Norwegian Meteorological Institute

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Assessing uncertainty in regional reanalysis by comparing against high-resolution gridded observations

Cristian Lussana and Ole Einar Tveito

28.01.15

Task 3.1 Coordinated uncertainty evaluation

- D 3.2 Common evaluation procedures (-> data source):
  - **A: feedback statistics**
    - Data source: radiosonde soundings
    - Parameters: T,Ws, RH
  - **B: point measurements**
    - Data source: station data
    - Parametes: Ws, Tmin, Tmax, number of days of threshold exceedance of T and RR
  - **C: gridded measurements**
    - Data source: gridded data
    - Parameters: RR, Tmin, Tmax
  - **D: satellite data products**
    - Data source: CM-SAF and CCI
    - Parameters: global radiation, total cloud cover, swe
  - **E: ensemble based comparison**
    - Data source: WP1 ensemble of gridded data
    - Parameters: RR, Tmin, Tmax
  - **F: user related models**

Task 3.1 Coordinated uncertainty evaluation

B: point measurements

- At which time scale can we find correlations between RAs and station observations (regarded as the truth)?
- Representativity issues
- Station data and RA fields are not independent
- Transformation from model-state space to observation space is a key issue

Task 3.1 Coordinated uncertainty evaluation

C: gridded measurements

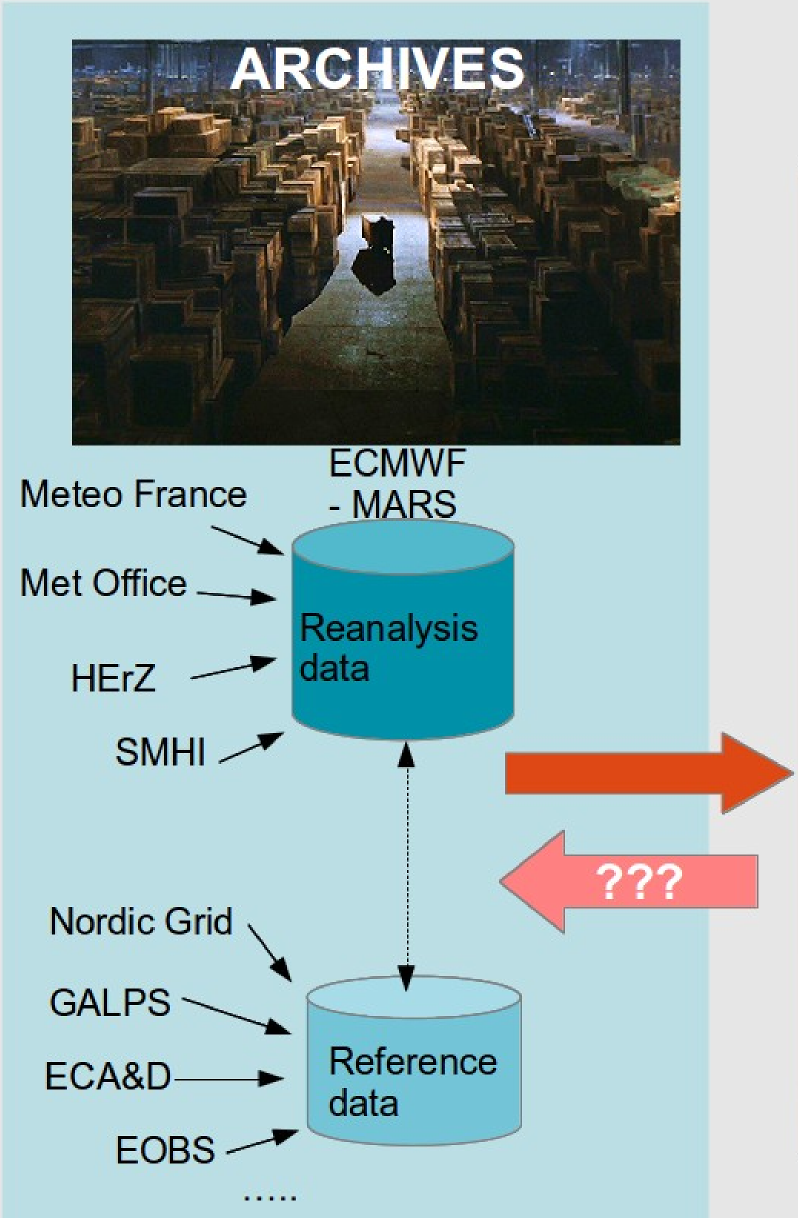
- Which scales (spatial and temporal) can be interpreted?
- Gridded datasets are the reference data for precipitation evaluation
- High-resolution regional grid will be particularly useful, although the question of correlation with RA fields is still present
- RA field and gridded datasets must be upscaled to a common grid

