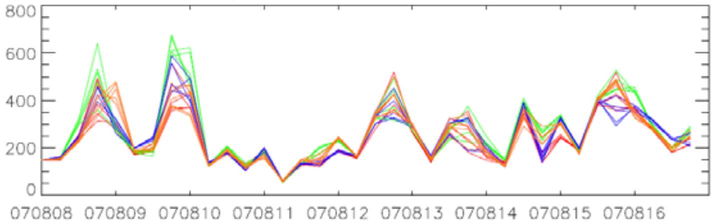
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

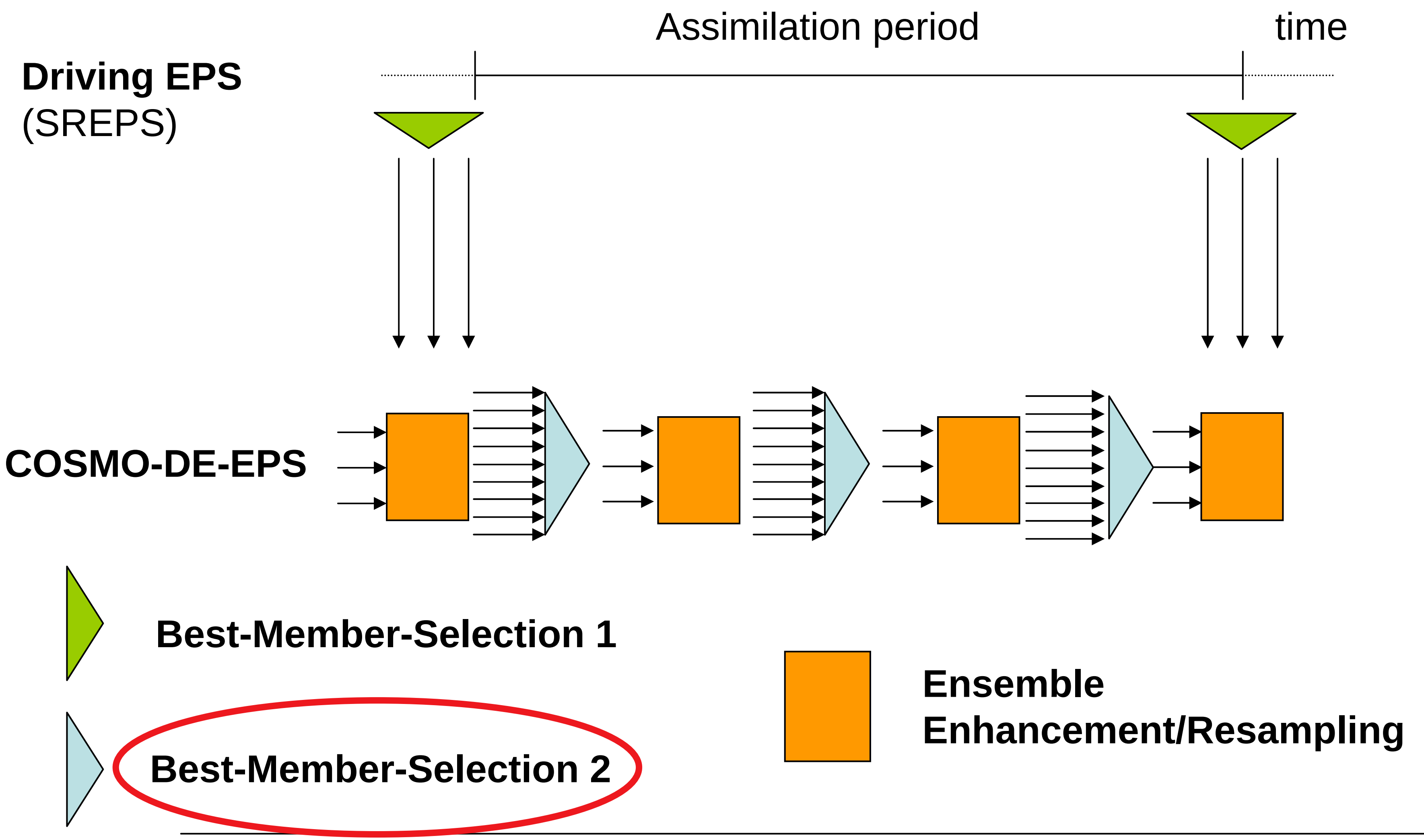
Best-member-selection as part of an ensemble based data assimilation approach

Tanja Weusthoff[1], Daniel Leuenberger[1],
Christian Keil[2], George Craig[2]
[1]MeteoSwiss, [2]DLR



COSMO/CLM User Seminar, Langen, 09.-11.03.2009

New data assimilation approach: SIRF



Objectives

WP1: Evaluation of classical and spatial metrics for the determination of weighting schemes for ensemble members (DLR, MeteoSchweiz)

WP1.1 Implement spatial metrics and evaluation on selected COPS events

FQM of Keil and Craig (2007), which is already available,

SAL (Wernli et al. 2007)

fuzzy verification package of Ebert (2007).

WP1.2 Implement classical metrics for conventional observations

WP1.3 Investigate correlation of metrics between models of different resolution

WP1.4 Compare object-oriented and classical metrics

WP1.5 Investigate persistence of skill in different metrics and different meteorological situations

Concept

Apply classical quadratical error measure for conventional observations:

$$[y - H(x)]^T R^{-1} [y - H(x)] \qquad (R^{-1})_{ij} = \begin{cases} 1/\sigma_{ij}^2 & i = j \\ 0 & else \end{cases}$$

In order to give more important observations a higher weighting („tuning“), an additional matrix A is multiplied to the error covariance matrix R containing the setting of the weights.

$$m = [y - H(x)]^T A R^{-1} [y - H(x)] \qquad A_{ij} = \begin{cases} \alpha_i & i = j \\ 0 & else \end{cases}$$

y = vector of observations
H(x) = vector of forecasts
H = forward operator,
H(x) = observation equivalent, calculated from model state x
R = covarianz-matrix of the observation errors, **Diagonalmatrix**
A = additional matrix for the weightings of parameters, **Diagonalmatrix**

Quality measure m

= distance of model member $x^{(k)}$ to observations

$$m = [y - H(x)]^T AR^{-1} [y - H(x)]$$

$$m^{(k)} = \left[\tilde{d}_i^{(k)}\right]^T AR^{-1} \tilde{d}_i^{(k)}$$

$$m^{(k)} = \sum_{i=1}^{Y} \alpha_i \frac{\left[\tilde{d}_i^{(k)}\right]^2}{\sigma_{ii}^2}$$

$$\tilde{d}_i^{(k)} = y_i - H(x^{(k)})$$

$$i = 1, \ldots Y$$

Y = # variables * # stations * # levels

$$k = 1, \ldots K$$

K = # member

$$m^{(k)} = D \cdot \alpha$$

optimal weighting α to be determined

Quality measure m

= distance of model member $x^{(k)}$ to observations

$$m^{(k)} = D \cdot \alpha$$

$$\begin{pmatrix} m^{(1)} \\ m^{(2)} \\ \cdots \\ \\ m^{K} \end{pmatrix} = \begin{pmatrix} d_1^{(1)} & d_2^{(1)} & d_3^{(1)} & \cdots & d_Y^{(1)} \\ d_1^{(2)} & d_2^{(2)} & \cdots & & \cdots \\ \cdots & \cdots & & & \\ & & & & \\ d_1^{(K)} & d_2^{(K)} & \cdots & & d_Y^{(K)} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \cdots \\ \\ \alpha_Y \end{pmatrix}$$

