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

R.E. Benestad

DNMI, November 5, 1998

Contents

1 Introduction **1**

2 CCA models **2**
2.1 Predictors and Predictands 2
2.2 Classical CCA 4
2.3 Linear relationships 6

3 Construction of Optimal Models **6**
3.1 Optimal number of predictors 6
3.2 Model skill 7
3.3 Cross-validation 8
3.4 Numerical tools 8

4 CCA Model Construction: SST models **9**
4.1 January SSTs and land surface Temperatures 9
4.1.1 Influence from North Sea SSTs 9
4.1.2 Influence from remote SSTs 12
4.2 April SSTs and land surface Temperatures 16
4.2.1 Influence from North Sea SSTs 16
4.2.2 Influence from remote SSTs 16
4.3 July SSTs and land surface Temperatures 24
4.3.1 Influence from North Sea SSTs 24
4.3.2 Influence from remote SSTs 24
4.4 October SSTs and land surface Temperatures 26
4.4.1 Influence from North Sea SSTs 26
4.4.2 Influence from remote SSTs 29
4.5 Discussion of the SST models 29

5 CCA Model Construction: SLP models **31**
5.1 Models based on the NCAR data 31
5.1.1 January North Atlantic SLPs and land surface Temperatures 31
5.1.2 April North Atlantic SLPs and land surface Temperatures 33
5.1.3 July North Atlantic SLPs and land surface Temperatures 33
5.1.4 October North Atlantic SLPs and land surface Temperatures 39
5.2 Models based on the NMC data 41

5.2.1 January North Atlantic SLPs and land surface Temperatures . . . 41
5.2.2 April North Atlantic SLPs and land surface Temperatures . . . 43
5.2.3 July North Atlantic SLPs and land surface Temperatures 43
5.2.4 October North Atlantic SLPs and land surface Temperatures . . . 51
5.3 Models based on the UEA data . . . 51
5.3.1 January North Atlantic SLPs and land surface Temperatures . . . 51
5.3.2 April North Atlantic SLPs and land surface Temperatures . . . 55
5.3.3 July North Atlantic SLPs and land surface Temperatures 58
5.3.4 October North Atlantic SLPs and land surface Temperatures . . . 58

6 Geopotential height models 58
6.1 January NMC 500hPa Geopotential heights and land surface Temperatures . . . 58
6.2 April NMC 500hPa Geopotential heights and land surface Temperatures . . . 63
6.3 July NMC 500hPa Geopotential heights and land surface Temperatures . . . 63
6.4 October NMC 500hPa Geopotential heights and land surface Temperatures . . . 69

7 Ice models 69
7.1 January ice and land surface Temperatures . . . 71
7.2 April ice and land surface Temperatures . . . 71
7.3 July ice and land surface Temperatures . . . 76
7.4 October ice and land surface Temperatures . . . 76

8 Model stationarity 81
8.1 Stationarity test for the SST models . . . 81
8.1.1 Local SST models . . . 81
8.1.2 North Atlantic SST models . . . 84
8.2 Stationarity test for the SLP models . . . 87
8.3 Discussion of the model stationarity . . . 87

9 Discussion and Conclusion 90

1 Introduction

One main objective of the RegClim project is to predict the effects of a global warming on regional scales for Norway. There are two common methods for using GCM projections to forecast local climate changes: dynamical and statistical downscaling. We will in this technical report discuss the construction of models for statistical downscaling based on a linear relationship between a predictor and a predictand. Henceforth, we will refer to the predictands, $\vec{y}$, as the quantity predicted and the predictors, $\vec{x}$, as the quantity used as input data in the prediction model.

This report is intended as a documentation of the CCA downscaling models developed at the DNMI. A large number of tables and figures describing the construction and testing of these models have therefore been included for future reference.

We begin with a discussion of technical details concerning the construction of the statistical models. The first section defines linear algebra notations employed here and is followed by sections where the equations used in Canonical Correlation Analysis (CCA) are derived. These sections are not essential for the understanding of the final results and may be skipped by those who are not interested in the model details. The section on cross-validation discusses the model results, i.e. the relationship between the predictor fields and the predictands. In this section, the model skill is evaluated. Where possible, a physical explanation is given as to how the predictors may influence the predictands. The section on model stationarity describes studies where the models have been constructed using the half of the data which approximately corresponds to the periods with lowest temperatures in the northern hemisphere, and subsequently used for prediction of the second half which is associated with warmer temperatures. This analysis is a crude sensitivity test to investigate if the assumption of constant relationship between predictors and predictands holds for a warming scenario. If the statistical downscaling models are to be used in the study of future climate change, it is important that the statistical relationship found for the training period also holds for the prediction period. The main findings are summarised in the last section.

2 CCA models

2.1 Predictors and Predictands

Linear downscaling assumes that local observations may be related to large scale circulation patterns though a simple linear statistical relationship, such as $\vec{y} = A\vec{x}$ (*Zorita & von Storch*, 1997; *von Storch* et al., 1993). In this case, the predictands and predictors contain several observations and are represented by vectors. We will use the matrix Y to refer to the time series of $\vec{y}(t)$, where the vectors are given as the columns of Y. The two data fields Y and X contain data which are sampled at p and q locations respectively over a time period with n measurements at each location:

$$\begin{aligned} Y &= [\vec{y}_1, \vec{y}_2, ..\vec{y}_n], \\ X &= [\vec{x}_1, \vec{x}_2, ..\vec{x}_n]. \end{aligned} \tag{1}$$

At time t, the data fields can be written as $\vec{y}_t = [y_{t1}, y_{t2}, ..y_{tp}]^T$. There are several techniques to find coupled patterns in climate data (*Bretherton* et al., 1992), and we will discuss the CCA method here.

The predictors discussed here only represent a small subset of all possible data sets. The discussion of models based on the few predictor data sets described here is extensive, and we try to limit the scope of this report by limiting the predictors to those quantities and levels which were available from the ECHAM4/OPYC model results. Most of the predictor data sets are described by *Benestad* (1998).

The predictands discussed here are temperature series from 24 stations obtained from the DNMI climate data base (*Hanssen-Bauer & Nordli*, 1998). We used monthly mean values of land surface temperature[1] from a number of stations, where only stations with long time series were selected, shown in figure 1. 4 stations were located in northern Norway (Vardø, Karasjok, Sihcajarvi and Tromsø), 3 sites were selected from mid Norway (Bodø, Skomvær fyr and Glomfjord), whereas the remaining 18 time series were from the southern part of Norway. There were 9 inland stations (Karasjok, Sihcajarvi, Røros, Kjøremsgrendi, Oppstryn, Lærdal, Åbjørsbråten, Flisa and Nesbyen) and 14 coastal stations. The period spanned by the predictands was 75 years, from 1923 to 1978. The reason why predictand data more up to date were not used in this study was that 1978 was the year that some of the stations with a long temperature record, such as Skomvær fyr, ended.

[1] Precipitation and other quantities will be discussed in later reports.

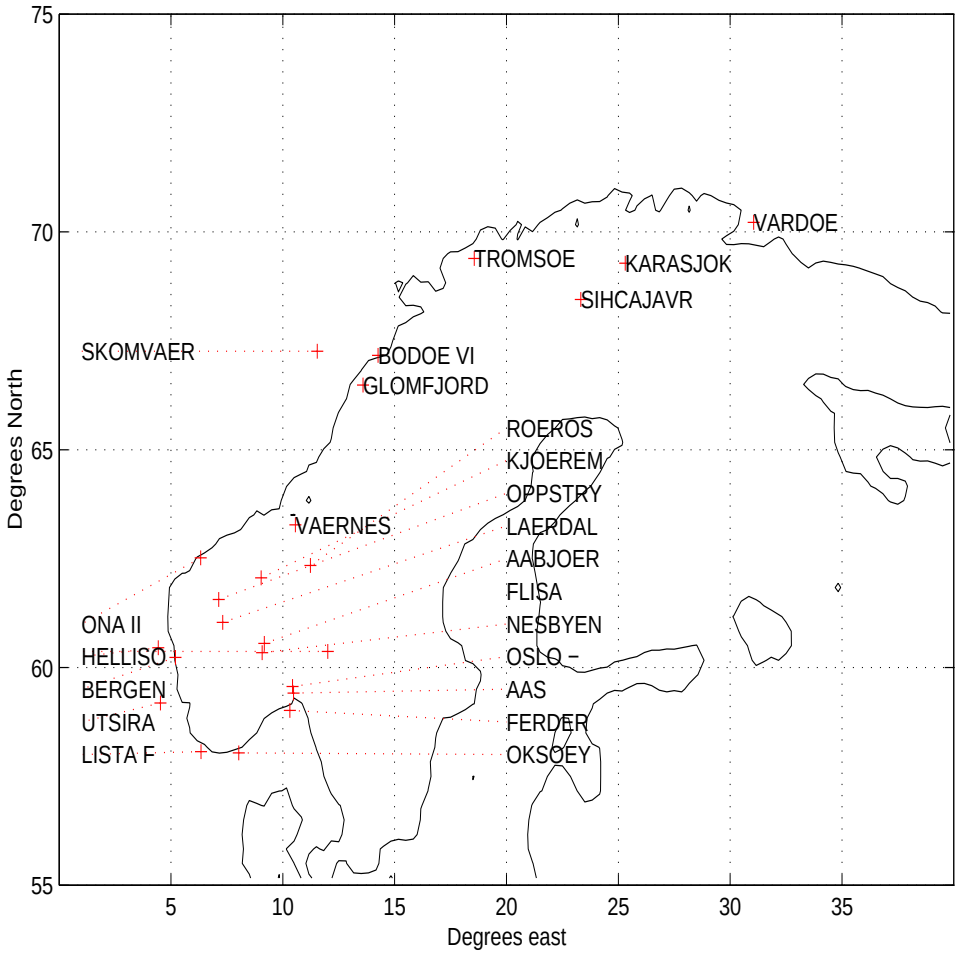

Figure 1: Map showing the location of the the stations (predictand locations) referred to in this report.

2.2 Classical CCA

CCA is a statistical method for finding spatially coherent patterns in different data fields that have the largest possible temporal correlation (*Wilks*, 1995; *Preisendorfer*, 1988). The climate data can be thought of as a linear superposition of spatially coherent patterns at any time, and the time evolution of each pattern is described by an index (the extension coefficients) that determine how much each pattern contributes to the climatic state. The CCA yields two sets of weights that give the combinations of the corresponding sets of patterns with the maximum temporal correlation. We want to find the spatial patterns that give the maximum temporal cross-correlation between Y and X. Thus we want to express the data fields as

$$\begin{aligned} Y &= GU^T, \\ X &= HV^T, \end{aligned} \tag{2}$$

where U and V, known as the *Canonical variates*, describe the time evolution that have the greatest possible correlations[2], and G and H are the spatial patterns associated with these. One important property of the canonical variates, U and V, is that each canonical variate is uncorrelated with all the canonical variates in the opposite set with the exception of the corresponding canonical variate (*Wilks*, 1995, p.400). The rotation matrices L and R, also referred to as the canonical correlation weights or canonical correlation vectors, satisfy the following properties: $L^T L = I$ and $R^T R = I$ (*Preisendorfer*, 1988, p.299). Mathematically, the analysis can be posed as a maximization problem, which can be expressed in the form of an eigenvalue equation. The temporal correlations are given as:

$$U^T V = LMR^T = C. \tag{3}$$

The correlation matrix M contains the correlation coefficients on its diagonal and all off-diagonal elements are zero when the columns in U and V are optimally correlated. By using the fact that the transpose of the rotation matrices equals their inverse (*Strang*, 1995; *Press* et al., 1989), equation 3 can written as:

[2]The leading column of each Canonical variate holds the time series which have the highest possible correlation, and the subsequent columns *must* be orthogonal to the respective leading column (*Bretherton* et al., 1992). The second Canonical variate would represent the highest possible correlation of the data if the first Canonical variates and corresponding patterns were excluded from the data. The third column gives the highest correlation if the first 2 leading Canonical variates were removed before analysis, and so on.

$$\begin{aligned} CR &= LM, \\ C^T L &= RM^T. \end{aligned} \tag{4}$$

By operating C on the the last of the equations 4, we get $CC^T L = CRM^T = LMR^T RM^T = LMM^T$, and we can now re-write the equations in the form of eigenequations where rotation vectors in the columns of L and R are the eigenvectors (*Preisendorfer*, 1988, p.302):

$$\begin{aligned} \left(C^T C\right) L &= L(MM^T), \\ \left(CC^T\right) R &= R(M^T M). \end{aligned} \tag{5}$$

In order to solve the eigenvalue equation, the normalised covariance matrices, which are subject to the maximization, must be estimated:

$$\begin{aligned} C_{YY} &= Y^T Y, \\ C_{XX} &= X^T X, \\ C_{YX} &= Y^T X. \end{aligned} \tag{6}$$

The matrix product, C is a normalised covariance matrix:

$$C = C_{YY}^{-0.5} C_{YX} C_{XX}^{-0.5}, \tag{7}$$

and can be diagonalised using the SVD algorithm (*Press* et al., 1989; *Strang*, 1995):

$$C = LMR^T. \tag{8}$$

In equation 8 L and R are left and right rotation matrices respectively which yield an optimal weighted combinations of the original time series. M is a diagonal matrix with the canonical correlation values in descending order on its diagonal. The CCA maps, H and G, can be calculated from the covariance and the rotation matrices:

$$H = C_{YY} C_{YY}^{-0.5} L, \tag{9}$$

$$G = C_{XX} C_{XX}^{-0.5} R. \tag{10}$$

CCA extension coefficients (describing time evolution) can be calculated from the rotation matrices and the original data:

$$U = C_{YY}^{-0.5} LY, \tag{11}$$
$$V = C_{XX}^{-0.5} RX. \tag{12}$$

2.3 Linear relationships

The statistical models described here are based on linear relationships between the predictors and predictands and can be expressed as (*Heyen* et al., 1996):

$$\begin{aligned} \hat{Y} &= GMV^T, \\ \hat{X} &= HMU^T. \end{aligned} \tag{13}$$

The first of equations 13 can be written as

$$\hat{Y} = \Psi X, \tag{14}$$

where the matrix Ψ is the statistical model that can be used for prediction. The canonical variates, U and V, the canonical correlation maps, G and H, and the correlation matrix, M, form the basis for the statistical model. The two data fields Y and X are related to these CCA products according to equation 2, and the canonical variate V can estimated as $V^T = (H^T H)^{-1} H^T X$. Y can therefore be predicted from X according to:

$$\hat{Y} = GM(H^T H)^{-1} H^T X \tag{15}$$

The projection of X onto Y gives the predicted values of Y, and is denoted as $\hat{Y}$. The CCA model is the matrix $\Psi = (GM(H^T H)^{-1} H^T)$, where G and H are the canonical patterns and M is the diagonal matrix with the canonical correlations along its diagonal.

3 Construction of Optimal Models

3.1 Optimal number of predictors

In the development of the statistical models, it is important to find the optimal number of predictands that yields the best prediction scores. The number and type of predictors must be selected carefully in order to maximise the skill and avoid overfitting (*Wilks*, 1995, p.185). For instance, combinations of noise or signals unrelated to the predicted quantity may give a good fit

to the data used in the training of the model, but will usually not produce good predictions. One method to construct models with optimal skill and avoid overfit involves cross-validation and the use of a screening technique to estimate the optimal number of predictors.

The predictors described in this report were the principal components (PC) of the gridded data (*Benestad*, 1998). A spatial weighting function $W = W_x \times W_y$, had been applied to the data prior to the PCA, where $W_x = |\sqrt{\cos(\pi\theta/360)}|$ and $W_y = |\sqrt{\cos(\pi(\phi - 60)/90)}|$. The reason for not removing the unwanted remote areas all together was that we wanted to retain all the spatial grid points in order to get a good estimate of the covariance matrix and hence a better estimate of the spatial patterns (We need more spatial data points than temporal data points). The data used in the models here were not subject to de-trending prior to the analysis. In other words, the empirical models are sensitive to slow climate changes[3].

3.2 Model skill

How do we define which model is the best? It may be *(i)* the model that yields the best score for one or a small number of stations, or *(ii)* the model which produces the highest average score for all stations. The choice depends on the nature of the problem we want to address with the models. It is for instance possible to construct empirical models with optimal skills near the larger cities for estimating scenarios for energy consumption associated with heating. We will in this report focus on the models that gives highest prediction scores for one or a small number of stations (the station with highest skill score), as the model skill for an area as large as Norway may not be very good.

There are different ways of measuring skill, such as root mean squared (RMS) errors, variance accounted for by prediction, or correlation coefficients[4]. Again, different skill measures are appropriate for different types of forecasts. In a global warming scenario, for instance, we may want to know how much the local winter temperatures will vary from year to year or how strong the maximum winds are going to be, e.g. the variance of the predictands. In this case, it is important to use models which skillfully predict the signal variance (i.e. where the predicted signal accounts for about 100% of the observations during the validation period). We will in this report use the

[3]De-trended data gives a better statistical interpretation, as the probability for correlation between low frequency variability being coincidental is higher than for high frequency signals.

[4]Other skill scores, such as linear error in probability space (LEPS) and the Brier Score for probability forecasts (*Wilks*, 1995) will not be discussed here.

highest correlation to define the optimal model, but also discuss RMS error (RMSE) as an additional skill measure.

3.3 Cross-validation

The cross validation approach excludes one data point in the construction of a statistical model and subsequently uses the model to predict the value of the predictand that was excluded from the model calibration (*Wilks*, 1995; *Kaas* et al., 1998). The data not used for calibration of the model are referred to as *independent data*. The process of excluding one data point is repeated N times, where N is the total number of observations, and for each iteration different data points are used as independent data. The cross-validation method therefore implies the construction of N different models which are based on different combinations of data.

3.4 Numerical tools

A Matlab script that estimates the CCA products is based on *Bretherton* et al. (1992)[5] and can be found in the Matlab file *cca.m*. This routine is called from the scripts *cmpcca.m*, *optmod.m*, and *crossval.m*. The Matlab script, *cca.m*, estimates Ψ and *crossval.m* performs a cross validation test. The latter script also employs the singular value decomposition (SVD) method and multivariate regression (MVR) methods[6], and tests have been carried out with monthly mean values of station observations as predictands and sea surface temperatures (SSTs), ice, geopotential heights (Φ), and sea level pressure (SLP) as predictors. The Matlab code estimates the mean and standard deviation of the model coefficients (weights) for the leading CCA, SVD or MVR patterns, G_{1i}. If the predictands are equally influenced by the leading CCA pattern and the higher order patterns (with almost similar correlation coefficients), then the models may be sensitive to the different combinations of data used in for calibration, as different patterns may contribute to the estimate of the leading CCA predictand weights. In this case, the weights are associated with large standard deviations. This situation may not necessarily imply poorly estimated model. If, however, the predictands are not

[5]Equations 12 differ from the equations given by *Bretherton* et al. (1992) in that we use $u_k = C_{YY}^{-0.5} L$ instead of their expression $u_k = C_{YY}^{0.5} L$. Their formulae did not give realistic results in Matlab as their expression was incorrect and the amplitudes of the predictions were too low by an order of 3. The canonical vectors are usually normalised, and $C_{XX}^{-0.5}$ in equation 12 ensures that $U^T U = I$ as $L^T L = I$. Since $C_{XX}^{0.5}$ has the same magnitude as X, the Canonical vectors in *Bretherton* et al. (1992) scale as X^2.

[6]The SVD and MVR models will be discussed in future issues of DNMI-report *KLIMA*.

well described by the predictor field, this could also give large standard deviations, maybe as a result of an overfit or lack of a physical relationship. It is also possible that the empirical relationships between circulation patterns and surface variables vary over decadal time scales (*Wilby*, 1997)[7], and therefore spread in the coefficient estimates may also imply non-stationary relationships between the predictors and predictands.

A Matlab script for constructing optimal empirical models is found in the file *modopt.m*. This routine loops through all possible (20) EOFs and tests whether the inclusion of each EOF increases the cross-validation prediction skill. The code in *modopt.m* calls the function *crossval3.m*, which performs the cross-validation analysis. The Matlab scripts *crossval2.m* (and *crossval.m*. The scripts *crossval.m*, *crossval2.m*, and *crossval3.m* are slightly different version of the same analysis.) was used to plot the results.

4 CCA Model Construction: 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, and will be referred to as the 'North Sea model', although SSTs from the Barent 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 Mediterranean, the Black Sea, the Labrador Sea, the Barent Sea, the Norwegian Sea, Skagerrak, Kattegat, the Baltic Sea, and Hudson Bay also are included.

4.1 January SSTs and land surface Temperatures

4.1.1 Influence from North Sea SSTs

CCA was applied to January mean SSTs from GISST2.2 and the January mean DNMI station temperatures, and table 1 gives a summary of the cross-validation prediction scores using CCA models based on different combinations of EOFs from the GISST2.2 January mean North Sea SSTs as predictors, and January mean temperatures from a selection of Norwegian stations as predictands. The left column indicates which EOFs were included in the predictor set for the particular model. The numbers listed in columns 2, 3 and 4 represent the scores for the model based on the EOF combination given in the first column. The second column gives the cross-validation correlation

[7] *Wilby* (1997) analysed daily observations, which may differ from monthly mean values.

Table 1: Evaluation of January temperature CCA model based on North Sea SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	FERDER FYR r= 0.64 rmse= 0.26	SKOMVÆR FYR r= 0.39 rmse= 0.17	TROMSØ r= 0.23 rmse= 0.25
1 2	FERDER FYR r= 0.70 rmse= 0.25	ONA II r= 0.51 rmse= 0.17	TROMSØ r= 0.16 rmse= 0.25
1 2 3 6 7 8	FERDER FYR r= 0.72 rmse= 0.24	SKOMVÆR FYR r= 0.49 rmse= 0.16	KARASJOK r= 0.40 rmse= 0.56

Table 2: Evaluation of January temperature CCA model based on North Atlantic SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	FLISA r= 0.60 rmse= 0.45	SKOMVÆR FYR r= 0.51 rmse= 0.15	UTSIRA FYR r= 0.31 rmse= 0.22
1 2	FLISA r= 0.62 rmse= 0.44	SKOMVÆR FYR r= 0.54 rmse= 0.15	UTSIRA FYR r= 0.40 rmse= 0.21
1 2 8 9 10	FLISA r= 0.70 rmse= 0.40	SKOMVÆR FYR r= 0.52 rmse= 0.15	OPPSTRYN r= 0.40 rmse= 0.27
2 8 9 10 11 14 17	LISTA FYR r= 0.55 rmse= 0.27	ONA II r= 0.36 rmse= 0.19	TROMSØ r= -0.08 rmse= 0.28
1 2 8 9 10 11 14 17	FLISA r= 0.75 rmse= 0.37	SKOMVÆR FYR r= 0.48 rmse= 0.16	BODØ VI r= 0.43 rmse= 0.26

and RMS error skill scores for the location with the highest correlation coefficient. The third column shows the scores for the station with the smallest RMS errors, and the right column gives the skill details of the stations with the lowest correlation scores, i.e. the 'worst' prediction.

The correlation analysis shown was based on the independent data, and table 1 suggests that a combination of EOFs $[1, 2, 3, 6, 7, 8]$ gave the optimal SST model for the coastal regions in the southern Norway.

Figure 2 shows the leading predictand CCA weights for the January mean land surface temperatures, estimated from a sample of N models constructed with different data combinations used for model calibration. Also shown is the spread of these mean estimates (error bars), and the larger the spread, the greater uncertainty in the estimate. The spread in estimated predictand weights was small when only the 'optimum' EOFs were included in the predictor fields. The strongest weights were found at the inland stations Nesbyen and Flisa, where there were large year to year temperature variations, and the smallest coefficients were estimated at the coastal stations Tromsø and Skomvær fyr, where the temperature amplitudes were smaller. The statistical models are expected to be stationary with respect to different data combinations used in the model construction, unless the relationships between the large scale patterns and the local climatic variable change over time (*Wilby*, 1997) or the model suffers from an overfit. The fact that the standard deviations were relatively small in figure 2, suggested that the relationship between large scale predictor and small scale predictand was approximately constant.

The (cross-validated) predicted temperatures at Ferder fyr, shown in figure 3, had the highest correlation with the corresponding observations, i.e. best correlation skill with $r = 0.72$, where the 96% significance level was estimated to be 0.32, and RMSE=0.24°C. It is evident from figure 3 that the local January SST model captured the late 1930s and early 1940s cold period extremely well, but missed many of the later smaller fluctuations. The prediction for Ferder fyr appeared to have higher correlation with the observations prior to 1943, when the model captured most of the small temperature variations.

Figure 4 shows the leading CCA SST pattern that accounted for the highest correlation with the station temperatures. The CCA weights were greatest near Skagerrak, Kattegat, and parts of the Baltic Sea. The predictions for most of the stations in the southern part of Norway were highly correlated with the observations, while those in the northern regions showed lower correlations (figure 2). The RMS errors in the southern part of Norway were less than 0.5°C (except for Røros), also indicative of high prediction skills. The temperatures in the coastal areas tended to have better prediction skills than the observations further away from the coast (with exception of

Flisa, figure 2). This pattern of prediction skills could be explained as the North Sea SSTs having a strong local effect.

4.1.2 Influence from remote SSTs

The influence of the large scale SST patterns in the North Atlantic on the Norwegian January temperatures was investigated by applying a similar CCA as described above to SST EOFs from the whole North Atlantic basin. The leading North Atlantic SST EOF, describing strong SST variability near Cape Hatteras on the east coast of the USA and the subtropical Atlantic, produced high correlation values with the January temperatures from all stations. Analysis with the leading SST EOF excluded, gave much lower correlation coefficients and accounted for less variance than if the leading EOF (*Benestad*, 1998) was included in the analysis (table 2). The leading CCA predictor pattern associated with the optimal model (We will also refer to the leading CCA structure as 'the predictor pattern') had strong weights near Cape Hatteras and the interior North Atlantic as well as in the Baltic Sea (figure 5). The fact that CCA models using SSTs from the entire North Atlantic produce higher correlation scores (figure 6) than models that only include the local SSTs provides strong observational evidence for large scale SST patterns similar to the coupled mode structure identified by *Grötzner* et al. (1998), *Sutton & Allen* (1997), *Latif* et al. (1996) and *Deser & Blackmon* (1993), affecting the large scale winter time atmospheric circulation. One possible explanation of the physical mechanism is that SST anomalies associated with the Gulf stream or north-south excursions of the Gulf stream off the American east coast affect the Norwegian winter temperatures by shifting the storm tracks location.

Many of the other inland stations in southern Norway scored high in the cross-validation analysis (figure 6), suggesting that it was the large scale North Atlantic SST anomalies which were responsible for the prediction skills seen here as opposed to local SSTs influencing only the coastal climate. Figure 7, showing the cross-validation predictions for Flisa, indicates that the North Atlantic SST model captured the major events, but had a tendency to exaggerate the smaller fluctuations. The correlation skill was 0.75, with RMS error of 0.37°C, and accounting for 66% of the variance.

The optimum North Atlantic SST models was based on EOFs 1, 2, 8, 9, 10, 11, 14, and 17. The lowest correlation skill, 0.43, which was still marginally greater than the 95% confidence level, was found in Bodø (table 2).

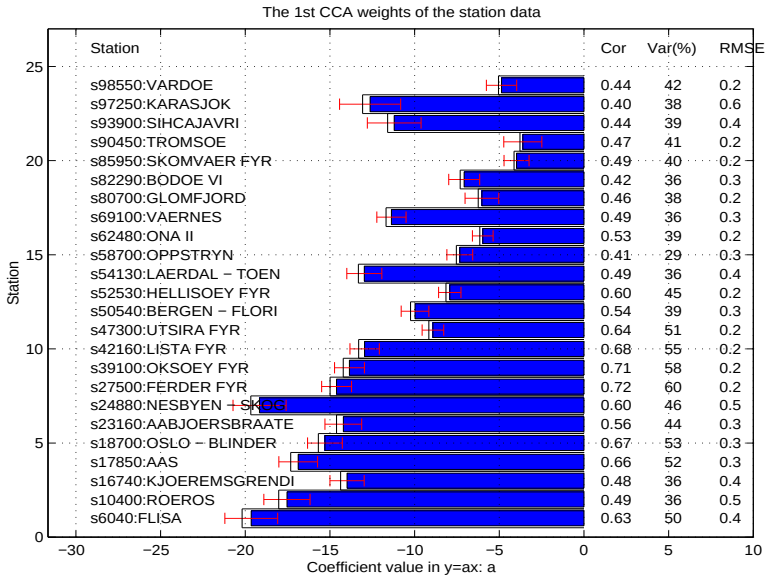

Figure 2: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading North Sea GISST2.2 January North Sea SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

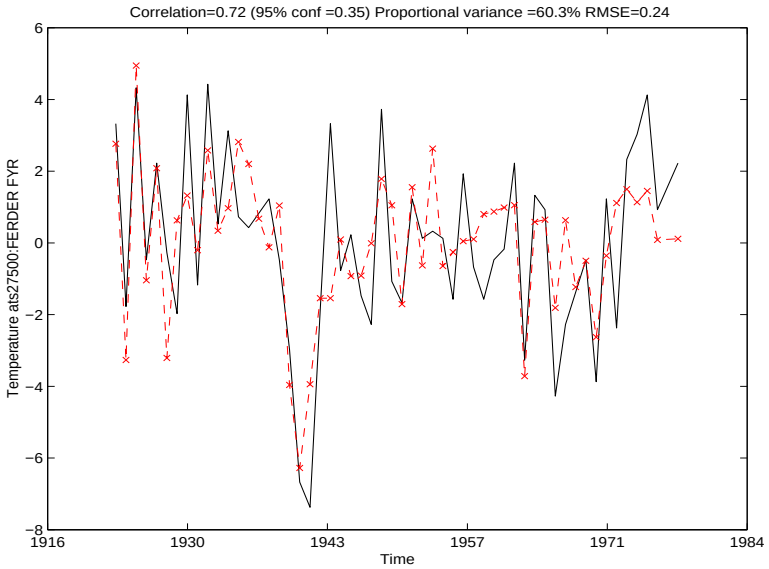

Figure 3: Time series of predicted January temperatures (dashed) at Ferder Fyr, employing the cross-validation method with GISST2.2 North Sea SSTs, shown with the observations (black solid line).

