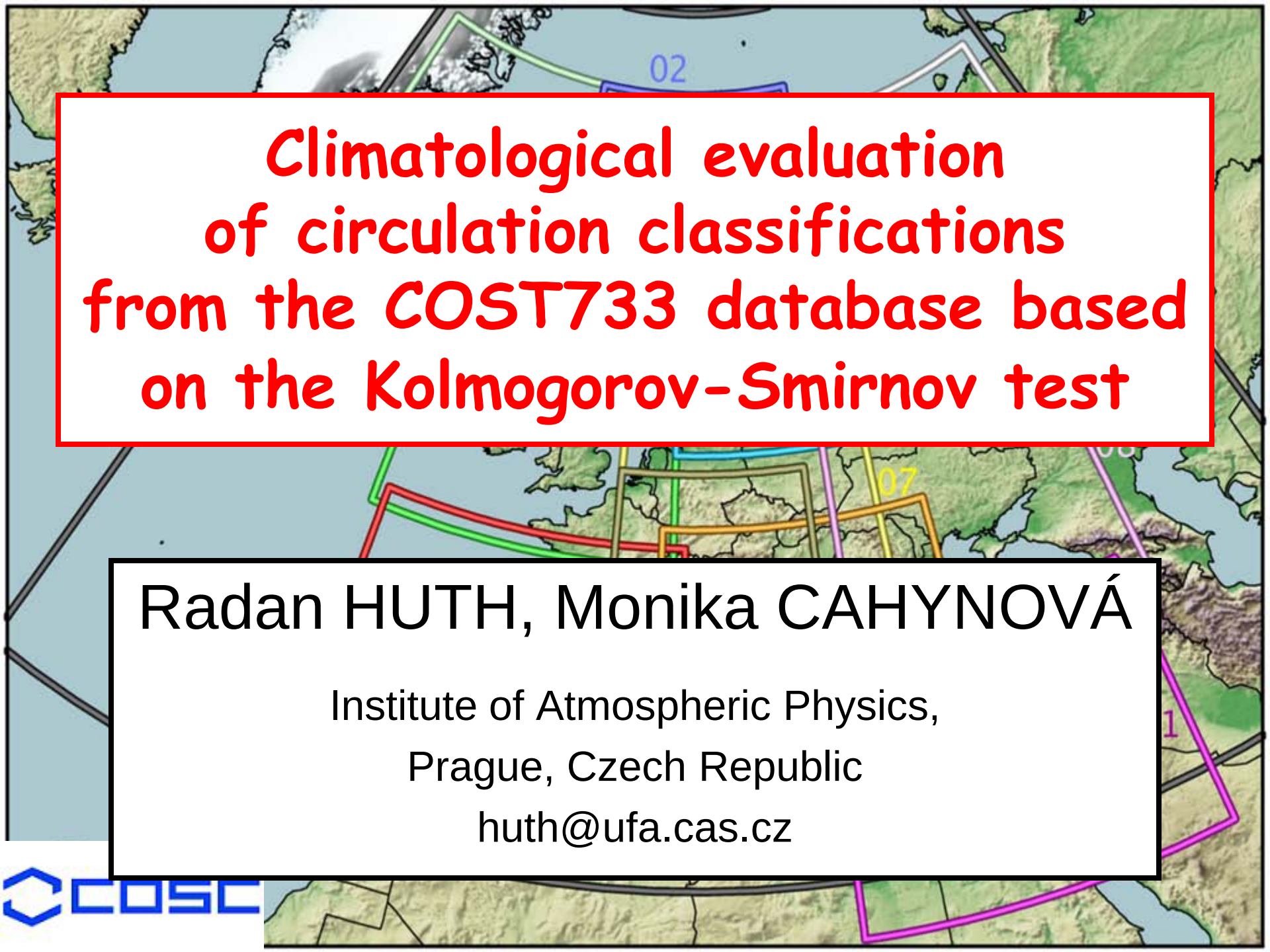


Climatological evaluation of circulation classifications from the COST733 database based on the Kolmogorov-Smirnov test

Radan HUTH, Monika CAHYNOVÁ

Institute of Atmospheric Physics,

Prague, Czech Republic

huth@ufa.cas.cz

What's new from my last presentation

- v2.0 of the database
- all domains
- DJF and JJA
- both temperature variables (Tmax & Tmin)
- precipitation

Which of my intentions haven't been fulfilled

- analysis in a gridded dataset
- i.e., station data are only analyzed

GOAL

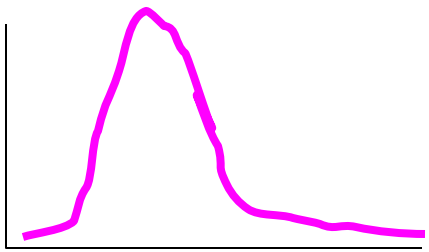
- assess the synoptic-climatological applicability of classifications
- i.e., how well they stratify surface weather (climate) conditions
- demonstrate effect of
  - selection of the classification method
  - number of types
  - sequencing
  - adding more variables
    - 500 hPa height
    - 500 hPa vorticity
    - 850/500 hPa thickness
  - seasonality of definition

