The challenge of finding „good“ reference data for verification

T. Gorgas, M.Dorninger

Department of Meteorology and Geophysics, University of Vienna, Austria

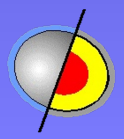
Why challenge?

„Verification“ is forecast against „truth“

find out about forecast errors and uncertainties

BUT

- Data defined as „truth“ only estimate the real state
- Measures derived from verification describe not only forecast errors, but also uncertainties in the reference data
- Even if both, reference data and forecasts are „right“ verification measures might result errors caused by matching problems between forecast data and reference data.
- Different resolutions, locations, representativity, topography, parameters…



What can we do?

Learn more about quality characteristics of reference data!

Find out about the usefulness of analyses and observations as reference data for verification!

Compare alternative data sets!

Data Sets applied at U Vienna: (Period: 2007)

- VERA (Vienna Enhanced Resolution Analysis)
  - arbitrary choice of target areas and resolutions (4,8,16km)
  - downscaling via „fingerprint“ method
  - „model independent“ – no NWP first guess field needed
- JDC (Joint D-PHASE and COPS) Observational Data Set
  - Operational surface observations of Central Europe for 2007

Alternative Data:

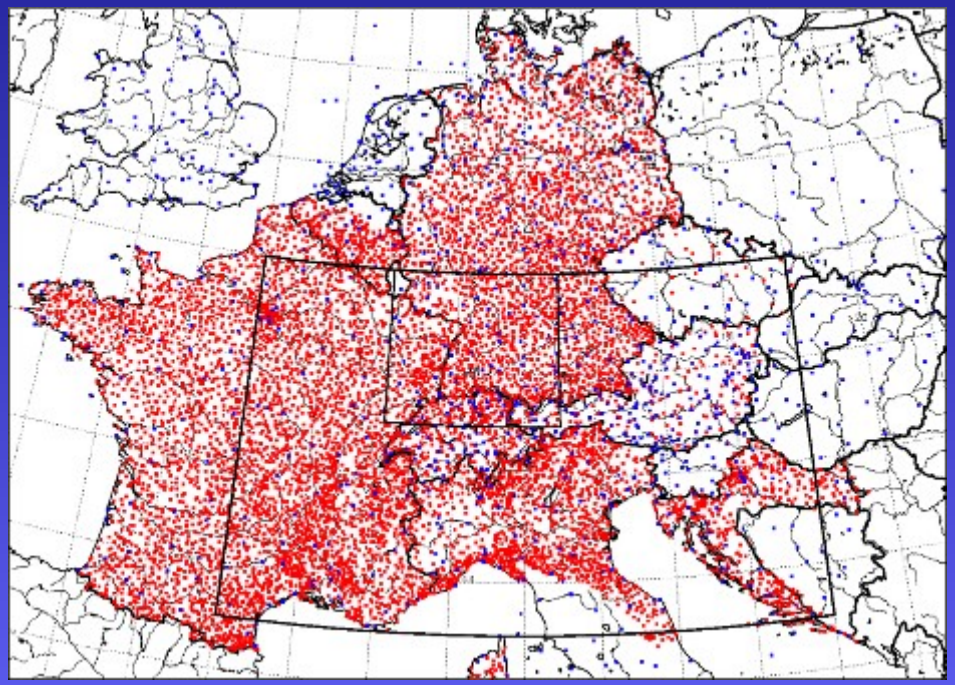
- NWP-model analyses
- Kriging
- Simple Analyses (p.e. Cressman)
- Area averaged VERA

Joint D-PHASE and COPS data set – Overview

Collection of operational networks of National Weather Services initiated in the framework of the WWRP programmes COPS and D-PHASE

Available at WDCC Hamburg according MAP Data Policy

Task performed in a cooperation of U Vienna (M.Dorninger, T. Gorgas) and U Hohenheim (T. Schwitalla)



- GTS-Stations: 1232
- NGTS-Stations: 15665
- NGTS-Stations minus double stations: 10811
- Mean station distance: (imagine a 1.6 mil. km^2 Central Europe):

