METreport

KliNoGrid RR-Rad

combination of radar-derived precipitation fields and raingauge observations

Cristian Lussana, Christoffer A. Elo and Snorre S. Rønning

METreport

Title		**Date**
KliNoGrid RR1-Rad and RR-Rad: combination of radar-derived precipitation fields and raingauge observations		December 29, 2016
Section		**Report no.**
Division for Climate Services		19/2016
Author(s)		**Classification**
Cristian Lussana, Christoffer A. Elo and Snorre S. Rønning		● Free ○ Restricted
Client(s)		**Client's reference**
The Norwegian Water Resources and Energy Directorate		Kolbjørn Engeland

Abstract

The report describes the work done on the combination of radar-derived precipitation fields and raingauge observations in the context of *Felles aktiviteter NVE-MET 2016 tilknyttet nasjonal flom- og skredvarslingstjeneste*. The objective is to establish the first experimental release of the Norwegian Climate Gridded dataset (KliNoGrid) RR1-Rad and RR-Rad products: an observation-based high resolution dataset of hourly (RR1) and daily (RR) precipitation for the Norwegian mainland covering the recent past (currently, from July 2014 to February 2016), which will be used in snow and hydrological modelling. The statistical interpolation scheme implemented is based on Optimal Interpolation and several useful diagnostic variables are included in the elaboration. The evaluation of the final products presented shows the added value of the combination compared to the original radar-derived estimates.

Disciplinary signature	2	Responsible signature
Hans Olav Hygen		Cecilie Stenersen

Contents

1 Introduction

This report describes the combination of radar-derived precipitation estimates and rain-gauge measurements, which is an activity included in the framework of MET-NVE common research activities for 2016 aimed at further promoting the use of radar-derived precipitation estimates in the NVE operational daily routine, not only in tasks where a qualitative information is needed (such as the monitoring of an ongoing extreme precipitation event) but also as a quantitative information to be used in snow- and hydrological modeling. The precipitation field estimates derived from the network of weather radars managed by MET do provide valuable information for the forecasting of floods, landslides and avalanches at NVE. The challenge is to achieve a realistic description for the uncertainties of radar-derived precipitation fields allowing for the correction of systematic errors and the combination of radar-derived estimates with other sources of information.

The radar data used is the hourly precipitation estimate obtained from the composite of Norwegian weather radars. The in-situ observations used are stored in the Climate Database at the Norwegian Meteorological Institute. Because of the dense network of daily precipitation gauge available compared to the less dense -but quickly growing- network of hourly gauge, our choice is to disaggregate the daily measurements to hourly time steps so we can benefit from the best possible in-situ data coverage over the spatial domain. An important point of this work is the development of automatic data quality control routines to identify observations affected by gross measurement errors before they enter the spatial interpolation procedure.

An Optimal Interpolation (OI: *Gandin and Hardin*, 1965) method is used for the combination of radar-derived estimates and raingauge measurements. The OI implementation presented in this report is similar to *Soci et al.* (2016); *Mahfouf et al.* (2007). The combined analysis fields constitute the first experimental release of the Norwegian Climate Gridded dataset (KliNoGrid) RR1-Rad and RR-Rad products: an observation-based high resolution dataset of hourly (RR1) and daily (RR) precipitation for the Norwegian mainland covering the recent past. Currently, the time interval spanned ranges from July 2014 to February 2016.

In detail:

- Precipitation day/hour. An observation is considered to measure precipitation if its value is greater than 0.1 mm both for hourly and daily precipitation.

- The three dimensional spatial coordinates associated with a generic point in the space $\mathbf{r}$ are specified as triplets (x, y, z).

