D12 supplementary analysis:
Downscaling Methods for the European Region

Torben Schmith, Danish Meteorological Institute
Lucia Benito, Fundación para la Investigación del Clima
Bo Christiansen, Danish Meteorological Institute
Malcolm Haylock, University of East Anglia
Jaime Ribalaygua, Fundación para la Investigación del Clima

1. Introduction

We report on an intercomparison of methods carried out for the whole European region, and supplementing the analysis reported in the D12 main document.

Three different statistical methods for downscaling extremes are applied to the NCEP/NCAR reanalysis data for the period 1958-2000. The intercomparison is done along the lines defined in the D12 main document

2. Data

Predictand data consists of all 481 stations contained in the 'FIC dataset'.

2. Methodology

Table 2: STARDEX Diagnostic Extreme Indices analysed in the study

Designation	Description
	a) Precipitation related indices
pav	Precipitation average (mm/day)
pq90	90th percentile of rainday amounts (mm/day)
px5d	Greatest 5-day total rainfall
pint	Simple Daily Intensity (rain per rainday)
pxcdd	Max no. of consecutive dry days
pfl90	% of total rainfall from events > long-term 90th percentile
pnl90	No. of events > long-term 90th percentile of raindays
	b) Temperature related indices
txav	Average Tmax (deg. C)

tnav	Average Tmin (deg. C)
txq90	Tmax 90th percentile (deg. C)
tnq10	Tmin 10th percentile (deg. C)
tnfd	Number of frost days Tmin < 0 °C
txhw90	Heat wave Duration (days)

3. Description of downscaling methods

UEA

Canonical Correlation analysis (CCA)

A downscaling method is being developed to be applied to the entire European regional dataset which models the STARDEX indices of rainfall extremes using large scale patterns of circulation. Rather than modelling the daily rainfall itself, then calculating the indices, this method will use seasonal measures of large-scale circulation variability to model the seasonal indices of extremes directly. The predictands will be the six STARDEX core indices of rainfall extremes calculated seasonally. Predictors will be selected from circulation variables, calculated over the entire European and East Atlantic region. From our work preparing D10, the best predictor seems to be sea level pressure but other variables will be considered including temperature, geopotential height and relative humidity at three atmospheric levels as well as sea surface temperature. The model uses canonical correlation analysis. Two methods to select predictors will be tested: one using just the variable with the best correlation with the indices (MSLP); and a cross-validation of all possible predictor combinations.

FIC

The Two step analogue method (2SA)

This method estimates daily precipitation amount and daily maximum and minimum temperature for a site for day "X" in two steps:

- The "n" most similar (analogous) days (according to the geostrophic fluxes at 1000 and 500 hPa) to the day "X" are selected from a reference dataset
- In a second step, precipitation and temperatures are obtained applying further analyses that searches (only in the "n" days population) for relationships between the predictands and some more predictors. For precipitation, averaging the 100 most similar days
(analogous) precipitation. For temperature, a multiple linear regression (only in the "n" days population) with forward and backward stepwise selection of predictors, using as potential predictors low troposphere thickness, the averaged temperatures of the previous days, and a sinusoid function of the day of the year.

DMI

Conditional weather generator (CWG)

We apply a weather generator approach based on daily values of an index describing the large scale circulation on which the probalistic characteristics of the precipitation are quantified. The relation between the circulation and the precipitation characteristics is calculated for each station and for each season.

For calibrating the model, a surface pressure pattern is obtained as the average pressure difference between rainy days and dry days measured at a given station. The circulation index is then calculated by regressing the daily surface pressure field on this pattern. We then divide the circulation index into a number of quantiles, usually between 5 and 10, and for each quantile the following precipitation characteristics are calculated: the probability for wet/dry days, the probabilities for a wet/dry day following a dry/wet day, and the two parameters describing the gamma-distribution that best approximates the probability density of the rain amount (only wet days).

When applying the model, the daily circulation index is calculated by regressing the daily surface pressure field on the pattern found above. Using the dependence of the probabilities on the circulation index a two-state Markov process is used to obtain the sequence of dry/wet days. Then, for each wet day the rain amount is drawn from a gamma-distribution with the parameters corresponding to the circulation index of that day.

5. Results of verifications and discussion

In this section, we will refer to figures, which can be found in a separate file showing the geographical distribution of the correlation coefficient as colour-coded circles. An open circle represents a non-significant (on the 95% level) while a close circle represents a significant result.

5.1 Precipitation

The index ‘pav’ is where the methods generally perform best, as would also be expected a priori. The FIC method I doing well both in both summer and winter (best for winter). Both The UEA and DMI methods have problems in winter over Russia and in summer over Russia and Southern Europe.

For the index ‘pint’ the performance is decreasing. Still there is a tendency for better skill during winter and overall, the FIC method is performing best. But even with this method and for winter, many stations have non-significant skill.

For the ‘pq90’ FIC also appears to have the overall best skill compared to the two other methods. For the ‘px5d’, ‘pcdd’, ‘pfl90’ and ‘pnl90’ indices similar conlusions can be drawn. But also for all these indices, many stations have non-significant skill.

5.2 Temperature

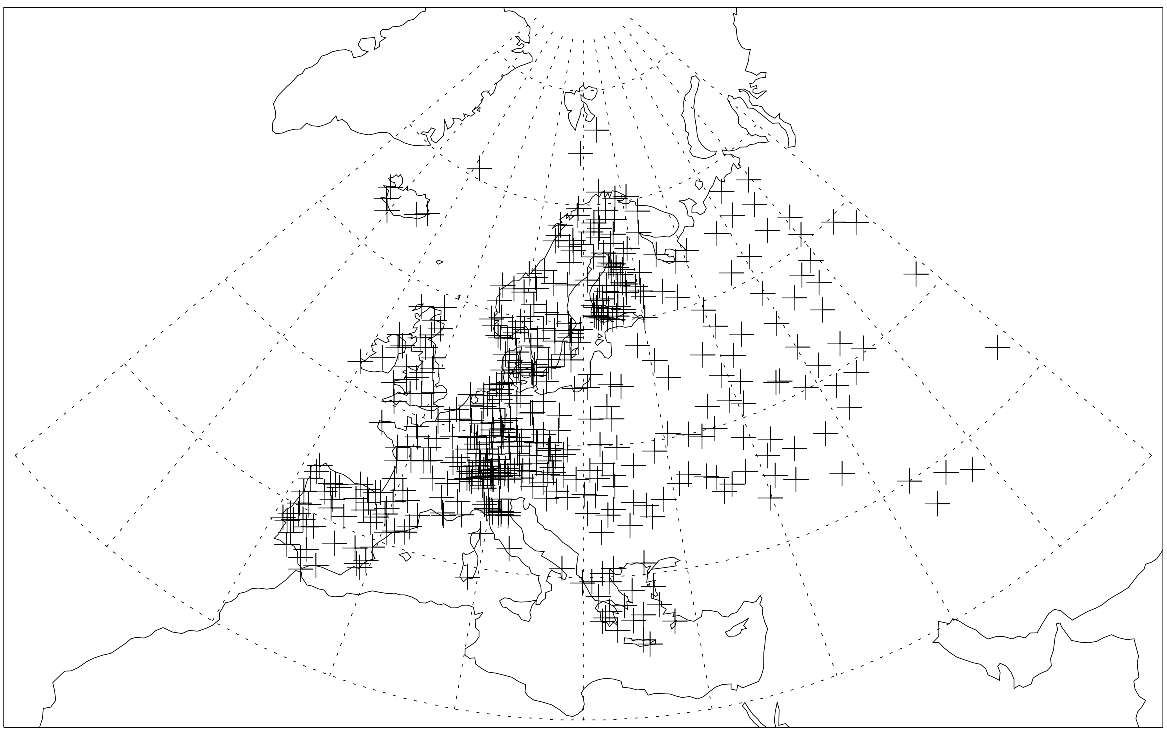
For temperature only the UEA and the FIC methods have been applied. Similarly than for precipitation, the FIC model appears to be superior to the UEA model.

6. Concluding remarks

From this analysis we must conclude that the ‘type 2’ (or indirect) method with many predictors performs best. The ‘type 1’ method of UEA does not use information at the daily level leading to inferior results. The DMI method uses a ‘type 2’ approach but with only one predictor. This appears to be a too simplistic approach as well.

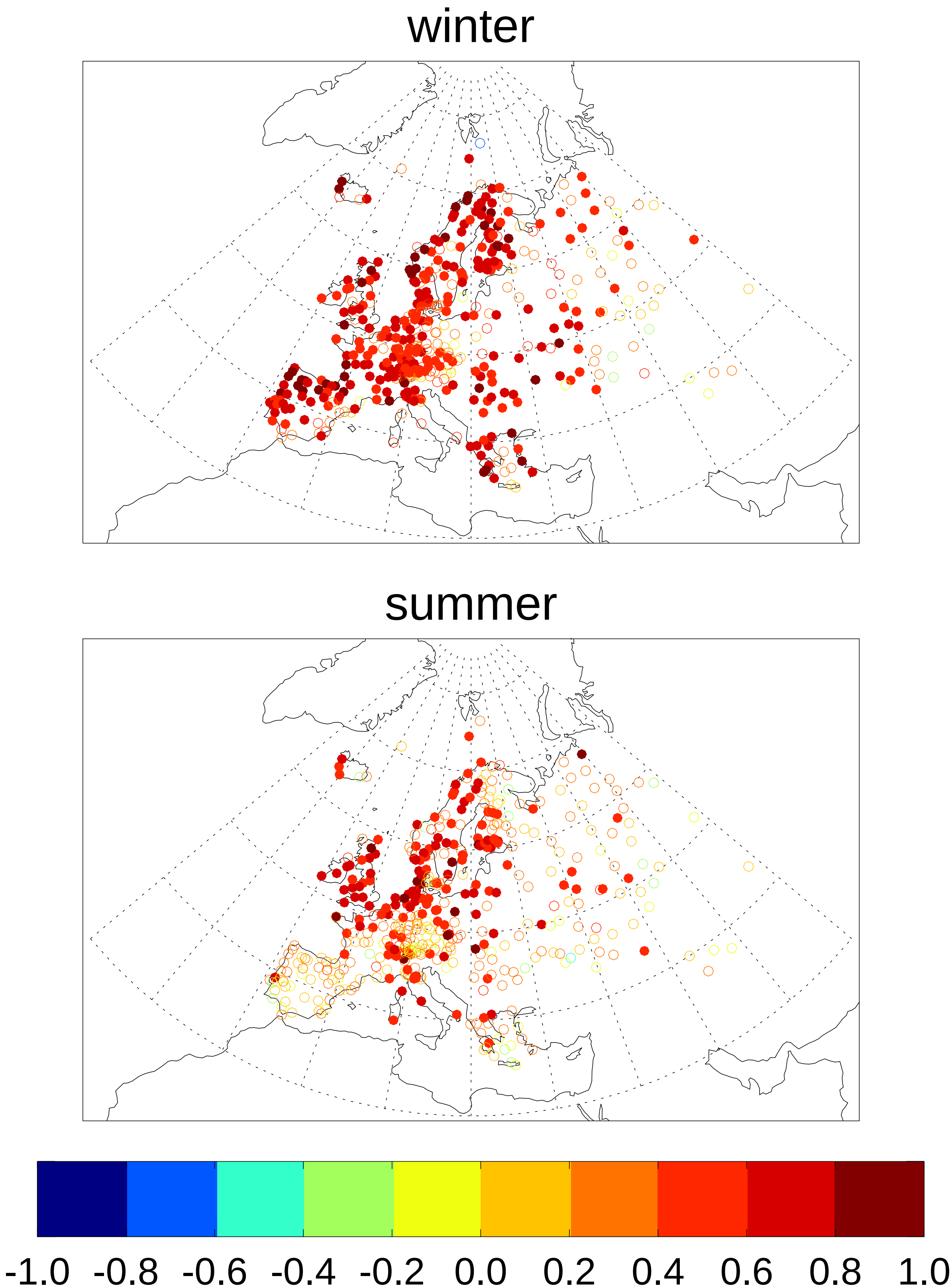
But many stations have non-significant skill even with the best method. We therefore must conclude, that on station-scale there is generally a poor downscaling potential using these methods. This could possibly be improved by considering area-aggregated precipitation measures.

FIC dataset
Station map
Period: 1958 - 2000

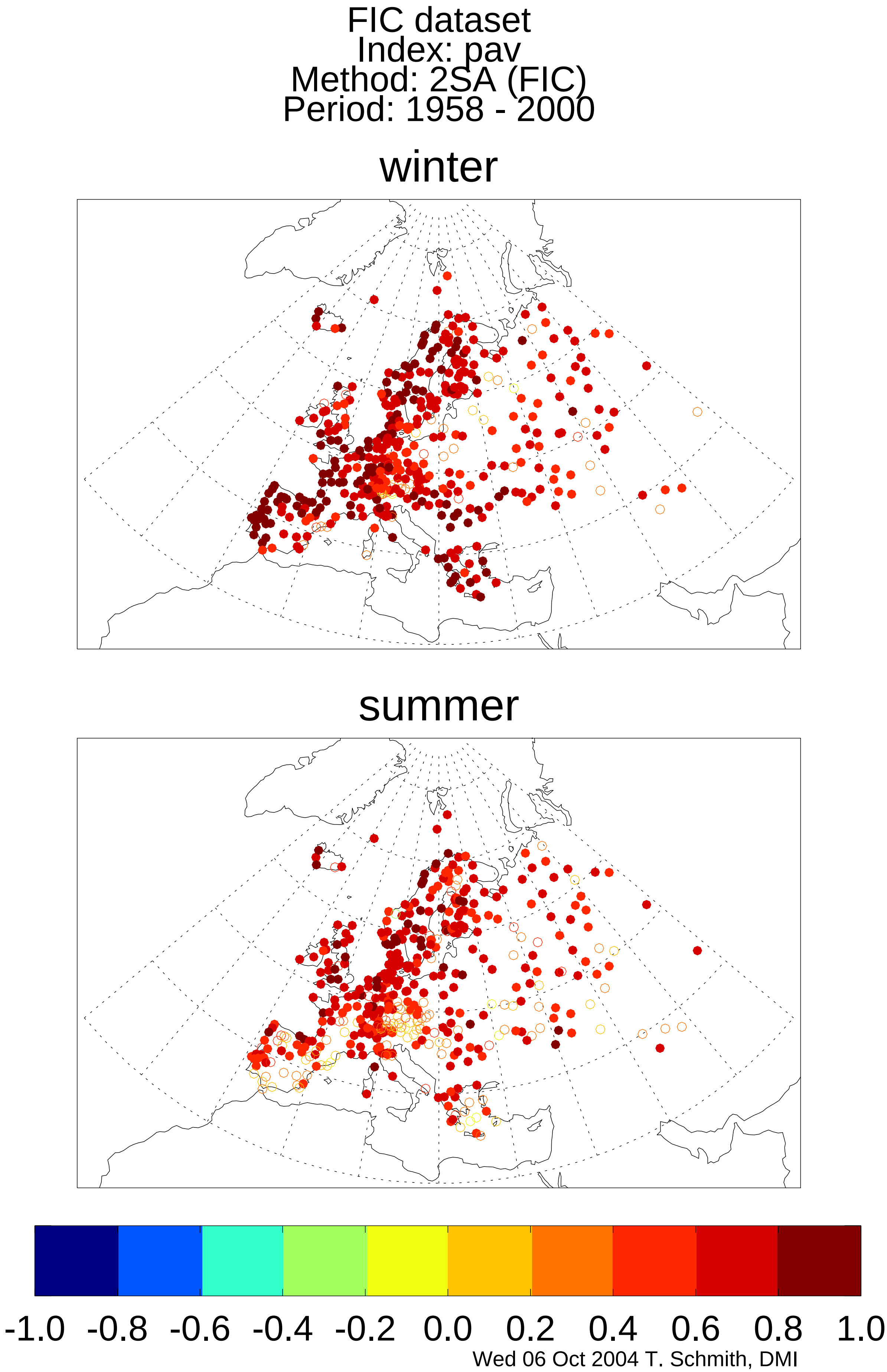


Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pav
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

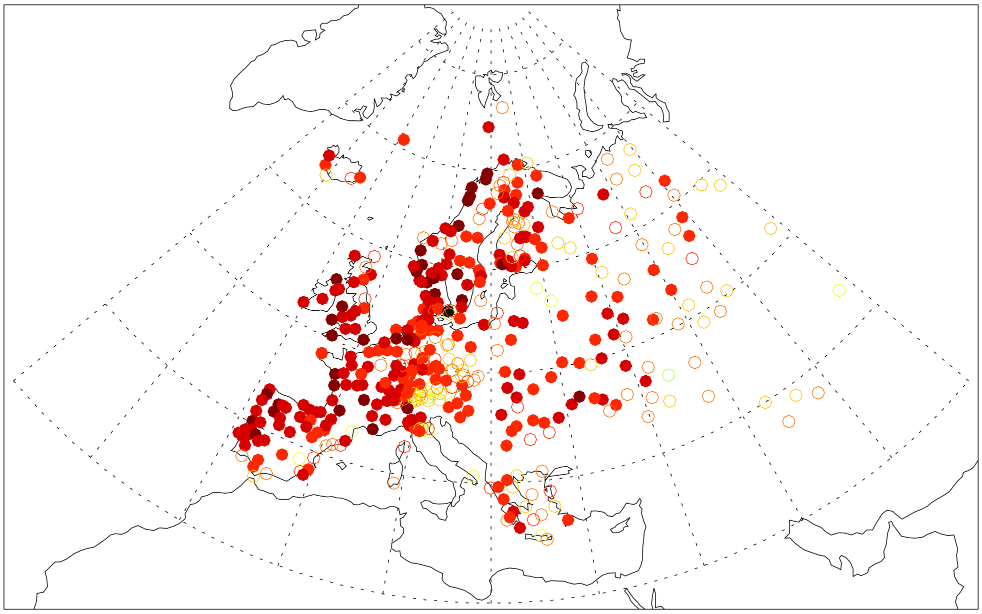


Wed 06 Oct 2004 T. Schmith, DMI

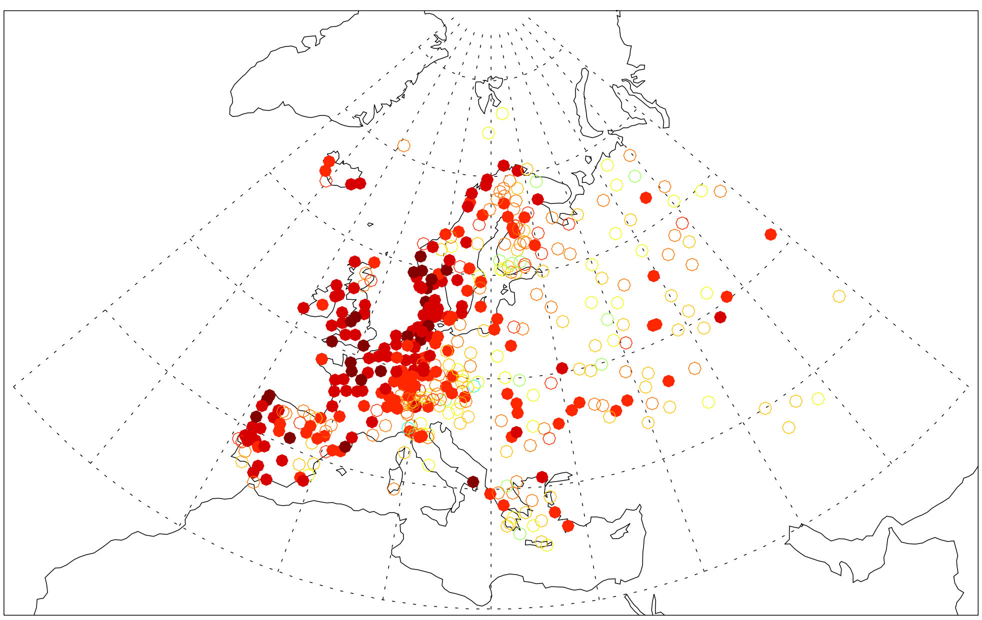

FIC dataset
Index: pav
Method: 2SA (FIC)
Period: 1958 - 2000
winter
summer
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
Wed 06 Oct 2004 T. Schmith, DMI

