cost 733

Harmonisation and applications of weather type Classifications for European Regions

Basic evaluation and comparison of circulation classifications from the COST733 database

Christoph Beck
Institute for Geography
University of Augsburg

christoph.beck@geo.uni-augsburg.de

Framework

Objectives

Data
&
Methods

Results

Conclusions

COST733 Action

„Harmonisation and applications of weather type classifications for European Regions“

Working Group 1
Existing methods and applications

Working Group 2
Implementation and development of weather types classification methods

Working Group 3
Comparison of selected weather type Classifications

Working Group 4
Testing methods for various applications

COST733 Action

„Harmonisation and applications
of weather type classifications for European Regions“

Working Group 3

Comparison of selected weather type Classifications

→ **Basic evaluation and comparison of circulation type classifications**

(According to the Memorandum of Understanding):

- **find or devise intercomparison tools**
  (for circulation type classifications)

- **statistical evaluation/comparison of classifications**
  (provided by WG2)

- **presentation and release of results**
  (to other WGs and scientific community)

- **recommend specifications for a new (common) method**
  (to WG2 together with WG4)

Basic evaluation/comparison of classifications

- estimate the performance of circulation type classifications in terms of their „discriminative power“
  - for varying climatic target variables (MSLP, Temp, Prec)
  - determine the relevance of the underlying basic methodological concepts
  - analyse the effect of varying classification “settings”

Basic evaluation/comparison of classifications

- estimate the performance of circulation type classifications in terms of their „discriminative power“
  - for varying climatic target variables (MSLP, Temp, Prec)
  - determine the relevance of the underlying basic methodological concepts
  - analyse the effect of varying classification “settings”

- are there superior individual methods?
- are there superior methodological concepts?
- are there superior classification settings?

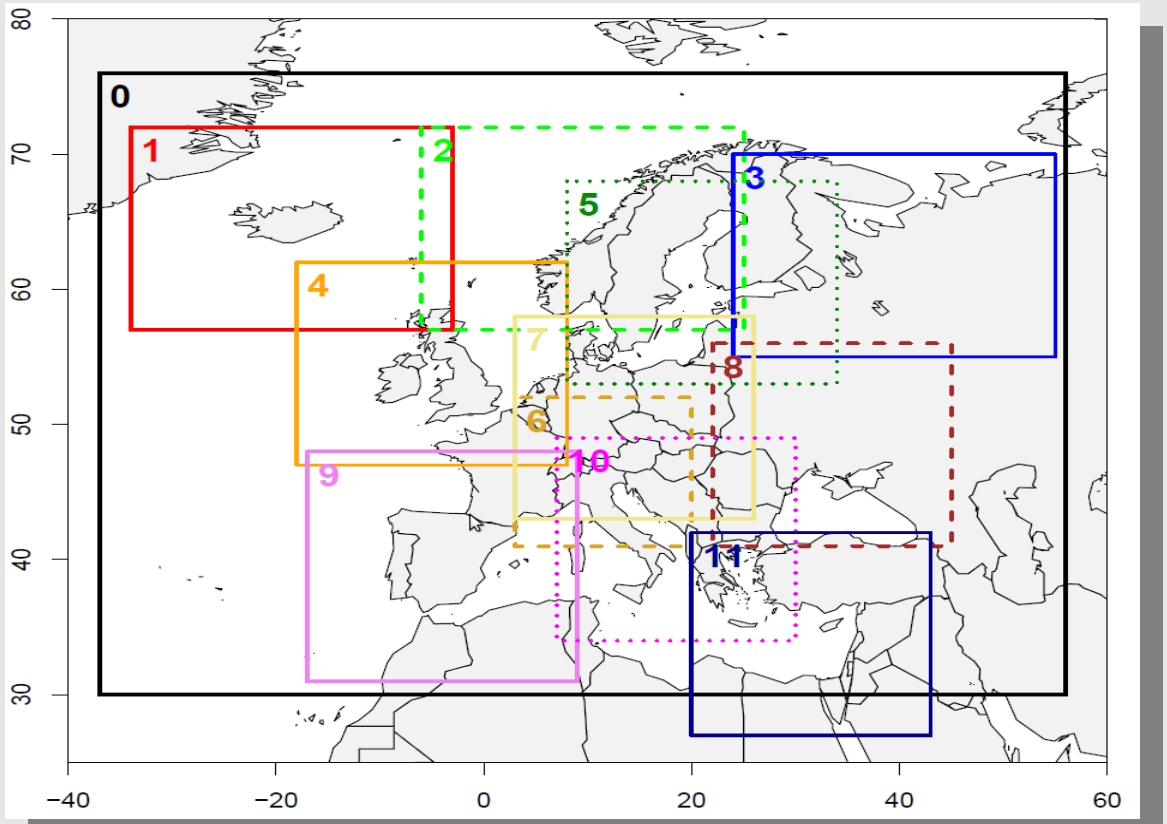
1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

- **for 12 spatial domains**



1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

- **for 12 spatial domains**
- **using ERA40 data**
- **for fixed numbers of types (9, 18, 27)**
- **using MSLP**
- **for single days**
- **on an annual basis**

1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

- **~ 20 methods, from 5 basic method groups**

Method Group	Specific Method					
SUBjective	Hess-Brezowsky	Peczely	Perret	Schueepp	ZAMG	
THReshold	GWT	JCT/LWT	LIT	WLK		
PCA	KRZ	PCT	PTT	PXE		
LeaDeR	ERP	KIR	LND	PTS		
OPTimization	CAP	CKM	NNW	PXK	SAN	SOM

1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

- **for 12 spatial domains**
- **using ERA40 data**
- **for fixed numbers of types (9, 18, 27)**
- **using MSLP** / **using additional variables**
- **for single days** / **for 4-day sequences**
- **on an annual basis** / **on a seasonal basis**

1) Circulation type classifications from the cost733cat 2.0 data base

~ 5000 classifications (provided by WG2)

enable the comparison of:

- **Methods**
  - is LUND better than KIRCHHOFER ?
- **Method Groups**
  - is OPT better than PCA ?
- **Settings**
  - is MSLP + GPH 500 better than MSLP alone ?

2) Climatic target variable data from ERA40

(Uppala et al., 2005)

- daily mean SLP (MSLP)
- daily mean 2-metre Temperature (2mT)
- daily precipitation sum (PREC)

- for the period 1957 - 2002

3) Evaluation criteria

- Explained Variation $EV = 1 - \frac{WSS}{TSS}$

- Pseudo-F statistic $PF = 1 - \frac{BSS/(k-1)}{WSS/(n-k)}$

- Within-type Standard Deviation $WSD = \sqrt{\frac{\sum_{k=1}^{K}(n_k - 1)\cdot SDI_k^2}{\sum_{k=1}^{K}(n_k - 1)}}$

- Pattern Correlation Ratio $PCR = \frac{PCI}{PCO}$

- (Fast)Silhouette Index $FSIL = \frac{1}{n}\sum_{i=1}^{n}\frac{fb_i - fa_i}{\max(fa_i, fb_i)}$

- Confidence Interval of the Mean $CIM = \frac{\sum_{k=1}^{K} z_{1-\alpha/2}\cdot\frac{SDI_k}{\sqrt{n_k}}\cdot n_k}{\sum_{k=1}^{K} n_k}$

SDI = standard deviation within class
PCI = mean pattern correlation within classes
PCO = mean pattern correlation between classes
fa_i = distance between case i and its own class centroid
fb_i = distance between case i and its nearest class centroid

TSS = total sum of squares
WSS = sum of squares within classes
BSS = sum of squares between classes
k = number of classes (types)
n = number of cases

4) Determination of Evaluation indices

- for gridded SLP-, Temperature- and Precipitation data

- for 12 different domains

- for individual months, seasons and the whole year

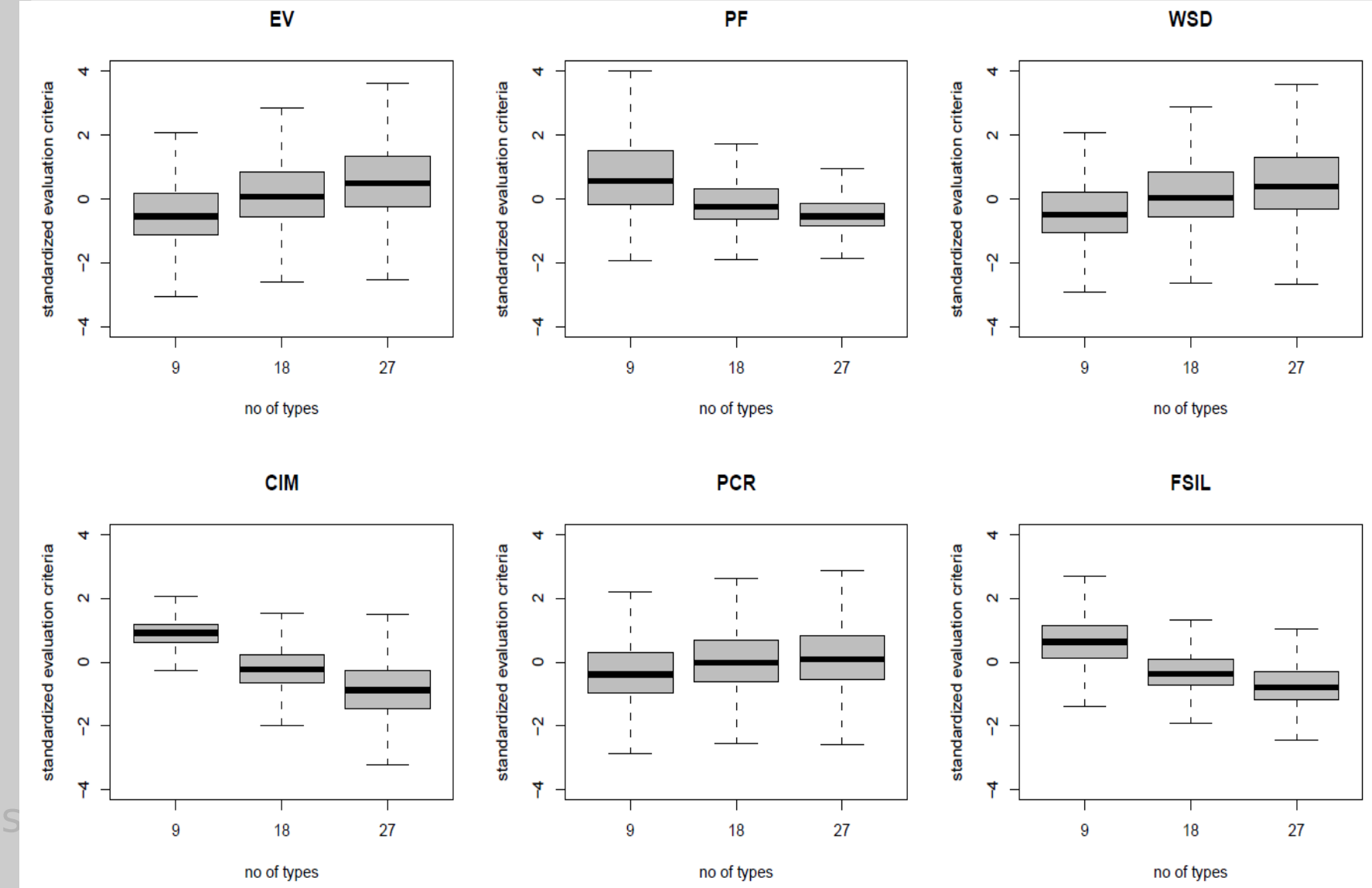
- for individual grid points / the whole gridded field

