Probabilistic verification of decadal CMIP5 hindcast experiments - VECAP verification tool -

International workshop on seasonal to decadal prediction
May 13th-16th, 2013

S. Stolzenberger[1], A. Hense[1], R. Glowienka-Hense[1]
T. Spangehl[2], A. Mazurkiewicz[2], M. Schröder[2]

[1]Meteorological Institute, University of Bonn

[2]Deutscher Wetterdienst

FONA Decadal Climate Prediction BMBF

MiKlip

Meteorological Institute University Bonn

DWD

Outline & Objectives

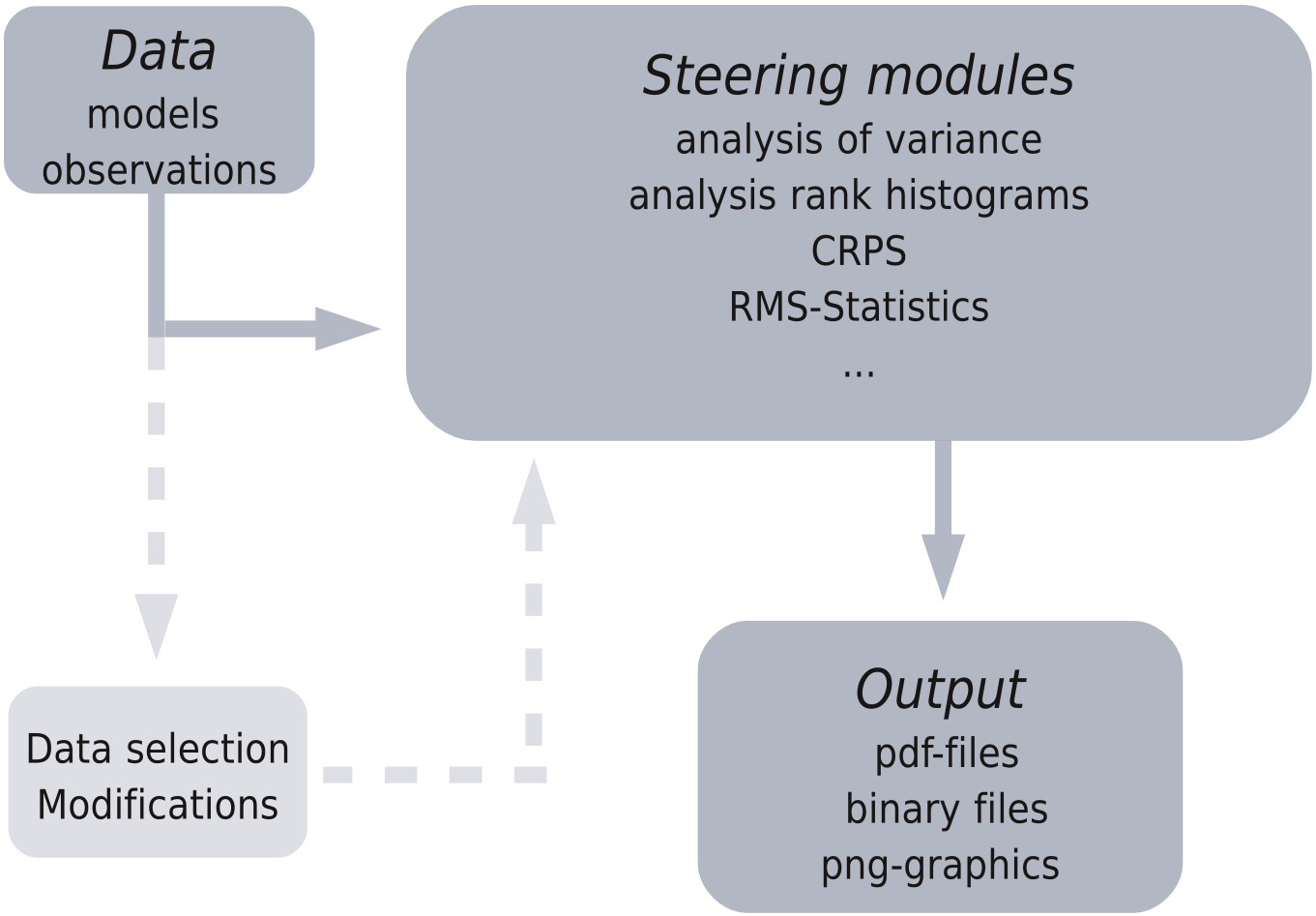
- Introduction to the VECAP verification tool
- Probabilistic verification of ensemble predictions