- the parameter $\text{rad.c}(\mathbf{r})$ named *radar attenuation coefficient* at point $\mathbf{r}$ has been introduced to take into account the effect on uneven radar data coverage over the Norwegian mainland on the spatial interpolation. To have an idea of $\text{rad.c}(\mathbf{r})$, see Figs. 1 and 2. Given the radar-derived accumulated precipitation over a long time period (in some cases we used 3 months, while in other 1 year has been used), we will indicate as $\mathbf{x}^{\text{RRlt}-\text{Rad}}$ the field over our grid points and as $x^{\text{RRlt}-\text{Rad}}(\mathbf{r})$ the precipitation at point $\mathbf{r}$ as extracted from $\mathbf{x}^{\text{RRlt}-\text{Rad}}$ (nearest neighbor). The first step to obtain $\text{rad.c}(\mathbf{r})$ is to smooth $\mathbf{x}^{\text{RRlt}-\text{Rad}}$ (we used *raster* library in R software: we calculate focal -moving window- values for the neighborhood of focal cells using a 5x5 matrix of weights all set to the value of 1); then $\text{rad.c}(\mathbf{r})$ is set as a number between 0 and 1, which represents the largest $\mathbf{x}^{\text{RRlt}-\text{Rad}}$ quantiles not greater than $x^{\text{RRlt}-\text{Rad}}(\mathbf{r})$.

- Observations

  - $\mathbf{y}^{\text{o}}$, column vector of dimension M (i.e. M-vector): observed daily precipitation (i.e. from 06 UTC of yesterday to 06 UTC of today):

  - $t = 1, \ldots 24$ and it is assumed that these are the 24 hours constituting the 24-hour interval spanned by $\mathbf{y}^{\text{o}}$.

  - $\widetilde{\mathbf{y}}^{\text{o}}(t)$, $\tilde{M}(t)$-vector: hourly precipitation actually observed for hour t.

  - $\mathbf{y}^{\text{o}}(t)$, $M(t)$-vector: best estimate of the hourly precipitation observed for hour t. It is the union of $\widetilde{\mathbf{y}}^{\text{o}}(t)$ and the disaggregation of $\mathbf{y}^{\text{o}}$ over hourly time steps.

  - In our case, $M(t) \geq M > \tilde{M}(t)$

- Background

  - $\mathbf{x}^{\text{b}}(t)$, I-vector: radar-derived hourly precipitation at grid points.

  - $\mathbf{y}^{\text{b}}(t)$, $M(t)$-vector: radar-derived hourly precipitation at station locations. The radar product is estimating hourly precipitation at grid points so we need to apply a function to transform estimates at grid points in estimates at station locations, such a function is called *observation operator*. In our case, the

5

observation operator is the nearest neighbor interpolation routine, which in linear algebra can be specified as a matrix $\mathbf{H}$:

$$\mathbf{y}^{\mathrm{b}}(t) = \mathbf{H}\mathbf{x}^{\mathrm{b}}(t)$$

where $\mathbf{H}$ is equal to zero almost everywhere except for one single element for each line, which is set to 1 and it selects the grid point closer to the correspondent location of interest.

2 KliNoGrid *RR1-Rad* and *RR-Rad* pseudocode

The pseudocode describing the spatial interpolation process is:

1. read: configuration file, grid specification, digital elevation model, grid mask (grid definition, see Appendix A)

2. read long-term radar-derived accumulated precipitation field (i.e. background field) and transform it to the radar attenuation coefficient field (see the Introduction and Figs. 1-2)

3. read daily and hourly observed values from the MET Norway's Climate Database

4. read hourly radar-derived precipitation fields (see Sec. 3)

5. observation operator is applied to obtain both the hourly radar-derived precipitation values and the radar attenuation coefficient values at the output locations (see the Introduction and Sec. 4)

6. automatic data quality control, plausibility tests (see Sec. 5.1)

7. automatic data quality control, compare observations with background (see Sec. 5.2)

8. simultaneous application of Spatial Consistency Test (SCT) and disaggregation of daily to hourly precipitation (see Appendix B)

9. Optimal Interpolation (OI) of hourly precipitation *RR1-Rad* (see Sec. 4 and Figs. 3,4,5)

10. OI of daily precipitation *RR-Rad* (see Sec. 4 and Figs. 6,7,8)

11. write output:

(a) analysis, background and Integral Data Influence (IDI) on a regular grid;

(b) analysis, leave-one-out cross-validated analysis, background and Integral Data Influence (IDI) at selected locations.