→ ~ 4000 „performance index“ samples

Aggregation of evaluation indices for varying groupings of classifications
(e.g. sequential classifications – non-sequential classifications)

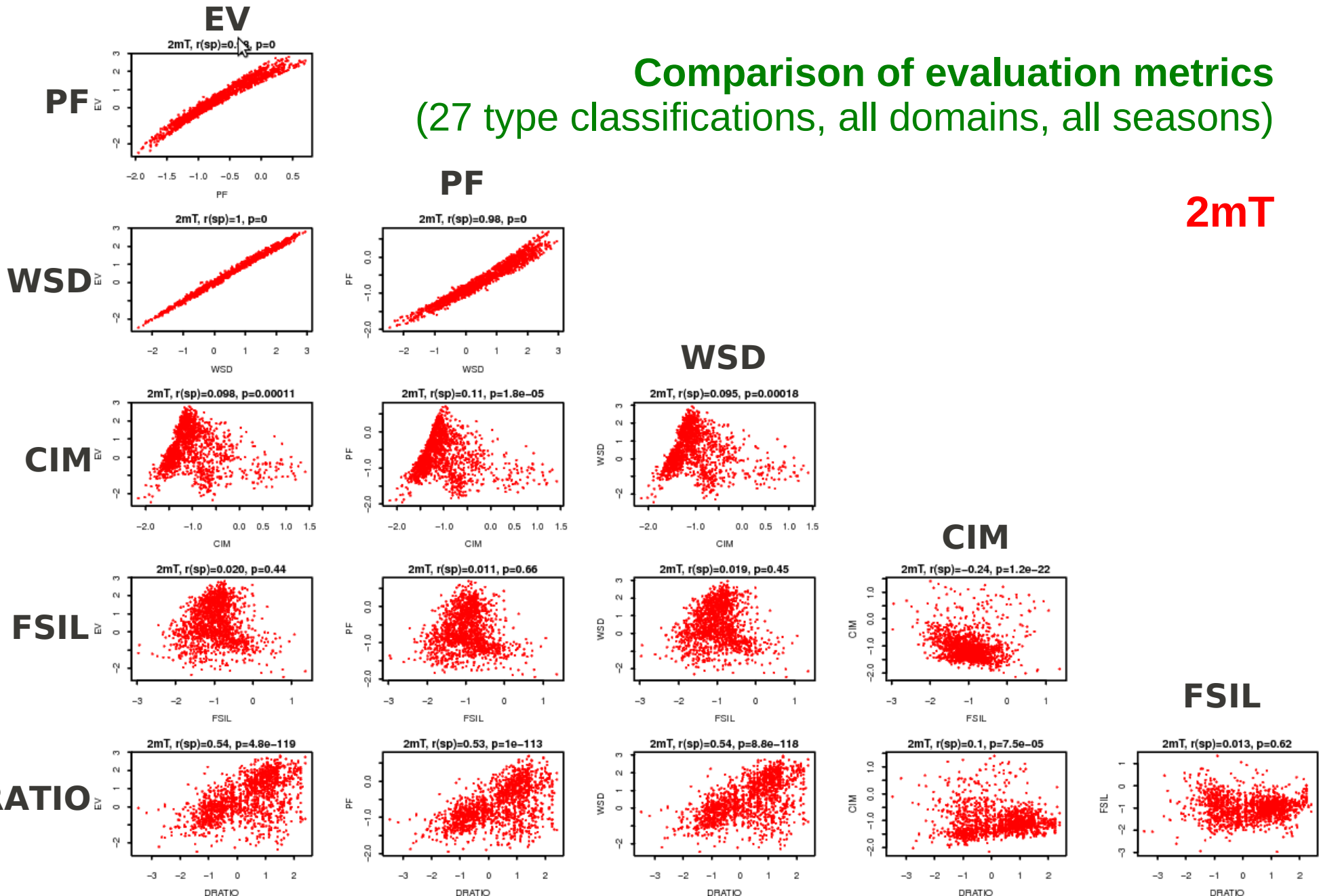
→ relevance of basic methodological approach / settings
(e.g. effect of classifying 4-day sequences)

Relevance of the number of types



Comparison of evaluation metrics
(27 type classifications, all domains, all seasons)

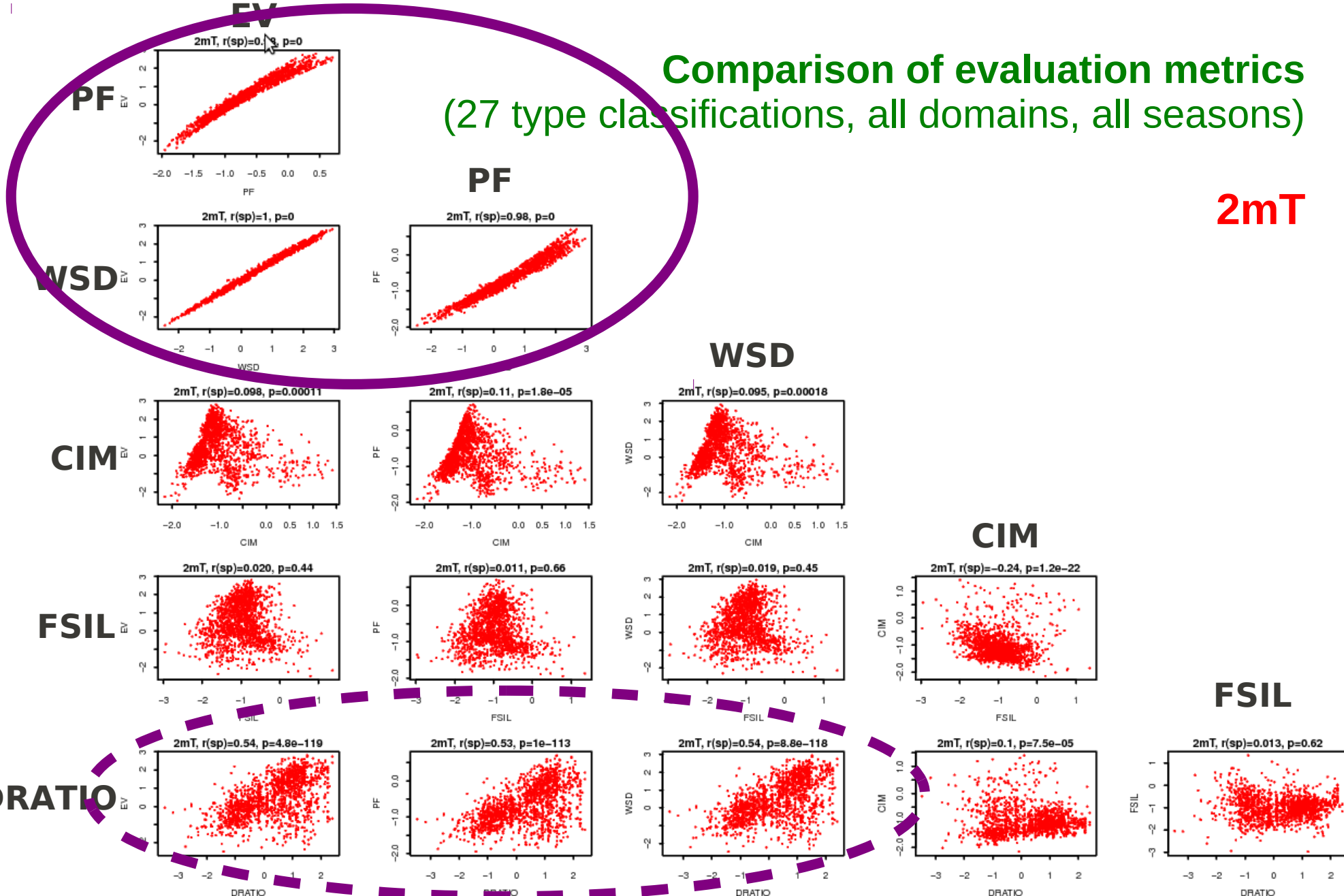
2mT



Comparison of evaluation metrics

(27 type classifications, all domains, all seasons)