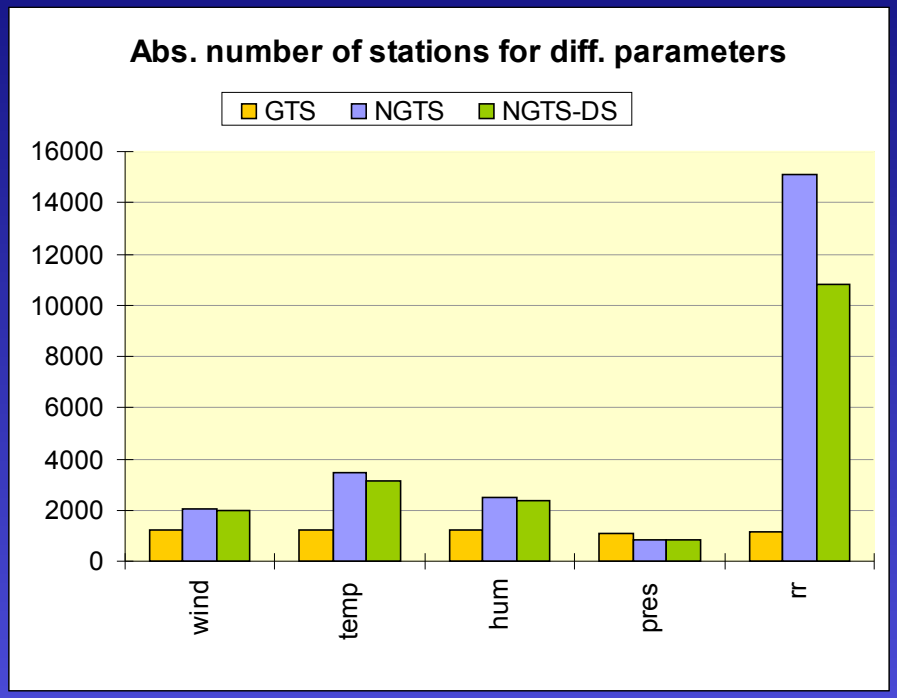
  GTS: ~ 36km
  GTS+Non-GTS: ~ 12km

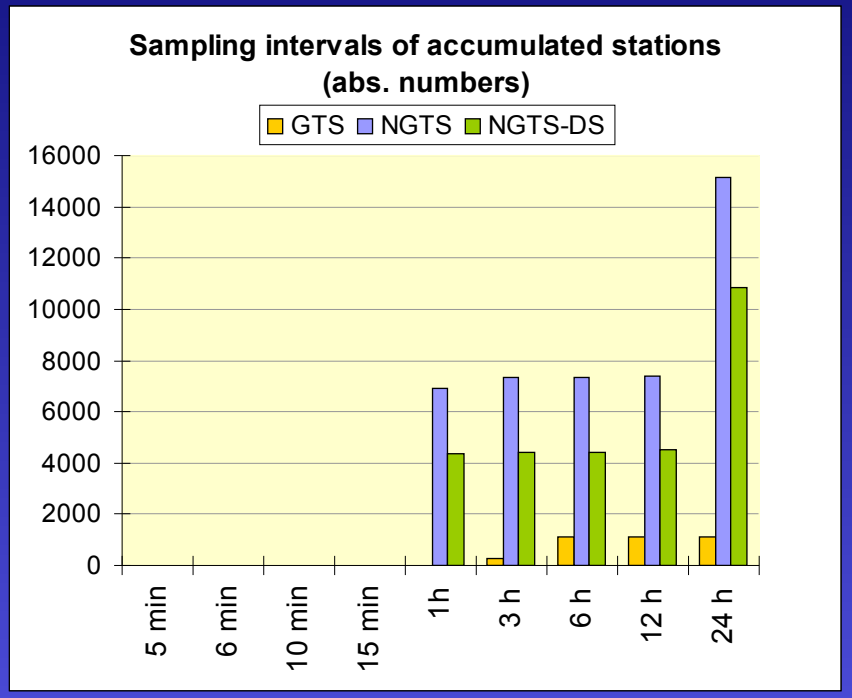
Frames: D-PHASE (large) & COPS (small) areas

