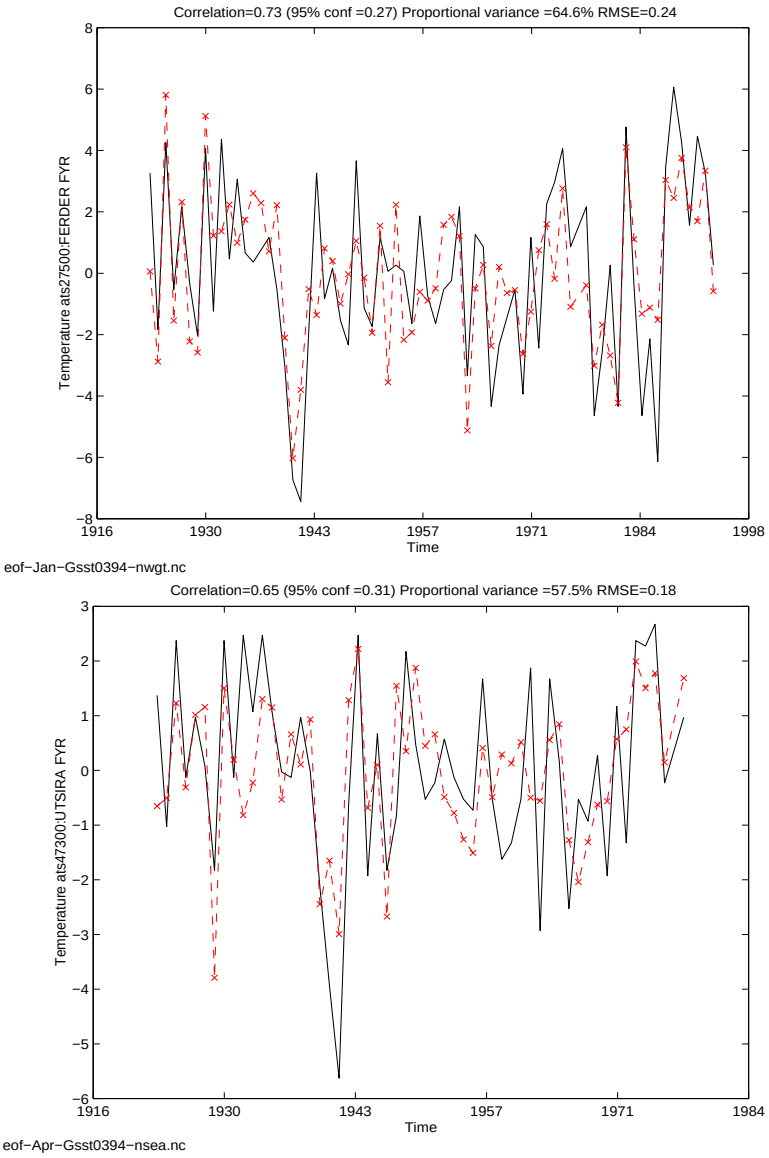

Figure 7: Same as figure 2, but for the North Atlantic January SST model (top) and The Nordic Seas April Model (bottom). The time series represent the temperatures from Ferder fyr in the upper panel and Utsira fyr in the lower panel.

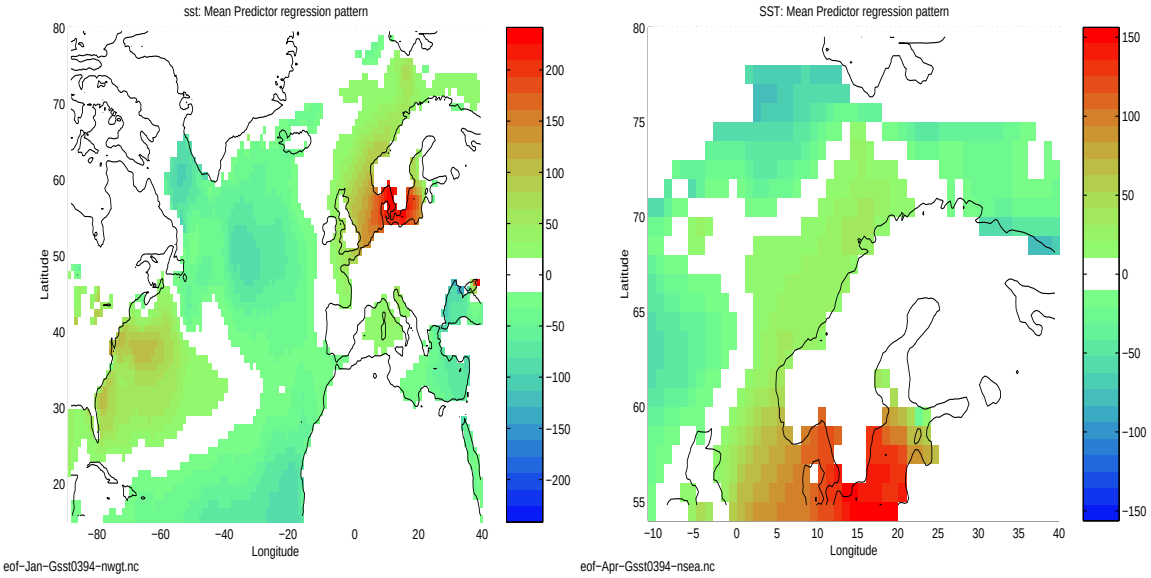

Figure 8: The mean January North Atlantic MVR GISST2.2 SST pattern associated with land surface temperatures at Ferder fyr (left). THe right panel shows the April SST pattern that is associated with Utsira fyr.

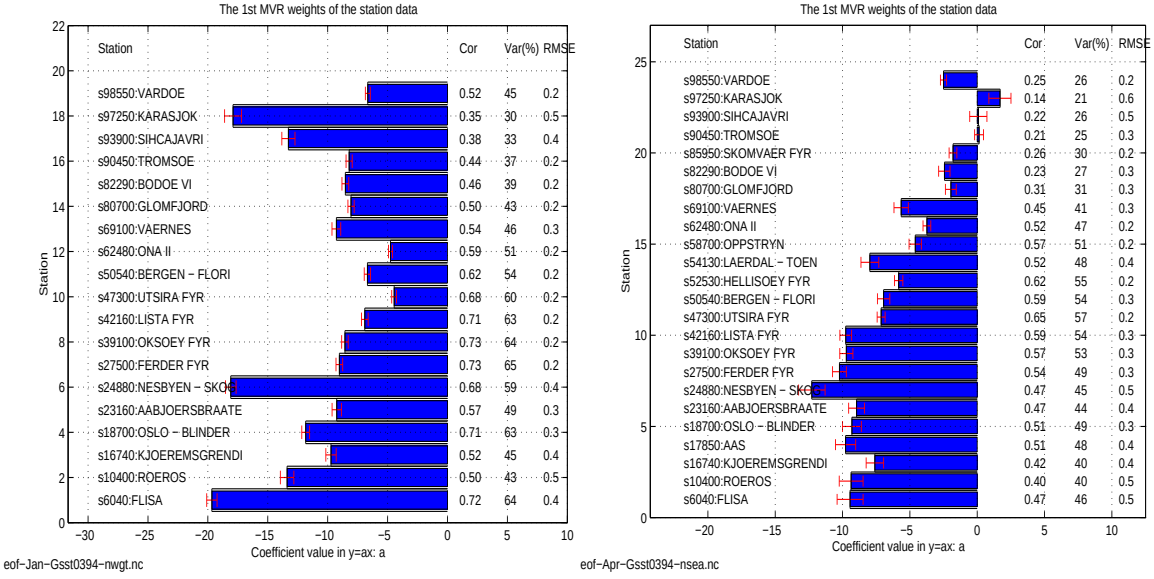

Figure 9: Same as figure 6, but showing the results for the January North Atlantic model (left) and the April Nordic Seas model (right).

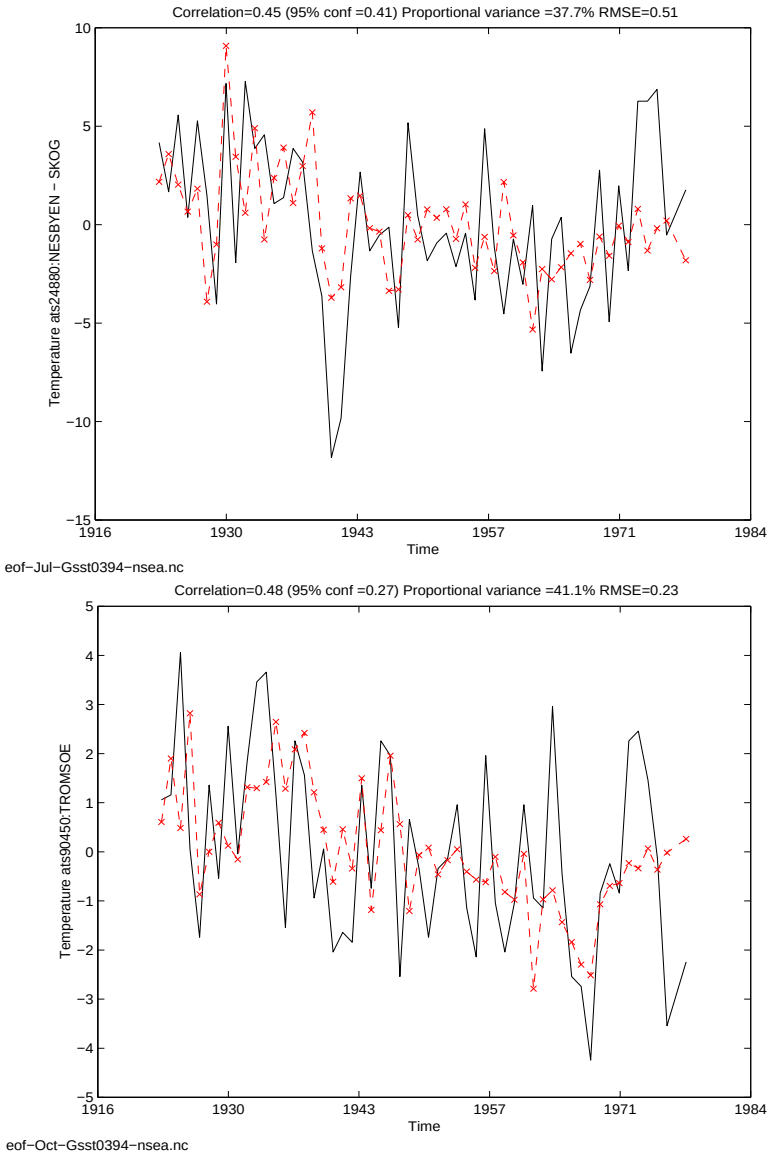

Figure 10: Same as figure 2, but for the North Atlantic January SST model (top) and The Nordic Seas April Model (bottom). The time series represent the temperatures from Ferder fyr in the upper panel and Utsira fyr in the lower panel.