ANALYSIS

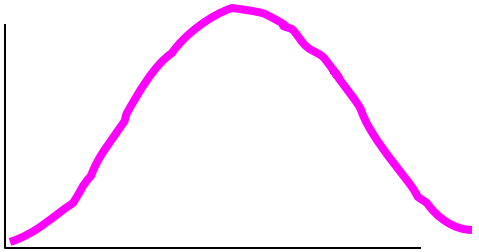
- variables
  - maximum temperature
  - minimum temperature
  - precipitation
- 126 stations from ECA&D database
- winter (DJF), summer (JJA)
- Jan 1961 – Dec 2000

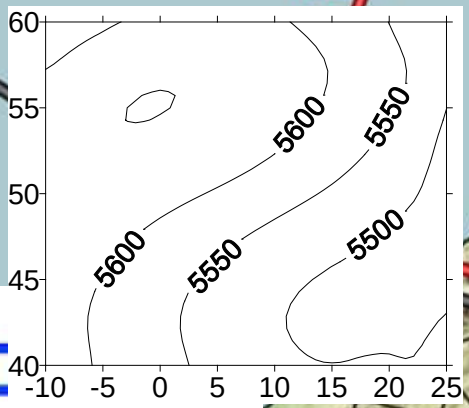
TOOL

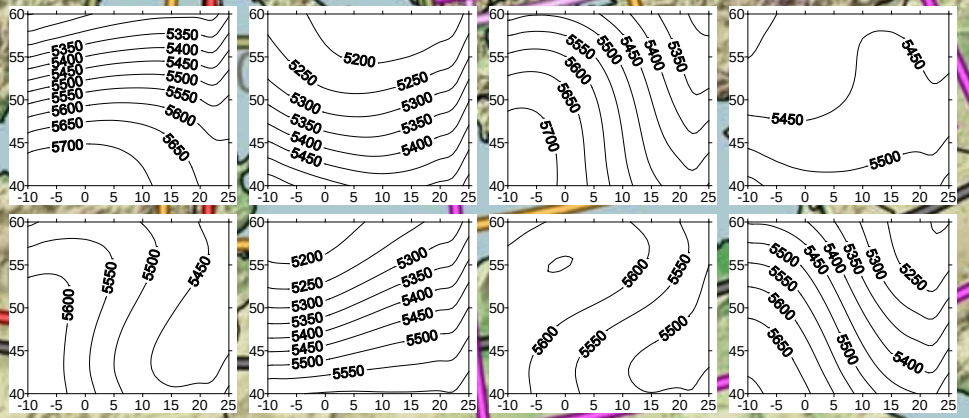
- 2-sample Kolmogorov-Smirnov test
- equality of distributions of the climate element under **one type** against under **all the other types**



X







TOOL

- at each station
- types for which the K-S test rejects the equality of distributions are counted
- the larger the count, the better the stratification, the better the synoptic-climatological applicability

RANKING OF CLASS'S

- at all stations individually:
  - for each classification: number of rejected K-S counted
  - classifications ranked by the %age of rejected K-S tests (= well separated classes)
  - higher %age ➔ better ➔ lower rank
- for each classification: ranks averaged over stations
- area mean rank ➔ ranking of the classification

Classifications examined

- 423 class's in each domain are ranked
- only a subset of class's enter the analysis
- omitted are
  - subjective class's & their objectivized versions
  - original class's provided by authors (those with 'o' in the name)
  - WLK method
  - SOM method
- 367 class's enter the competition

Classifications examined

- 4 methods with 3 classifications, differing in
  - number of types (9, 18, 27)
- 6 methods with 30 classifications, differing in
  - sequencing (no x 4 days)
  - additional variables (Z500, THICK850/500, VOR500, all together)
  - number of types (9, 18, 27)
- 5 methods with 35 classifications, differing in
  - as above
  - seasonal definition
- infrequent types (frequency < 10 days in the given season) are omitted