Red: Non-GTS stations

Blue: GTS-stations

Joint D-PHASE and COPS data set – Parameters

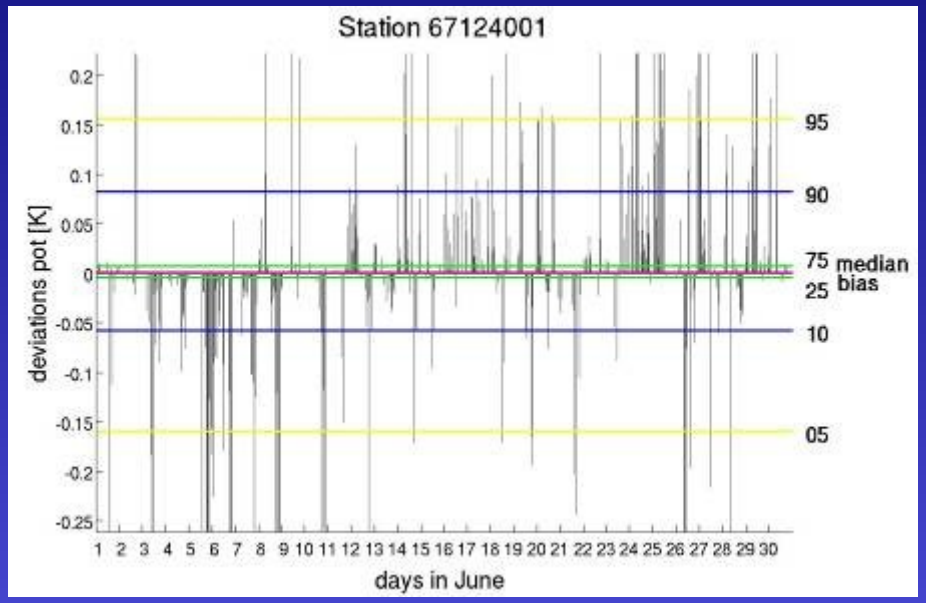




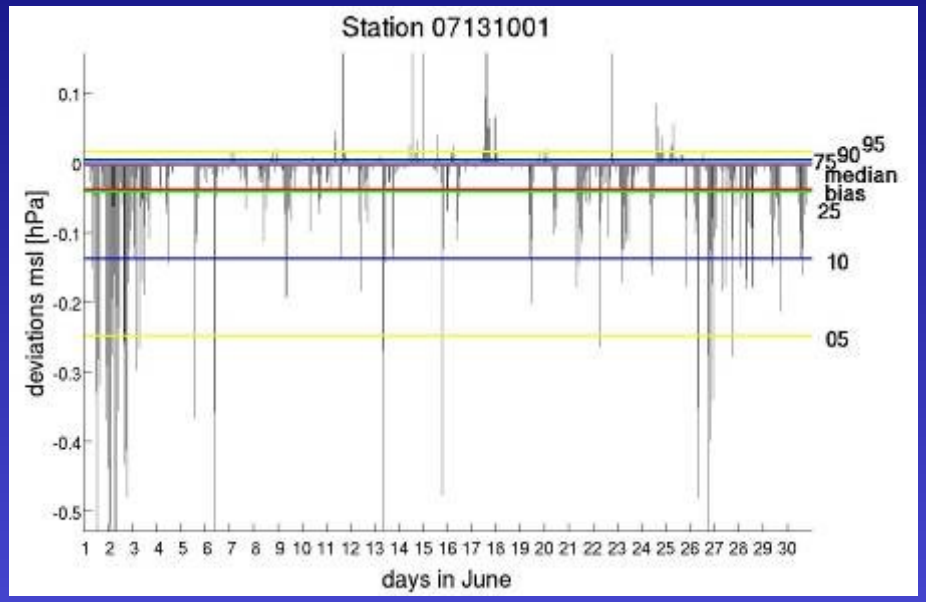
Precipitation-Stations available:	Up to 1h:	Up to 6h:	Up to 24h:
	6900 Non-GTS Stations Only Non-GTS available	1232 GTS + 7335 Non-GTS Typical interval of GTS	1232 GTS + 15132 Non-GTS + Climatological networks

Quality Control of Parameters to Analyse

deviations of potential temperature for 06/2007



deviations of msl-pressure for 06/2007



Quality Control Scheme, U Vienna

(D.Mayer, R.Steinacker)

- **Operates on a variational basis, 2D**
- **Generates weighted deviations for each station & each parameter at any given analysis time**
- **For analysis use: values are corrected if deviations exceed certain thresholds**
- **Mean deviations (biases) can be applied to improve analysis quality**

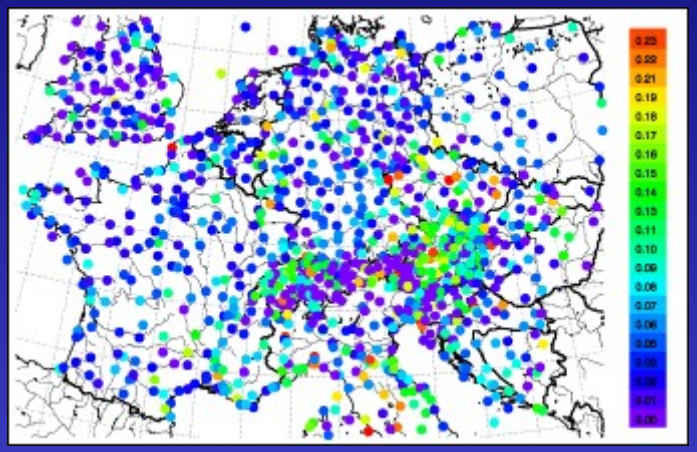
Estimated variabilities in observations

Standard deviations of hourly quality control outputs for 2007/06

MSL-pressure

Range:

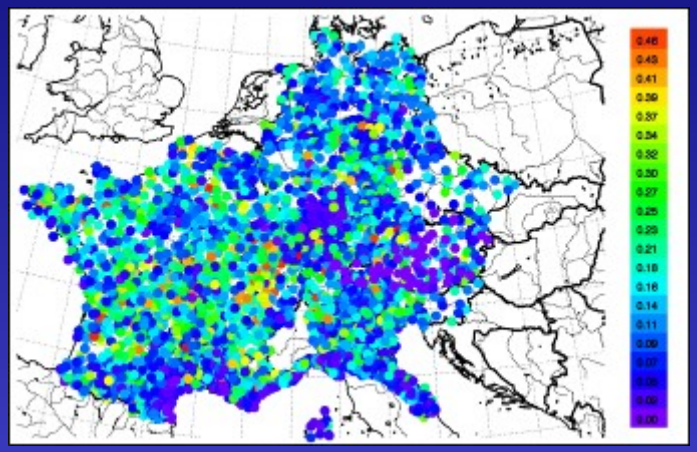
0.00 hPa
0.23 hPa



RR 1h acc.