3 Radar-derived precipitation fields

In this work, the source of our information for radar data is the composite of the Norwegian weather radars managed and post-processed by MET. The composite of radar reflectivity fields is available at a sampling rate of 7.5 minutes, which means that each hourly accumulated precipitation field is obtained by using 8 of these composite fields. The transformation from reflectivity to rainfall rate is based on the classic *Marshall and Palmer* (1948) relation, which is:

$$Z = 200 R^{1.6} \tag{1}$$

where Z ($mm^6 m^{-3}$) is the reflectivity factor and R ($mm\,h^{-1}$) is the rainfall rate (*American Meteorological Society*, 2016). The hourly accumulated precipitation field is then computed from the 8 corresponding 7.5 minutes rainfall rate fields by assuming a constant rain-rate in between two radar measurements.

4 Optimal Interpolation

For hourly precipitation, it is assumed that both the observations and the background field are estimates of an unknown truth, which is denoted by the superscript t. The truth either at grid points or station locations is obtained from the true hourly precipitation continuous field as describe in *Daley* (1991), for example. We assume an additive error model:

$$\mathbf{y}^{\text{o}}(t) = \mathbf{H}\mathbf{x}^{\text{t}}(t) + \varepsilon^{\text{o}}(t) \tag{2}$$

$$\mathbf{x}^{\text{b}}(t) = \mathbf{x}^{\text{t}}(t) + \eta^{\text{b}}(t) \tag{3}$$

where: $\mathbf{H}$ is the linear observation operator, which in our case is a matrix of dimensions $M(t)$ x I (see the Introduction); $\varepsilon^{\text{o}}(t)$ is the observation error, $M(t)$-vector; $\eta^{\text{b}}(t)$ is the background error at the grid points, I-vector. The errors are assumed to be Gaussian, with mean value equal to zero everywhere, moreover observation and background errors are assumed to be independent, then:

$$\varepsilon^{\mathrm{o}}(t) \quad \sim \quad N(\mathbf{0}, \mathbf{R}) \tag{4}$$

$$\eta^{\mathrm{b}}(t) \quad \sim \quad N(\mathbf{0}, \mathbf{B}) \tag{5}$$

$$\left\langle \varepsilon^{\mathrm{o}}(t) \eta^{\mathrm{b}}(t) \right\rangle \quad = \quad 0 \tag{6}$$

where $\mathbf{R}$ ($M(t)$ x $M(t)$ matrix) and $\mathbf{B}$ (I x I matrix) are the observation and background, respectively, error covariance matrices. Please note that since we're using hourly observation obtained through temporal disaggregation of daily observations (a process involving the use of radar data, see Sec B), then Eq. (6) should be regarded as an approximation.

Given our assumption, the best linear unbiased estimate of the truth $\mathbf{x}^{\mathrm{t}}(t)$ on the grid is the analysis $\mathbf{x}^{\mathrm{a}}(t)$ (*Daley* (1991); *Jazwinski* (2007); *Gandin and Hardin* (1965)):

$$\mathbf{x}^{\mathrm{a}}(t) = \mathbf{x}^{\mathrm{b}}(t) + \mathbf{K}\left[\mathbf{y}^{\mathrm{o}}(t) - \mathbf{H}\mathbf{x}^{\mathrm{b}}(t)\right] \tag{7}$$

where the Kalman gain matrix is:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{8}$$

See Figs. 3 and 4 for an example of analysis and background fields.