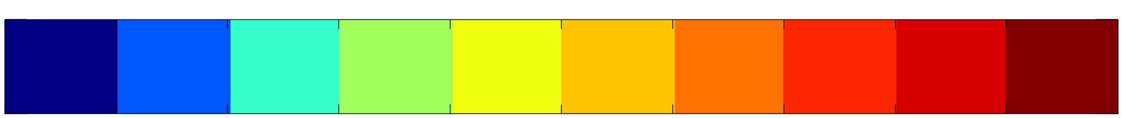
FIC dataset
Index: pav
Method: CWG (DMI)
Period: 1958 - 2000

winter



summer

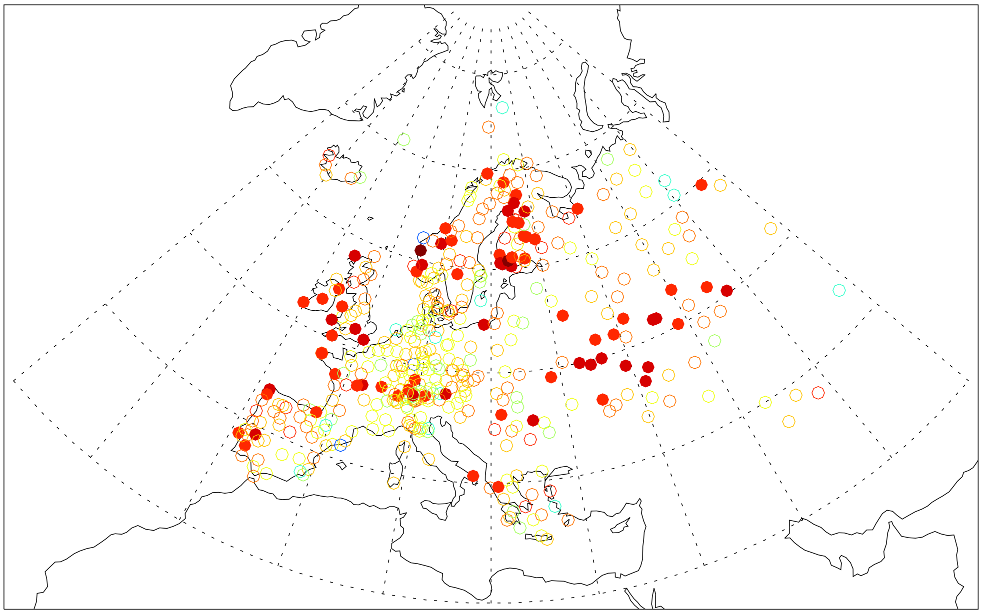




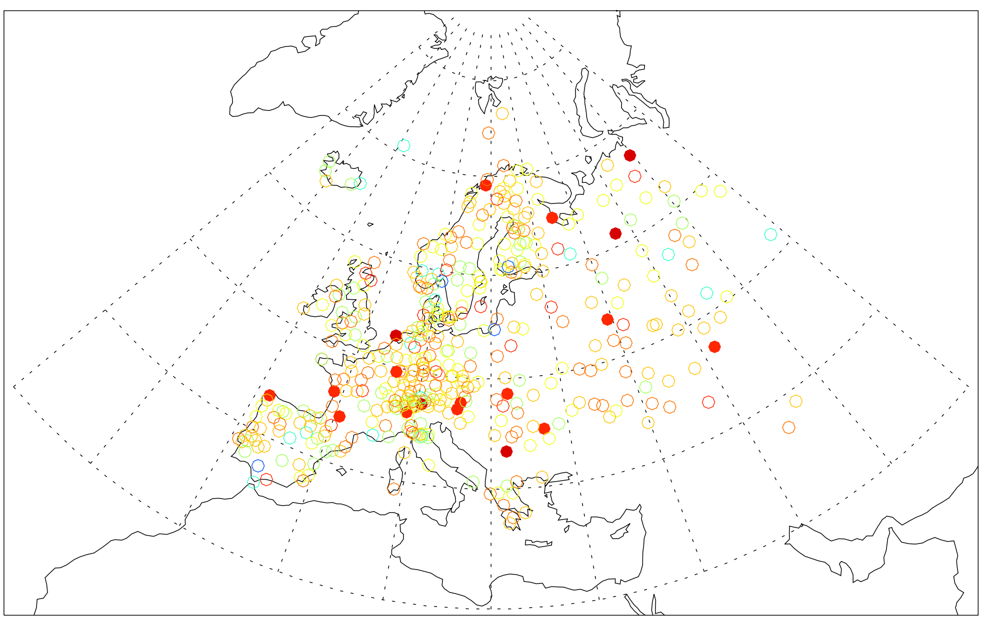
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pint
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

winter

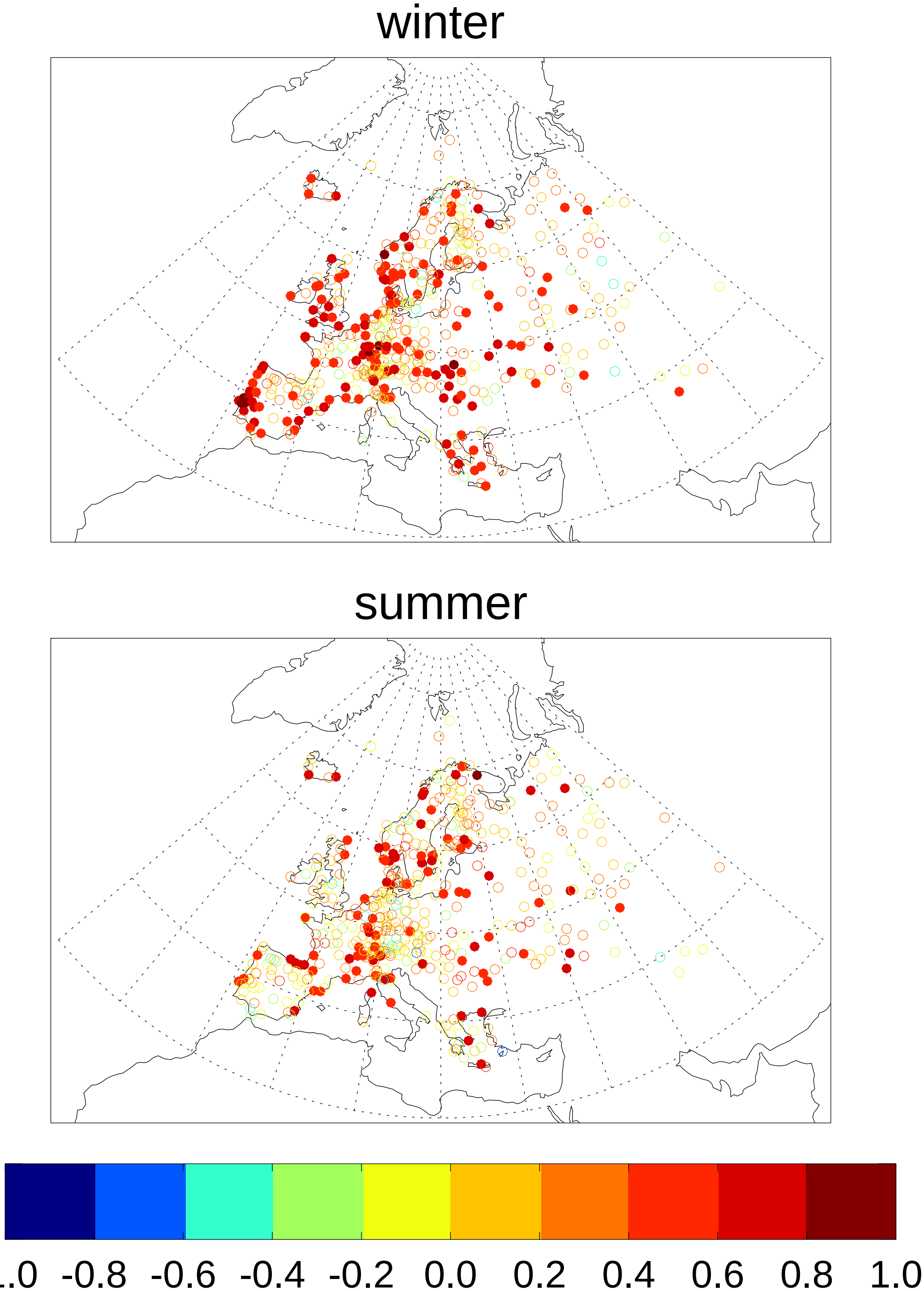


summer



Wed 06 Oct 2004 T. Schmith, DMI

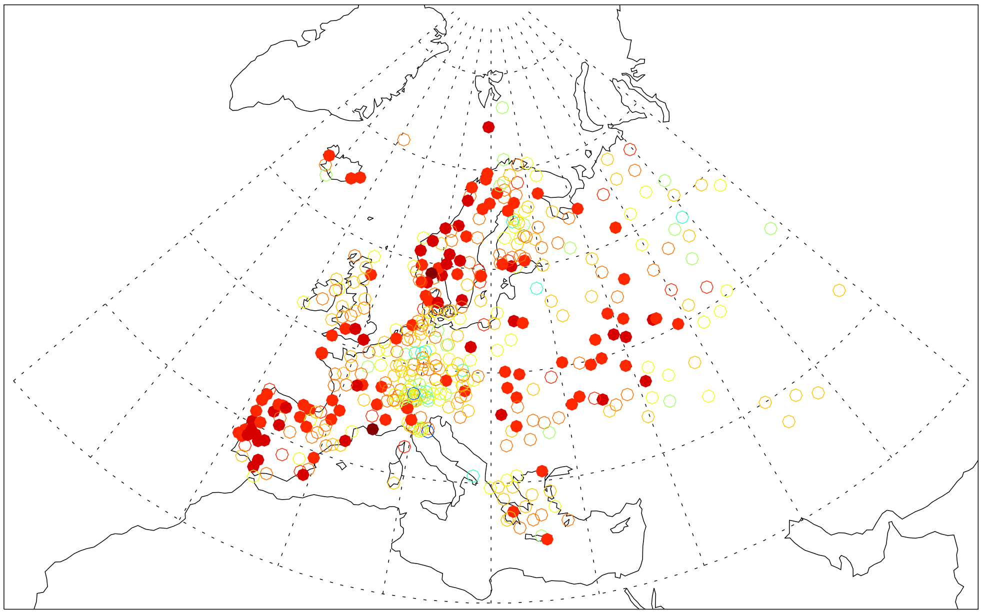
FIC dataset
Index: pint
Method: 2SA (FIC)
Period: 1958 - 2000



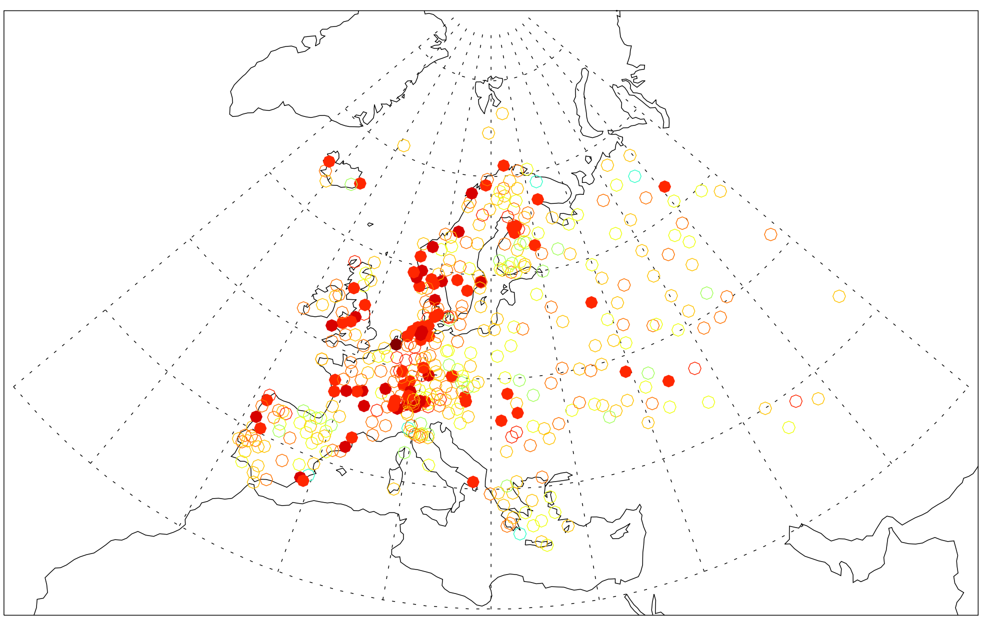
Wed 06 Oct 2004 T. Schmith, DMI

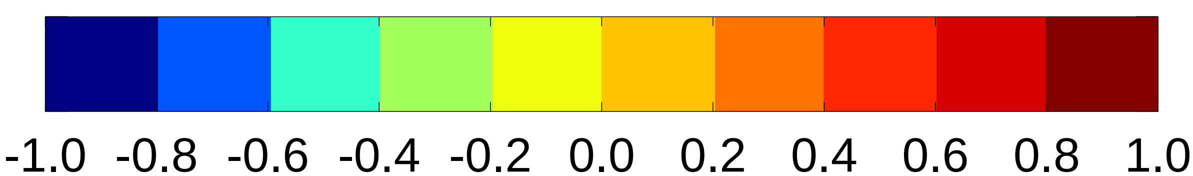
FIC dataset
Index: pint
Method: CWG (DMI)
Period: 1958 - 2000