Range:

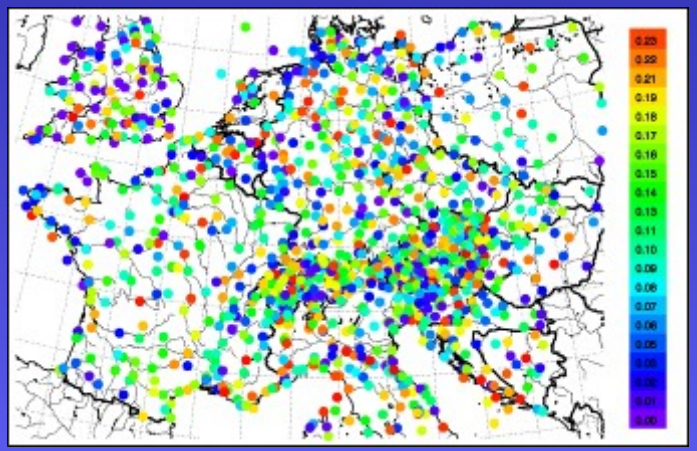
0.00 mm
0.46 mm



Eq.Pot. Temp.

Range:

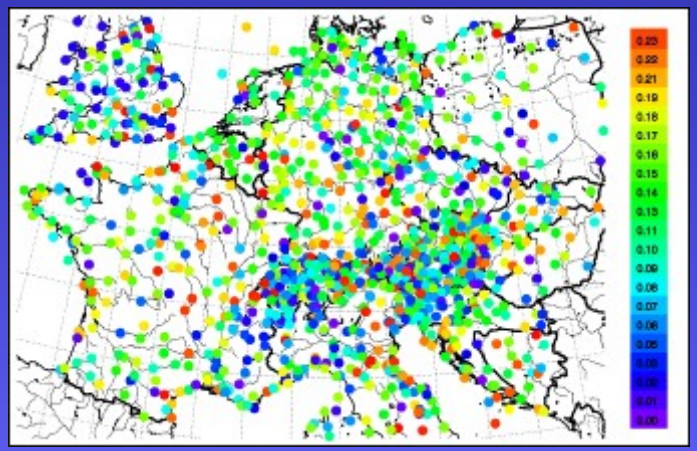
0.00 K
0.23 K



Pot. Temp.

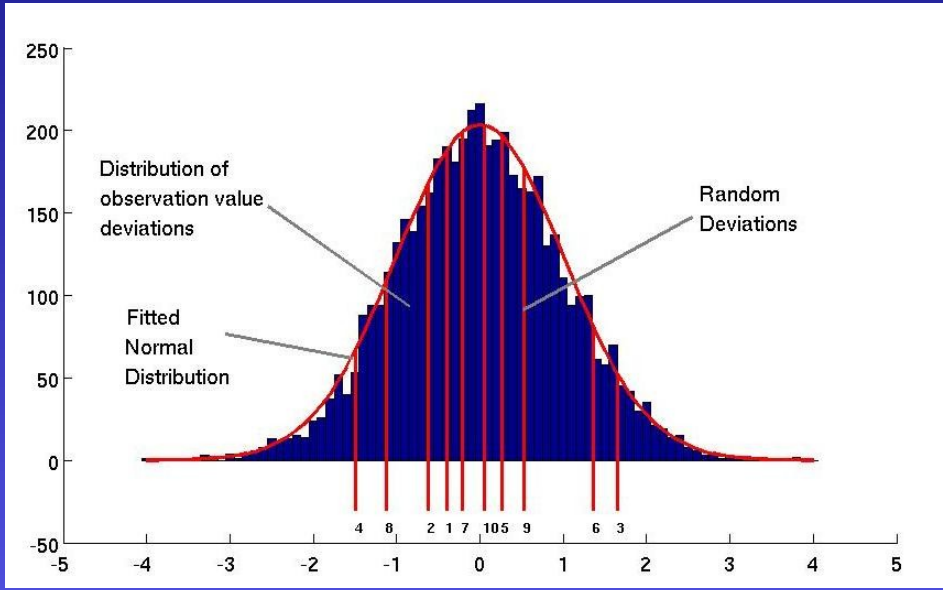
Range:

0.00 K
0.23 K



Estimation of uncertainties in VERA analyses - a very first approach

Schematic randomisation procedure performed for each station and parameter

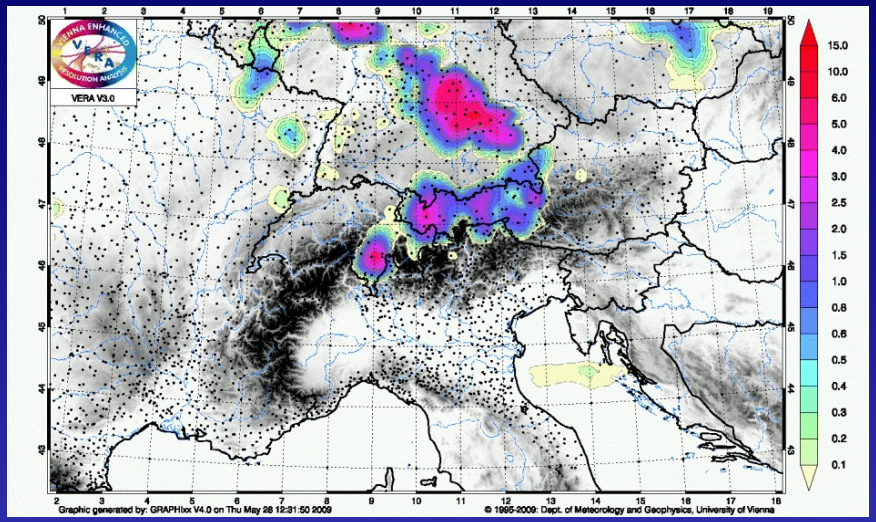


First experiments: Choose sets for 10 Ensemble Members

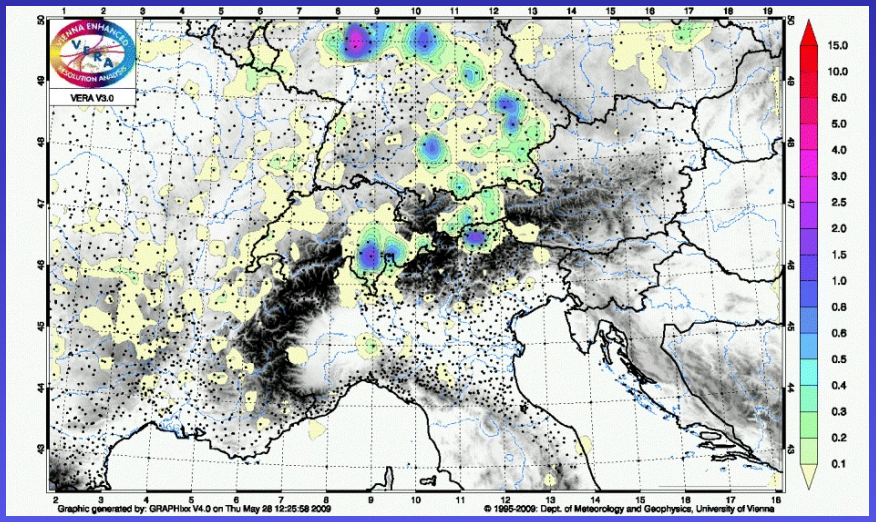
Steps towards ensemble analyses

- Correct station observation values by removing biases dereived from deviations proposed by quality control
- Analyse bias-corrected observations = reference analysis
- Generate normal distribution fitted to distribution of quality control outputs
- Create a number of sets of (gaussian) randomized observation values
- Use perturbated data to create ensemble analyses

2007062112 8km RR [mm/h]

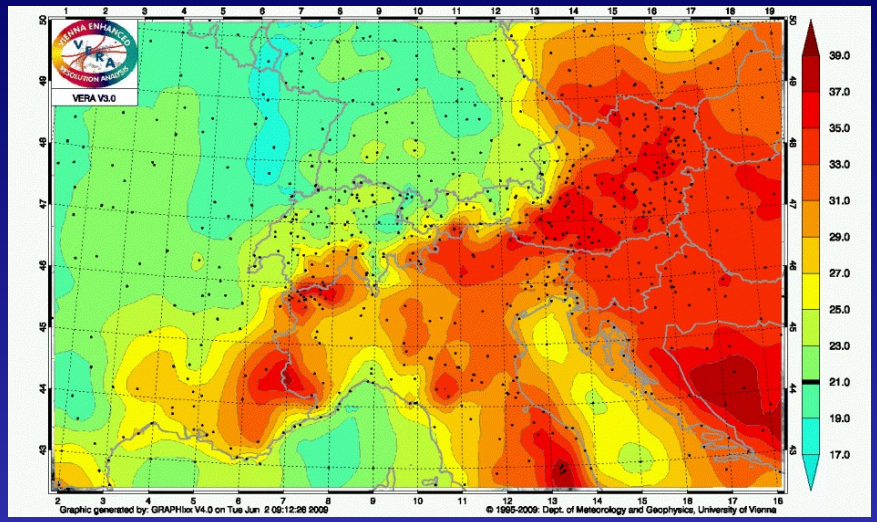


Analysis RR 1h acc.

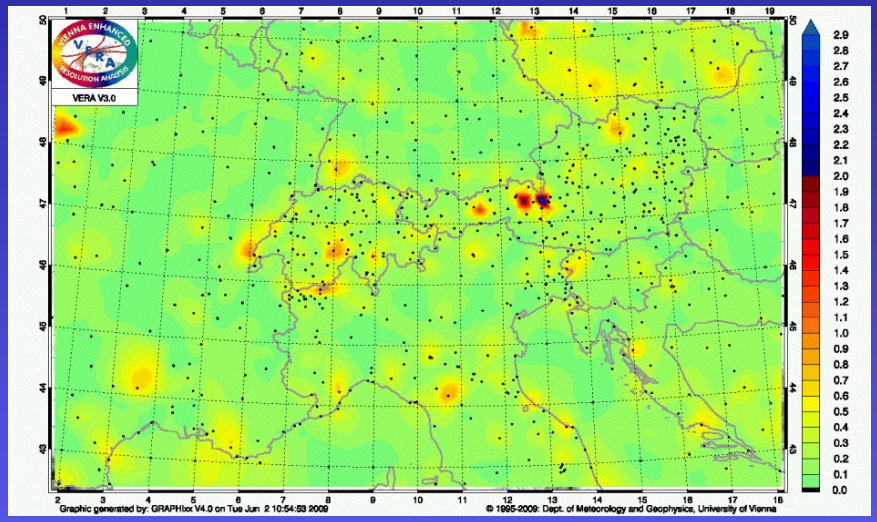


Stdev. of Ens. Members (10)

2007062112 8km Pot. Temp. [K]



Analysis Pot. Temp.