Analogously, the best linear unbiased estimate of the truth $\mathbf{H}\mathbf{x}^{\mathrm{t}}(t)$ at the station locations (or at any selected number of points on the domain) is the analysis $\mathbf{y}^{\mathrm{a}}(t)$:

$$\mathbf{y}^{\mathrm{a}}(t) = \mathbf{H}\mathbf{x}^{\mathrm{b}}(t) + \mathbf{W}\left[\mathbf{y}^{\mathrm{o}}(t) - \mathbf{H}\mathbf{x}^{\mathrm{b}}(t)\right] \tag{9}$$

where:

$$\mathbf{W} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{10}$$

For the daily precipitation, we proceed as for the hourly precipitation except that the background field is set to:

$$\mathbf{x}^{\mathrm{b}} \quad = \quad \sum_{t=1}^{24} \mathbf{x}^{\mathrm{a}}(t) \tag{11}$$

Then the analyses can be written as:

8

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}\left[\mathbf{y}^{\mathrm{o}} - \mathbf{H}\mathbf{x}^{\mathrm{b}}\right] \tag{12}$$

$$\mathbf{y}^{\mathrm{a}} = \mathbf{H}\mathbf{x}^{\mathrm{b}} + \mathbf{W}\left[\mathbf{y}^{\mathrm{o}} - \mathbf{H}\mathbf{x}^{\mathrm{b}}\right] \tag{13}$$

See Figs. 6 and 7 for an example of analysis and background fields.

We need to specify the error covariance matrices. For the hourly precipitation, we assume:

$$
\begin{align}
\mathbf{R} &\equiv \sigma_{\mathrm{o}h}^2(t)\mathbf{I} \tag{14}\\
\mathbf{B} &\equiv \sigma_{\mathrm{b}h}^2(t)\widetilde{\mathbf{B}} \tag{15}\\
\widetilde{\mathbf{B}}_{ij} &= \left[1 + \frac{\Delta\mathrm{dist}_H\left(\mathbf{r}_i,\mathbf{r}_j\right)}{D_H}\right] \cdot \exp\left[-\frac{\Delta\mathrm{dist}_H\left(\mathbf{r}_i,\mathbf{r}_j\right)}{D_H}\right] \cdot \left\{1 - \left[\Delta\mathrm{rad.c}\left(\mathbf{r}_i,\mathbf{r}_j\right)\right]^2\right\} \tag{16}\\
\varepsilon^2 &\equiv \frac{\sigma_{\mathrm{o}h}^2(t)}{\sigma_{\mathrm{b}h}^2(t)} = 0.1 \tag{17}
\end{align}
$$

where: $\Delta\mathrm{dist}_H\left(\mathbf{r}_i,\mathbf{r}_j\right)$ indicates the horizontal distance between the i-th and j-th points; $\Delta\mathrm{rad.c}\left(\mathbf{r}_i,\mathbf{r}_j\right)$ indicates the radar attenuation coefficient difference between the i-th and j-th points. When setting the spatial correlation function based on distance only (within the same radar attenuation class), it is implicitly assumed that far from any radar location the background field is equal to zero and the station network is dense enough to yield useful information for hydrological application. It's worth remarking that the combined field of precipitation is not a simple product to handle by the final user because it includes different spatial scales for different domain areas.

Analogously, for the daily precipitation, we assume:

$$
\begin{align}
\mathbf{R} &\equiv \sigma_{\mathrm{o}d}^2\mathbf{I}\\
\mathbf{B} &\equiv \sigma_{\mathrm{b}d}^2\widetilde{\mathbf{B}}\\
\varepsilon^2 &\equiv \frac{\sigma_{\mathrm{o}d}^2}{\sigma_{\mathrm{b}d}^2} = 0.1
\end{align}
$$

In practice, we're using the same station distribution and error covariance matrices for daily and hourly precipitation. We use the same value for the decorrelation distance D_H both for daily and hourly OI, which is set to $20\,\mathrm{Km}$. However, it's important to remark that the error covariance matrices differ in the specification of the error variances, nevertheless

their ratio ε^2 is set in both cases to 0.1 (i.e. we trust far more the observations than the background).

