Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

cen

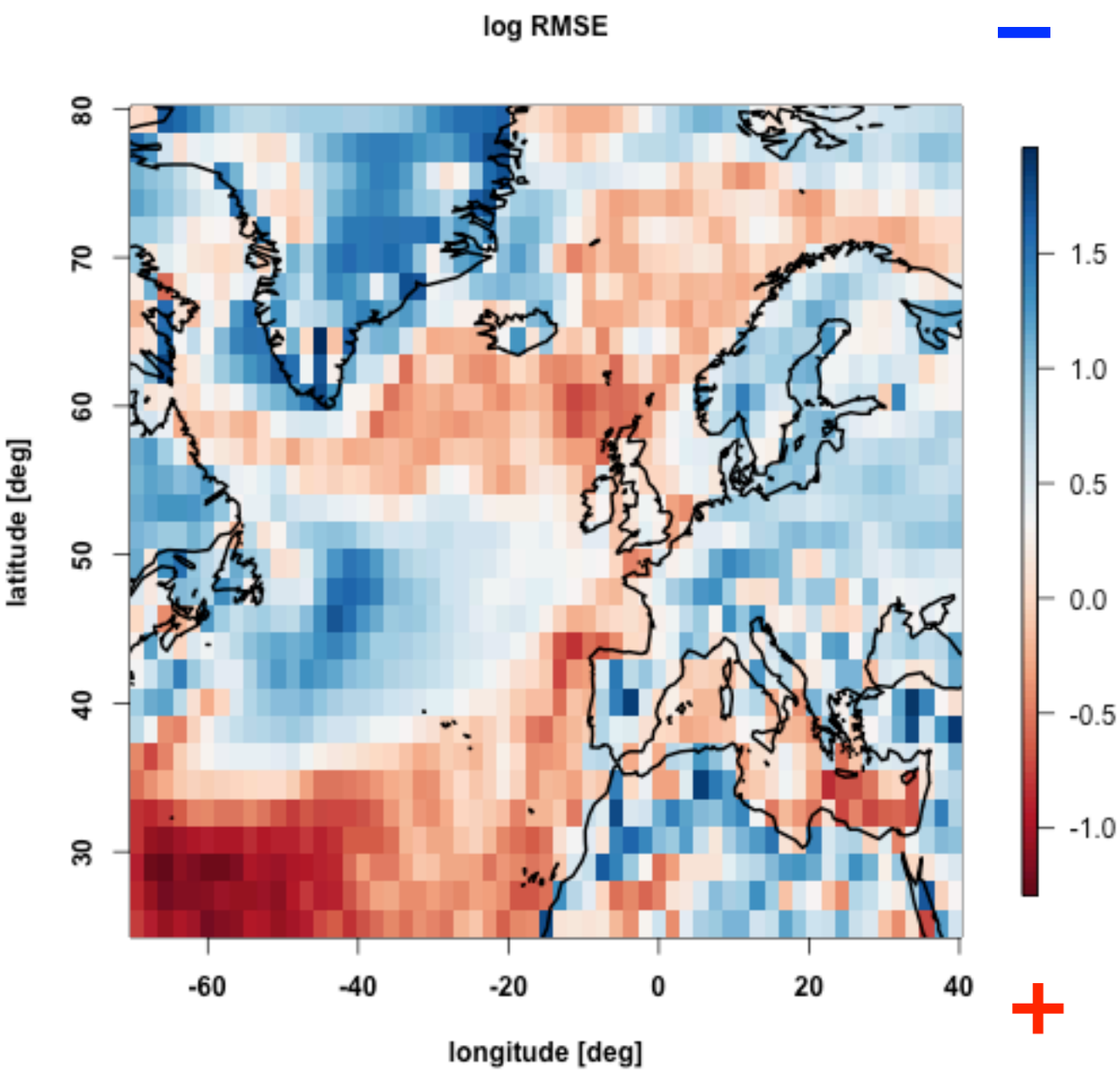
An advanced score for evaluating seasonal forecast skill

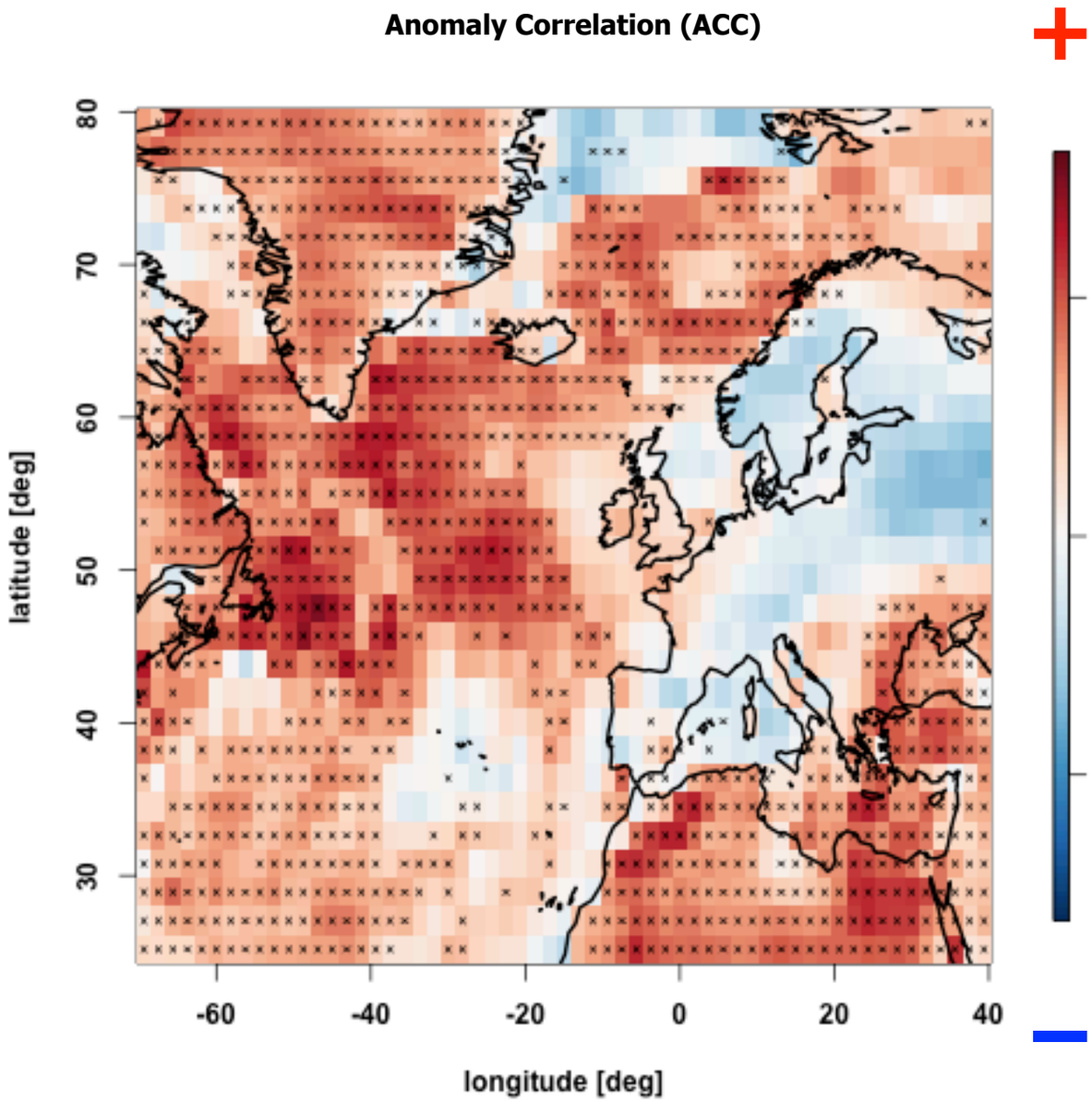
André Düsterhus & Sabrina Wahl

andre.duesterhus@uni-hamburg.de

S2D 2018
Boulder, 19th September 2018

Seasonal Prediction

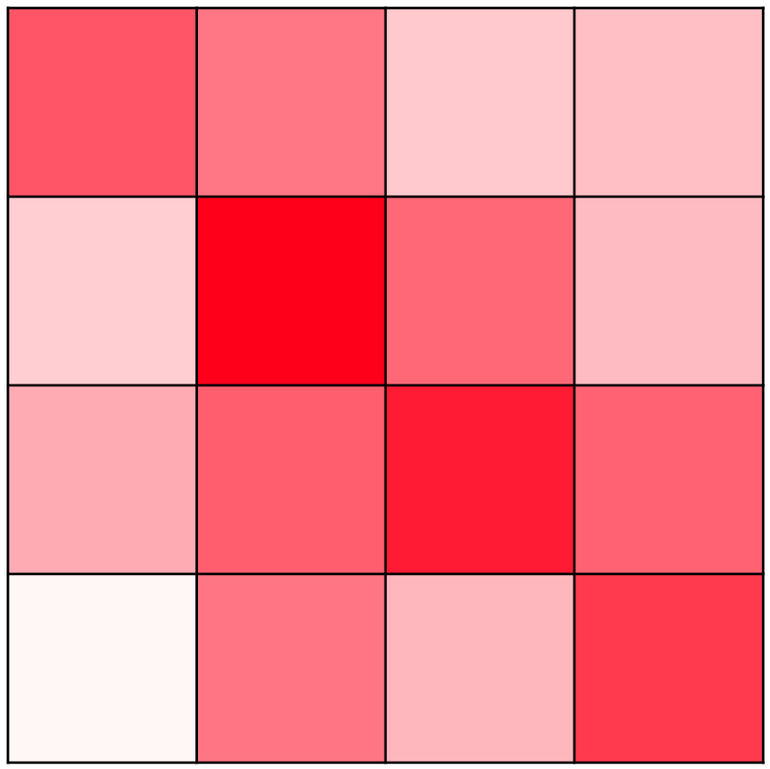




MPI-ESM-SPS, MR, 30 member ensemble mean, May initialisation, mean surface temperature JA, 1982-2015