Stdev. of Ens. Members (10) – Max: 2.9 K

Implement quantified uncertainties in basic verification?

First Experiments with Gaussian Error Propagation

$$\delta G \approx [\,(\frac{\partial f}{\partial x}\delta x)^2 + (\frac{\partial f}{\partial y}\delta y)^2 + (\frac{\partial f}{\partial z}\delta z)^2 + \ldots\,]^{1/2}$$

- Use ensemble standard deviations as error estimation for verification data
- Treat each grid point value as independent component with a particular error estimate
- First approach: Neglect models‘ interpolation errors for comparison

Dependencies considered for partial derivatives

- Bias … (yi)
- RMSE … (yi)
- Stdev (yi) … (yi)
- Corr. … (yi, Stdev(yi))

yi … values of the analysis fields

Simple comparison of NWP-model and VERA

- Calculate appropriate met. parameters from direct model output
- Interpolate model gridpoint values on the VERA grid using inverse distance
- Use adaption procedures for different topographies

Comparison of COSMO2 -12h forecast (2.2km) to VERA at 2007062112

	VERA 4km			VERA 8km			VERA 16km		
	Pot. Temp	MSL-Pressure	Wind Vel	Pot. Temp	MSL-Pressure	Wind Vel	Pot. Temp	MSL-Pressure	Wind Vel.
Bias	0.08	-0.04	0.69	0.08	-0.04	0.69	-0.03	0.02	0.65
EU. Bias*	1.2	1.6	0.2	2.5	3.5	0.4	14	18	0.9
RSME	2.08	1.55	1.60	2.14	1.56	1.67	2.18	1.57	1.66
EU. RMSE*	0.06	0.06	0.07	0.14	0.13	0.16	0.22	0.20	0.36
Correlation	0.92	0.88	0.66	0.91	0.88	0.64	0.91	0.88	0.65
EU. Corr.*	0.02	0.03	0.14	0.04	0.04	0.27	0.08	0.08	0.6
Stdev. V	5.13	3.22	1.59	5.14	3		.24		1.70
Stdev. M	5.12	2.58	1.86	5.15	2		.58		1.91
EU. Stdev. V*	0.02	0.02	0.12	0.04	0		.07		0.51

Estimated Uncertainties (EU) in %

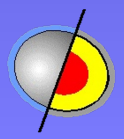
With manipulation to avoid „factor 2“:

4km: 0.12%
8km: 0.16%
16km: 0.18%

Larger differences between 4,8,16km than amounts of uncertainties

Outlook

- Continue with ensemble analyses with more members and different perturbation methods
  - Experiments with the magnitude of station value perturbation
  - Try variation of station density
  - Cross validation
- Find methods to estimate and implement interpolation errors in the model forecast fields
- Compare with alternative analysis methods and observation networks
- Find out about the spatial representativity of station observations
- Get more detailed information about the scales represented in analyses, forecast fields and observations.



Deutsche Forschungsgemeinschaft DFG

SPP 1167

Study of the process chain and predictability of precipitation by combining the D-PHASE ensemble and the COPS data sets in the COPS domain

FWF Der Wissenschaftsfonds.