Forecast: COSMO-DE-EPS

- High resolution ensemble with 20 members, 00 UTC runs

- 4 driving models:
  - Member 1-5 EZMWF (red)
  - Member 6-10 GME (green)
  - Member 11-15 NCEP (blue)
  - Member 16-20 UKMO (yellow)

- 5 perturbations of model physics
  - Members 1, 6,11,16 Entrainment rate (entrscv=0.002)
  - Members 2, 7,12,17 Cloud cover at saturation (clc_diag=0.5)
  - Members 3, 8,13,18 Laminar layer depth (rlam_heat=50)
  - Members 4, 9,14,19 Laminar layer depth (rlam_heat=0.1)
  - Members 5,10,15,20 Turbulent lengh scale (tur_len=150)

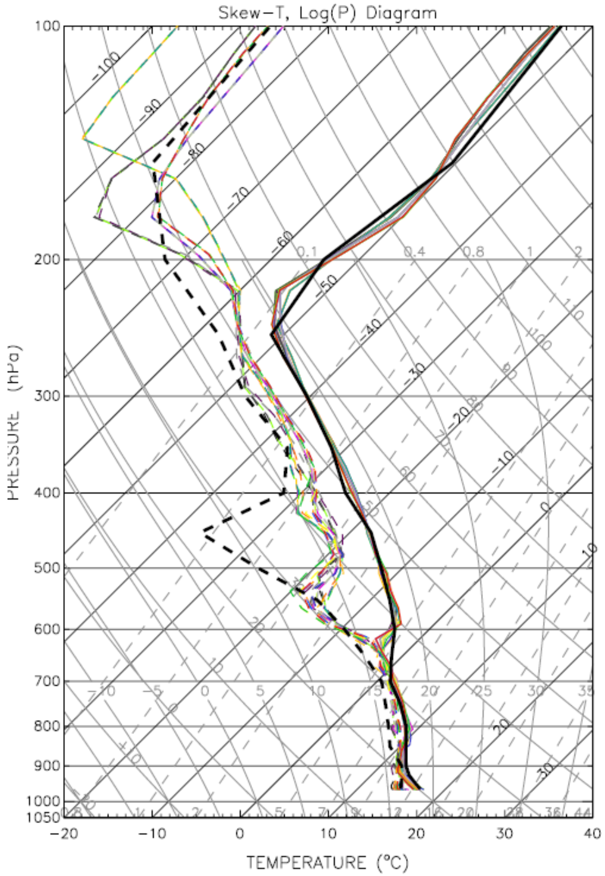
Observations

Best-member-selection based on standard observations like Synop (hourly), Radiosondes (6 hourly or less), …

First investigation with radiosondes

- Variables on pressure levels:
  - temperature,
  - relative humidity,
  - u- and v- wind components
  - …

- Integrated variables:
  - CAPE
  - CIN
  - integrated water vapour (IWV)
  - …



Weighting …

Seven variables

Temperature T,
relative humidity RH,
Wind components U and V } 21 levels

CAPE
IWV
CIN } single levels

Three level regions

1000 – 700 hPa
650 – 400 hPa
350 – 100 hPa

→ Weightings of variables and levels are multiplied!

Initial Weighting → everything is set to the same weight 1

T	RH	U	V	CAPE	IWV	CIN
1	1	1	1	1	1	1

1000-700	650-400	350-100
1	1	1

… & Evaluation

Q: Do the target quality measure and our m identify the same best members?

1. **Correlation** between target quality r and m

$$cor = \frac{\text{cov}_{r,m}}{\sigma_r \cdot \sigma_m}$$

2. **Distance** between the two quality measures:

$$\left| r - m \right|$$

→ Need $m^{(k)}$ and $r^{(k)}$ in the same range, e.g. [0,1] in order to determine the optimal weighting α = [0,1] such that

$$\left| r - m \right| = \left| r - D \cdot \alpha \right| = \min$$

Here: **Target quality r = DAS**

Target Quality: DAS

- The target quality r of the ensemble is provided by the quality measure with respect to precipitation (radar) → the Displacement and Amplitude Score DAS (Keil and Craig, 2009)

- DAS is based on an areal image matcher using classical optical flow technique & has 2 components:
  - displacement error field DIS
  - amplitude error AMP (RMS of observed and morphed forecast imagery)

$$DAS = \frac{DIS}{D_{\max}} + \frac{AMP}{I_0}$$

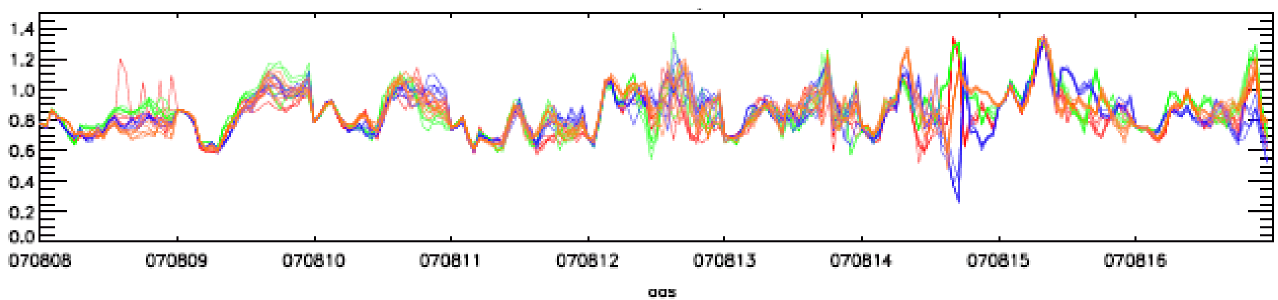
D_{max} maximum search distance
I_0 characteristic intensity

Target Quality: DAS

EZMWF
GME
NCEP
UKMO

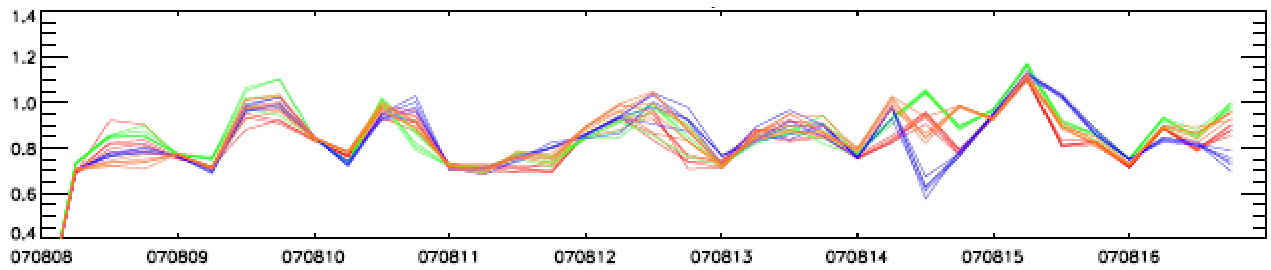
→ DAS at specific time is highly correlated with 6-h mean