Multi-Categorical Contingency Tables

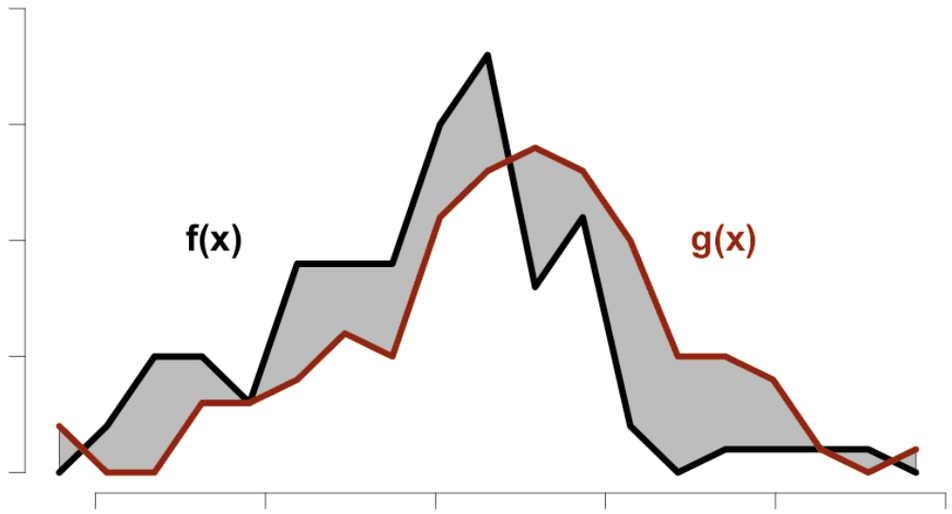
observations



prediction

- Categories for variables have to be ordered
- Common for precipitation, climate indices

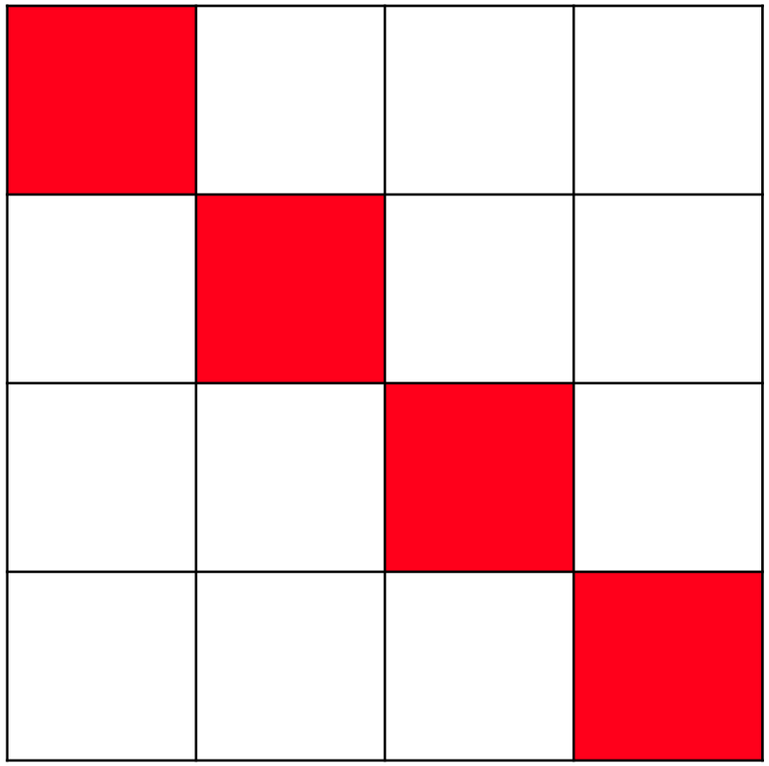
Earth Mover's Distance



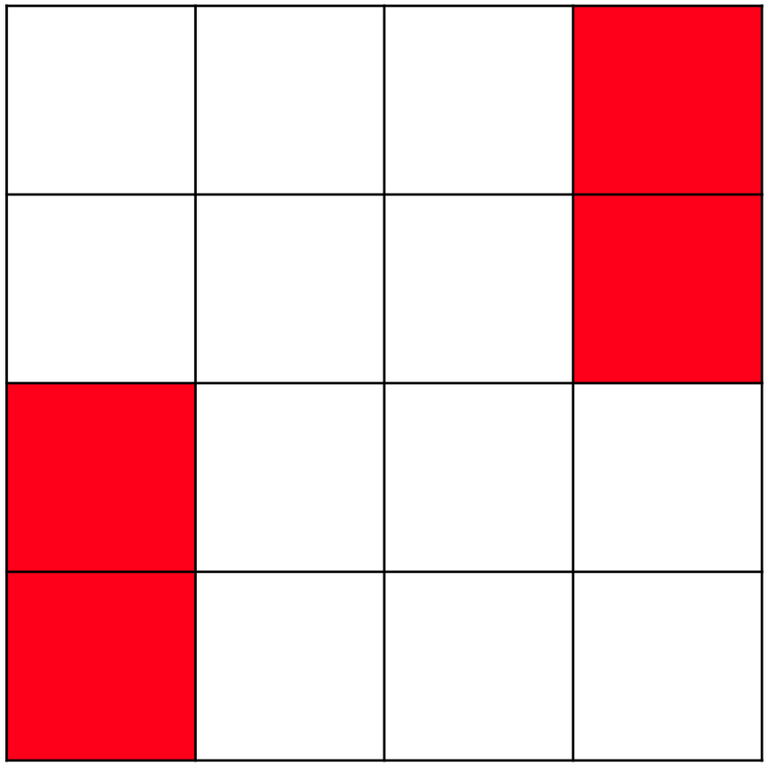
- We transform probability fields
- Equivalent to the Wasserstein metric
- Optimal solution via an optimisation approach

Y. Rubner et al. (2001): Empirical Evaluation of Dissimilarity Measures for Color and Texture, CVIU, 84, 25–43

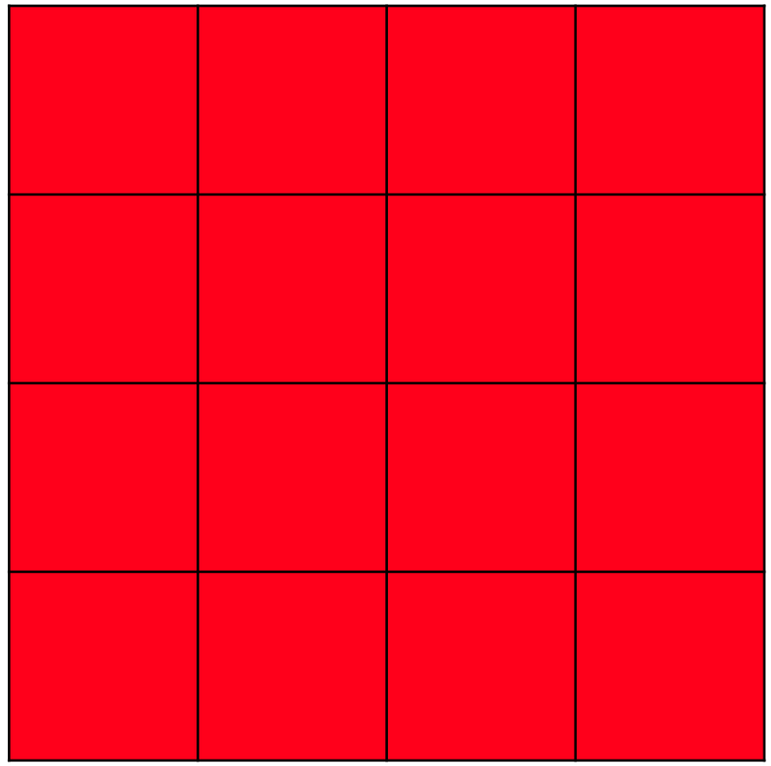
Important MCTs



Optimal



Worst



Equal

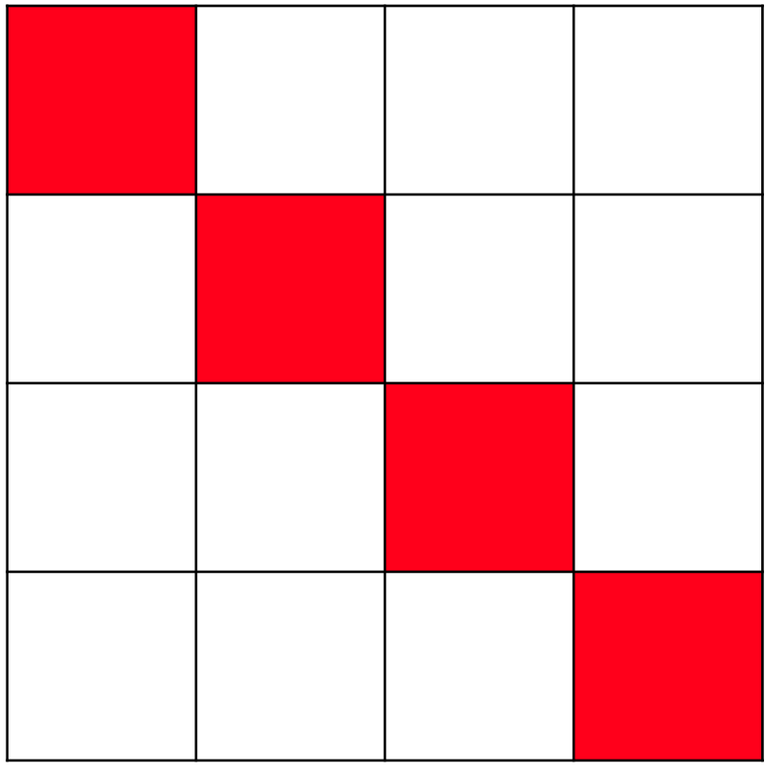
Categorical 2D-EMD score

- The distance D is the 2D-EMD between two MCTs
- We define the score as:

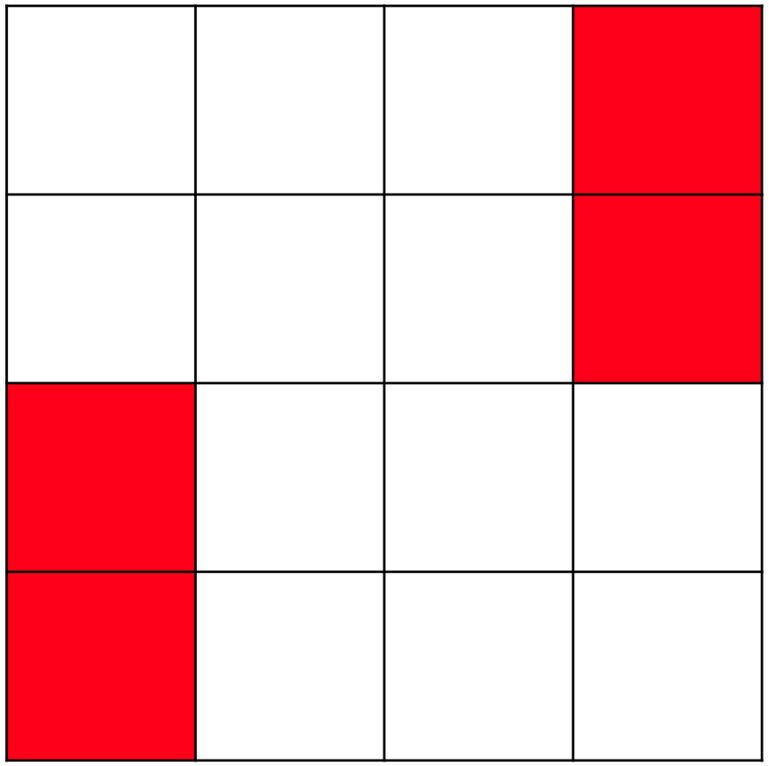
$$S_{EMD}(\mathbb{A}) = 1 - 2 \cdot \frac{D(\mathbb{A}, \mathbb{O})}{D(\mathbb{W}, \mathbb{O})}$$

$\mathbb{A}$	Test matrix
$\mathbb{O}$	Optimal matrix
$\mathbb{W}$	Worst matrix

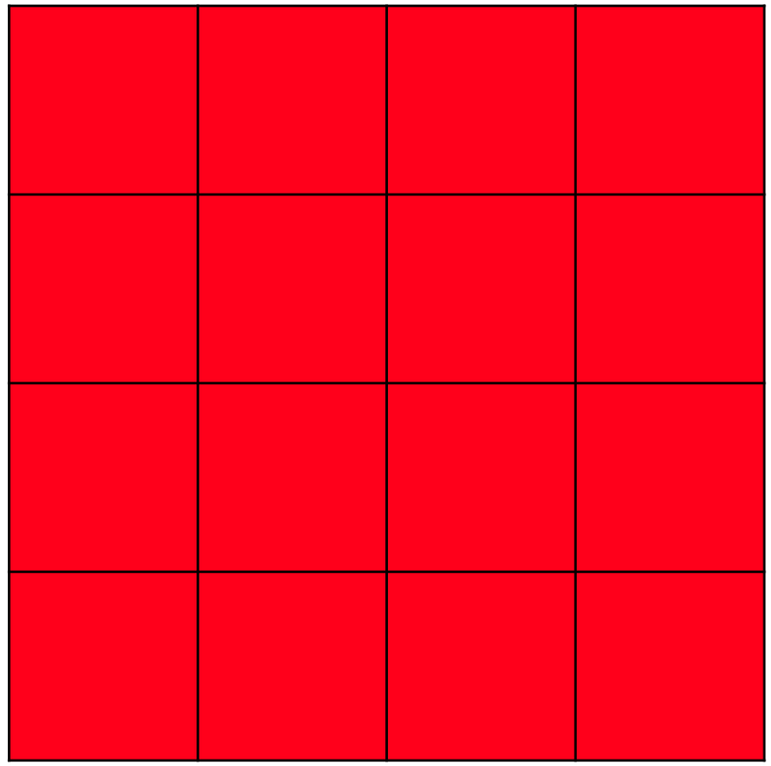
Important MCTs



Optimal
1



Worst
-1



Equal
≈0

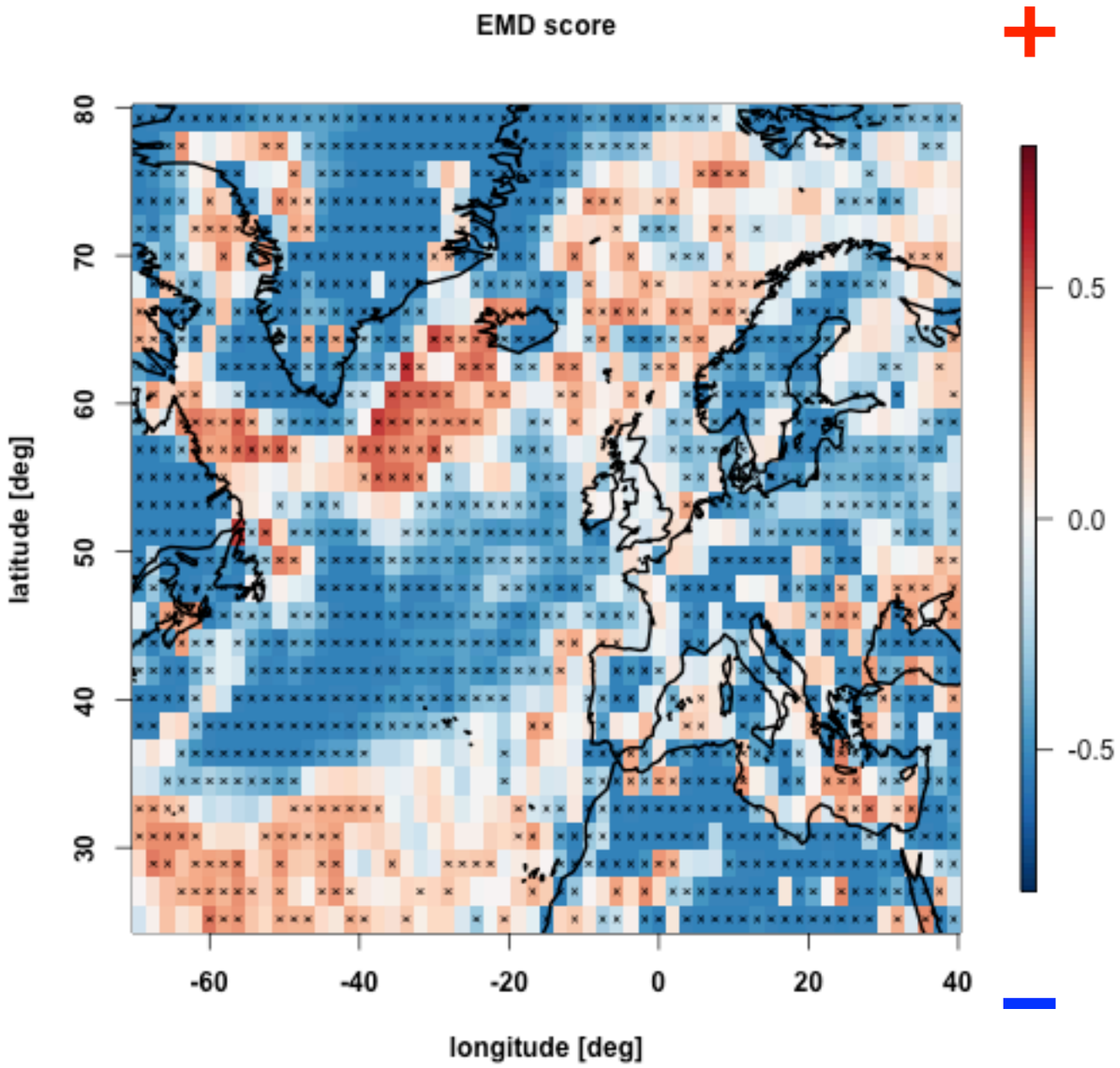
Continuous to categorical

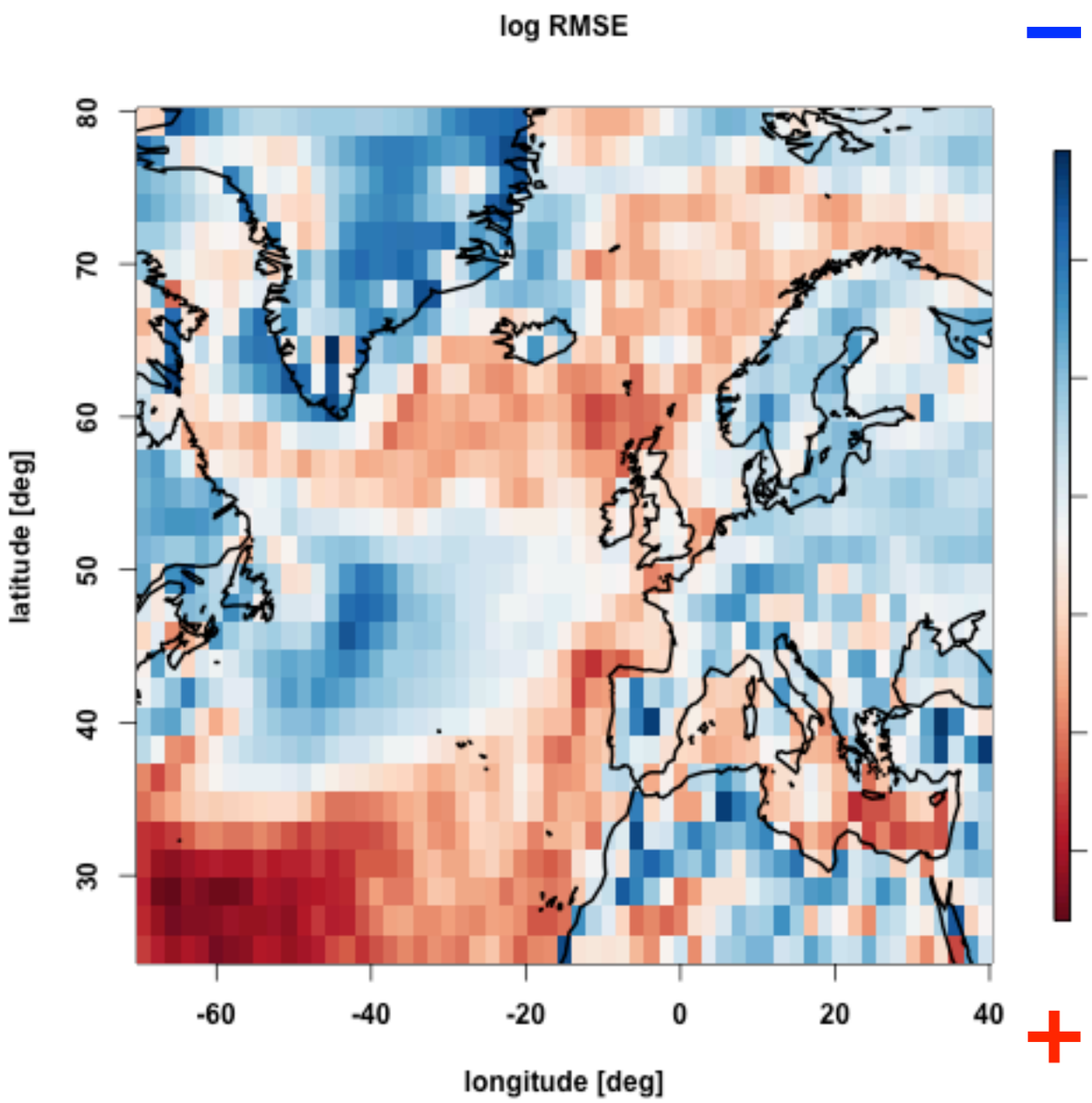
- Task: Transforming continuous variables into categorical
- Solution: Using deciles as separators
- Consequence: Limited sensitivity to extreme values