Result 1: Ranking of methods

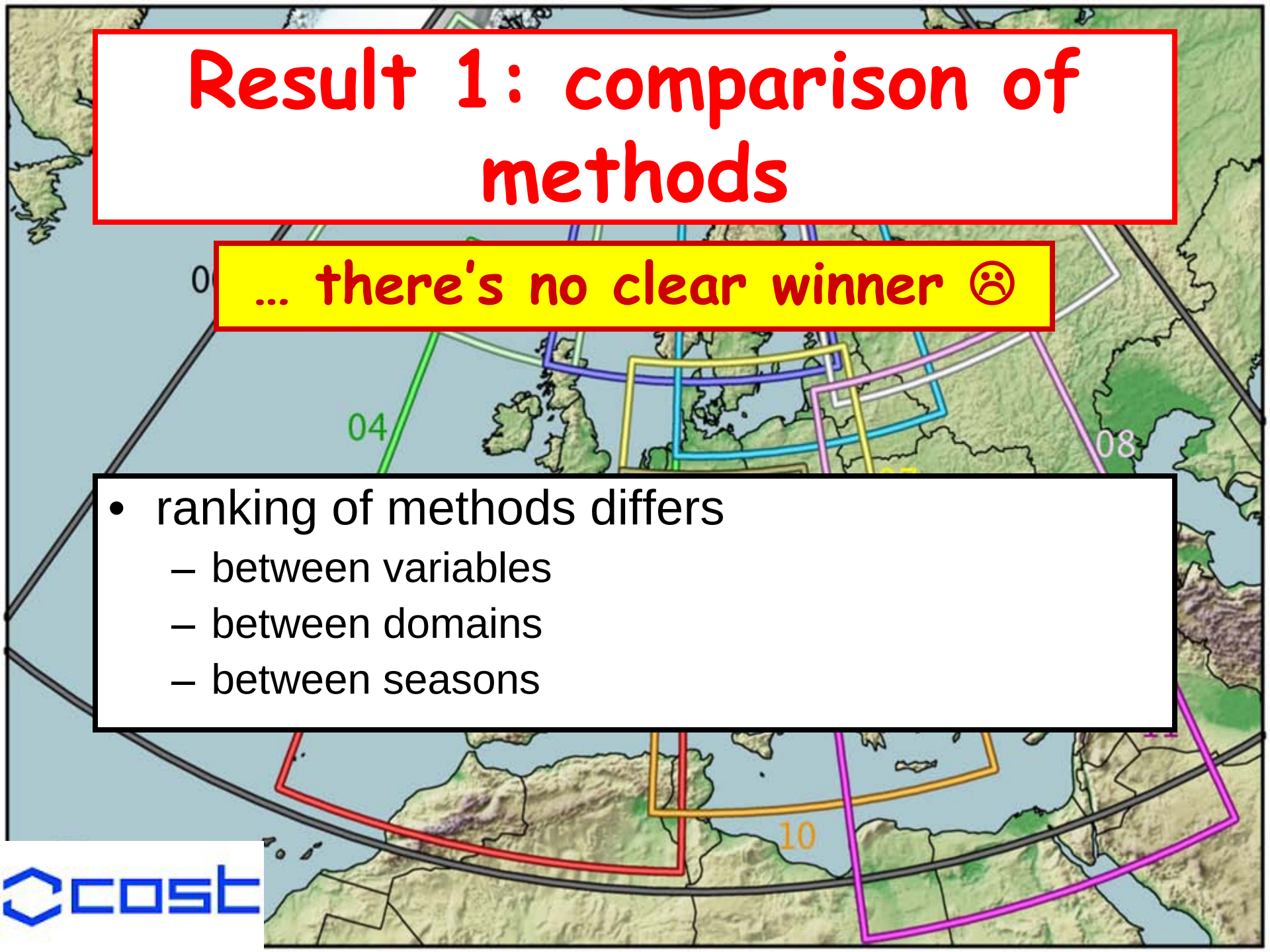
- area mean ranks averaged over 3 realizations with different numbers of types (~9, ~18, ~27) of each of 15 methods
- result: order of the method, independent of the number of types