winter



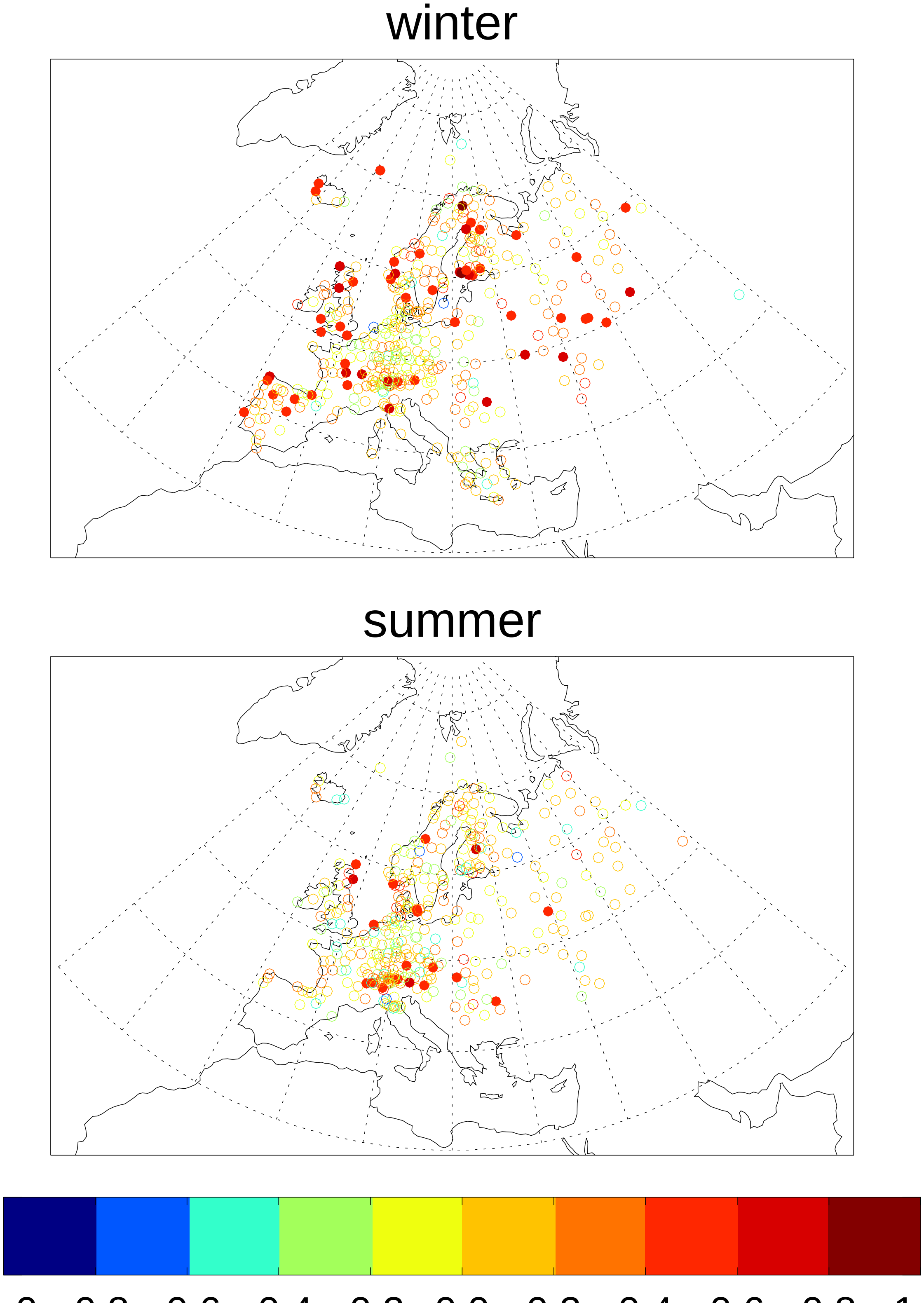
summer





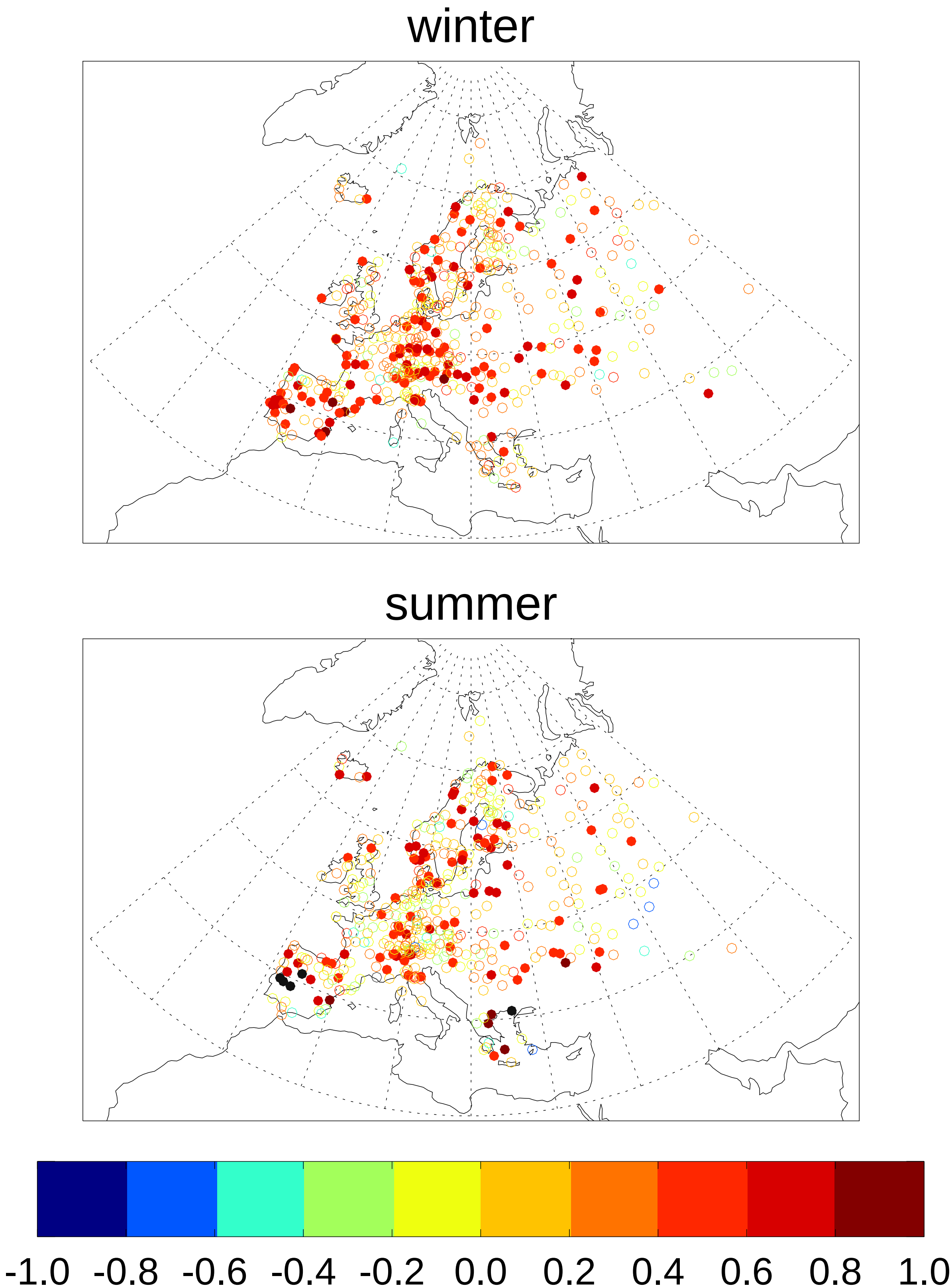
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pq90
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000



Wed 06 Oct 2004 T. Schmith, DMI

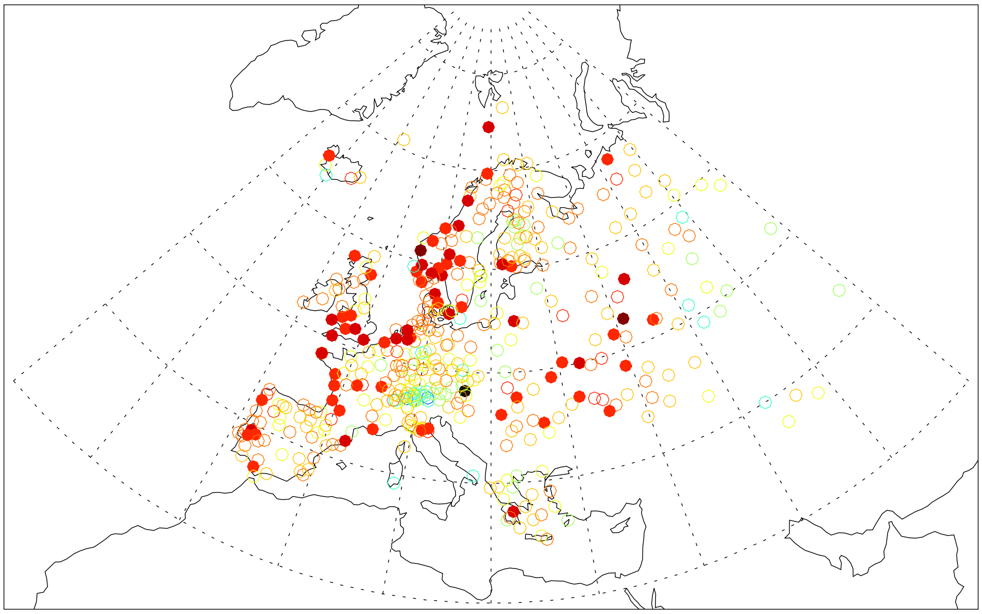
FIC dataset
Index: pq90
Method: 2SA (FIC)
Period: 1958 - 2000



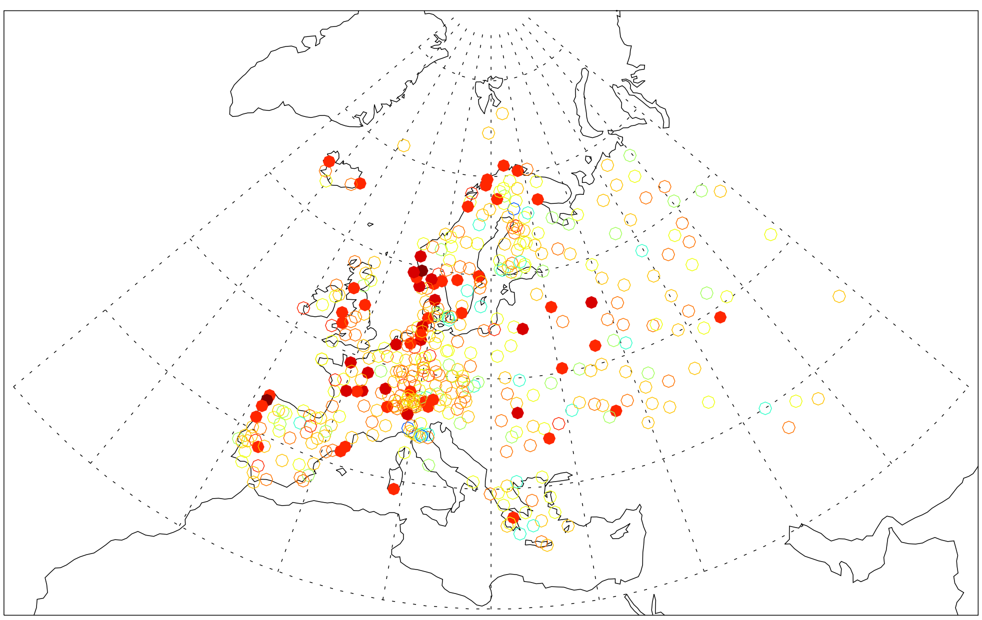
Wed 06 Oct 2004 T. Schmith, DMI

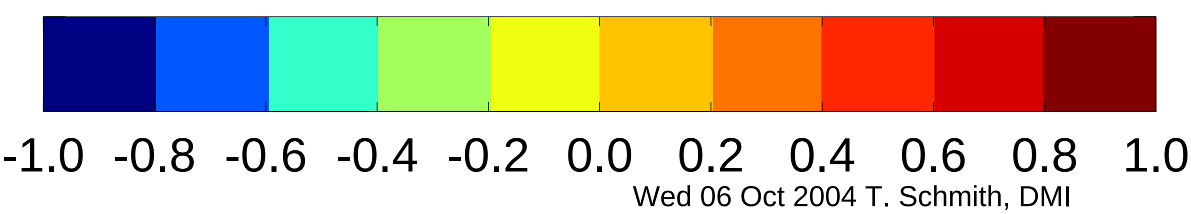
FIC dataset
Index: pq90
Method: CWG (DMI)
Period: 1958 - 2000

winter



summer

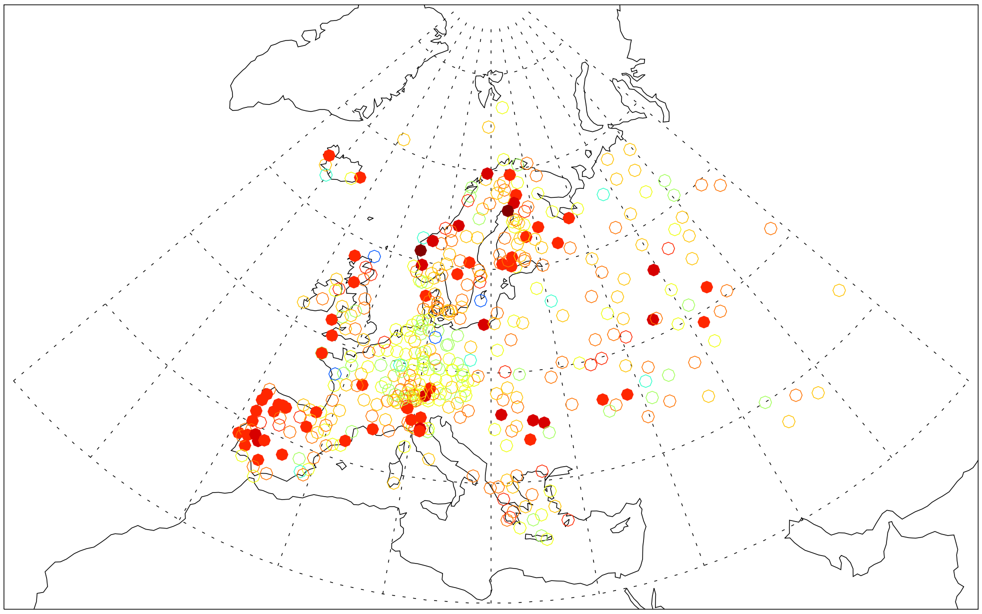




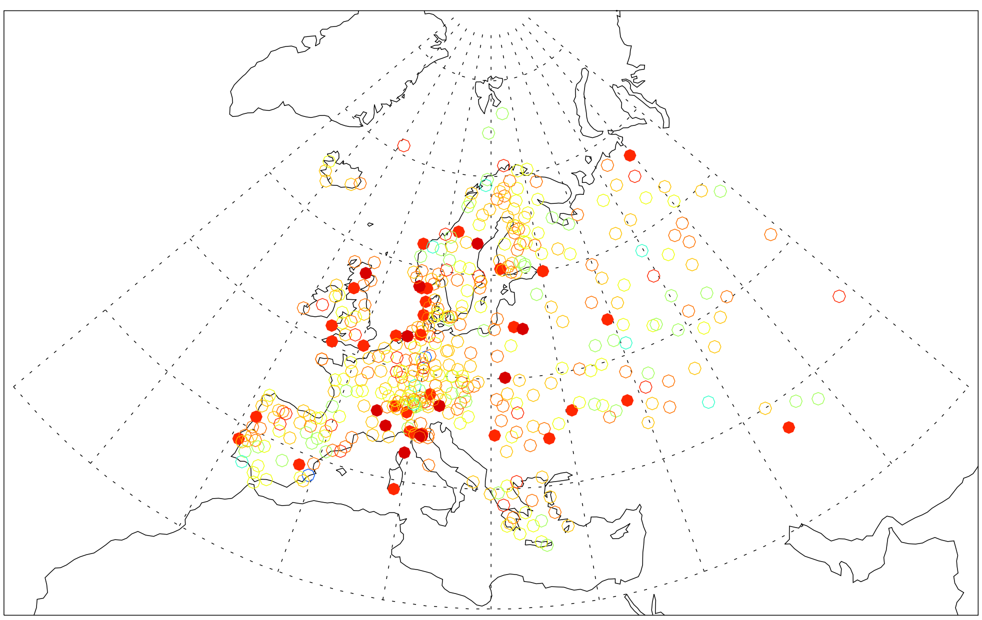
Wed 06 Oct 2004 T. Schmith, DMI

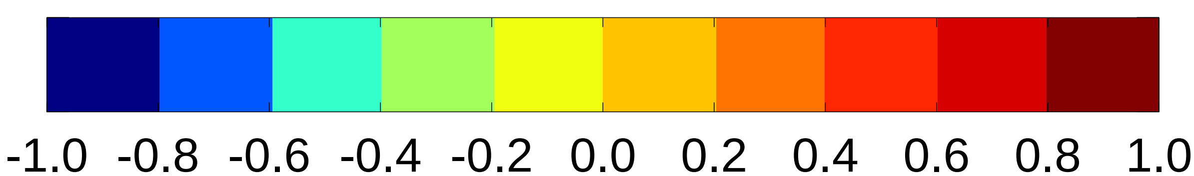
FIC dataset
Index: px5d
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

winter



summer

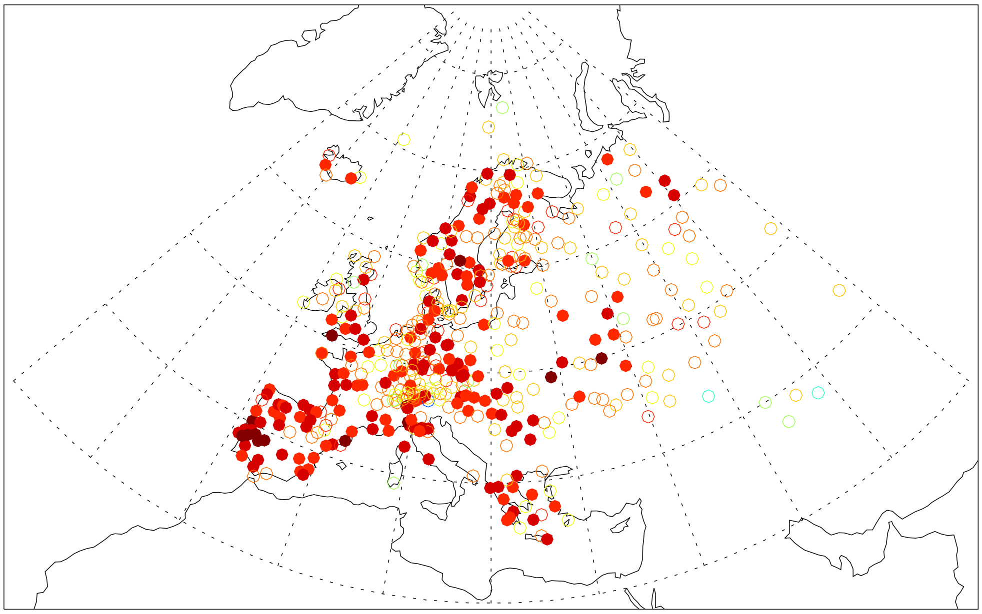




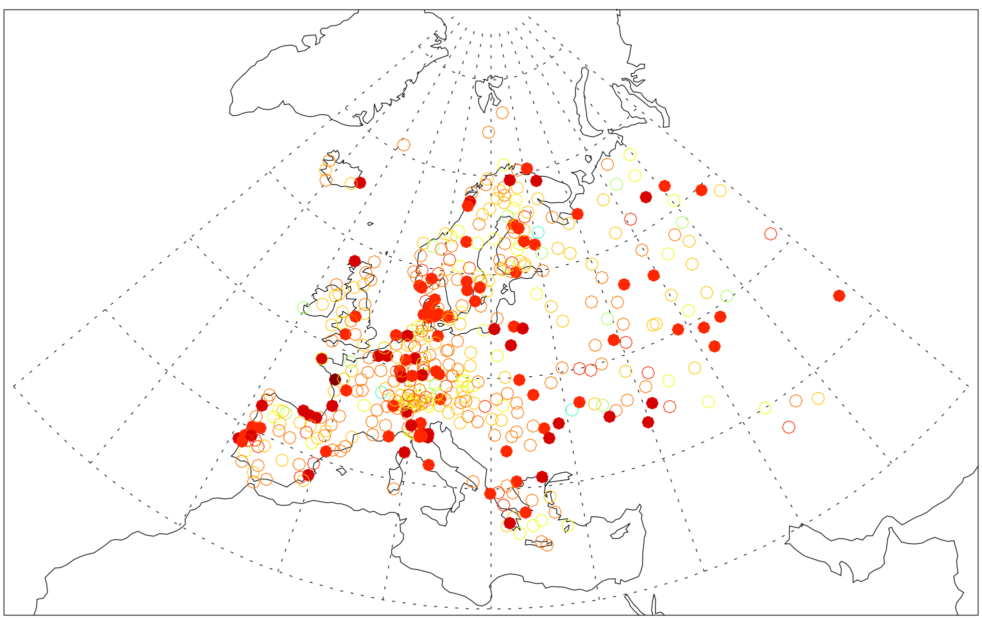
Wed 06 Oct 2004 T. Schmith, DMI

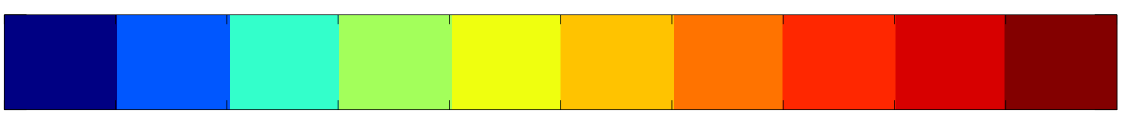
FIC dataset
Index: px5d
Method: 2SA (FIC)
Period: 1958 - 2000