Given our assumption on the covariance matrices, Eqs. (8) and (10) can be rewritten as (*Uboldi et al.* (2008)):

$$\mathbf{K} = \widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} \left(\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} + \varepsilon^2\mathbf{I} \right)^{-1} \tag{18}$$

$$\mathbf{W} = \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} \left(\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} + \varepsilon^2\mathbf{I} \right)^{-1} \tag{19}$$

As a result, the gain matrices depends only on our choices for D_H and ε^2.

Consider the m-th observation (it doesn't matter here if hourly or daily), the correspondent leave-one-out Cross-Validated (CV) analysis $\breve{y}_m^{\mathrm{a}}$ is defined as the analysis estimate obtained for the m-th observation by using all observations except the m-th observation itself. The CV analysis vector $\breve{\mathbf{y}}^{\mathrm{a}}$ is the M-vector having the CV analysis as components.

Given a subset of N data to evaluate, the CV-score is defined as the root mean squared difference:

$$CV_{\mathrm{score}} = \sqrt{\frac{1}{N} \sum_{m=1}^{N} (\breve{y}_m^{\mathrm{a}} - y_m^{\mathrm{o}})^2} \tag{20}$$

The CV score represents an estimate of the analysis error based on the idea that each observation is used as an independent verification of the analysis field. The estimate is conservative because of the implicit degradation of the local resolution of the observational network. The CV analysis is also useful for quality control tests on the observations.

The Integral Data Influence (IDI) of all the observations on the analysis at the m-th observation location is introduced as:

$$y_m^{\mathrm{IDI}} = \sum_{k=1}^{M} \frac{\partial y_m^{\mathrm{a}}}{\partial y_k^{\mathrm{o}}} = \sum_{k=1}^{M} \mathbf{W}_{mk} \tag{21}$$

The IDI vector $\mathbf{y}^{\mathrm{IDI}}$ is the M-vector having $y_m^{\mathrm{IDI}}, m = 1, \ldots, M$ as components.

Similarly, $\mathbf{x}^{\mathrm{IDI}}$ is the I-vector having $x_i^{\mathrm{IDI}}, i = 1, \ldots, I$ as components, where the IDI of all the observations on the analysis at the i-th grid point is:

$$x_i^{\mathrm{IDI}} = \sum_{k=1}^{M} \frac{\partial x_i^{\mathrm{a}}}{\partial y_k^{\mathrm{o}}} = \sum_{k=1}^{M} \mathbf{K}_{ik} \tag{22}$$

From a practical point of view, the IDI field could be conveniently seen as the analysis field obtained when all observed values are set to 1 and all background values are set to 0. See Figs. 5 and 8 for examples.

The fusion of leave-one-out cross-validation and IDI concepts yields to the introduction of $\check{y}_m^{\text{IDI}}$, which is the the Cross-Validated IDI (CV-IDI) of all the observations except the m-th on the analysis at the m-th observation location:

$$\check{y}_m^{\text{IDI}} = \sum_{k=1}^{M} \frac{\partial \check{y}_m^{\text{a}}}{\partial y_k^{\text{o}}} = 1 + \frac{1}{1 - \mathbf{W}_{mm}} \left(y_m^{\text{IDI}} - 1 \right) \tag{23}$$

The CV-IDI vector $\check{\mathbf{y}}^{\text{IDI}}$ is the M-vector having $\check{y}_m^{\text{IDI}}$ as components.

As a remark, the vectors $\mathbf{y}^{\text{IDI}}$, $\check{\mathbf{y}}^{\text{IDI}}$ and $\mathbf{x}^{\text{IDI}}$ are independent of the actual observed values and they depend only on the station distribution, given the matrix elements.

5 Automatic Data Quality Control routines

We are using only daily and hourly observations flagged in the KDVH as:

- 0: Original value is checked and found OK. User value: OK

- 1: Value is controlled and corrected, or value is missing and interpolated. User value: OK

- 2: Original value is not checked. User value: LU (slightly suspect)

Beside, as a further quality control condition we use only observations that are not blacklisted by the *seNorge ver 2.0* automatic data quality control routines.

All the observations flagged as *suspect* by the following tests are not used in the OI. The information is part of the program output for the users.