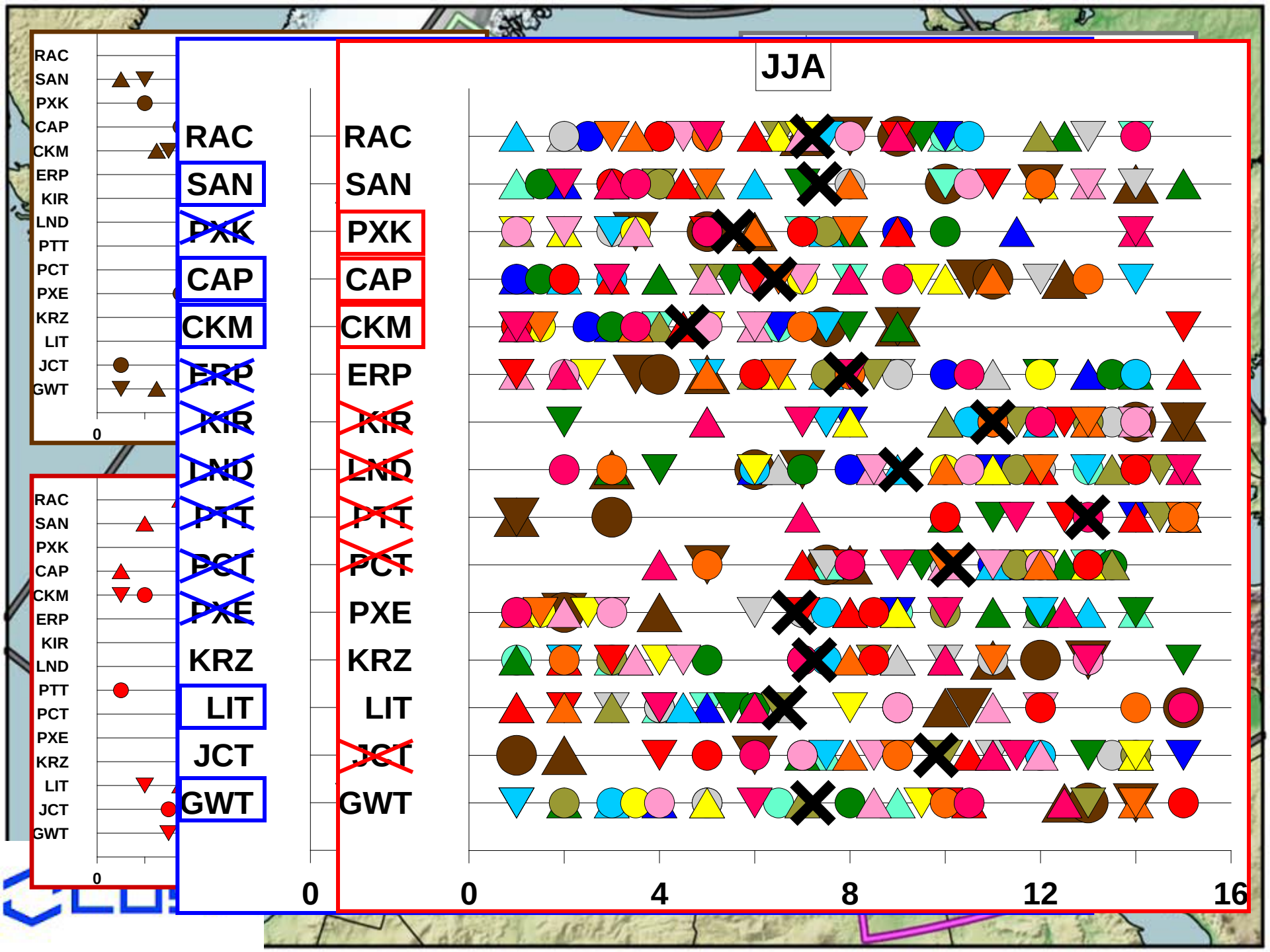
Result 1: comparison of methods

Result 1: comparison of methods

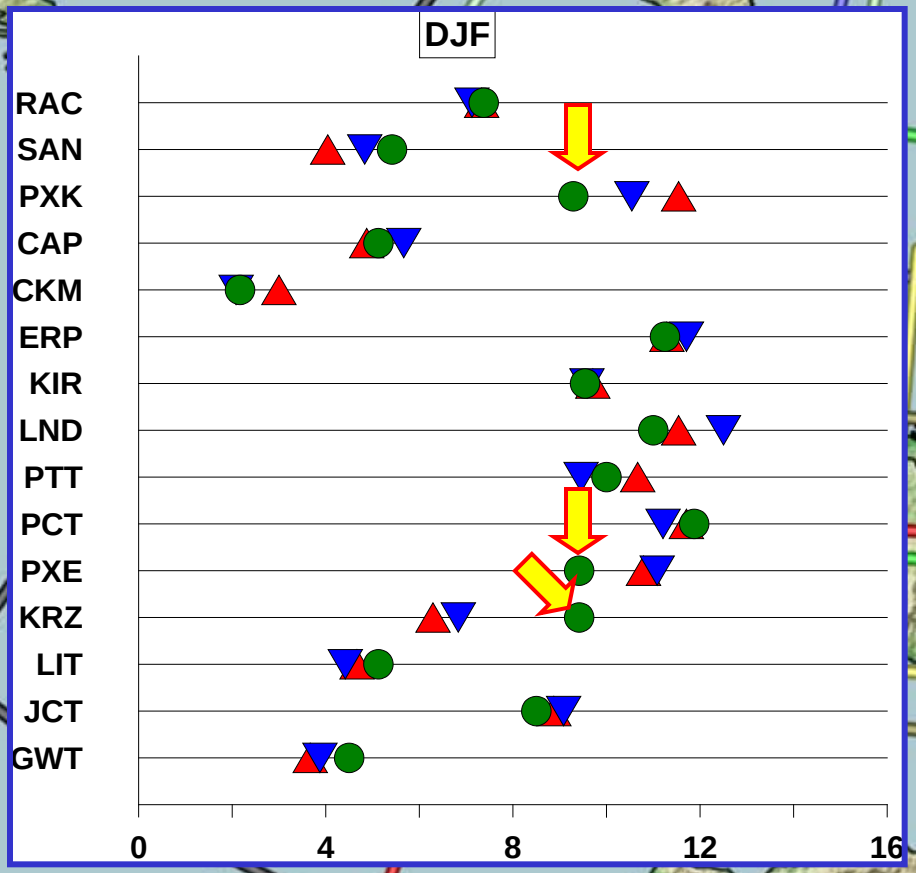