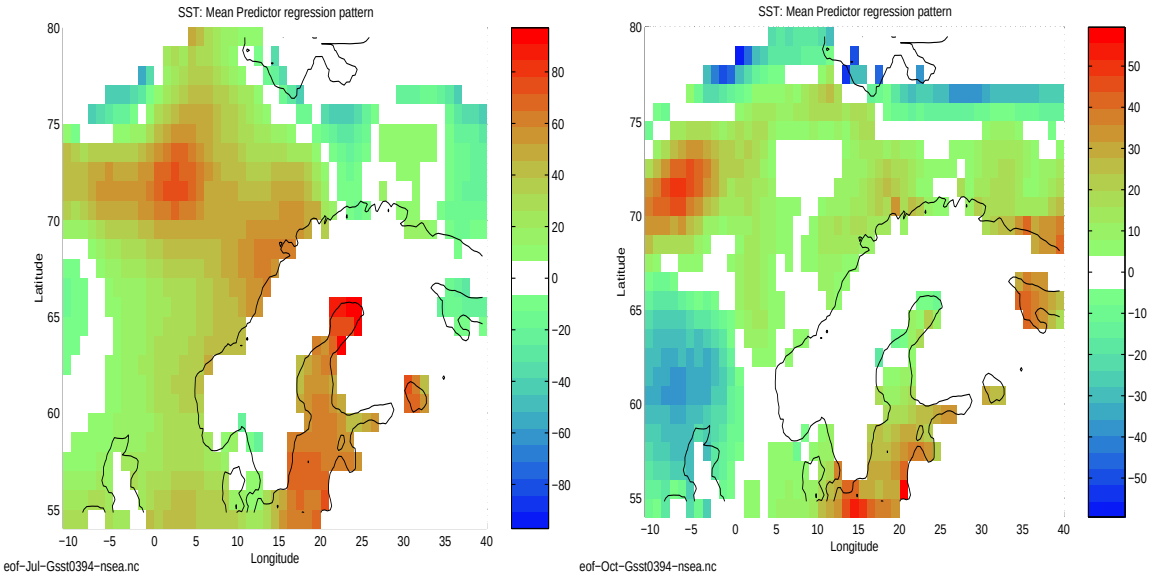

Figure 11: The mean January North Atlantic MVR GISST2.2 SST pattern associated with land surface temperatures at Ferder fyr (left). THe right panel shows the April SST pattern that is associated with Utsira fyr.

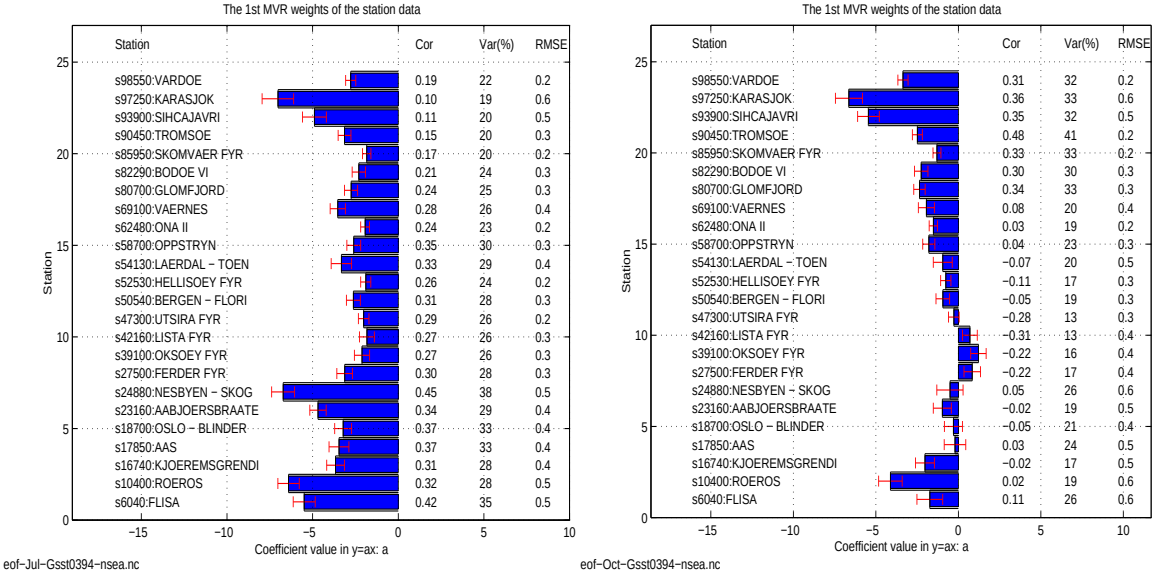

Figure 12: Same as figure 6, but showing the results for the January North Atlantic model (left) and the April Nordic Seas model (right).

- The predictor and predictand weights for the North Atlantic model are small compared to that of the CCA weights: reason is that the regression weights indicate the importance of the leading SST EOF pattern while they are associated with the leading CCA pattern for the CCA analysis. This also explains the random polarities of the predictand weights, as the leading North Atlantic SST EOF was unrelated to the Norwegian temperatures.

The cross-validation results for a number SST MVR models based on differnt seasons, areas, time prediods, and EOF combination in the predictor set are shown figures 2 to 12. The screening procedure that found the optimal EOF combination for the predictor set, gave the same combinations as for the CCA models (*Benestad*, 1998a) but different combinations to those of the SVD models (*Benestad*, 1998c). Figures 3 and 3 show the temperature predictions for Ferder fyr using the January MVR and SVDMVR models respectively. Figure 4 shows the January MVR predictions at Ferder fyr up to 1994, demonstrating that this model also captured the more recent temperature anomalies. Figures 6 show the cross-validation scores for the January MVR model, where the left panel represent the model calibrated and validated over the same time preiod as in *Benestad* (1998a) and *Benestad* (1998c) and the right panel shows the scores for the model calibrated on the 1923-1994 period[1]. The cross-validation prediction scores for the best predictions were higher (r=0.77 for 1923-1994, compared to r=0.72 for 1923 to 1978) for the longer time period.

The predictand weights shown in these figures are the strength of the regression coefficient (slope) that relates the leading SST EOF and the temperatures. The cross-validation skill-scores (table 1) were similar as for the CCA models and the location of the best prediction was in general similar as for the Nordic Sea CCA models.

The upper panel in figure 8 shows the predictions using the January North Atlantic model. As for the CCA models, the January North Atlantic model gave higher best correlation skill than the Nordic Sea model, but the best MVR prediction was found at Ferder fyr as opposed to Flisa for the CCA predictions. The reason why the two kinds of models produced best predictions for different regions may be due to the fact that the MVR model is sensitive to the temperature amplitudes while the CCA models are insensitive to signal amplitudes, and smaller amplitudes, for instance in coastal climates, are assocaited with smaller RMS errors. The left panel in figure 9 shows the regression coefficients between the temperatures and the leading North Sea

[1]Notice that some stations have been excluded in order to get more up to date observations.

SST EOF (*Benestad*, 1998b), suggesting that the leading SST EOF had strongest influence on the inland Norwegian January temperatures.

The lower panel in figure 8 shows the April results for the MVR model, indicating a reduction in prediction skill during the spreing season. The seasonal variation in MVR prediction skill was similar to the skill seasonality of the CCA models. The regression coefficients in the right panel of figure 9 indicates a clear link between the Norwegian January temperatures and the leading Nordic SST EOF with strongest loadings in Skagerrak and Kattegat.

4 SLP Models

4.1 The NMC ds195.5 models

Figures 13 to 15 show the cross-validation results for the NMC SLP MVR models. The left panel in figure 13 shows the best prediction for January, which was for the temperatures at Utsira fyr. The left panel in figure 14 shows the SLP predictor pattern assocaited with the predicted Utsire temperature anomalies. The right hand panels of figures 13 and 14 show the corresponding results for Flisa April temperatures. There was a significant drop in the prediction skill from January to April, similar to the reduction in skill seen for the CCA models. In both months, however, the SLP predictor pattern resembles the NAO structure with a north-south dipole whose northern maximum is centered over the vicinity of Iceland and Greenland. The southern maximum during January is centered over the Iberian peninsula, however, during April the southern branch is displaced westward by almost 40 degrees.

The regression coefficients, shown in figure 15 suggests that the leading SLP EOF in general had slightly stronger influence in southern Norway than in Northern Norway during January (left), but the largest coefficients were found in northern Norway during April (right). The cross-validation skill-scores in the north were significantly lower than in the south during January, while in April, the skill scores were similar in north and south[2].

The prediction scores increased from spring to Autumn 13, and the summertime SLP pattern was characterised with strong anomalies over Scandinavia and Greenland (figure 17, left panel). In October, 89% of the Tromsø temperature anomalies were accounted for by a west-east dipole pattern assocaited with significant variations in the meridional wind component (along shore winds). The leading SLP EOFs had weak influence on the northern

[2]The geographical distribution of prediction skill corresponded roughly with that of the CCA models (*Benestad*, 1998a)

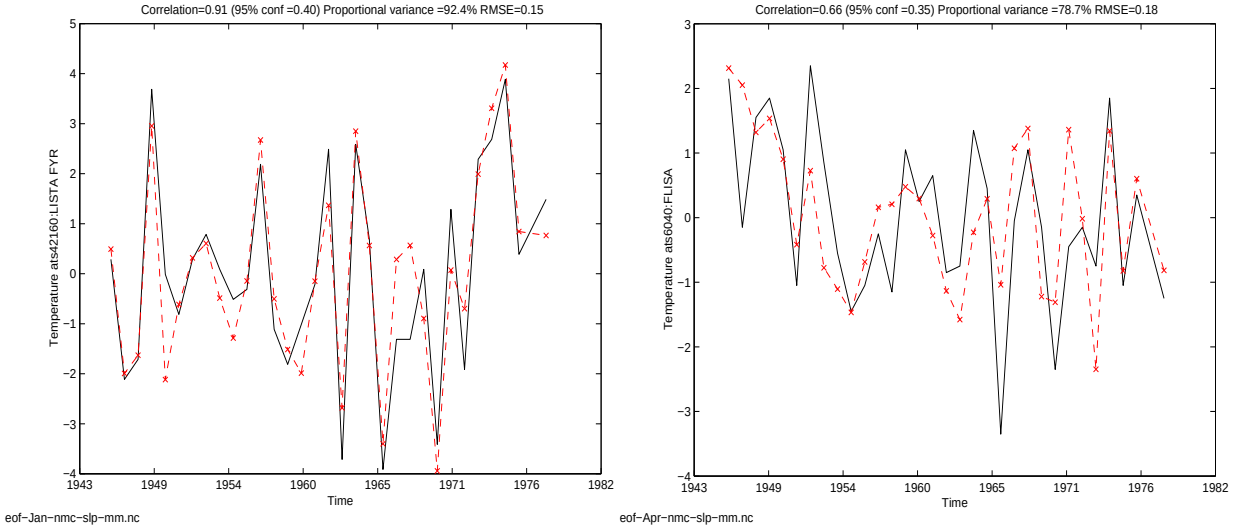

Figure 13: Time series of predicted January (left) and April (right) temperatures (dashed) at Lista fyr (left) and Flisa (right), employing the cross-validation method with NMC SLPs, shown with the observations (black solid line).