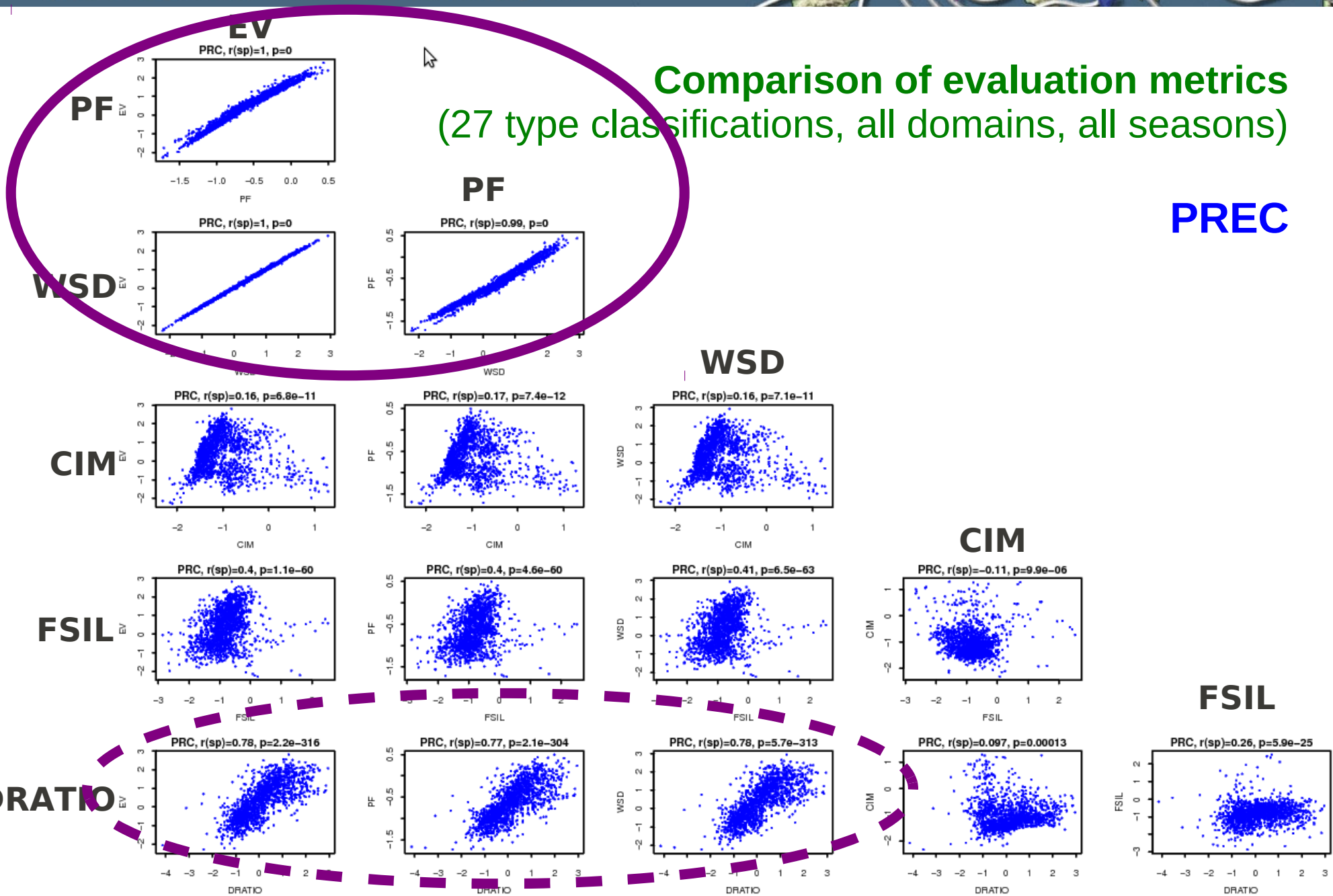
2mT



Comparison of evaluation metrics

(27 type classifications, all domains, all seasons)