… there's no clear winner ☹

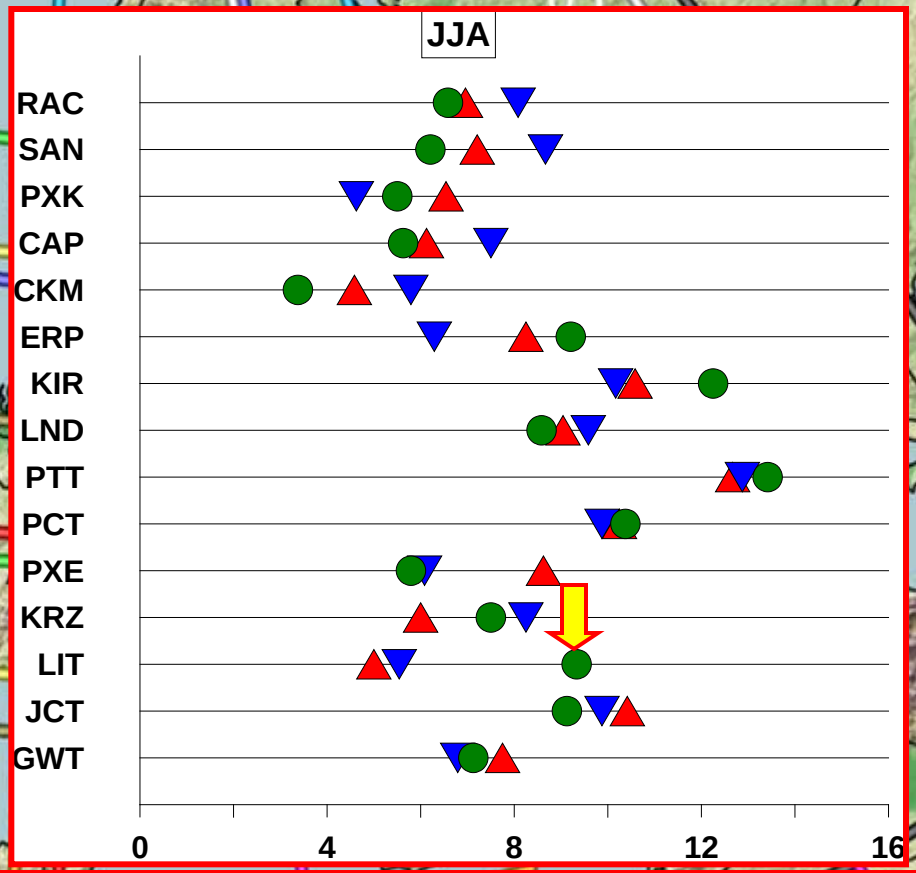
- ranking of methods differs
  - between variables
  - between domains
  - between seasons