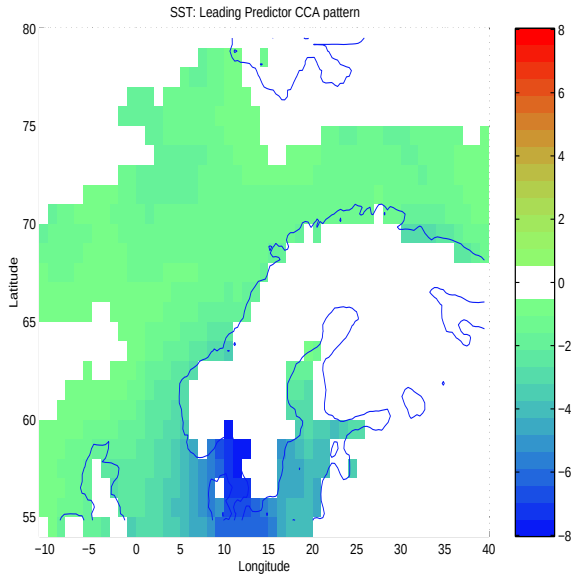

Figure 4: The mean leading January CCA GISST2.2SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA North Sea SST pattern estimates are greater than 1°C are not shown.

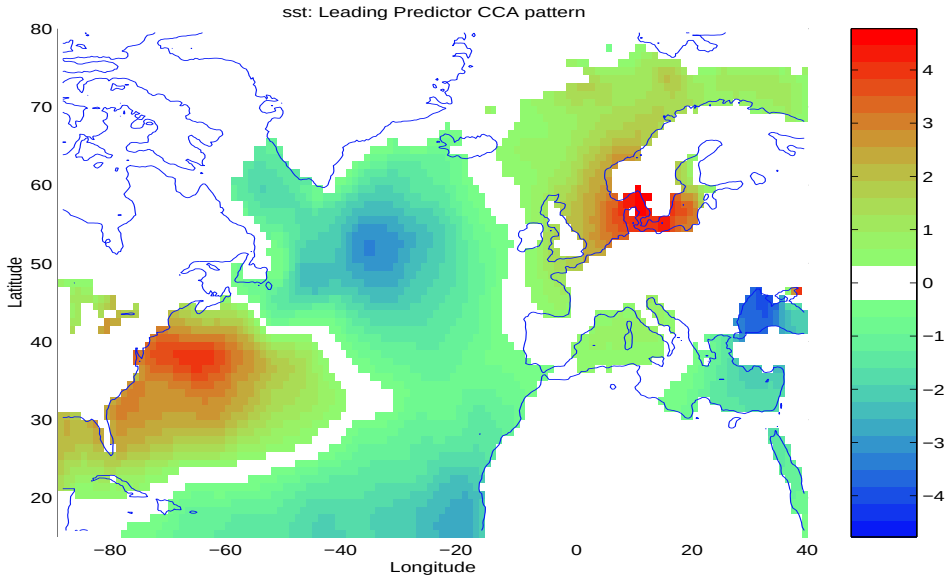

Figure 5: The mean leading January CCA GISST2.2 SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA North Atlantic SST pattern estimates are greater than 1°C are not shown.

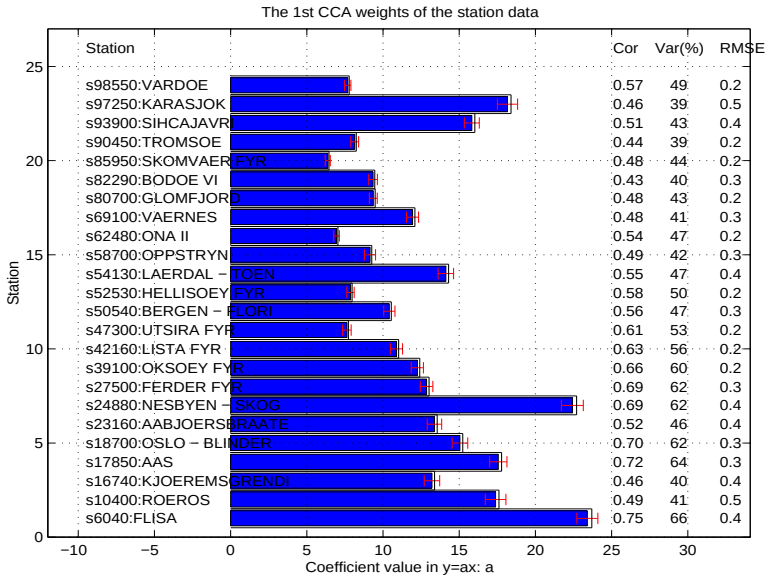

Figure 6: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading GISST2.2 January North Atlantic SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

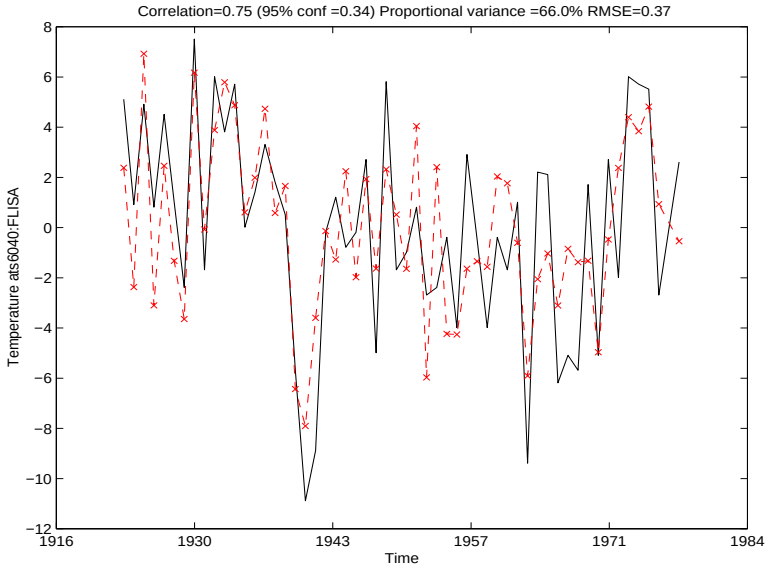

Figure 7: Time series of predicted January temperatures (dashed) at Flisa, employing the cross-validation method with GISST2.2 North Atlantic SSTs, shown with the observations (black solid line).

4.2 April SSTs and land surface Temperatures

4.2.1 Influence from North Sea SSTs

The optimal April North Sea SST model included EOFs: 1 2 3 6 8 10 17, and 18, and the highest correlation score of 0.65 was found at Utsira fyr (figures 8 and 9), and was marginally lower than for the corresponding January model. The worst prediction skill was r=0.15 at Karasjok (table 4). The estimates of the CCA weights were associated with large standard deviations at all locations, and the CCA correlation coefficients were: 0.8846, 0.8472, 0.7805, 0.6808, 0.6560, 0.6033, 0.5871, and 0.4141. It is therefore unlikely that the spread in the weight estimates was due to contribution from several unresolved CCA patterns, but rather reflected the weak influence of the April SSTs on the land temperatures. The leading CCA predictor pattern, shown in figure 10 was characterised by large weights along the west coast of Norway. The mean predictor pattern was associated with large standard deviations, and the Skagerrak, Kattegat, and Baltic Sea have been masked out due to high uncertainties in the predictor weights.

4.2.2 Influence from remote SSTs

The large scale features in the leading CCA predictor pattern were similar to the January SST structures, but with stronger SST anomalies in the Mediterranean and weaker anomalies off Cape Hatteras, in the Black Sea and in the interior North Atlantic (figure 11). The April correlation skill scores (figure 12) were lower than the January skill scores, and a comparison between regional (North sea) and large scale (North Atlantic) SSTs, reveals a stronger relationship between the Norwegian April temperatures and local SSTs than remote SSTs. Figure 12 indicates small spread in the model coefficient estimates, and like January the weights tended to be larger in the south. The correlation skills were highest on the west coast of southern Norway, and the correlation scores in Tromsø dropped from 0.44 in January to merely 0.09 in April (table 4). The prediction of the April temperatures at Ferder fyr are shown in figure 13, with r=0.66, RMS error of 0.24°C and accounting for 62% of the variance. The North Atlantic SST model captured the major events, such as the cooling in the late 1930s and early 1940s and the warming in the 1970s, but not some of the smaller fluctuations.

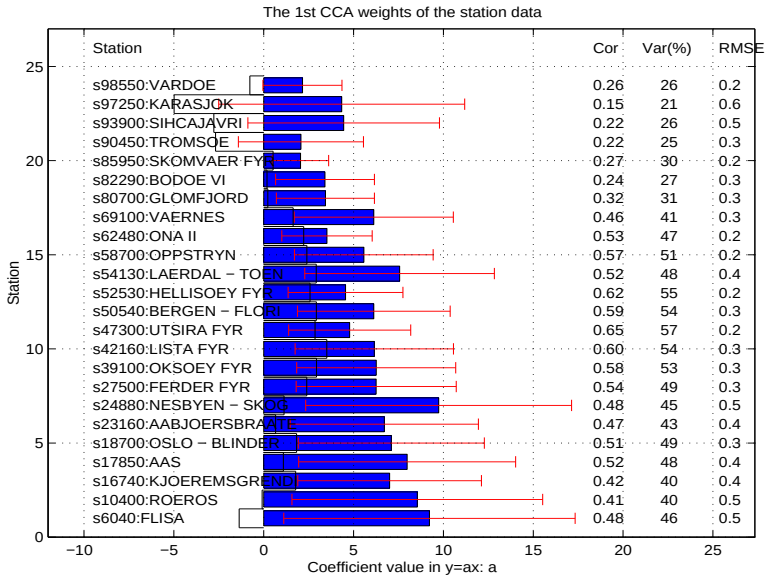

Figure 8: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading North Sea GISST2.2 April North Sea SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

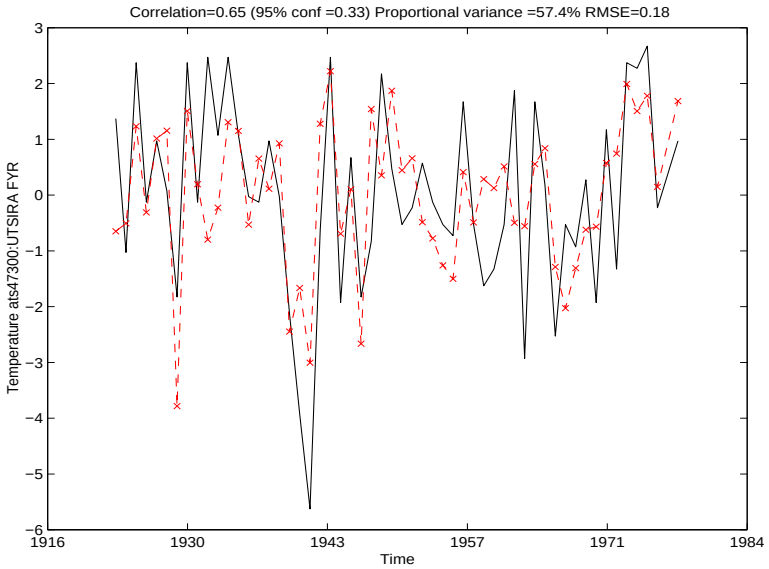

Figure 9: Time series of predicted April temperatures (dashed) at Utsira fyr, employing the cross-validation method with GISST2.2 North Sea SSTs, shown with the observations (black solid line).

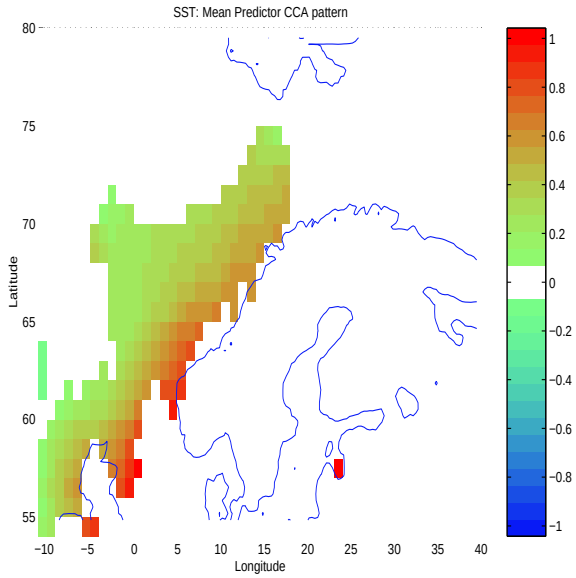

Figure 10: The mean leading April CCA North Sea GISST2.2 SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

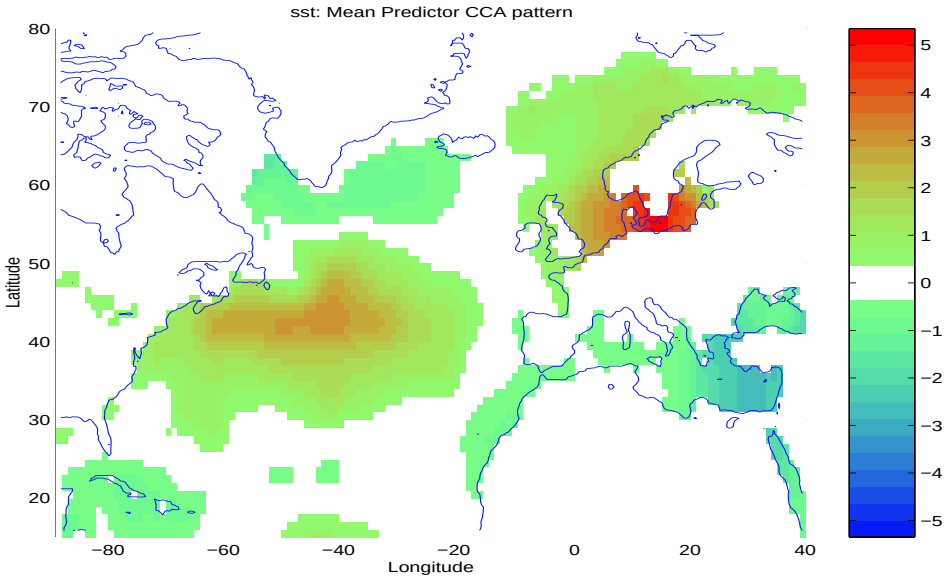

Figure 11: The mean leading April CCA North Sea SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

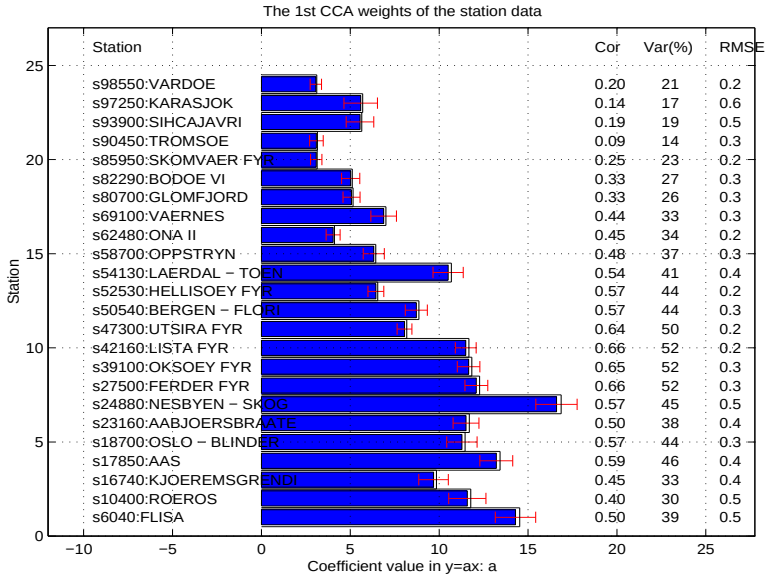

Figure 12: The mean leading April CCA GISST2.2 North Atlantic SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

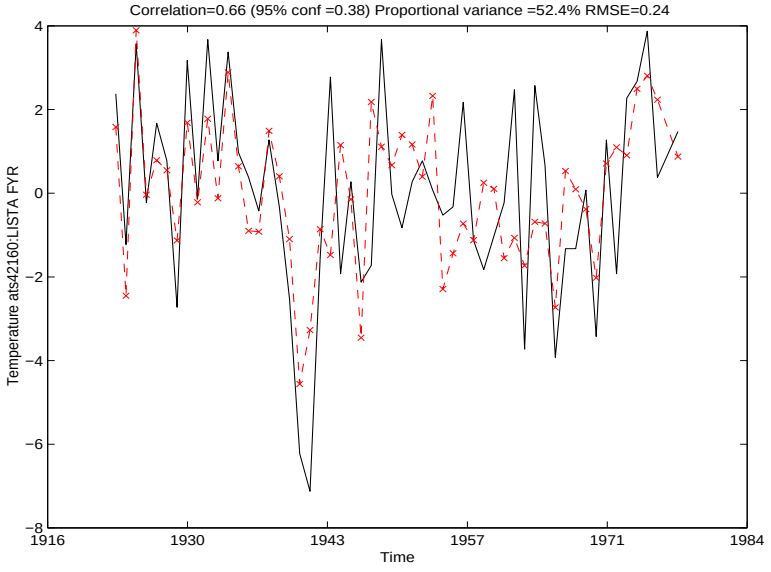

Figure 13: Time series of predicted April temperatures (dashed) at Lista fyr, employing the cross-validation method with GISST2.2 North Atlantic SSTs, shown with the observations (black solid line).

Table 3: Evaluation of April temperature CCA model based on North Sea SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	UTSIRA FYR r= 0.42 rmse= 0.21	SKOMVÆR FYR r= 0.08 rmse= 0.18	KARASJOK r= -0.36 rmse= 0.61
1 2 3 6 8 10 17 18	UTSIRA FYR r= 0.65 rmse= 0.18	ONA II r= 0.53 rmse= 0.17	KARASJOK r= 0.15 rmse= 0.62

Table 4: Evaluation of April temperature CCA model based on North Atlantic SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VÆRNES r= 0.25 rmse= 0.36	SKOMVÆR FYR r= 0.15 rmse= 0.18	VARDØ r= 0.03 rmse= 0.21
1 2 3	FERDER FYR r= 0.60 rmse= 0.28	SKOMVÆR FYR r= 0.27 rmse= 0.17	TROMSØ r= 0.16 rmse= 0.26
1 2 3 6 8 10	LISTA FYR r= 0.66 rmse= 0.24	ONA II r= 0.45 rmse= 0.18	TROMSØ r= 0.09 rmse= 0.26

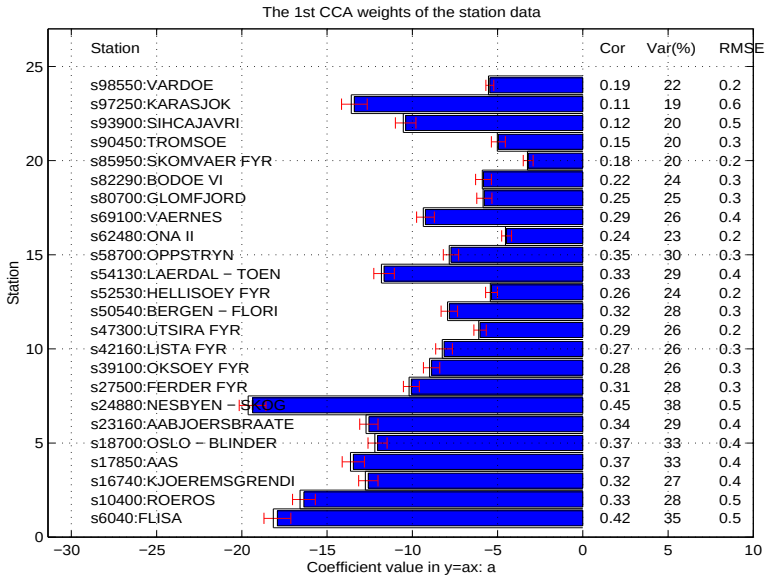

Figure 14: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading North Sea GISST2.2 July SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

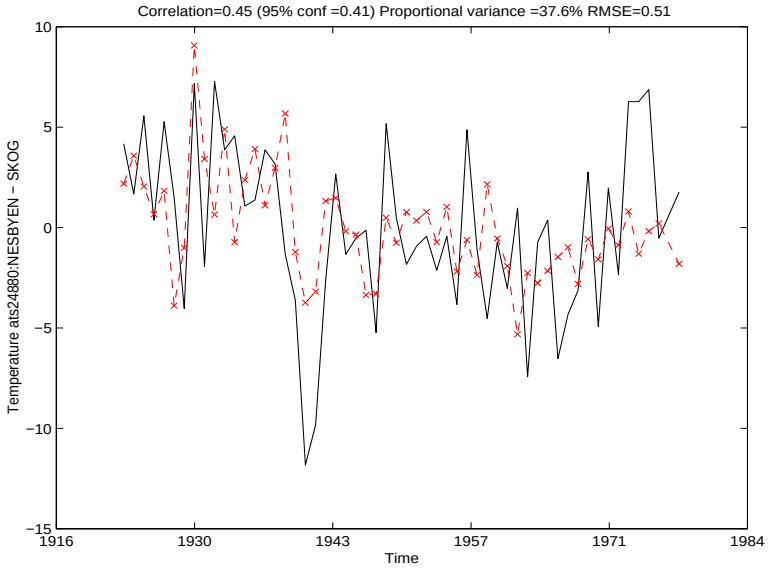

Figure 15: Time series of predicted July temperatures (dashed) at Utsira fyr, employing the cross-validation method with GISST2.2 North Sea SSTs, shown with the observations (black solid line).

Table 5: Evaluation of July temperature CCA model based on North Sea SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	NESBYEN - SKOGLUND r= 0.24	SKOMVÆR FYR r= 0.11	LISTA FYR r= 0.09
1 4 10 11 18 20	NESBYEN - SKOGLUND r= 0.45 rmse= 0.51	SKOMVÆR FYR r= 0.18 rmse= 0.18	KARASJOK r= 0.11 rmse= 0.63

Table 6: Evaluation of July temperature CCA model based on North Atlantic SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
2	TROMSØ r= 0.26 rmse= 0.25	SKOMVÆR FYR r= -0.07 rmse= 0.18	NESBYEN - SKOGLUND r= -0.65 rmse= 0.58
2 4 5 6 9 10 14 15 16 17	KARASJOK r= 0.61 rmse= 0.49	SKOMVÆR FYR r= 0.54 rmse= 0.15	FERDER FYR r= 0.16 rmse= 0.38

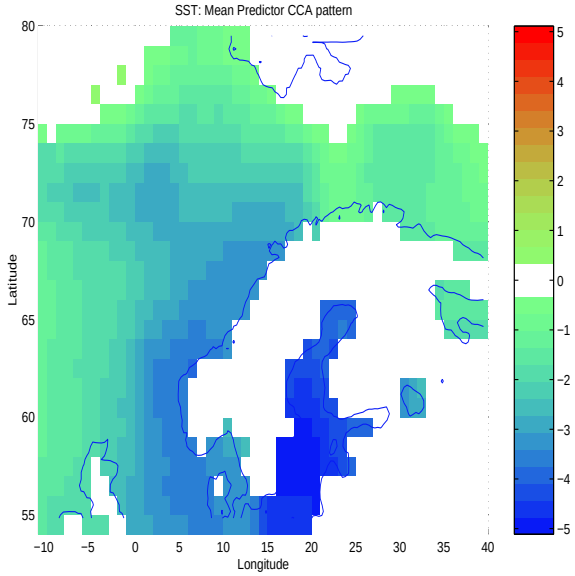

Figure 16: The mean leading July CCA North Sea GISST2.2SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

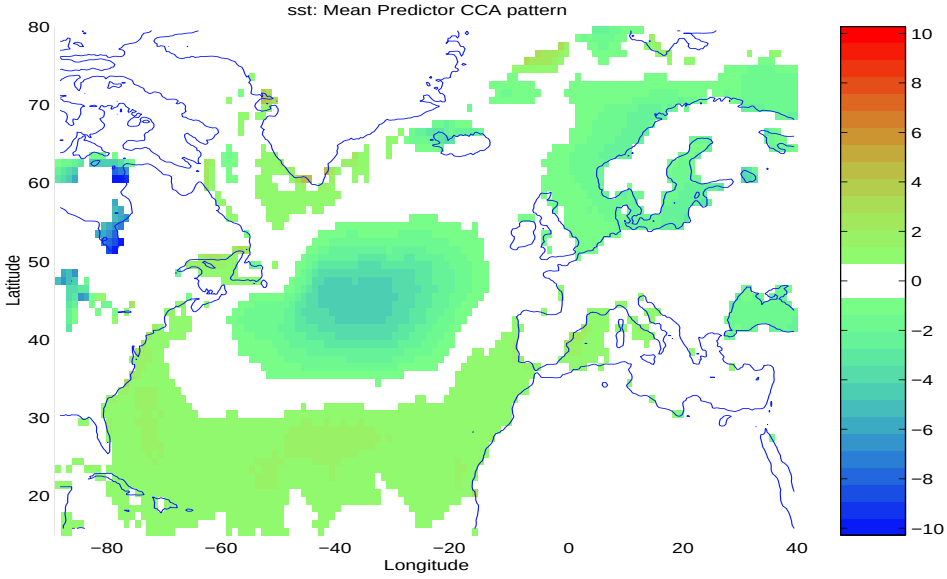

Figure 17: The mean leading July CCA North Atlantic SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

4.3 July SSTs and land surface Temperatures

4.3.1 Influence from North Sea SSTs

The relationship between the July SSTs and land temperatures was weak (maximum r=0.45, table 5), with the optimum EOF combination in the predictor set consisting of EOFs 1, 4, 10, 11, 18, and 20. The CCA weight estimates had low standard deviations, but the correlation skill was generally low and barely significant for the best predictions (figure 14). The highest prediction scores were found at Nesbyen (figure 15) and Flisa, both inland stations, and the CCA predictor pattern indicated strong SST weights in the Baltic Sea (figure 16). This result suggests that warm Baltic SSTs were associated with warm Norwegian July temperatures, although the predictions could only account for about 30-40% of the July temperature variability. The highest correlation score was 0.45, the July North Sea SST model could only account for a fraction of the temperature anomalies during the major events, but did not describe all the smaller fluctuations (figure 15). The model also predicted some spurious smaller events in the 1960s. It is possible that both the SSTs and the land temperatures were both forced by the same large scale circulation pattern (i.e. High pressure system).

4.3.2 Influence from remote SSTs

The leading CCA predictor pattern for the July North Atlantic SSTs was characterised by strongest SST anomalies over a large area of the interior Northern Atlantic (figure 17), and the best July skill scores (figure 18) were lower than those for January (figure 6) and April (figure 12). However, the highest prediction scores for the North Atlantic SSTs (r=0.61, figure 19) were higher than for the regional SSTs (r=0.45, figure 15), indicating at least some influence from the distant large scale maritime regions. The lowest correlation was 0.16 and found at Ferder fyr.

The estimates of the July model indicated substantially larger weights in northern Norway, where the correlation scores also were the highest. The best prediction skill was found at Karasjok (figure 19), where the cross-validation correlation coefficient was 0.61, the RMS error was 0.49°C, and the predicted signal accounted for 64% of the variability. Figure 19 indicates that the July North Atlantic SST model captured most of the major events, but did predict too warm temperatures during the 1970 warm event and missed some of the smaller peaks. The reduction of prediction skill and weights to the south in figure 18 is quite remarkable. Experiments with dynamical models are needed in order to give a physical explanation for this observation.

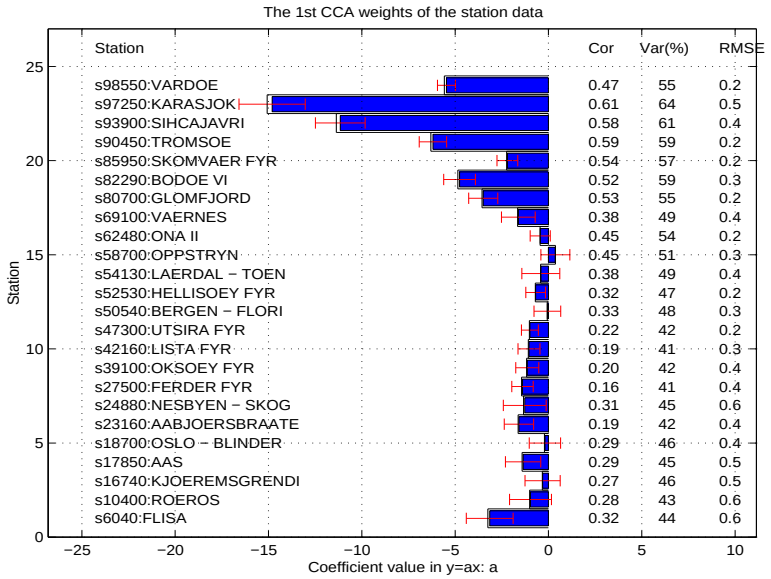

Figure 18: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July GISST2.2 North Atlantic SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

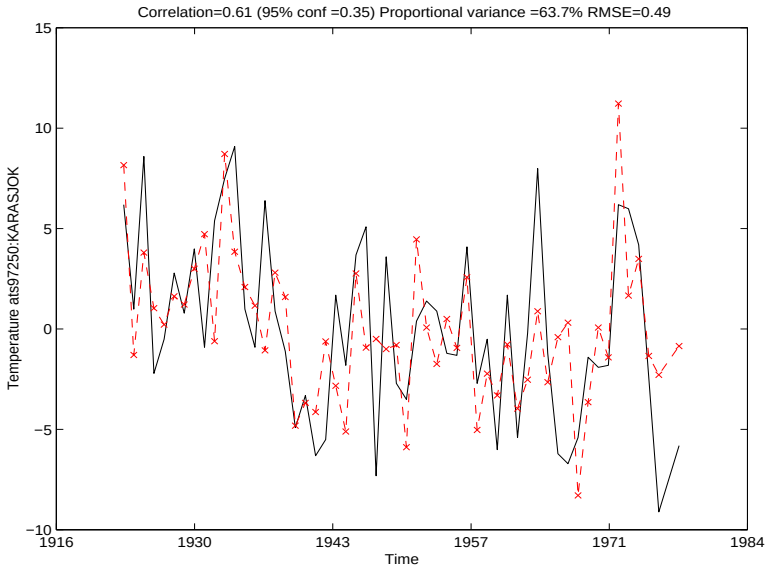

Figure 19: Time series of predicted July temperatures (dashed) at Karasjok, employing the cross-validation method with GISST2.2 North Atlantic SST, shown with the observations (black solid line).