Task 3.1 Coordinated uncertainty evaluation

E: ensemble based comparison

- Size of ensemble: up to 20
- Ensemble mean and spread
- *How does one validate prediction of variance?*
- *Impact of observation error? (from WP1)*
- *Is it possible to validate probabilistic prediction of rare events (such as high-impact weather)?*

General Outlook



Tasks:

- Retrieve/Read data
  - Are all the data in MARS?
  - Common data format?
- Statistics of reanalysis
- Statistics of reference data
- Aggregation in time/Re-gridding
  - See UERRA Deliverable D3.2
- Verification
- Visualization

Verification

- **Traditional, dichotomous (2-category) evaluation** (e.g. Precipitation amount above or below a threshold)
  - Threat Score (TS) or Critical Success Index (CSI)
    - TS=CSI= Hits / (Hits+Misses+False Allarm)
    - *How well did the forecast Yes events correspond to the observed Yes events?*
  - Equitable Threat Score (ETS) or Gilbert Skill Score (GSS)
    - ETS=GSS= Hits - Hits_{random}/(Hits+Misses+False Allarm- Hits_{random})
    - Hits_{random}= (Hits+False Allarm) (Hit+Misses)/Total
    - *How well did the forecast Yes events correspond to the observed Yes events (accounting for hits that would be expected by chance)?*
  - Frequency Bias
    - Bias= (Hits+Misses+False Allarm) / (Hits+Misses+False Allarm)
    - *How similar were the frequencies of Yes forecasts and Yes observations?*

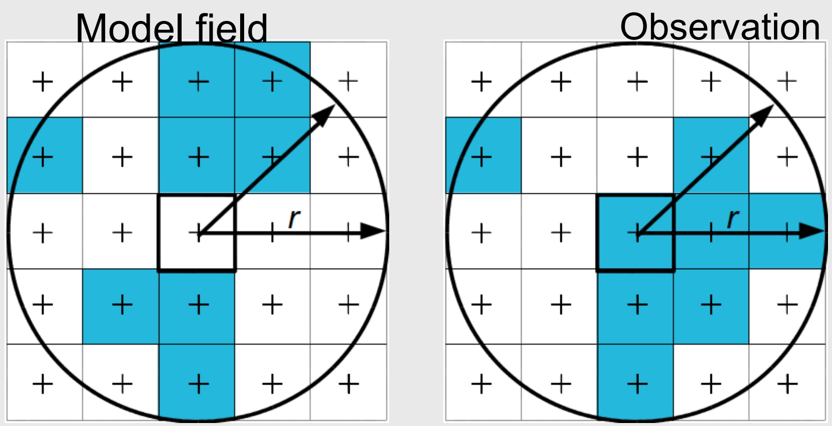
Verification

- **3-category evaluation (Precipitation)**
  - Stable Equitable Error in Probability Space (SEEPS) (Rodwell et al., 2010, QJRMS; Haiden et al., 2012, MonWeatherRev)
    - 24h accumulated precipitation. At each observation location, the weather is partitioned into three categories: “dry”, “light precipitation” and “heavy precipitation”
    - The boundary between “light” and “heavy” is determined by the station climatology
    - The SEEPS score evaluates the performance of the forecast across all three categories.

- **Spatial Verification**, SAL (Structure - Amplitude – Location) quality measure (Wernli et al., 2008, Mon. Wea. Rev.)

Verification

- **Continuos Probabilistic Evaluation**
  - Fractions Skill Score (FSS) (Schwartz et al., 2010, Wea.Forecasting, after work by Robert and Lean, 2008, Mon.Wea.Rev)
    - FSS= 1 – FSB / FSB_{worst} (FSB Fractions Brier Score)



$$FBS = \frac{1}{N} \sum_{i=1}^{N} \left[NP_{M(i)} - NP_{O(i)} \right]^2$$

- *Statistical Interpolation methods can be used to smooth the observed data according to a predefined spatial scale. Then, by comparing the original model fields with the smoothed observed gridded dataset it should be possible to identify the effective model resolution (at least in data dense areas). <u>Can the FSS be a suitable tool for such a study?</u>*

Verification

- **Continuos Probabilistic Evaluation**
  - Continuos Rank Probability Score (CRPS) (see for example Hersbach 2000 and references therein)

    $$CRPS = \frac{1}{Ncases} \sum_{i=1}^{Ncases} \int_{-\infty}^{\infty} [P_i^M(x) - P_i^O(x)]^2 dx$$