winter



summer

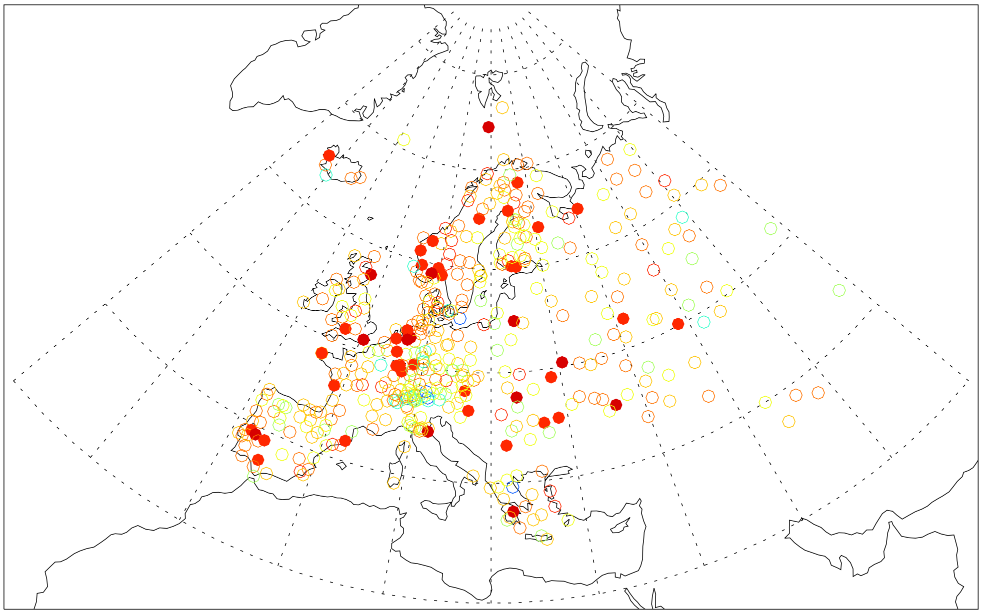




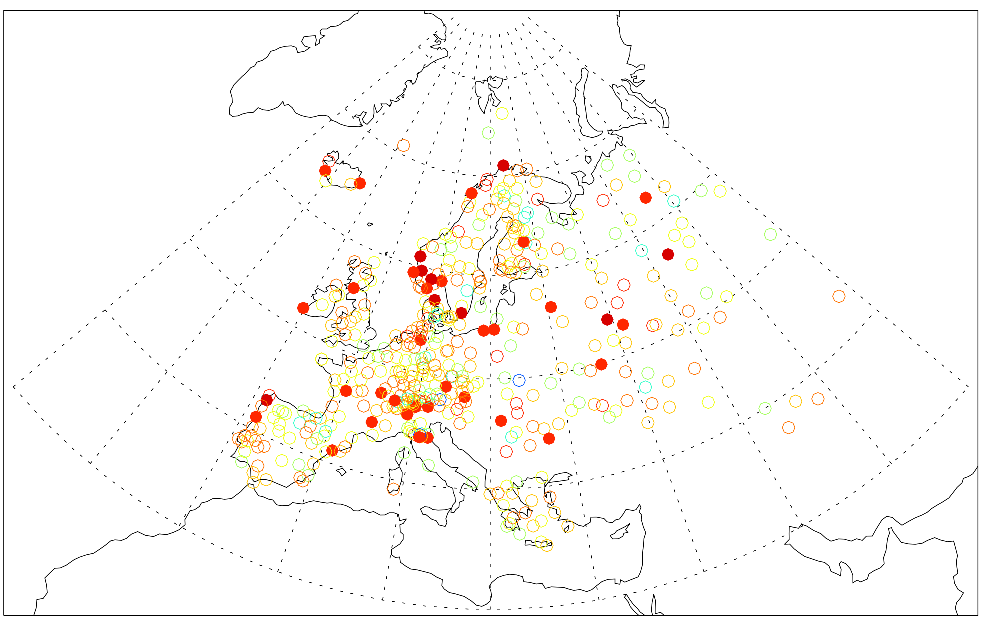
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: px5d
Method: CWG (DMI)
Period: 1958 - 2000

winter

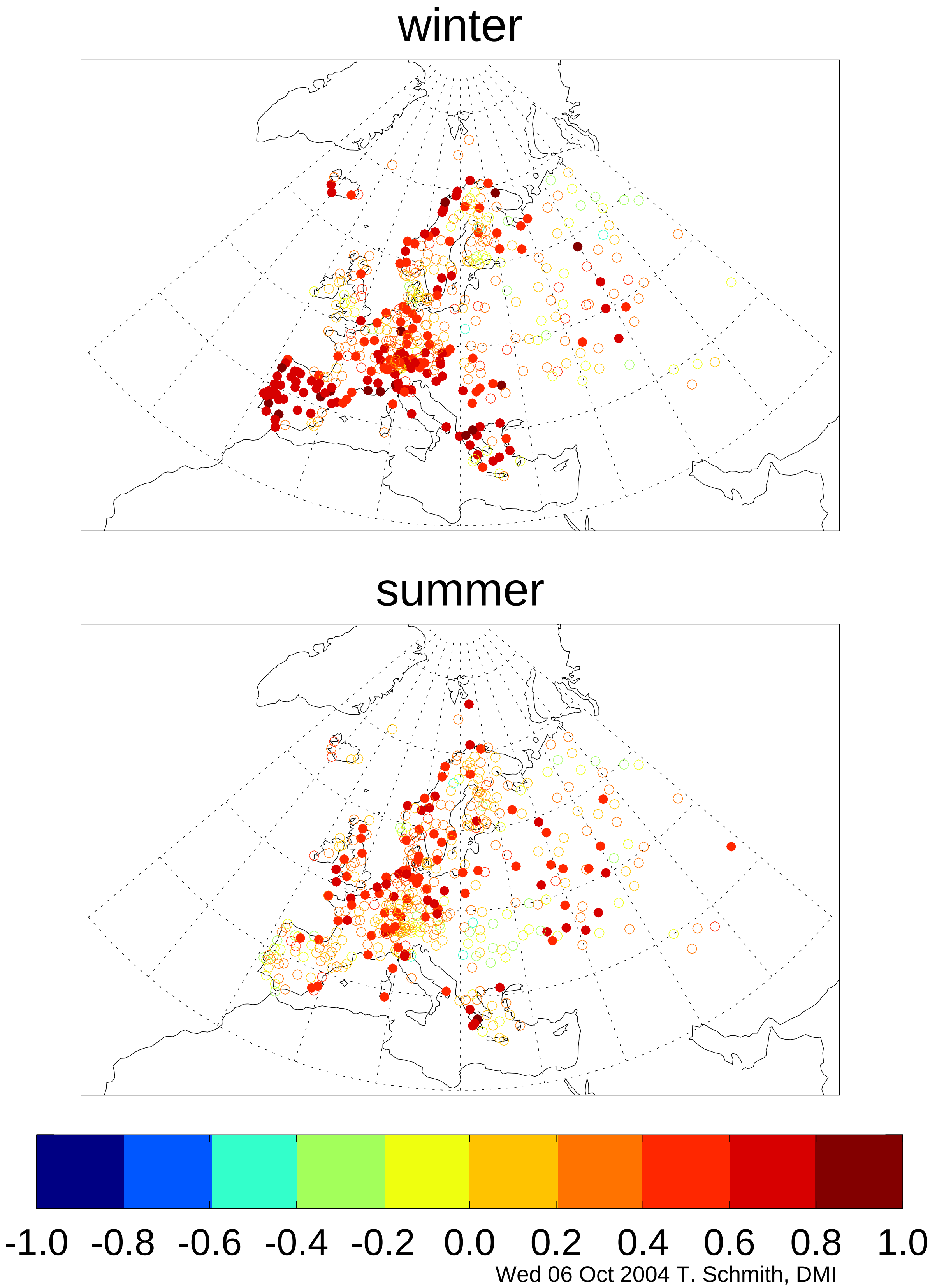


summer

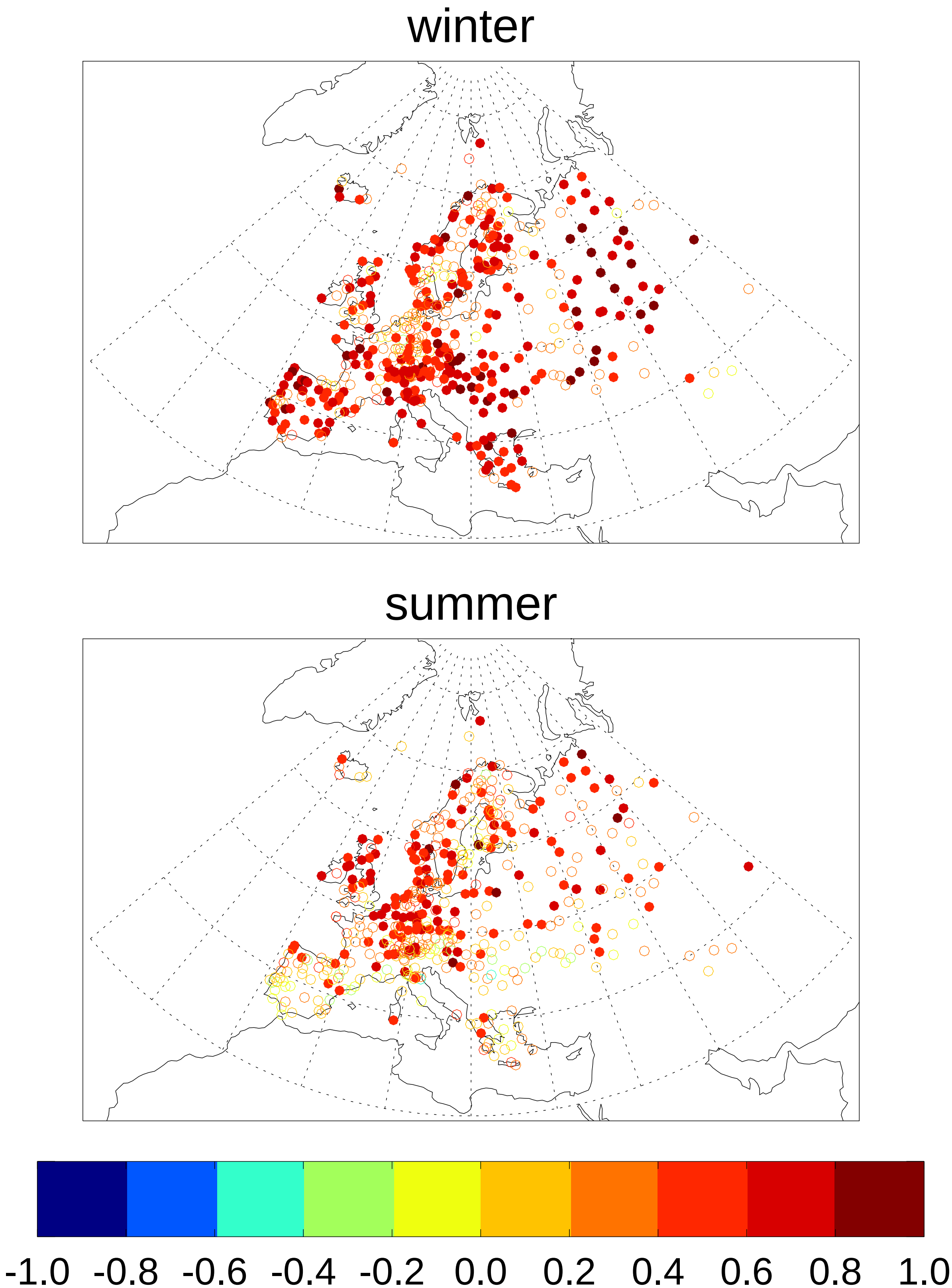


Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pxcdd
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

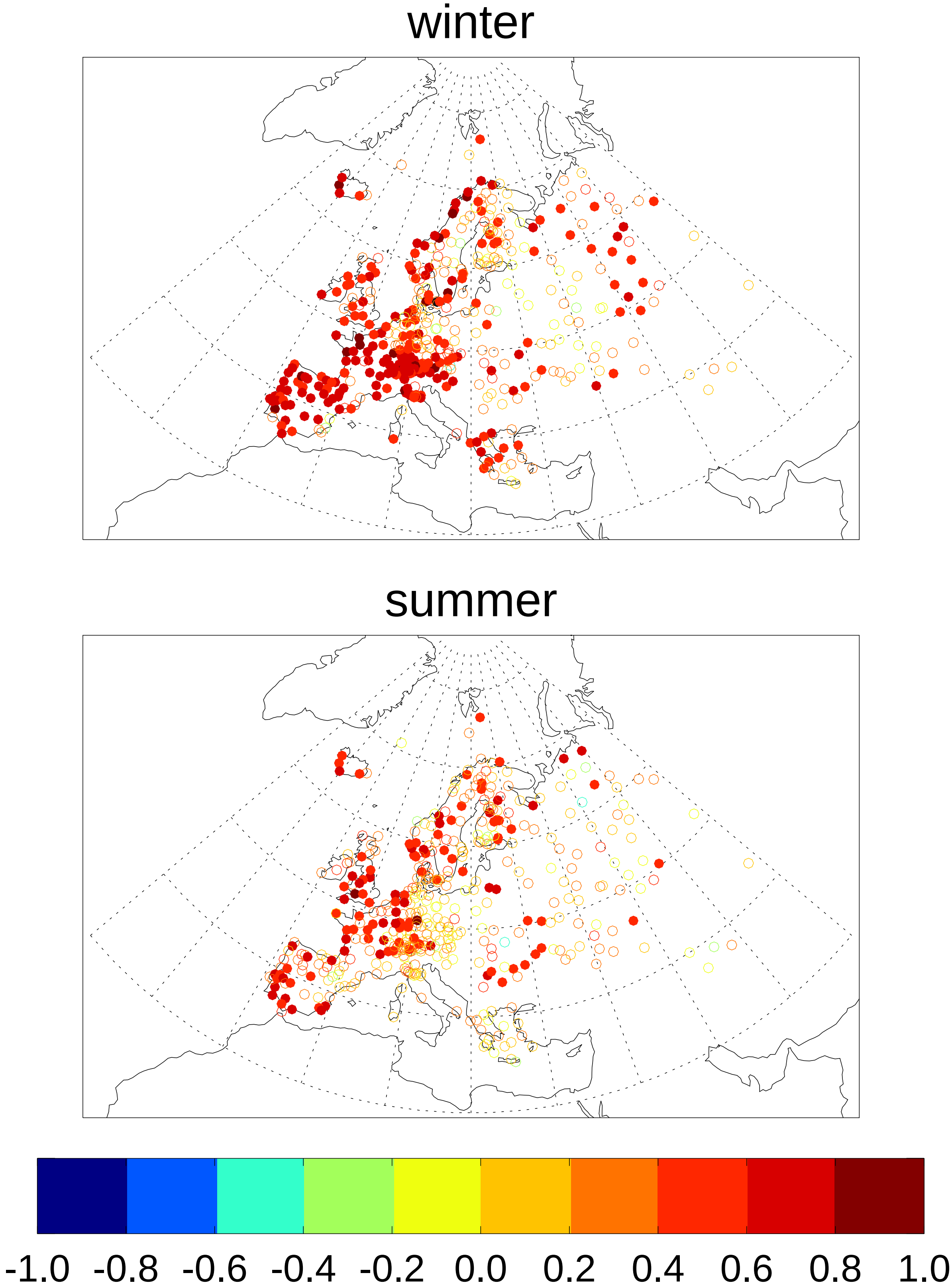


FIC dataset
Index: pxcdd
Method: 2SA (FIC)
Period: 1958 - 2000



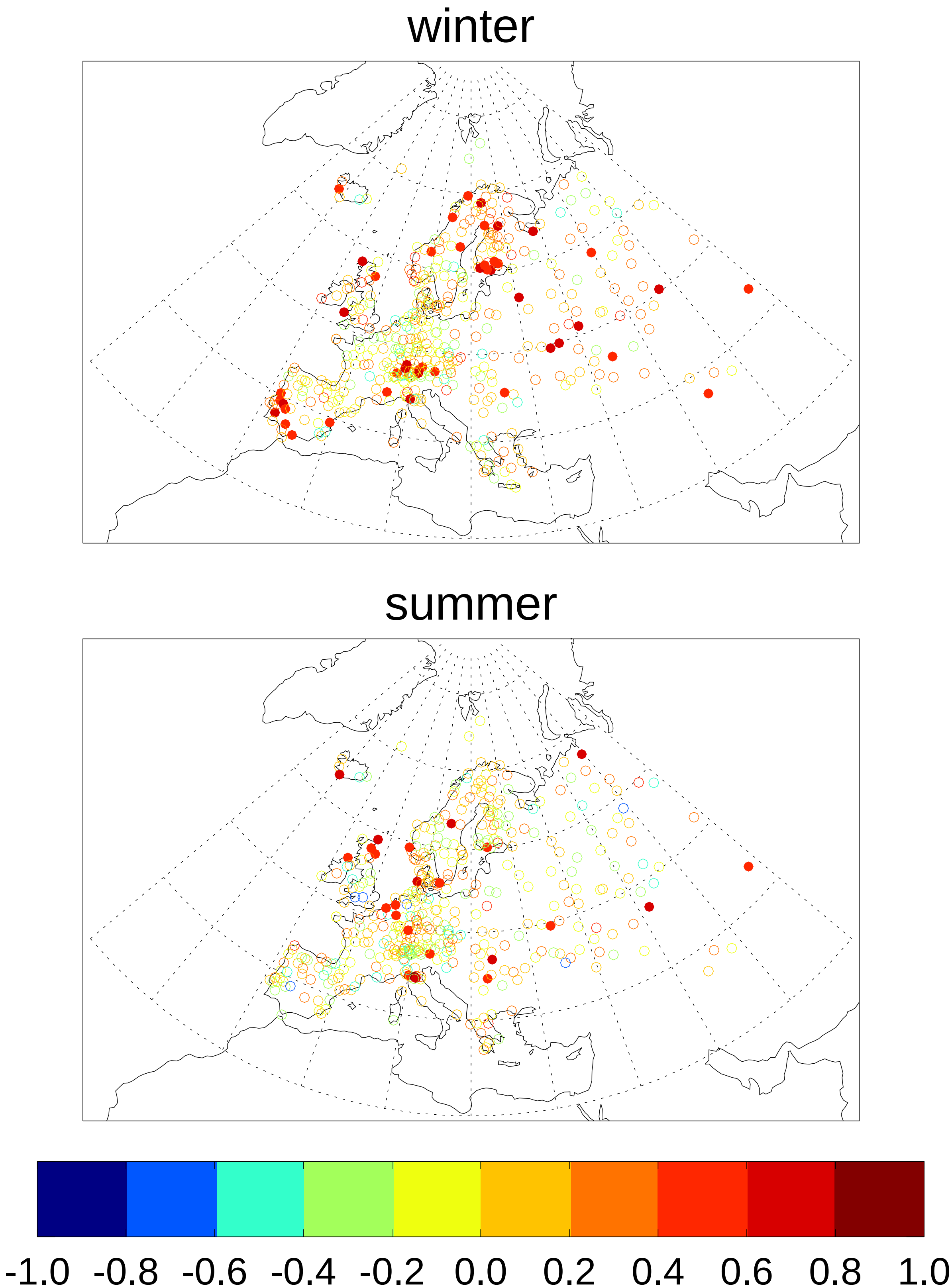
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pxcdd
Method: CWG (DMI)
Period: 1958 - 2000