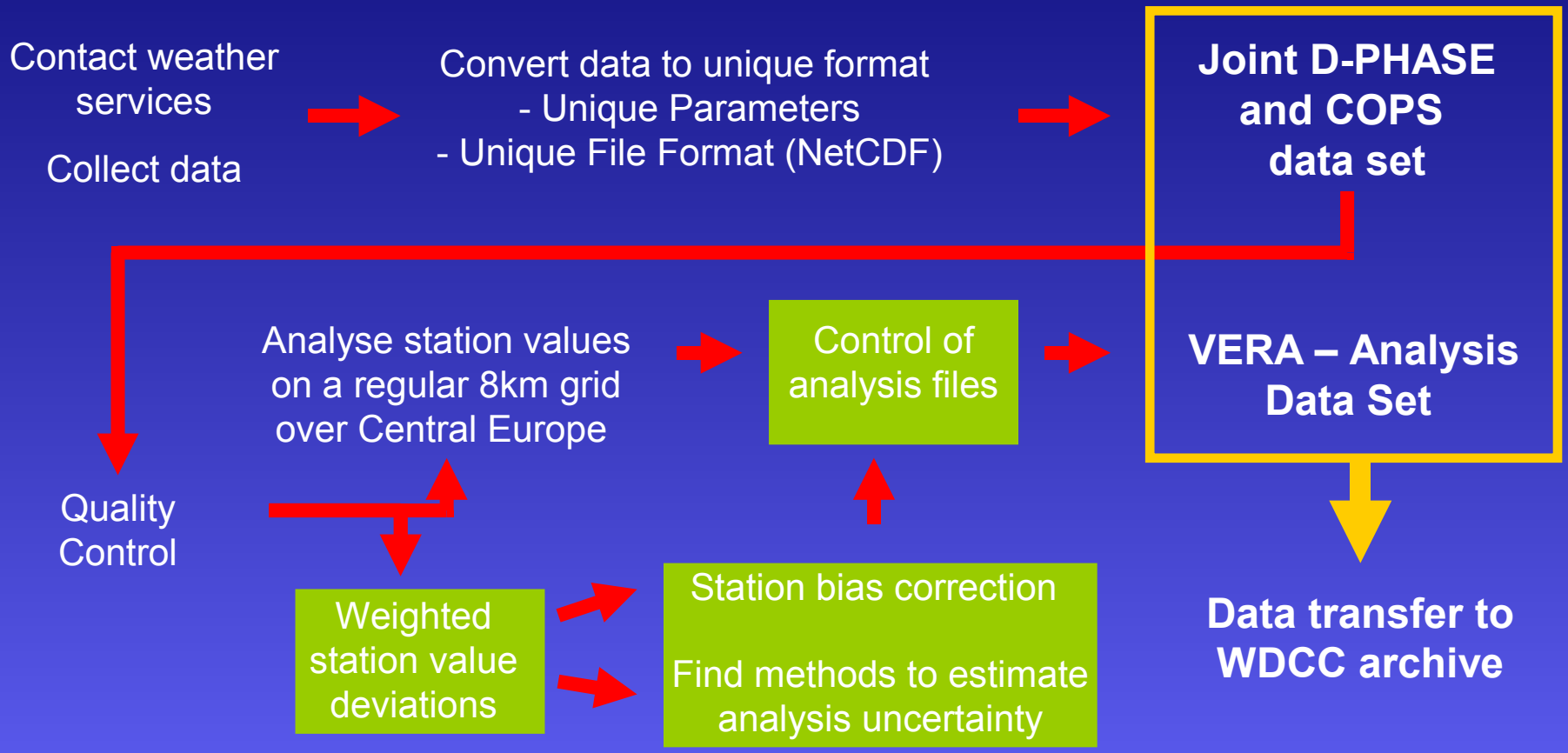
VERITA

(NWP model verification over complex terrain with VERA)
(P20925 – funded by FWF)

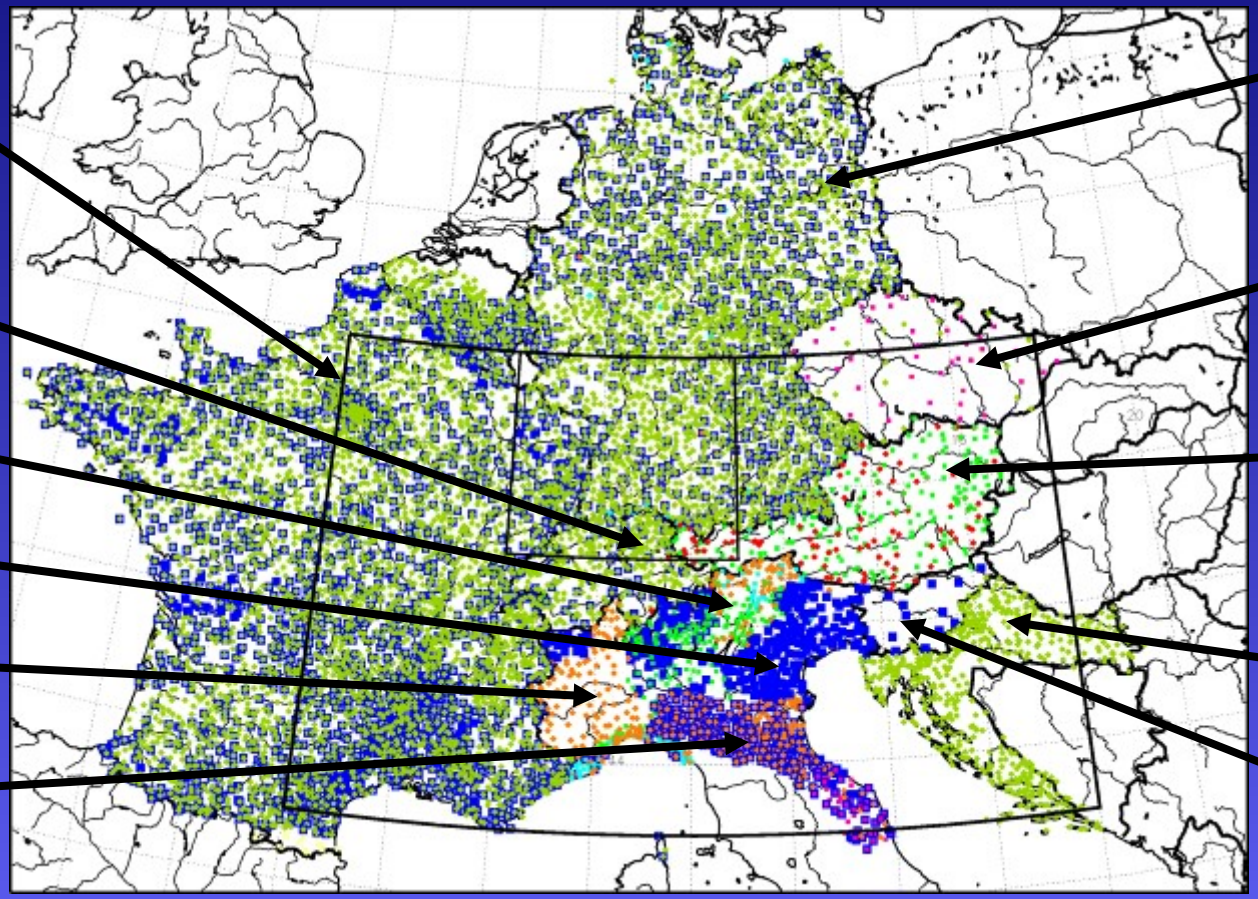
Ongoing projects

Thank you for your attention !