Philosophy of VECAP

Verification, **C**alibration and **A**ssessment of **P**redictability of medium-range climate predictions using satellite data

- Compare
  - model with observations: **verification**
  - model with another model: **model sensitivity**
  - Model A with data vs Model B with data: **model assessment**
- Use standard univariate methods
- Understand prediction system (zonal means, indices)
- Use satellite retrieval data (Poster T. Spangehl)

VECAP tool


Data
models
observations
Steering modules
analysis of variance
analysis rank histograms
CRPS
RMS-Statistics
...
Data selection
Modifications
Output
pdf-files
binary files
png-graphics

Models and observations

- ECHAM6-T63L47/MPIOM CMIP5 simulations
  - hindcast experiments, yearly initalization in January
  - Baseline 0 and Baseline 1 (3-10 realizations)

- ERA-Interim (since 1979)

  interpolation on ECHAM6-T63 Gauss grid (1.875° x 1.875°) on standard pressure levels

Sharpness and reliability

one-way/two-way anova of ensemble simulations

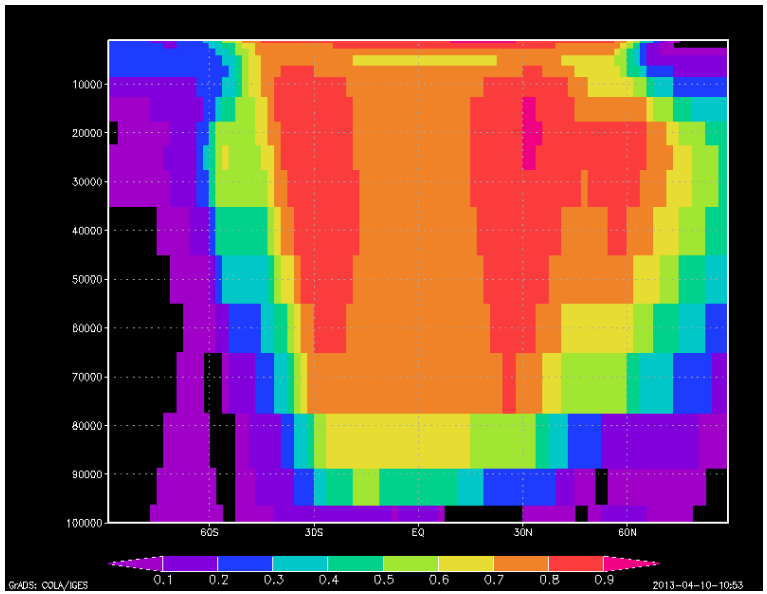
- ratio between internal and external variability
- (potential) sharpness exists if external variability is larger than zero (rejection of null hypothesis)