DAS, hourly calculated



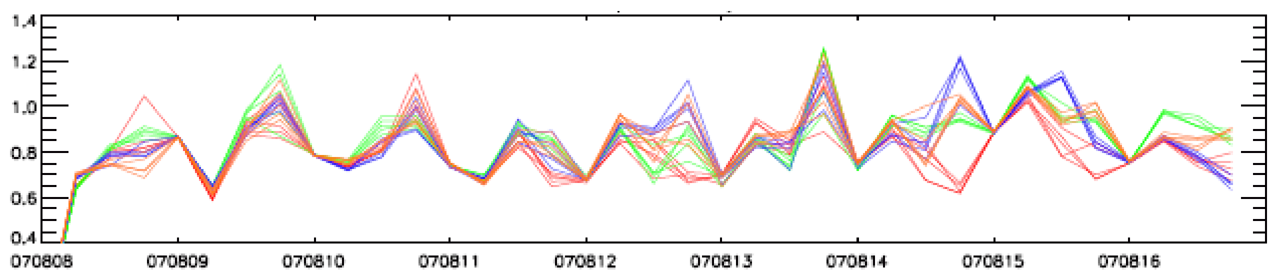
DAS,

6-hourly mean
(forward averaged)



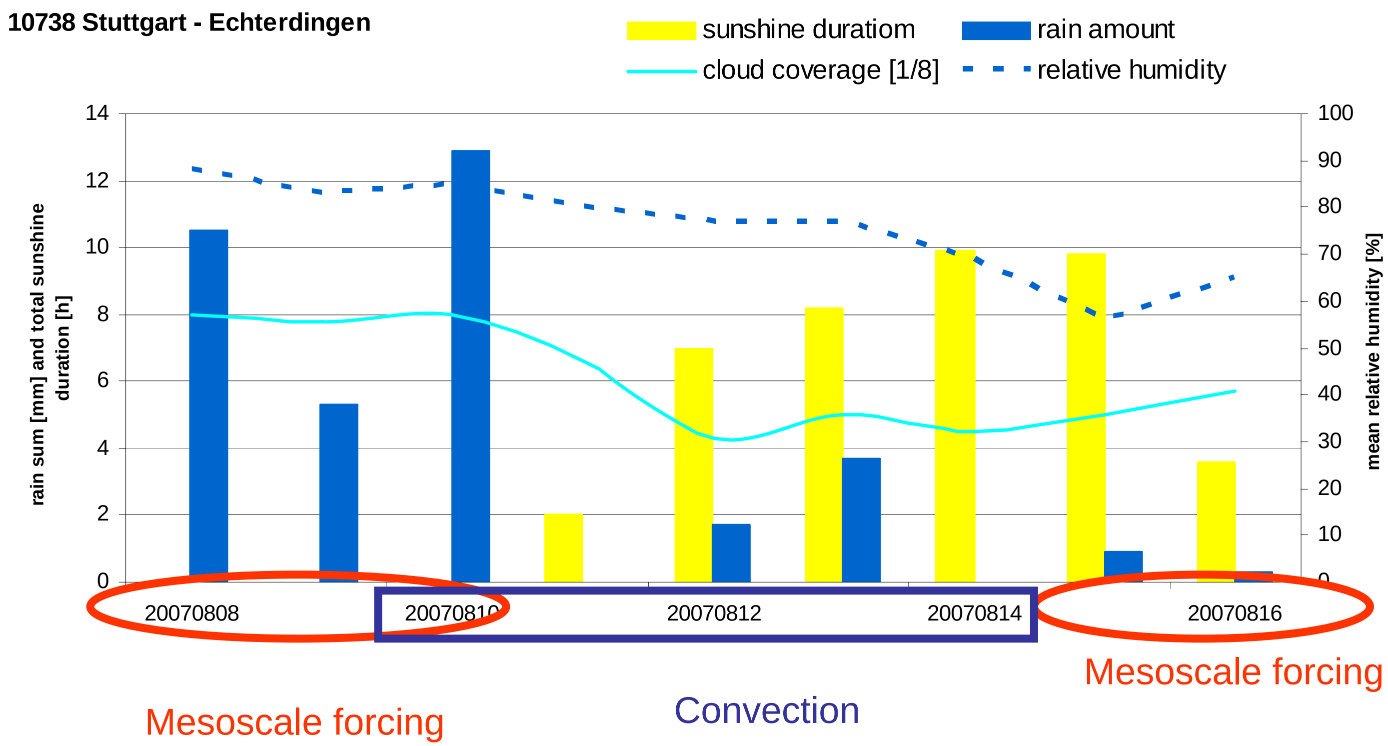
DAS,

every 6 h
(at start time of averaging period)