Choice of Categories

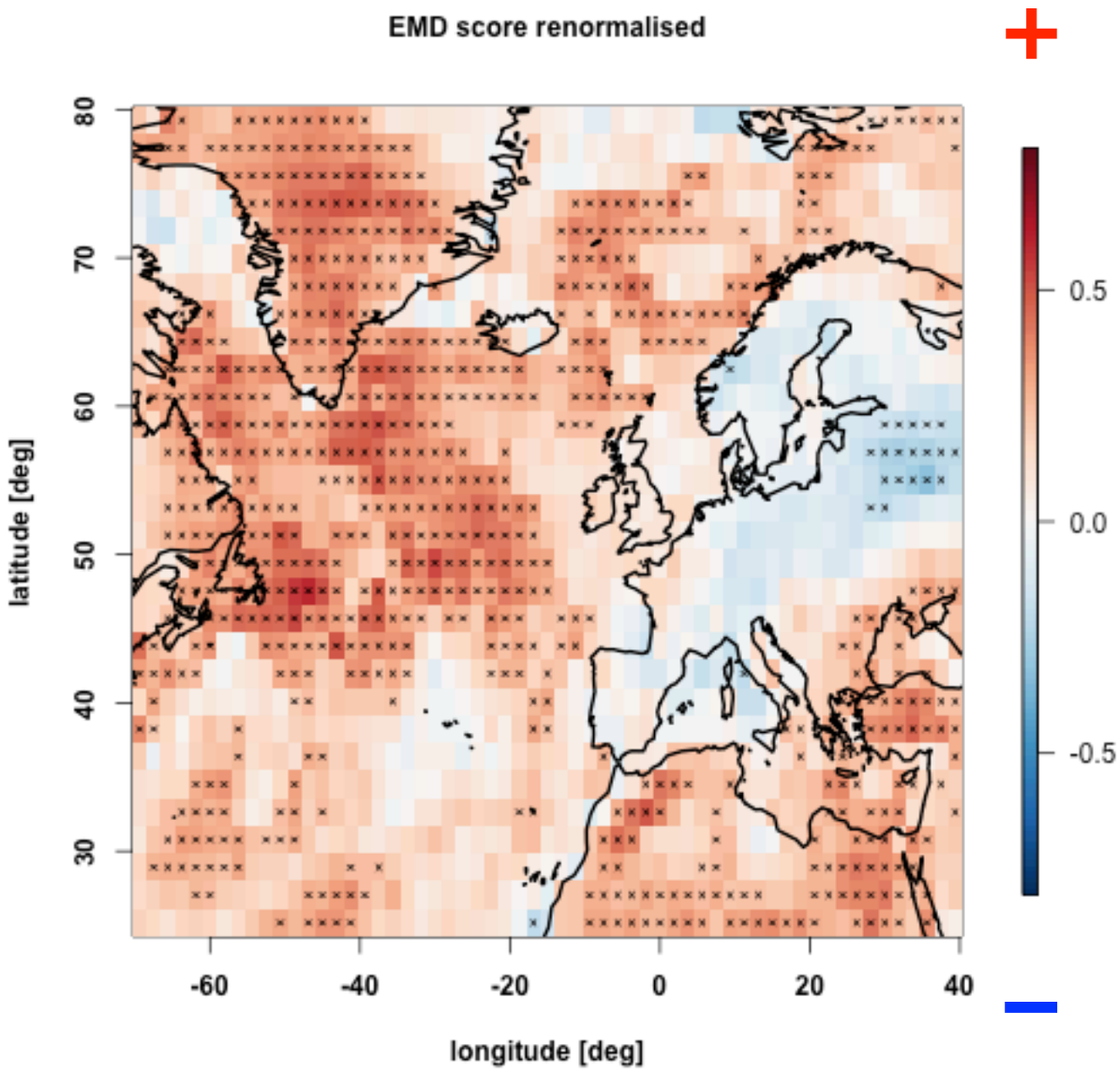
	detrend/ normalisation	Observation/ Reanalysis	Prediction
1		■	■
2	√	■	■
3	√	■	■

1 (log) RMSE

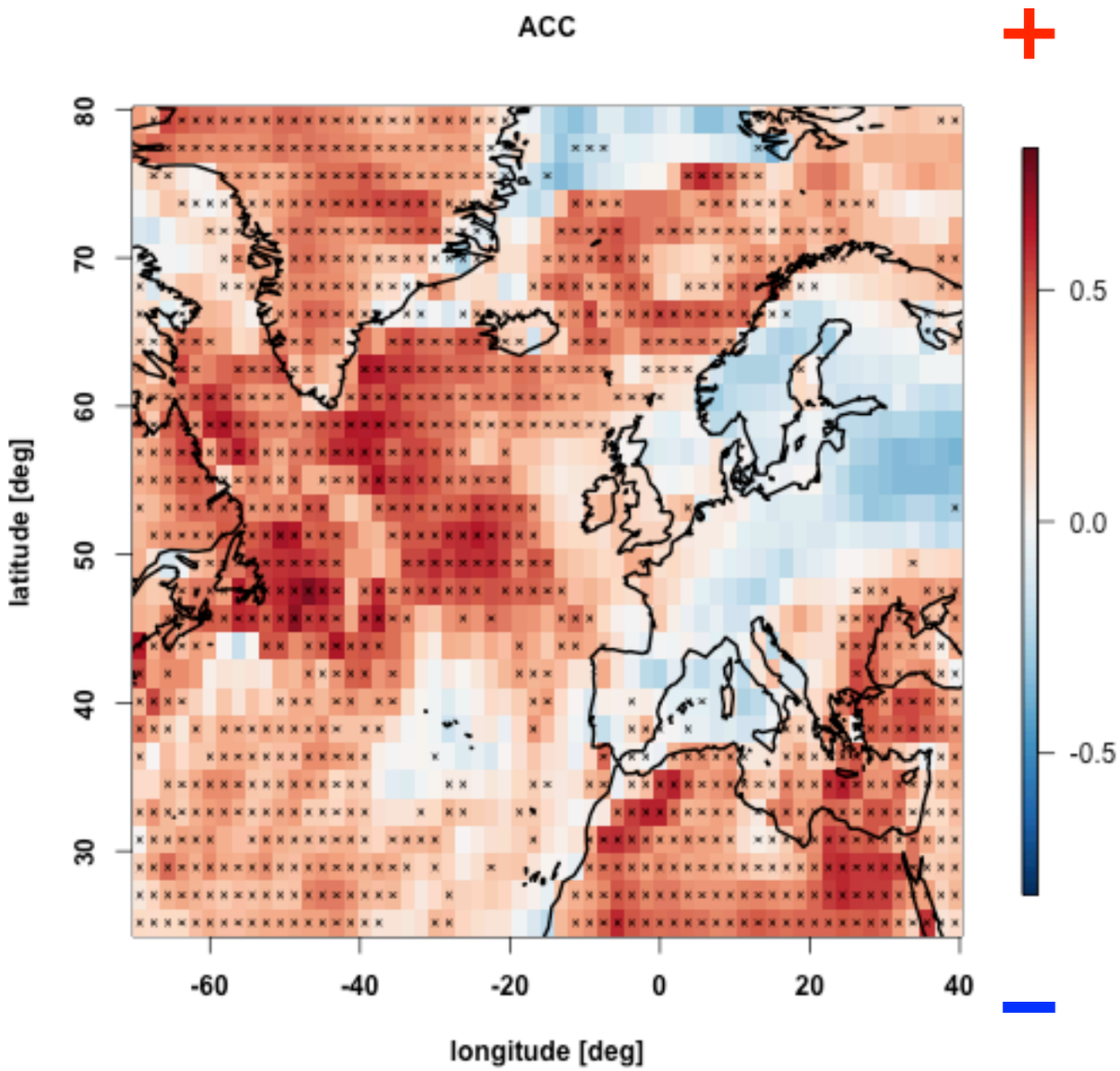

EMD score
latitude [deg]
longitude [deg]
80
70
60
50
40
30
-60
-40
-20
0
20
40
0.5
0.0
-0.5


log RMSE
latitude [deg]
longitude [deg]
80
70
60
50
40
30
-60
-40
-20
0
20
40
1.5
1.0
0.5
0.0
-0.5
-1.0