Table 7: Evaluation of October temperature CCA model based on North Sea SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.07 rmse= 0.21	SKOMVÆR FYR r= -0.27 rmse= 0.18	ÅS r= -0.55 rmse= 0.45
1 2 3 5 7 9 10 15 20	TROMSØ r= 0.48 rmse= 0.23	SKOMVÆR FYR r= 0.34 rmse= 0.17	LISTA FYR r= -0.30 rmse= 0.37

Table 8: Evaluation of October temperature CCA model based on North Atlantic SSTs from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	TROMSØ r= 0.12 rmse= 0.25	SKOMVÆR FYR r= -0.38 rmse= 0.18	ÅBJØRSBRÅTEN r= -0.73 rmse= 0.42
1 2	TROMSØ r= 0.24 rmse= 0.25	SKOMVÆR FYR r= 0.07 rmse= 0.18	UTSIRA FYR r= -0.12 rmse= 0.24
1 2 6 8 13	TROMSØ r= 0.37 rmse= 0.24	SKOMVÆR FYR r= 0.15 rmse= 0.18	LISTA FYR r= -0.19 rmse= 0.35

4.4 October SSTs and land surface Temperatures

4.4.1 Influence from North Sea SSTs

The highest correlation score for the October North Sea SST model was 0.48 in Tromsø, and the worst prediction was for Lista fyr where r=-0.30 (table 7)[8]. Figure 20 shows a slight cooling trend for Tromsø for both observations and predictions and the correlation skill score may have been 'artificially inflated' as a result of this common trend. The October North Sea model missed the

[8]Negative correlation coefficients do not represent any skill in validation of predictions, but merely indicate the fact that the model predicts warm events during cold months and vice versa. Thus negative correlation coefficients from the cross-validation analysis indicate that the predictions were 'misleading'.

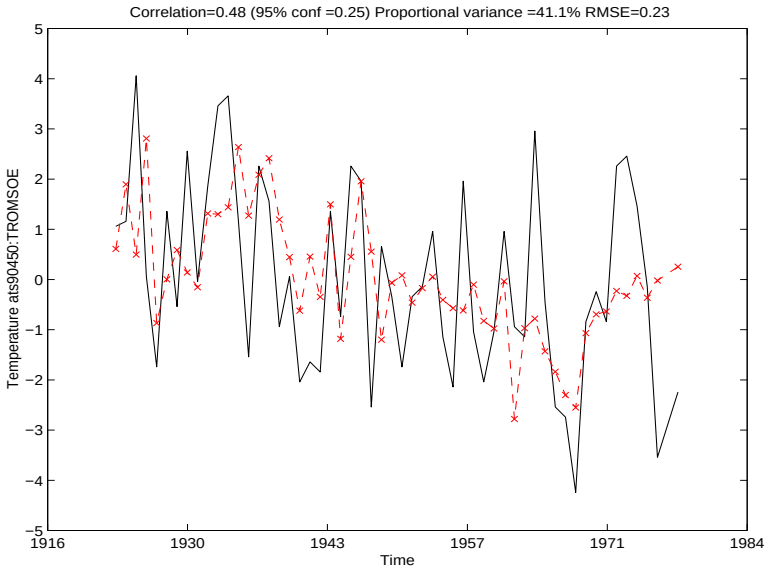

Figure 20: Time series of predicted October temperatures (dashed) at Karasjok, employing the cross-validation method with North Sea GISST2.2 SST, shown with the observations (black solid line).

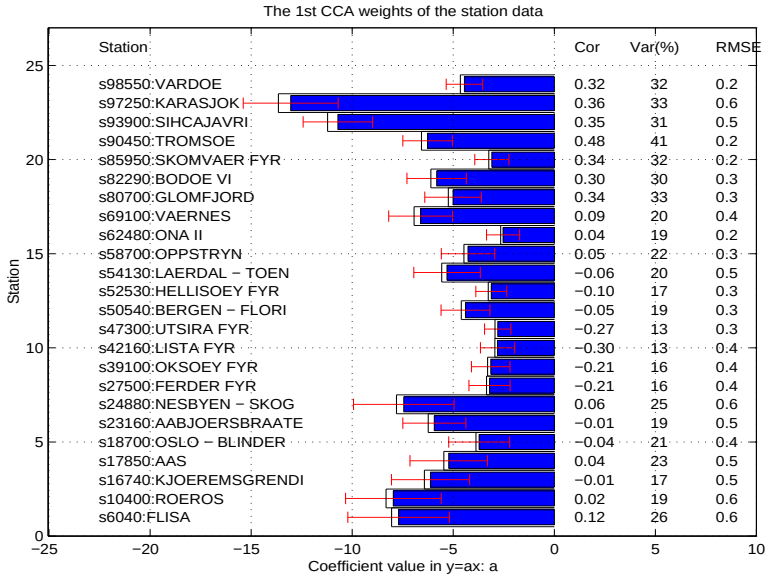

Figure 21: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October GISST2.2 North Sea SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

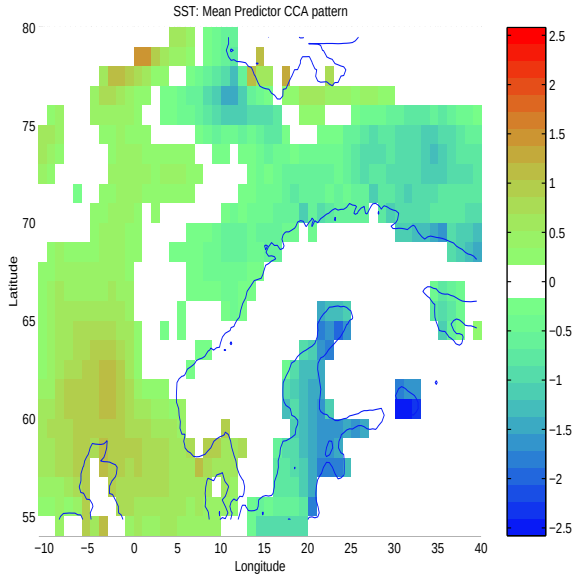

Figure 22: The mean leading October CCA GISST2.2 North Sea SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA North Sea SST pattern estimates are greater than 1°C are not shown.

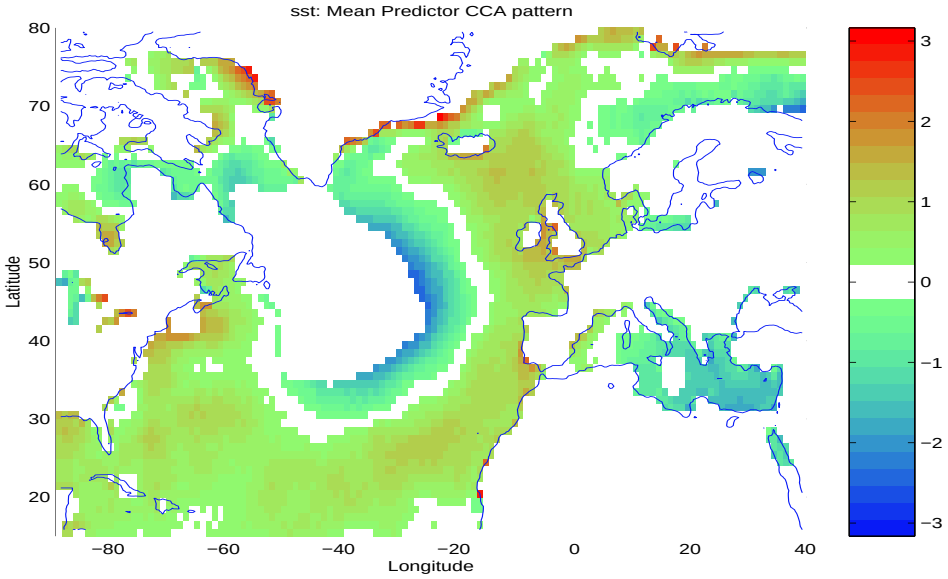

Figure 23: The mean leading October CCA North Atlantic SST pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA pattern estimates are greater than 1°C are not shown.

major warming events, but captured one of the cool periods during the 1960's. Evidently, the relationship between the land temperatures and the regional SSTs was weak during October (figure 21), and only marginal prediction skills were found in northern Norway. The leading October North Sea SST CCA pattern was characterised by strongest positive weights north of Scotland and negative weights in the Baltic Sea (figure 22).

4.4.2 Influence from remote SSTs

The October CCA predictor pattern in figure 23 was associated with strong weights along the Greenland coast, in the eastern Mediterranean, and in the North Atlantic interior. Because the predictor weights over a large region in the North Atlantic were associated with large standard deviations, a large area is masked out (white) east of New Foundland.

The October North Atlantic SST model gave in general poor predictions (figures 24 and 25), and the variance accounted for by the predictions was in general low. Figure 24 shows the predictions with highest correlation score of 0.37 for Tromsø, which is still above the 95% confidence level. It is evident that most of the large events were not captured by the model. The prediction skill for the whole North Atlantic was lower than when only the regional SSTs were used for model calibration. In general, the skills were marginally greater in the north than in the south. The worst prediction was for Lista fyr, where the cross-validation correlation was -0.19 (table 8). This result suggests that the North Atlantic SSTs contribute little to the Norwegian autumn temperatures.

4.5 Discussion of the SST models

In summary, the SST model is promising for the prediction of January mean land temperatures in Norway, and may be suitable for seasonal forecasting for the winter (the SSTs vary slowly). There is a deterioration of the North Atlantic SST model skill in the summer and autumn seasons, especially in the south for the July month. This observation brings up speculations as to whether the SST model performs better in colder climates and is more unreliable when it is warm. The reason why the model skill varies so much with seasons is unclear, but some explanations may include the effect of the seasonally varying mixed layer in the ocean or boundary layer in the atmosphere, or smaller ocean-atmosphere heat fluxes during summer due to smaller differences between SSTs and air temperatures. The model sensitivity to the seasons may be an indication that non-stationarity can be a problem for downscaling of global warming scenarios if warming over the

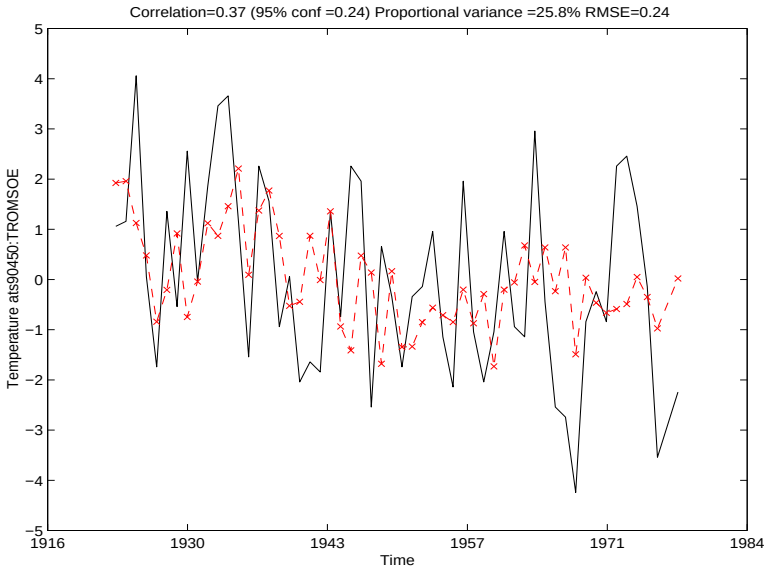

Figure 24: Time series of predicted October temperatures (dashed) at Karasjok, employing the cross-validation method with GISST2.2 North Atlantic SST, shown with the observations (black solid line).

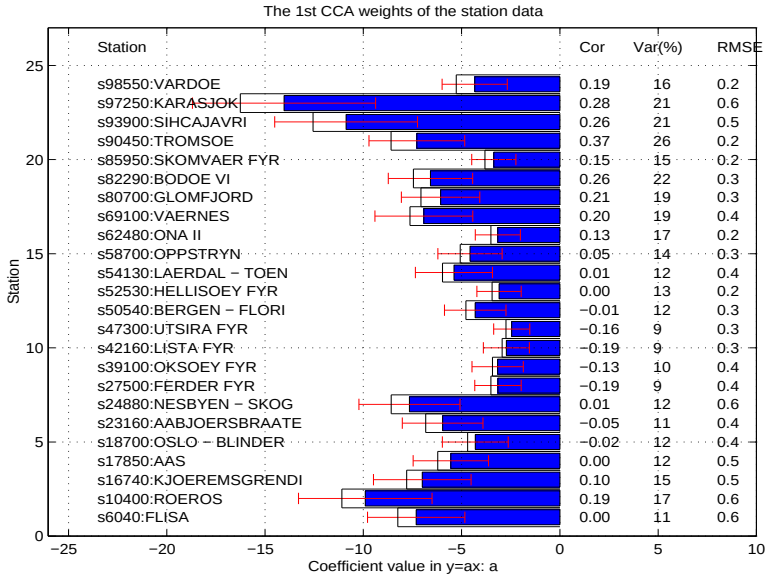

Figure 25: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October GISST2.2 North Atlantic SST CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Nordic countries make future winters look more like autumns or spring of present climate. The confidence of the SST model results for future climate predictions also depends on how well the SSTs are represented in the GCMs. A common problem for low-resolution ocean models is to give a sufficiently realistic description of the Gulf stream, as the models tend to reproduce realistical global circulation, and often simulate too weak north-south temperature gradients and place the location where the current separates from the east coast of the USA too far north or too far south (*Cane & Prasad*, 1995).

5 CCA Model Construction: SLP models

Large scale pressure systems, or geostrophic circulation, may have a strong influence on the surface temperatures in Norway. In order to investigate the relationship between SLP and temperatures, CCA models were constructed using SLP as the predictor fields. 3 different data sets, with different spatial resolution and spanning different time periods, were used as predictors: NCAR, NMC, and UEA.

5.1 Models based on the NCAR data

5.1.1 January North Atlantic SLPs and land surface Temperatures

The CCA results using the NCAR ds010.0 SLPs as predictor produced predictions with maximum cross-validation correlation of 0.84 and RMSE of 0.2°C (figure 26). The January NCAR SLP model captured most of the major and smaller anomalies, although it did not predict sufficiently low temperatures during the cold winter in the early 1940s. The leading CCA predictor pattern describing the SLP structure which had highest correlations with land temperatures is shown in figure 27. This pattern resembled the North Atlantic Oscillation (NAO), and as much as 80% of the January mean temperatures variability in some locations (Oksøy fyr) could be accounted for by enhanced westerlies due to a deepening of the Icelandic low and strengthening anti-cyclone over the Azores (figure 26). The high prediction skills and low model coefficient spread (figure 28) suggested that the NCAR SLPs were ideal predictors for most of the stations, except those in the far north. Table 9 suggests that the optimum choice of predictors was a combination of EOFs 1,4,6,7,8, and 12.

The data north of 70°N had been excluded from the analysis which produced figure 27, but similar results were obtained when the Arctic region was

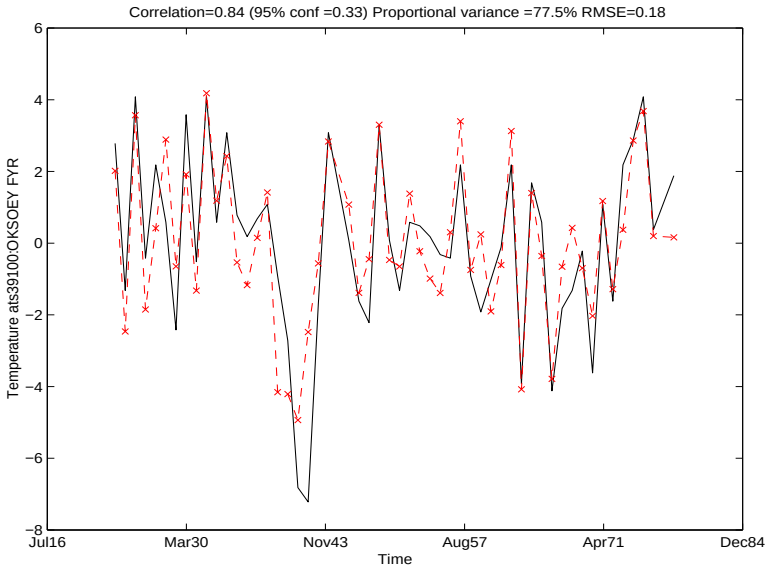

Figure 26: Time series of predicted January temperatures (dashed) at Oksøy Fyr, employing the cross-validation method with NCAR SLP, shown with the observations (black solid line).

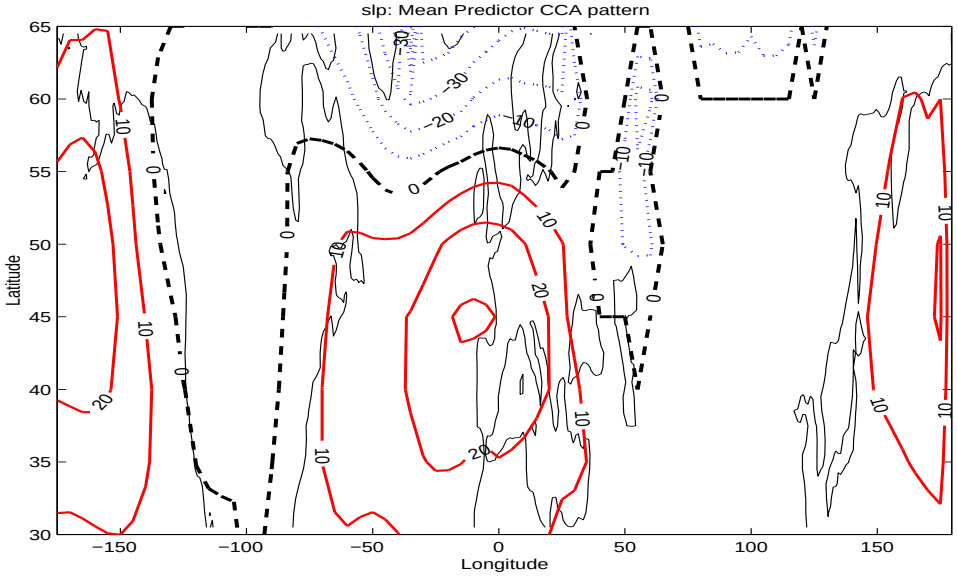

Figure 27: The mean leading January CCA NCAR SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

included (not shown). The NCAR data set contained a number of spatial points with 'bad' data (*Benestad*, 1998), especially near the Arctic but also over Russia, and it is questionable whether any Arctic climatic signals that may be correlated with the Norwegian temperatures is present in the NCAR data set.

Figure 28 indicates small spread in the model coefficient estimates. All weights had the same sign, and Nesbyen and Flisa were associated with the largest weights.

5.1.2 April North Atlantic SLPs and land surface Temperatures

The April NCAR SLP model was associated with lower correlation skill than the corresponding January model. The highest skills in April were found at Ona II as opposed to Oksøy fyr in January. Figure 29 indicates negative signs for the weights at Oksøy, Ferder fyr, and Vardø, while the remaining stations had positive weights. The predictor pattern described negative (positive) SLP anomalies west of the British isles, associated with warm (cold) temperature anomalies at all stations except for the south and east coast stations (figure 30). Figure 31 indicates that the April model had a tendency to produce some spurious cold temperatures at Oksøy fyr, but did capture most of the warm Aprils.

Table 10 indicates that the optimal model used EOFs 1, 2, 4, 6, 9, 10, 14, and 15 as predictors.

5.1.3 July North Atlantic SLPs and land surface Temperatures

The NCAR July model gave lower skill scores than the January model, but made skillful predictions nevertheless (table 11). The optimum combination of EOFs included EOFs 1, 2, 4, 6, 10, and 11. The predictor pattern (figure 32) indicated a high (low) pressure system over the North Sea, which was associated with higher (lower) temperatures in the south and cold anomalies in the north. The best prediction was found at Åbjørsbråten, with r=0.72, RMS error of 0.14°C, and variance of 62% (figure 33). The July model captured most of the warm episodes, but missed some of the cold events. The predictions of the peak values were were often too large or too small. The predictands with poorest skill was found at Ona, where r=0.37. Figure 34 shows the predictor weight estimates, and the weights were negative with relative large standard deviations in the north, suggesting that higher order CCA patterns may have played a role here. The best predictions were found along the west coast of Norway.

Table 9: Evaluation of January temperature CCA model based on January SLPs from NCAR ds010.0 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	OKSØY FYR r= 0.67 rmse= 0.25	SKOMVÆR FYR r= 0.48 rmse= 0.16	KARASJOK r= 0.33 rmse= 0.58
1 4	BODØ VI r= 0.73 rmse= 0.19	SKOMVÆR FYR r= 0.70 rmse= 0.13	OPPSTRYN r= 0.51 rmse= 0.25
1 4 6 7 8 12	OKSØY FYR r= 0.84 rmse= 0.18	SKOMVÆR FYR r= 0.74 rmse= 0.12	VARDØ r= 0.55 rmse= 0.18

Table 10: Evaluation of April temperature CCA model based on April SLPs from NCAR ds010.0 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.32 rmse= 0.16	UTSIRA FYR r= -0.23 rmse= 0.12	LÆRDAL - TØNJUM r= -0.43 rmse= 0.18
1 2	ONA II r= 0.55 rmse= 0.10	ONA II r= 0.55 rmse= 0.10	FERDER FYR r= 0.15 rmse= 0.17
1 2 4	ONA II r= 0.61 rmse= 0.10	ONA II r= 0.61 rmse= 0.10	FERDER FYR r= 0.11 rmse= 0.17
1 2 4 6 9	ONA II r= 0.62 rmse= 0.09	ONA II r= 0.62 rmse= 0.09	ÅS r= 0.01 rmse= 0.20
1 2 4 6 9 10 14 15	ONA II r= 0.66 rmse= 0.09	ONA II r= 0.66 rmse= 0.09	FERDER FYR r= 0.22 rmse= 0.17

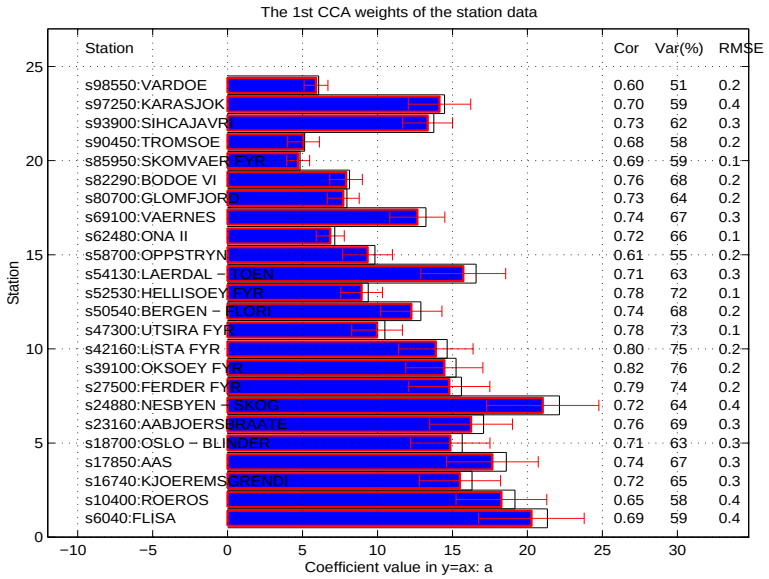

Figure 28: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading January NCAR SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

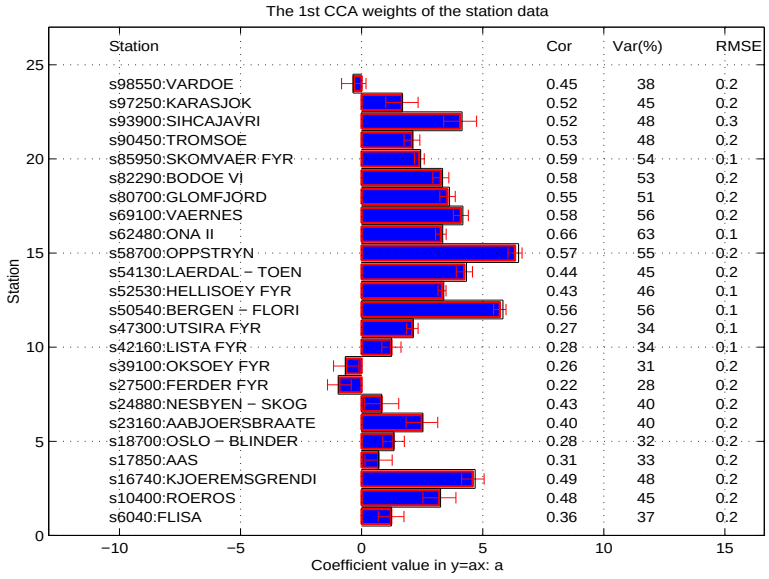

Figure 29: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading April NCAR SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

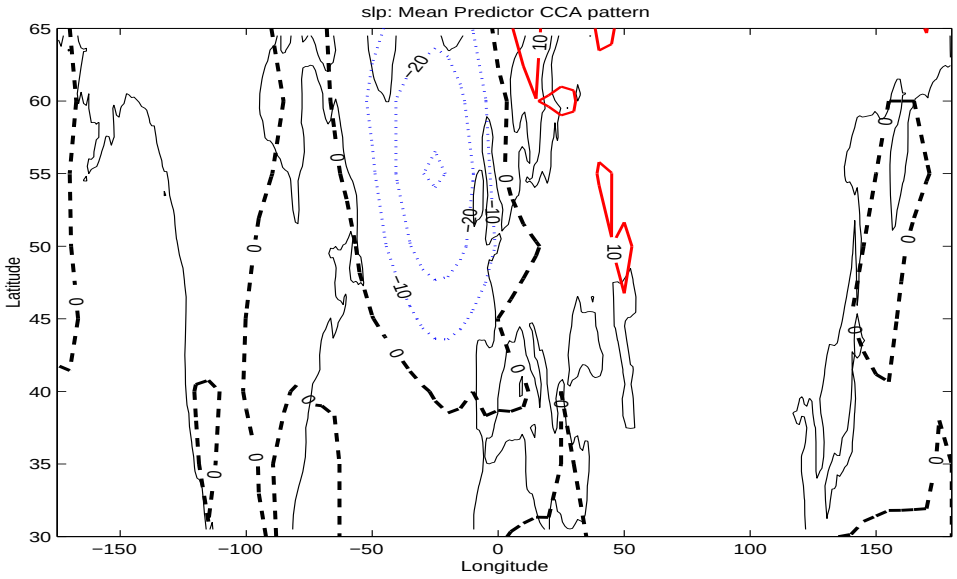

Figure 30: The mean leading April CCA NCAR SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

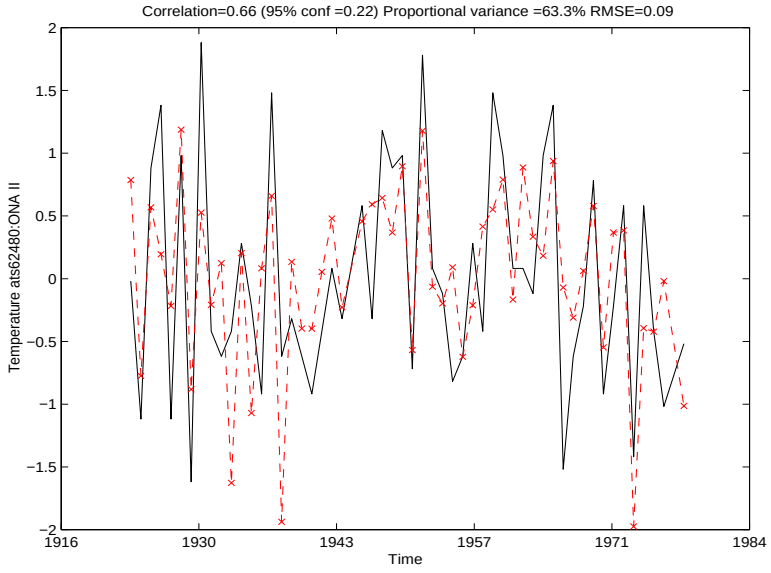

Figure 31: Time series of predicted April temperatures (dashed) at Ona II, employing the cross-validation method with NCAR SLP, shown with the observations (black solid line).