First Results

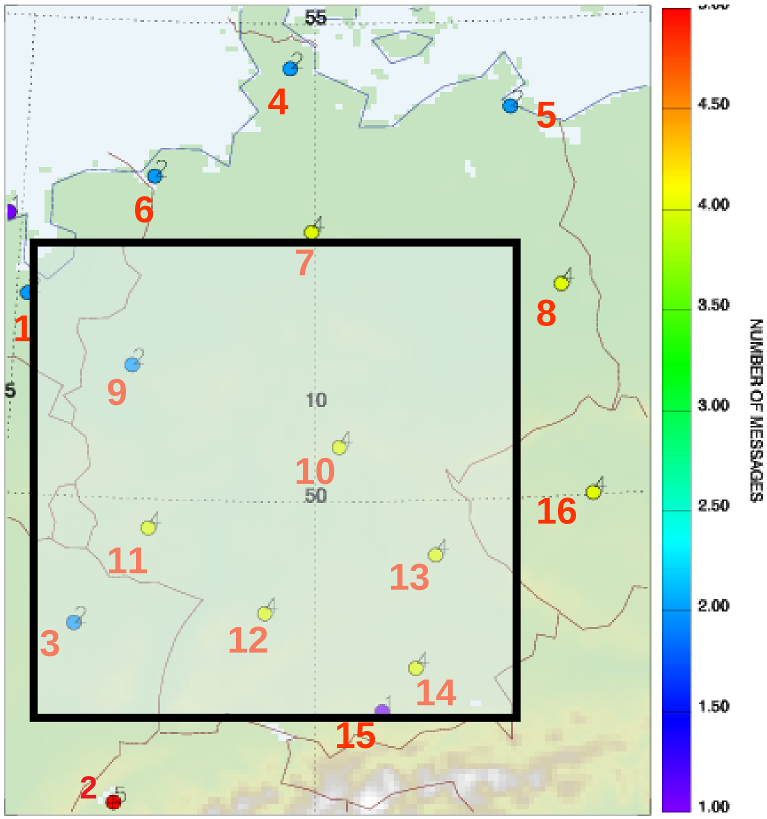
… investigated period: 08.-16.08.2007



First Results

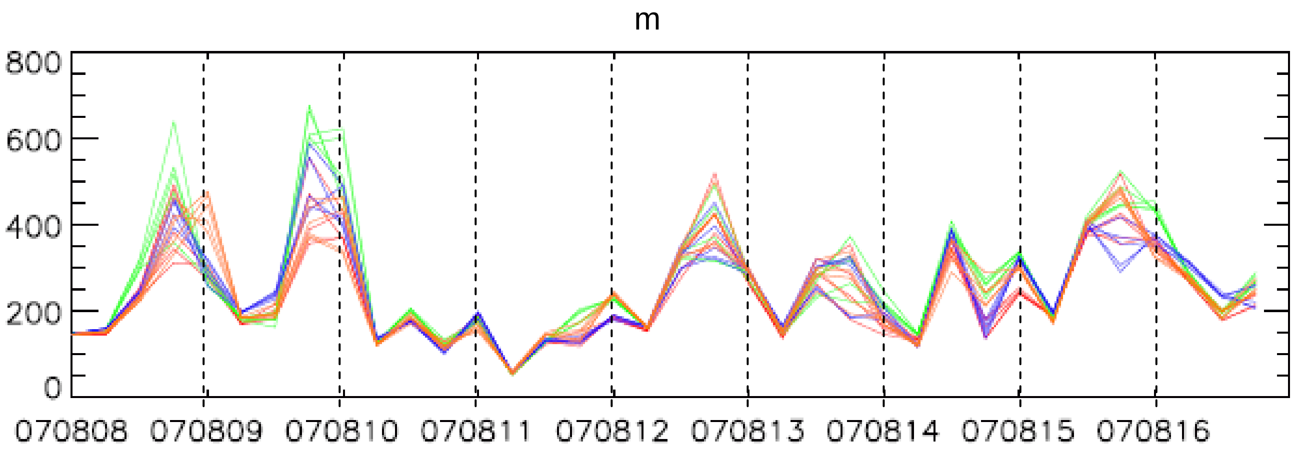
… investigated period: 08.-16.08.2007

… 8 radiosondes within considered DAS area

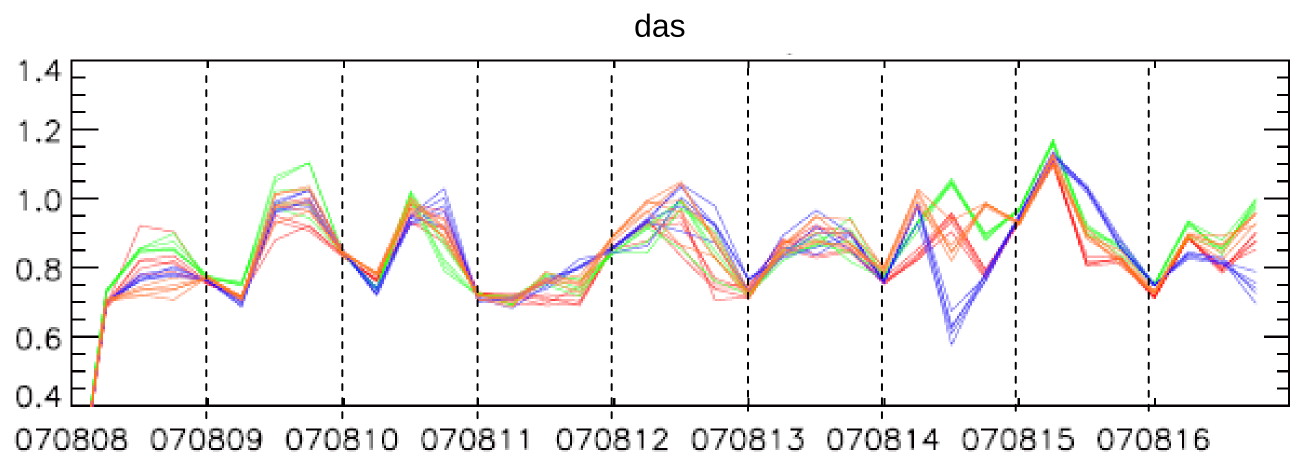


1 DeBilt
2 Payerne
3 Nancy
4 Schleswig
5 Greifswald
6 Emden
7 Bergen
8 Lindenberg
9 Essen
10 Meiningen
11 Idar-Oberstein
12 Stuttgart
13 Kuemmersbruck
14 Muenchen
15 Hohenpeissenberg
16 Prag

Equally weighted variables and levels



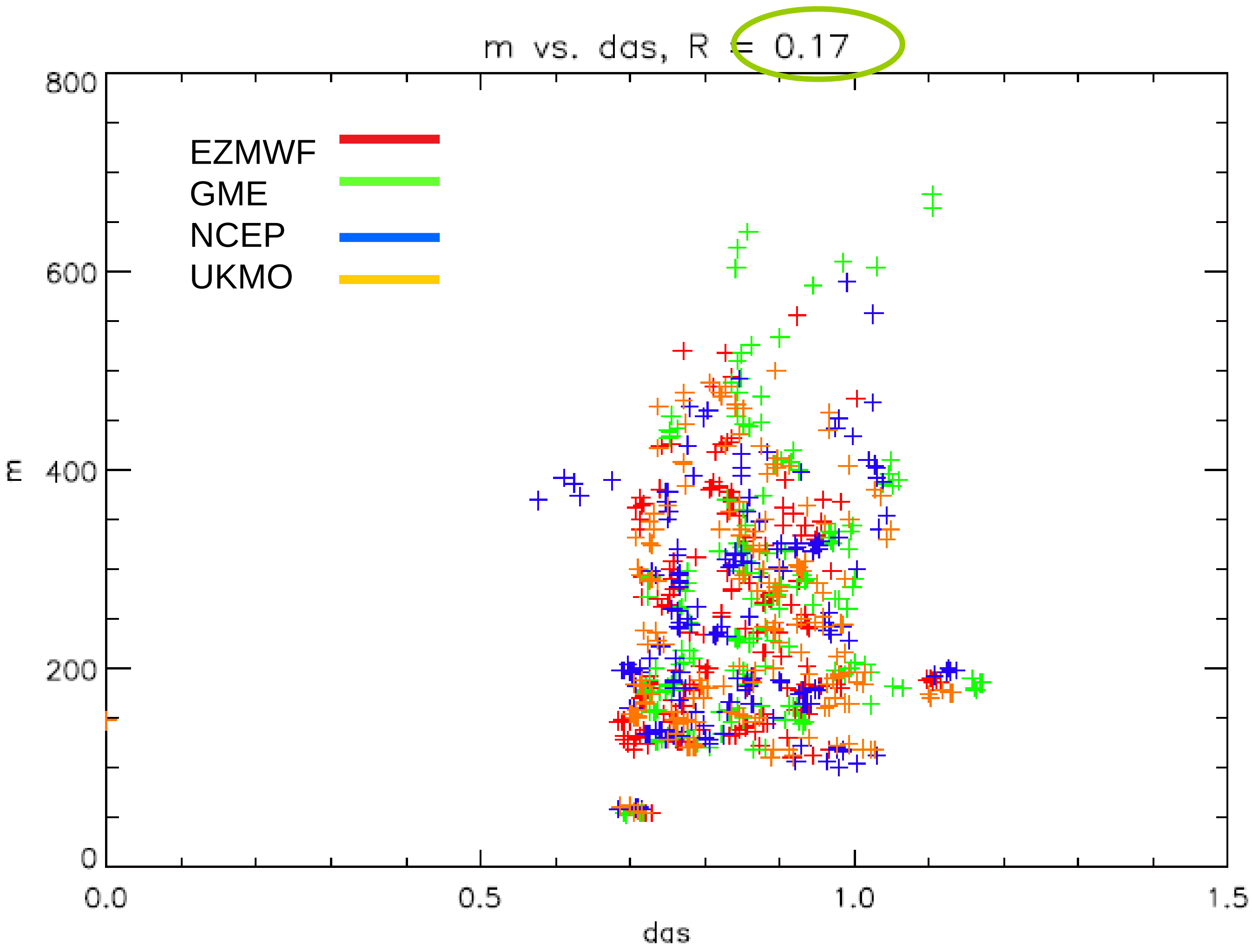
EZMWF
GME
NCEP
UKMO



Maximum in the afternoon, minimum in the morning (new run);