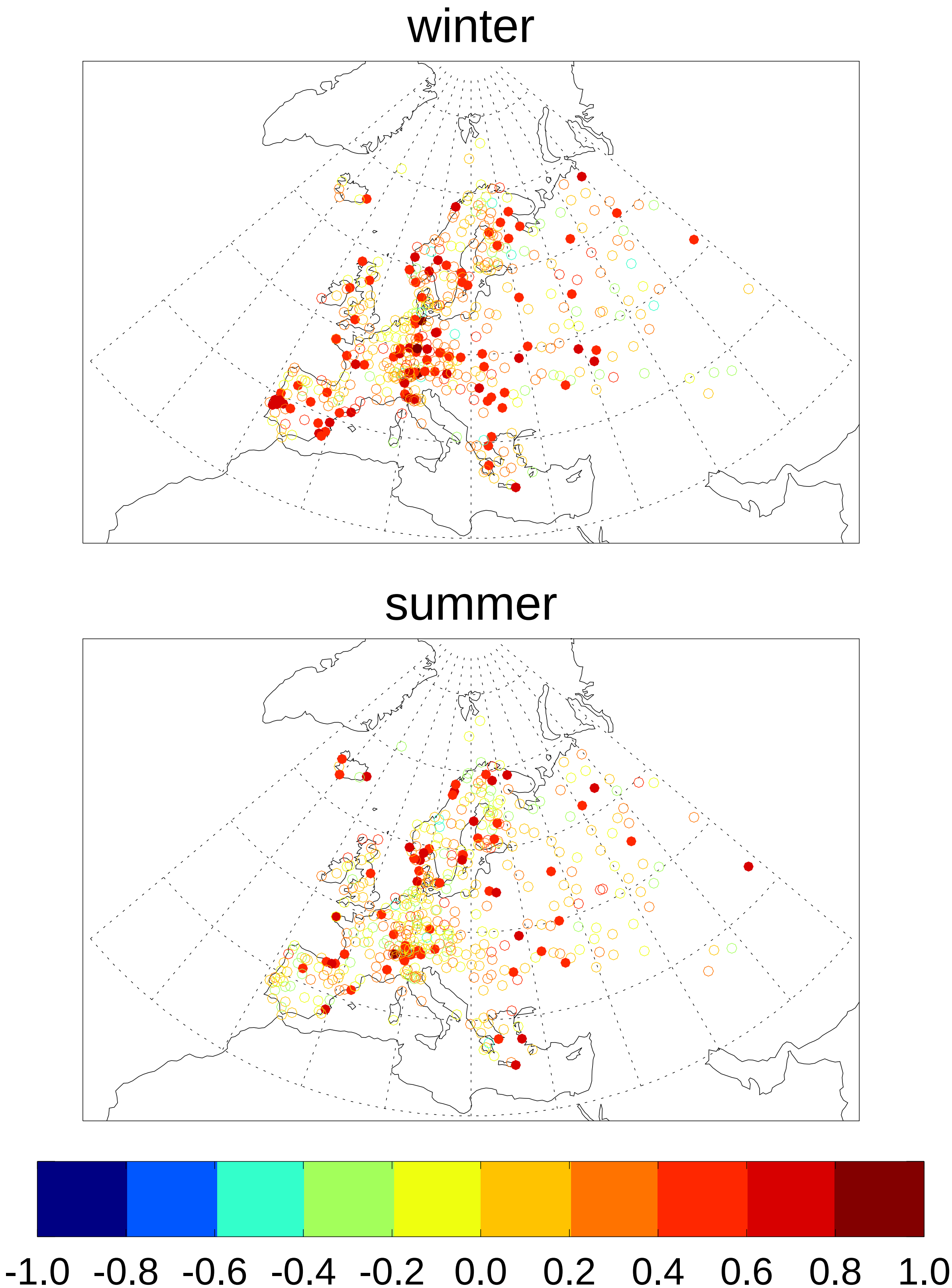
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pfl90
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000



Wed 06 Oct 2004 T. Schmith, DMI

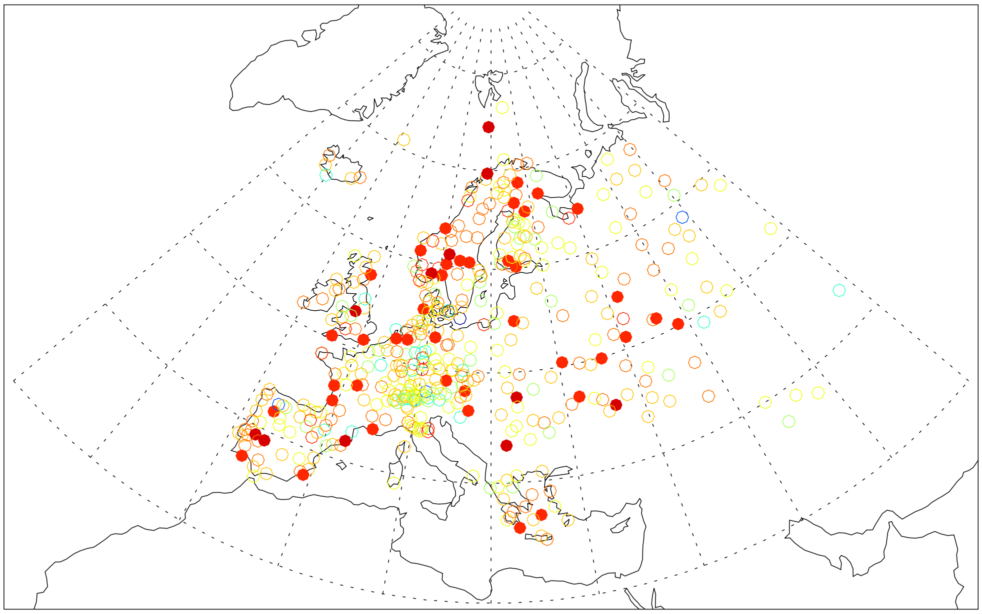
FIC dataset
Index: pfl90
Method: 2SA (FIC)
Period: 1958 - 2000



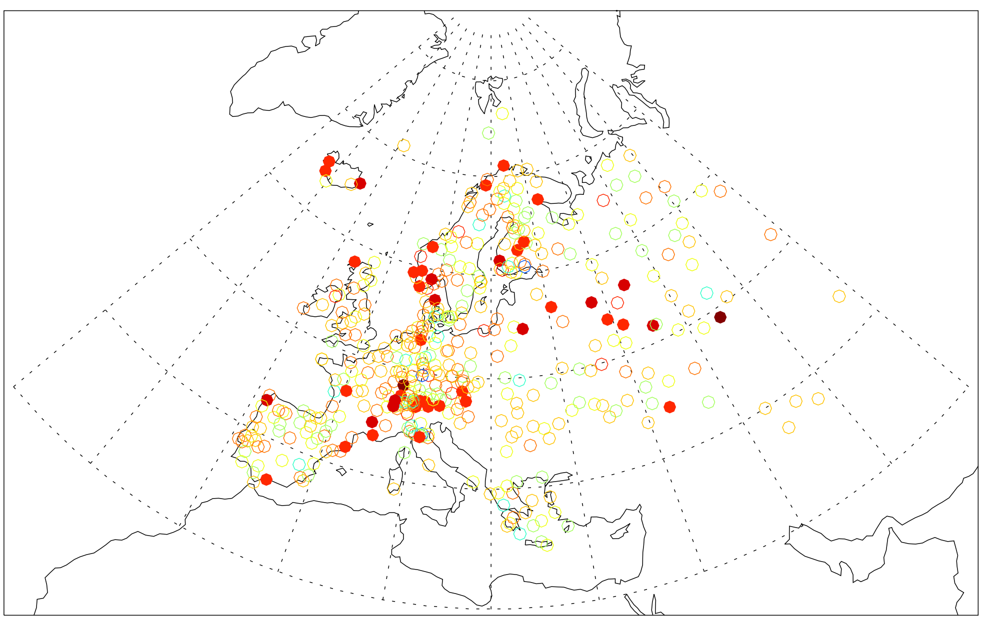
Wed 06 Oct 2004 T. Schmith, DMI

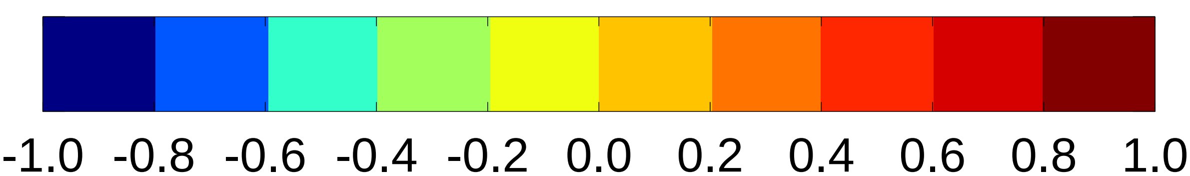
FIC dataset
Index: pfl90
Method: CWG (DMI)
Period: 1958 - 2000

winter



summer

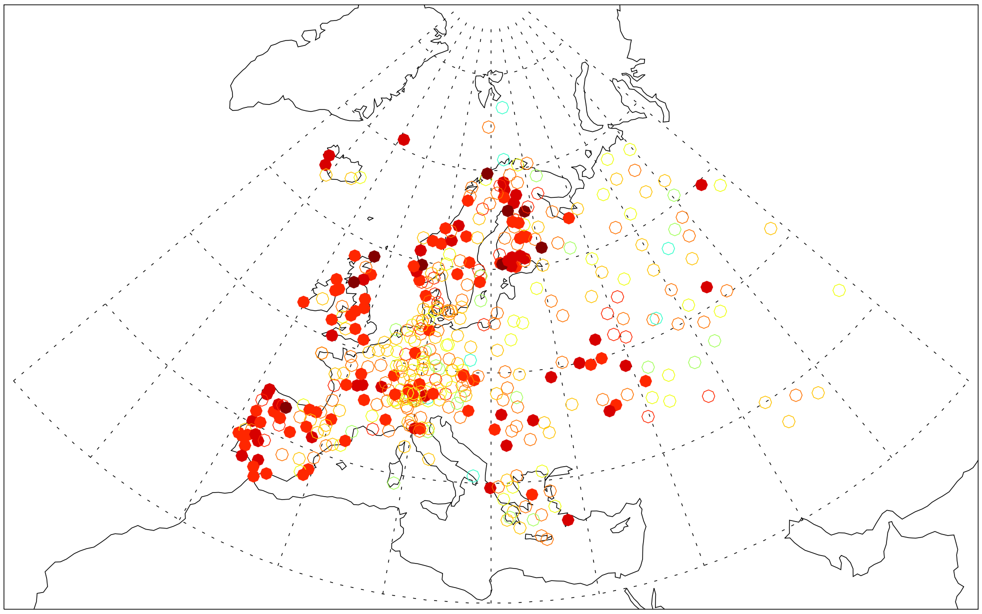




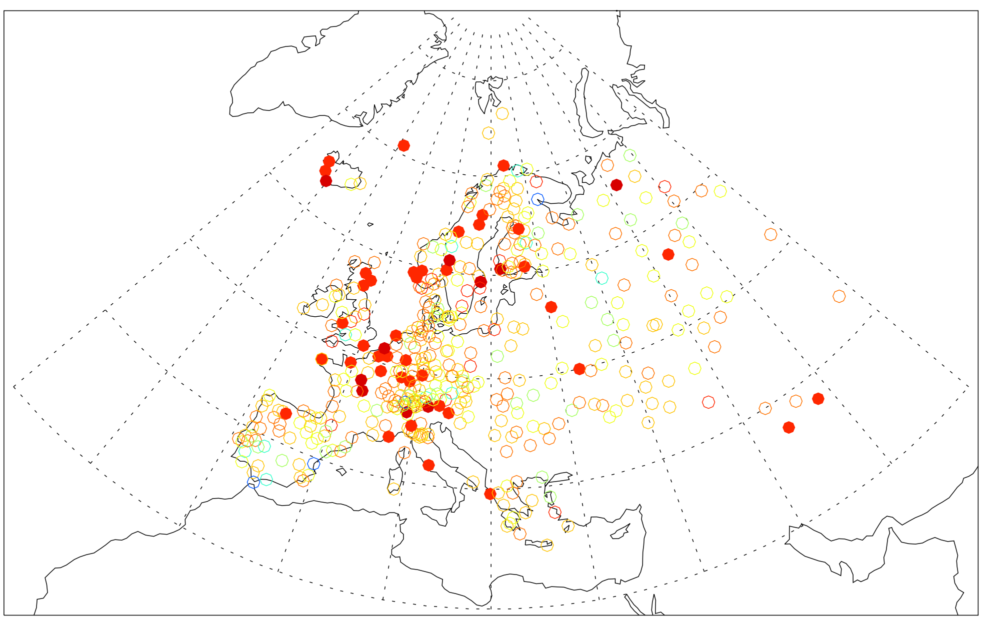
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pnl90
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

winter

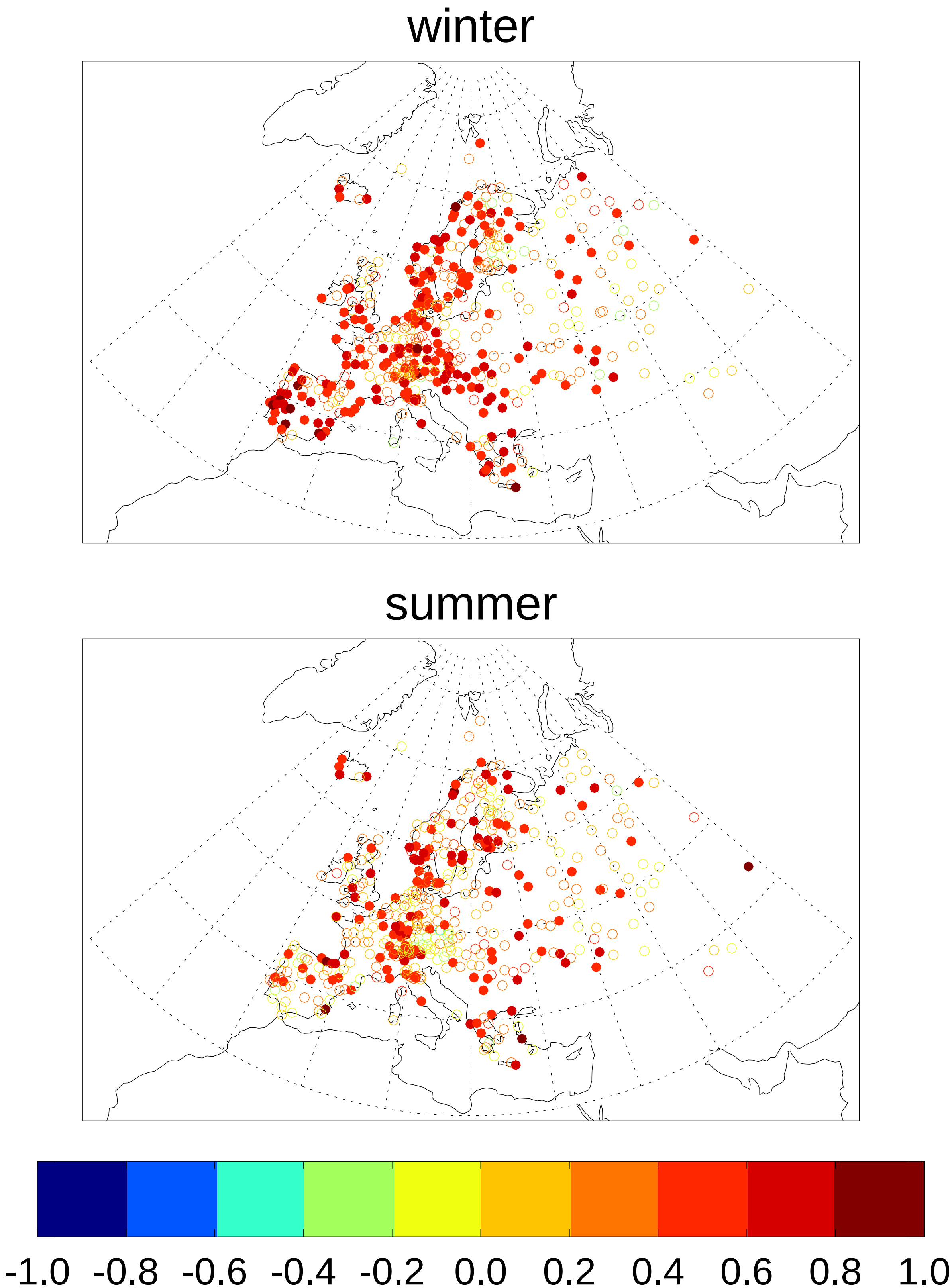


summer



Wed 06 Oct 2004 T. Schmith, DMI

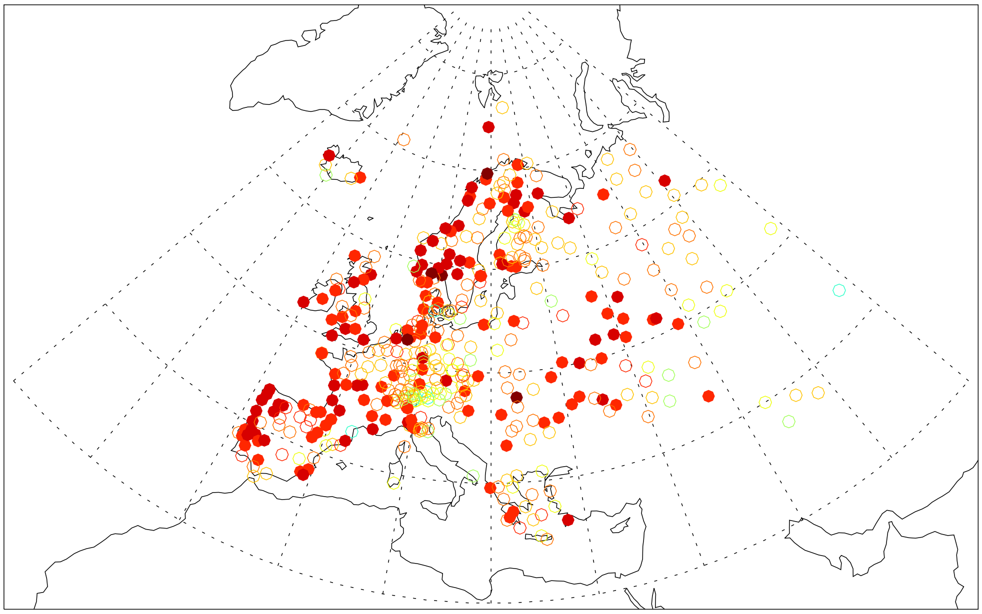
FIC dataset
Index: pnl90
Method: 2SA (FIC)
Period: 1958 - 2000