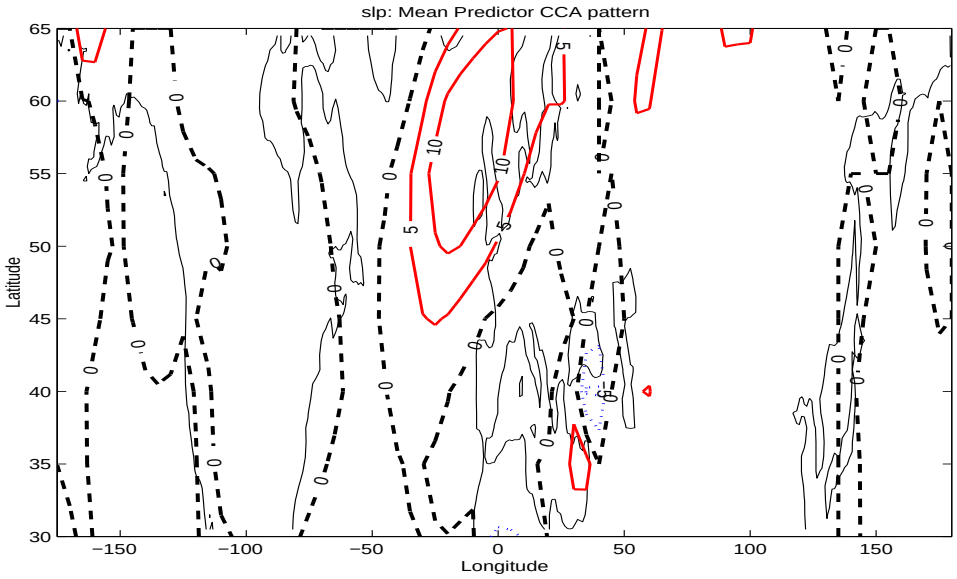

Figure 32: The mean leading July CCA NCAR SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

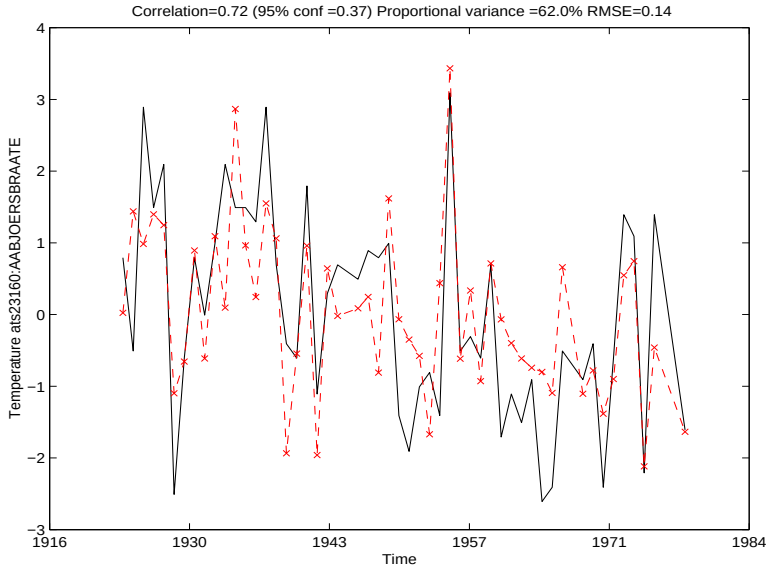

Figure 33: Time series of predicted July temperatures (dashed) at Åbjørsbråten, employing the cross-validation method with NCAR SLP, shown with the observations (black solid line).

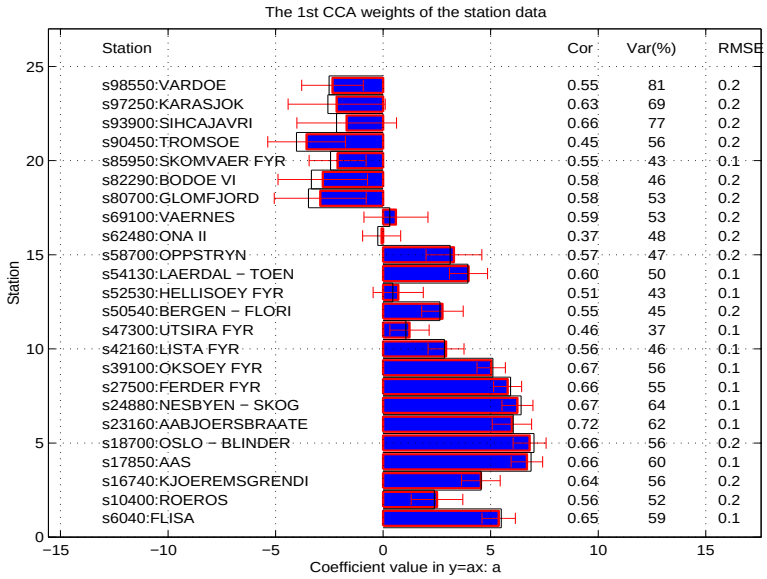

Figure 34: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July NCAR SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

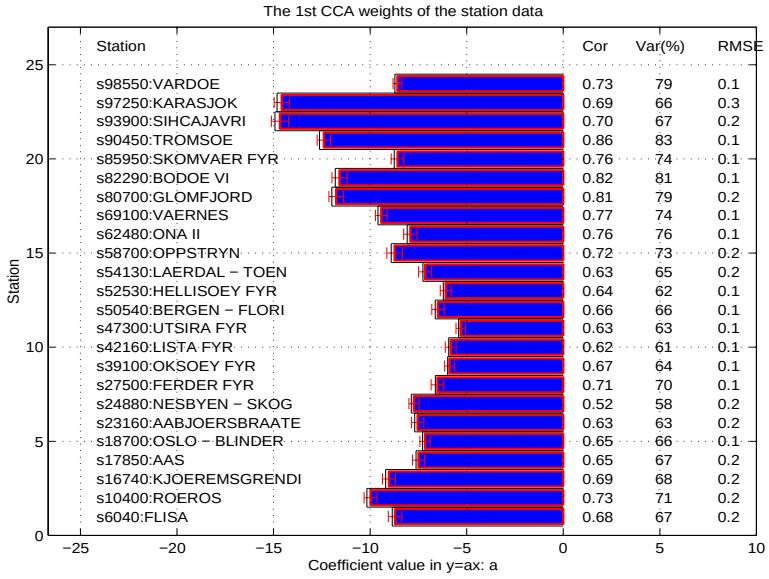

Figure 35: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October NCAR SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Table 11: Evaluation of July temperature CCA model based on July North Sea SLPs from NCAR ds010.0 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	SIHCAJAVRI r= 0.33 rmse= 0.28	ONA II r= 0.15 rmse= 0.16	FERDER FYR r= -0.06 rmse= 0.19
1 2	GLOMFJORD r= 0.47 rmse= 0.23	ONA II r= 0.24 rmse= 0.15	VARDØ r= -0.25 rmse= 0.24
1 2 4 6	KARASJOK r= 0.62 rmse= 0.24	SKOMVÆR FYR r= 0.52 rmse= 0.15	UTSIRA FYR r= 0.30 rmse= 0.15
1 2 4 6 10 11	ÅBJØRSBRÅTEN r= 0.72 rmse= 0.14	OKSØY FYR r= 0.67 rmse= 0.12	ONA II r= 0.37 rmse= 0.15

5.1.4 October North Atlantic SLPs and land surface Temperatures

The optimal October model includes EOFs 1, 2, 3, 4, 5, 7, 9, 11, 12, 17, and 20 in the predictors (table 12). The predictor weights in figure 35 all had the same sign and were associated with small standard deviations. The predictor pattern described a strong pressure system over Iceland (figure 36), and low pressure was associated with warmer weather over Norway. The best prediction was found at Tromsø (r=0.86, figure 37), where the October temperatures were highly correlated with the southerly geostrophic wind. The October NCAR model captured virtually all the peaks in figure 37. The lowest skill was seen at Nesbyen, where the correlation skill was 0.52 and was significant at the 95% level.

The NCAR models could skillfully predict the land surface temperatures, but with slightly lower skill during April and July. It is unlikely that the weak sensitivity to the seasons will be important for studies of future warming scenarios.

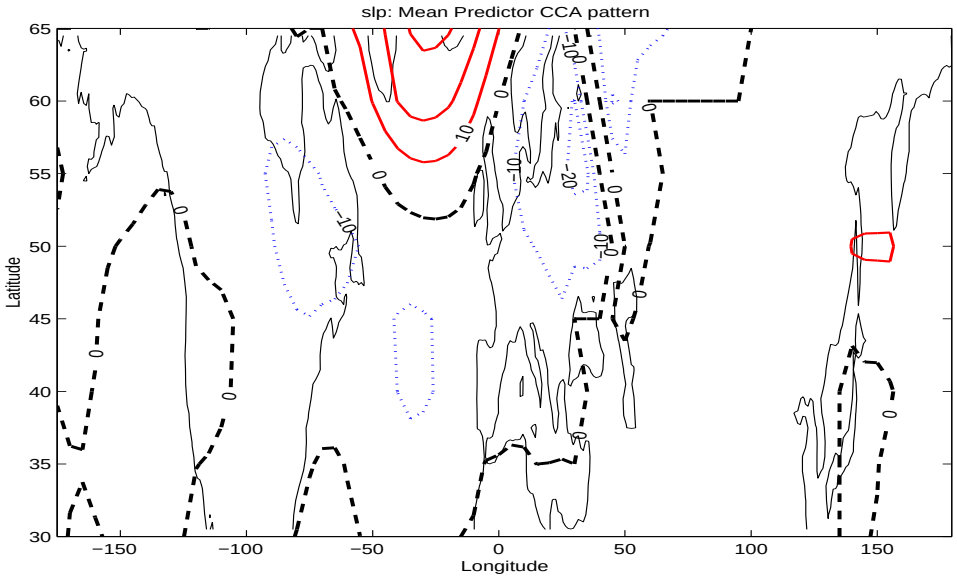

Figure 36: The mean leading October CCA NCAR SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

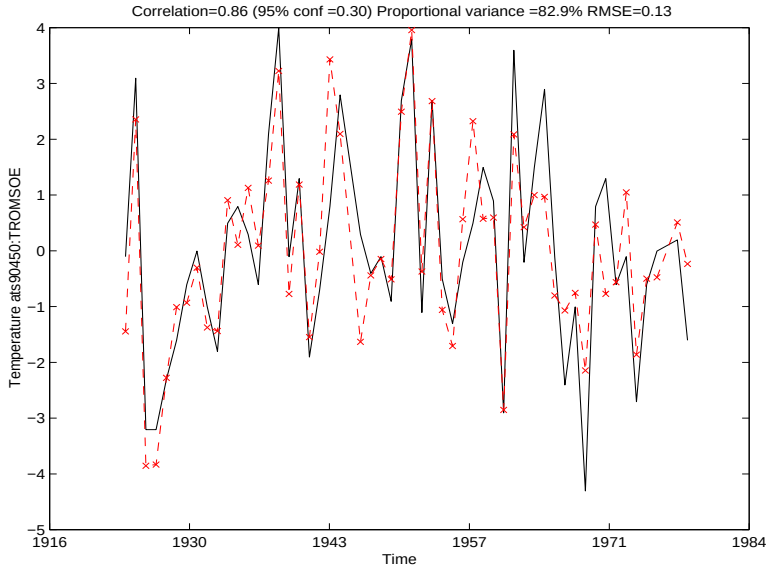

Figure 37: Time series of predicted October temperatures (dashed) at Tromsø, employing the cross-validation method with NCAR SLP, shown with the observations (black solid line).

Table 12: Evaluation of October temperature CCA model based on October North Sea SLPs from NCAR ds010.0 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	OPPSTRYN r= 0.35 rmse= 0.22	UTSIRA FYR r= 0.27 rmse= 0.14	FERDER FYR r= -0.07 rmse= 0.18
1 2 3 4 5 7 9 11 12 17 20	TROMSØ r= 0.86 rmse= 0.13	UTSIRA FYR r= 0.63 rmse= 0.11	NESBYEN - SKOGLUND r= 0.52 rmse= 0.19

5.2 Models based on the NMC data

5.2.1 January North Atlantic SLPs and land surface Temperatures

The NMC ds195.5 SLP data set spanned a relatively short time period (1946-1994), but had a high spatial resolution ($2.5° \times 2.5°$) compared to the other SLP data sets. Figure 38 shows the leading CCA predictor pattern for the NMC ds195.5 SLPs, which described a north-south dipole structure resembling the NAO. The optimum number of EOFs in the CCA was 8, which included EOFs 1, 5, 6, 11, 13, 15, 16 and 20 (table 13). The NMC CCA pattern differed from the corresponding NCAR pattern by describing a weaker dipole structure, with the southern maximum of the north-south dipole pattern covering a larger area. There was weaker variability over the Bay of Biscaya in the NMC pattern, and stronger SLP anomalies over Finland and north eastern Russia. The station with the highest correlation skills associated with the NMC model was Lista fyr (figure 39), with a correlation coefficient of 0.91, RMSE of 0.14°C, and accounting for 92% of the variance. The NMC model was poor at predicting the January temperatures in northern Norway, despite using data from the Arctic (figure 40). The January NMC model predicted all the major events, although the peak values were sometimes slightly too large or too small.

The different time periods, different errors in the predictor data (*Benestad*, 1998), and different spatial resolution may decide which locations have the greatest prediction skill and what predictor pattern is associated with these temperatures. Figure 40 indicates that all the weights had same sign and that the standard deviations were small. Hence the leading CCA pattern, associated with enhanced westerly geostrophic flow, was important for

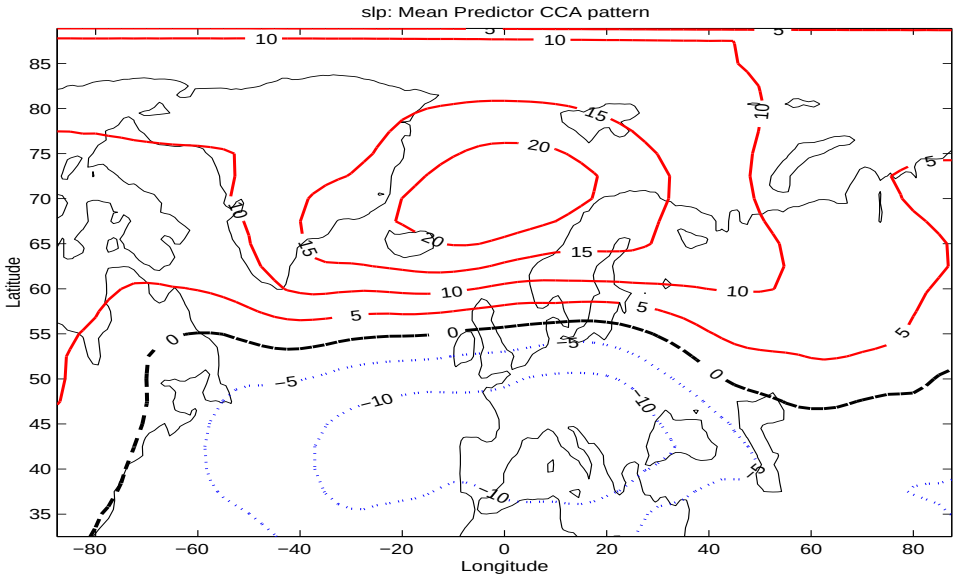

Figure 38: The mean leading January CCA NMC SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

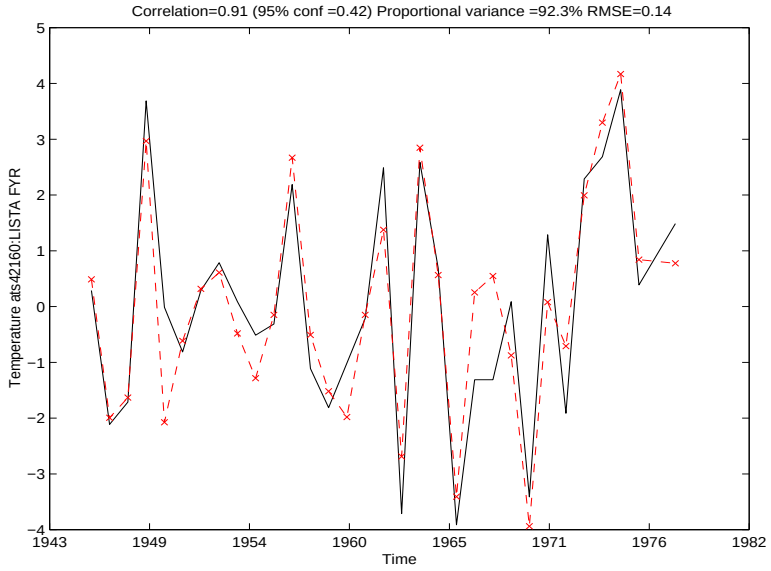

Figure 39: Time series of predicted January temperatures (dashed) at Lista Fyr, employing the cross-validation method with NMC SLP, shown with the observations (black solid line).

all the stations.

5.2.2 April North Atlantic SLPs and land surface Temperatures

The April NMC model produced predictions with lower skills than the January model for the locations with highest skills. The predictand weight estimates in figure 41 were small compared to the January weights, suggesting weaker influence of April SLPs on the temperatures. Flisa, with the highest correlation score, had almost zero mean weights but relatively large spread, which suggests that higher order patterns were likely to be important. The CCA correlations for the whole time series were high for the 8 CCA patterns: 1.0000, 0.9918, 0.9845, 0.9652, 0.9159, 0.9008, 0.833 and 0.6608.

The predictor pattern for April was substantially different to the January leading CCA pattern, where the April temperatures were correlated with a west-east dipole pressure system in addition to a north-south dipole over western Atlantic (figure 42), and warm temperature anomalies in Scandinavia were associated with southerly geostrophic flow. The highest prediction skill for April was found at Flisa, with r=0.67, RMS error of 0.18°C, and accounting for 79% of the variance. Figure 43 demonstrates that the NMC April model could reproduce the timing of most anomalies, but misjudged the amplitudes.

5.2.3 July North Atlantic SLPs and land surface Temperatures

Figure 44 shows the leading NMC CCA predictor pattern for July, with a tripole structure with maximum SLP variability north of the British isles and maxima with opposite polarity over subtropical Atlantic and east of Svalbard. This system was associated with southerly and northerly geostrophic winds over southern Norway.

The highest skills in July were seen along the west coast in southern Norway. The best prediction, which was for Hellisøy fyr, had a correlation score of 0.79, RMS error of 0.11°C, and accounted for 75% of the variance (figure 45).

The NMC July model gave relatively high correlation skills over all of Norway, and there were small uncertainties in the estimates of predictand coefficients for the leading CCA pattern (figure 46). The predictand coefficients were large and positive in the north and small and negative in southern Norway.

Table 13: Evaluation of January temperature CCA model based on January SLPs from NMC ds195.5 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	OKSØY FYR r= 0.75 rmse= 0.23	UTSIRA FYR r= 0.71 rmse= 0.18	KARASJOK r= 0.12 rmse= 0.77
1 2	OKSØY FYR r= 0.72 rmse= 0.25	UTSIRA FYR r= 0.68 rmse= 0.19	KARASJOK r= 0.03 rmse= 0.79
1 5 6 11 13 15 16 20	LISTA FYR r= 0.91 rmse= 0.14	UTSIRA FYR r= 0.86 rmse= 0.13	VARDØ r= 0.00 rmse= 0.29

Table 14: Evaluation of April temperature CCA model based on April SLPs from NMC ds195.5 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.40 rmse= 0.19	SKOMVÆR FYR r= 0.35 rmse= 0.14	OPPSTRYN r= -0.05 rmse= 0.23
1 2 3 5 6	ÅBJØRSBRÅTE r= 0.49 rmse= 0.24	UTSIRA FYR r= 0.34 rmse= 0.15	BODØ VI r= -0.07 rmse= 0.24
1 2 3 5 6 7	ÅBJØRSBRÅTE r= 0.62 rmse= 0.21	HELLISØY FYR r= 0.49 rmse= 0.14	GLOMFJORD r= 0.30 rmse= 0.22
1 2 3 5 6 7 15 17	FLISA r= 0.67 rmse= 0.18	ONA II r= 0.52 rmse= 0.14	GLOMFJORD r= 0.28 rmse= 0.23

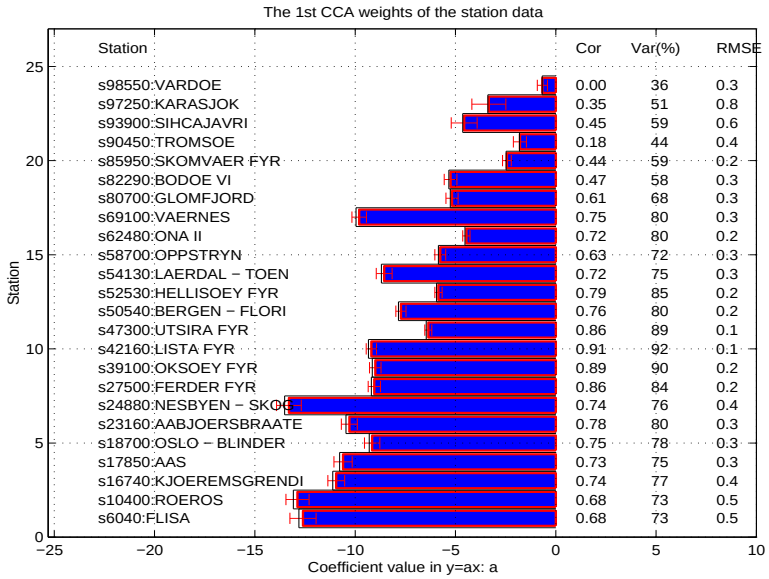

Figure 40: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading January NMC SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

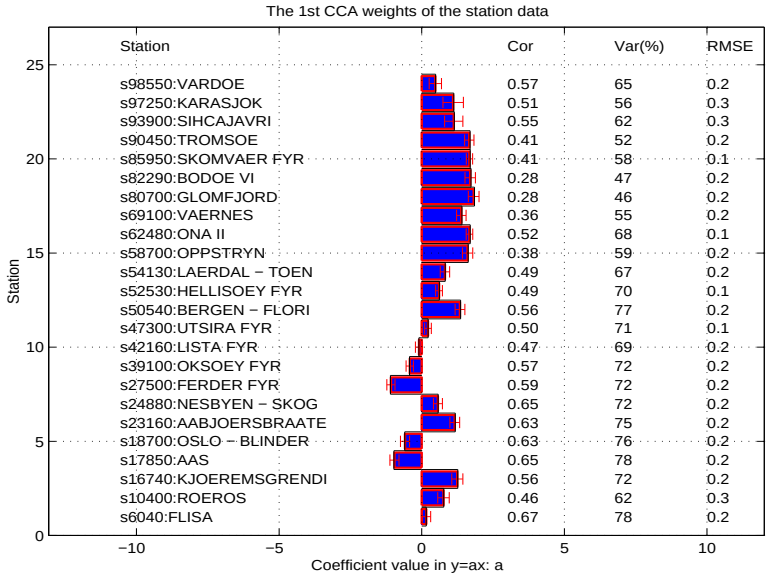

Figure 41: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading April NMC SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

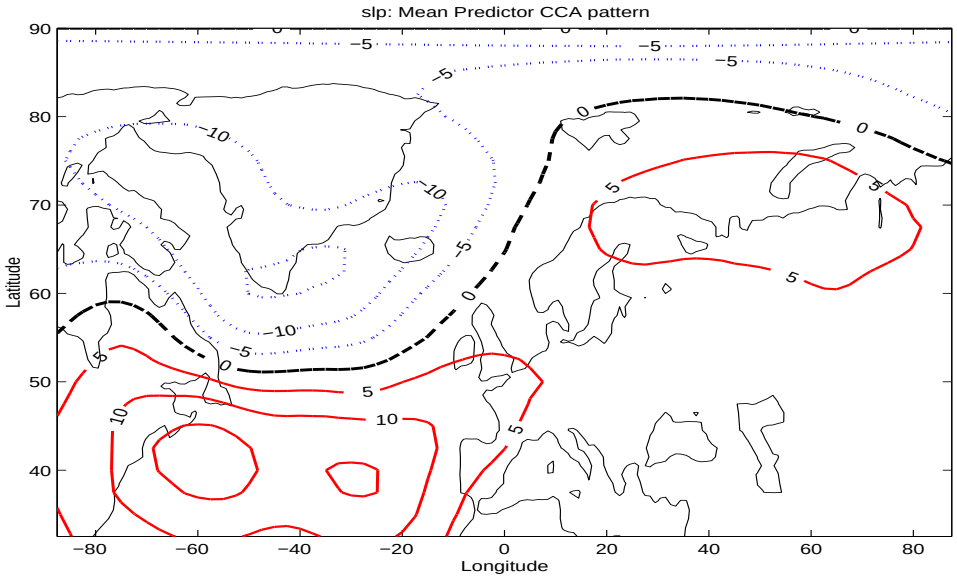

Figure 42: The leading April CCA NMC SLP pattern associated with the land surface temperatures.

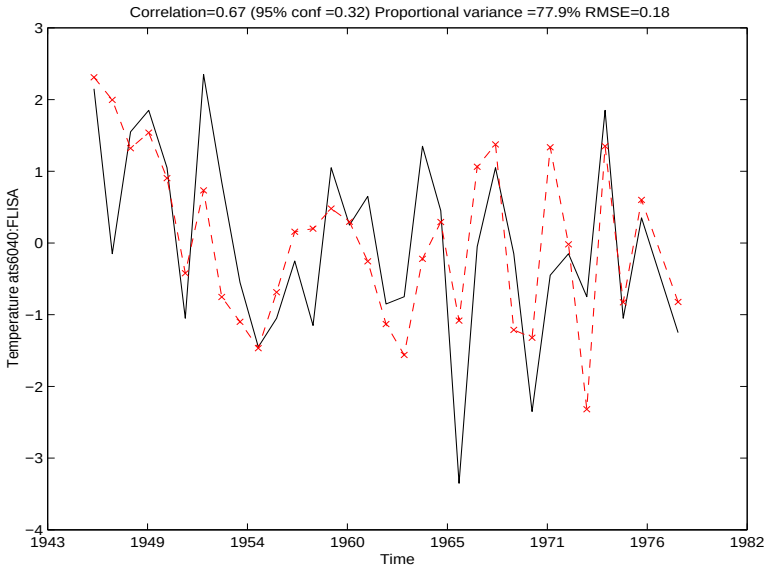

Figure 43: Time series of predicted April temperatures (dashed) at Flisa, employing the cross-validation method with NMC SLP, shown with the observations (black solid line).

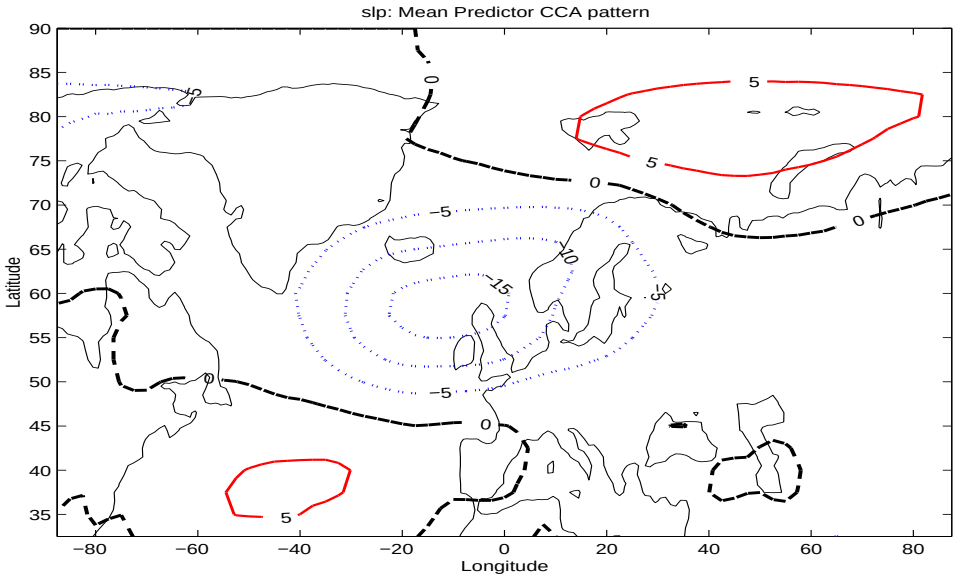

Figure 44: The leading July CCA NMC SLP pattern associated with the land surface temperatures.

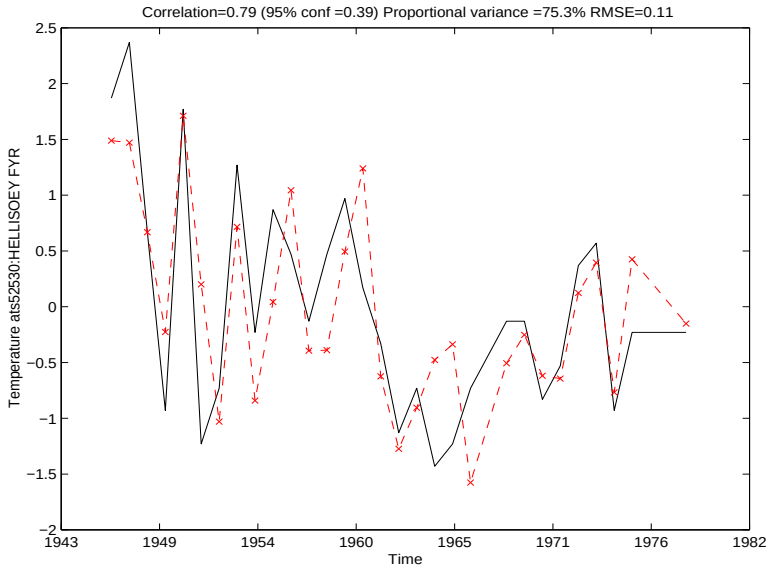

Figure 45: Time series of predicted July temperatures (dashed) at Hellisøy fyr, employing the cross-validation method with NMC SLP, shown with the observations (black solid line).