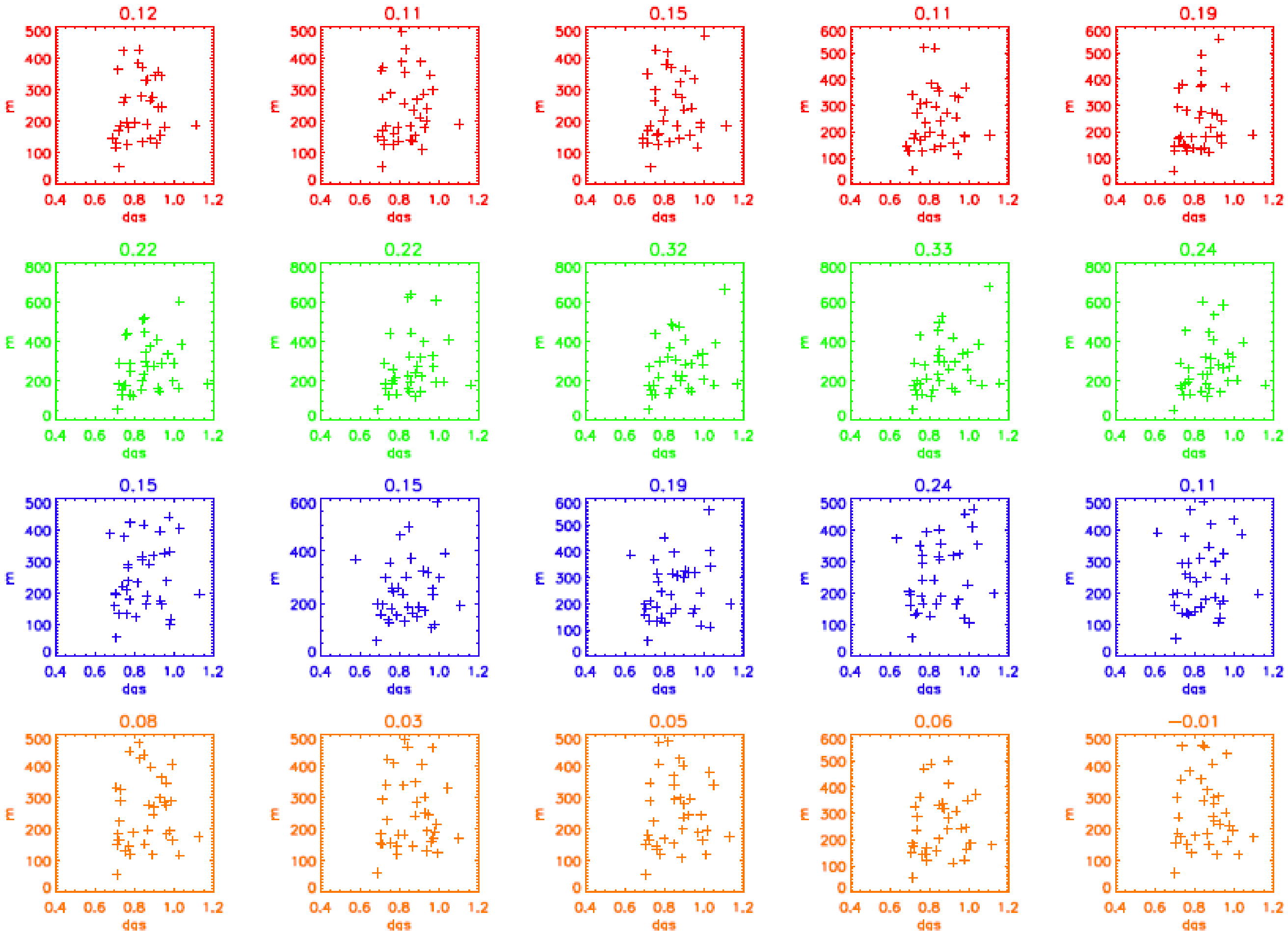
Partly clustering members with same driving models

Scatterplots (m vs DAS)