2 Anomaly Correlation (ACC)

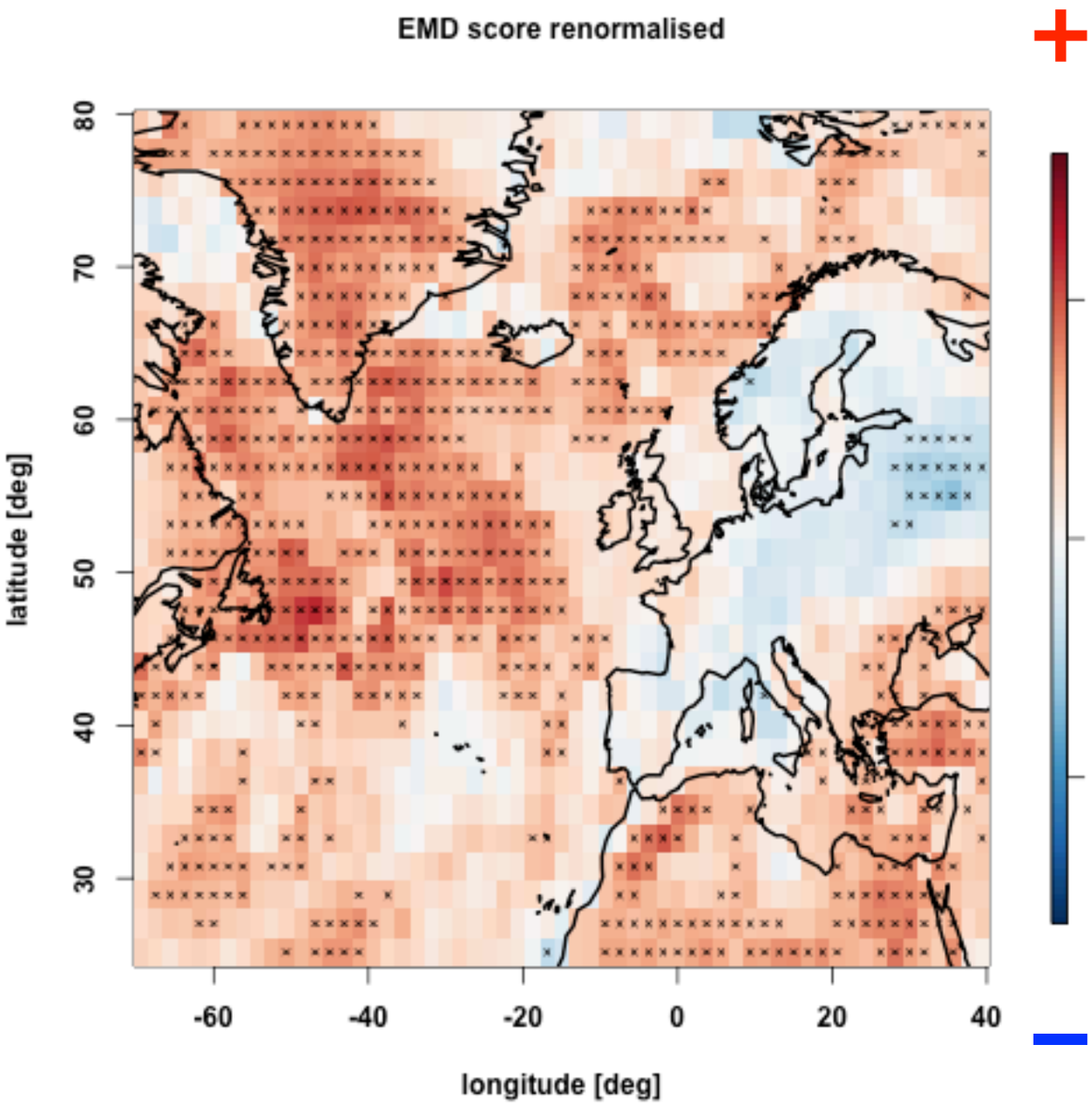


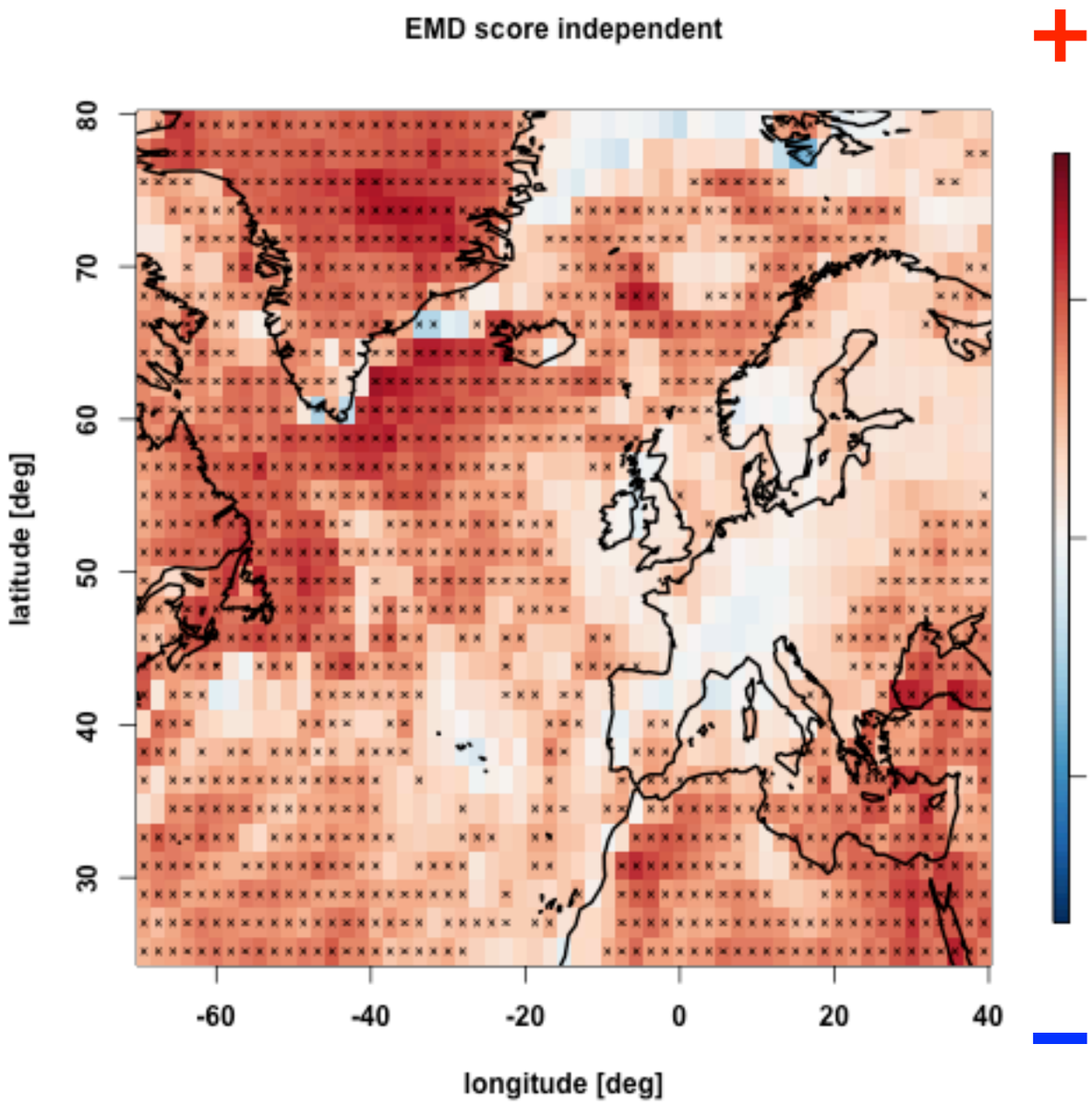
+



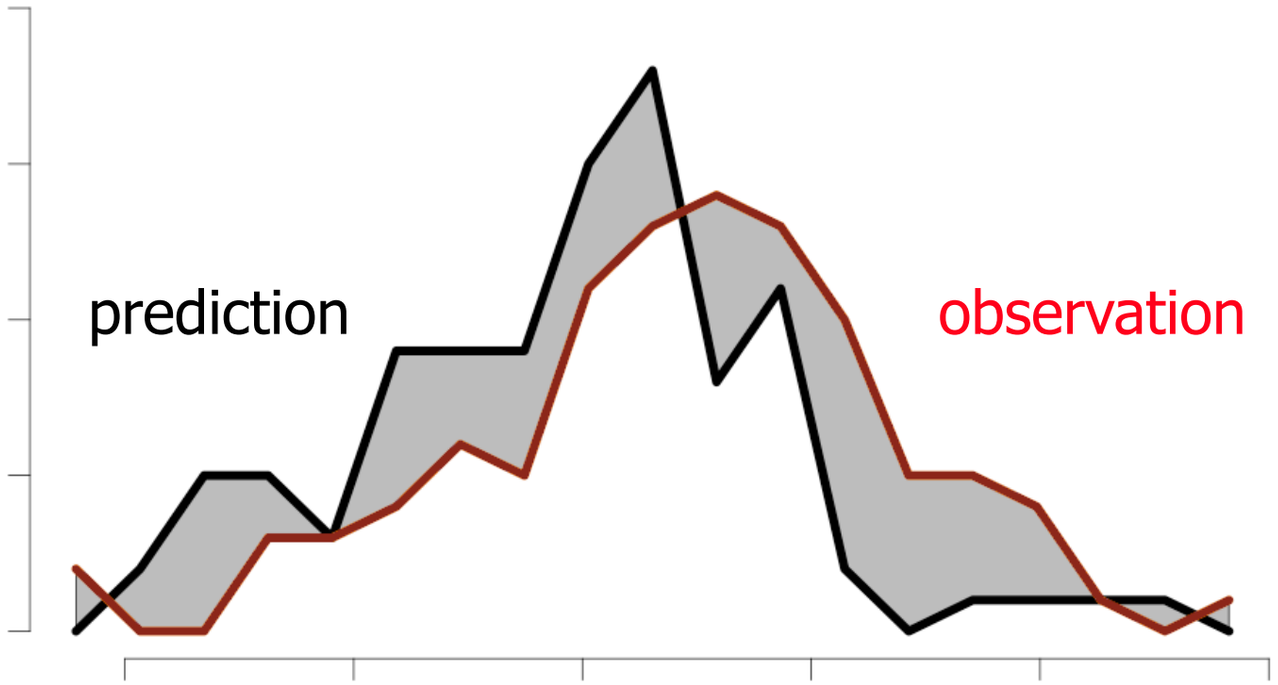
+

3 non-parametric ACC





1d-EMD-score



$\mathbb{C}$ Climatology

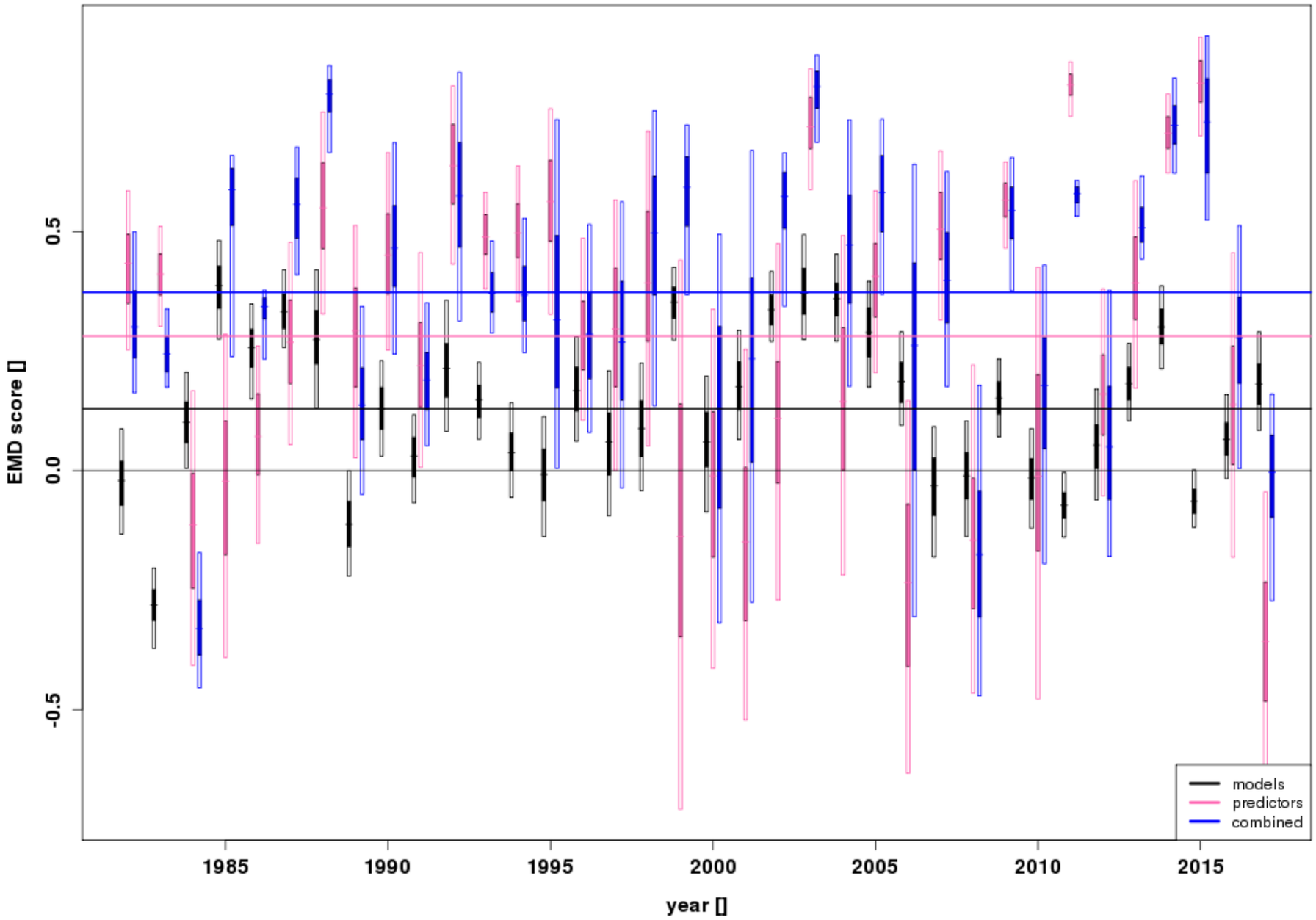
$\mathbb{O}$ Observation

$\mathbb{P}$ Prediction

$$S(\mathbb{P}) = 1 - \frac{D(\mathbb{P}, \mathbb{O})}{D(\mathbb{C}, \mathbb{O})}$$

$$S(\mathbb{O}) = 1$$
$$S(\mathbb{C}) = 0$$

1d-EMD-score


0.5
0.0
-0.5
EMD score []
1985
1990
1995
2000
2005
2010
2015
year []
models
predictors
combined

Summary

- A new score basing on the Earth Mover's Distance
- Designed to access MCTs
- Allows to replicate patterns of RMSE and ACC and offer a non-parametric version of ACC

A. Düsterhus & S. Wahl (in review): An advanced score for multi-categorical contingency tables

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

cen

Advanced score for the evaluation of prediction skill

André Düsterhus & Sabrina Wahl

andre.duesterhus@uni-hamburg.de

S2D 2018
Boulder, 19th September 2018

2D EMD

$$D(\mathbb{A},\mathbb{B}) = \frac{\sum_{i,j} h_{i,j} f_{i,j}}{\sum_{i,j} f_{i,j}}$$