JJA
RAC
SAN
PXK
CAP
CKM
ERP
KIR
LND
PTT
PCT
PXE
KRZ
LIT
JCT
GWT
0
4
8
12
16

Variables: ranks averaged over domains

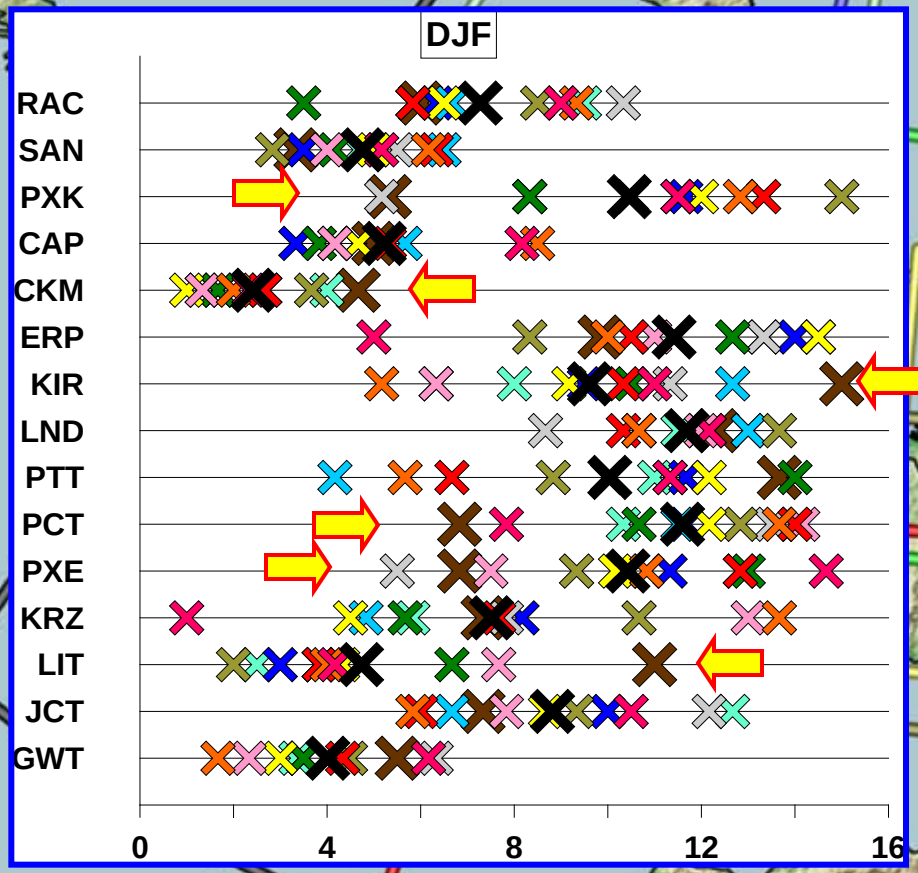


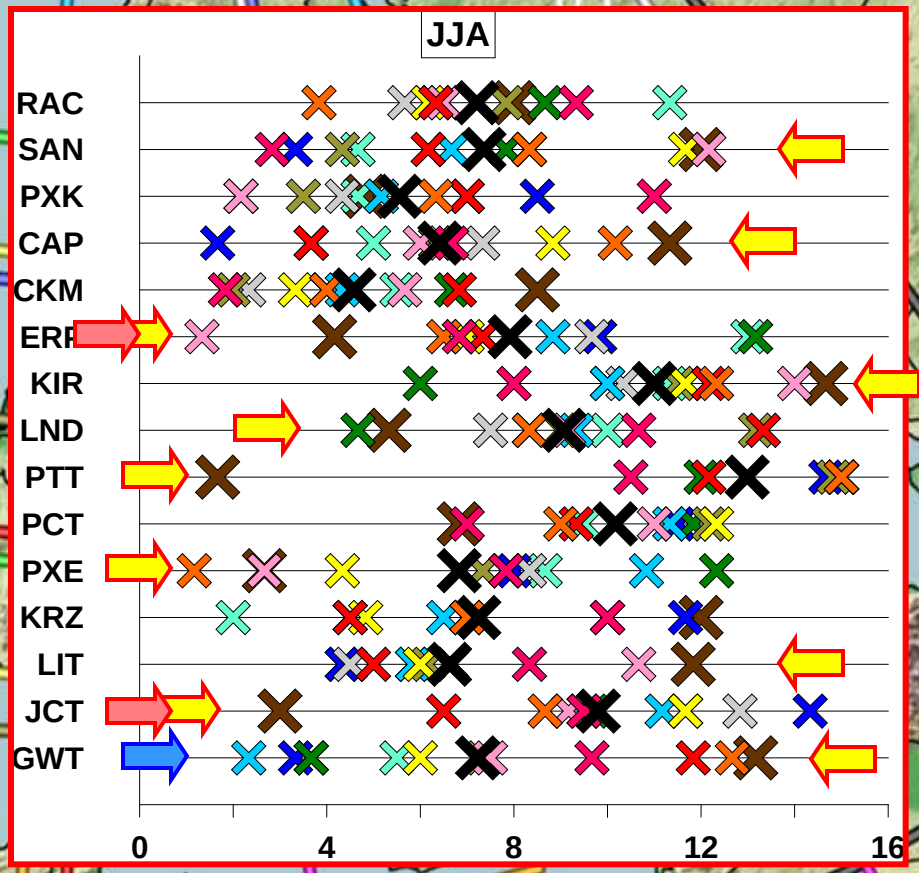


ranks of Tmin, Tmax close to each other
precip somewhat different

larger spread of ranks, differences even between Tmin x Tmax
pattern more chaotic ➔ firm conclusions hard to draw

Domains: ranks averaged over variables





- no apparent geographical dependence
- sensitivity to the size of the domain

- sensitivity to the domain size apparent for more class's
- some regional dependence

Overall ranking

method	DJF	JJA
GWT	2	6
JCT	8	12
LIT	3	4
KRZ	7	8
PXE	11	5
PCT	14	13
PTT	10	15
LND	15	11
KIR	9	14
ERP	13	10
CKM	1	1
CAP	5	3
PXK	12	2
SAN	4	9
RAC	6	8