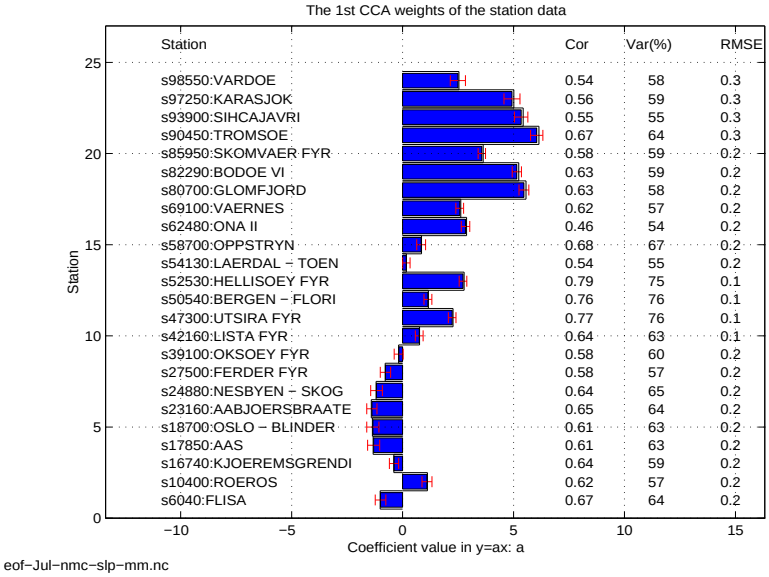

Figure 46: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July NMC SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

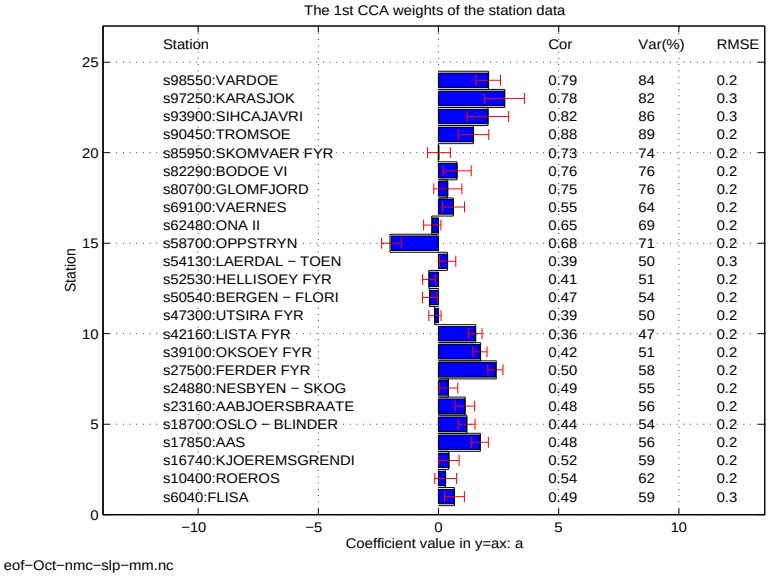

Figure 47: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October NMC SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Table 15: Evaluation of July temperature CCA model based on July SLPs from NMC ds195.5 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	GLOMFJORD r= 0.36 rmse= 0.29	HELLISØY FYR r= 0.13 rmse= 0.18	VARDØ r= -0.44 rmse= 0.32
1 2	GLOMFJORD r= 0.41 rmse= 0.28	HELLISØY FYR r= 0.31 rmse= 0.17	LÆRDAL - TØN r= -0.59 rmse= 0.20
1 2 3	UTSIRA FYR r= 0.74 rmse= 0.12	UTSIRA FYR r= 0.74 rmse= 0.12	VÆRNES r= 0.26 rmse= 0.20
1 2 3 6 8 15	HELLISØY FYR r= 0.79 rmse= 0.11	HELLISØY FYR r= 0.79 rmse= 0.11	ONA II r= 0.46 rmse= 0.17

Table 16: Evaluation of October temperature CCA model based on October SLPs from NMC ds195.5 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	FERDER FYR r= 0.42 rmse= 0.20	UTSIRA FYR r= -0.13 rmse= 0.16	SKOMVÆR FYR r= -0.64 rmse= 0.26
1 2 3	OPPSTRYN r= 0.54 rmse= 0.22	UTSIRA FYR r= 0.45 rmse= 0.14	KARASJOK r= 0.16 rmse= 0.47
1 2 4	TROMSØ r= 0.85 rmse= 0.17	UTSIRA FYR r= 0.54 rmse= 0.14	LÆRDAL - TØN r= 0.45 rmse= 0.23
1 2 4 8 15 19 20	TROMSØ r= 0.88 rmse= 0.15	TROMSØ r= 0.88 rmse= 0.15	LISTA FYR r= 0.36 rmse= 0.20

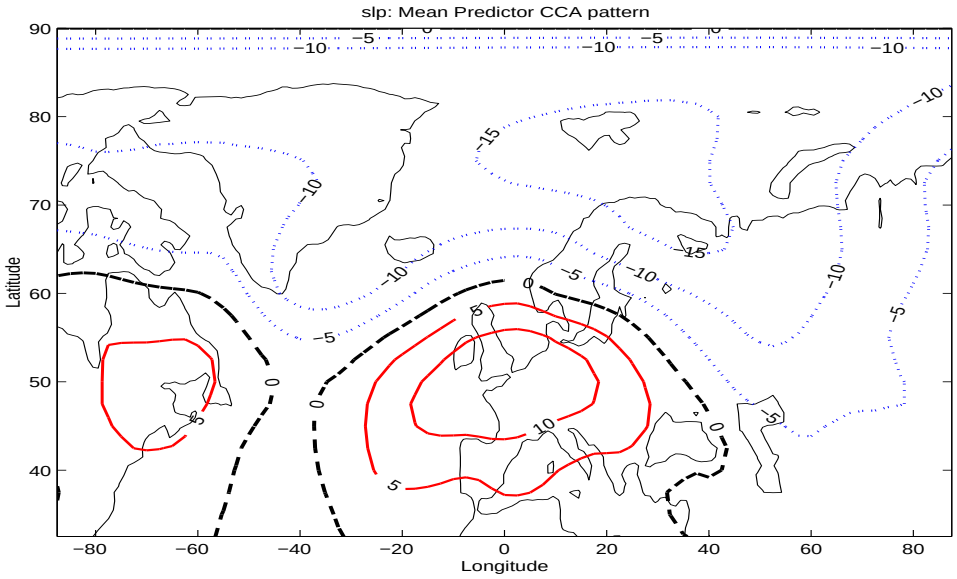

Figure 48: The leading October CCA NMC SLP pattern associated with the land surface temperatures.

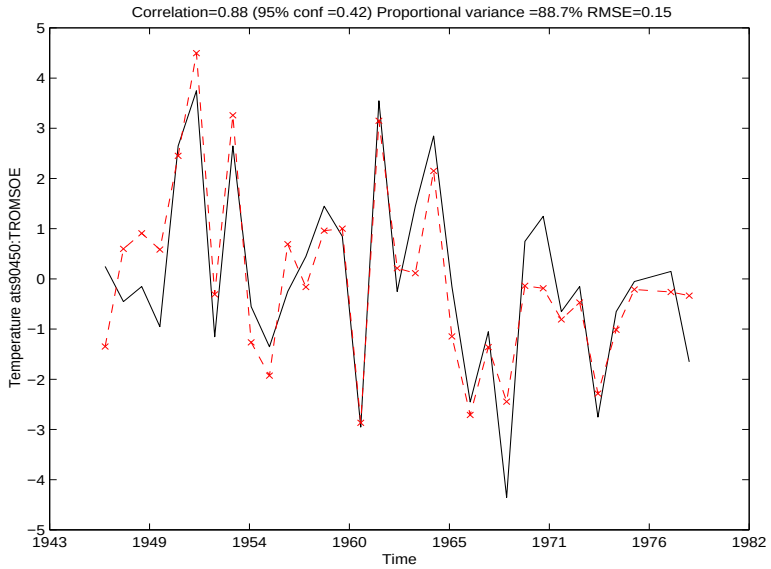

Figure 49: Time series of predicted October temperatures (dashed) at Tromsø, employing the cross-validation method with NMC SLP, shown with the observations (black solid line).

5.2.4 October North Atlantic SLPs and land surface Temperatures

The October NMC model gave high skill scores in the north, with the best prediction in Tromsø where r=0.88, and worst prediction at Lista fyr with r=0.36 (table 16). Figure 47 shows that the weights in the north were relative large, and that the standard deviation at all stations were small. The leading CCA pattern described a north-south dipole structure with maxima over the British isles and Svalbard (figure 48), and zonal component of the geostrophic flow was highly correlated with temperature anomalies in northern Norway. The prediction for Tromsø, shown in figure 49, demonstrates that the October model captured most of the temperature signal.

The model results for the NMC SLP models can be summarised as relatively good prediction skills for most seasons, but with January being the month with the highest and April the month with the lowest prediction scores.

5.3 Models based on the UEA data

5.3.1 January North Atlantic SLPs and land surface Temperatures

The leading UEA CCA SLP pattern (figure 50) was similar to the leading NMC CCA SLP pattern, but the southern dipole maximum was displaced further east. The optimum predictor combination consisted of the EOFs 1,2,3,4,7,8,10,11,12 and 18 (table 17), and the highest skills were found at Glomfjord (figure 51), with a correlation coefficient of 0.89 and RMSE of 0.12°C. The UEA CCA SLP pattern implied a south-westerly geostrophic flow, and most of the temperature variability at Glomfjord could be predicted by the UEA January model.

Contrary to the NMC model, however, the UEA model also produced relatively skillful predictions for northern Norway (figure 52). The UEA data was cropped north of north of 70%, where the data quality was highly questionable (*Benestad*, 1998). The fact that the prediction skills were so high without data from the Arctic may suggest that small scale Arctic SLP anomalies did not have much impact on the Norwegian January temperatures.

The predictand weights in figure 52 were associated with small standard deviations, which may suggest that the leading CCA pattern was well resolved from the higher order CCA structures. The implication of a dominant leading CCA pattern was that most of the Norwegian January temperatures were highly correlated with the meridional wind component.

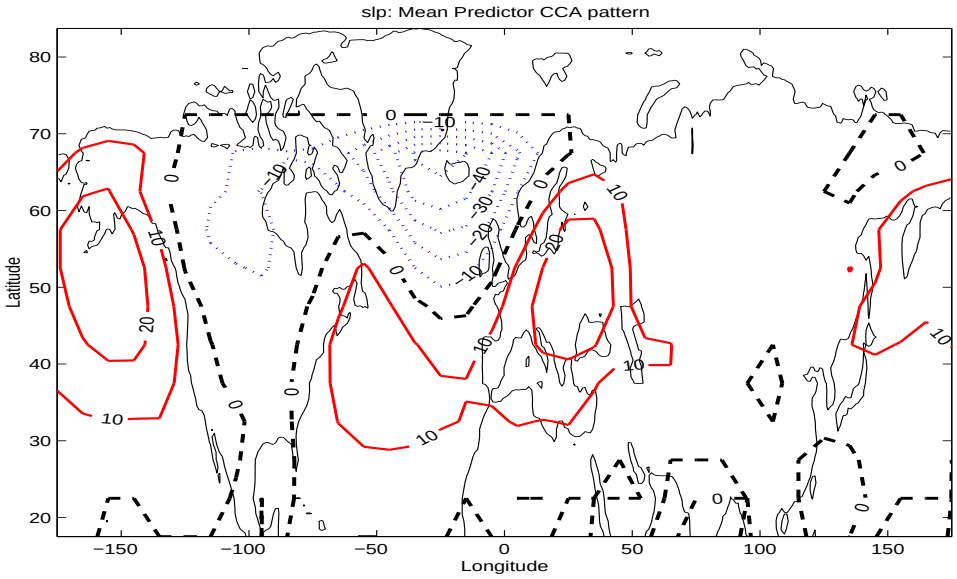

Figure 50: The mean leading January CCA UEA SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

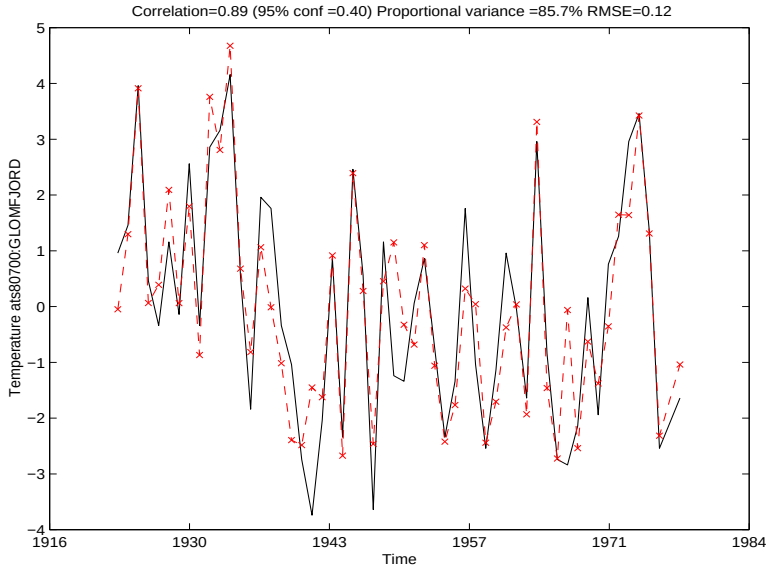

Figure 51: Time series of predicted January temperatures (dashed) at Glomfjord, employing the cross-validation method with UEA SLP, shown with the observations (black solid line).

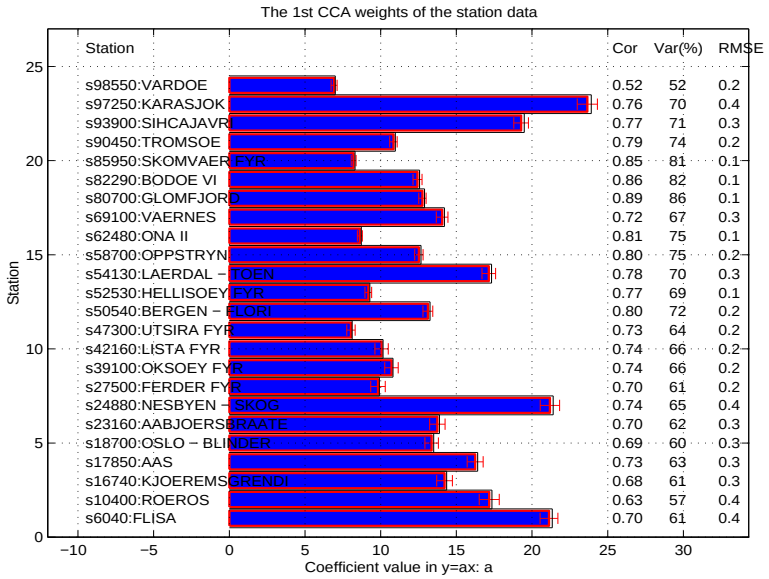

Figure 52: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading January UEA SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

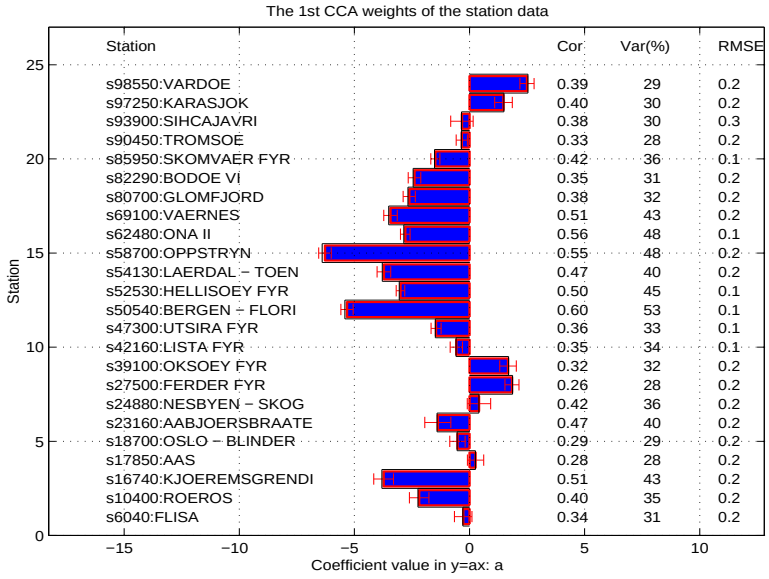

Figure 53: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading April UEA SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Table 17: Evaluation of January temperature CCA model based on January SLPs from UEA and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	OKSØY FYR r= 0.68 rmse= 0.24	SKOMVÆR FYR r= 0.50 rmse= 0.16	KARASJOK r= 0.38 rmse= 0.56
1 2 3 4 7	GLOMFJORD r= 0.80 rmse= 0.16	SKOMVÆR FYR r= 0.80 rmse= 0.11	RØROS r= 0.49 rmse= 0.49
1 2 3 4 7 8 10 11 12 18	GLOMFJORD r= 0.89 rmse= 0.12	SKOMVÆR FYR r= 0.85 rmse= 0.10	VARDØ r= 0.52 rmse= 0.18

Table 18: Evaluation of April temperature CCA model based on April SLPs from UEA and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.32 rmse= 0.16	UTSIRA FYR r= -0.11 rmse= 0.12	HELLISØ Y FYR r= -0.44 rmse= 0.12
1 2	ONA II r= 0.49 rmse= 0.10	ONA II r= 0.49 rmse= 0.10	BODØ VI r= 0.14 rmse= 0.19
1 2 3	BERGEN - FLORIDA r= 0.50 rmse= 0.14	ONA II r= 0.46 rmse= 0.11	ÅS r= 0.11 rmse= 0.19
1 2 3 4 8 9 10	BERGEN - FLORIDA r= 0.60 rmse= 0.13	ONA II r= 0.56 rmse= 0.10	FERDER FYR r= 0.26 rmse= 0.17

Table 19: Evaluation of July temperature CCA model based on July SLPs from UEA and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	GLOMFJORD r= 0.47 rmse= 0.23	ONA II r= 0.27 rmse= 0.15	BJRSBRTEN r= -0.54 rmse= 0.20
1 3 6 7 9 10 12 14 20	FLISA r= 0.79 rmse= 0.11	FLISA r= 0.79 rmse= 0.11	VARDØ r= 0.18 rmse= 0.21

Table 20: Evaluation of October temperature CCA model based on October SLPs from UEA and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.18 rmse= 0.20	UTSIRA FYR r= 0.01 rmse= 0.14	SKOMVÆR FYR r= -0.53 rmse= 0.20
1 2 3 4 5 6 8 10 12 15 17	TROMSØ r= 0.83 rmse= 0.14	UTSIRA FYR r= 0.63 rmse= 0.11	NESBYEN - SKOGLUND r= 0.47 rmse= 0.19

5.3.2 April North Atlantic SLPs and land surface Temperatures

The UEA April model had lower correlation skills (table 18) than the corresponding January model, and the predicted variance was generally less than for January predictions (figure 53). The weights had no longer same sign: Oksøy, Ferder, Vardø and Karasjok (Ås and Nesbyen had same sign as these stations, but the standard deviations were large.) had opposite sign to the other stations. The leading CCA weights were similar to those of the NCAR April model, but different to the NMC results. The fact that the standard deviations were small for all stations, suggested that the predictor pattern in figure 54 was important for all stations. North-westerly winds were associated with warmer weather along the west coast, but lower temperatures in the north and the southern coast. The best prediction, shown in figure 55 was found in Bergen, and the UEA April model captured most of the variability, although the peak values often were underestimated or exaggerated.

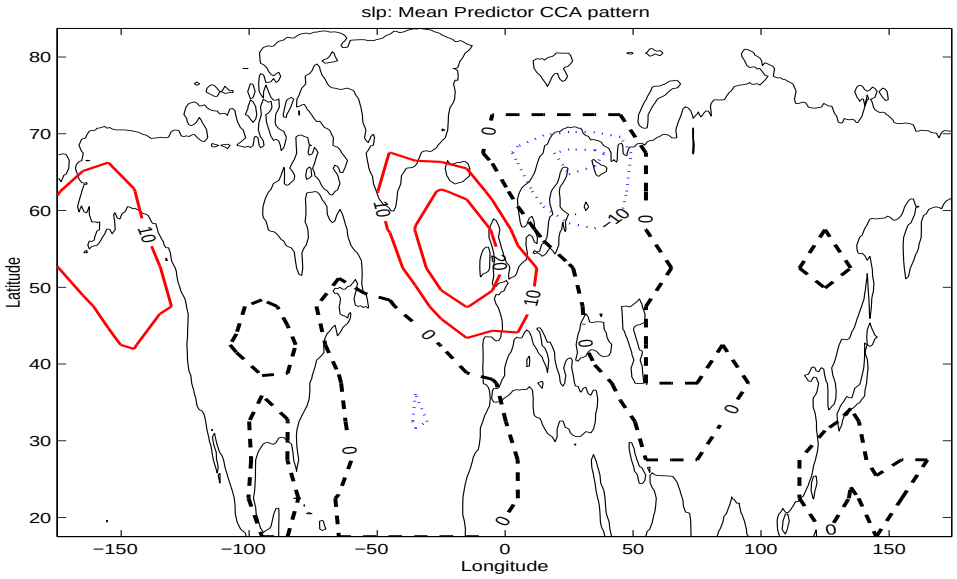

Figure 54: The mean leading April CCA UEA SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

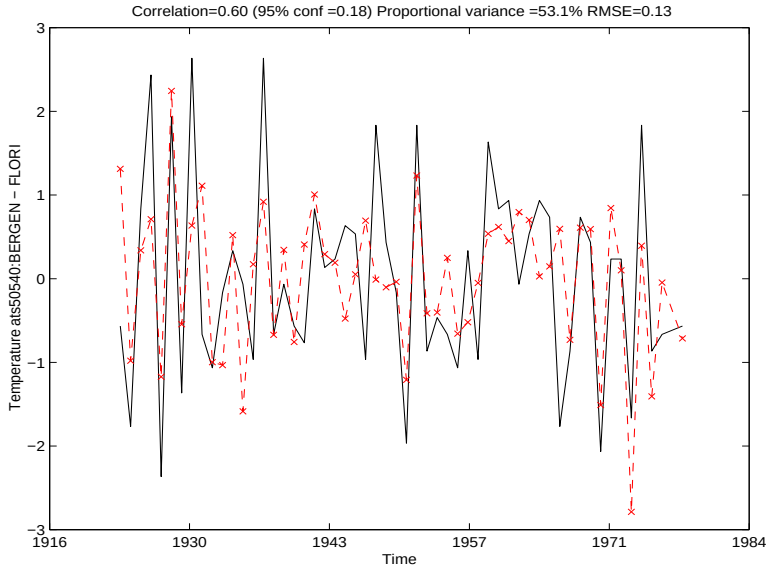

Figure 55: Time series of predicted April temperatures (dashed) at Bergen, employing the cross-validation method with UEA SLP, shown with the observations (black solid line).

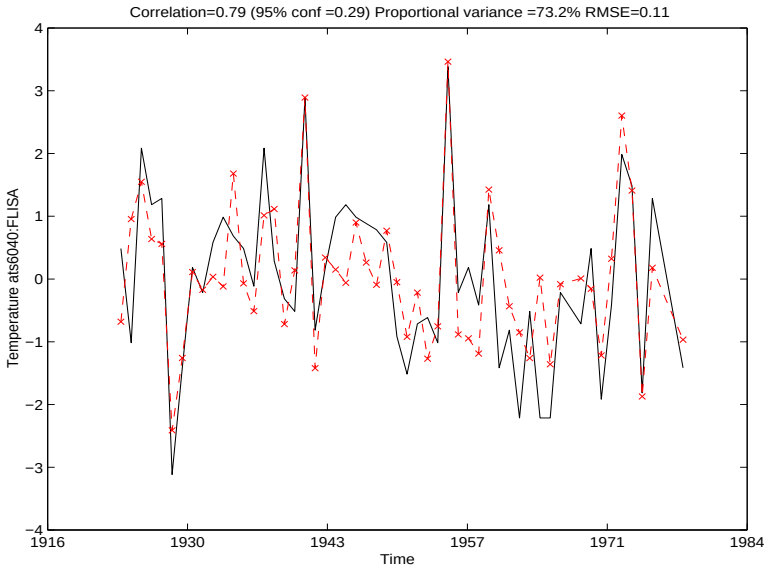

Figure 56: Time series of predicted July temperatures (dashed) at Flisa, employing the cross-validation method with UEA SLP, shown with the observations (black solid line).

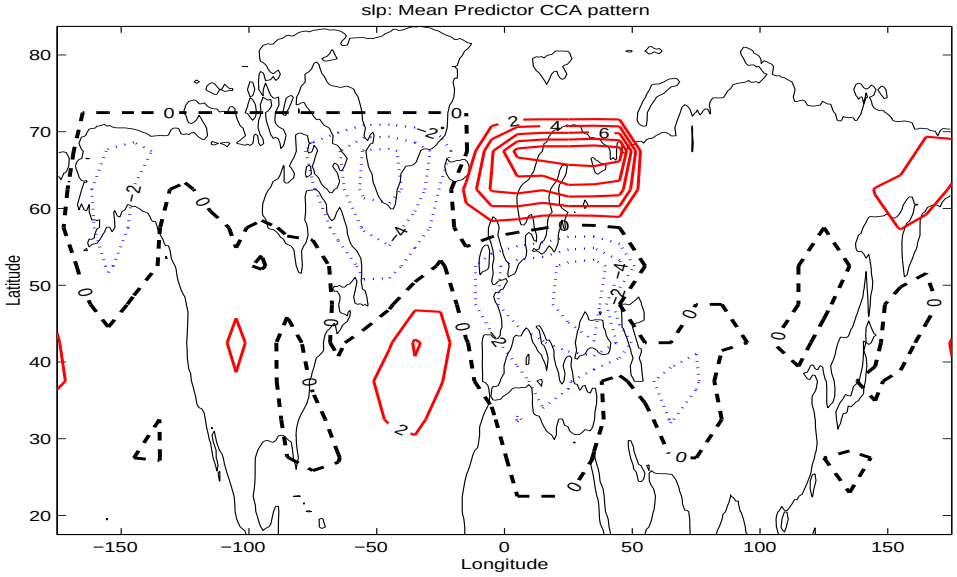

Figure 57: The mean leading July CCA UEA SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

5.3.3 July North Atlantic SLPs and land surface Temperatures

The UEA July model had high prediction skills, although a bit lower than for the January model. The highest correlation score was seen at Flisa, with r=0.79 (figure 56), and the lowest scores at Vardø, with cross-validation correlation of 0.18 (table 19). Figure 56 demonstrates that the July UEA model captured essentially all the warm episodes, but underestimated or overestimated some cold events. The leading CCA pattern, shown in figure 57 described a SLP tripole pattern, with strong weights over northern Scandinavia and two opposite poles over southern Greenland and Ukraine. The leading model coefficients were all positive with small standard deviations, and high temperatures were associated with high pressure over Scandinavia (figure 58). The highest skill scores were found at the inland stations in southern Norway.

5.3.4 October North Atlantic SLPs and land surface Temperatures

All leading CCA weights of the UEA October model in figure 59 had the same sign and small spread, indicating the importance of the leading CCA predictor pattern. The leading UEA CCA pattern consists of a west-east dipole pattern (figure 60). The greatest prediction skill was seen at Tromsø, where all the events were captured by the model despite the peak amplitudes being out by as much as 1°C. The implication of the leading CCA pattern for October being the dominating feature was that the temperatures were strongly correlated with southerly geostrophic winds. The worst predictions were for Nesbyen, with r=0.47.

In summary, the NMC models indicated small variations of skill with the seasons, with January giving best prediction scores. This weak seasonal sensitivity may suggest that the SLP models are unlikely to have problems with non-stationarity when this model is used for downscaling of global warming scenarios.

6 Geopotential height models

6.1 January NMC 500hPa Geopotential heights and land surface Temperatures

The NMC 500hPa geopotential heights ($\Phi(500hPa)$) were used as predictors in the CCA model to see if upper air parameters could be related to local surface climate variability. Figure 62 shows the best prediction for the

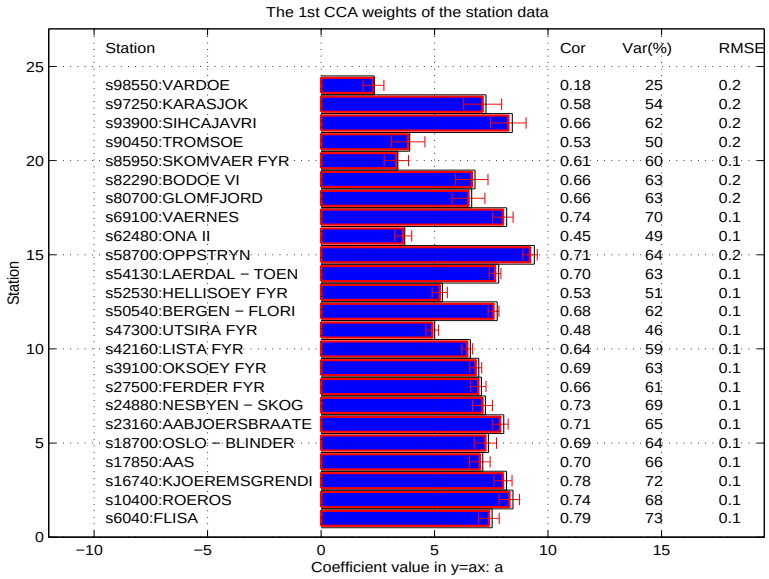

Figure 58: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July UEA SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

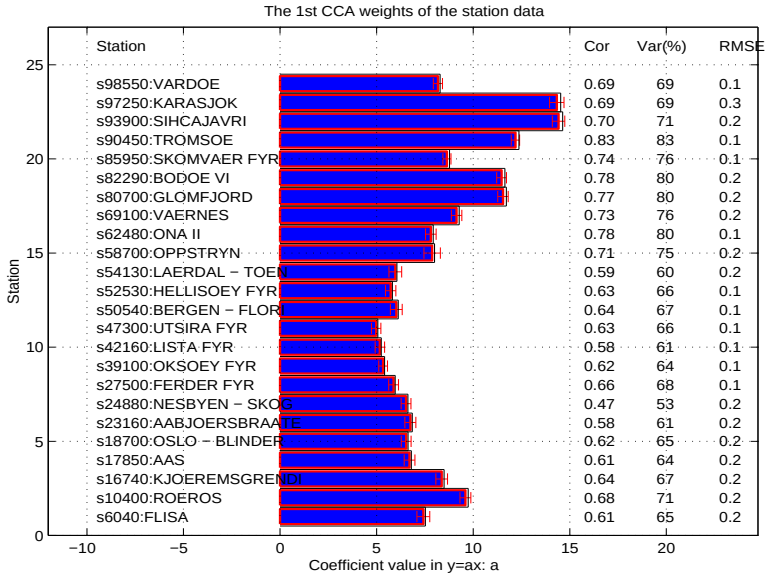

Figure 59: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October UEA SLP CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

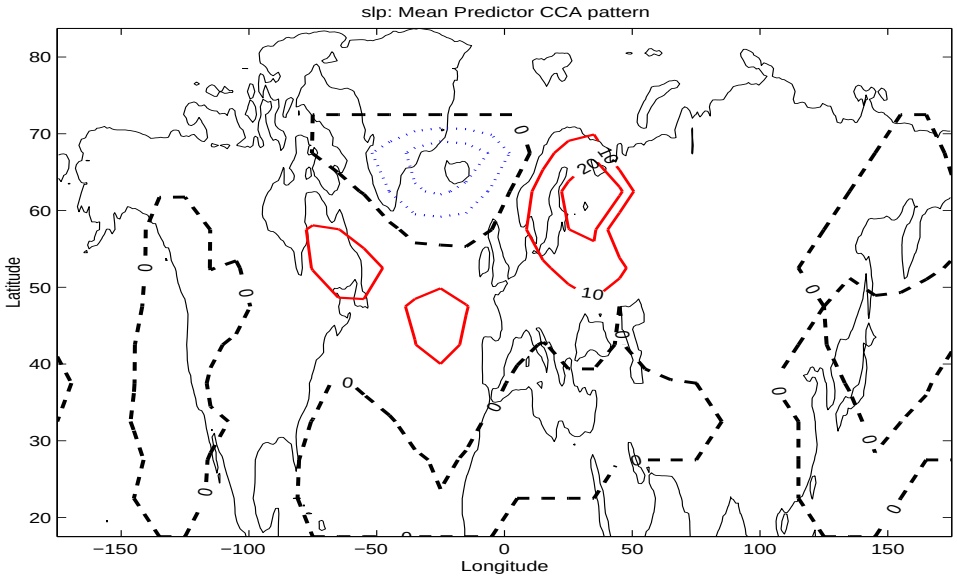

Figure 60: The mean leading October CCA UEA SLP pattern associated with the land surface temperatures. Weights in the regions where the standard deviation of the leading CCA SLP pattern estimates are greater than 10hPa are not shown.