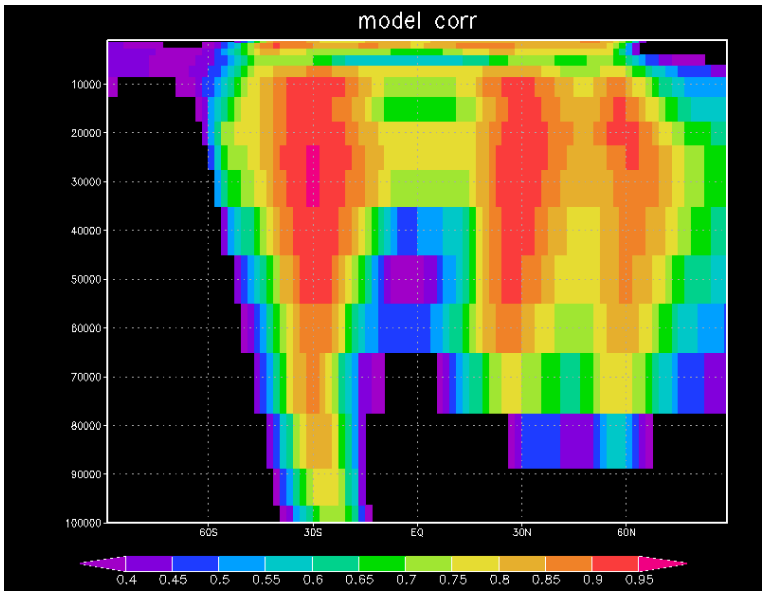
analysis rank histogram

- check if the ensemble is calibrated
- distribution function of the random variable M "model prediction" is identical to the CDF of the random variable B "observation" $p_t = Prob(b_t \leq M) = F_{M,t}(m = b_t)$
- estimate PDF of p_t e.g. as histogram or as $\beta-$density

Anova and correlations

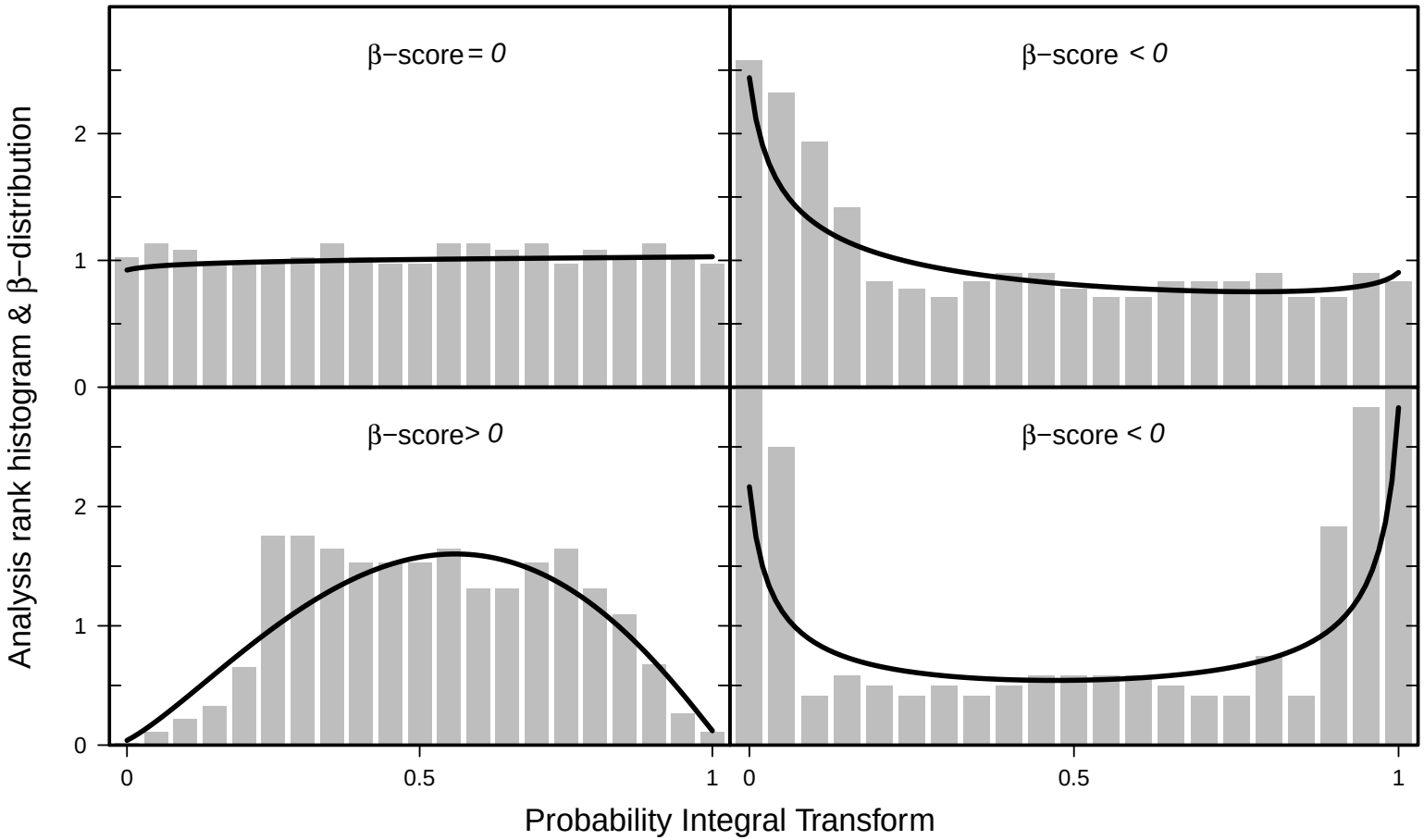
- geopotential - annual mean - 1979-2012
- 17 levels, zonal means
- prediction year 2-5