    $P^M_i(x)$ is the forecast probability CDF for the ith forecast case, and $P^O_i(x)$ is the obsevration for the ith case expressed as a CDF.
  - Mean Square Skill Score (MSSS) (Murphy 1988 Mon.Wea.Rev.; Lin et al 2008 Mon.Wea.Rev.)

    $$MSSS(T) = 1 - \frac{MSE_M(T)}{MSE_{Clim}}$$

    MSE=Mean Square Error
  - Relative Operative Characteristic (ROC) curves (Mason 1982, Stanski et al. 1989)
  - Relative Economic Value (Richardson 1998, 2000) economic cost-loss analysis

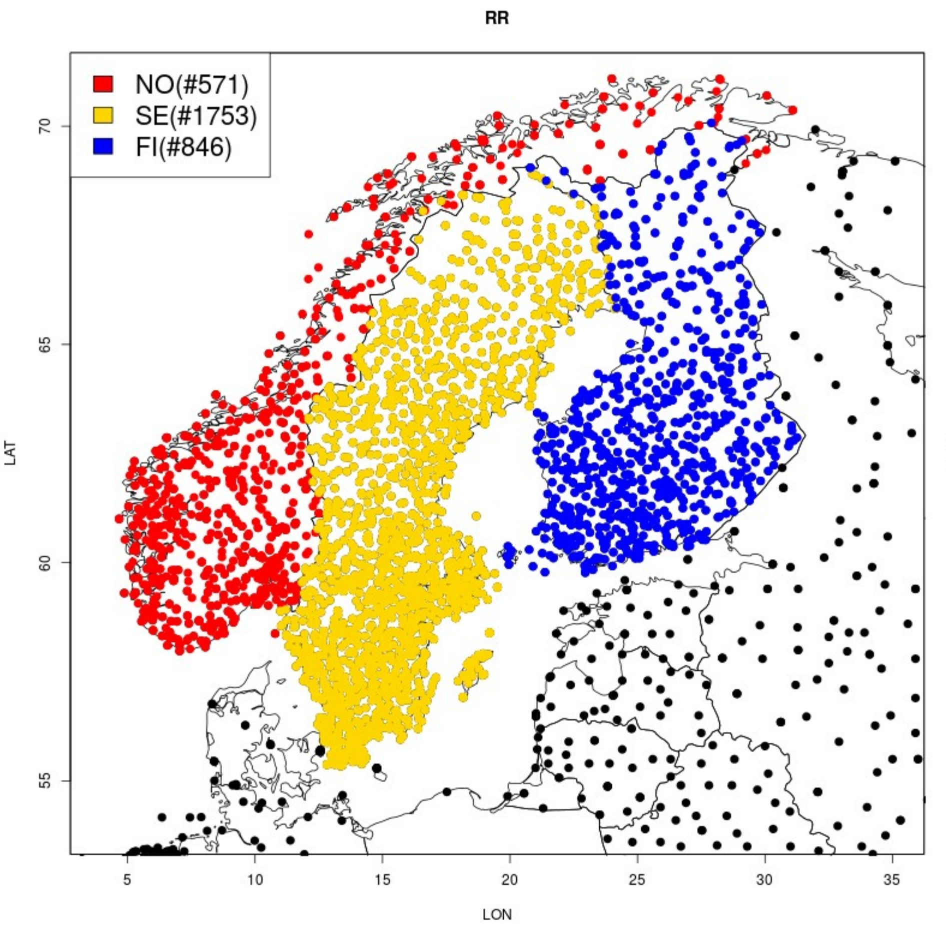
Related Projects / Existing Software

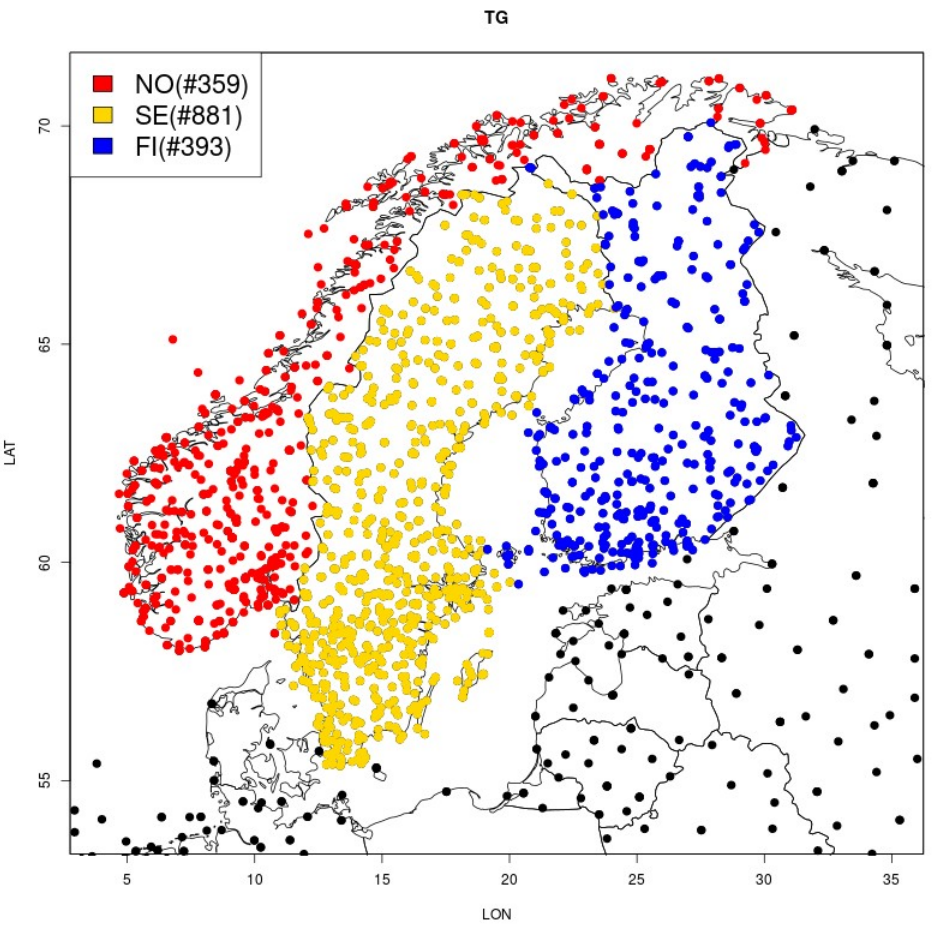
- Empirical Statistical Downscaling ESD (Benestad R. and Mezghani A.)
    - PCA based on EOF decomposition applied on the different datasets -> the first PCs should show similar patterns
- Hirlam-Aladin R package for verification HARP (Singleton A.) developed by the Hirlam / Aladin communities, most recent presentation: *http://www.cnrm.meteo.fr/aladin/IMG/pdf/harp_bucharest_2014.pdf*
    - Includes Verification / SpatialVx packages (related to ICP-MEsoVICT project *http://www.ral.ucar.edu/projects/icp/*)
- UERRA partners [?] and EURO4m results/algorithms
- NOTE on building software togheter:
    - GitHub: code review and management for open source and private projects. *Need private repositories?*

NGCD - Nordic Gridded Climate Dataset

- An observation gridded dataset for temperature and precipitation covering Finland, Sweden and Norway.
  - Spatial resolution 1Km x 1Km
  - CRS: EPSG Projection 3035 - ETRS89 / ETRS-LAEA
  - Temporal resolution: daily
  - Time range: 1971(?) - 2010
  - Data sources: ECA&D, eklima.met.no, SMHI, FMI
- Nordic observation gridded dataset will be an outcome of the Nordic Framework for Climate Services (SMHI, FMI, MI, (DMI,IMO))
- NGCD - November 2014, first versions (2 from MET Norway, 1 from FMI and 1 from SMHI?)