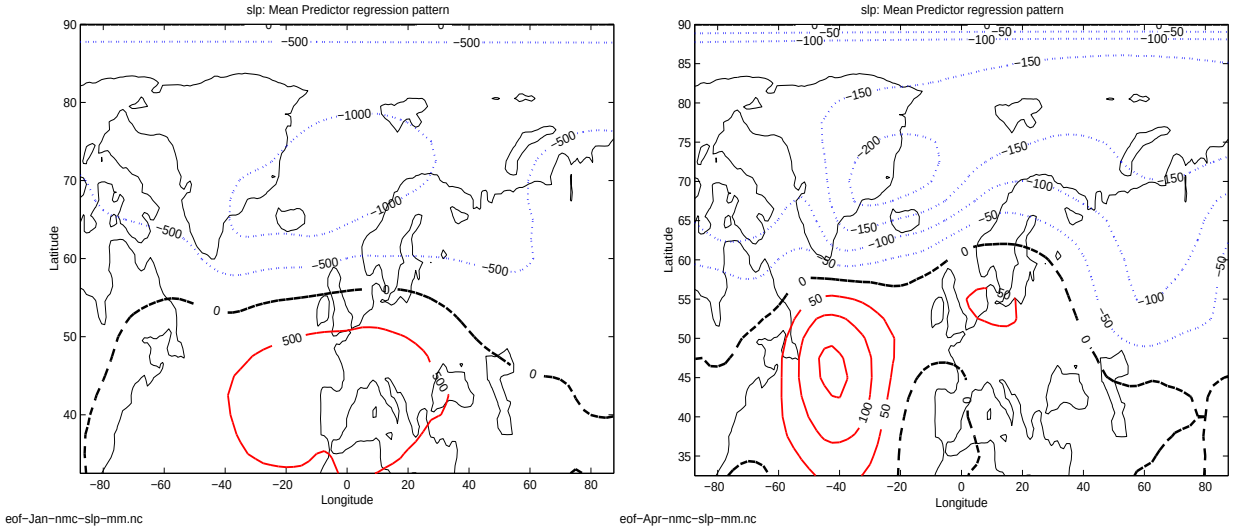

Figure 14: The mean January (left) and April (right) MVR NMC SLP patterns associated with the corrsponding temperatures in figure 13.

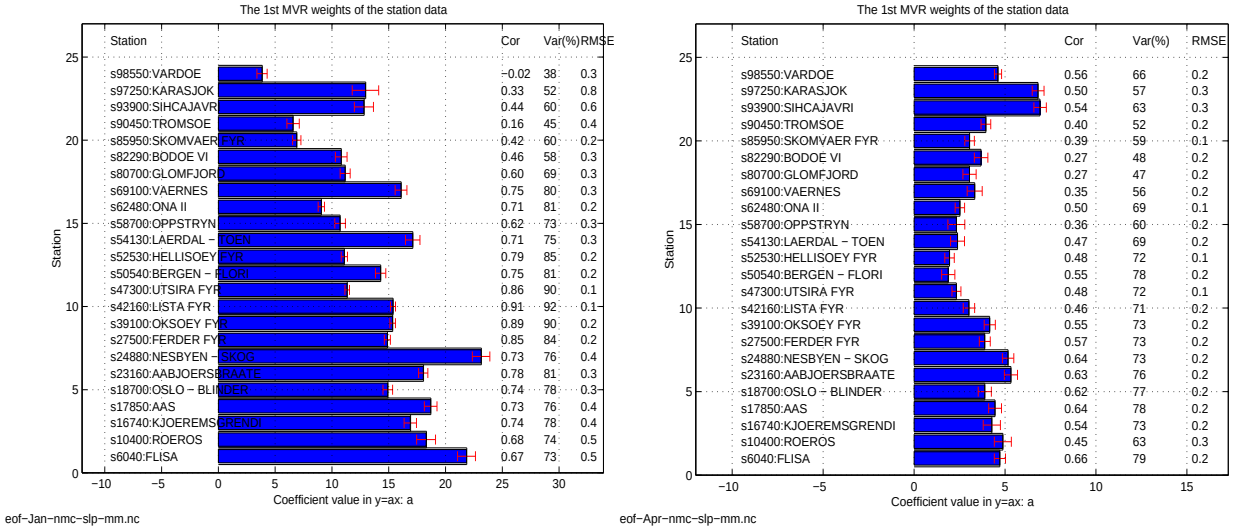

Figure 15: The mean January (left) and April (right) weights associated with EOF1 from the cross-validation analysis shown in filled bars indicate the mean MVR predictand weights for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation results from the cross-validation analysis are given on the right hand side.

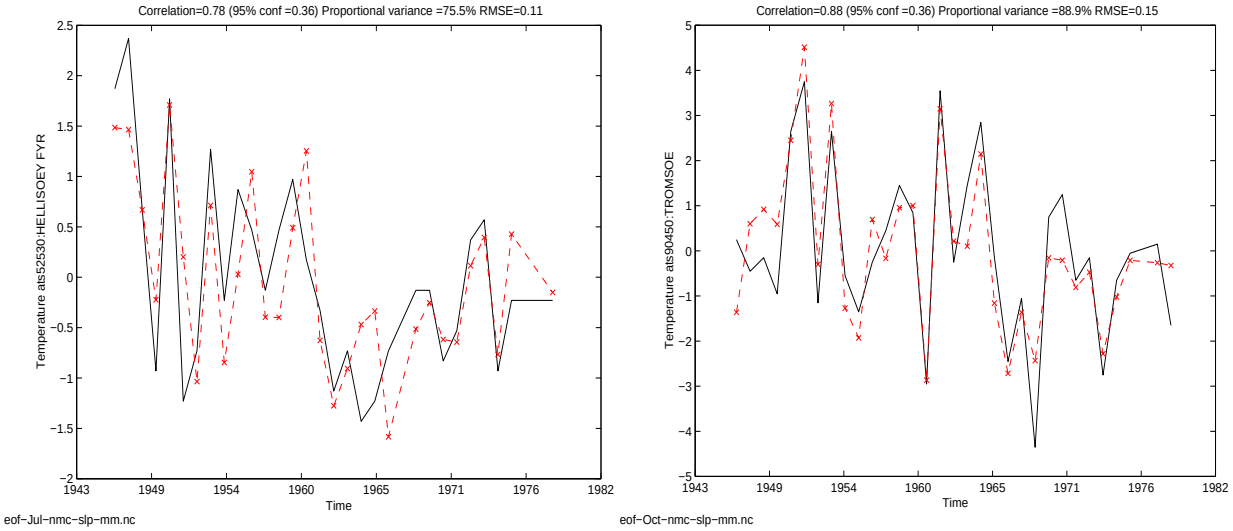

Figure 16: Time series of predicted July temperatures at Hellisøy (left) and October temperatures at Tromsø (right), employing the cross-validation method with NMC SLPs. Predictions are shown as dashed lines and observations as solid line.

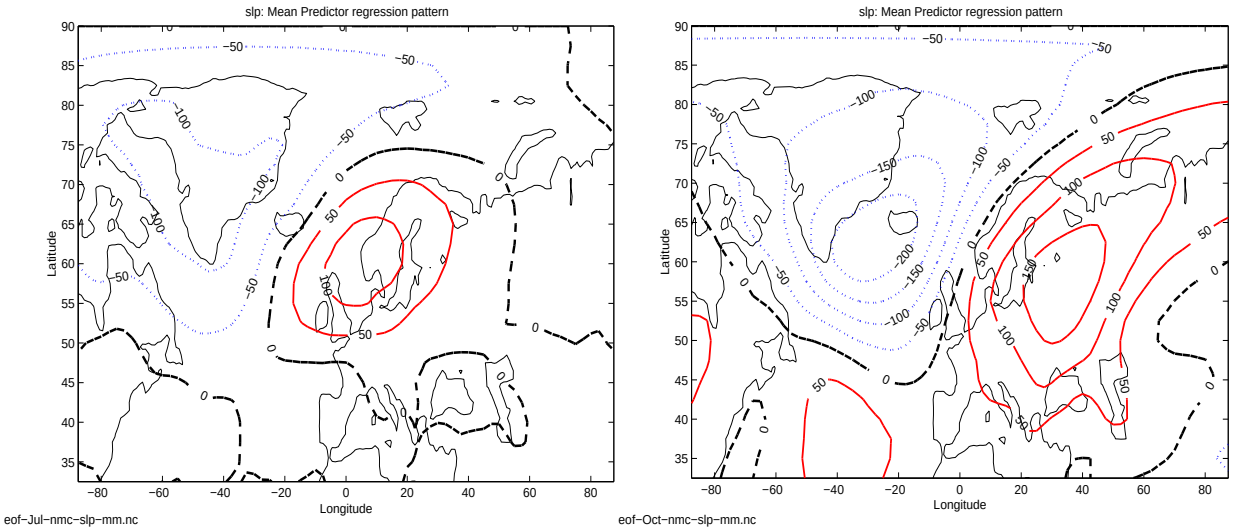

Figure 17: Same as figure 14, but for the July (left) and October (right) NMC SLP models: The pattern associated with the Hellingsøy temperatures is shown in the left panel and Tromsø in the right panel.