Analysis rank histogram and β-score

β-score$= 1 - \frac{1}{\sqrt{\alpha\beta}}$



β-scores for temperature at different p-levels

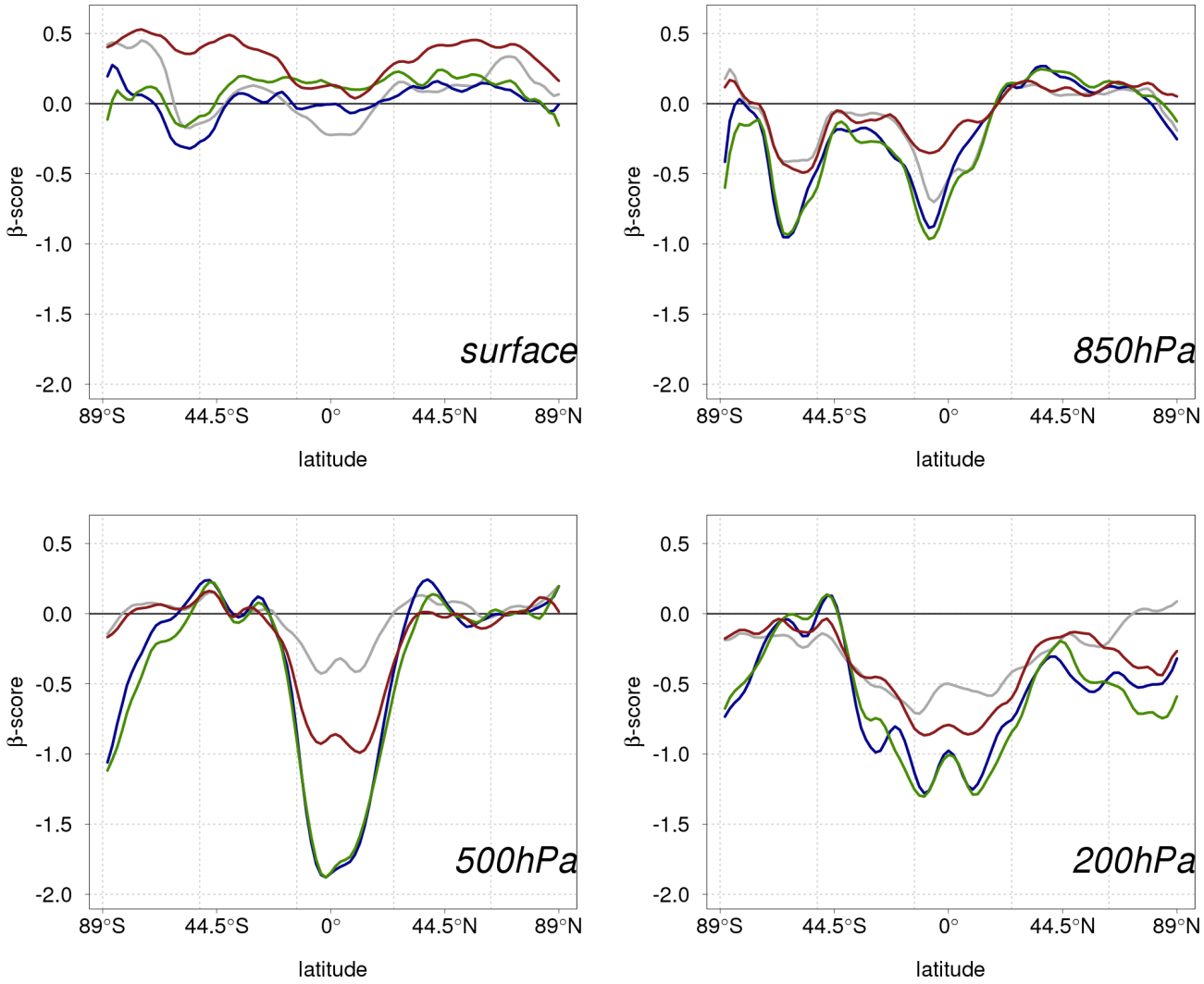
1980-2011

prediction year 1

prediction year 2-5

prediction year 6-9

prediction year 10



Score and score function