Scatterplots (m vs DAS),

Individual members



Range of correlation…

EZMWF

0.11 – 0.19

GME

0.22 – **0.33**

NCEP

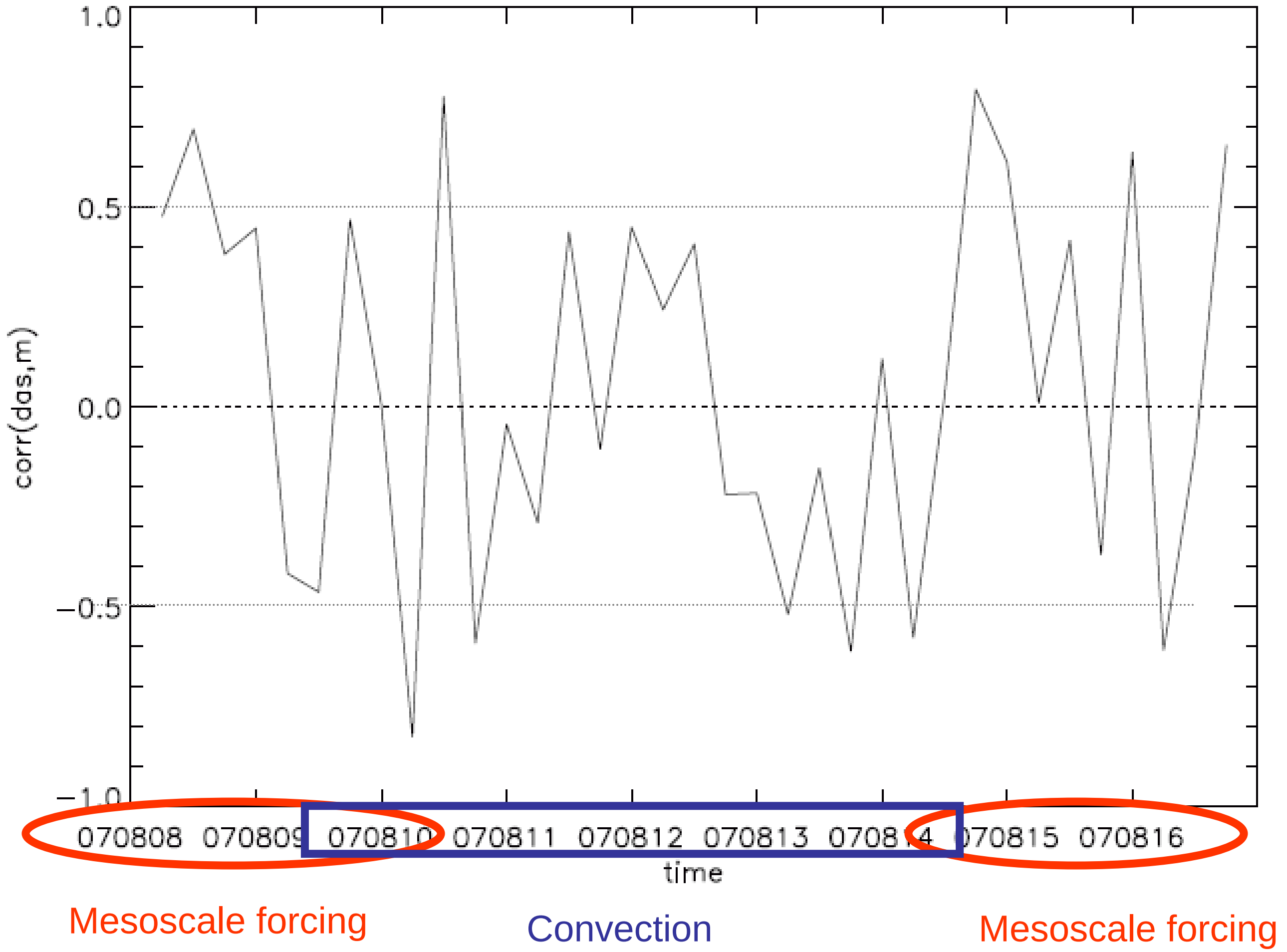
0.11 – 0.24

UKMO

-0.01 - 0.08

Correlation

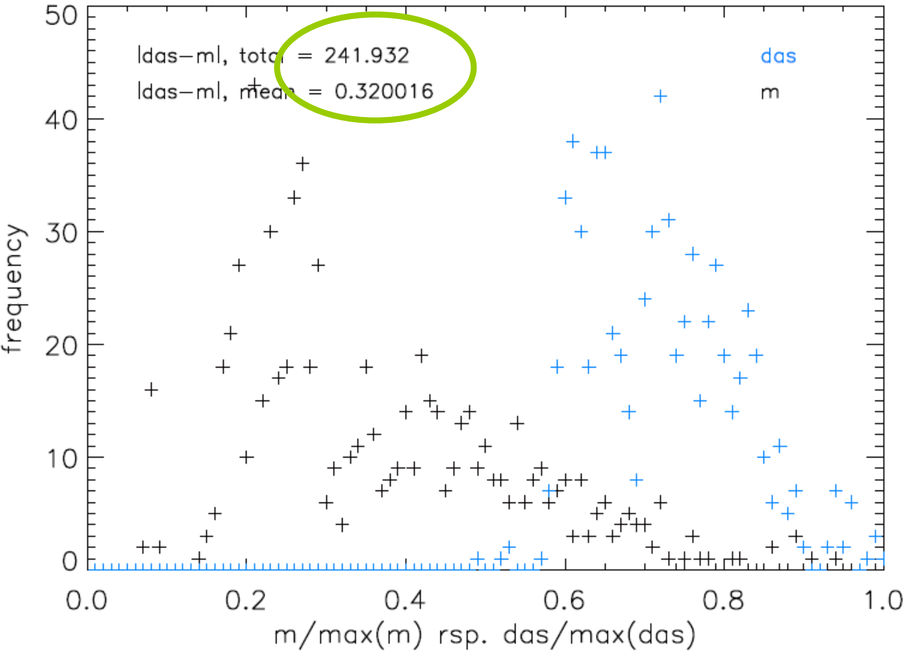
Individual time steps…



→ large variation of correlation with time

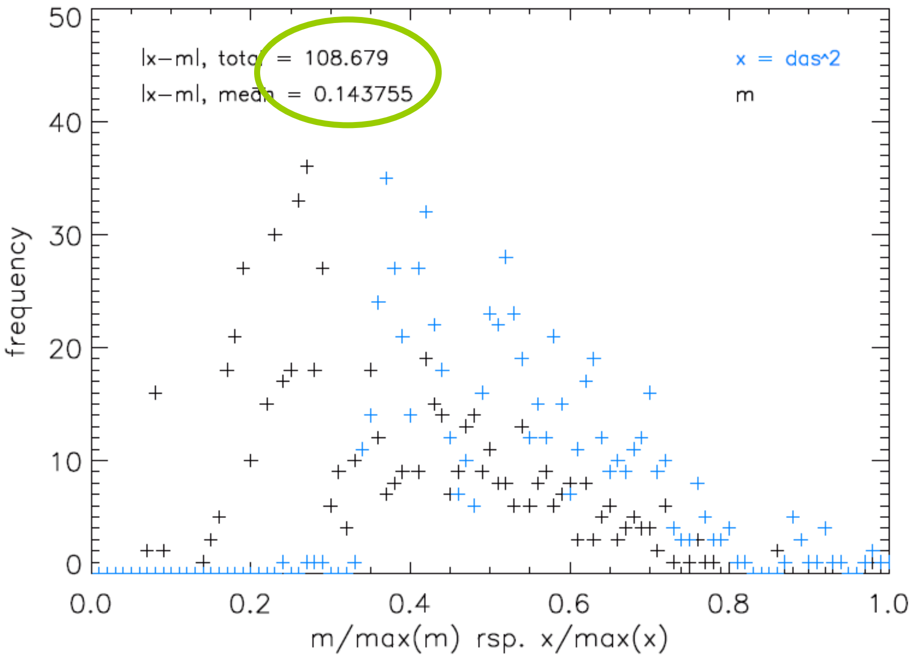
→ correlation apparently higher for mesoscale driven days…

Frequency distribution



histogram of normalized quality measures:

das/max(das) and m/(max(m))



same as above, but: das → $x=das^2$

→ PDF become more similar,

→ Decrease in |das-m|

Sensitivity studies

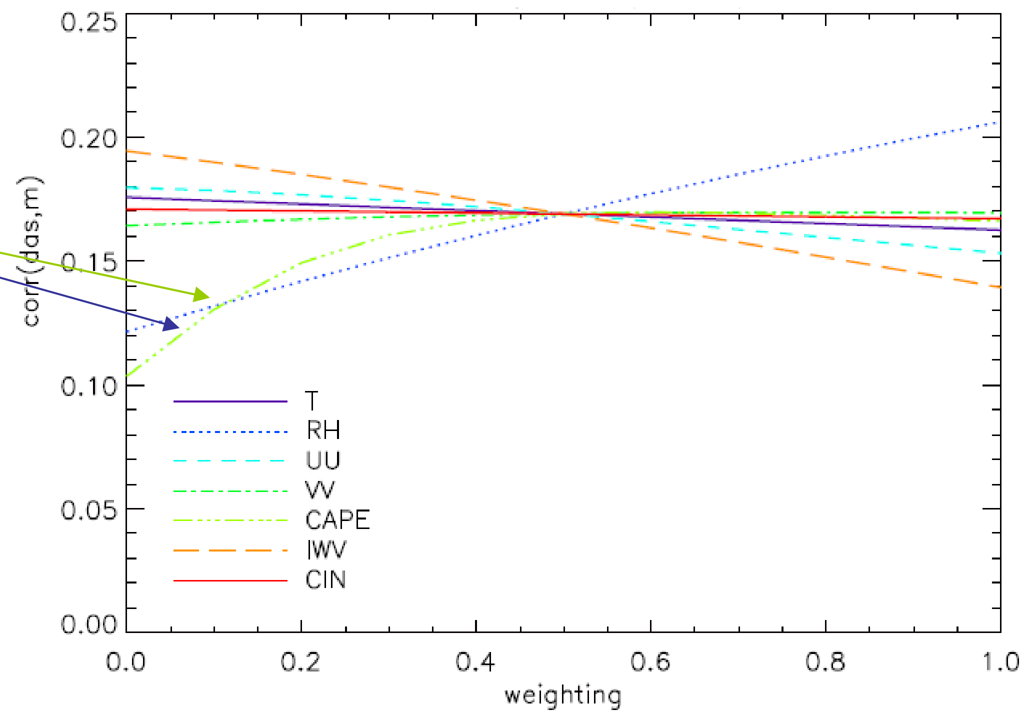
1. Variables

T	RH	U	V	CAPE	IWV	CIN
0.5	0.5	0.5	0.5	0.5	0.5	0.5

1000-700	650-400	350-100
1	1	1

Procedure: Set all variables to 0.5, vary each variable from 0 to 1.0 in steps of 0.1. Level weightings stay 1.