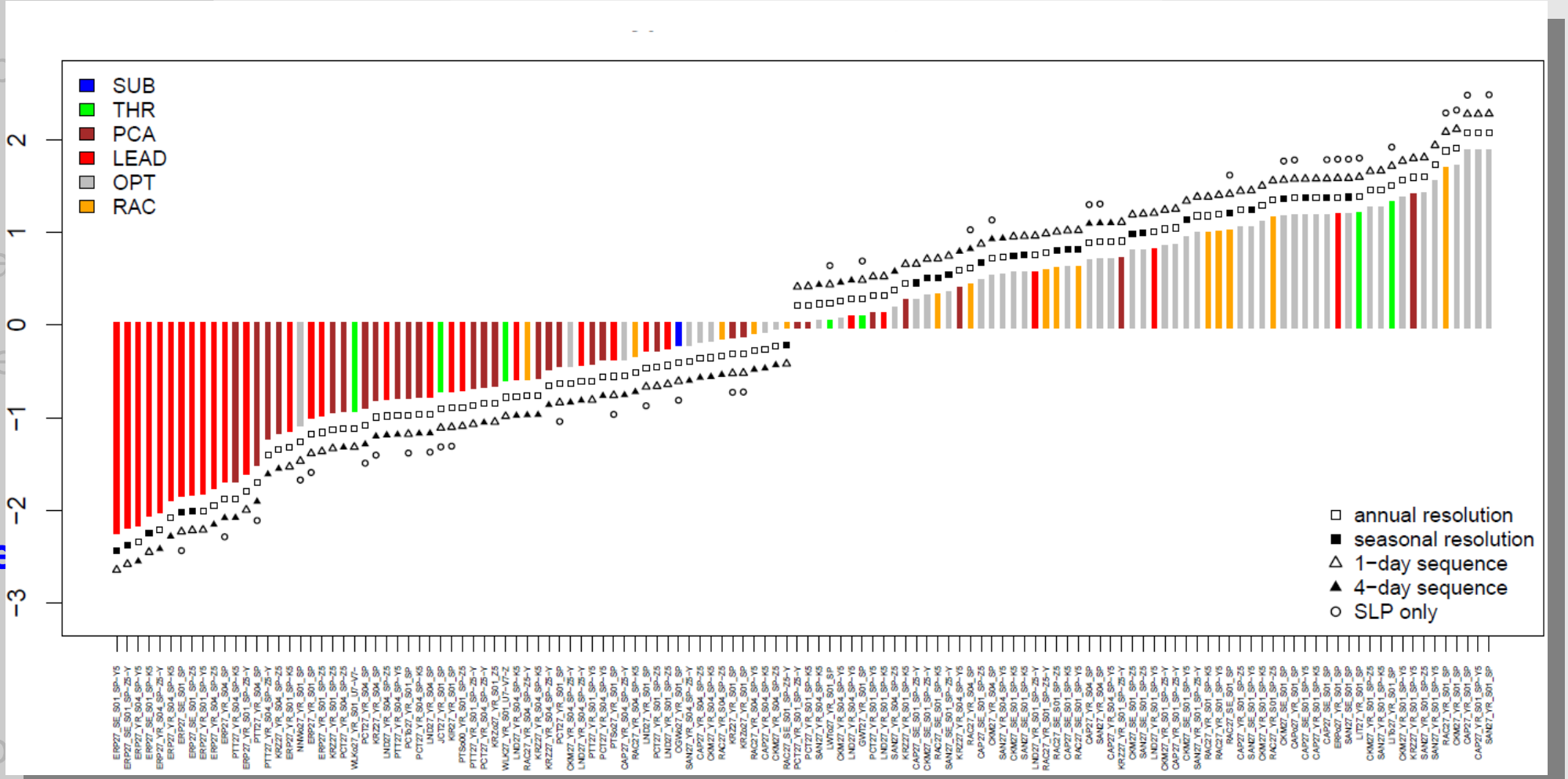
PREC



Framework

Comparison / “ranking” of individual methods

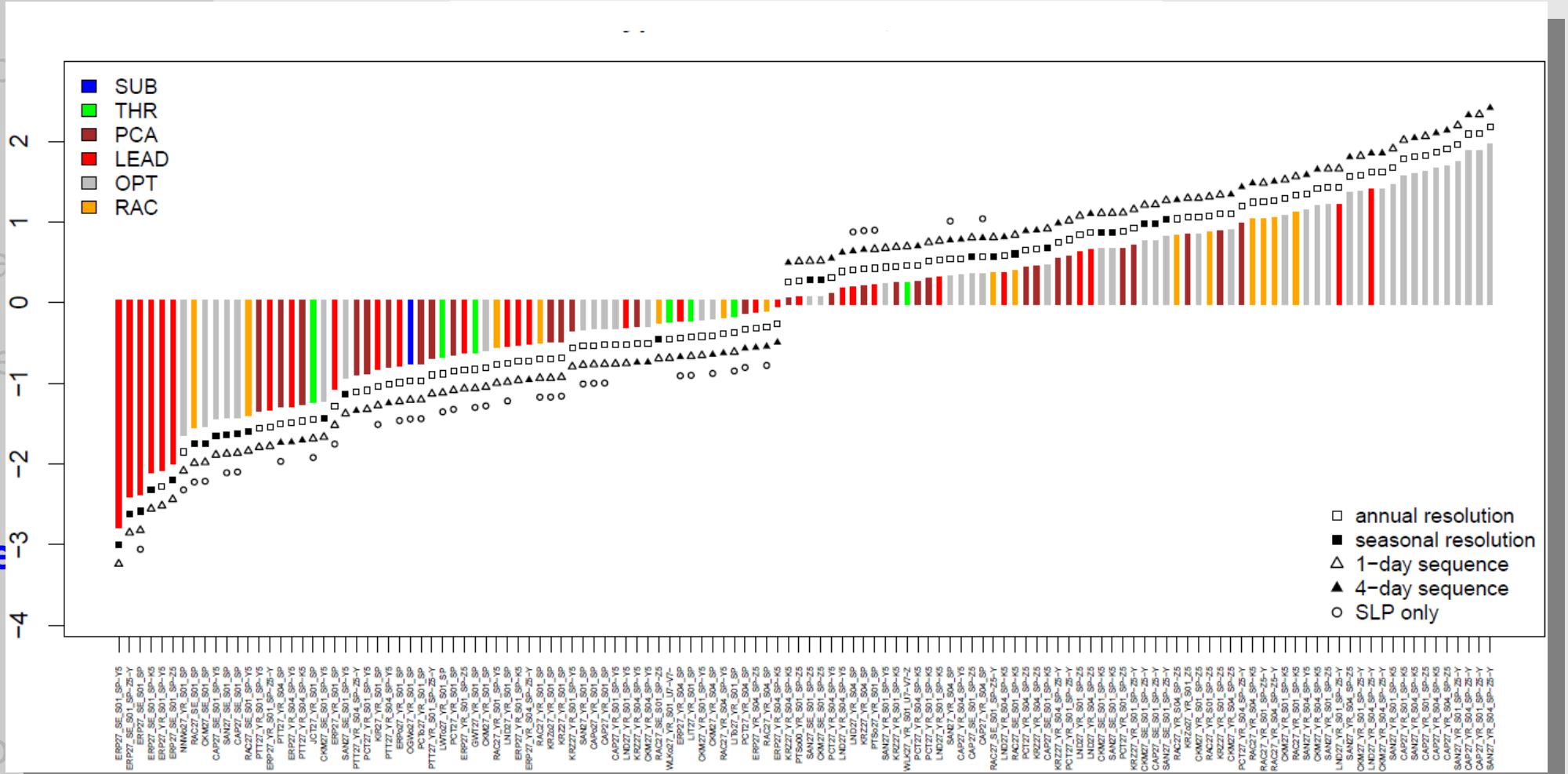
27 types, over all domains and all seasons, EV for MSLP



Framework

Comparison / “ranking” of individual methods

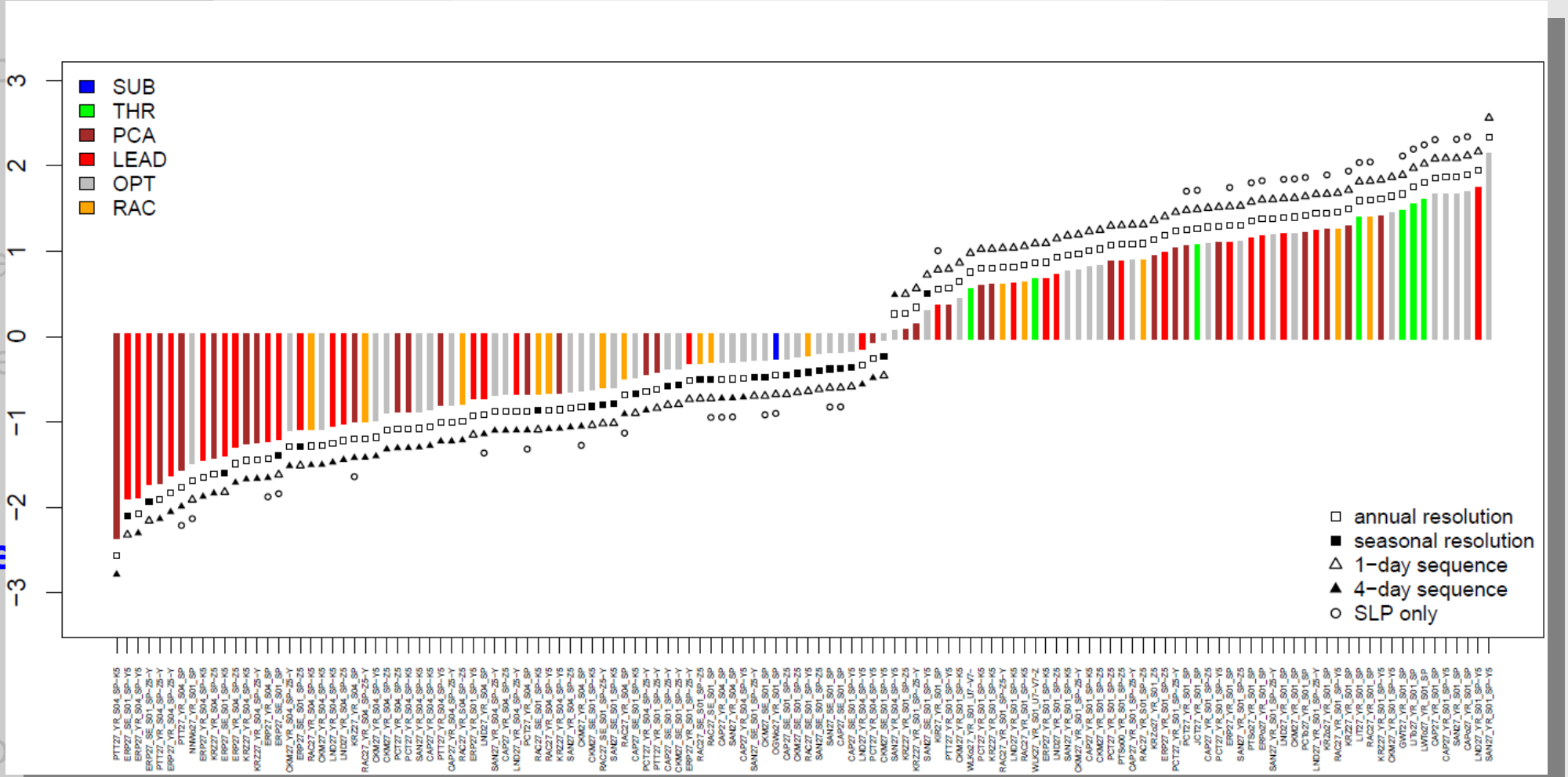
27 types, over all domains and all seasons, EV for 2mT



Framework

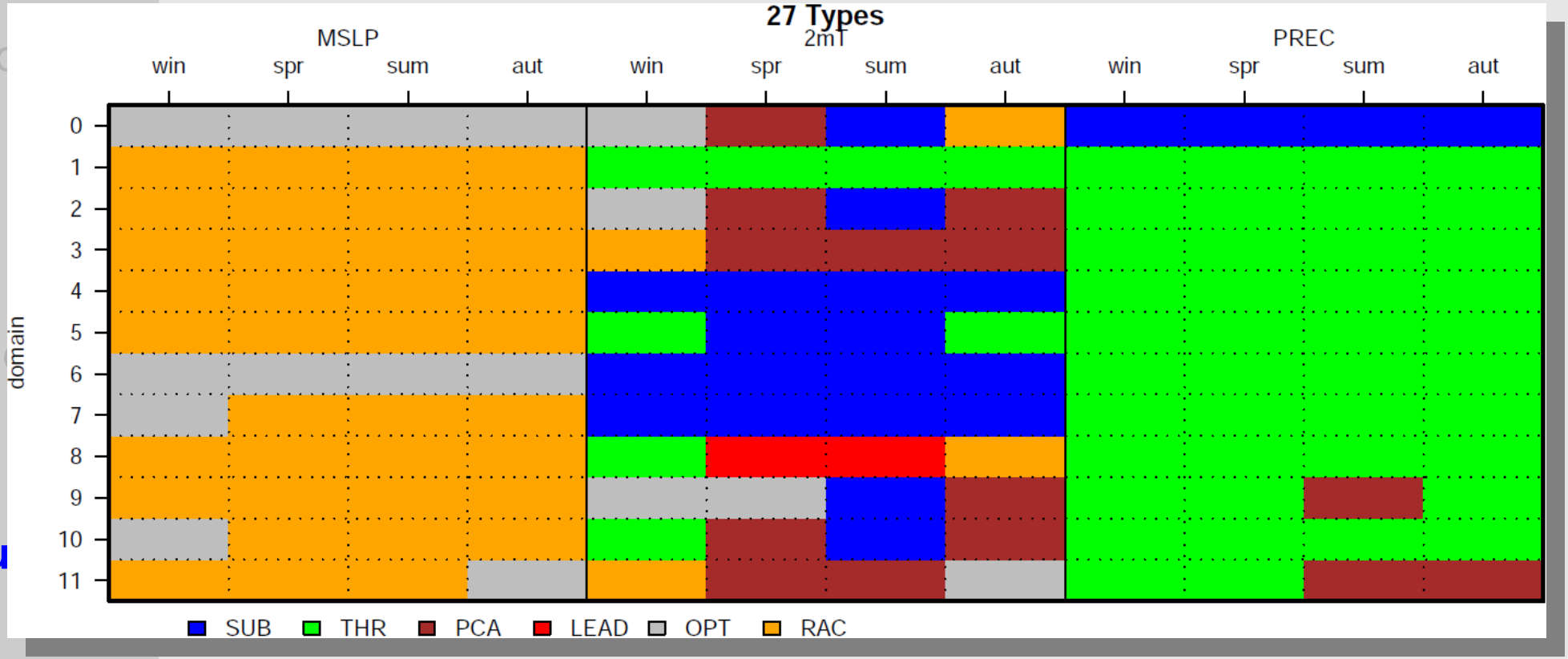
Comparison / “ranking” of individual methods