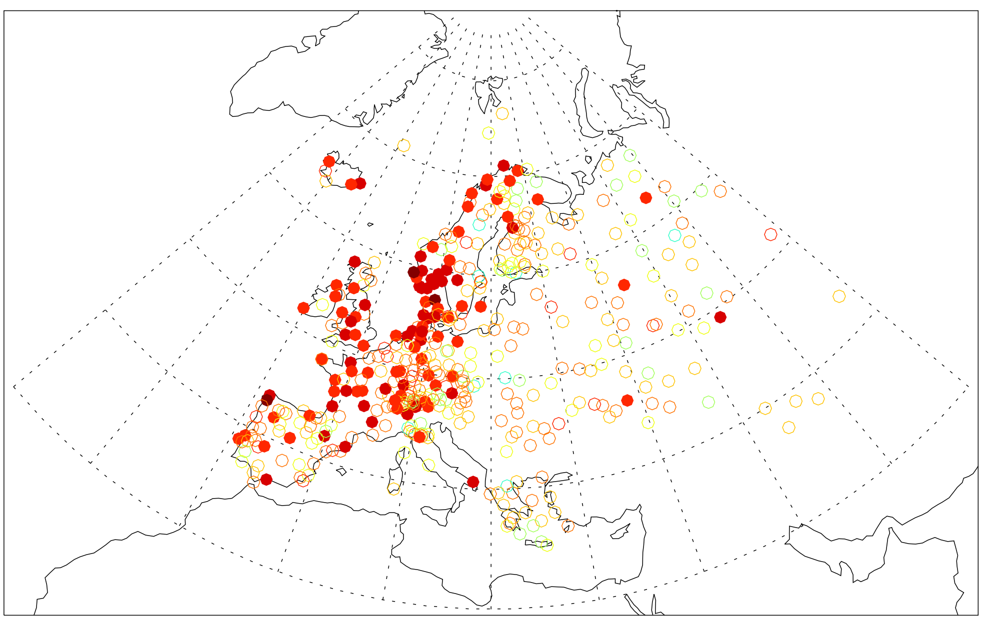
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: pnl90
Method: CWG (DMI)
Period: 1958 - 2000

winter

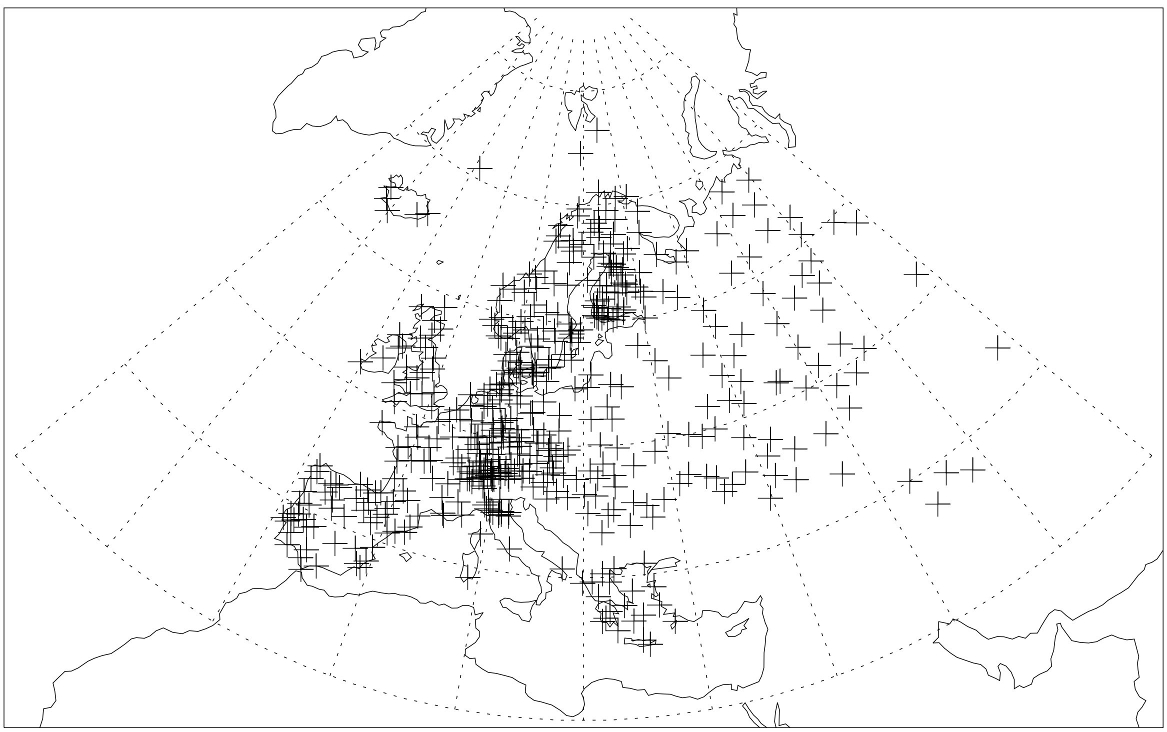


summer

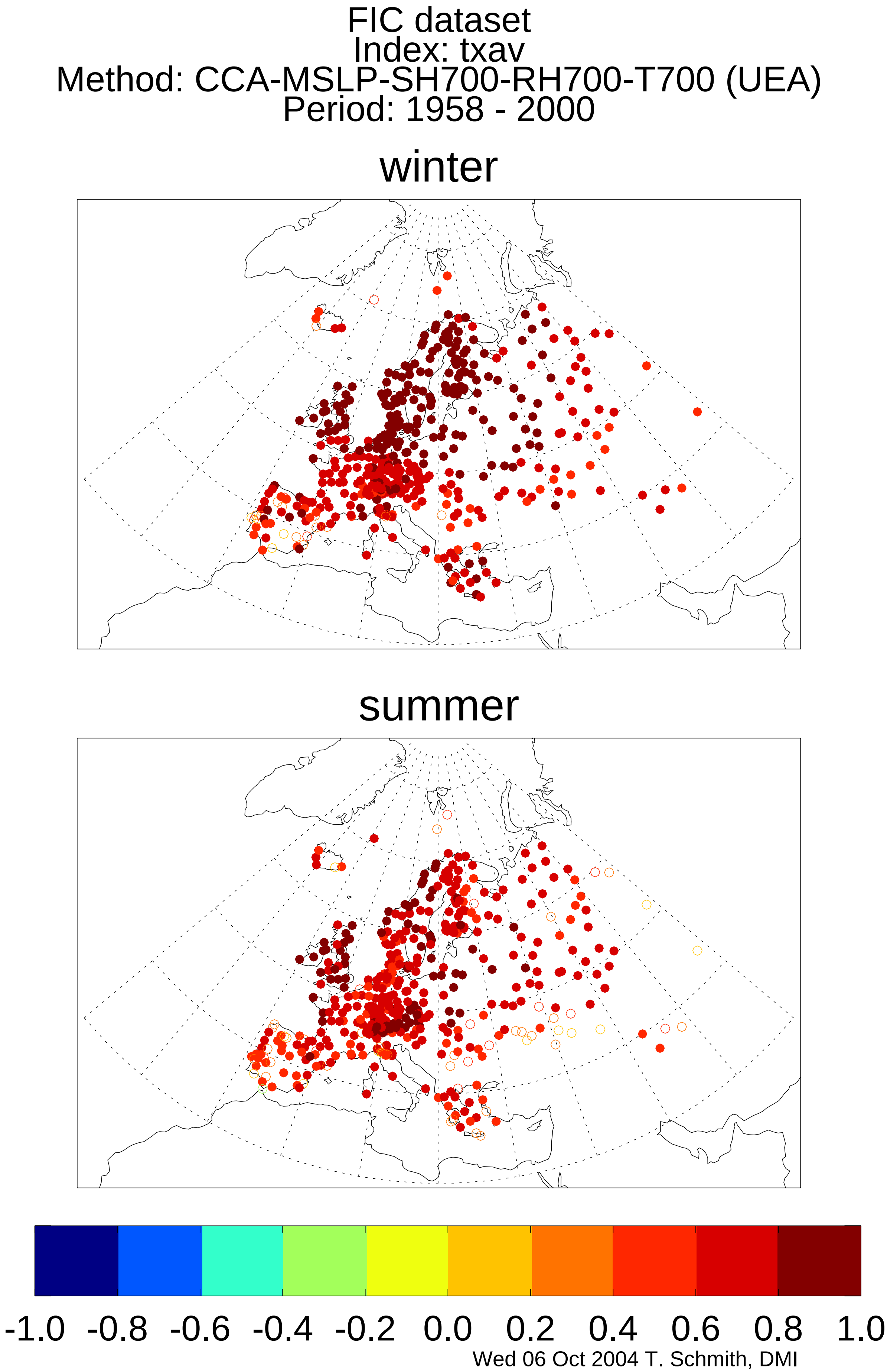


Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Station map
Period: 1958 - 2000

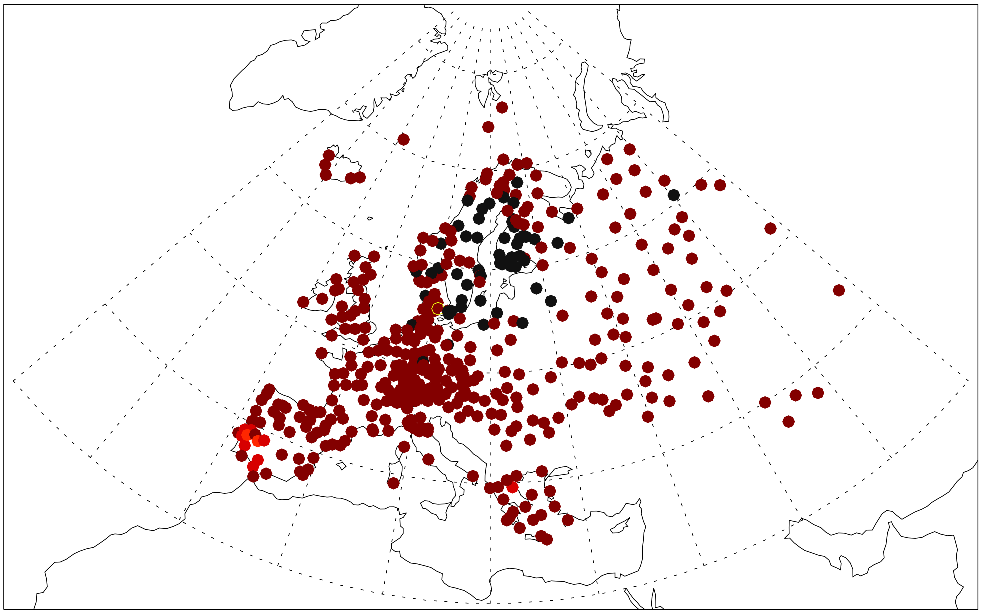


Wed 06 Oct 2004 T. Schmith, DMI

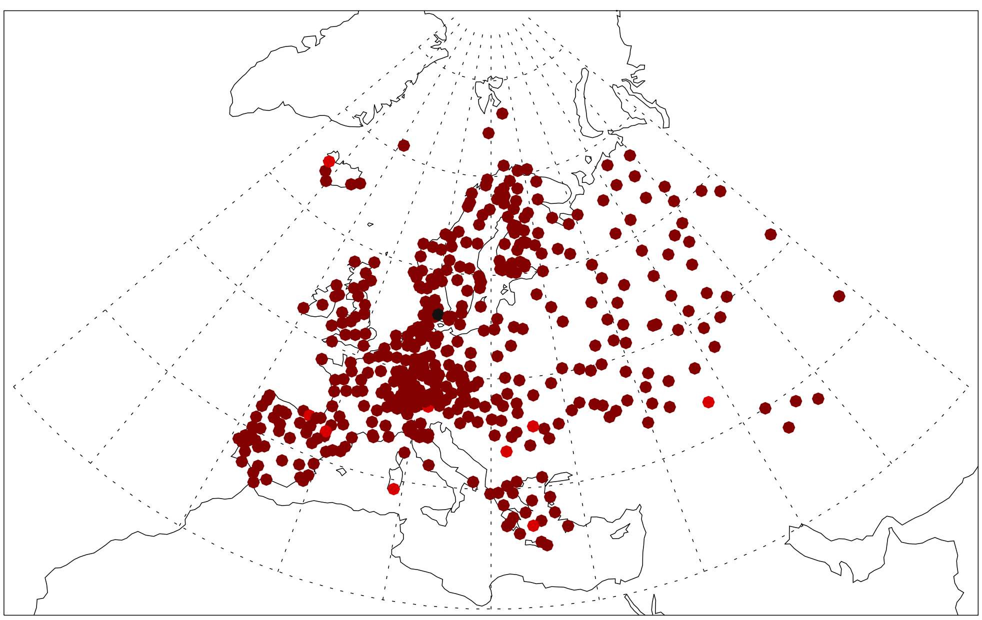

FIC dataset
Index: txav
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000
winter
summer
-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
Wed 06 Oct 2004 T. Schmith, DMI

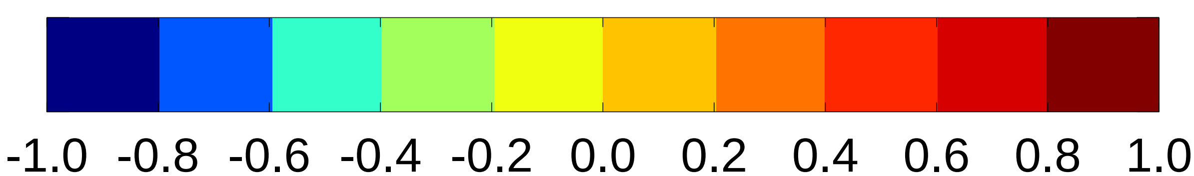
FIC dataset
Index: txav
Method: 2SA (FIC)
Period: 1958 - 2000

winter



summer

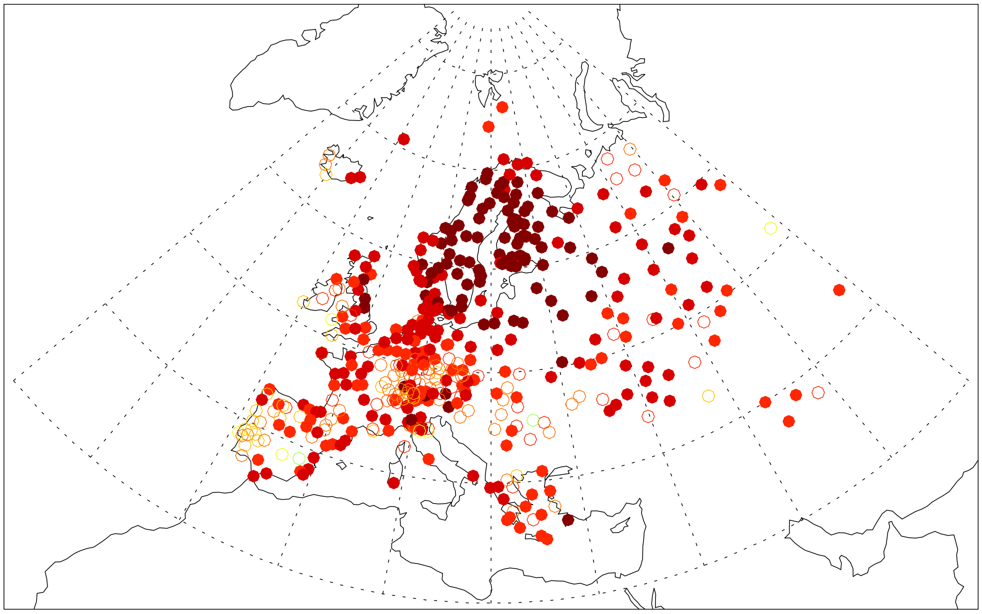




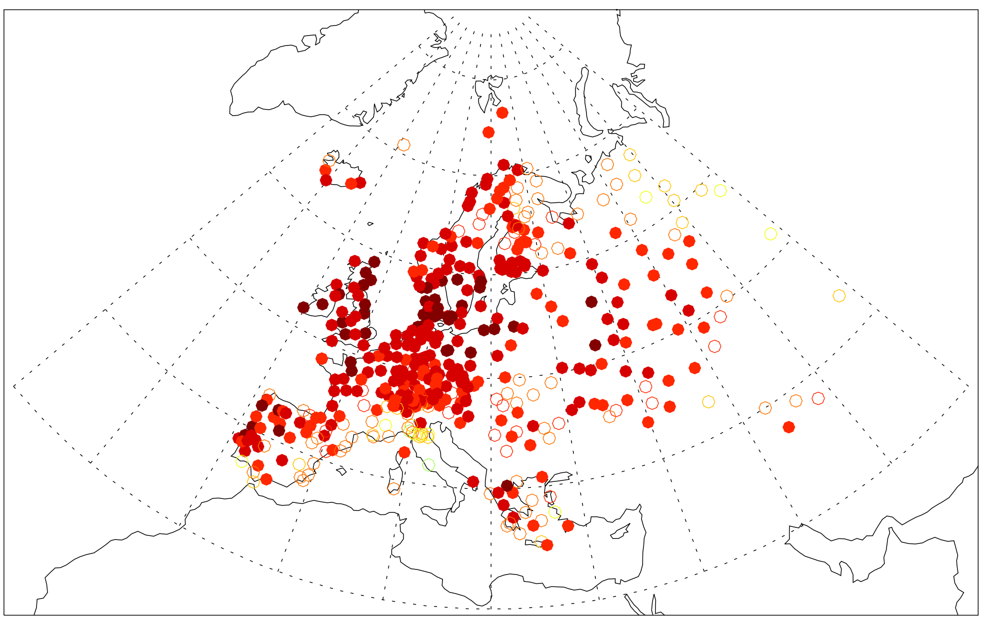
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: txq90
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