RH and CAPE show strongest (positive) response to weighting



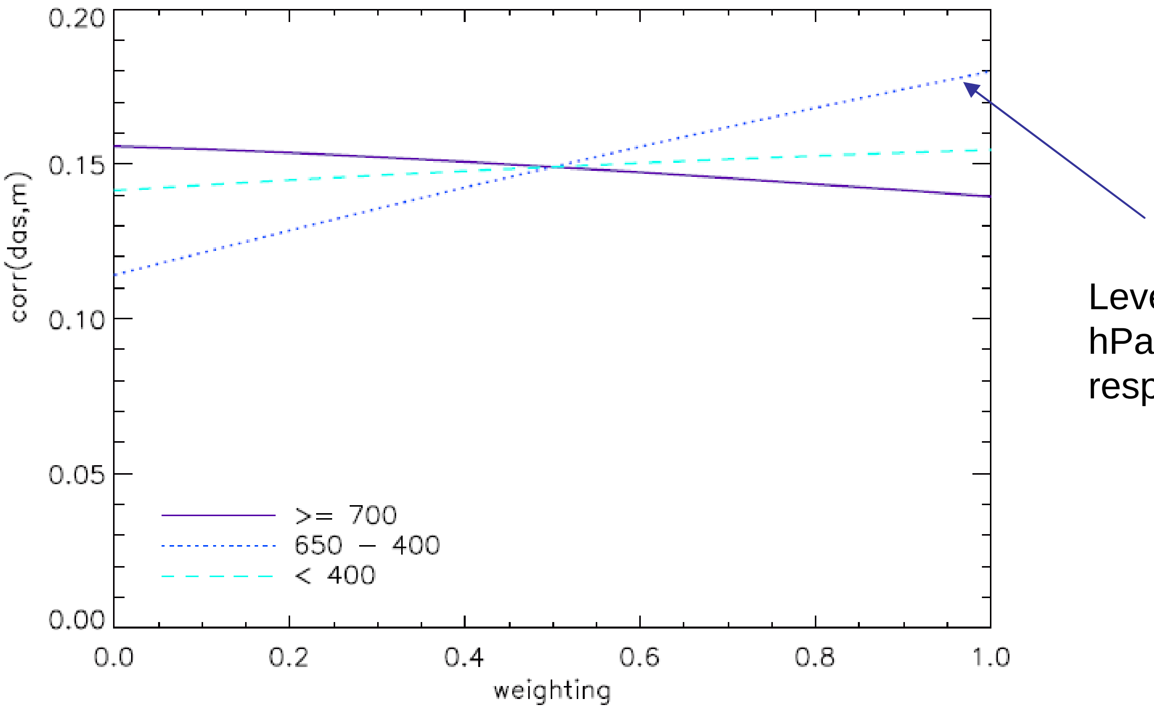
Sensitivity studies

2. Levels

T	RH	U	V	CAPE	IWV	CIN
1	1	1	1	1	1	1

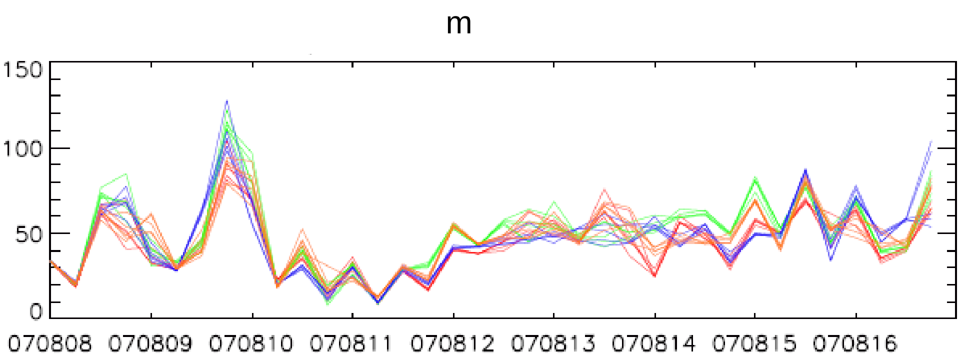
1000-700	650-400	350-100
0.5	0.5	0.5

Procedure: Set all levels to 0.5, vary each level from 0 to 1.0 in steps of 0.1. Level weightings stay 1.



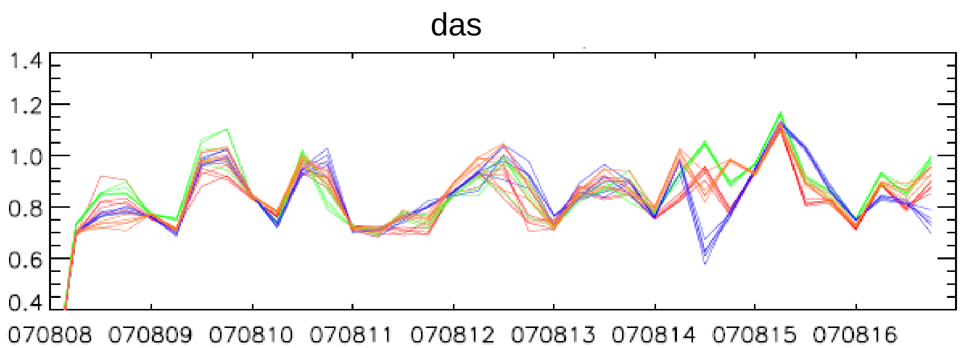
Levels between 650 and 400 hPa show strongest (positive) response to weighting