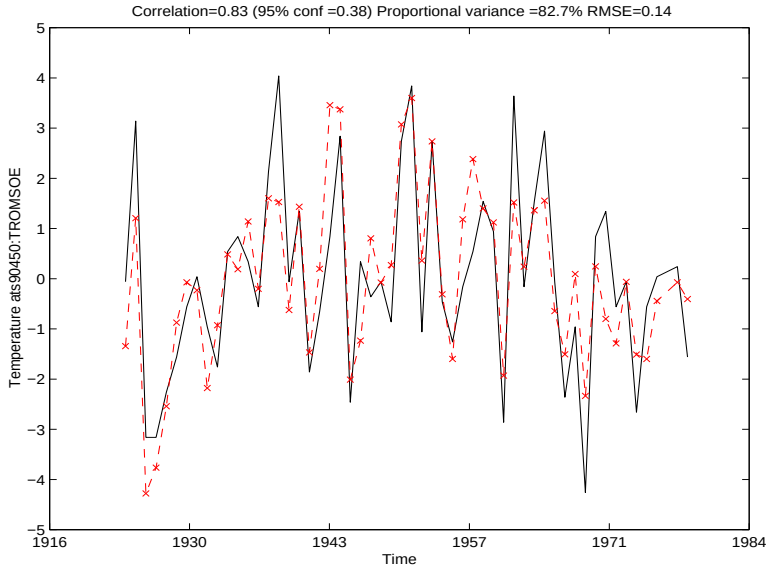

Figure 61: Time series of predicted October temperatures (dashed) at Tromsø, employing the cross-validation method with UEA SLP, shown with the observations (black solid line).

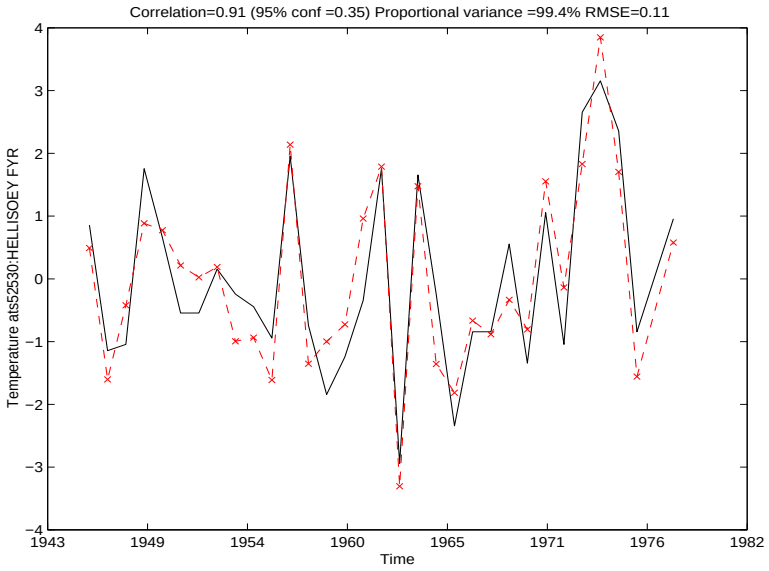

Figure 62: Time series of predicted January temperatures (dashed) at Hellisøy fyr, employing the cross-validation method with 500 hPa geopotential heights, shown with the observations (black solid line).

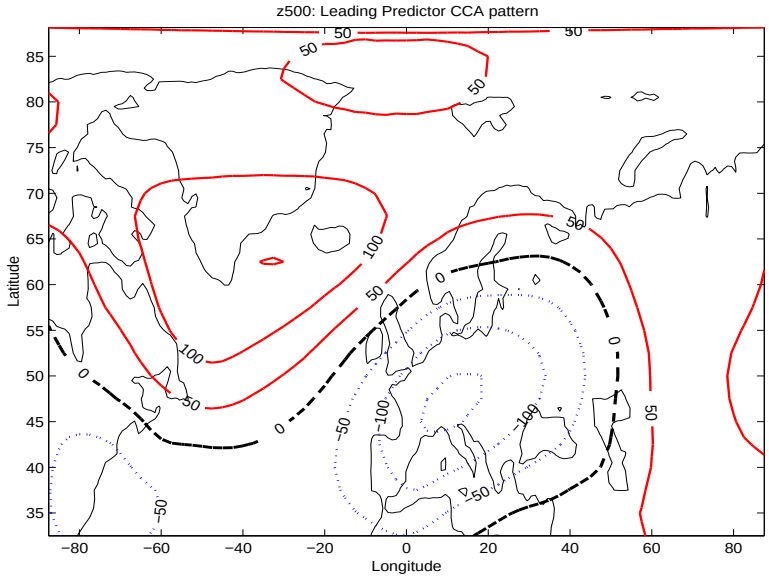

Figure 63: The mean leading January CCA 500 hPa geopotential height pattern associated with the land surface temperatures.

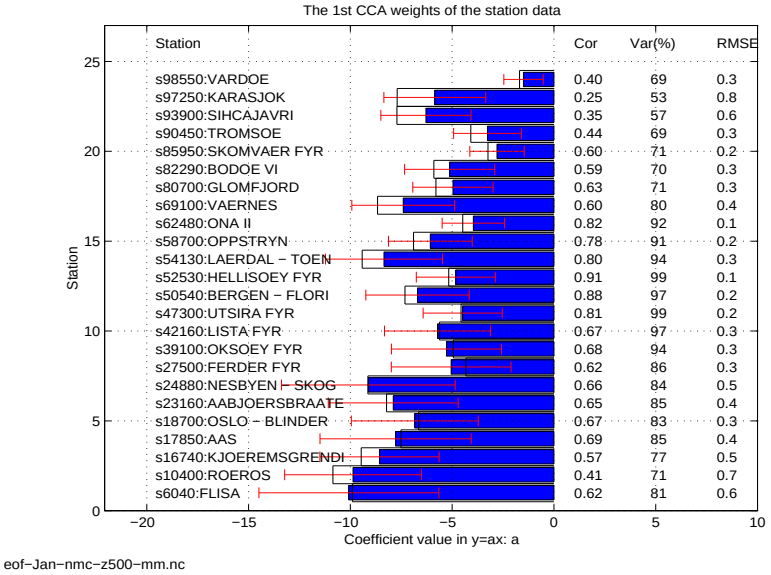

Figure 64: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading January NMC Z(500hPa) CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series.The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

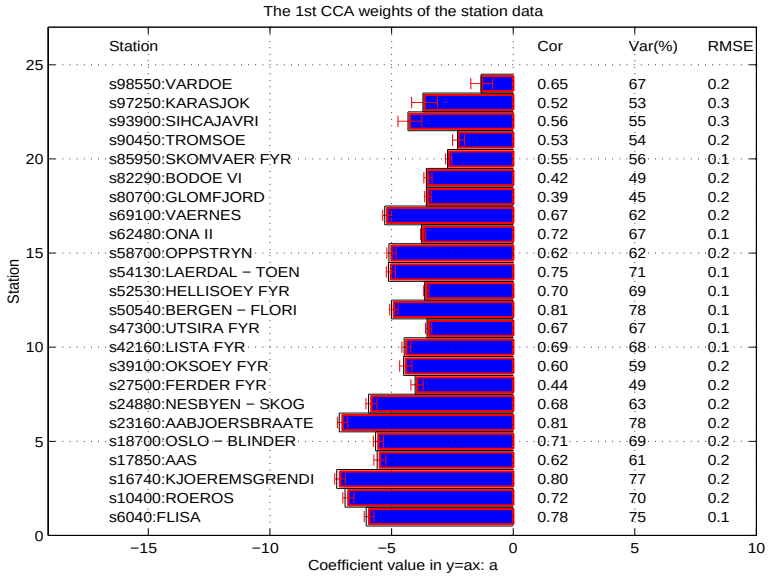

Figure 65: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading April NMC Z(500hPa) CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

January temperatures, which was for at Hellisøy fyr, and it is evident that the geopotential height model had high prediction skills. The worst prediction, which was for Karasjok, had a correlation score of 0.25 (table 21). The January CCA predictor pattern, shown in figure 63, was characterised by a dipole pattern with a positive centre over Iceland and eastern Greenland, and negative weights over the Alps. Figure 64 indicates that more SLP structures than the first pattern contributed to a significant part of the variability at all stations, as the standard deviations were large for all predictions. The CCA correlation coefficients for the entire period, r_{cca}, were 1.0000, 1.0000, 1.0000, 0.9991, 0.9629, 0.9618, 0.9036, 0.8552, 0.7313 and 0.5429. All the leading CCA weights had the same sign, suggesting that positive anomalies, $\Phi(500hPa)'$, over central Europe were correlated with high temperatures in Norway.

6.2 April NMC 500hPa Geopotential heights and land surface Temperatures

The maximum prediction skill of the April $\Phi(500hPa)$ model was lower maximum score for January. All the predictor weights had the same sign and small standard deviations (figure 65). The highest scores were seen in the south and the west coast of Norway, while northern Norway had marginally lower skills. In April, the best prediction was for Åbjørsbråten (figure 66), as opposed to Hellisøy fyr in January. The April geopotential model predicted most of the major events, although the peak values were slightly wrong. The lowest prediction skill was seen at Glomfjord, where r = 0.39. The leading CCA predictor pattern in figure 67 had a structure with four maxima. Table 22 indicates that the optimal number of EOFs for the April geopotential height model was 7.

6.3 July NMC 500hPa Geopotential heights and land surface Temperatures

The July geopotential model gave the best prediction in Karasjok, where r=0.89, RMS error was 0.17°C, and accounted for 93% of the variance (figure 68). The leading CCA predictor pattern in figure 69 had strongest anomalies over the Barents Sea and the British Channel. The leading July weights in figure 70 were in general small with large standard deviation, although the correlation scores were high. A significant part of the variability in the predictand may therefore have been due to other $\Phi(500hPa)$ structures described by higher order CCA patterns. The CCA correlation coefficients,

Table 21: Evaluation of January temperature CCA model based on January 500 hPa temperatures from NMC and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	ONA II r= 0.75 rmse= 0.15	ONA II r= 0.75 rmse= 0.15	RØROS r= 0.35 rmse= 0.62
1 4	HELLISØY FYR r= 0.84 rmse= 0.14	ONA II r= 0.79 rmse= 0.14	KARASJOK r= 0.38 rmse= 0.71
1 4 9 11 13 14 15 16 17 19	HELLISØY FYR r= 0.91 rmse= 0.11	HELLISØY FYR r= 0.91 rmse= 0.11	KARASJOK r= 0.25 rmse= 0.83

Table 22: Evaluation of April temperature CCA model based on April 500 hPa temperatures from NMC and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	BERGEN - FLORI r= 0.46 rmse= 0.17	HELLISØY FYR r= 0.30 rmse= 0.14	BODØ VI r= -0.50 rmse= 0.22
1 2	VARDØ r= 0.67 rmse= 0.16	HELLISØY FYR r= 0.24 rmse= 0.14	FERDER FYR r= -0.19 rmse= 0.23
1 2 4 10 12 18 20	ÅBJØRSBRÅTE r= 0.81 rmse= 0.15	HELLISØY FYR r= 0.70 rmse= 0.11	GLOMFJORD r= 0.39 rmse= 0.21

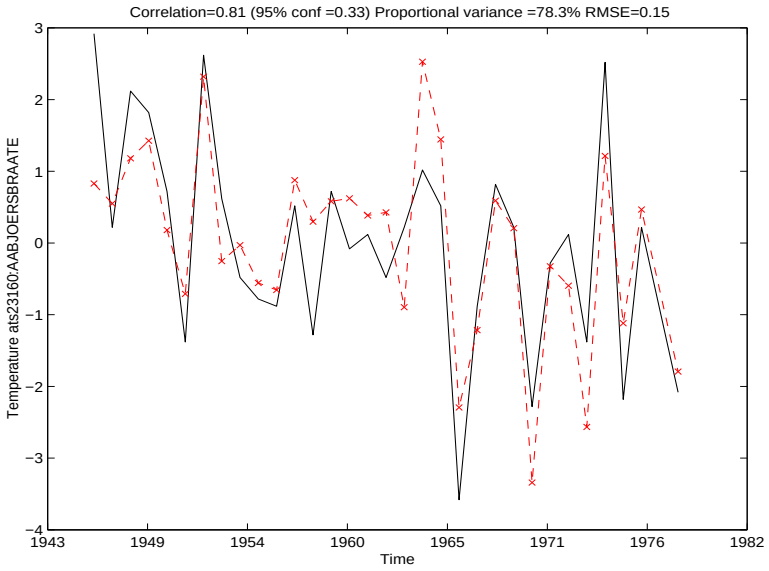

Figure 66: Time series of predicted April temperatures (dashed) at Åbjørsbråten, employing the cross-validation method with 500 hPa geopotential heights, shown with the observations (black solid line).

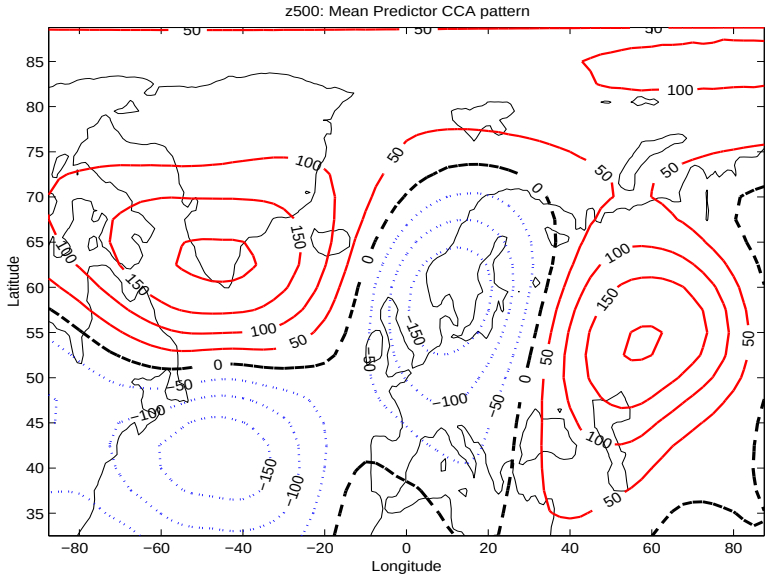

Figure 67: The mean leading April CCA 500 hPa geopotential height pattern associated with the land surface temperatures.

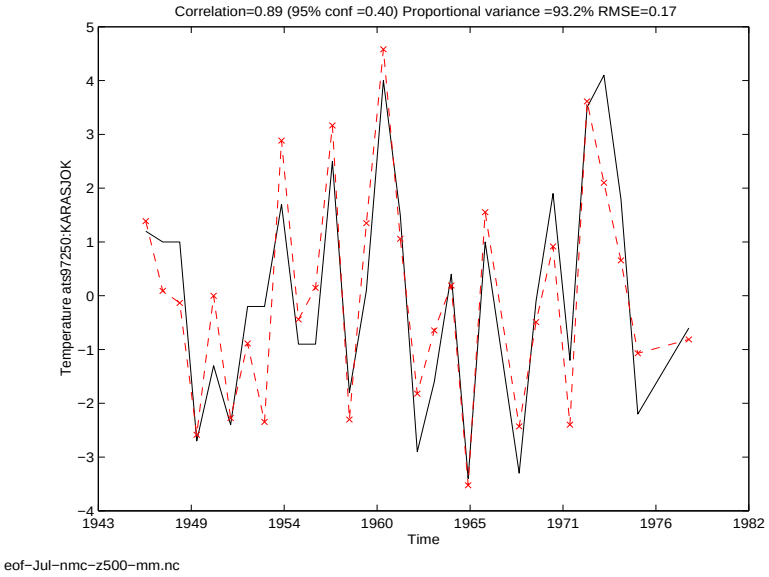

Figure 68: Time series of predicted July temperatures (dashed) at Karasjok, employing the cross-validation method with 500 hPa geopotential heights, shown with the observations (black solid line).

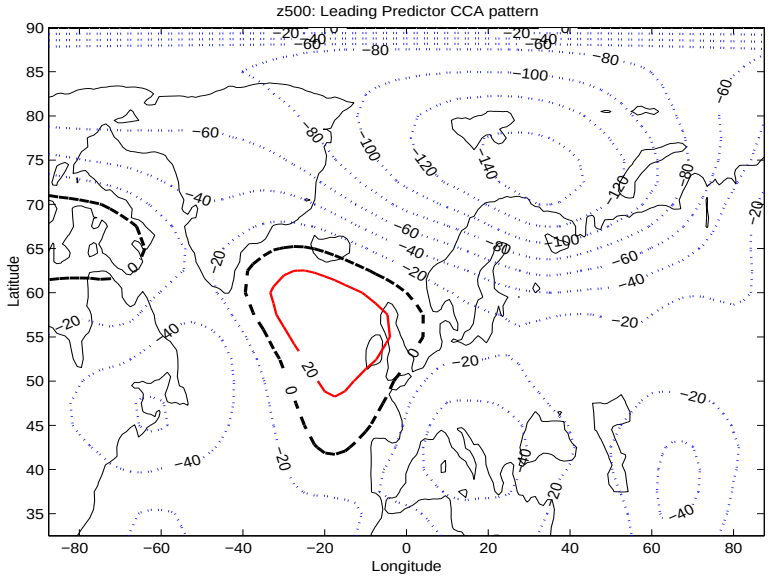

Figure 69: The mean leading July CCA 500 hPa geopotential height pattern associated with the land surface temperatures.

Table 23: Evaluation of July temperature CCA model based on July 500 hPa temperatures from NMC and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	ÅS r= 0.19 rmse= 0.27	HELLISØY FYR r= 0.12 rmse= 0.18	KARASJOK r= -0.64 rmse= 0.39
1 2	ÅS r= 0.62 rmse= 0.21	LISTA FYR r= 0.44 rmse= 0.16	ONA II r= -0.36 rmse= 0.20
1 2 3 4 6 8 9 11 13	KARASJOK r= 0.89 rmse= 0.17	BERGEN - FLORI r= 0.80 rmse= 0.12	ONA II r= 0.36 rmse= 0.20

Table 24: Evaluation of October temperature CCA model based on October 500 hPa temperatures from NMC and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	OKSØY FYR r= 0.39 rmse= 0.18	UTSIRA FYR r= 0.20 rmse= 0.16	SIHCAJAVRI r= -0.65 rmse= 0.47
1 2	TROMSØ r= 0.71 rmse= 0.23	UTSIRA FYR r= 0.63 rmse= 0.13	NESBYEN - SKOGLUND r= 0.57 rmse= 0.22
1 2 3 4	TROMSØ r= 0.89 rmse= 0.15	UTSIRA FYR r= 0.69 rmse= 0.12	NESBYEN - SKOGLUND r= 0.55 rmse= 0.23
1 2 3 4 6 7 10 15 18	TROMSØ r= 0.93 rmse= 0.12	ONA II r= 0.87 rmse= 0.10	NESBYEN - SKOGLUND r= 0.57 rmse= 0.24

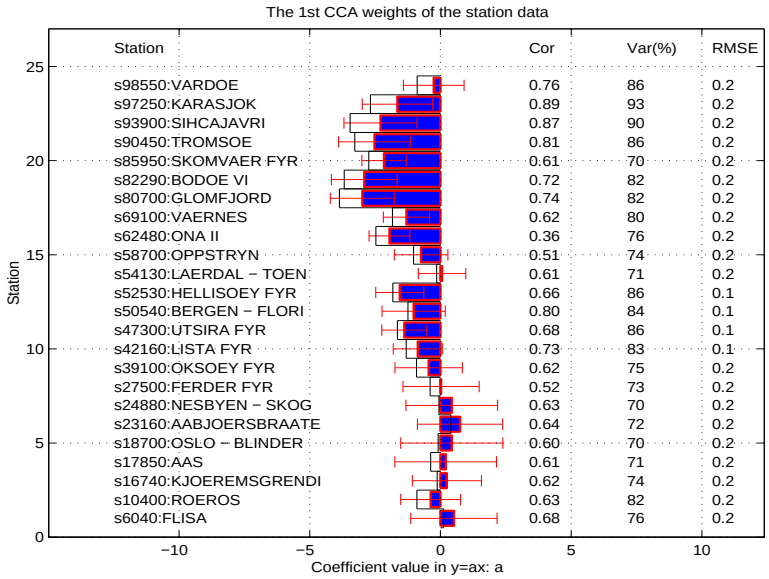

Figure 70: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July NMC Z(500hPa) CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

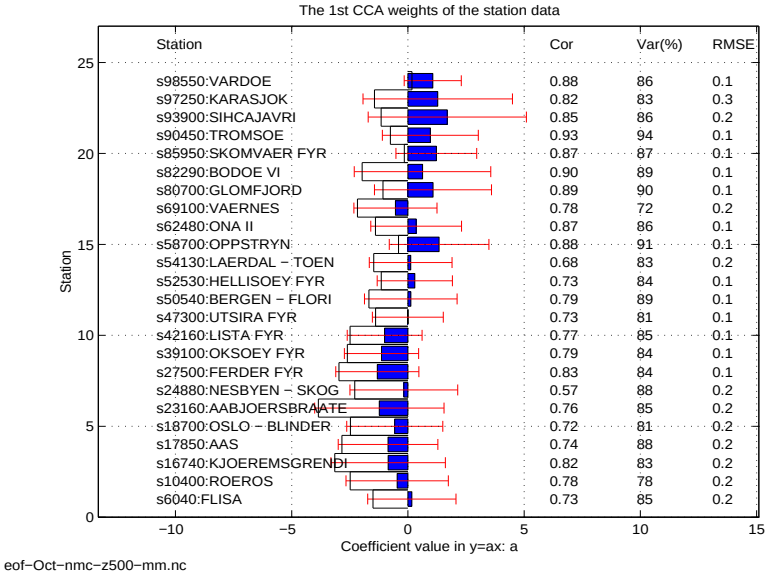

Figure 71: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading October NMC Z(500hPa) CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

r_{cca}, were 1.0000, 1.0000, 1.0000, 1.0000, 0.9916, 0.9756, 0.8994, 0.8249 and 0.6784. Table 23 indicates that the optimal model was based on 9 EOFs and the lowest skill score of the optimal model was r=0.36 and was found at Ona II.

6.4 October NMC 500hPa Geopotential heights and land surface Temperatures

The October $\Phi(500hPa)$ model had high correlation skills, with the best scores of r = 0.93 in Tromsø (figure 71) and lowest skill in Nesbyen where r= 0.57 (table 24). Virtually all the October temperature variability in Tromsø could be predicted by the model, and the predictions got the peak values nearly right (figure 72). The leading CCA prediction pattern in figure 73 described strong anomalies over the North Sea, and the large uncertainties in the model coefficient estimates could be explained as a competition of the leading CCA patterns which all had correlation near 1: r_{cca} = [1.0000, 1.0000, 0.9980, 0.9878, 0.9751, 0.9055, 0.8444, 0.7475, 0.6591]

In summary, the geopotential height model was best at predicting October and January temperatures, but the skill during the other seasons were also considered as good. The large spread in the estimates of the leading model coefficients could be attributed to the fact that several of the leading CCA all had correlations coefficients close to 1. The high prediction scores from the cross-validation analysis supported this explanation. The weak model sensitivity to the seasons is not likely to cause big non-stationarity problems with regards to downscaling future global warming scenarios.

7 Ice models

Although prognostic ice data were not available from the ECHMA4/OPYC integrations, we wanted to investigate the influence of the ice variability on the Norwegian climate. If changes in the ice conditions are important then it is important to take this into account in downscaling of global warming scenarios, by for example representing this type of forcing by a stochastic model with realistic power spectra and serial correlation. It is also possible, however, that the ice extent is correlated with the Norwegian land temperatures if both quantities are affected by similar climatic forcings. For instance, large scale cooling over the northern hemisphere may be associated with both lower local temperatures and extension of the ice sheets. Nevertheless, the result of this analysis may be useful for further studies in the relationship between sea ice, ocean and atmosphere coupling. Ice sheets act as heat sinks

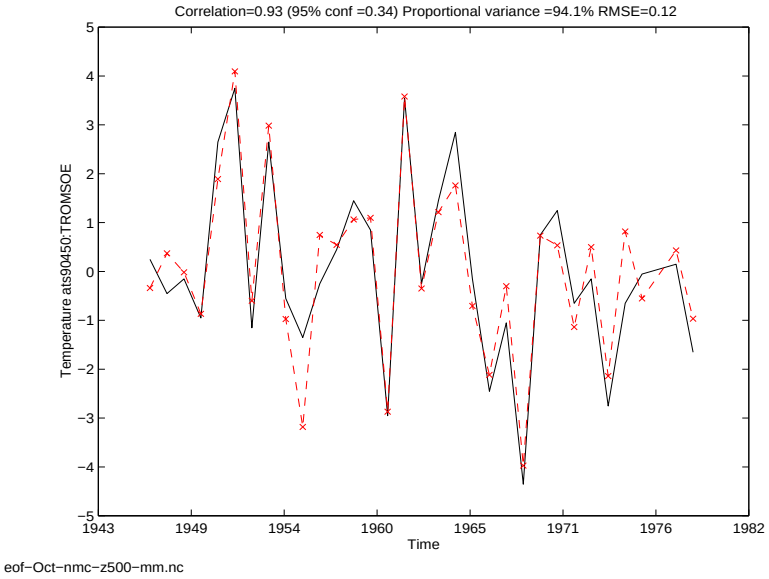

Figure 72: Time series of predicted October temperatures (dashed) at Tromsø, employing the cross-validation method with 500 hPa geopotential heights, shown with the observations (black solid line).

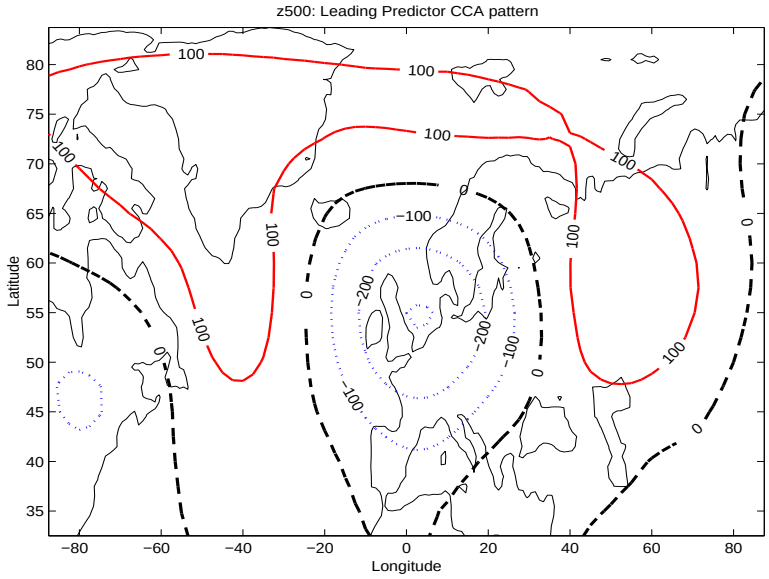

Figure 73: The mean leading October CCA 500 hPa geopotential height pattern associated with the land surface temperatures.

for both the atmosphere and the ocean through high albedo and large latent energy of melting and variation of the ice cover may have a significant effect on the planetary albedo (*Peixoto & Oort*, 1992). The ice cap in the Arctic also insulates the relatively warm ocean from the cold polar atmosphere, and the ice cover may be affected by the oceanic circulation and stratification and topographic effects as well as atmospheric forcing.

7.1 January ice and land surface Temperatures

The influence of changes in the ice extent on the Norwegian land temperatures was investigated by constructing an ice CCA model.

The optimum model had EOFs 1, 2, 3, 4, 7, 13, and 16 as predictors (table 25) and the best prediction skills were found at Vardø (figure 74), with correlation of 0.64, and RMSE of 0.21°C. The Ice model captured some of the low frequency variability for January, but missed several of the shorter episodes. The peak of the observed cold event was recorded in the late 1960s, but the lowest model temperatures were predicted 4 years later, and the prediction skills for southern Norway, shown in figure 75, were almost nil for most places. Some correlation, however, was found between the ice predictors and the temperatures at Skomvær fyr, Bodø, Glomfjord, Lista fyr, Ferder, Oksøy, Blindern, and Ås[9]. Figure 75 shows large uncertainties in the model coefficients and small correlations in the south.

The leading ice predictor CCA pattern is shown in figure 76 which describes variability near the Arctic ice edge, the Great lakes, and in the Hudson bay, with the ice edge along the east coast of Greenland and in the Labrador Sea contracting and expanding in antiphase.