$f_{i,j}$ optimal flow

$h_{i,j}$ Distance between histograms

$$\sum_{i,j} h_{i,j} f_{i,j} \stackrel{!}{=} min$$

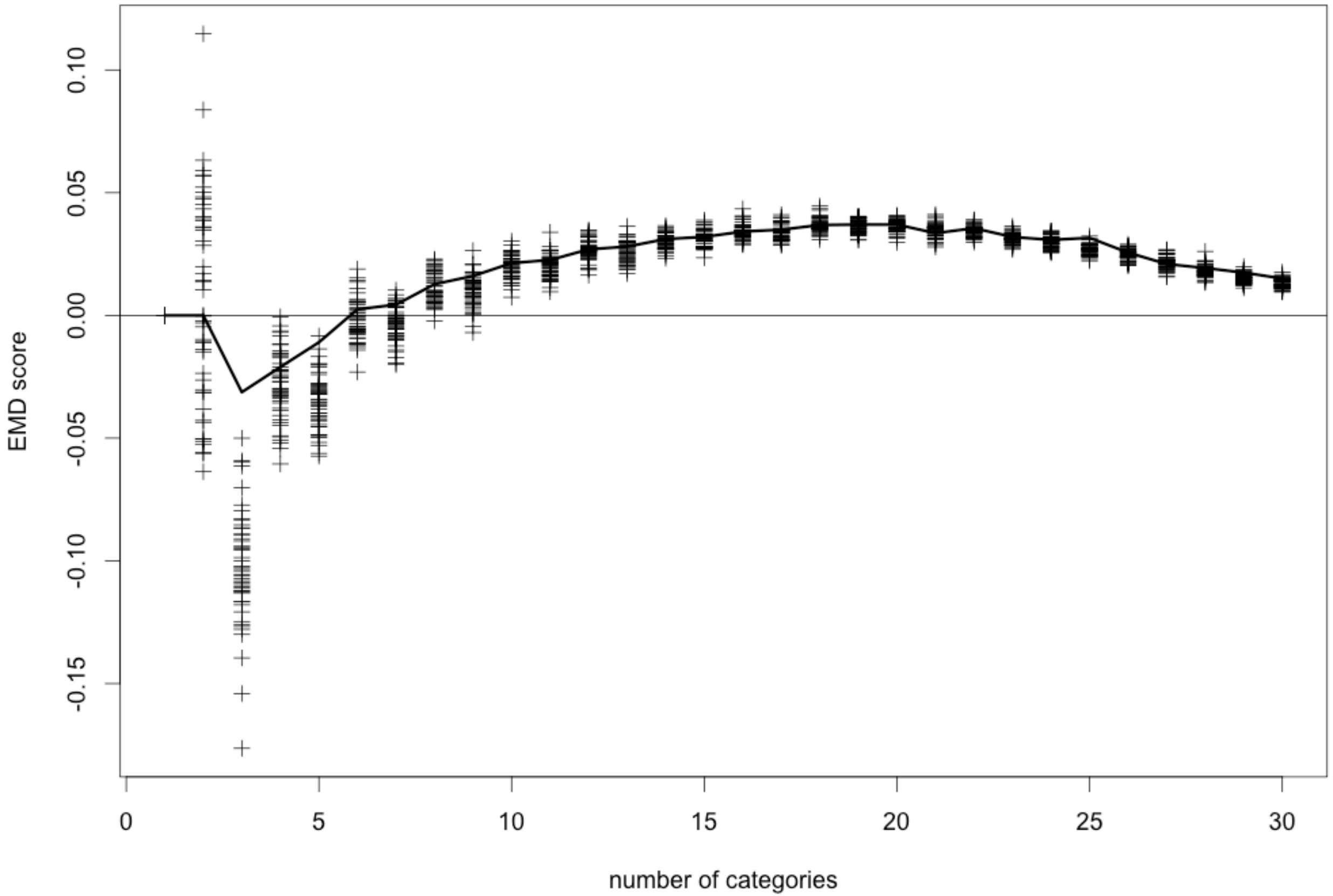
$$\mathbb{A} = (w_{\mathbb{A}})_{i,j}$$

$$\mathbb{B} = (w_{\mathbb{B}})_{i,j}$$

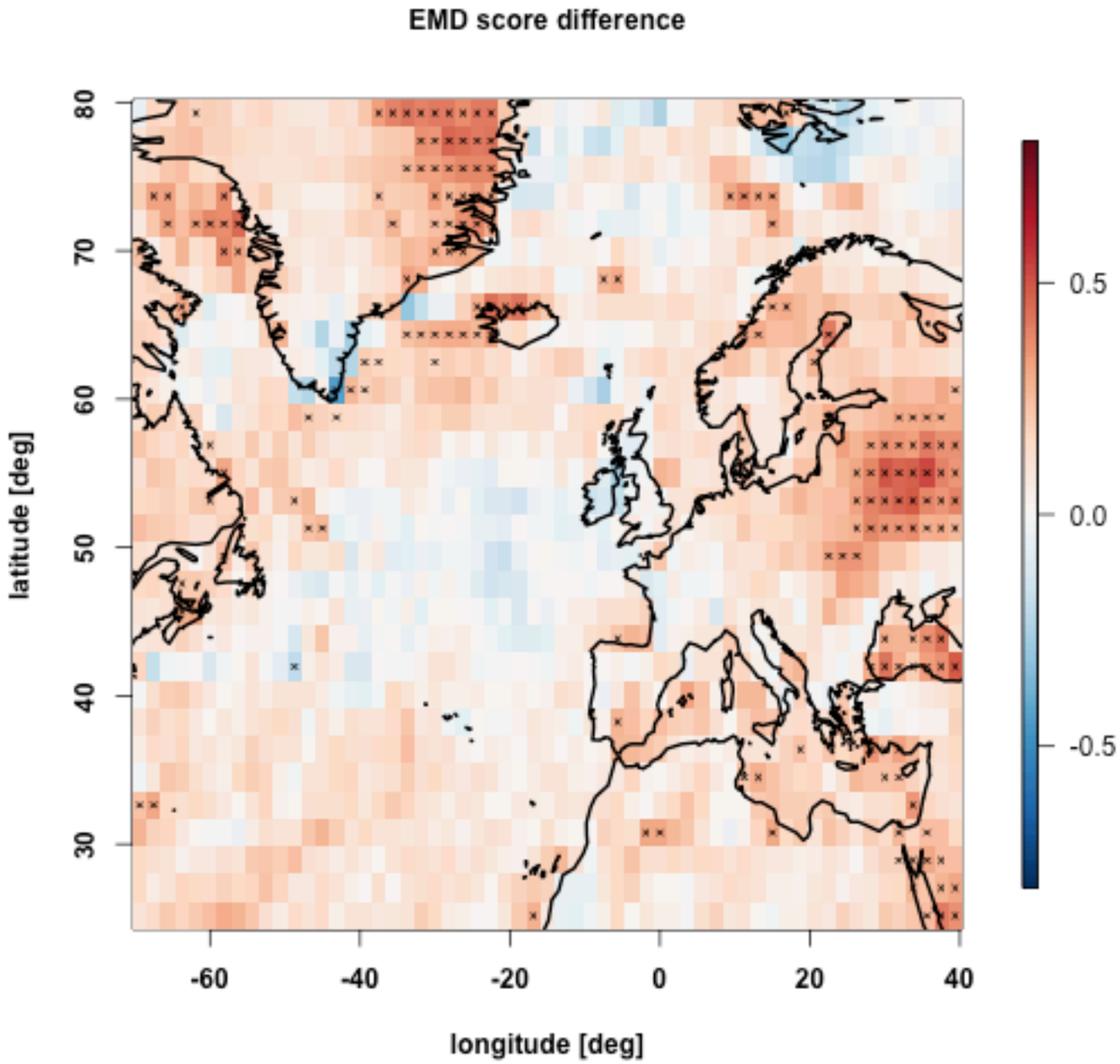
under the conditions:

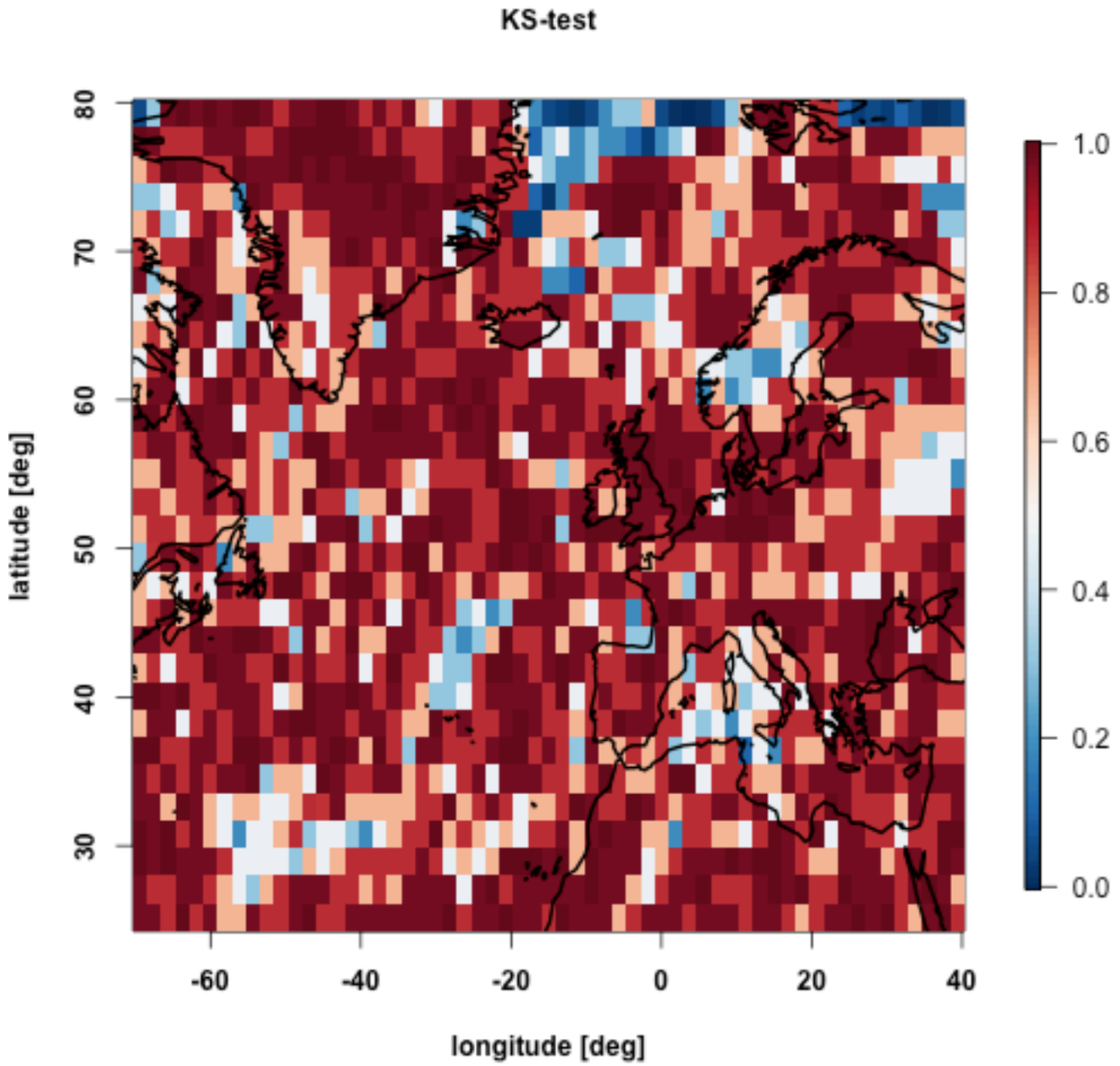
$$\sum_i f_{i,j} \leq w_{\mathbb{B}}(j) \qquad \sum_j f_{i,j} \leq w_{\mathbb{A}}(i))$$