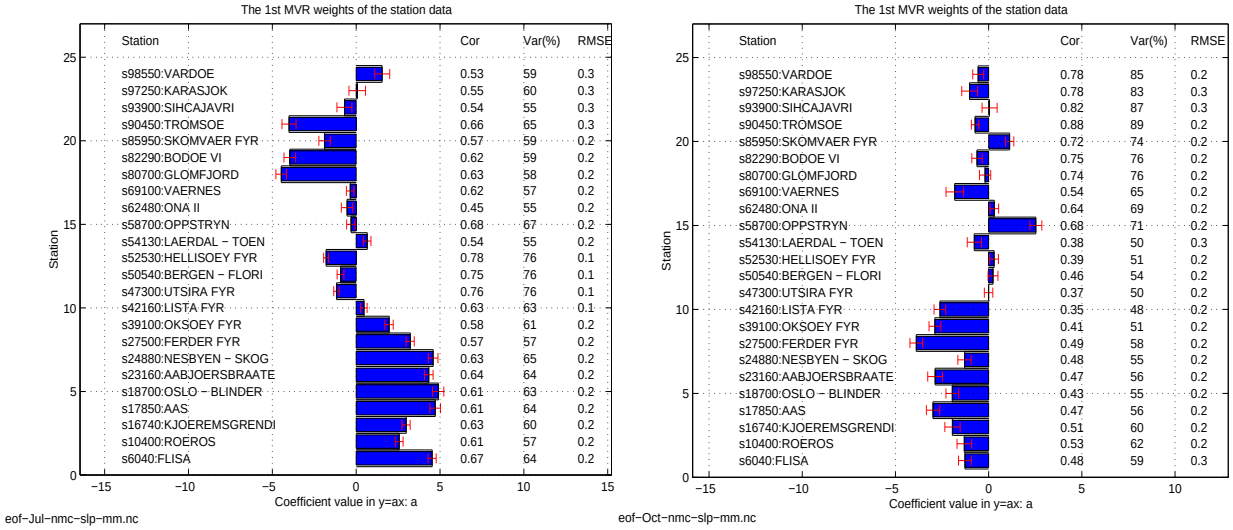

Figure 18: Same as figure 15, but for the July (left) and October (right) NMC SLP models.

Norwegian temperatures during summer and autumn, with the exception near Tromsø, Bodø, and Glomfjord in July.

4.2 The UEA models

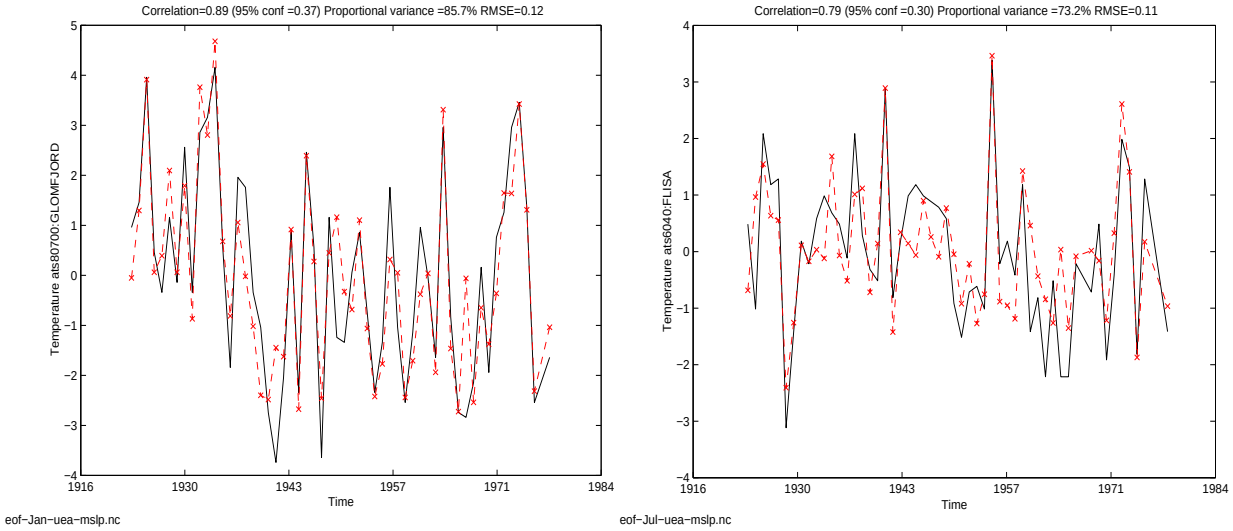

Figure 19: Time series of predicted January temperatures at Glomfjord (left) and July temperatures for Flisa (right), employing the cross-validation method with UEA SLPs. The predictions are shown as a dashed line and the observations as a solid line.

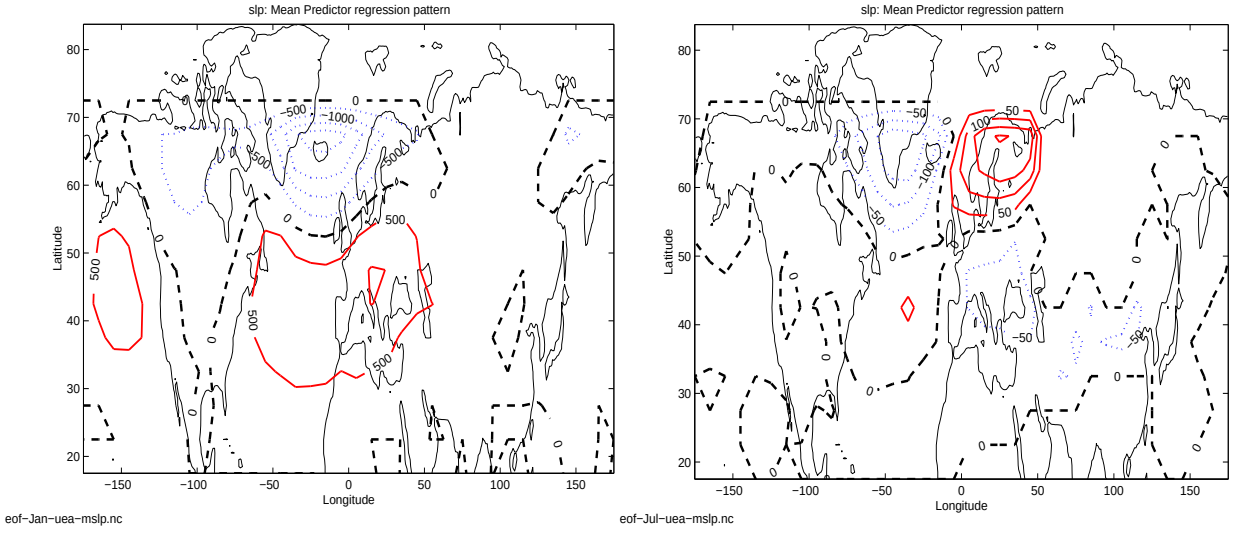

Figure 20: The mean January MVR UEA SLP pattern associated with the land surface temperatures in Glomfjord (left) and Flisa (right).

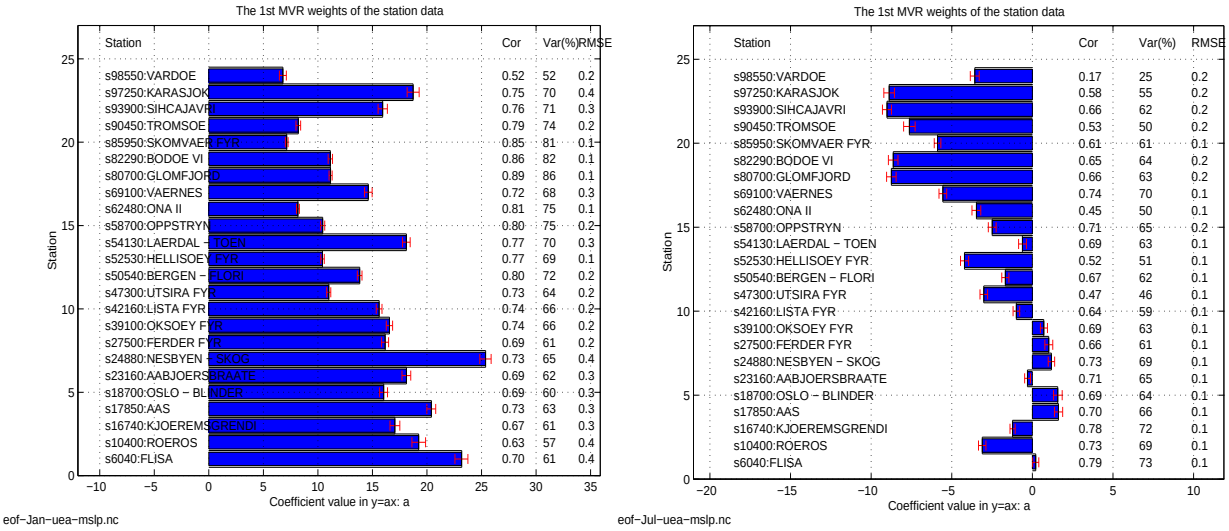

Figure 21: Same as figure 15, but for the January (left) and July (right) UEA SLP models.

The UEA SLP MVR models gave similar skill scores as the UEA model (figures 19 to 21), but the locations of best prediction skill were not the same as for thr MVR models. A summary of the best and wors skill scores are given in table 2. A comparison between the January UEA MVR model and January UEA CCA results shows good agreement between the predictor pattern, skill scores and location of best prediction (*Benestad*, 1998a, p.52, fig. 51). The July results were also similar, although not identical.