Overview of tasks and work steps

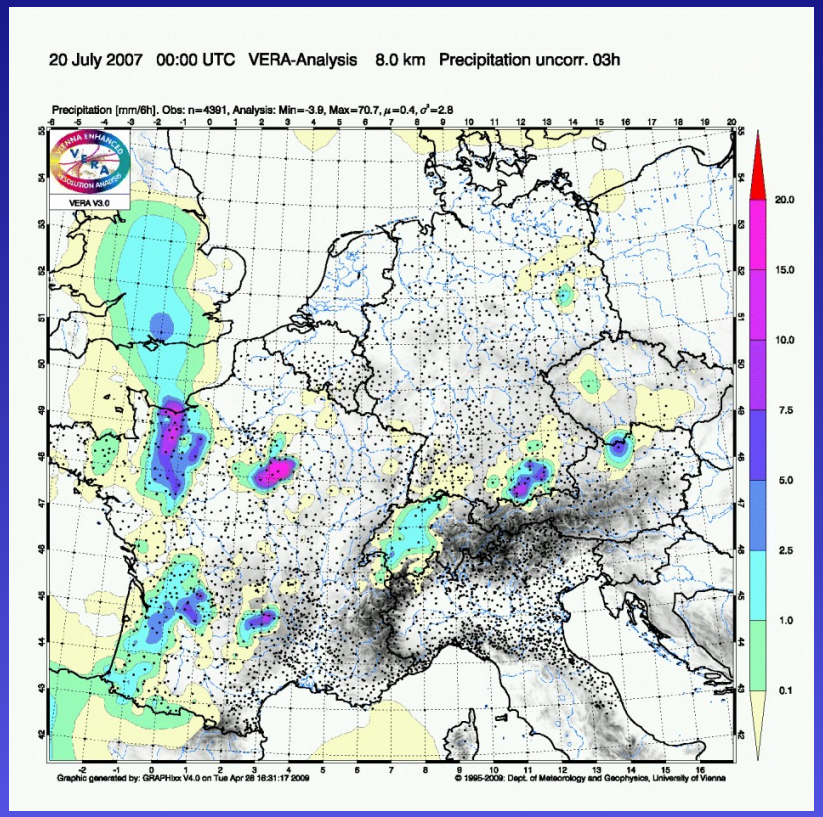


Sampling intervals of precipitation observations

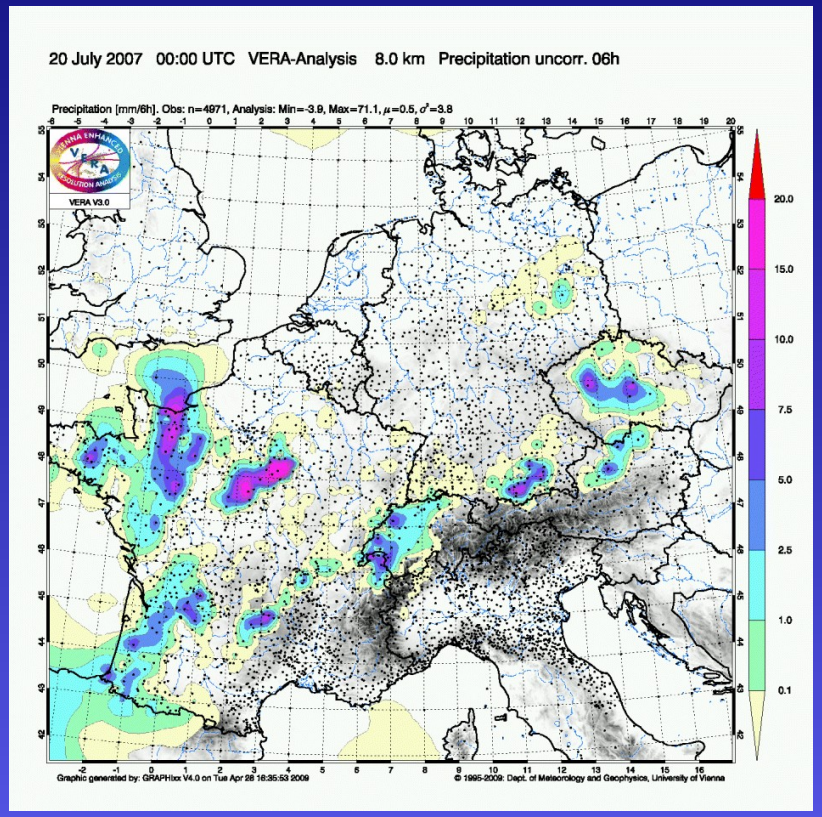


VERA Analyses for IOP9c - Precipitation

20.07.2007



3-hourly, 00 – 21 UTC



6-hourly, 00 – 18 UTC