Observation distribution

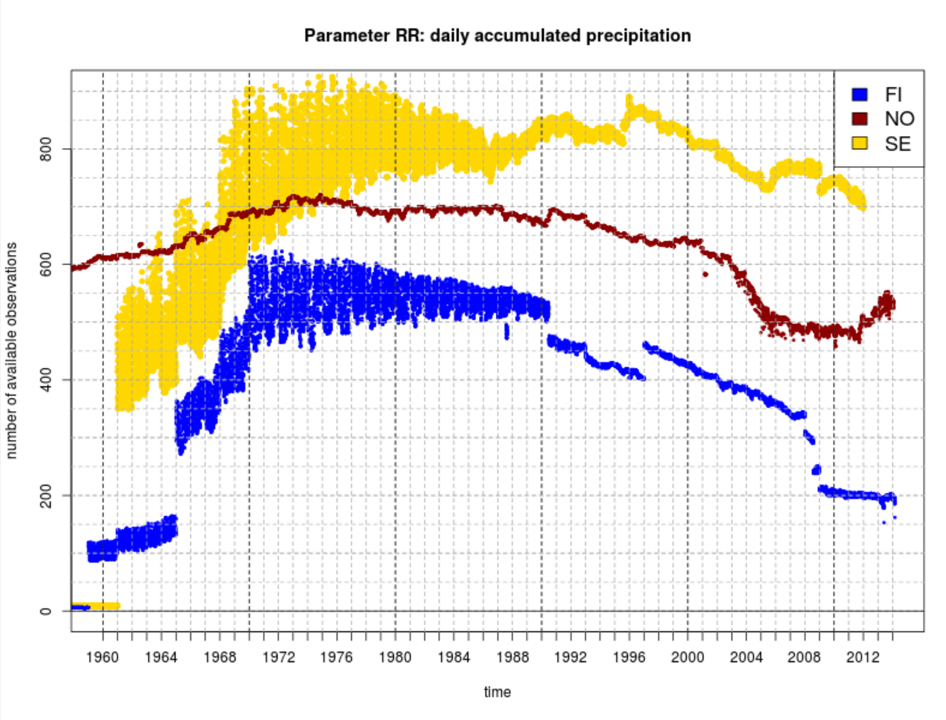




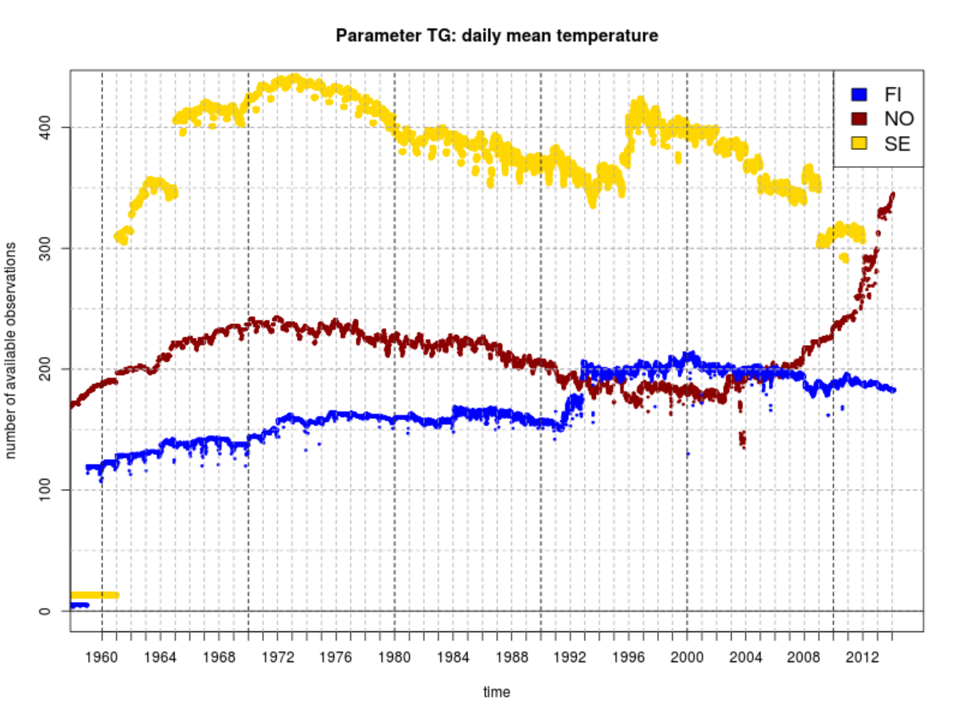