27 types, over all domains and all seasons, EV for PREC



Relevance of the basic methodological approach

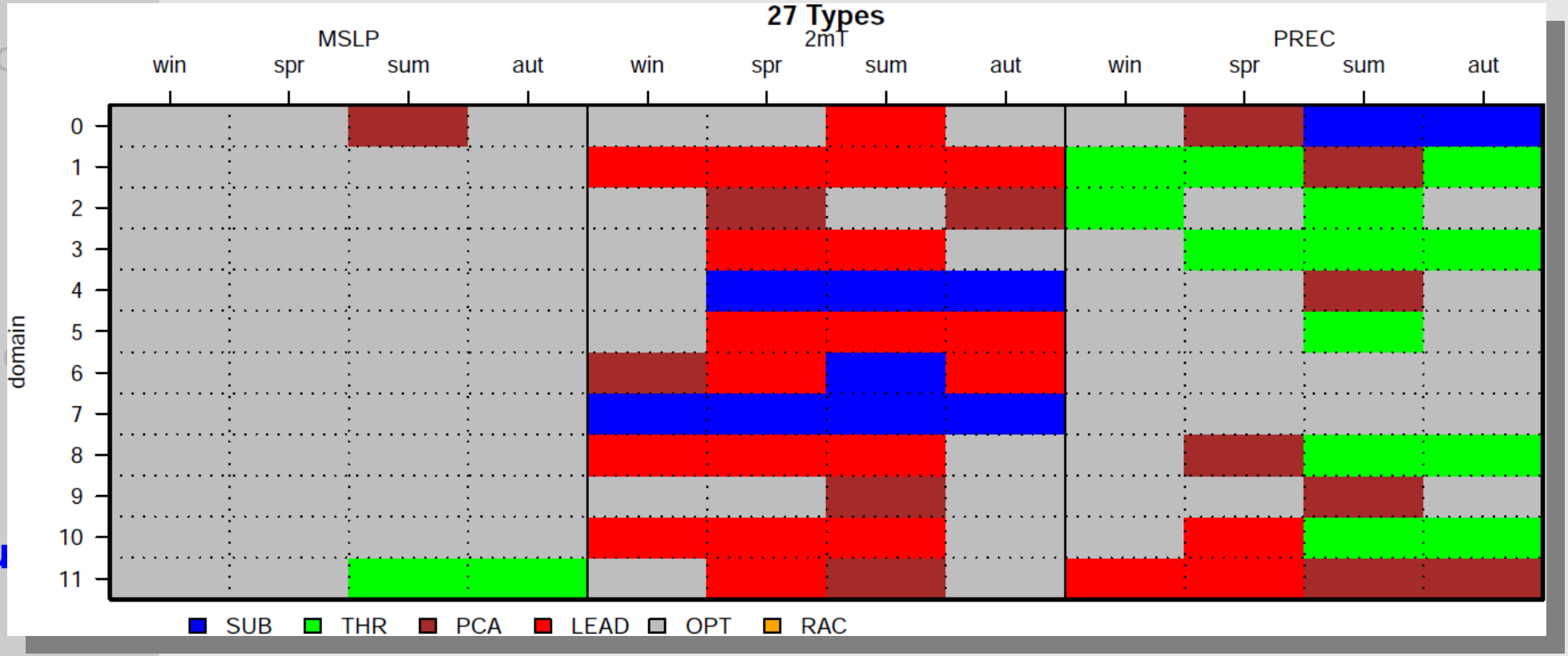
Method group with highest mean EV



SUB = Subjective methods
THR = Threshold based methods
PCA = PCA based methods
LDR = Leader based methods
OPT = Optimisation methods
RAC = Random centroid methods

Relevance of the basic methodological approach

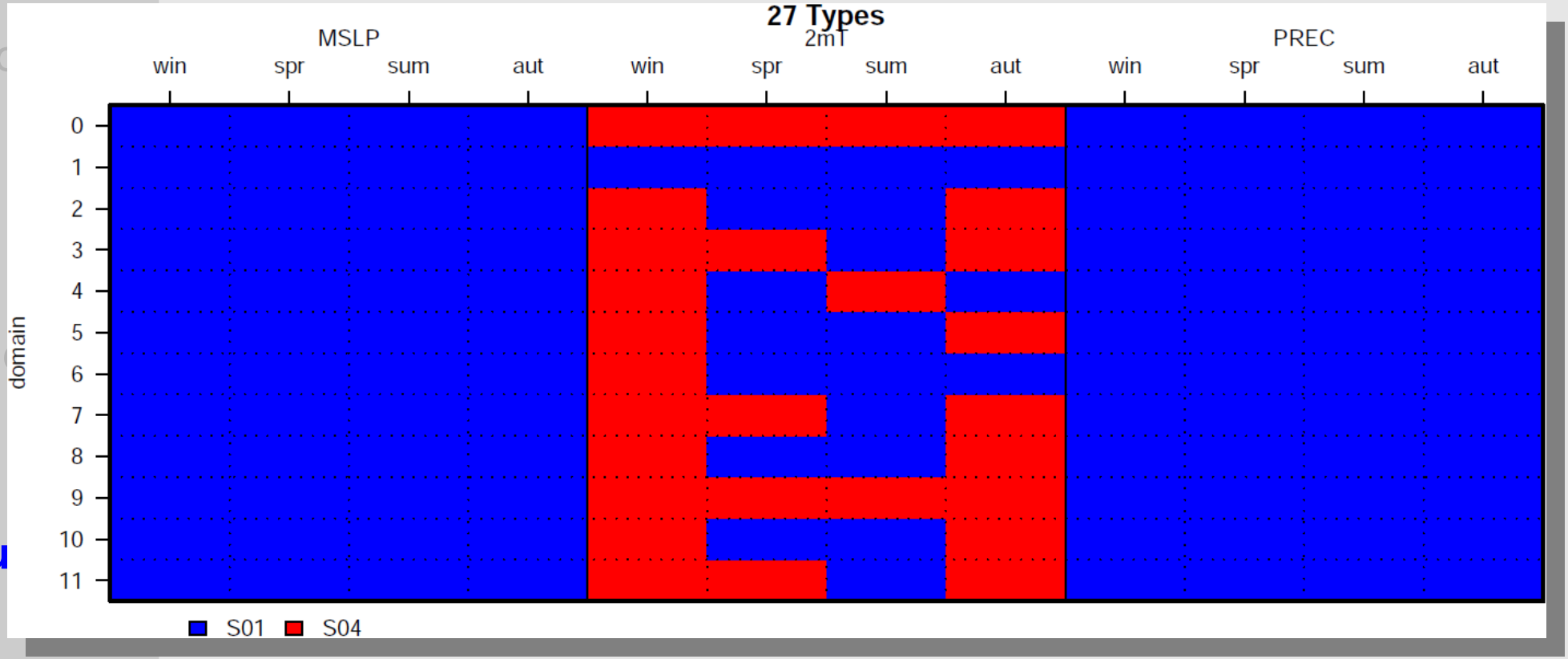
Method affiliation of best performing CTC



SUB = Subjective methods
THR = Threshold based methods
PCA = PCA based methods
LDR = Leader based methods
OPT = Optimisation methods
RAC = Random centroid methods