- A Score $\mathcal{S}$ is the expectation of a score function S with respect to the observation density

$$\mathcal{S}(f_M, f_B) = \int S(f_M(m), b) f_B(b) db$$

- A (negatively oriented) score is called **proper** if

$$\mathcal{S}(f_B, f_B) \leq \mathcal{S}(f_M, f_B)$$

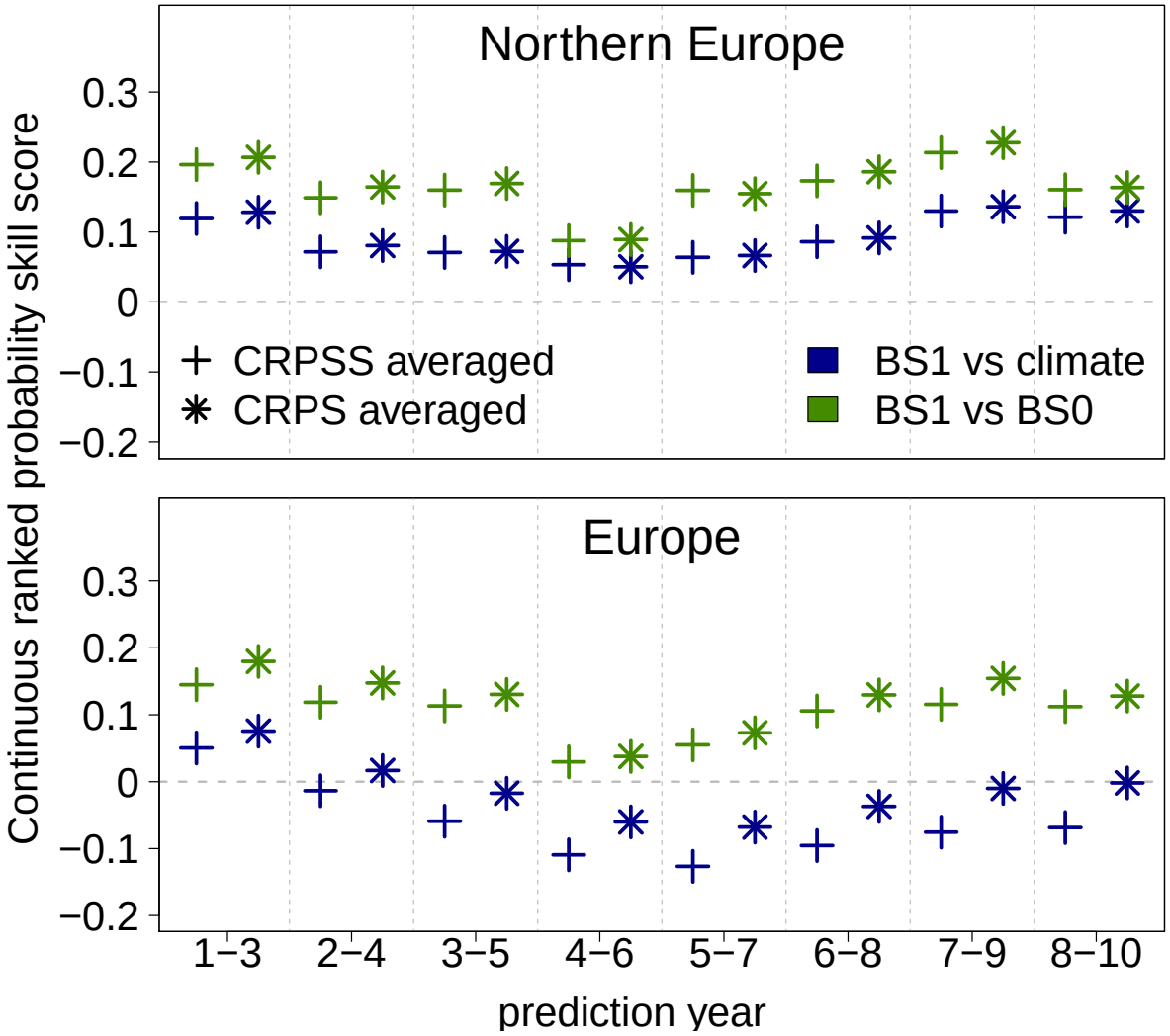
Implementation of a score function in the VECAP tool