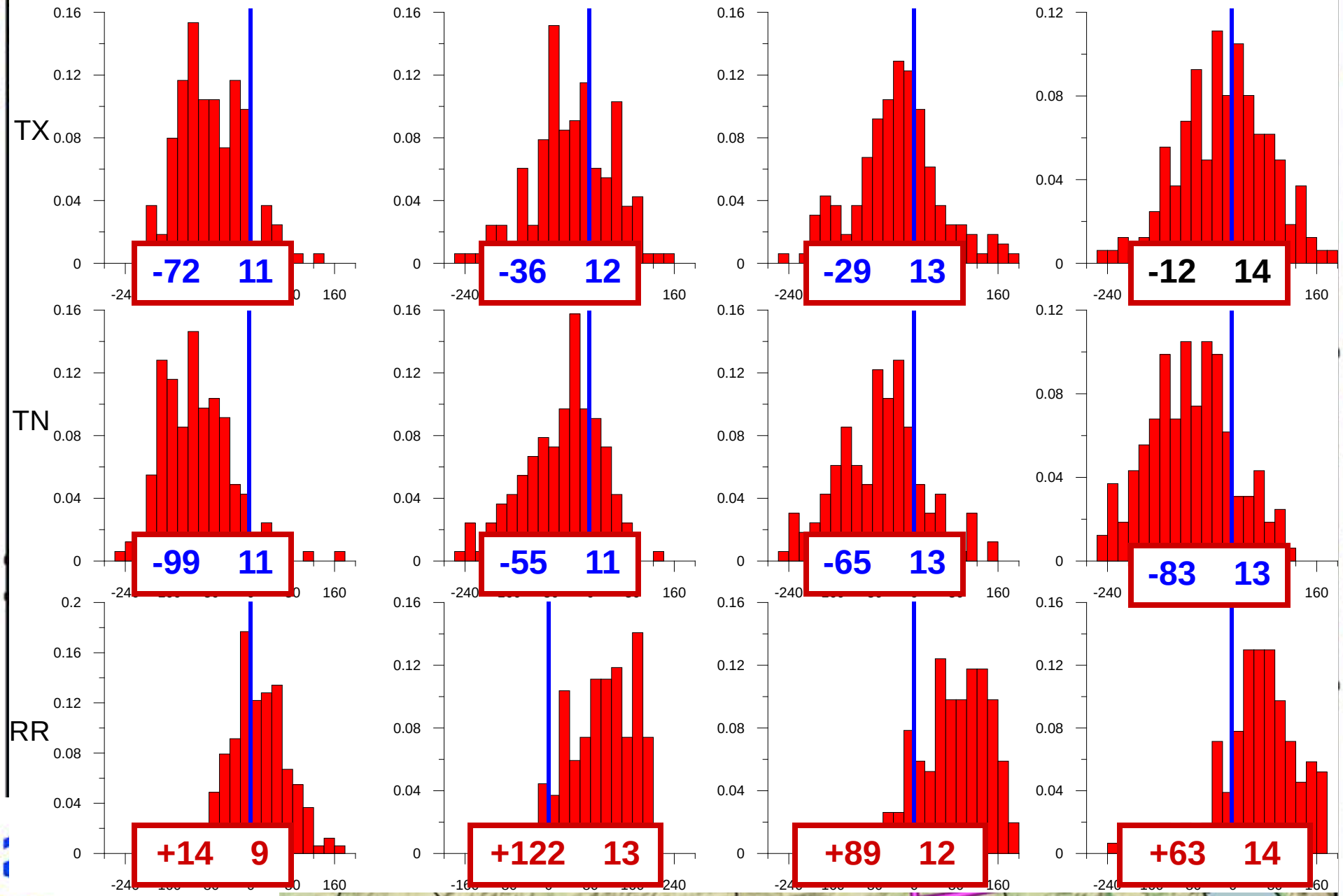
Short summary

- rankings vary among target variables, across domains, between seasons
- caution: results are contaminated (potentially biased) by unequal numbers of really occurring (enough populated) types; the contamination is stronger in JJA
- several well-performing methods can be identified: CKM, CAP, LIT, GWT
- several methods cannot be recommended: PCT, PTT, LND, KIR, ERP

Result 2: effect of sequencing

- all pairs of classifications
  - differing in sequencing (no vs. 4-days)
  - with all other attributes equal
- difference in rank is calculated
- histogram of differences
- t-test: equality of the difference to zero

EFFECT OF SEQUENCING, DJF



Effect of sequencing

	DJF			JJA		
domain	TX	TN	RR	TX	TN	RR
00	**-72.1**	**-99.3**	*+14.1*	**-75.5**	**-118.8**	*+41.4*
01	+4.7	**-35.2**	*+83.0*	+40.5	**-29.0**	*+107.1*
02	**-59.0**	**-75.3**	*+113.0*	**-13.8**	**-77.8**	*+94.7*
03	**-35.9**	**-54.6**	*+121.7*	+2.3	-9.5	*+88.8*
04	-6.0	**-69.4**	*+94.1*	-1.3	**-76.5**	*+90.5*
05	**-21.2**	**-40.5**	*+113.0*	+6.9	**-40.3**	*+82.9*
06	**-33.7**	**-58.7**	*+103.3*	*+30.1*	**-34.4**	*+108.0*
07	**-29.5**	**-65.1**	*+88.9*	+2.0	**-37.4**	*+102.0*
08	**-34.4**	**-59.9**	*+80.9*	+9.1	-9.4	*+99.9*
09	-11.6	**-83.2**	*+62.7*	*+22.7*	**-78.8**	*+79.3*
10	**-29.2**	**-53.8**	*+72.2*	-8.9	**-41.1**	*+92.4*
11	**-14.3**	**-37.4**	*+96.9*	*+42.2*	**-30.4**	*+124.7*

- improves stratification for temperature
- improvement larger for Tmin
- deteriorates it for precip
- improvement is largest / deterioration smallest for the large domain

- positive effect on temperature is weaker & less ubiquitous
- deterioration for precip similar to DJF
- same effect for D00 as in DJF

Result 3: sensitivity to the number of types

- all pairs of classifications
  - differing in no. of types
    - 9 vs. 18
    - 18 vs. 27
  - with all other attributes equal
- difference in rank is calculated
- histogram of differences
- t-test: equality of the difference to zero

Result 3: sensitivity to the number of types

- significantly better results for lower nos. of types in almost all cases
- the only few (=3) exceptions, all for JJA and for the comparison of 9 vs. 18 types

effect of

adding height or thickness:
- improvement for temperature – though spatially variable
- varied response for precip
- height more effective than thickness