5 Geopotential height Models

The NMC 500hPa geoptential MVR model results are show in figures 22 to 27. The January prediction for Hellingsøy (figure 22, left panel) were associated with slightly lower cross-validation correlation skills than the corresponding CCA prediction. However, on the whole, there was good agreement between the CCA and MVR models. The April MVR cross-validation results corresponded even better with the CCA model, giving alsmost identical skill scores for the best prediction (Åbjørsbråten). The best MVR July prediction were associated with slighlt lower RMS errors than the corresponding CCA predictions, but the MVR predictor pattern for Karasjok was characterised by strongest anomalies over southern Sweden (figure 26, left panel) while the leading CCA pattern indicated strongest variability over the Barents Sea and North-East Atlantic (*Benestad*, 1998a, p.66, fig. 69). It is important to note, however, that the two predictor patterns are not directly comparable,

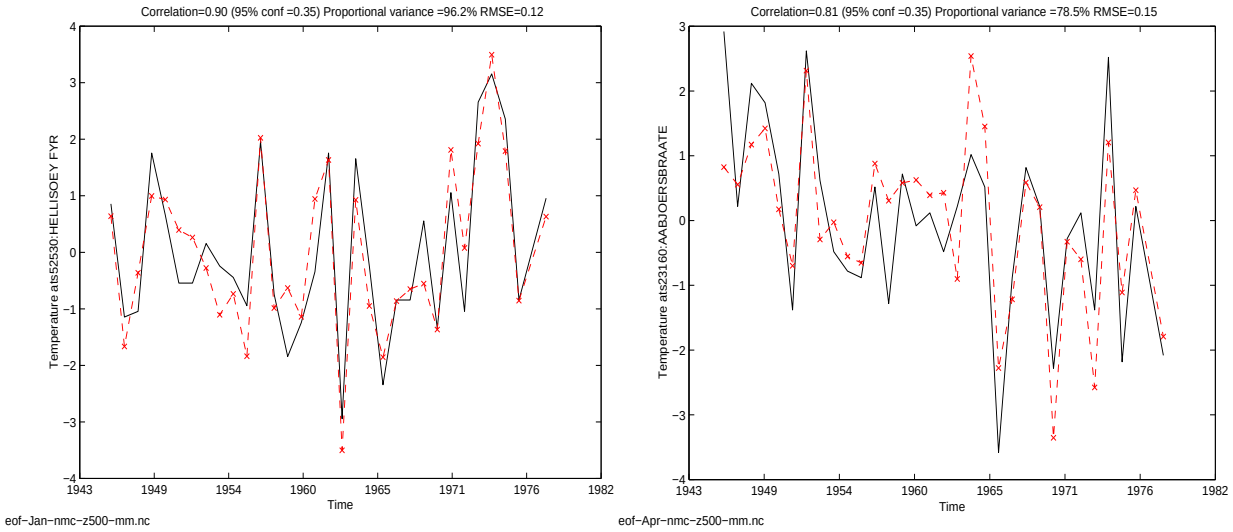

Figure 22: Time series of predicted January temperatures at Hellisøy fyr (left) and April temperatures for Åbjørsbråten (right), employing the cross-validation method with NMC z(500hPa). The predictions are shown as a dashed line and the observations as a solid line.

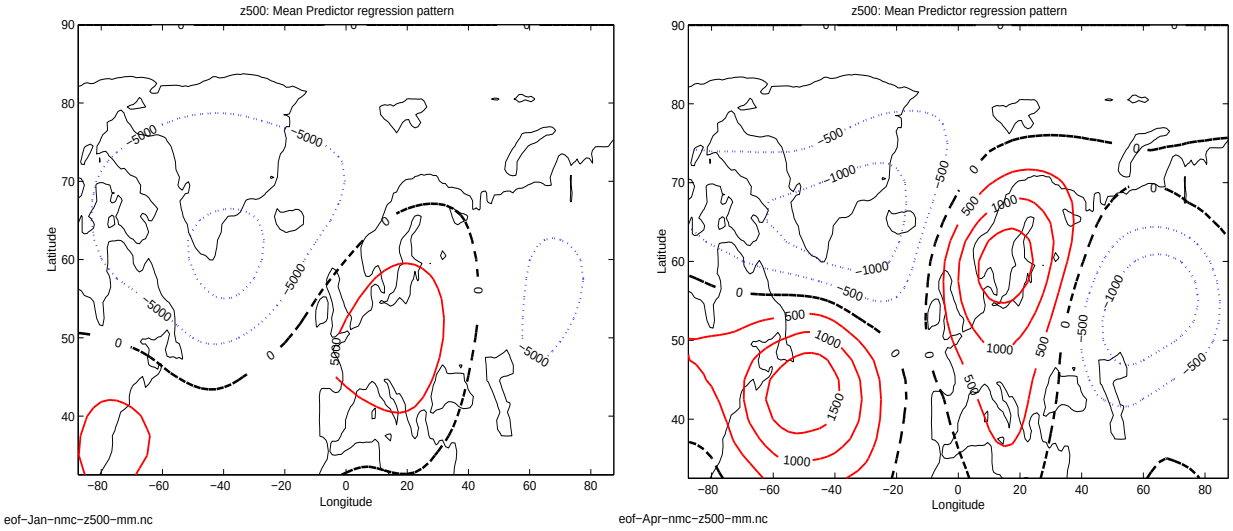

Figure 23: The mean January (left) and April (right) MVR NMC z(500hPa) patterns associated with the land surface temperatures at Hellingsøy (left) and Åbjørsbråten (right).

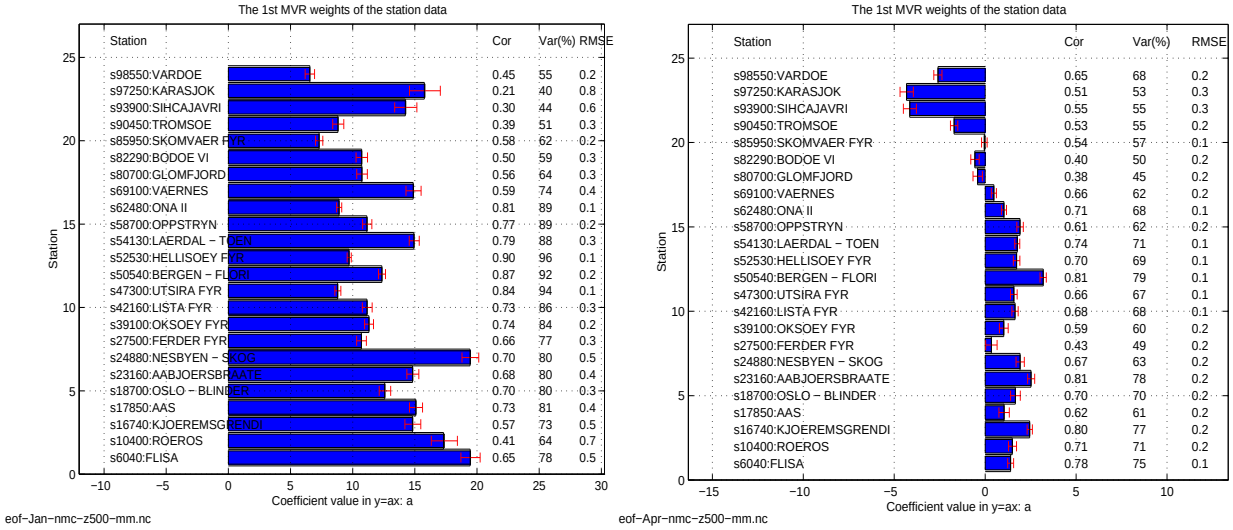

Figure 24: The mean January (left) and April (right) predictand weights for the land surface temperatures at Hellingsøy and Åbjørsbråten. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation results from the cross-validation analysis are given on the right hand side.

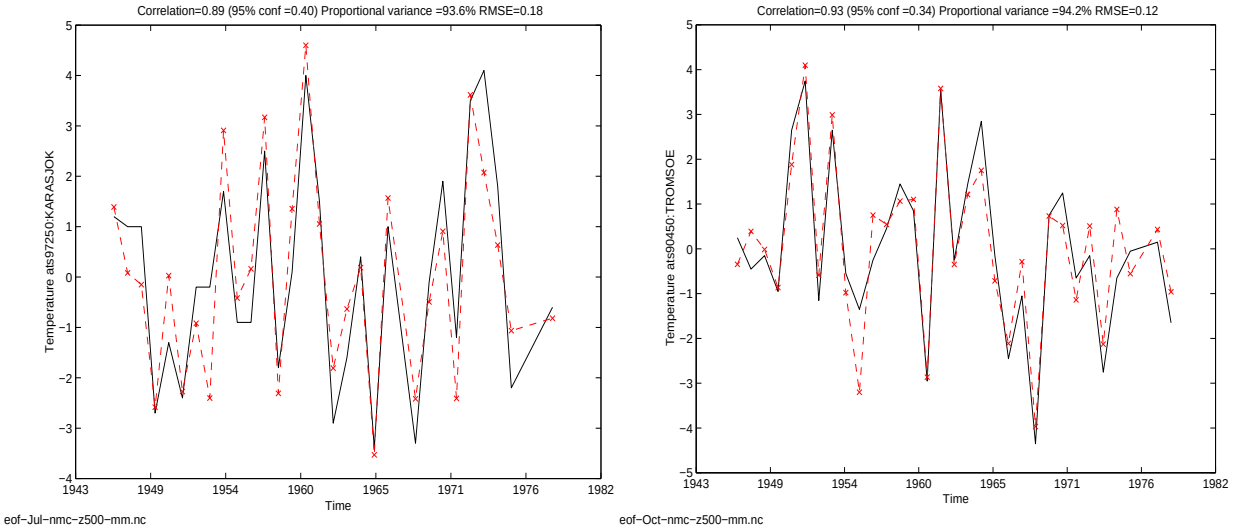

Figure 25: Time series of predicted July temperatures at Karasjok (left) and October temperatures for Tromsø (right), employing the cross-validation method with NMC z(500hPa). The predictions are shown as a dashed line and the observations as a solid line.