winter

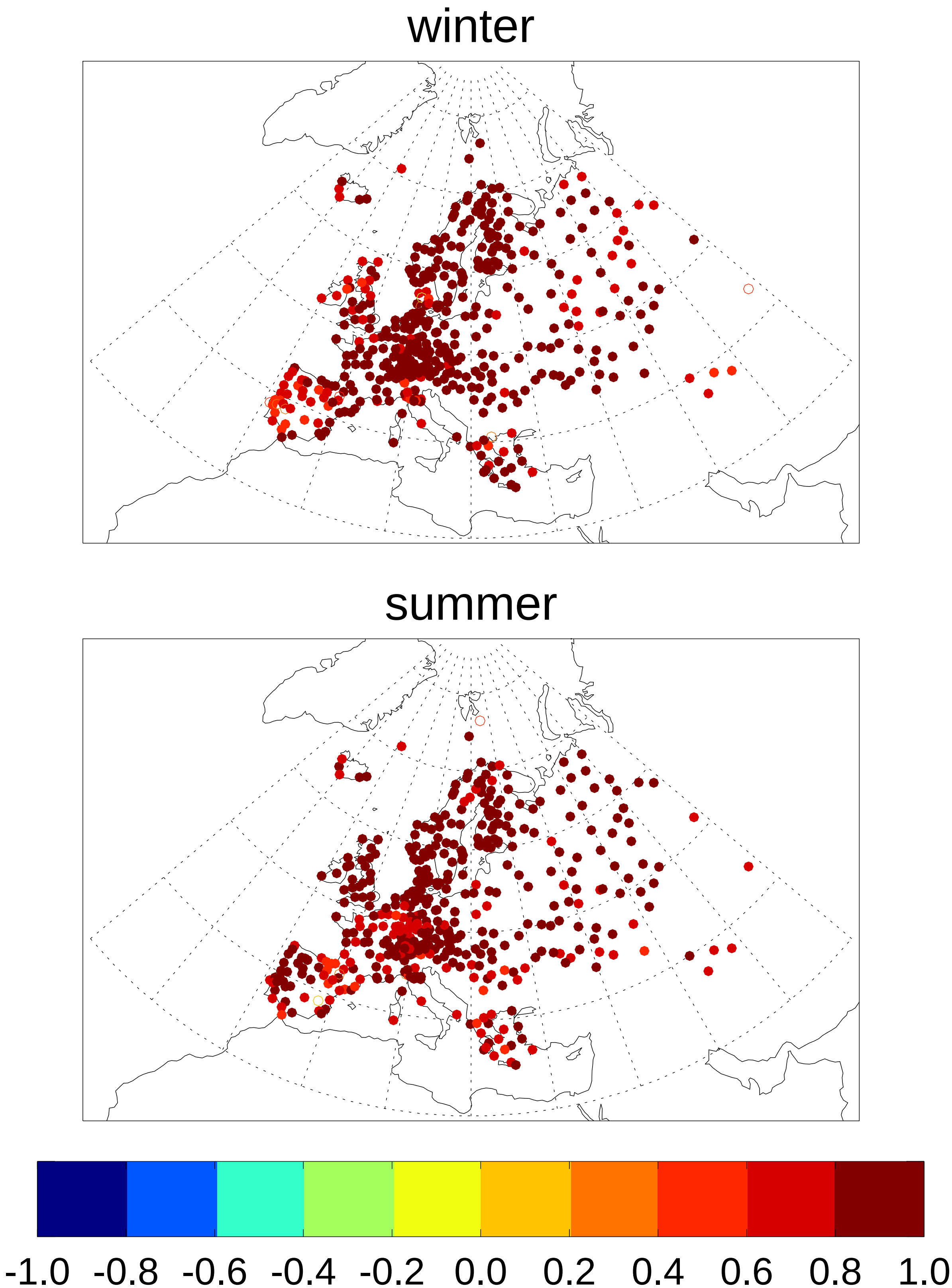


summer



Wed 06 Oct 2004 T. Schmith, DMI

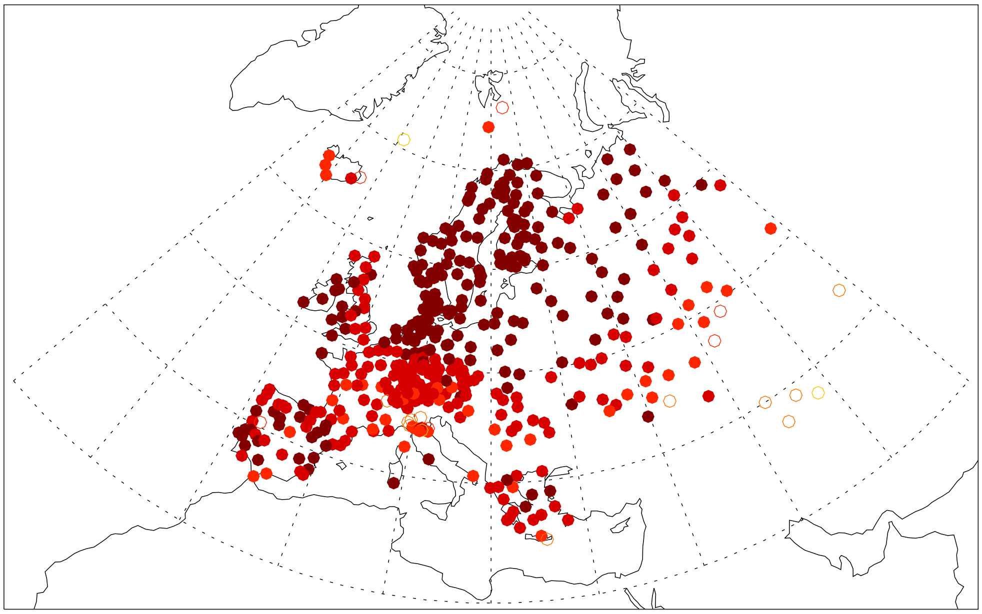
FIC dataset
Index: txq90
Method: 2SA (FIC)
Period: 1958 - 2000



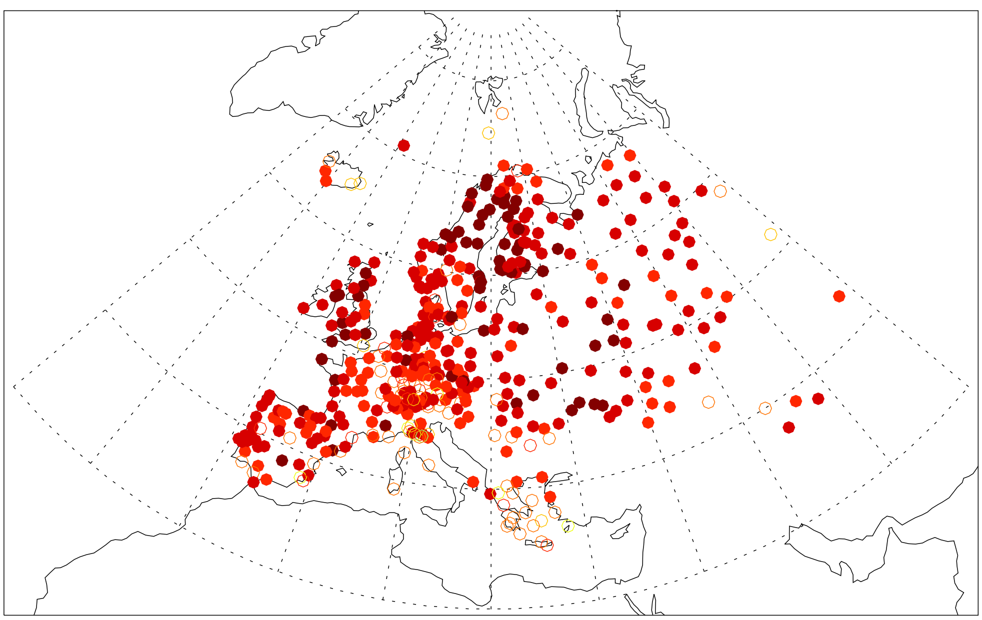
Wed 06 Oct 2004 T. Schmith, DMI

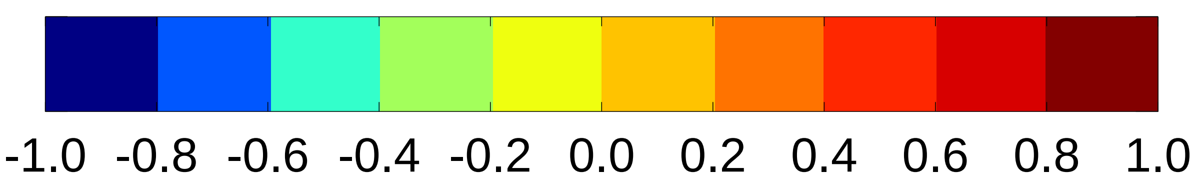
FIC dataset
Index: tnav
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000

winter



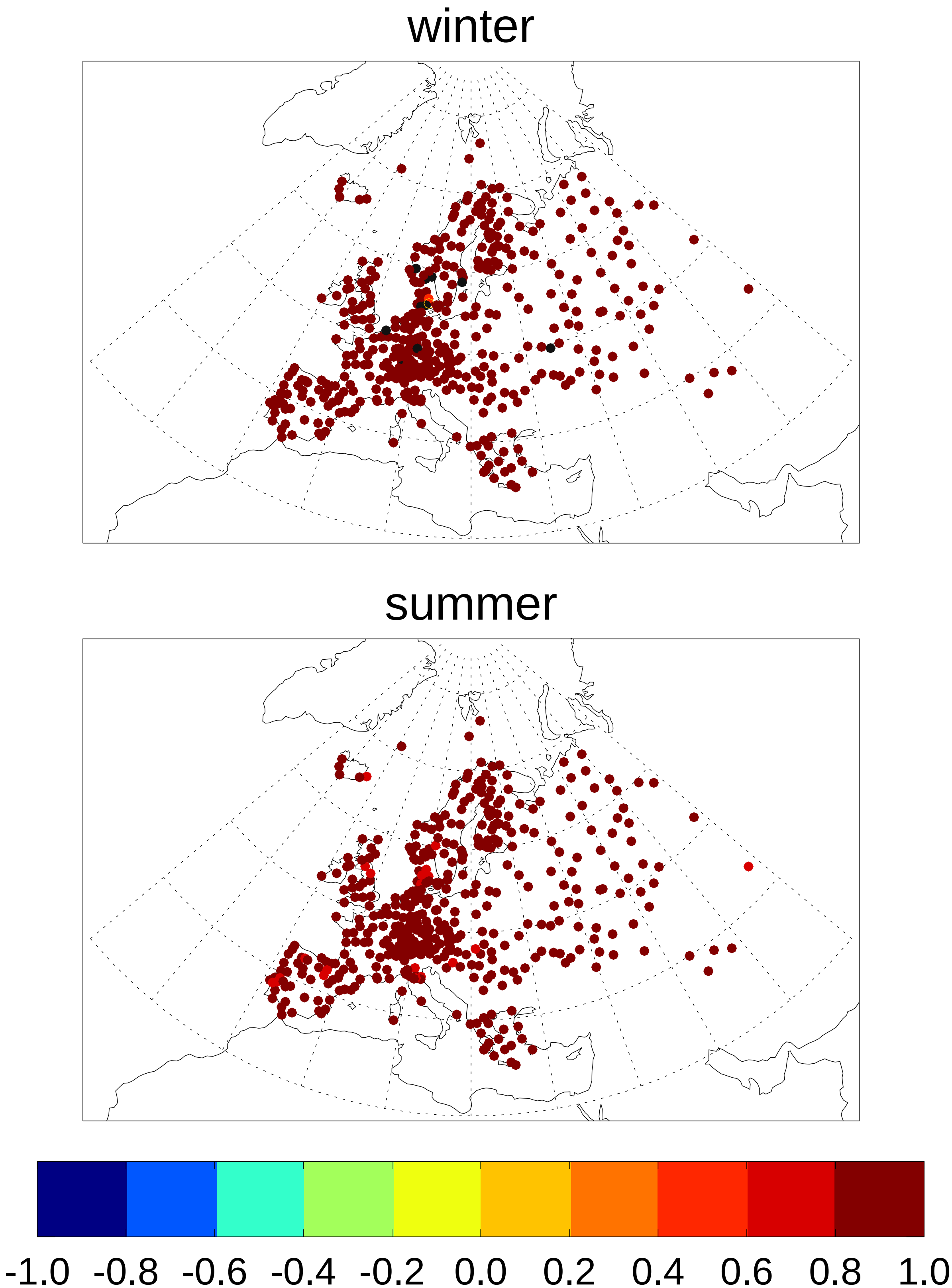
summer





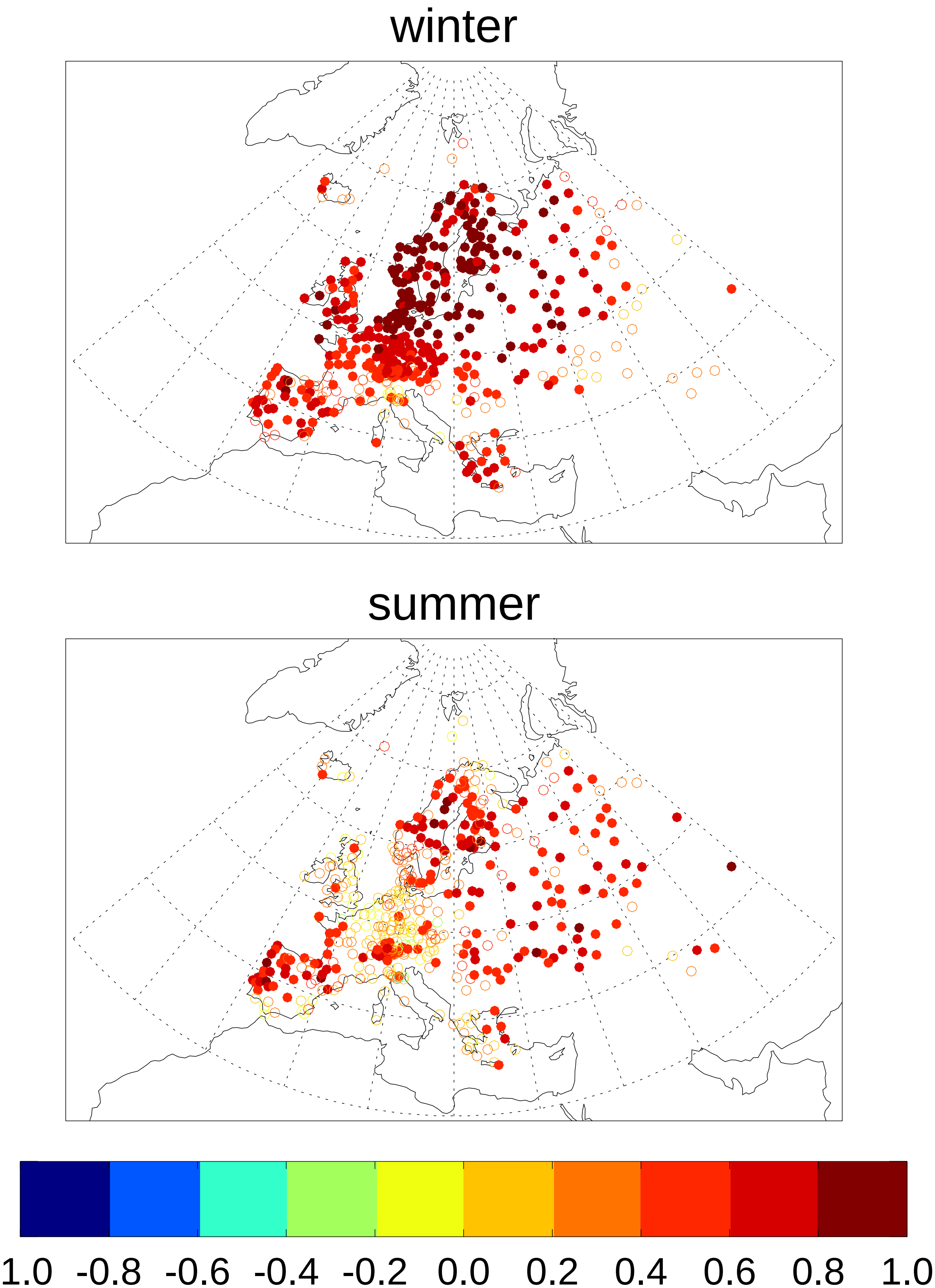
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: tnav
Method: 2SA (FIC)
Period: 1958 - 2000