One example:

Continuous ranked probability score $\mathcal{S}_{CRPS}$

- based on the score function
  $S(F_{M,t}(m), b)_{crps} = \int_{-\infty}^{\infty}(F_{M,t}(m) - \mathcal{H}(m - b_t))^2 dm$
  ($\mathcal{H}$: Heavyside function)
- analytical solution of the integral for standard Gaussian

Baseline 1 vs Baseline 0



$$CRPSS = 1 - CRPS/CRPS_{ref}$$

- geopotential 500hPa
- annual mean
- 1980-2011

Summary

- Introduction to the VECAP verification tool
  - verification/validation of individual data sets
  - output as graphics and binary files
- Probabilistic verification of ensemble predictions
  - necessary tests: anova and analysis rank histograms
  - sufficient tests: correlations, scores/skill scores etc.

Probabilistic verification of decadal CMIP5 hindcast experiments - VECAP verification tool -

International workshop on seasonal to decadal prediction
May 13th-16th, 2013

S. Stolzenberger[1], A. Hense[1], R. Glowienka-Hense[1]
T. Spangehl[2], A. Mazurkiewicz[2], M. Schröder[2]

[1]Meteorological Institute, University of Bonn

[2]Deutscher Wetterdienst

FONA Decadal Climate Prediction BMBF

MiKlip

Meteorological Institute University Bonn

DWD

RMS-Statistics

- Mean observations
- Standard deviation observations
- Mean model data
- Standard deviation model data
- Mean difference: model minus observation(*Bias* would be squared difference but with loss of information). Confidence-test: Student-t-test.
- Quotient observation/model (=ensemble mean) standard deviation.
- Correlation between model and observational time series. Confidence test: Student-t-test.

Sea level pressure indices

- Southern oscillation indices:
  - Standardized slp difference Tahiti-Darwin
  - Standardized slp Darwin
  - Equatorial SOI
- NAO indices:
  - Standardized slp differences Lisbon-Stykkisholmur
  - Gibraltar-Reykjavik
  - A pc-based index over the Atlantic sector (20°N-80°N, 90°W-40°E) Deviating from Hurrels (1995) suggestion, the pc are calculated from the correlation matrix and all months of the year are stringed togethter. This reduces the noise for small data series and localizes the pattern.
- North Pacific Index NPI
  - SLP (30° – 65°N, 160°E-140°W)
- Antarctic Oscillation (AAO) indices
  - zonal mean SLP difference 40°S minus 65°S
  - zonal mean SLP difference 40°S minus 70°S