adding vorticity:
- general deterioration

				/500 hPa thickness			500 hPa vorticity		
domain	TX	TN	RR	TX	TN	RR	TX	TN	RR
00	**-56.0**	**-58.4**	**-77.7**	**-69.5**	**-73.7**	**-76.9**	+12.3	**-14.6**	+11.5
01	**-31.7**	**-39.6**	-8.4	-34.5	**-45.7**	+10.8	-2.5	**-22.4**	+19.1
02	**-29.7**	**-52.3**	-16.4	**-30.5**	**-47.5**	*+17.9*	+13.7	+11.3	+2.7
03	**-55.0**	**-60.3**	**-32.5**	**-28.3**	**-28.2**	+3.4	*+29.2*	*+34.0*	*+22.8*
04	+8.3	*+23.2*	-10.4	-7.7	+.6	-7.7	*+25.5*	*+33.1*	-9.4
05	**-63.1**	**-54.1**	+4.6	-13.5	-9.6	*+28.8*	*+29.9*	*+41.2*	*+47.7*
06	**-67.8**	**-66.6**	**-19.9**	-29.2	**-34.9**	-16.8	+10.0	*+30.4*	+5.8
07	-15.9	**-29.3**	-2.3	+3.7	**-28.0**	+8.1	*+34.9*	*+25.7*	+10.5
08	**-77.7**	**-89.4**	**-30.1**	**-66.9**	**-85.8**	**-23.0**	*+32.8*	*+21.1*	*+22.9*
09	**-72.0**	**-34.4**	-25.0	**-38.1**	-10.3	-2.1	**-23.9**	*+28.6*	*+26.0*
10	**-76.0**	**-92.5**	**-28.1**	**-39.2**	**-70.3**	**-20.5**	*+27.0*	*+28.2*	+11.5
11	**-75.8**	**-78.1**	**-82.2**	**-62.0**	**-67.1**	**-61.0**	-6.3	+15.1	+13.2

adding height or thickness:
• improvement for temperature; stronger in SE half of Europe; stronger for Tmin
• deterioration / improvement for precip in N+NW / S+SE Europe

adding vorticity:
• improvement for temperature in most domains
• varied response for precip

				500 hPa thickness			500 hPa vorticity		
domain	TX	TN	RR	TX	TN	RR	TX	TN	RR
00	**-60.5**	**-78.4**	**-36.6**	**-42.4**	**-54.1**	+19.6	-5.5	-.7	-4.8
01	+31.3	**-74.9**	*+71.5*	*+61.1*	-33.9	*+87.4*	+10.5	-14.9	*+47.0*
02	**-36.2**	**-98.2**	*+86.7*	-5.6	**-88.9**	*+92.8*	**-33.6**	**-27.4**	+1.8
03	**-48.1**	**-107.9**	*+51.9*	-17.2	**-85.8**	*+52.3*	**-66.2**	**-50.6**	+20.1
04	-19.9	**-51.8**	*+56.5*	+29.2	**-53.3**	*+107.0*	+15.5	+14.2	+4.2
05	**-65.1**	**-124.6**	*+67.2*	-12.2	**-90.9**	*+79.5*	**-38.2**	-16.9	+12.0
06	**-151.0**	**-183.6**	-16.3	**-85.5**	**-148.7**	*+28.6*	**-61.8**	**-29.9**	-1.5
07	**-119.0**	**-157.7**	-12.8	**-96.4**	**-136.3**	+12.2	**-44.7**	**-22.7**	**-18.5**
08	**-137.1**	**-141.4**	-16.5	**-109.2**	**-136.9**	+14.7	**-37.0**	-7.5	-7.5
09	**-83.3**	**-126.1**	**-123.6**	**-80.3**	**-132.5**	**-87.3**	-8.3	-.8	**-42.6**
10	**-162.6**	**-172.5**	-27.5	**-148.2**	**-172.5**	-27.5	**-42.0**	-14.7	-11.7
11	**-104.6**	**-92.2**	**-61.5**	**-99.4**	**-101.4**	**-35.0**	-21.5	-5.7	-17.7

Result 5: Effect of seasonality

	DJF			JJA		
domain	TX	TN	RR	TX	TN	RR
00	**-44.0**	**-71.2**	**-46.4**	**-132.7**	**-94.5**	**-156.4**
01	**-40.6**	-28.2	**-17.1**	-44.2	**-54.7**	**-92.6**
02	**-75.2**	**-63.5**	**-35.1**	**-46.5**	-14.6	**-87.0**
03	-15.5	-25.7	**-55.4**	+3.4	-17.0	**-58.1**
04	**-37.8**	**-45.3**	**-36.2**	**-62.2**	**-49.4**	-20.1
05	**-42.5**	-32.6	**-21.9**	-14.9	-29.0	**-58.8**
06	**-26.1**	**-35.8**	**-30.5**	-12.3	+18.6	-9.4
07	**-27.1**	**-39.4**	**-28.9**	**-52.5**	**-39.5**	**-55.0**
08	**-36.8**	**-47.5**	-19.6	**-44.7**	**-57.3**	**-40.1**
09	-38.1	**-56.4**	**-61.6**	**-90.7**	**-84.0**	-15.7
10	-16.9	**-25.7**	**-25.5**	**-70.7**	**-64.5**	**-45.1**
11	**-44.9**	**-45.5**	-17.0	-18.8	*+36.9*	+8.2

- general improvement
- some geographical variability

BUT:
systematic difference in the no. of types
(7 seasonal vs. 9 non-seasonal)
➔ improvement may be partially an artifact of this difference

instead of CONCLUSIONS

- what else might have been done
- or has been done by someone else and might be nice to be combined with this study

- gridded dataset (Ensembles, NCEP or ERA40)
- other criteria of stratification