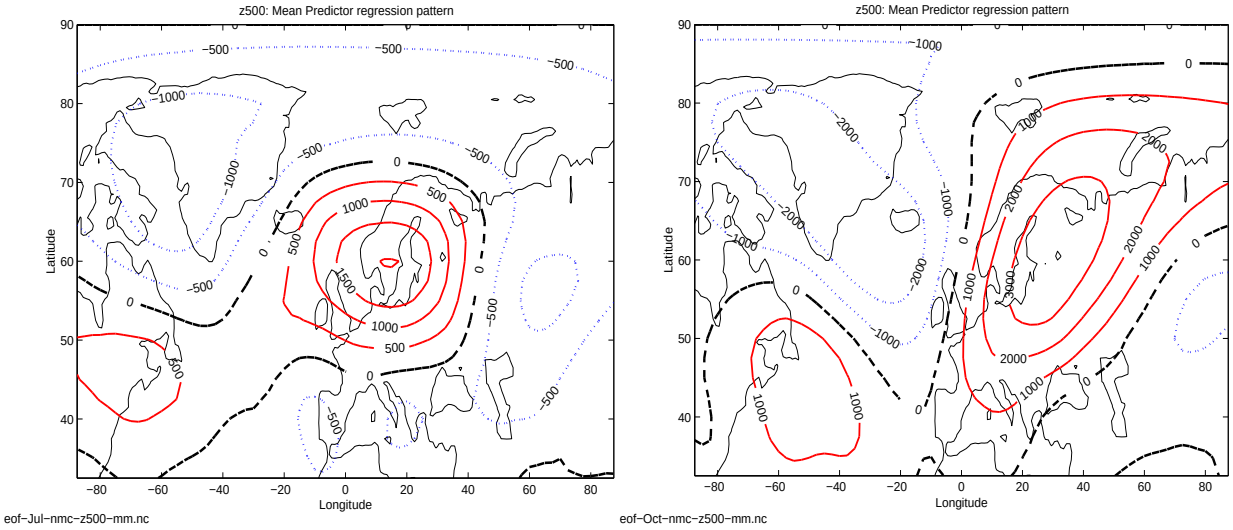

Figure 26: The mean July (left) and October (right) MVR NMC z(500hPa) pattern associated with the land surface temperatures at Karasjok and Tromsø respectively.

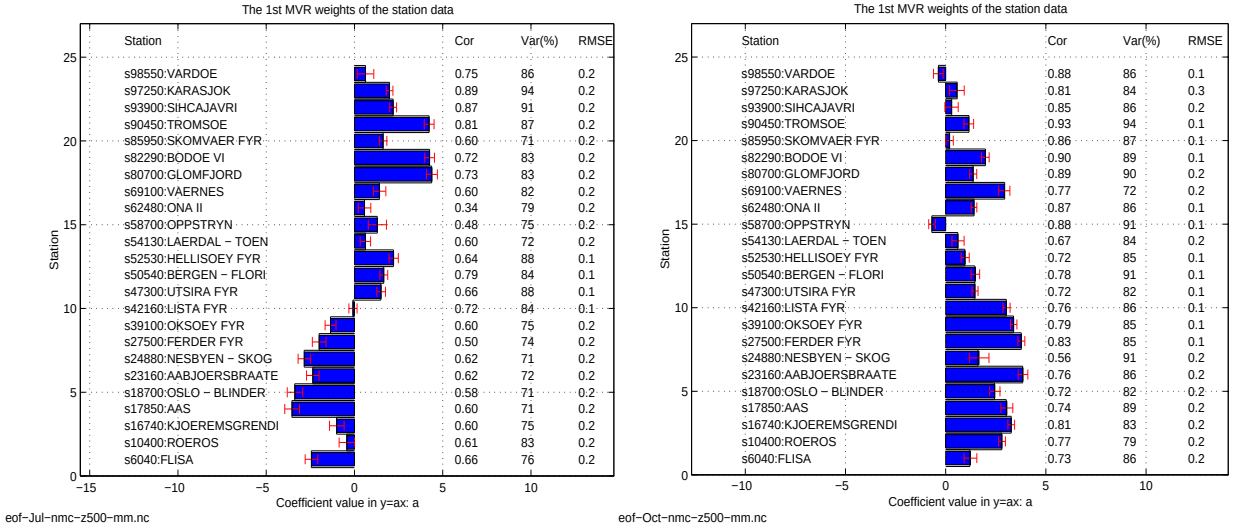

Figure 27: Same as for figure 30, but for the July (left) and October (right) NMC 500 hPa geopotential height models.

as a superposition of several CCA patterns may affect the predictand at a location while the regression analysis only gives one pattern for each station. Both the MVR and the CCA models gave best predictions for Tromsø, and both models gave cross-validation correlation scores or 0.93. The geographical distribution of skill was similar for both kinds of models.

6 The 500hPa Temperature Models

Benestad (1998c) suggested that the 500hPa temperatures over the North Atlantic region may be a good predictor for the Norwegian temperatures. Here, we discuss the results from a similar analysis as the cross-validation in *Benestad* (1998c), but here we extended the time period (1960-1994) of analysis by excluding the stations that did not have up to date observations. We have also included the best and worst prediction skills for analysis done for the same time period (1960-1978) as for the SVD analysis in *Benestad* (1998c) in table 4, just for comaprison with the SVD model results. Because the 1960-1978 is short, the regression results may be subject to large numerical errors due to badly scaled matrices of near-singluarity problems. The location of the best predicitons were not the same as for the SVD results. Nervertheless, the cross-validation results gave extremely high skill scores, which were systematically higher than the corresponding best SVD scores (*Benestad*, 1998c, p. 35, table. 4). The model results for the 1960-1994 were lower than those of the 1960-1978 period, however, the former results were more robust.

The cross-validation analysis of the NMC 500hPa temperature MVR models demonstrated that the 500hPa temperatures were suitable predictors for the Norwegian temperatures (figures 28 to 33). The surface temperatures were in general associated with strongest 500hPa temperature anomalies over scandinavia. In January, high Sihcajarvi temperatures (figure 28, left) coincided with warm anomalies aloft (figure 29, left). High April temperatures at Flisa were also associated with a warming near the 500hPa isobar. In January, the best skills were seen in the north while in April there were good predictions both in the south and the north (figure 30). The best July prediction was for Røros (figure 31, left) but in October the highest prediction score was found at Bodø, and the corresponding predictor patterns in figure 32 both suggested a relatively deep structure of the temperature anomaly. Figure 33 suggests that the leading July T(500hPa) EOF had little effect on the Norwegian July temperatures, and that the October EOF had a stronger influence.

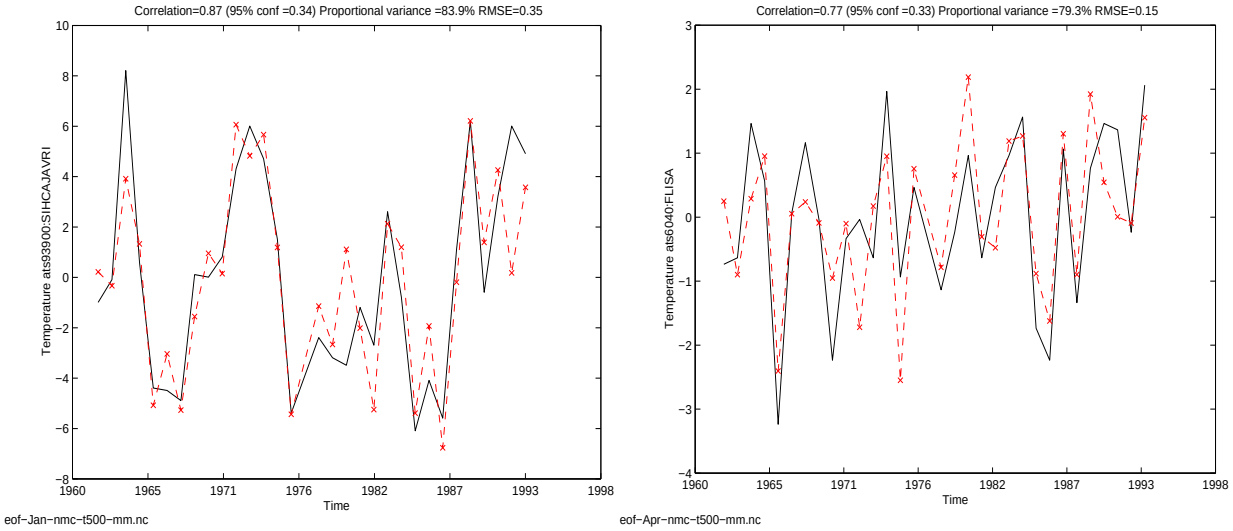

Figure 28: Time series of predicted January temperatures at Sihcajarvi (left) and April temperatures for Flisa (right), employing the cross-validation method with NMC T(500hPa). The predictions are shown as a dashed line and the observations as a solid line.

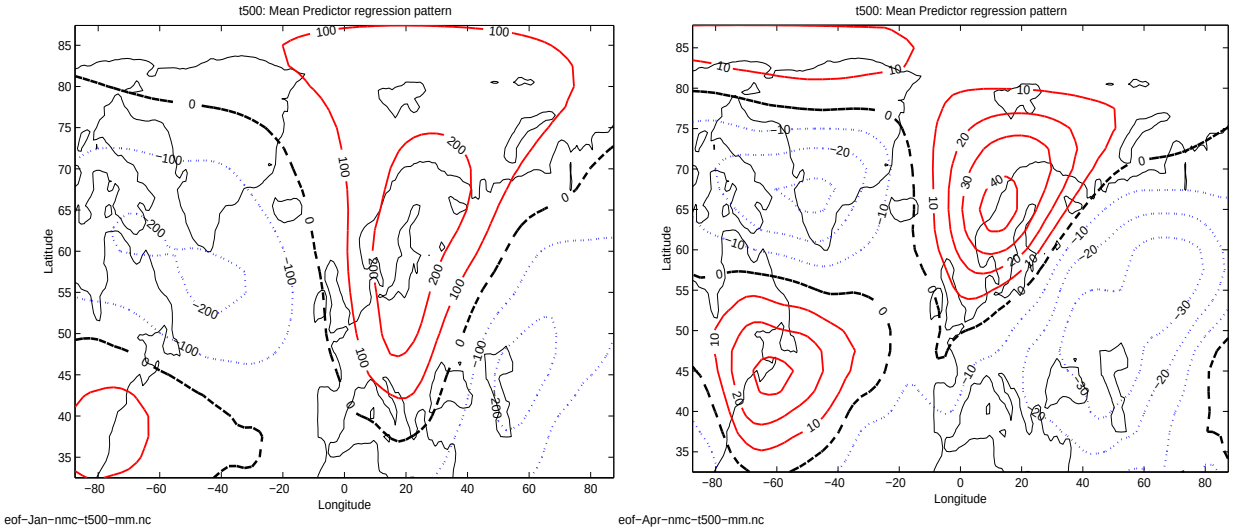

Figure 29: The mean January (left) and April (right) MVR NMC T(500hPa) patterns associated with the land surface temperatures at Sihcajarvi and Flisa respecitvely.