7.2 April ice and land surface Temperatures

The April ice predictor pattern resembled that of January, but lacked the antiphase relationship between the east Greenland coast and the Labrador Sea (figure 77). The pattern also had stronger weights in the Baltic Sea and the White Sea off the Kola peninsula. The model skill had almost disappeared in northern Norway for April (table 26), but the correlation scores were substantially higher along the west coast of southern Norway (figure 78). The best prediction was for Oksøy Fyr, where r= 0.81, the variance 84%, and the RMS error 0.23°C (figure 79). Most of the temperature variability in

[9]Some of these stations may be related to ice extent variability in Baltic or the White Sea, which may be explained by higher order CCA patterns. The CCA correlations were 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 0.9636, 0.9202, and CCA patterns 2-4 had weights in the Baltic Sea while CCA patterns 4-5 in the White Sea.

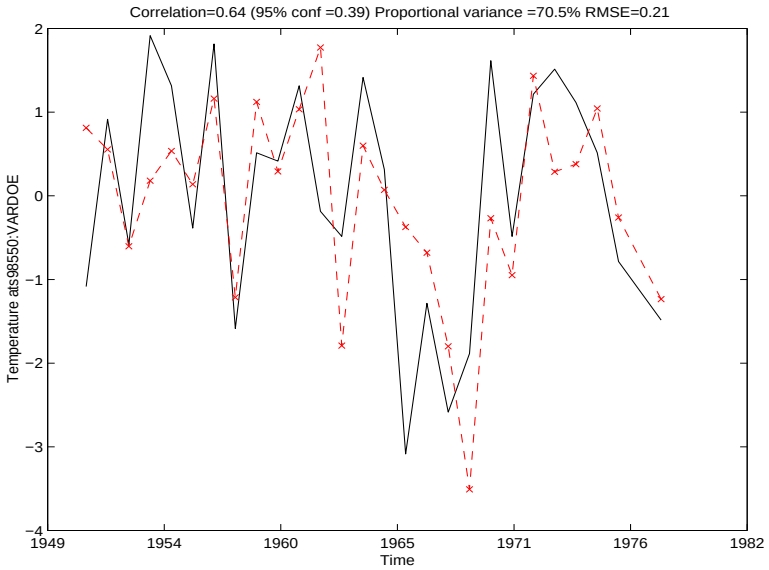

Figure 74: Time series of predicted January temperatures (dashed) at Vardø, employing the cross-validation method with GISST2.2 ice, shown with the observations (black solid line).

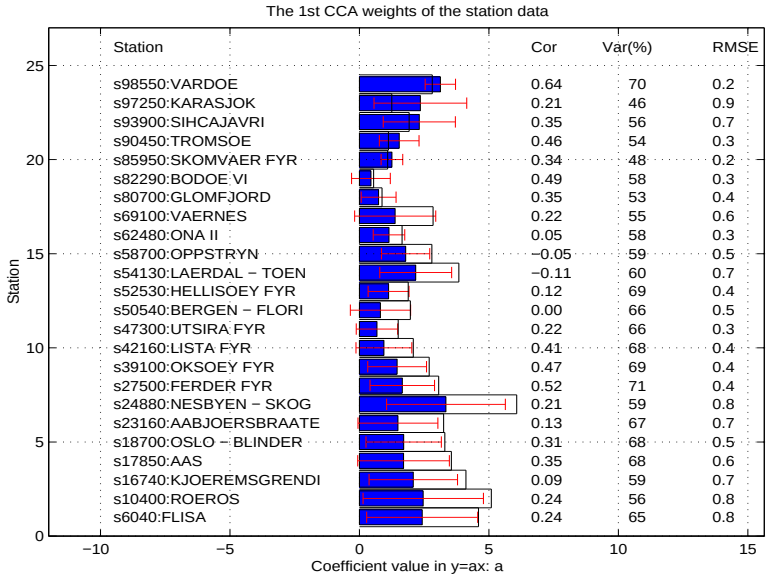

Figure 75: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading January GISST2.2 ice CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Table 25: Evaluation of January temperature CCA model based on January ice from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	FERDER FYR r= 0.41 rmse= 0.37	SKOMVÆR FYR r= -0.54 rmse= 0.25	TROMSØ r= -0.58 rmse= 0.37
1 2 3 4 7 13 16	VARDØ r= 0.64 rmse= 0.21	VARDØ r= 0.64 rmse= 0.21	LÆRDAL - TØNJUM r= -0.11 rmse= 0.68

Table 26: Evaluation of April temperature CCA model based on April ice from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	NESBYEN - SKOGLUND r= 0.53 rmse= 0.61	ONA II r= 0.32 rmse= 0.25	KARASJOK r= -0.58 rmse= 0.87
1 3 5 6 9	OKSØY FYR r= 0.65 rmse= 0.30	SKOMVÆR FYR r= 0.39 rmse= 0.23	OPPSTRYN r= 0.33 rmse= 0.38
1 3 5 6 9 14 15	FERDER FYR r= 0.77 rmse= 0.26	UTSIRA FYR r= 0.63 rmse= 0.23	VARDØ r= 0.11 rmse= 0.30
1 3 5 6 9 14 15 19	OKSØY FYR r= 0.81 rmse= 0.23	UTSIRA FYR r= 0.73 rmse= 0.20	VARDØ r= 0.09 rmse= 0.31

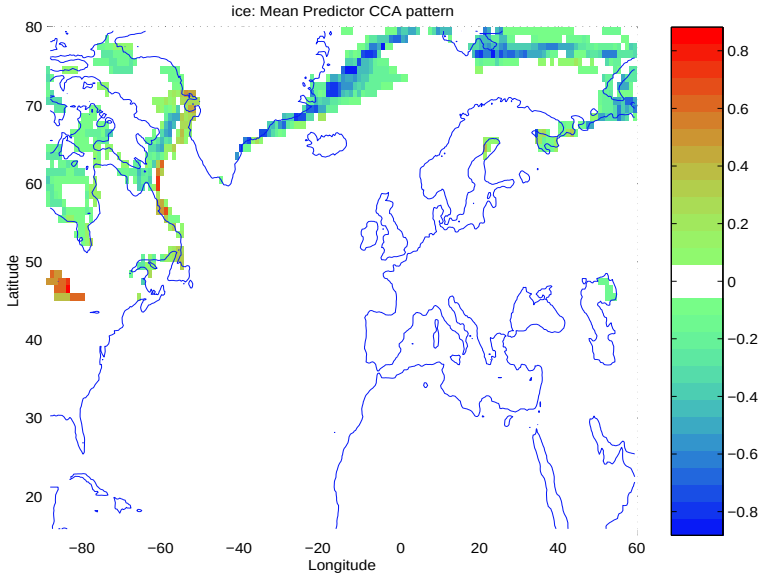

Figure 76: The mean leading January CCA Ice pattern associated with the land surface temperatures.

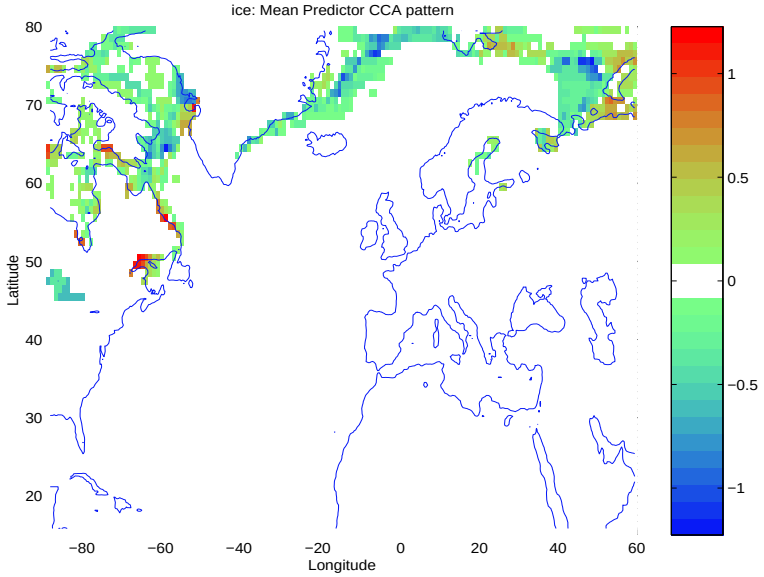

Figure 77: The mean leading April CCA Ice pattern associated with the land surface temperatures.

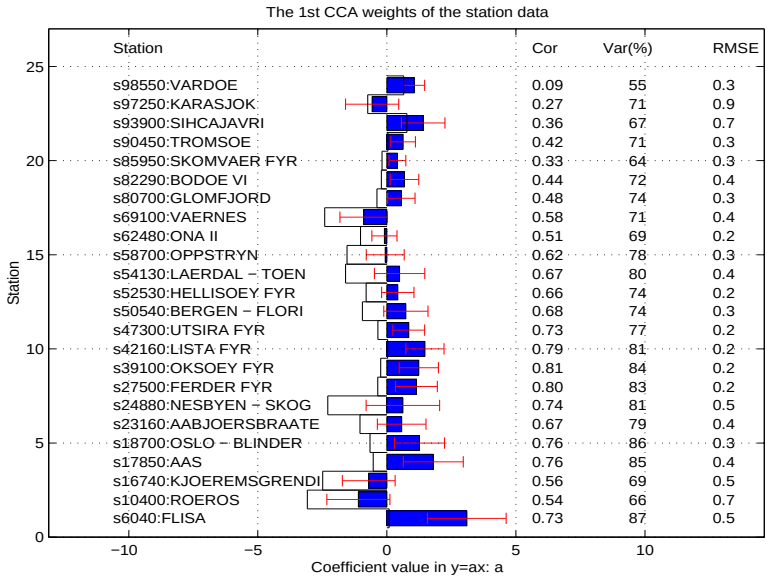

Figure 78: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading April GISST2.2 ice CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

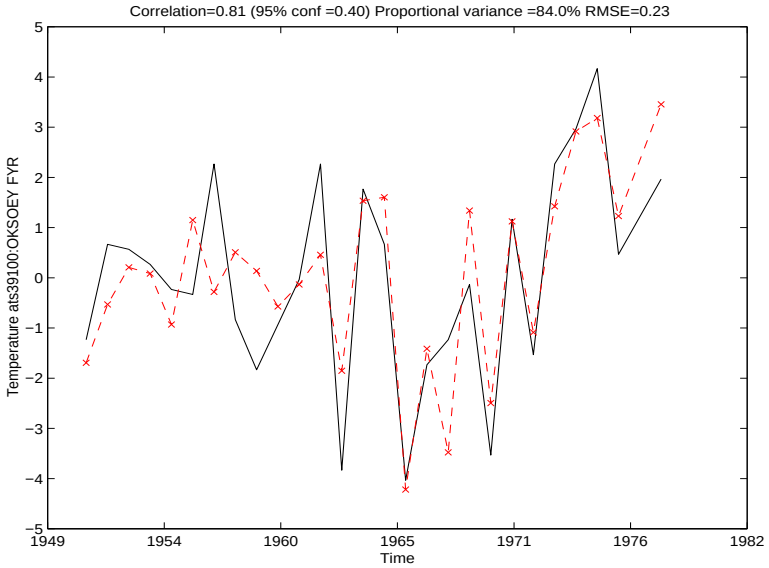

Figure 79: Time series of predicted April temperatures (dashed) at Oksøy Fyr, employing the cross-validation method with GISST2.2 ice, shown with the observations (black solid line).

figure 79 was captured by the ice model, and the stations along the west coast of southern Norway were generally characterised by high correlation scores. The large standard deviations and high prediction skill may suggest that other CCA patterns with high correlation also play a role. The CCA correlation coefficients were 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 1.0000, 0.9354, and 0.8235.

7.3 July ice and land surface Temperatures

The July Ice model was associated with medium correlation skills along the west coast of southern Norway, with maximum score at Oksøy fyr of r=0.62 (figure 80). In the rest of the country, the prediction skill was almost nil (figure 81). The predictand weights in figure 81 were associated with large standard deviations and small weights. The July ice predictor pattern described variability only along the Arctic ice edge (figure 82), and it is not clear how the ice extent and the temperatures along the west coast were interrelated. It is possible that both were affected by large scale oceanic or atmospheric circulation. Table 27 indicates that the optimal July ice model was based on 7 EOFs.

7.4 October ice and land surface Temperatures

The leading ice CCA pattern for October, shown in figure 83, had similar features to July, but with less variability along the ice edge. The predictions for the October temperatures had high correlation skills for Karasjok (r=0.80, figure 84) and Sihcajarvi, but figure 85 indicates large spread in the estimates of the model coefficient. It is likely that more than one CCA pattern played a role for the October temperatures in northern Norway. In general, inland stations in northern Norway were associated with good prediction skills, and the coastal stations marginal skill. The skill scores in southern Norway were almost nil.

The ice model demonstrated that the ice extent was correlated with temperatures over parts of Norway. The relation between the ice and the local climate was strongest in northern Norway for October and January months, but about non-existent during April and July. On the other hand, the temperatures on the south coast of Norway were correlated with the ice variability during April and July, but not during October. The strong model sensitivity to the seasons suggested that the model may be strongly non-stationary with respect to a green house gas scenario. The fact that the ice model only gave good predictions over smaller regions appears to be inconsistent with

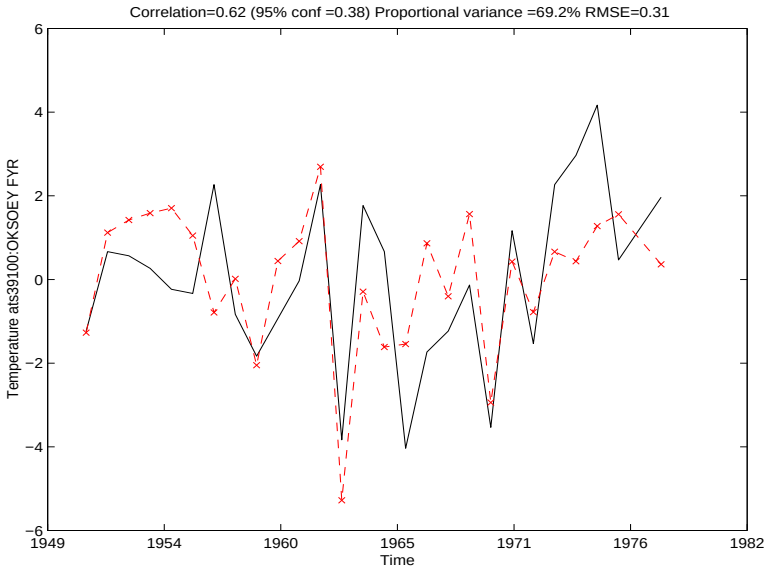

Figure 80: Time series of predicted July temperatures (dashed) at Oksøy fyr, employing the cross-validation method with GISST2.2 ice, shown with the observations (black solid line).

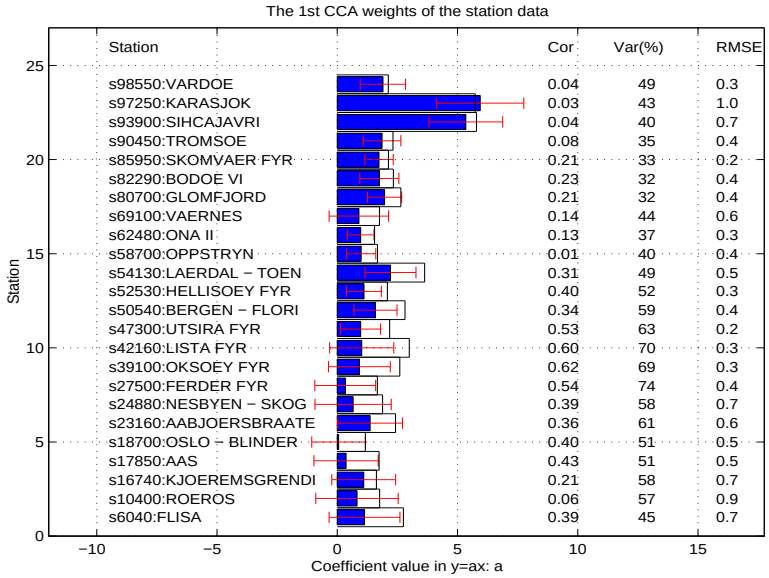

Figure 81: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading July GISST2.2 ice CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

Table 27: Evaluation of July temperature CCA model based on July ice from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	FERDER FYR r= 0.27 rmse= 0.39	SKOMVÆR FYR r= -0.03 rmse= 0.24	VARDØ r= -0.36 rmse= 0.27
1 4 10 11 13 15 16	OKSØY FYR r= 0.62 rmse= 0.31	SKOMVÆR FYR r= 0.21 rmse= 0.25	OPPSTRYN r= 0.01 rmse= 0.44

Table 28: Evaluation of October temperature CCA model based on October ice from GISST2.2 and surface temperatures from DNMI's climate data base

EOFs included	Maximum correlation location (independent data)	Minimum RMSE (predictand)	Smallest correlation ('Worst prediction')
1	VARDØ r= 0.32 rmse= 0.25	SKOMVÆR FYR r= 0.28 rmse= 0.23	KJØREMSGRENDI r= -0.03 rmse= 0.59
1 2 4 5 6 7 8 10 13 16 18	KARASJOK r= 0.80 rmse= 0.51	SKOMVÆR FYR r= 0.62 rmse= 0.20	LISTA FYR r= -0.13 rmse= 0.56

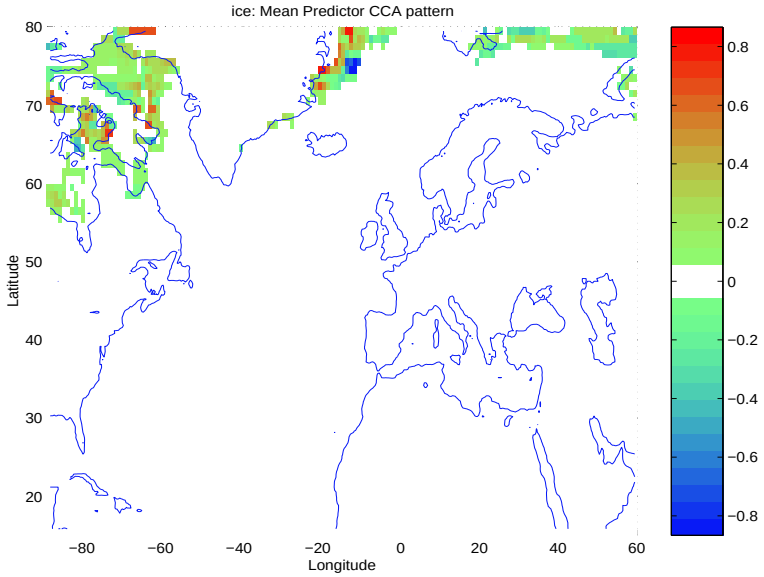

Figure 82: The mean leading July CCA Ice pattern associated with the land surface temperatures.

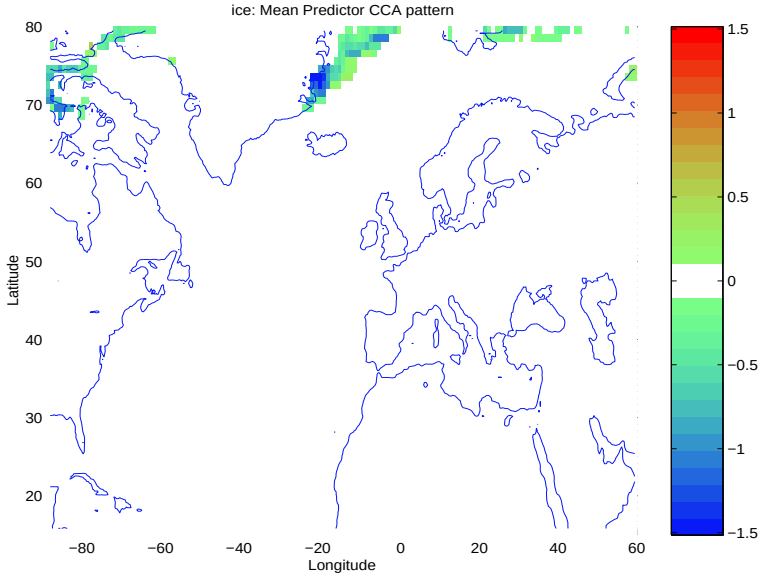

Figure 83: The mean leading October CCA Ice pattern associated with the land surface temperatures.

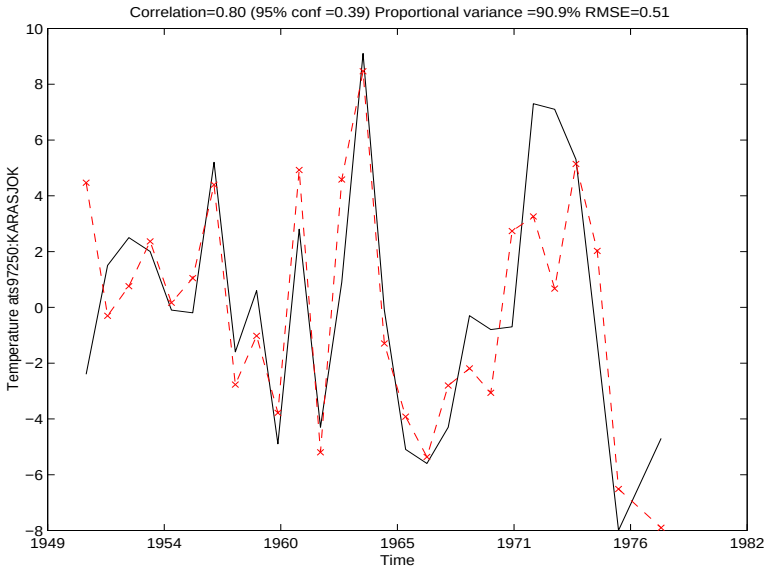

Figure 84: Time series of predicted October temperatures (dashed) at Karasjok, employing the cross-validation method with GISST2.2 ice, shown with the observations (black solid line).

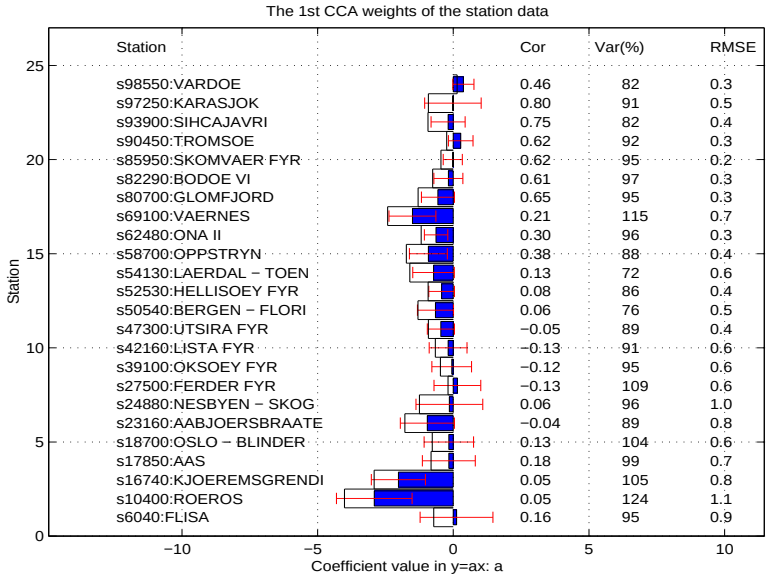

Figure 85: The mean weights (model coefficients) from the cross-validation analysis shown as filled bars, indicate the importance of the leading OctoberGISST2.2 ice CCA pattern for the land surface temperatures. The empty black boxes show the weights from a model trained on the whole time series. The error bars indicate the standard deviation and hence the spread in samples of each coefficient. The correlation, variance and RMSE results from the cross-validation analysis are given on the right hand side.

the hypothesis of both ice coverage and Norwegian station temperatures being forced by a planetary scale cooling. The reason why ice correlates with temperatures in southern part of Norway during summer is not entirely understood to our knowledge, but deserves further investigation, for instance by making numerical experiments.

8 Model stationarity

We have already mentioned model stationarity, and by investigating the uncertainties in the model coefficient estimates from different calibration data combinations, a crude picture of how the models depend on the training period is emerging. The fact that model skill varies with the season, demonstrates that the relationship between predictors and predictands is not constant throughout the year.

Another problem is that the PCA results for the present climate, on which our models are based, may not span the data space that corresponds to a global warming scenario. We have not tested the stationarity of the PCA products in this report, but the issue of non-stationary EOFs may be a problem if there is a sudden change in the large scale circulation in the future.

In order to investigate the model stationarity further, the data were split in two, where the first subset was used for model calibration and the second subset for evaluation. The global temperatures for calibration period were lower than the validation period with the exception of the 1940s (*Bengtsson*, 1996, p.298), and thus using a warmer validation period provided a good test to investigate the stationarity of the models in a slowly warming climate. Only the longest data sets (GISST2.2, NCAR, UEA) will be discussed here.

8.1 Stationarity test for the SST models

Because only the January SST models were associated with high skill scores, only the January models were examined here. Both North Sea and North Atlantic models were tested.

8.1.1 Local SST models

The stationarity of the January North Sea SST model was tested by calibrating the model over 3 different periods, 1923-1955, 1923-1945, and 1923-1965, and using the remaining period up to 1978 for validation. The predictions for the 1955-1978 validation period using the January North Sea SST model

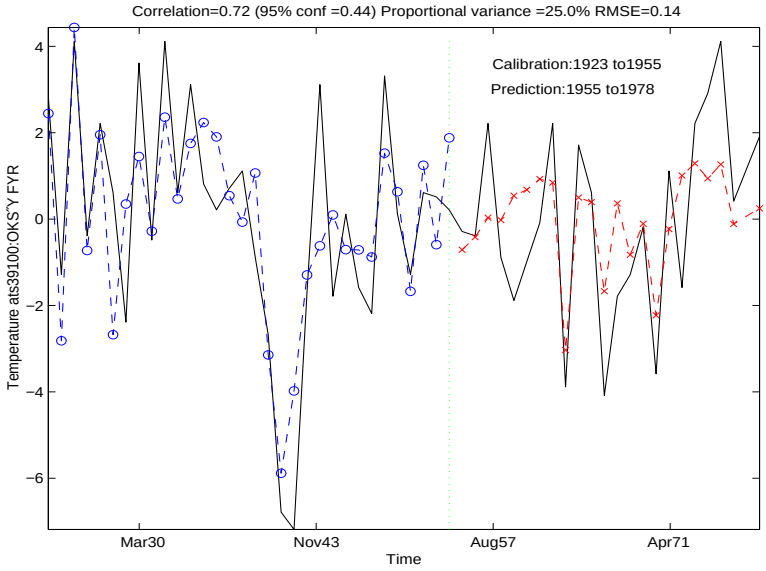

Figure 86: Time series of predicted January temperatures using the January North Sea SST model at Oksøy fyr, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

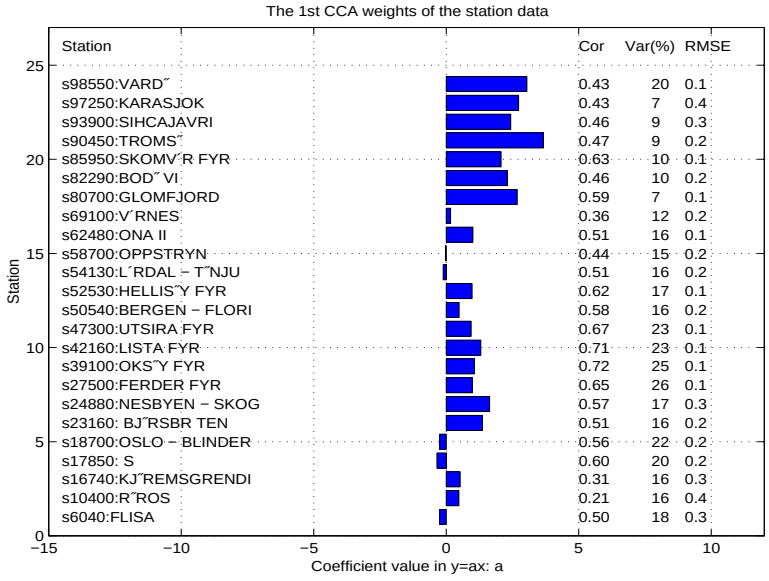

Figure 87: The model weights for the leading January North Sea SST CCA pattern and the correlation, variance, and RMS skill scores for the validation period 1955-1978.