Relevance of sequence length

Sequence length of best performing CTC

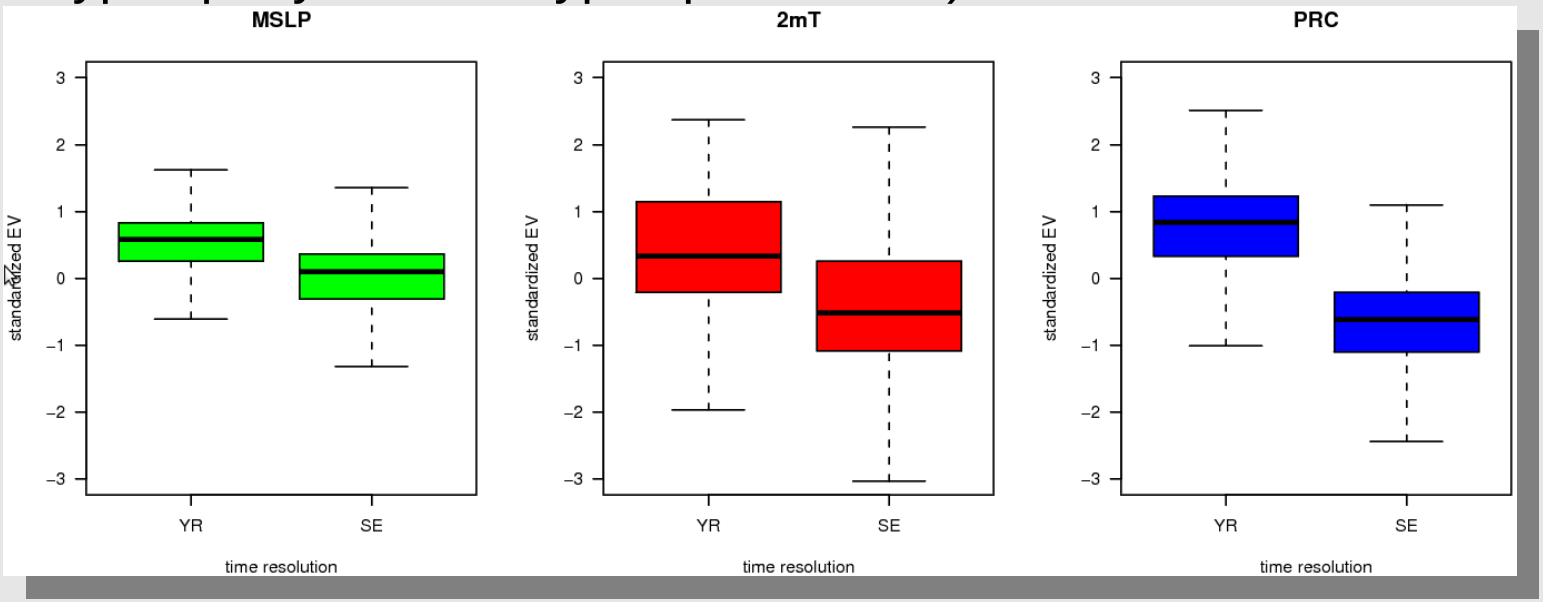


S01 = 1-day sequence

S04 = 4-day sequence

Relevance of seasonality

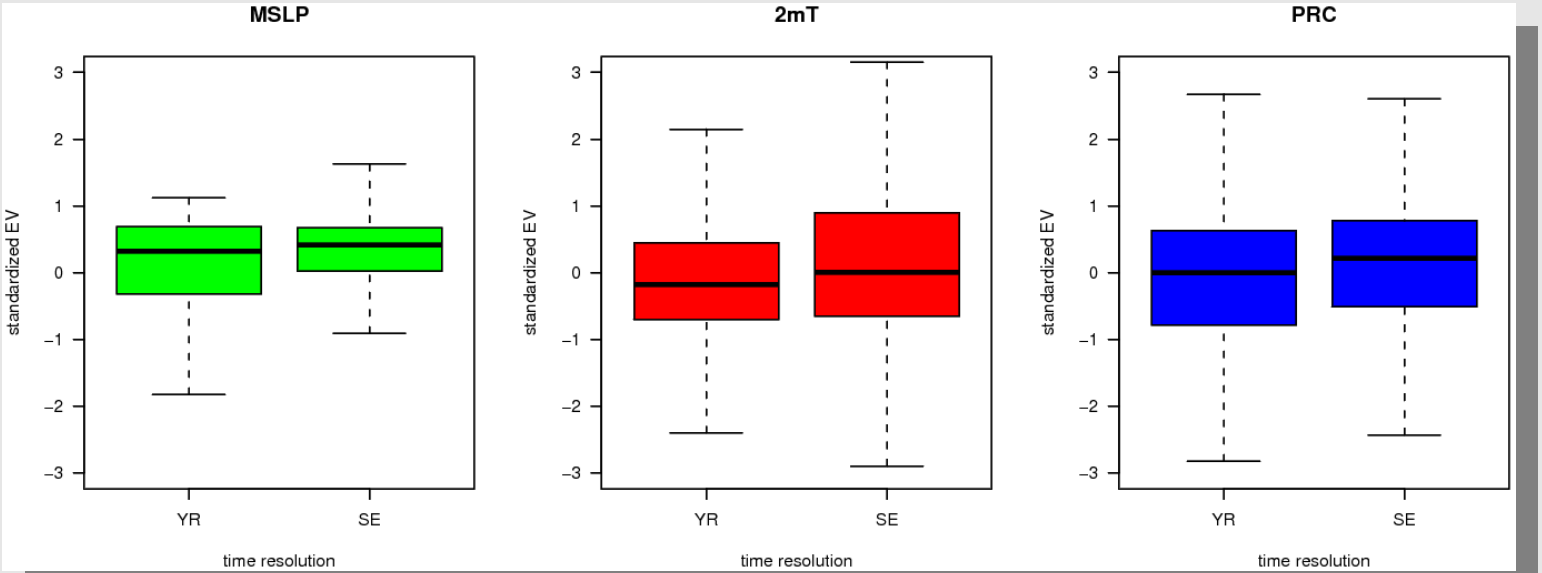
Mean EV over all CTCs performed on an annual or seasonal basis
(**27** types per year vs. **7** types per season)



YR = annual classification **SE = seasonal classification**

Relevance of seasonality

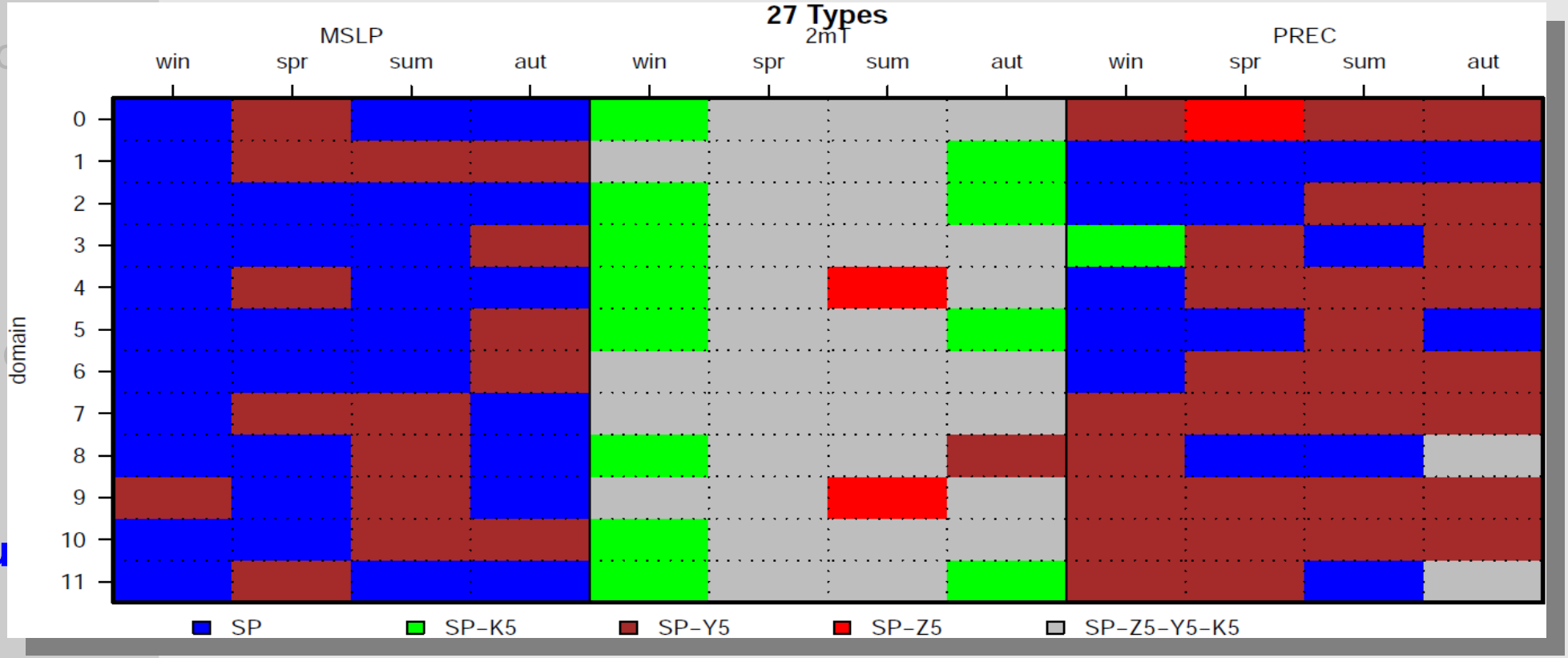
Mean EV over all CTCs performed on an annual or seasonal basis
(**9** types per year vs. **7** types per season)



YR = annual classification **SE = seasonal classification**

Relevance of additional input variables

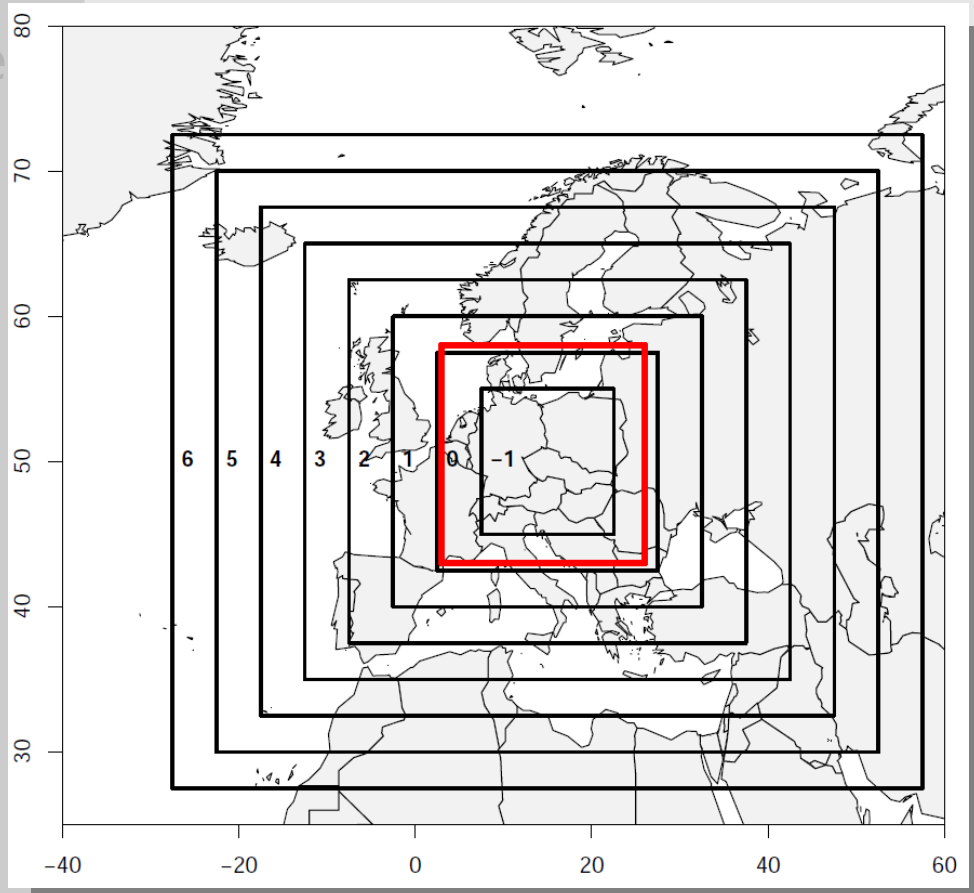
Input variable(s) used by the best performing CTC