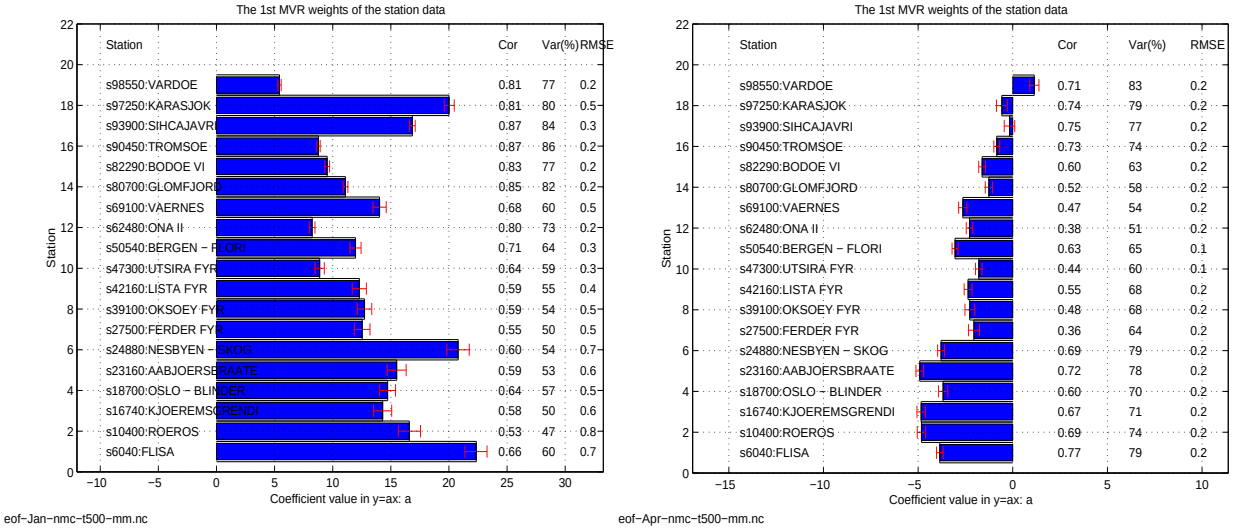

Figure 30: The mean January (left) and April (right) predictand weights associated with EOF1 for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation results from the cross-validation analysis are given on the right hand side.

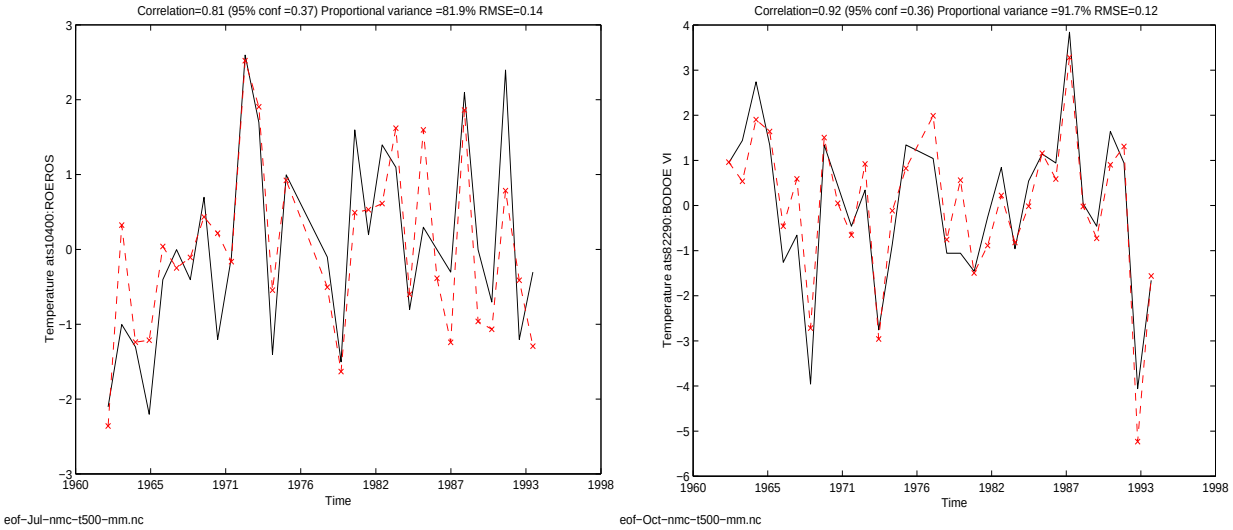

Figure 31: Time series of predicted July temperatures at Røros (left) and October temperatures for Bodø (right), employing the cross-validation method with NMC T(500hPa). The predictions are shown as a dashed line and the observations as a solid line.

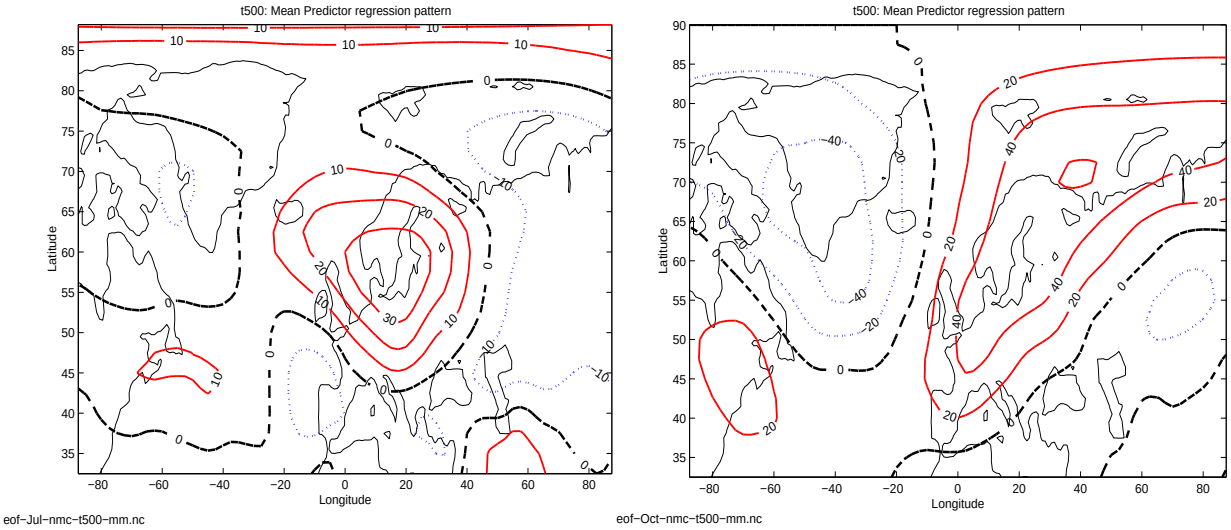

Figure 32: The mean July MVR NMC T(500hPa) pattern associated with the land surface temperatures at Røros and Bodø respectively.

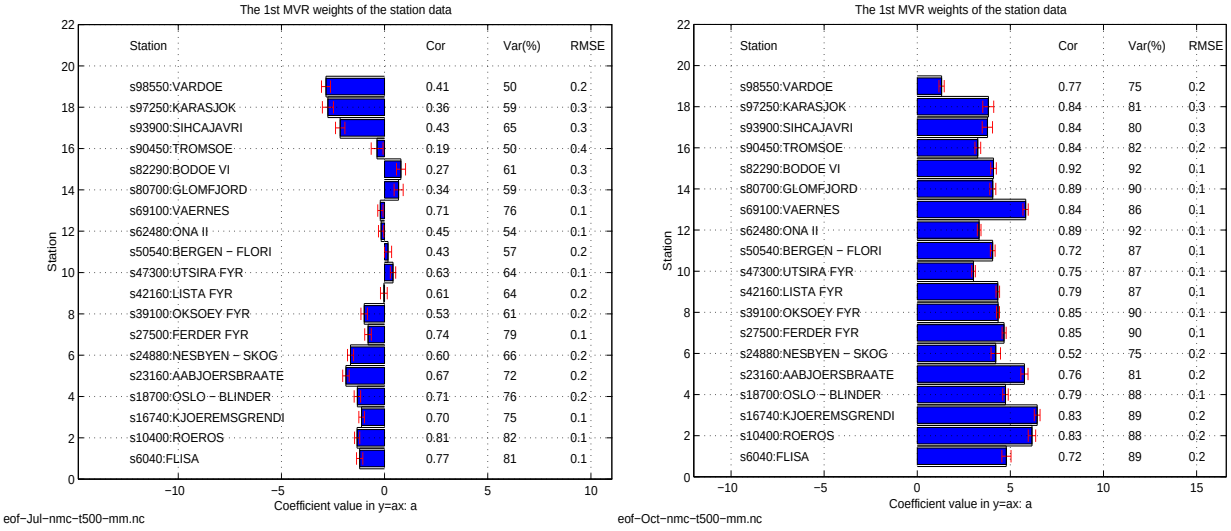

Figure 33: Same as for figure 30, but for the July (left) and October (right) NMC 500 hPa temperature models.

Table 1: Model scores

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
Regional SST 1 2 3 6 7 8	Optimal CCA/MVR predictors FERDER FYR r= 0.72 rmse= 0.24	Jan SKOMVÆR FYR r= 0.49 rmse= 0.16	 KARASJOK r= 0.40 rmse= 0.57
Regional SST 1 2 3 6 8 10 17 18	Optimal CCA/MVR predictors UTSIRA FYR r= 0.65 rmse= 0.18	Apr ONA II r= 0.52 rmse= 0.17	 KARASJOK r= 0.14 rmse= 0.63
Regional SST 1 4 10 11 18 20	Optimal CCA/MVR predictors NESBYEN - SKOG r= 0.45 rmse= 0.51	Jul SKOMVÆR FYR r= 0.17 rmse= 0.18	 KARASJOK r= 0.10 rmse= 0.63
Regional SST 1 2 3 5 7 9 10 15 20	Optimal CCA/MVR predictors TROMSØ r= 0.48 rmse= 0.23	Oct SKOMVÆR FYR r= 0.33 rmse= 0.17	 LISTA FYR r= -0.31 rmse= 0.37
Norh Atlantic SST 1 2 3 8 9 10 11 12 18	Optimal MVR predictors FERDER FYR r= 0.73 rmse= 0.24	Jan ONA II r= 0.59 rmse= 0.15	1922-1994 KARASJOK r= 0.35 rmse= 0.52
Regional SST 1 2 5 10 11 18 19	Optimal SVD predictors FERDER FYR r= 0.60 rmse= 0.28	Jan Ona II r= 0.47 rmse= 0.17	 TROMSØ r= -0.12 rmse= 0.28
Regional SST 1 2 3 6	Optimal MVR predictors FERDER FYR r= 0.77 rmse= 0.22	Jan ONA II r= 0.64 rmse= 0.14	1922-1994 VARDØ r= 0.38 rmse= 0.17
Regional SST 1 2 3 10 11 13	Optimal MVR predictors NESBYEN - SKOG r= 0.42 rmse= 0.49	Jul VARDØ r= 0.33 rmse= 0.18	1922-1994 KARASJOK r= 0.21 rmse= 0.54