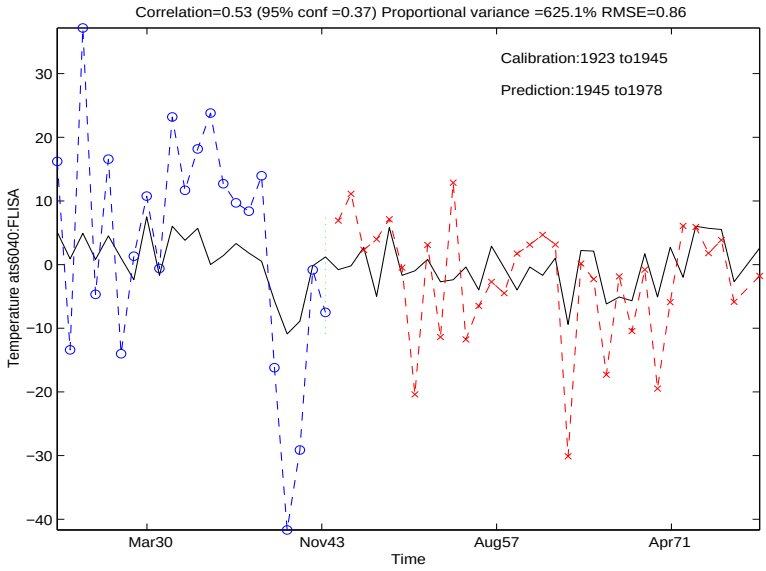

Figure 88: Time series of predicted January temperatures using the January North Sea SST model at Flisa, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

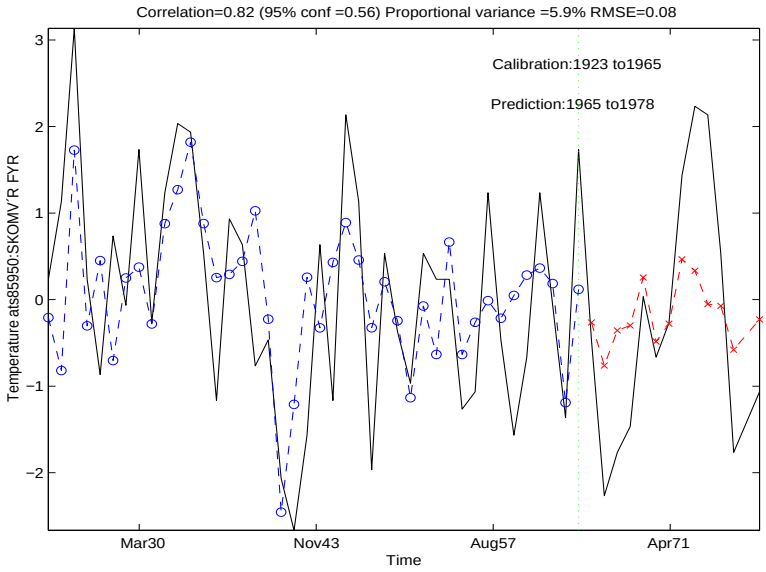

Figure 89: Time series of predicted January temperatures using the January North Sea SST model at Skomvær fyr, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

had a correlation coefficient of 0.72 with the observations, and is shown in figure 86. The observations are shown in solid black and the predictions using the *dependent* data from the calibration period as circle-dashed. The predicted *independent* data are shown as cross-dashed lines. Although the model did not capture the 1958 cold event and the the warm peak in the 1970s, the results were promising. Figure 87 shows the validation period skill scores for all the stations, and the correlation scores for several stations were higher than 0.60. It is therefore reasonable to assume that the January North Sea SST model calibrated on the 1922-1978 period[10] will give reasonable for at least the next 20-30 years.

Tests with shorter calibration periods, 1922-1945, gave poor results when the model training data set was too short for model calibration and the square root of the covariance was complex (figure 88). A longer calibration period increased the correlation scores (figure 89), although the validation period was now reduced and one should strictly only compare the score of validation periods of equal lengths and at same locations. It is also evident that the SST model does not capture the 1970's warm peak very well.

8.1.2 North Atlantic SST models

Figure 90 shows the best prediction for the 1923-1955, where 32 years of data went into the model calibration and 23 years were used for validation. It is evident that the model produces realistic simulations of the 'future' climate, although some smaller events were not captured by the model. The statistics for all the stations are given in figure 91.

The 22-year period spanning 1923-1945 was clearly too short for the model, and both the predictions for the calibration as well as the validation period was poor (not shown). As with the corresponding North Sea model, the square root of the covariance matrices was complex as a result of too short calibration period. The model calibrated over the 1923-1950 period gave reasonable predictions, although it did not capture all the small fluctuations (figure 92). Longer calibration period of 44 years gave improved predictions for the validation period (figure 93). It is therefore reasonable to believe that a North Atlantic SST model trained on the 1922-1978 period will give realistic simulations for at least the subsequent 30 years if the warming trend continues as it has done in the past. We therefore do not believe

[10]Unfortunately, some of the stations in the DNMI climate data base with long records were closed down in the late 1970s. We hope to be able to extend the calibration period to 1997 in the future by excluding the stations that to not have recent data or use the NACD data as predictors. As this report is about the construction and testing of the empirical models, the choice of time period is not crucial.

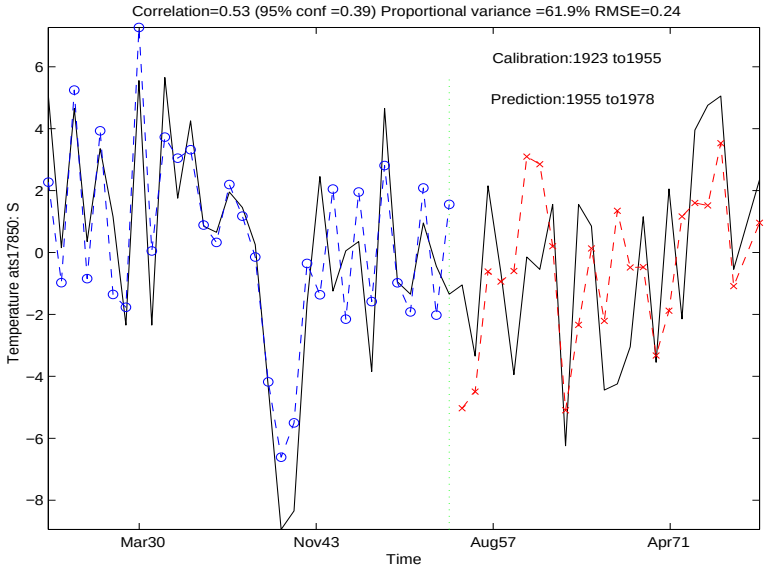

Figure 90: Time series of predicted January temperatures using the January North Atlantic SST model at Ås, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

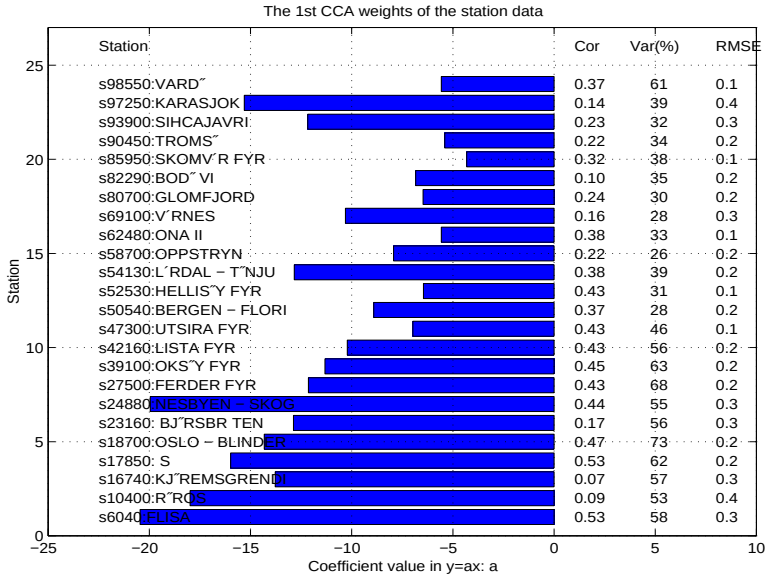

Figure 91: The model weights for the leading January North Atlantic SST CCA pattern and the correlation, variance, and RMS skill scores for the validation period 1955-1978.

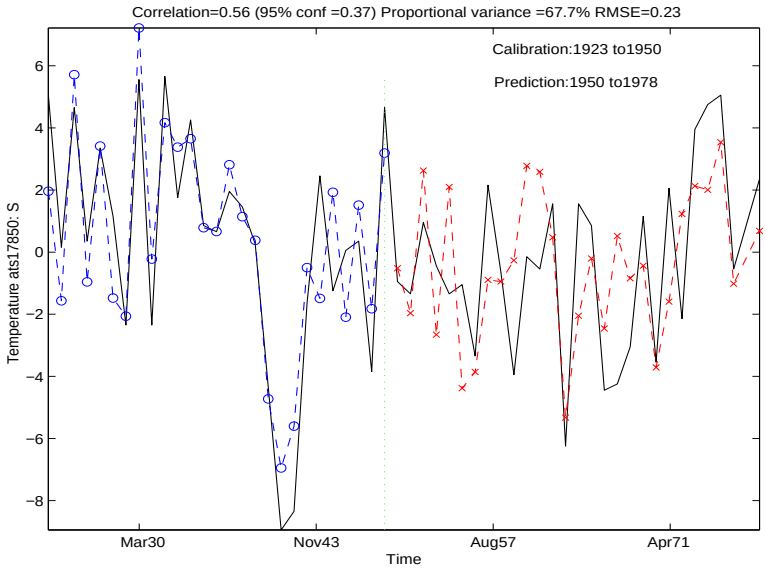

Figure 92: Time series of predicted January temperatures using the January North Atlantic SST model at Ås, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

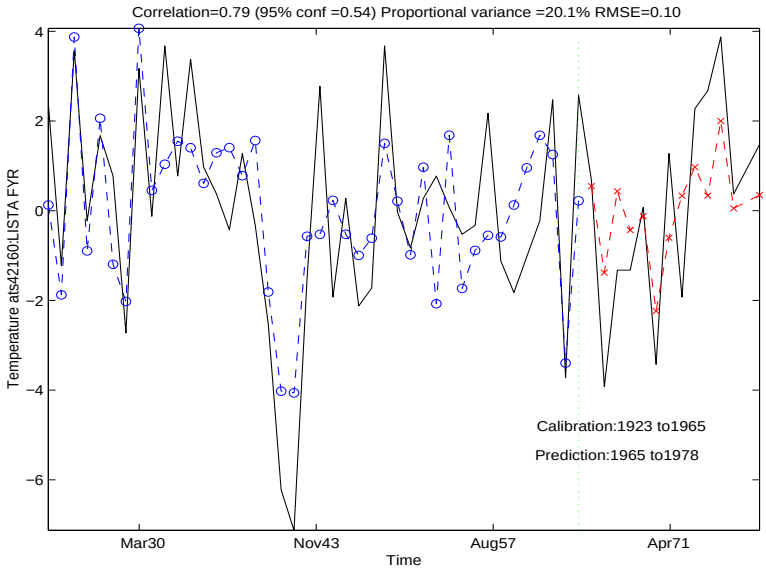

Figure 93: Time series of predicted January temperatures using the January North Atlantic SST model at Lista fyr, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

that non-stationarity will be a serious problem for predictions of near-future scenarios.

The North Atlantic SST models captured more of the 1970's warm event than the North Sea models, although the correlation scores of the January 1923-1955 North Atlantic model in general were lower than those for the corresponding North Sea model.

8.2 Stationarity test for the SLP models

The UEA model was tested for stationarity in a similar fashion as the SST models described above. The 3 models were calibrated using the 1923-1955 period (figure 94), 1923-1950 period (figure 95), and the 1923-1965 period (figure 96) respectively. Figure 97 shows the skill scores of all the stations for the 1923-1955 model. It is evident that all these models captured most of the variability, and even the 1923-1950 model had high prediction skills.

The same test of UEA April and July models revealed no prediction bias for the 1950 to 1978 validation period, although the predictions were not as good as for January (not shown). In particular, the July model missed the cold episode around 1965 while the April model underestimated the peak values. The October model gave better predictions than the April and July models, but did not obtain as high correlation scores as the January model. There were no indications of changing relationship between the large scale SLP patterns and the local temperatures.

8.3 Discussion of the model stationarity

In summary, the SST models gave reasonable simulations of temperature variability at 30 years lead time, but did not describe the peak values of all the events. No evidence for non-stationarity was found with regards to the SST models, and it is reasonable to assume that they can give a good description of near-future climate scenarios.

The UEA SLP model captured virtually all the major events in the validation period, regardless of calibration and validation period lengths. The results from the model stationarity tests demonstrated that this type of models is unlikely to give distorted results as a result of non-stationarity if they are used as downscaling models for global warming scenarios with lead times up to at least 50 years, but there are no reasons to expect the model skill to deteriorate rapidly beyond 50 year lead time.

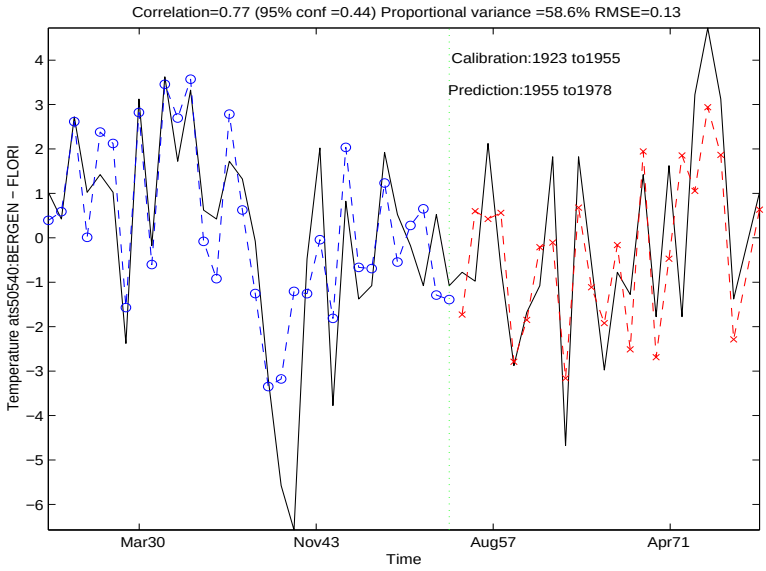

Figure 94: Time series of predicted January temperatures using the January UEA SLP model in Bergen, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

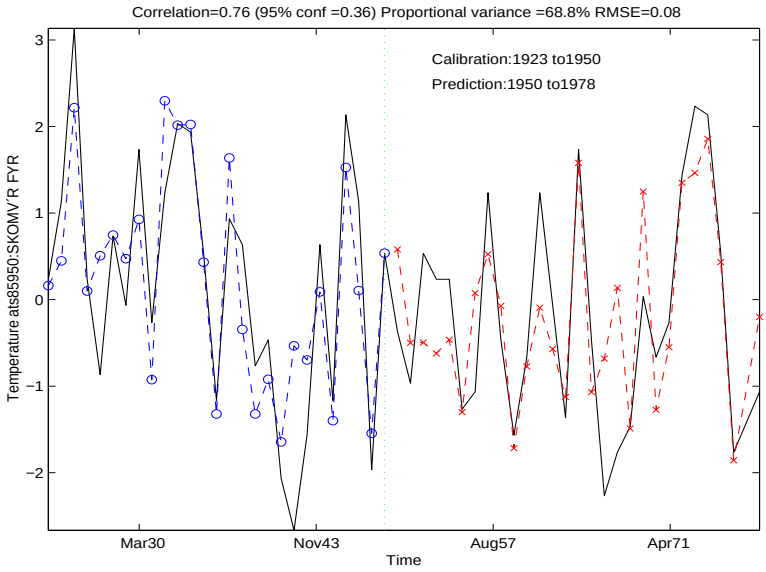

Figure 95: Time series of predicted January temperatures using the January UEA SLP model at Skomvær fyr where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

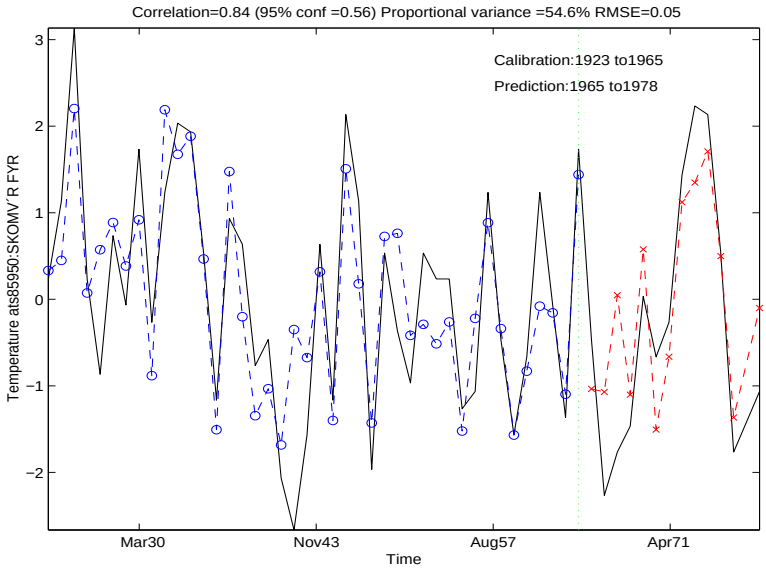

Figure 96: Time series of predicted January temperatures using the January UEA SLP model at Skomvær fyr, where the predicted values for the calibration period (circle dashed) and validation period (cross dashed) are shown with the observations (black solid line).

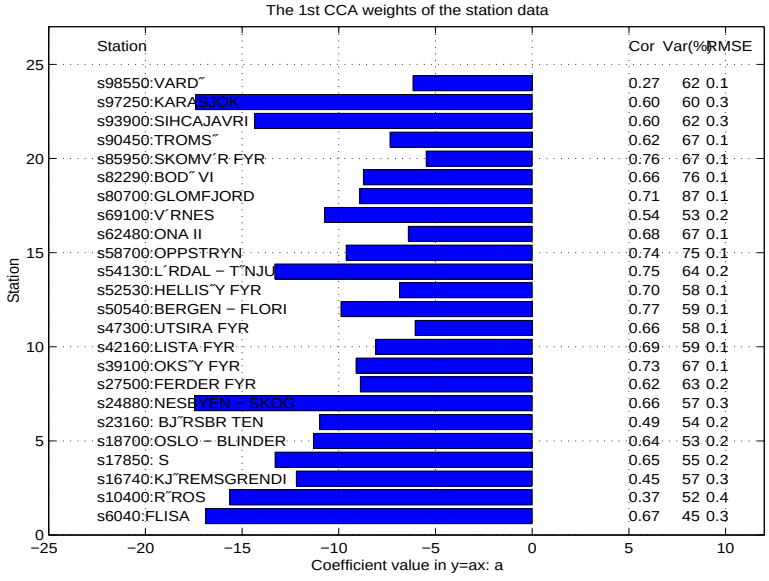

Figure 97: The model weights for the leading January UEA SLP CCA pattern and the correlation, variance, and RMS skill scores for the validation period 1955-1978.

9 Discussion and Conclusion

The Norwegian surface air temperatures are strongly influenced by the large scale SLP and geopotential height patterns for all seasons. It has been demonstrated that the SLP and geopotential models were suitable as downscaling models for future scenarios, as the relationships between the large scale SLP patterns and the local temperatures were stationary.

CCA analysis indicates that Norwegian winter time temperatures are influenced by both local SSTs (in the North Sea, Skagerrak, Kattegat, and the Baltic Sea) and remote SSTs (large SST anomalies in the North Atlantic interior and off the coast of New England), but that summer and autumn SSTs have little influence on the Norwegian temperatures. The fact that the winter time SSTs seem to play a role is consistent with coupled North Atlantic mode hypothesis (*Sutton & Allen*, 1997; *Latif* et al., 1996; *Deser & Blackmon*, 1993), which involves both SSTs and SLP. It was remarkable that the local SST model could nearly predict the extreme low temperatures over southern Norway in the late 1930s and early 1940s, while the SLP predictors underestimated the cold temperatures by almost a factor of 2. This result suggests that extreme cold events may be influenced by the local SST anomalies rather than just the large atmospheric circulation. The role of SSTs has not been completely explored, as advection of warm water by the North Atlantic current may be responsible for ice-free conditions off the coast of Northern Norway, and a reversal of this current may result in significantly cooler SSTs and affect the Scandinavian climate (*Steig* et al., 1998; *Cane*, 1998; *Broecker*, 1991).

The normal ice variability along the Arctic ice sheet and in the Baltic and White Sea appeared to be correlated with the Norwegian temperatures. In January and October, the highest correlations were found in Northern Norway while during April and July, the temperatures in southern Norway showed significant correlation with the ice predictors. Of all the predictors, ice gave the best correlation scores during April, and there is no obvious explanation for this relationship, although high Atlantic ice extent may be associated with low global temperatures.

It is important to keep in mind the limitations of statistical downscaling, especially when applied to model results from Green House Gas (GHG) integrations using global coupled General Circulation Models (GCMs). The statistical models are based on historical data, and there are not guarantee that the past statistical relationships between different data fields will hold in the future. In the Model stationarity section, we have applied a simple sensitivity test to see if the statistical model can be extrapolated to situations where the northern hemisphere generally is warmer than during the

model training period. The results from these sensitivity studies suggested that the SLP models may justifiably be applied to global warming scenarios to predict likely regional climate changes. The January SST model also passed the stationarity test, however, the predictions from the SST model were generally not as good as those of the SLP models. It is unlikely that the relationships between the local climatic parameters and different large scale predictor quantities change in a similar fashion for all types of predictors. Cross-validation for different seasons suggested best skills during the winter period and generally lowest skills during the spring and summer for all except the SST models, which gave the lowest scores in October (table 29). The seasonal variations in model skill was strongest for the SSTs, and less marked for NMC SLPs. If in a severe global warming scenario the winter season resembles the present summer season, then the comparison between seasonal model skills may indicate a deterioration in the SST model reliability.

Other short-comings of the empirical models may be that the ones described here only can describe coupled patterns that are approximately linearly related. If the relationship is strongly nonlinear, then the same models can be applied to transformations of the original data, i.e. $y = Mx^2 \rightarrow y = Mx'$ where $x' \equiv x^2$.

One should also be concerned about the uncertainties associated with the GHG GCM results as well as those of the downscaling methods themselves (*Wilby* et al., 1998). It is well known that these low resolution models are far from perfect, and that they have problems associated with for instance cloud representation, atmosphere-ocean coupling, and artificial climate drift (*Bengtsson*, 1996; *Anderson & Carrington*, 1994; *Treut*, 1994). Part of the problems are due to incomplete understanding of the climate system. The important mechanisms causing variability such as ENSO and NAO are for instance still not understood (*Sarachik* et al., 1996; *Anderson & Carrington*, 1994; *Philander*, 1989). Due to discretisation and gridding of data, it is unlikely that the global GCMs will simulate regional details realistically (*Crane & Hewitson*, 1998; *Zorita & von Storch*, 1997; *von Storch* et al., 1993; *Robinson & Finkelstein*, 1991). However, because a wide range of global GCMs predict similar global behaviour, it is believed that the GCMs may be useful for predicting large scale features.

The dynamical downscaling approach provides an alternative to the statistical downscaling, but without assuming that historical relationships between large scale circulation and local climate remain constant. In theory, these nested dynamical models are physically consistent representation of a small region of the atmosphere, and it is indeed remarkable that the dynamical climate models reproduce the main features of the climate as realistically as they do, considering that they are based on merely fundamental physical

laws. However, the dynamical models are not perfect and there are some draw backs associated with dynamical downscale models, such as:

(i) The dynamical downscaling models are tuned for the present climate. Cloud schemes are parameterised and based on empirical relationships (*Bengtsson*, 1996; *Heyen* et al., 1996) and the parameterisation of cloud radiation is notoriously difficult to implement in climate and weather forecast models (*Palmer*, 1996). Furthermore, we do not know if these parameterisation schemes will be valid in a global warming scenario even if they were appropriate for the present climate. This issue also applied to the global GCMs, whose results are used as predictors for the downscaling models.

(ii) The dynamical downscaling models are to date extremely expensive to run, and only a few integrations can be afforded. This inhibits the use of dynamical models for long integrations and extensive hypothesis testing.

(iii) Upscaling instabilities (*Palmer*, 1996; *Lorenz*, 1967, 1963) are filtered out: these may not be important in operational weather forecast integrations, which are integrated over a shorter period, but upscaling instabilities can cause inconsistencies in longer integrations between the model boundary values and the internal dynamics. The SSTs taken from the coupled GCM run may also not be appropriate as boundary values in a nested model if these are sensitive to local ocean forcing (like Ekman pumping).

(iv) Schemes to counter numerical instabilities are usually needed due to the fact that the model consists of discrete values on a discrete grid. Furthermore, there are no perfect numerical intergation schemes, and results from dynamical models are subject to round off errors as well as numerical diffusion and inadequate conservation properties (*Palmer*, 1996; *Press* et al., 1989).

Merits of statistical downscaling:

(i) They are cheap to run, which means that we can apply these statistical models to results from a number of different global coupled GCMs. We can therefore get an idea of the uncertainties associated with the GCMs.

(ii) They can be tailored for specific use, and the statistical models can be optimised for the prediction of certain parameters at specified locations, as for instance specified by customers. This makes the statistical model approach ideal for end users.

(iii) Dynamical downscaling models still have a low spatial resolution in terms of impact studies, and one may still have to apply some kind of downscaling/MOS technique to the dynamical model results.

(iv) The statistical models are appropriate in MOS application and are also promising tools for seasonal forecasting where for instance the SSTs may influence the local climate (during the winter season).

(v) The statistical models can be used to find coupled patterns between

Table 29: Summary of best 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	ICE	Max/Min
Jan	0.72	0.75	0.84	0.91	0.89	0.91	0.62	0.89 / 0.64
Apr	0.65	0.66	0.66	0.67	0.60	0.81	0.62	0.82 / 0.60
Jul	0.61	0.48	0.72	0.79	0.79	0.89	0.62	0.90 / 0.48
Oct	0.48	0.37	0.86	0.88	0.83	0.93	0.62	0.94 / 0.37
avg.	0.62	0.57	0.77	0.81	0.78	0.89	0.72	

two different climatic parameters, and hence provide a basis for analysing both historical data as well as the results from dynamical downscaling. An alternative approach to improve our physical understanding is to run model experiments with GCMs or nested models, however, this kind of numerical experiments requires substantial computer resources and is expensive.

In summary, both downscaling approaches have their weaknesses and strengths, and it is difficult to say which is superior. It is clear that different methods are appropriate for different use, and it is therefore important to apply both dynamical and statistical downscaling to global GHG GCM results. A comparison between the two fundamentally different techniques can give us a measure of uncertainty associated with the predictions. If both methods give similar answers, then we may at least have *some* confidence in the results.

References

Anderson, D.L.T, & Carrington, D. 1994. Simulation of Tropical Variability as a Test of Climate Models. *In:* Speranza, A., & nd R. Fantechi, S.Tibaldi (eds), *Global Change.* Environment and Quality of Life, vol. EUR 15158 EN. European Commission.

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

Bengtsson, L. 1996. The Climate response to the Changing Greenhouse Gas Concentration in the Atmosphere. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Bretherton, C.S, Smith, C., & Wallace, J.M. 1992. An Intercomparison of Methods for finding Coupled Patterns in Climate Data. *Journal of Climate*, **5**, 541–560.

Broecker, W.S. 1991. The great ocean conveyor. *Oceanography*, **4**, 79–89.

Cane, M. 1998. A Role for the Tropical Pacific. *Science*, **282**, 92–95.

Cane, R. Murtuguddeand M., & Prasad, V. 1995. *A Reduced Gravity, Primitive Equation, Isopycnal Ocean GCM: Formulation and Simulations.* Tech. rept. NASA/GSFC.

Crane, R.G., & Hewitson, B.C. 1998. Doubled CO_2 Precipitation Changes for the Susquehanna Basin: Downscaling from the Genesis General Circulation Model. *International Journal of Climatology*, **18**, 65–76.

Deser, C., & Blackmon, M.L. 1993. Surface climate variations over the North Atlantic ocean during winter: 1900-1989. *Journal of Climate*, **6**, 1743–1753.

Grötzner, A., Latif, M., & Barnett, T.P. 1998. A Decadal Climate Cycle in the North Atlantic Ocean as Simulated by the ECHO Coupled GCM. *Journal of Climate*, **11**, 831–847.

Hanssen-Bauer, I., & Nordli, P.Ø. 1998. *Annual and seasonal temperature variations in Norway 1896-1997.* Klima 25/98. DNMI.

Heyen, H., Zorita, E., & von Storch, H. 1996. Statistical downscaling of monthly mean North Atlantic air-pressure to lea level anomalies in thr Baltic Sea. *Tellus*, **48A**, 312–323.

Kaas, E., Christensen, O.B., & Christensen, J.H. 1998. Dynamical Versus Empirical Downscaling. *Tellus*, **submitted**.

Latif, M., Grötzner, A., M-Münich, Maier-Reimer, E., Venzke, S., & Barnett, T. P. 1996. A Mechanism for Decadal Climate Variability. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Lorenz, E. 1963. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20**, 130–141.

Lorenz, E. 1967. *The Nature and Theory of the General Circulation of the Atmosphere.* Publication 218. WMO.

Palmer, T.N. 1996. Predictability of the Atmosphere and Oceans: from Days to Decades. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Peixoto, J.P., & Oort, A.H. 1992. *Physics of Climate.* AIP.

Philander, S.G. 1989. *El Niño, La Niña, and the Southern Oscillation.* N.Y.: Academic Press.

Preisendorfer. 1988. *Principal Component Analysis in Meteorology and Oceanology.* Elsevier Science Press.

Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. 1989. *Numerical Recepies in Pascal.* Cambridge University Press.

Robinson, P.J., & Finkelstein, P.L. 1991. The development of impact-oriented scenarios. *Bull. Am. Met. Soc.*, **4**, 481–490.

Sarachik, E.S., Winton, M., & Yin, F.L. 1996. Mechanisms for Decadal-to-Centennial Climate Variability. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Steig, E.J., Brook, E.J., White, W.C., Sucher, C.M., Bender, M.L., Lehman, S.J., Morse, D.L., Waddington, E.D., & Clow, C.D. 1998. Synchronous Climate Change in Antartica and the North Atlantic. *Science*, **282**, 92–95.

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

Sutton, R.T., & Allen, M.R. 1997. Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, **388**, 563–567.

Treut, H. Le. 1994. Cloud Representation in Large-Scale Models: Can we Adequately Constrain them through Observed Data? *In:* Speranza, A., & nd R. Fantechi, S.Tibaldi (eds), *Global Change.* Environment and Quality of Life, vol. EUR 15158 EN. European Commission.

von Storch, H., Zorita, E., & Cubasch, U. 1993. Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime. *Journal of Climate*, **6**, 1161–1171.

Wilby, R.L. 1997. Non-stationarity in daily precipitation series: Implications for GCM down-scaling using atmospheric circulation indices. *International Journal of Climatology*, **17**, 439–454.

Wilby, R.L., Hassan, H., & Hanaki, Keisuke. 1998. Statistical downscaling of hydrometeorological variables using general circulation model output. *Journal of Hydrology*, **205**, 1–19.

Wilks, D.S. 1995. *Statistical Methods in the Atmospheric Sciences.* Orlando, Florida, USA: Academic Press.

Zorita, E., & von Storch, H. 1997. *A survey of statistical downscaling results.* Tech. rept. 97/E/20. GKSS.