RR

TG

Observation availability in time

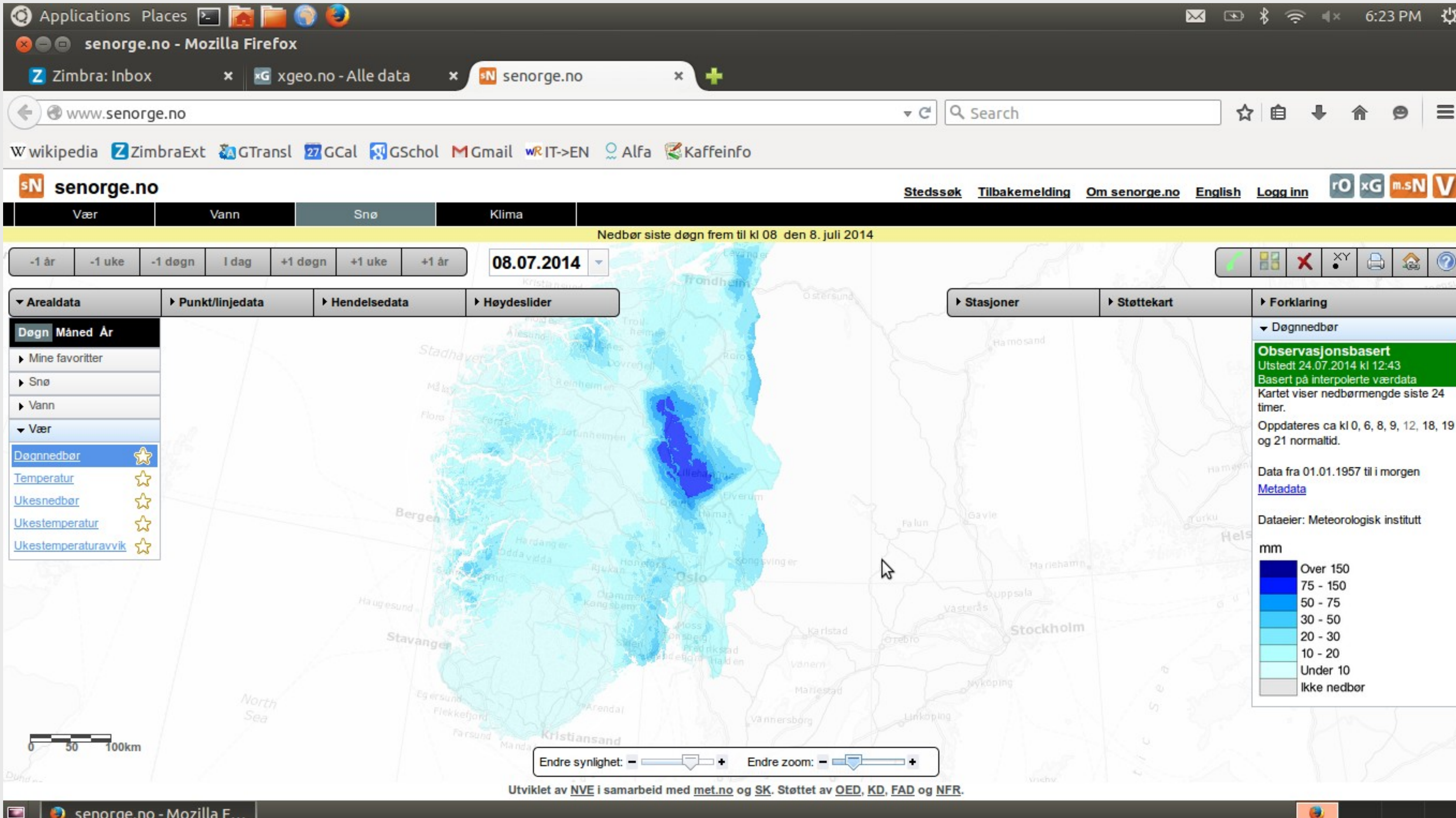


RR

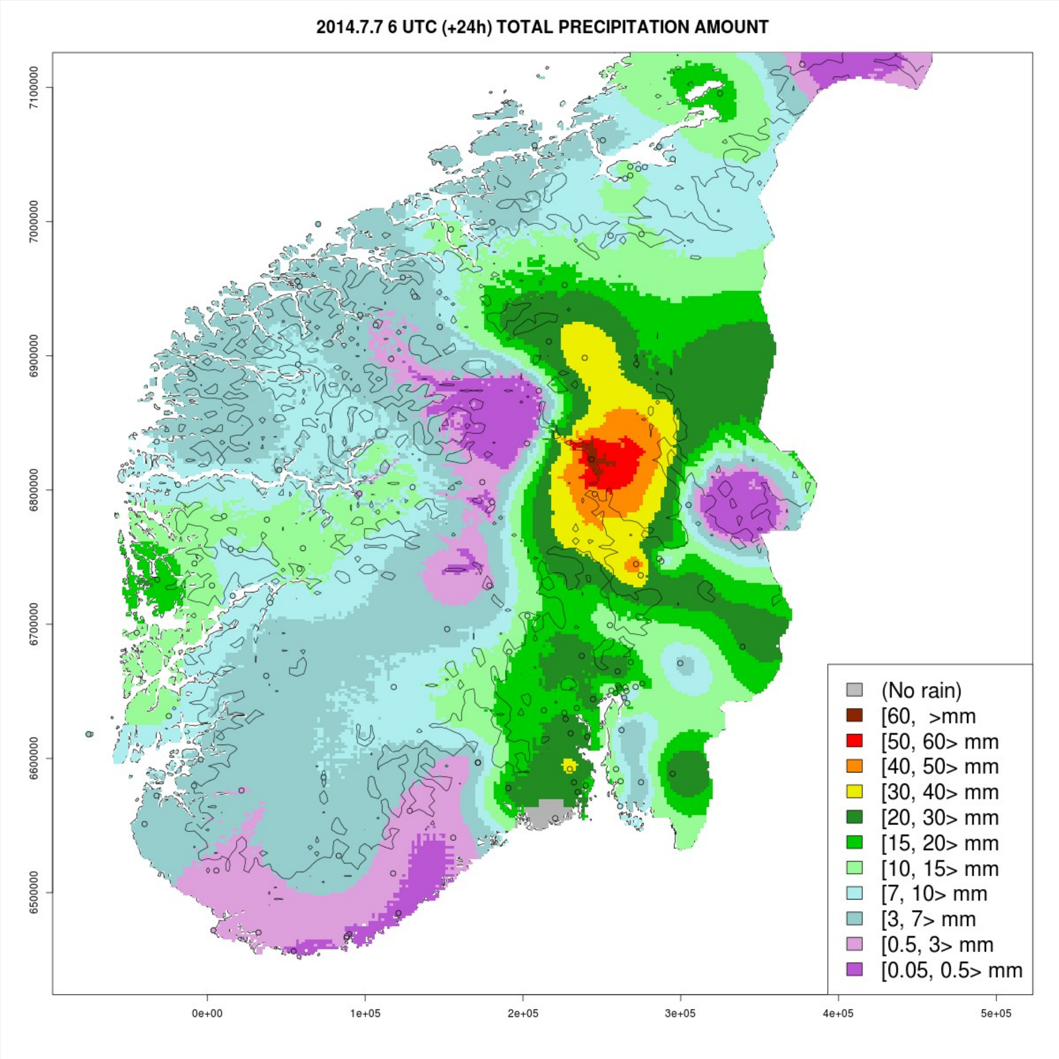