SP = MSLP
Y5 = 500hPa vorticity

K5 = 1000/500hPa thickness
Z5 = 500hPa height

Relevance of domain size



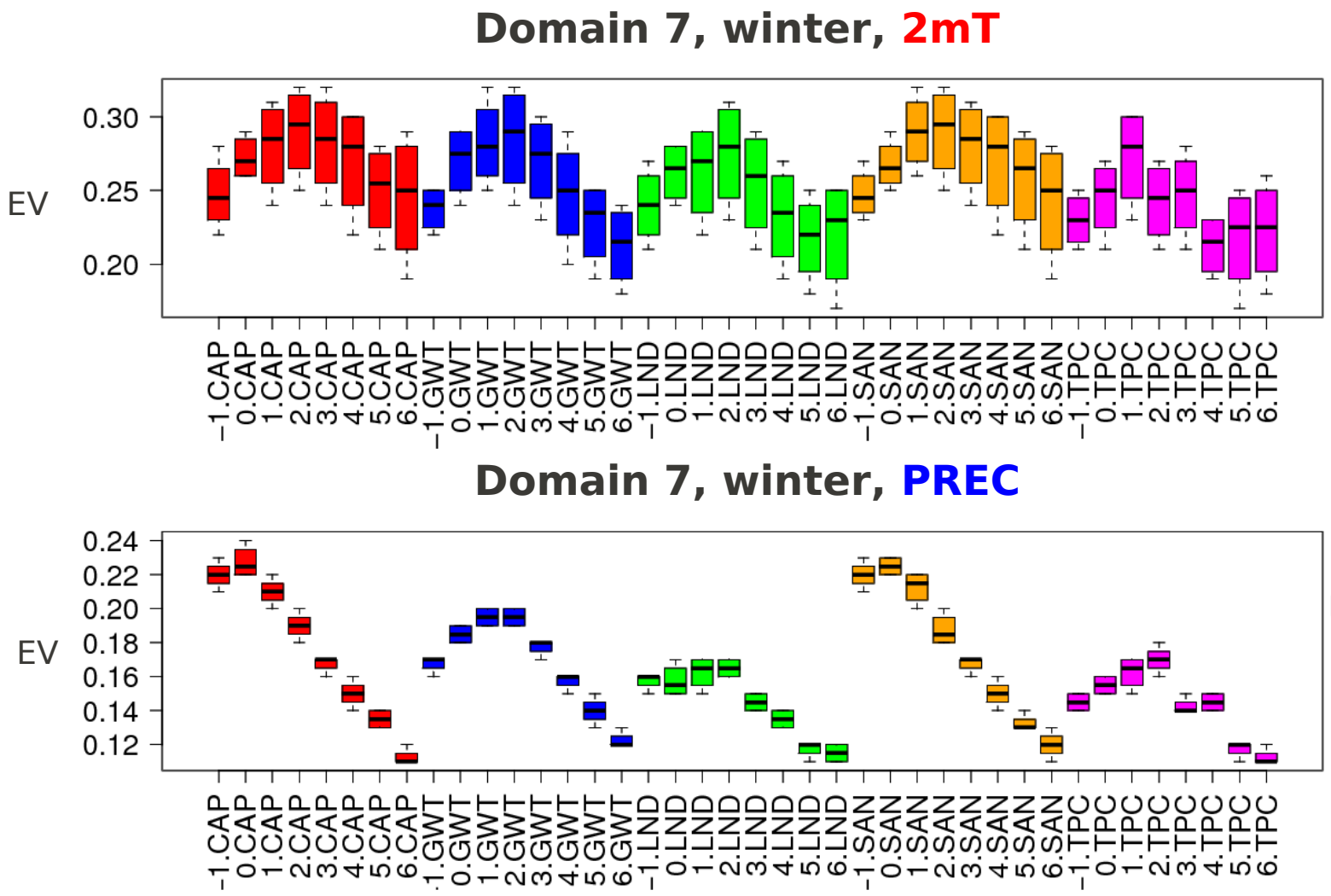
1) **CTCs for varying spatial domains centered over Central Europe** (from largest domain 6 to smallest domain -1)

2) **Evaluation of each CTC for COST733 domain 7** („original size“ domain 0)

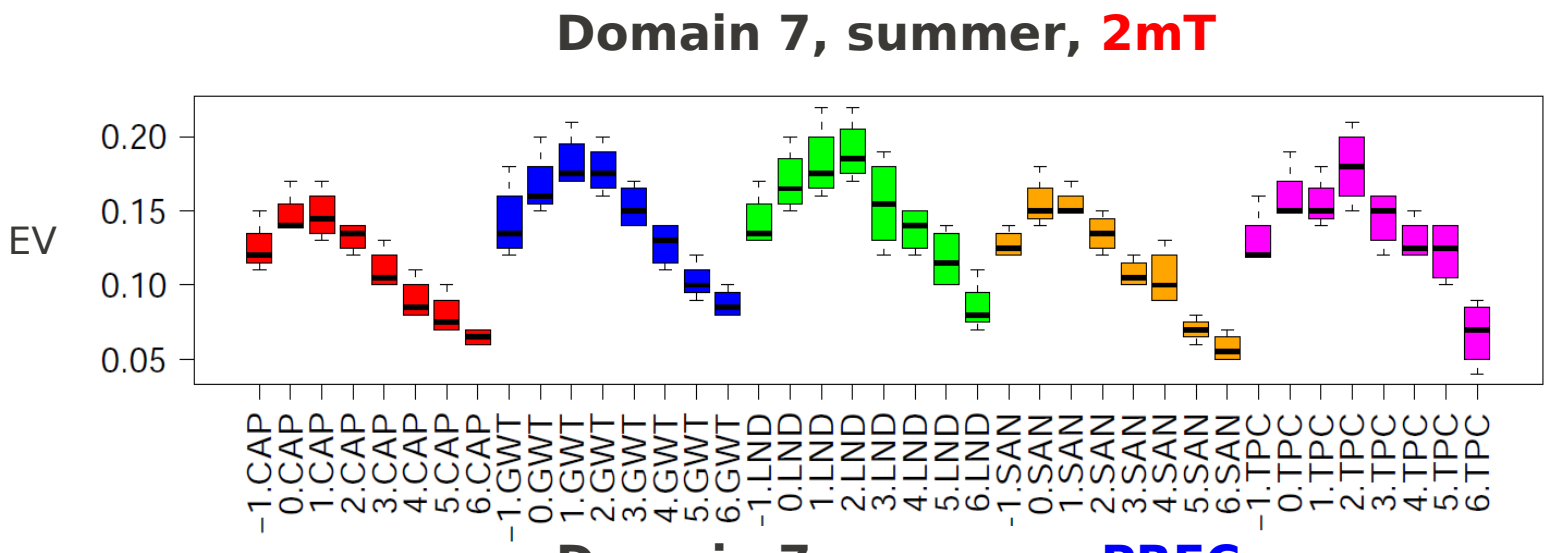
Only for a subsample of CTCs:

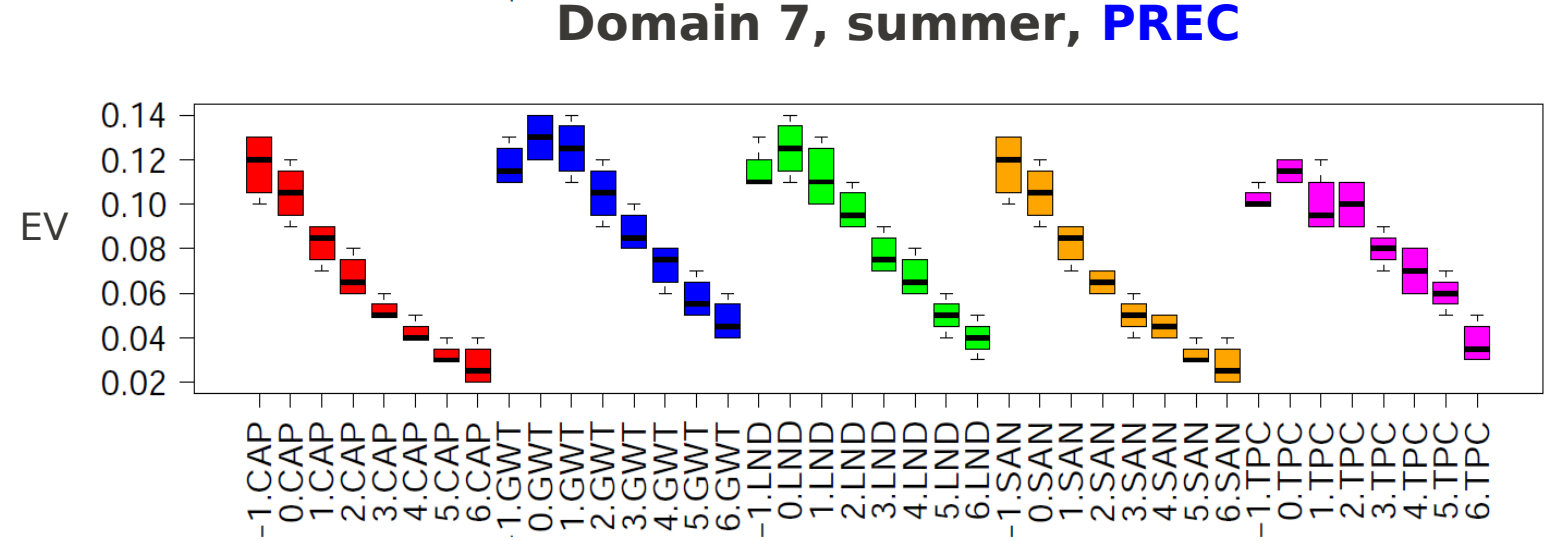
- CAP
- GWT
- LND
- SAN
- TPC

Relevance of domain size



Relevance of domain size





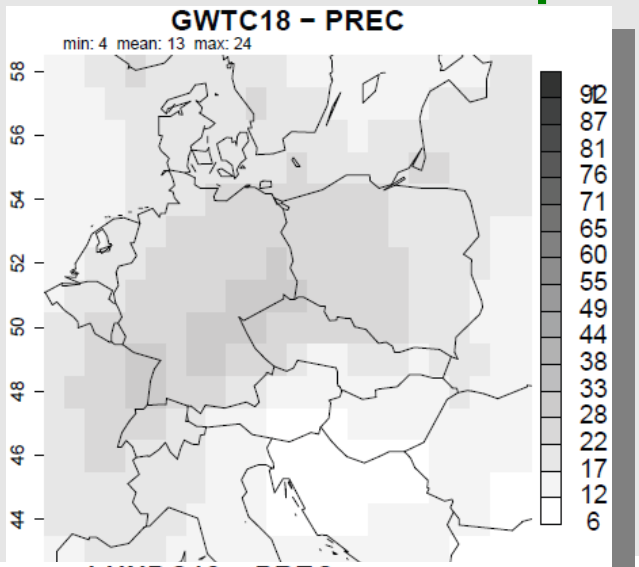
Performance variations within spatial domains