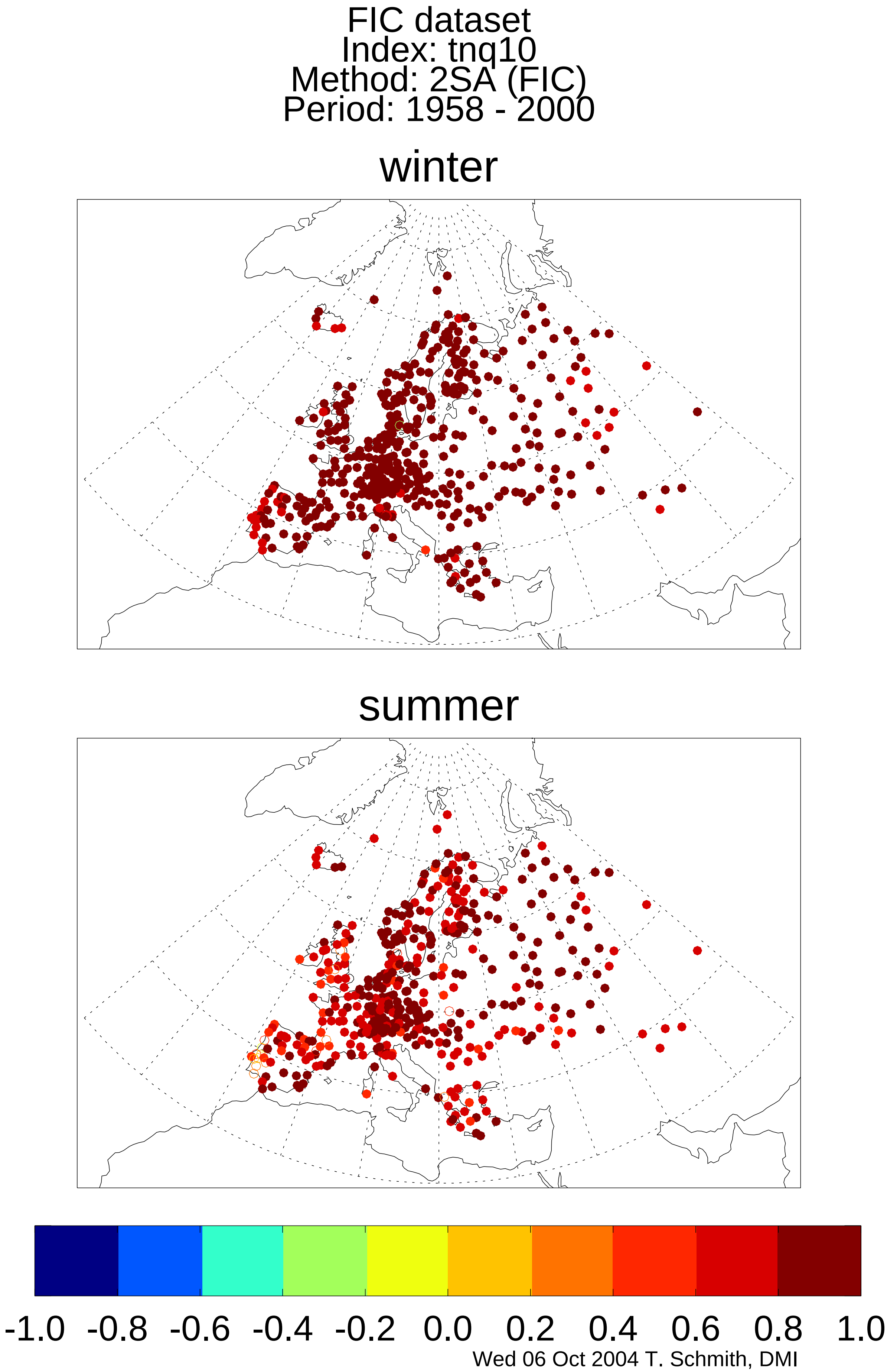


Wed 06 Oct 2004 T. Schmith, DMI

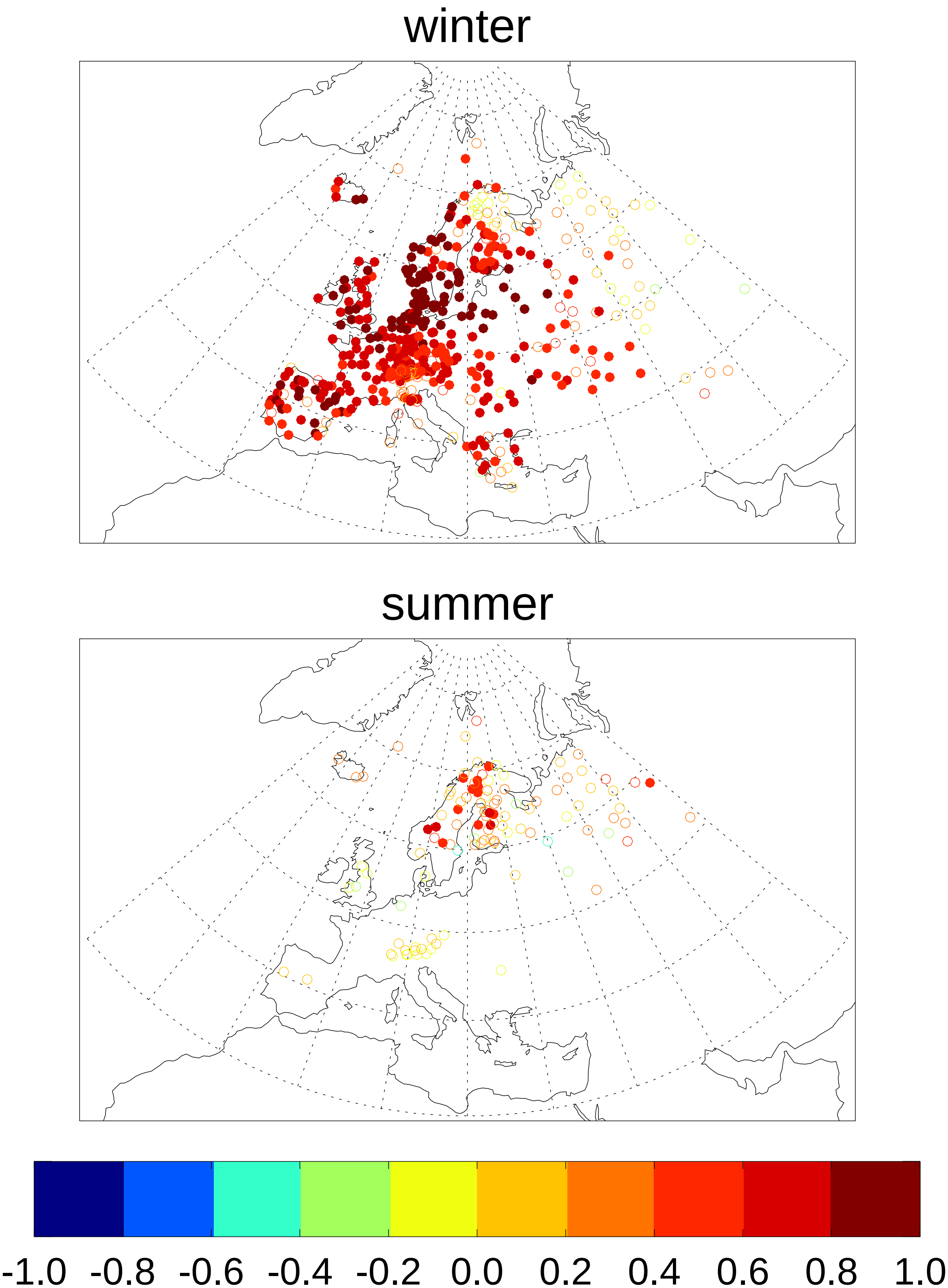
FIC dataset
Index: tnq10
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000



Wed 06 Oct 2004 T. Schmith, DMI

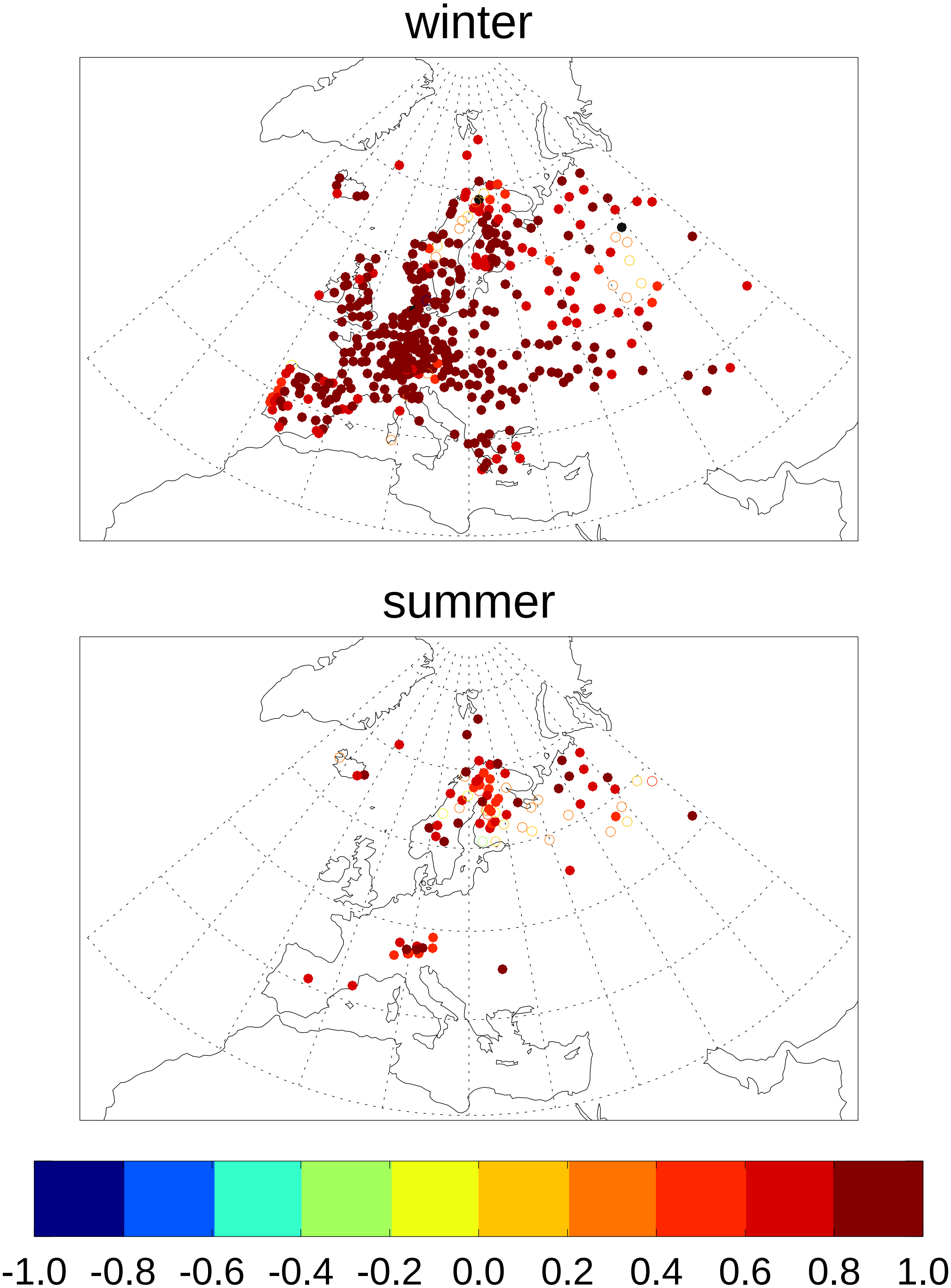

FIC dataset
Index: tnq10
Method: 2SA (FIC)
Period: 1958 - 2000
winter
summer
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: tnfd
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000



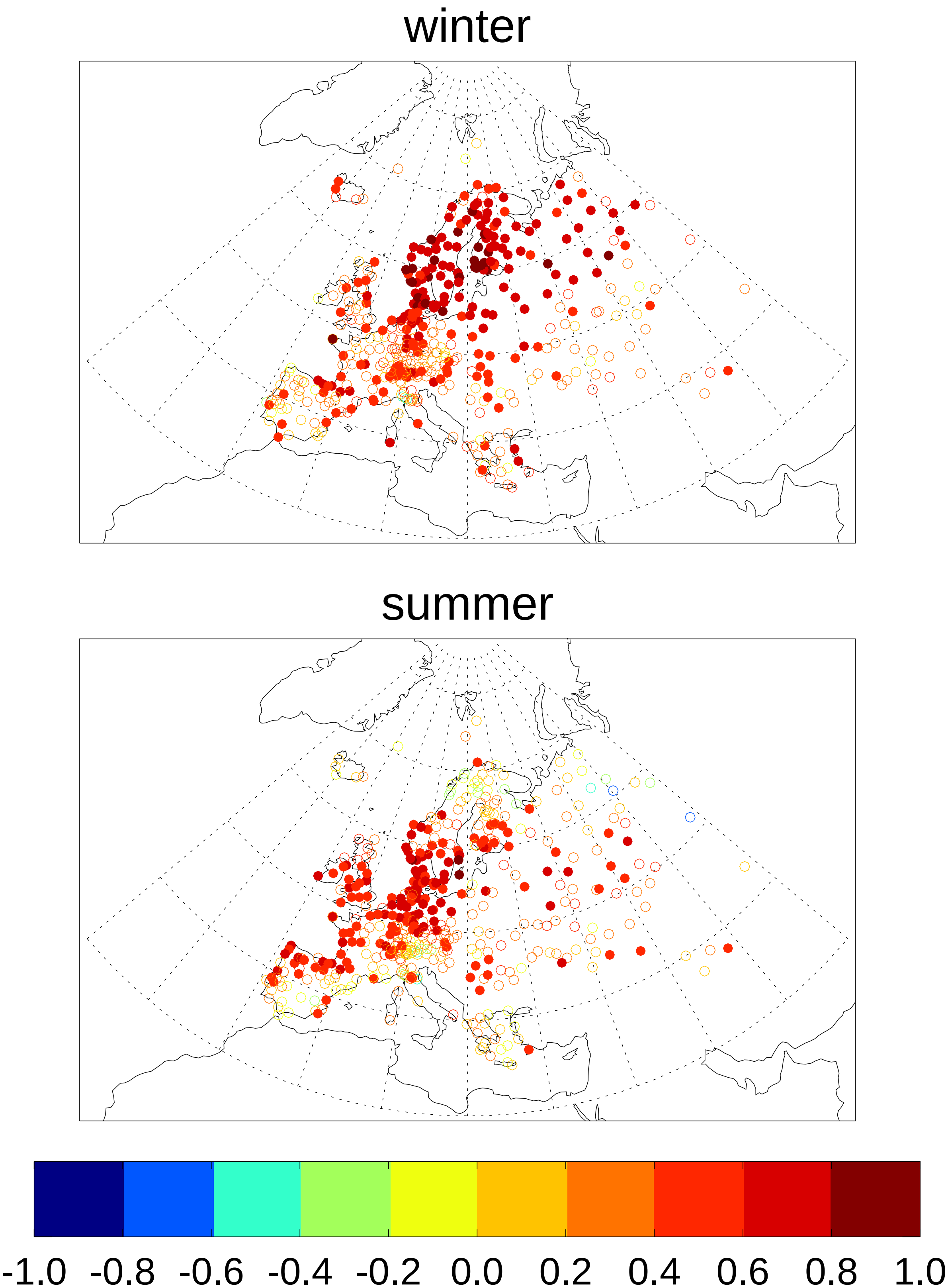
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: tnfd
Method: 2SA (FIC)
Period: 1958 - 2000



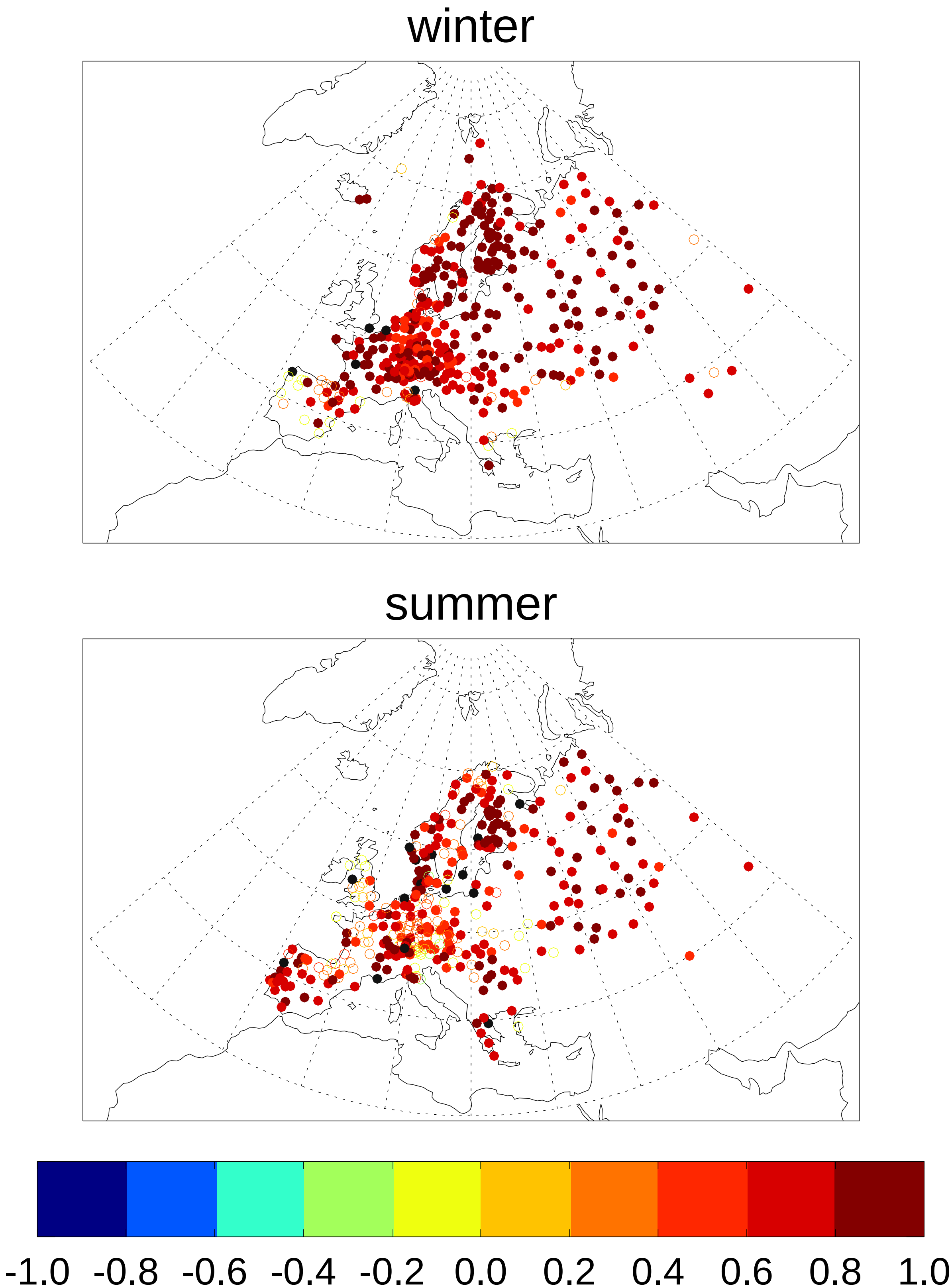
Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: txhw90
Method: CCA-MSLP-SH700-RH700-T700 (UEA)
Period: 1958 - 2000



Wed 06 Oct 2004 T. Schmith, DMI

FIC dataset
Index: txhw90
Method: 2SA (FIC)
Period: 1958 - 2000



Wed 06 Oct 2004 T. Schmith, DMI