$$\sum_i f_{i,j} = \min(w_{\mathbb{B}}(j), w_{\mathbb{A}}(i))$$

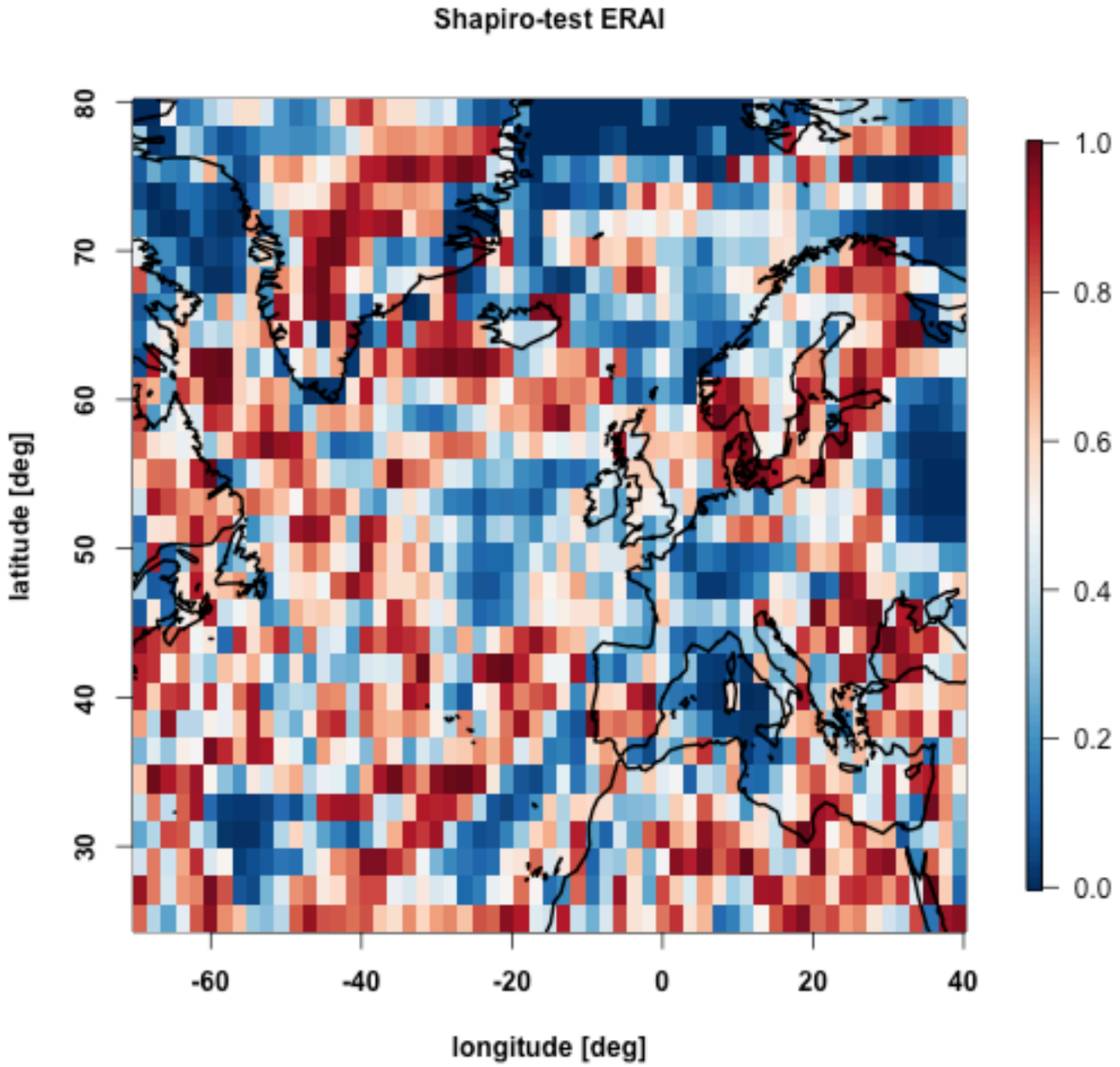

EMD score
0.10
0.05
0.00
-0.05
-0.10
-0.15
0
5
10
15
20
25
30
number of categories

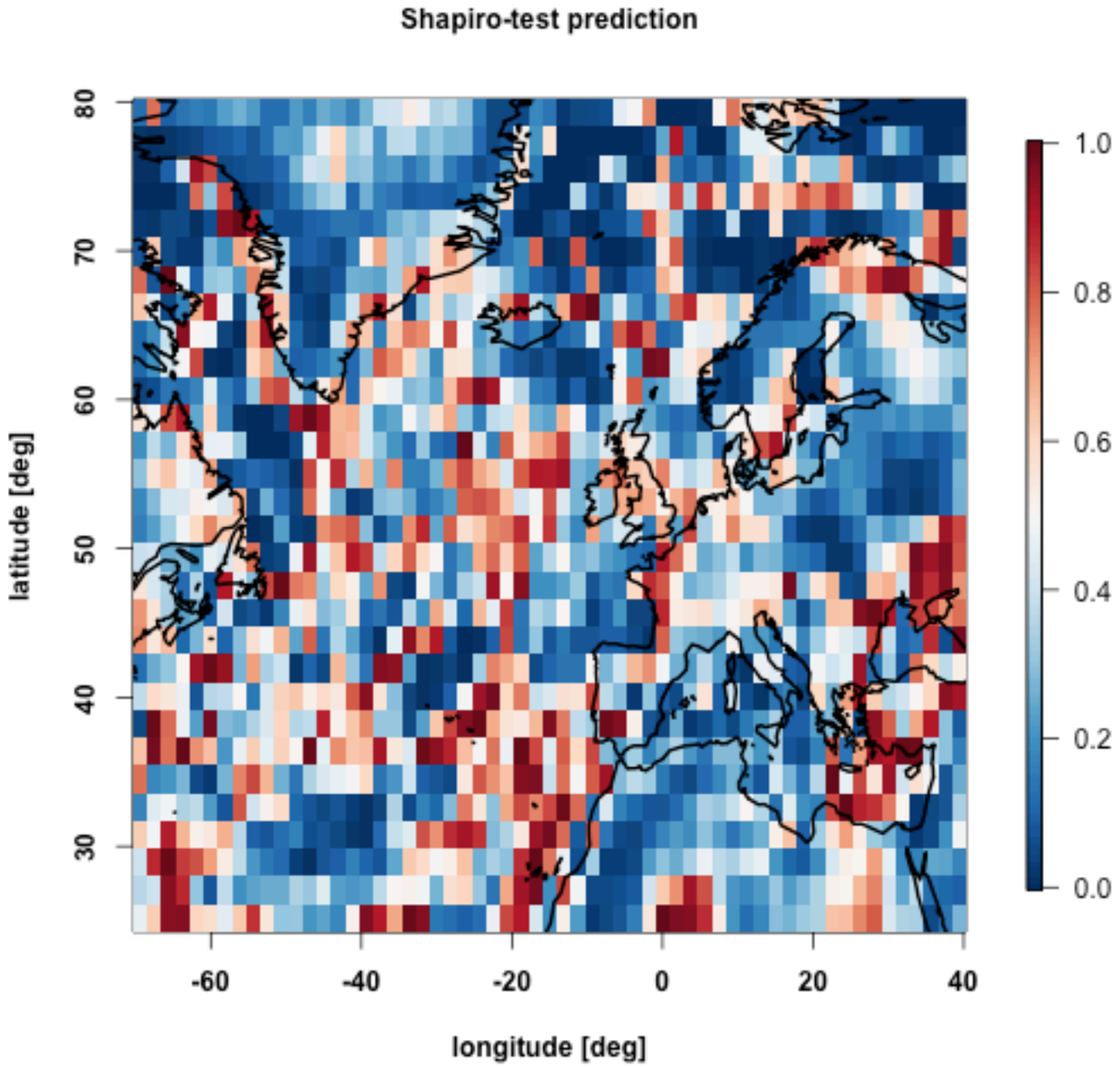
KS-Test


EMD score difference
80
70
60
50
40
30
latitude [deg]
-60
-40
-20
0
20
40
longitude [deg]
0.5
0.0
-0.5


KS-test
80
70
60
50
40
30
latitude [deg]
-60
-40
-20
0
20
40
longitude [deg]
1.0
0.8
0.6
0.4
0.2
0.0

Shapiro-Test



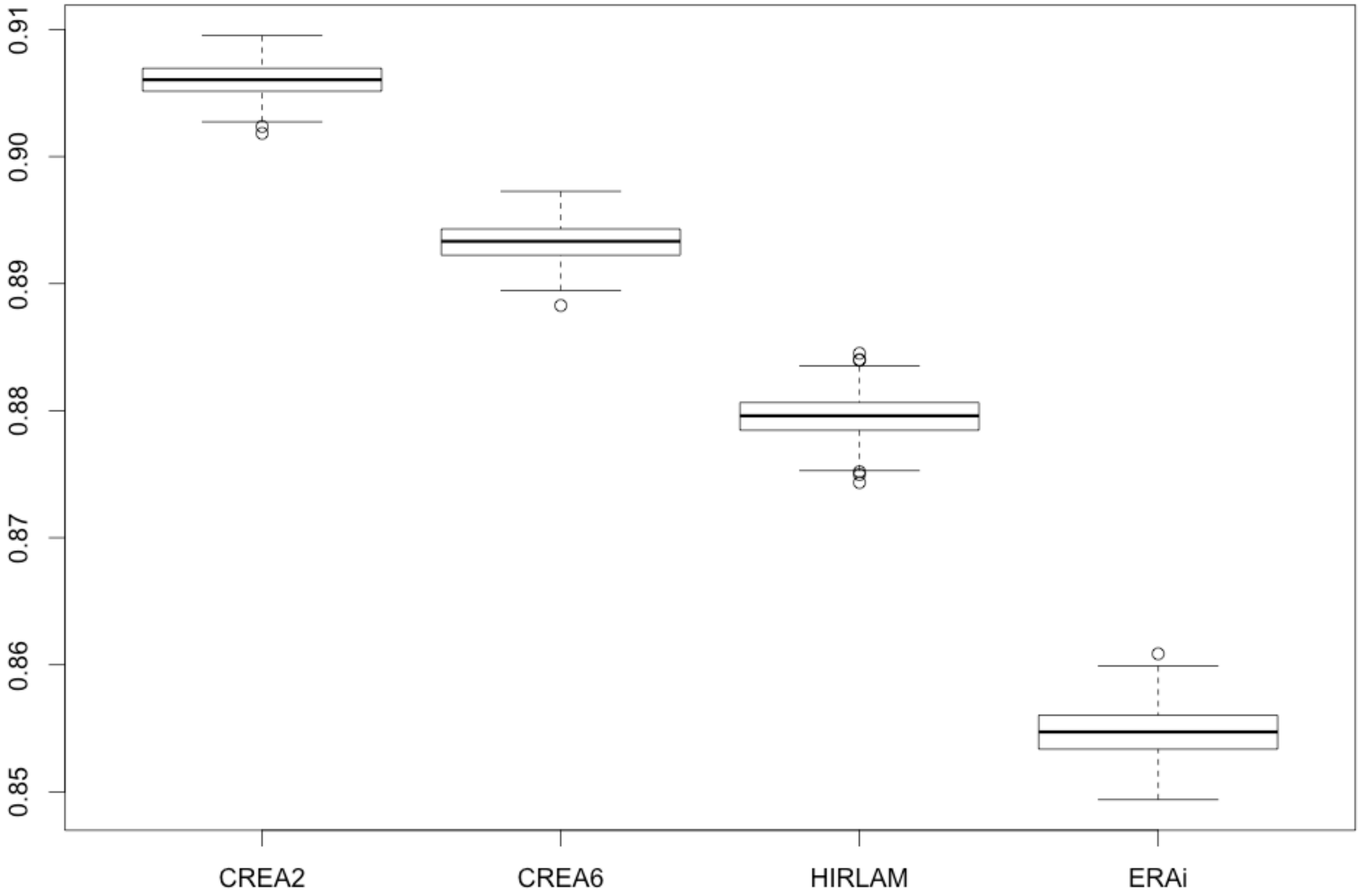


EMD-Score []

0.0
0.2
0.4
0.6
0.8
1.0

2
4
6
8
10

lead time [days]


0.91
0.90
0.89
0.88
0.87
0.86
0.85
CREA2
CREA6
HIRLAM
ERAi