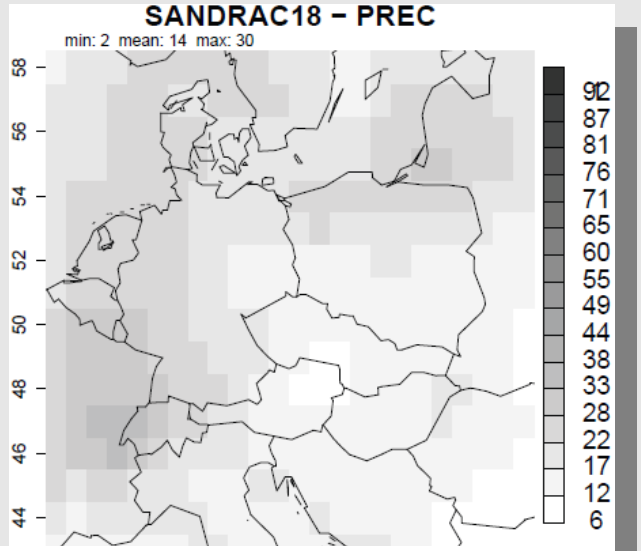
Domain 7

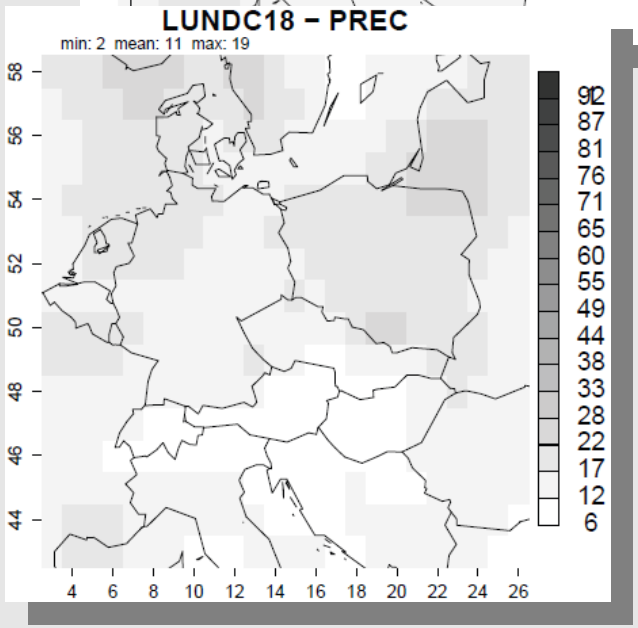
winter

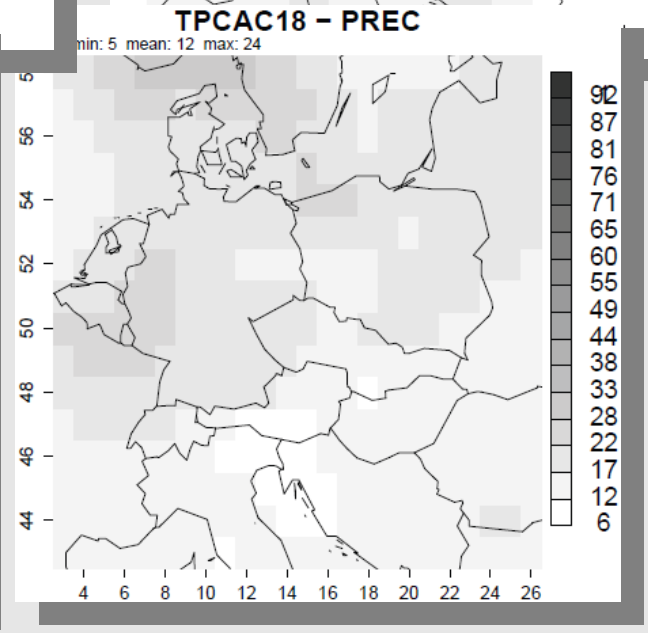
PREC

EV at gridboxes









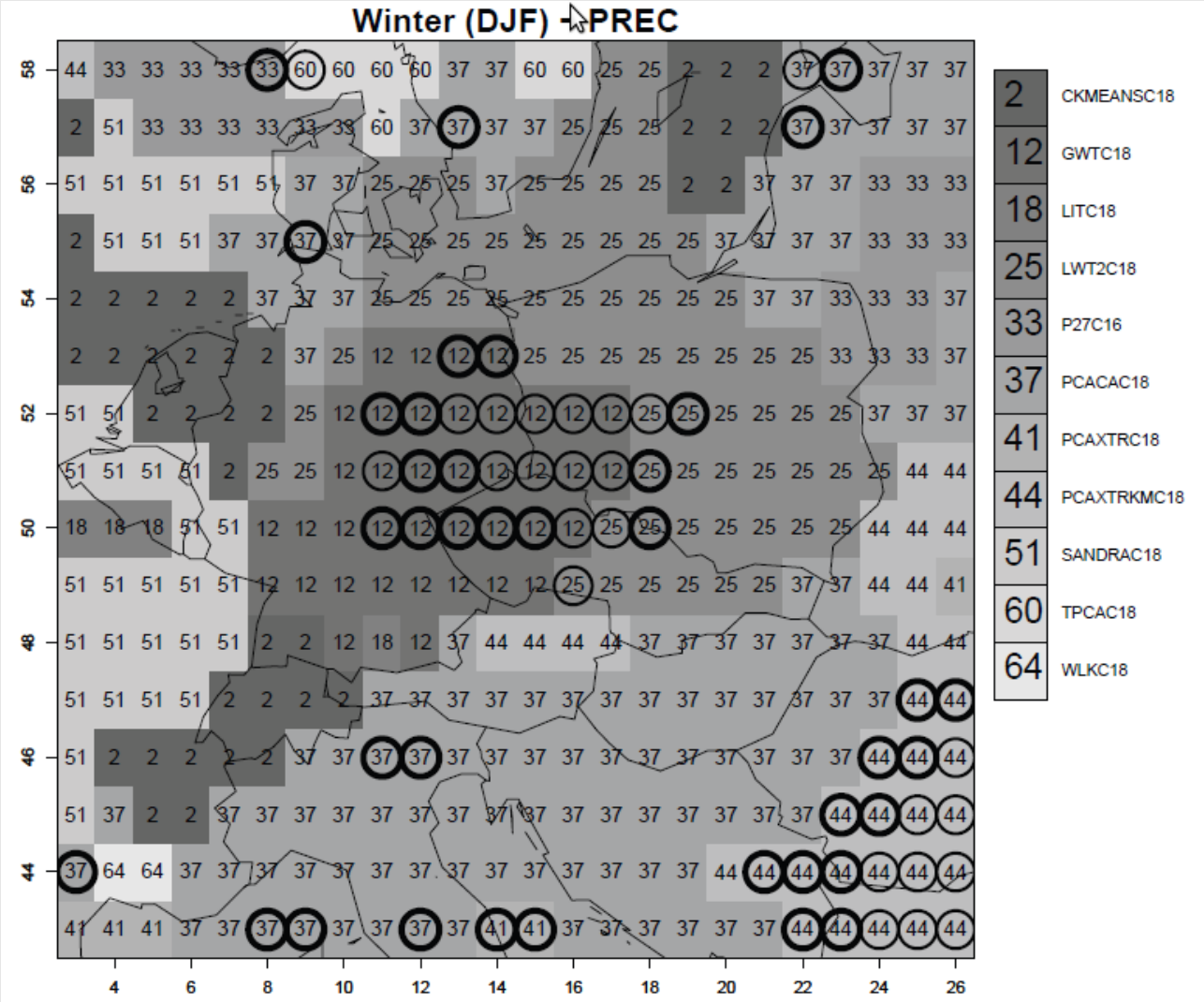
Performance variations within spatial domains

Domain 7

winter

PREC

best performing CTC at each gridbox



What can we learn from comparison studies in COST733 ?

- there is not one generally best individual CTC
- there is not one generally best basic approach
- distinct variations in performance among and within basic method groups
- performance shows marked seasonal and spatial variations (among and within spatial domains)
- high performance for the classified variable or for one target variable are not necessarily related to high performance for other target variables

What can we learn from comparison studies in COST733 ?

- **indications on strengths and weaknesses of basic methodological approaches:**
  - **OPT + MSLP**
  - **LDR/PCA + 2mT**
  - **THRES + PREC**

- **effect of varying classification settings on performance:**
  - **sequencing: + 2mT / - PREC**
  - **seasonal classifications: + / - ?**
  - **additional input variables**
    - **Vorticity: +PREC**
    - **1000/500 thickness: +2mT**
  - **importance of domain size: the smaller the better?**

What can we learn from comparison studies in COST733 ?

- **there is still some work left ...**