TG

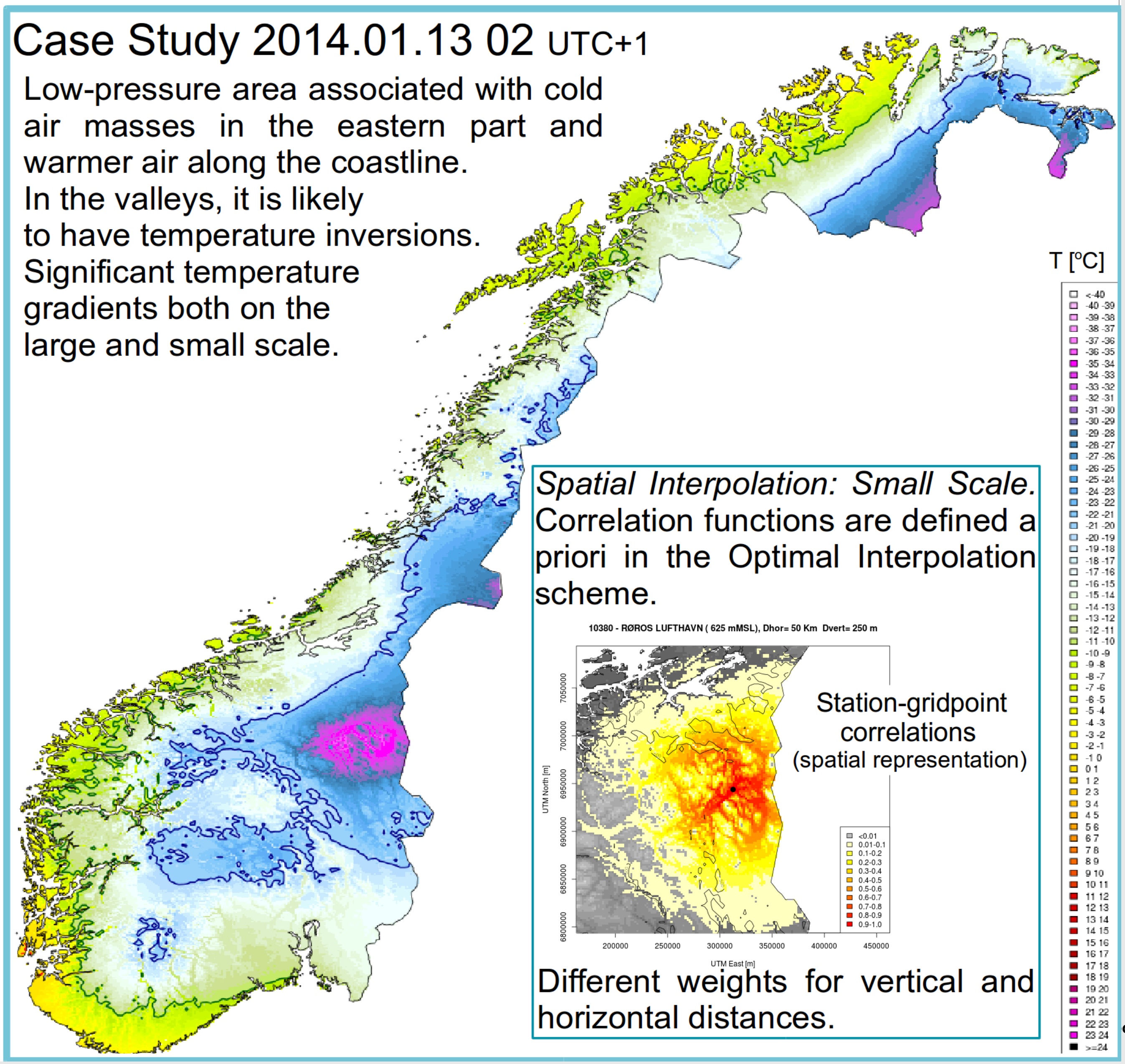
SeNorge 1.1
daily precipitation 2014.07.08