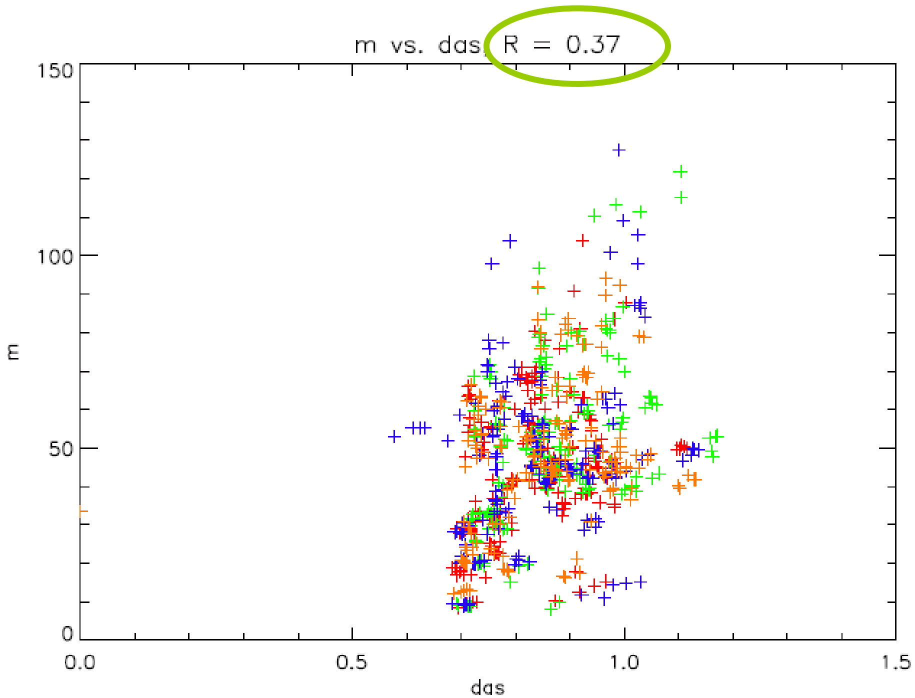
Combination of variables and levels …



T	RH	U	V	CAPE	IWV	CIN
0.4	1	0.4	0.6	0.1	0.1	0.1

1000-700	650-400	350-100
0	1	0





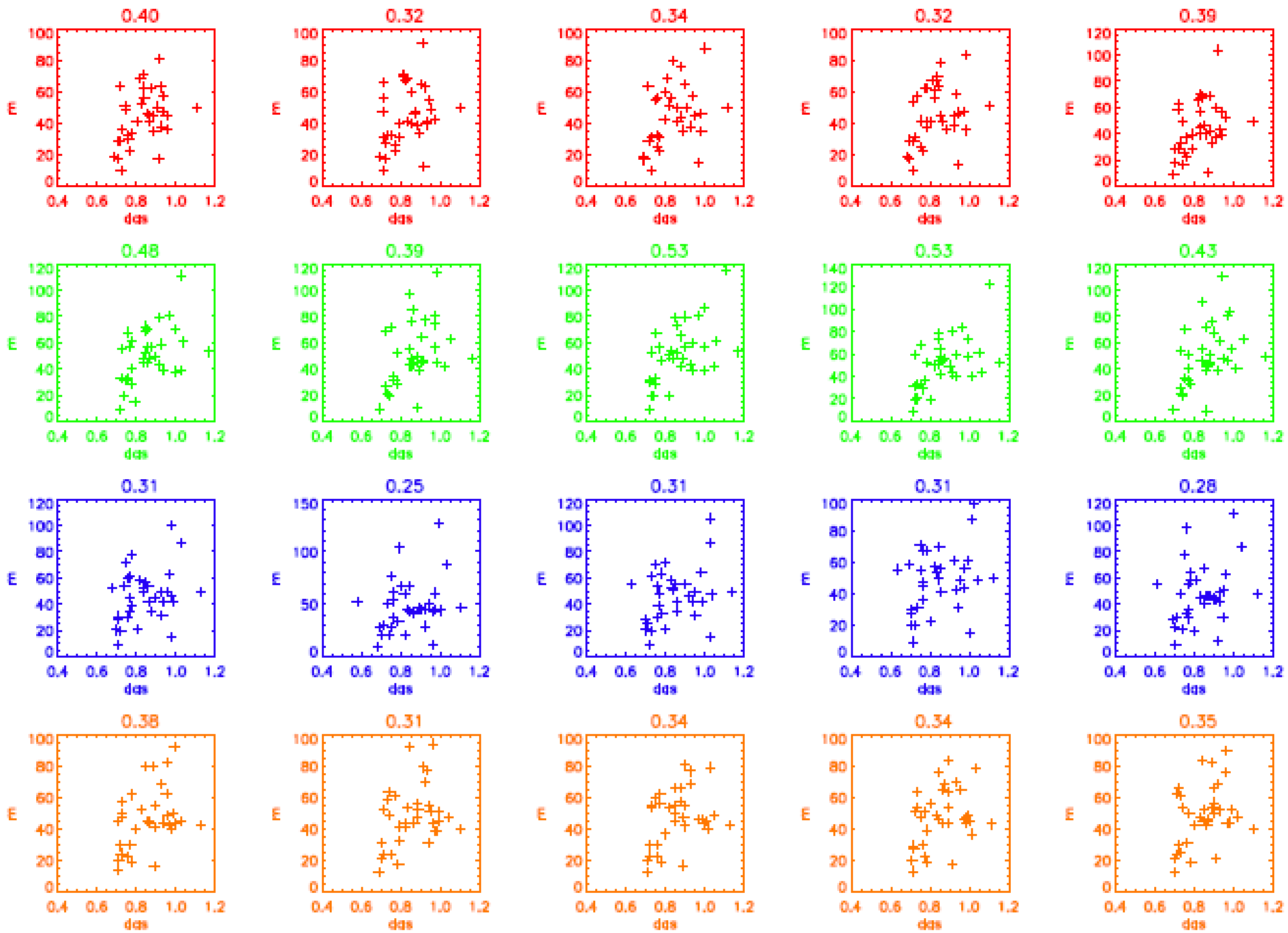
EZMWF
GME
NCEP
UKMO

Scatterplots (m vs DAS),

Individual members

Range of correlation…

(for equally weighted)



EZMWF

0.32 – 0.40
(0.11 – 0.19)

GME

0.39 – **0.53**
(0.22 – 0.33)

NCEP

0.25 – 0.31
(0.11 – 0.24)

UKMO

0.31 - 0.38
(-0.01 - 0.08)

Comparison

- Equally weighted

T	RH	U	V	CAPE	IWV	CIN
1	1	1	1	1	1	1

1000-700	650-400	350-100
1	1	1

- Combined weighting

T	RH	U	V	CAPE	IWV	CIN
0.4	1	0.4	0.6	0.1	0.1	0.1

1000-700	650-400	350-100
0	1	0

	Equally weighted	Combined weighting
Corr(m,das)	0.17	0.37
TOTAL(|das-m|)	241.93	242.62
Corr(m,das²)	0.16	0.36
TOTAL(|das²-m|)	108.68	109.36

→ Rarely a change in distance |das - m|, but increase in correlation!

Preliminary Fazit

- Radiosondes do give information on best members in terms of precipitation forecast (as measured by DAS).
- Test with radiosoundings reveal a sensitivity of the best member selection with respect to parameter and heights.
- The correlation with DAS could be improved by using radiosondes data with different weightings for variables and heights.
- The correlation between the two quality measures DAS and m is apparently time-dependent.
- The correlations are in the order of 0.4, while DAS values at starting time of averaging period are highly correlated with DAS averaged over next 6 hours

  → the precipitation measure DAS provides a good basis for the best member selection. The use of standard observation as alternative quality measure seems to be appropriate.

Outlook

- find suitable combination of only a few parameters from radiosondes and synop stations (e.g. total energy)
- determine optimal weighting of those parameters with respect to target quality measure r
- expand best member selection to driving ensemble

Thank you for the attention!

The Displacement and Amplitude Score DAS (I)

- DAS is based on an areal image matcher using classical optical flow technique

- DAS has two components
  - displacement error field(of observed and forecast imagery)

$$\overline{DIS_{obs}} = \frac{1}{n_{obs}} \sum_{A} DIS_{obs}(x,y)$$

  - amplitude error (RMS of observed an morphed forecast imagery)

$$\overline{AMP_{obs}} = \left[\frac{1}{n_{obs}} \sum_{A} AMP_{obs}(x,y)^2 \right]^{1/2}$$

The Displacement and Amplitude Score DAS (II)

- Application in forecast space to account for false alarms

- The total amplitude error AMP and displacement error DIS are then defined by weighted averages:

$$AMP = \frac{1}{(n_{obs} + n_{fct})} \left(n_{obs} \overline{AMP_{obs}} + n_{fct} \overline{AMP_{fct}} \right)$$

$$DIS = \frac{1}{(n_{obs} + n_{fct})} \left(n_{obs} \overline{DIS_{obs}} + n_{fct} \overline{DIS_{fct}} \right)$$

Reference: Keil, C. and G. C. Craig, 2009: A displacement and amplitude score employing an optical flow technique, submitted to *Wea. and Forecasting*.

Calculation of variables

Integrated Water Vapour

$$IWV = \int \rho_w dz = \rho \int q\, dz$$

q = spezifische Feuchte

CAPE, CIN…

(max CAPE)