Table 2: Model scores

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
NCAR SLP 1 6 7 10 11 17 20	Optimal MVR/CCA predictors LISTA FYR r= 0.85 rmse= 0.16	Jan UTSIRA FYR r= 0.81 rmse= 0.13	TROMSØ r= 0.34 rmse= 0.22
NCAR SLP 1 2 4 6 9 14 17 20	Optimal MVR/CCA predictors ONA II r= 0.67 rmse= 0.08	Apr ONA II r= 0.67 rmse= 0.08	OSLO - BLINDER r= 0.27 rmse= 0.14
NCAR SLP 1 2 4 5 6 7 8 10 11 12 15 19	Optimal MVR/CCA predictors ÅBJØRSBRÅTE r= 0.74 rmse= 0.11	Jul FLISA r= 0.74 rmse= 0.11	ONA II r= 0.34 rmse= 0.12
NCAR SLP 1 2 3 4 5 7 8 9 11 12 17 20	Optimal MVR/CCA predictors TROMSOE r= 0.86 rmse= 0.11	Oct UTSIRA FYR r= 0.65 rmse= 0.09	NESBYEN - SKOG r= 0.59 rmse= 0.15
NMC SLP 1 5 6 11 13 15 16 20	Optimal MVR/CCA predictors LISTA FYR r= 0.91 rmse= 0.15	Jan UTSIRA FYR r= 0.86 rmse= 0.13	VARDØ r= -0.02 rmse= 0.30
NMC SLP 1 2 3 5 6 7 15 17	Optimal MVR/CCA predictors FLISA r= 0.66 rmse= 0.18	Apr ONA II r= 0.50 rmse= 0.14	BODØ VI r= 0.27 rmse= 0.22
NMC SLP 1 2 3 6 8 15	Optimal MVR/CCA predictors HELLISØY FYR r= 0.78 rmse= 0.11	Jul HELLISØY FYR r= 0.78 rmse= 0.11	ONA II r= 0.45 rmse= 0.18
NMC SLP 1 2 4 8 15 19 20	Optimal MVR/CCA predictors TROMSØ r= 0.88 rmse= 0.15	Oct TROMSØ r= 0.88 rmse= 0.15	LISTA FYR r= 0.35 rmse= 0.20

Table 3: Model scores

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
UEA SLP	Optimal MVR/CCA predictors	Jan	
1 2 3 4	GLOMFJORD	SKOMVÆR FYR	VARDØ
7 8 10	r= 0.89	r= 0.85	r= 0.52
11 12 18	rmse= 0.12	rmse= 0.10	rmse= 0.18
UEA SLP	Optimal MVR/CCA predictors	Apr	
1 2 3	BERGEN - FLORI	ONA II	FERDER FYR
4 8 9	r= 0.60	r= 0.55	r= 0.26
10	rmse= 0.13	rmse= 0.10	rmse= 0.17
UEA SLP	Optimal MVR/CCA predictors	Jul	
1 3 6	FLISA	FLISA	VARDØ
7 9 10	r= 0.79	r= 0.79	r= 0.17
12 14 20	rmse= 0.11	rmse= 0.11	rmse= 0.21
UEA SLP	Optimal MVR/CCA predictors	Oct	
1 2 3 4	TROMSØ	UTSIRA FYR	NESBYEN - SKOG
5 6 8 10	r= 0.82	r= 0.62	r= 0.47
12 15 17	rmse= 0.14	rmse= 0.11	rmse= 0.20
NMC z500	Optimal MVR predictors	Jan	
1 4 9	HELLISØY FYR	HELLISØY FYR	KARASJOK
11 13 14	r= 0.90	r= 0.90	r= 0.21
15 19	rmse= 0.12	rmse= 0.12	rmse= 0.82
NMC z500	Optimal MVR/CCA predictors	Apr	
1 2 4	ÅBJØRSBRÅTE	ONA II	GLOMFJORD
10 12 18	r= 0.81	r= 0.71	r= 0.38
20	rmse= 0.15	rmse= 0.11	rmse= 0.21
NMC z500	Optimal MVR/CCA predictors	Jul	
1 2 3	KARASJOK	BERGEN - FLORI	ONA II
4 6 8	r= 0.89	r= 0.79	r= 0.34
9 11 13	rmse= 0.18	rmse= 0.12	rmse= 0.20
NMC z500	Optimal MVR/CCA predictors	Oct	
1 2 3	TROMSØ	ONA II	NESBYEN - SKOG
4 6 7	r= 0.93	r= 0.87	r= 0.56
10 15 18	rmse= 0.12	rmse= 0.10	rmse= 0.24

Table 4: Model scores

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
hNMC t500	Optimal MVR predictors	Jan	Matrix close to singular
1 2 5 6	TROMSØ	SKOMVÆR FYR	RØROS
7 9 11	r= 0.95	r= 0.91	r= -0.04
12 14 15	rmse= 0.16	rmse= 0.15	rmse= 2.08
NMC t500	Optimal MVR predictors	Jan	1960-1993
1 6 8 14 20	SIHCAJAVRI	VARDØ	RØROS
	r= 0.87	r= 0.81	r= 0.53
	rmse= 0.35	rmse= 0.16	rmse= 0.76
NMC t500	Optimal MVR predictors	Apr	Matrix close to singular
1 5 9	BERGEN - FLORI	SKOMVÆR FYR	VARDØ
12 18 19	r= 0.92	r= 0.86	r= -0.33
	rmse= 0.11	rmse= 0.10	rmse= 0.43
NMC t500	Optimal MVR predictors	Apr	1960-1993
1 4 6	FLISA	UTSIRA FYR	FERDER FYR
7 8 9	r= 0.77	r= 0.44	r= 0.36
18 20	rmse= 0.15	rmse= 0.14	rmse= 0.22
NMC t500	Optimal MVR	Jul	Matrix close to singular
1 2 3	RØROS	RØROS	VARDØ
4 7 8	r= 0.98	r= 0.98	r= 0.27
11 14 20	rmse= 0.07	rmse= 0.07	rmse= 0.82
NMC t500	Optimal MVR	Jul	1960-1994
1 2 3	RØROS	UTSIRA FYR	TROMSØ
5 9 18	r= 0.81	r= 0.63	r= 0.19
20	rmse= 0.14	rmse= 0.13	rmse= 0.37
NMC t500	Optimal MVR	Oct	Matrix close to singular
1 2 5	GLOMFJORD	ONA II	NESBYEN - SKOG
6 8 19	r= 0.89	r= 0.87	r= 0.00
	rmse= 0.19	rmse= 0.11	rmse= 0.48
NMC t500	Optimal MVR	Oct	1960-1994
1 2 3	BODØ VI	ONA II	NESBYEN - SKOG
4 6 8	r= 0.92	r= 0.89	r= 0.52
11 12	rmse= 0.12	rmse= 0.09	rmse= 0.22

Table 5: Summary of best MVR model skills for the different models and different seasons. The score should not be directly compared as the location with highest scores varied with model and seasons. The cross-validation period for the NMC models was shorter than the for the other models, which may also give misleading results in a model score comparison.

Month	Local SST	NCAR SLP	NMC SLP	UEA SLP	NMC z	NMC T_{500hPa}	Max/Min
Jan	0.72	0.85	0.91	0.89	0.90	0.95	0.95/ 0.72
Apr	0.65	0.67	0.66	0.60	0.81	0.92	0.92/ 0.65
Jul	0.45	0.74	0.78	0.79	0.89	0.98	0.98 / 0.45
Oct	0.48	0.86	0.89	0.82	0.93	0.89	0.93 / 0.48
avg.	0.58	0.78	0.81	0.78	0.88	0.93	

Table 6: Summary of best SVD model skills for the different models and different seasons. The score should not be directly compared as the location with highest scores varied with model and seasons. The cross-validation period for the NMC models was shorter than the for the other models, which may also give misleading results in a model score comparison.

Month	Local SST	NCAR SLP	NMC SLP	UEA SLP	NMC z	NMC T_{500hPa}	Max/Min
Jan	0.68	0.72	0.75	0.61	0.72	0.87	0.87 / 0.61
Apr	0.47	0.48	0.66	0.58	0.68	0.86	0.86 / 0.47
Jul	0.43	0.61	0.68	0.58	0.57	0.85	0.85 / 0.43
Oct	0.44	0.77	0.78	0.82	0.54	0.81	0.82 / 0.44
avg.	0.51	0.65	0.71	0.65	0.63	0.85	

Table 7: Summary of best CCA model skills for the different models and different seasons. The score should not be directly compared as the location with highest scores varied with model and seasons. The cross-validation period for the sea ice models and the NMC models was shorter than the for the other models, which may also give misleading results in a model score comparison.

Month	Local SST	N.Atl. SST	NCAR SLP	NMC SLP	UEA SLP	NMC Z	Max/Min
Jan	0.72	0.75	0.84	0.91	0.89	0.91	0.89 / 0.64
Apr	0.65	0.66	0.66	0.67	0.60	0.81	0.82 / 0.60
Jul	0.61	0.48	0.72	0.79	0.79	0.89	0.90 / 0.48
Oct	0.48	0.37	0.86	0.88	0.83	0.93	0.94 / 0.37
avg.	0.62	0.57	0.77	0.81	0.78	0.89	

References

Benestad, R.E. 1998a. *CCA applied to Statistical Downscaling for Prediction of Monthly Mean Land Surface Temperatures: Model Documentation.* Klima 28/98. DNMI.

Benestad, R.E. 1998b. *Description and Evaluation of the Predictor Data sets used for Statistical Downscaling in the RegClim.* Klima 24/98. DNMI.

Benestad, R.E. 1998c. *SVD applied to Statistical Downscaling for Prediction of Monthly Mean Land Surface Temperatures: Model Documentation.* Klima 30/98. DNMI.

Strang, G. 1995. *Linear Algebra and its Application.* San Diego, California, USA: Harcourt Brace & Company.