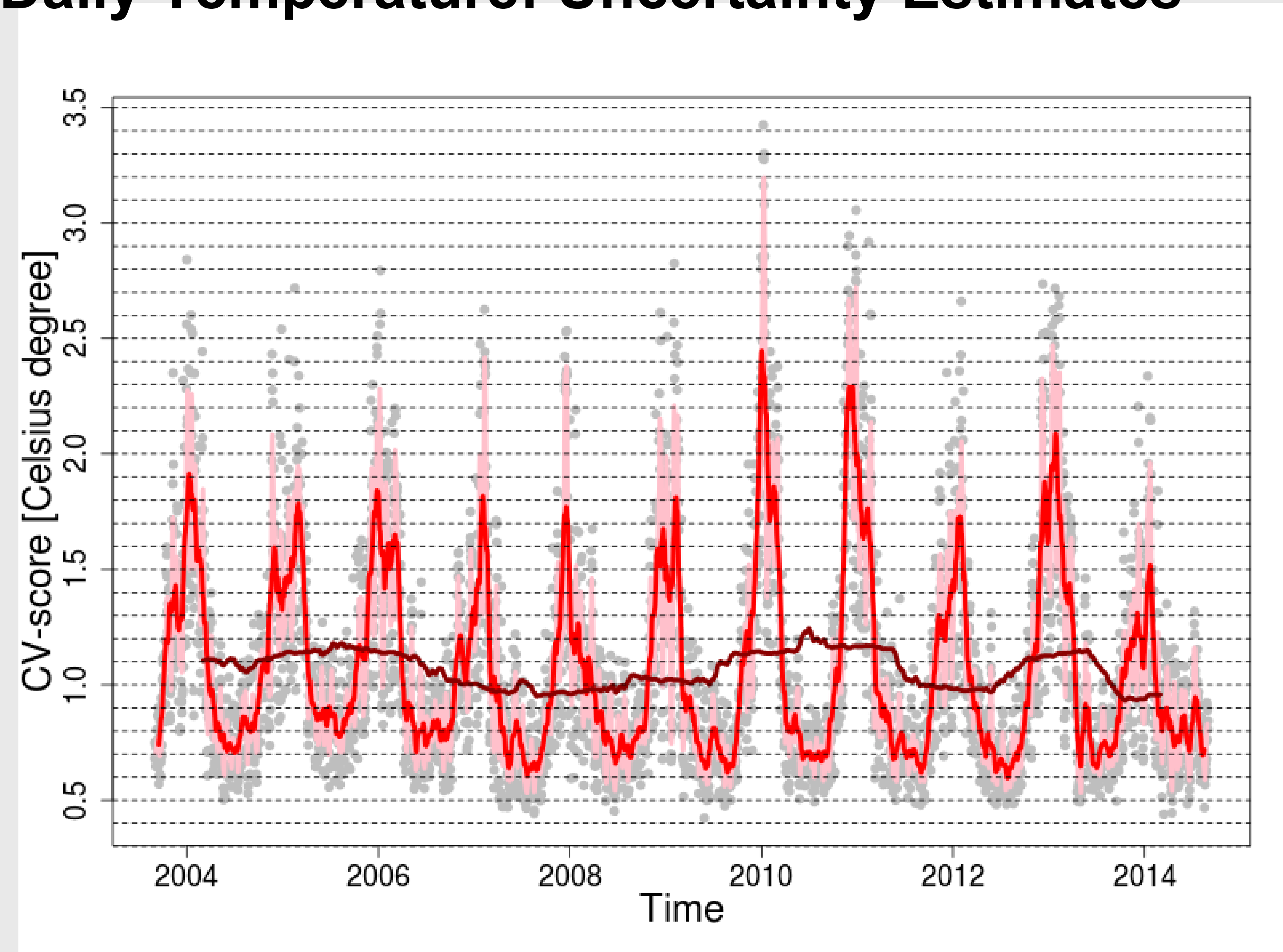
SeNorge 2.0
daily precipitation 2014.07.08



Temperature



Daily Temperature: Uncertainty Estimates




Norwegian
Meteorological
Institute

Norwegian
Meteorological
Institute