5.1 plausibility tests

For daily observations:

- a range test is performed with lower and upper thresholds set to $0 \, \text{mm} \cdot \text{h}^{-1}$ and $500 \, \text{mm} \cdot \text{h}^{-1}$, respectively.

For hourly observations:

- a range test is performed with lower and upper thresholds set to $0\,\mathrm{mm\cdot h^{-1}}$ and $200\,\mathrm{mm\cdot h^{-1}}$, respectively.

- time consistency test. The total time series of 24 observations referring to a single precipitation day is flagged as suspect if:

  - the percentage of missing observations is greater than 90%

  - there are more than 22 hours recording precipitation

5.2 Compare observations with the background (i.e. radar-derived data)

An hourly observation is flagged as *suspect* if:

- the observation doesn't measure precipitation and the background estimates precipitation

- the observation measure precipitation and the background doesn't estimates precipitation and the observation is located in a position r where we assume the radar coverage is good enough (i.e. $\mathrm{rad.c\,(\mathbf{r})} \geqslant 0.8$)

For the case of daily observations, we apply exactly the same rules and the background values are extracted from the field of $\sum_{t=1}^{24} \mathbf{Hx^b}\,(t)$.

5.3 Spatial Consistency Test based on OI

The SCT applied here is described in *Lussana et al.* (2010) and in Appendix C. Our particular application is detailed in Appendix B.

As stated in Sec. 4, the specification of the error covariance matrices differ between daily and hourly OI because of the different error variances. As a consequence, we should use different thresholds for the hourly and daily SCT, which are denoted as T^h and T^d in Appendix B.

At this point of our work, we're still using the same value for both daily and hourly observations but we plan to investigate more this issue and to set different threshold values.

The current setup is:

$$T_i^h \equiv T_i^d = \begin{cases} 20\,\mathrm{mm^2\cdot h^{-2}} & ,y_i^o < 10\,\mathrm{mm\cdot h^{-1}} \\ y_i^o & ,y_i^o \geqslant 10\,\mathrm{mm\cdot h^{-1}} \end{cases}$$

6 Evaluation

The evaluation presented here is based on the separation of predictions (i.e. analysis, cv-analysis or background) and observations in "yes, an event will happen" or "no, the event will not happen" (dichotomous prediction). A threshold is specified to separate "yes" and "no". The verification is based on categorical statistics computed from elements in contingency tables.

We focus on the following scores (see *WWRP/WGNE Joint Working Group on Forecast Verification Research* http://www.cawcr.gov.au/projects/verification, which is reported in Appendix D) : Probability of detection (POD); False alarm ratio (FAR); Bias score (BIAS) and Equitable threat score (ETS).

The verification scores are shown in: Figs. 9-12 for the whole period; Figs. 10-13 for the summer (JJA, June-July August 2015); Figs. 11-14 for the winter (DJF, December 2015-January 2016-February 2016). The dataset used for verification includes all the data available over the spatial domain. The "yes"/"no" separation threshold is indicated on the x-axis. The red curves (CV-Ya) refer to the verification of the leave-one-out cross validated analysis values $\breve{y}^a$ against the observations y^o and it is representative of the analysis performance at grid points. The brown curves (Ya) refer to the verification of the analysis values y^a against the observations y^o and it is representative of the analysis performance at station locations (i.e. due to the observation representativity error, for example). The yellow curves (Yb) refer to the verification of the background values y^b against the observations y^o and it is representative of the background performances at station locations. It's worth remarking that intense precipitation is a very rare event compared to the most common case of light precipitation, for this reason some of the Figures show noisy results for intense precipitation because of the very few events considered.

RR1-Rad In general, there is a clear evidence of the added value of the combined product compared to the raw radar-derived one: the red and brown curves always shows better results than the yellow one. The performances at station locations (Ya) and at grid points (CV-Ya) are not too different, which might indicate reasonable choices for the OI parameters. In the ETS scores RR1-Rad outperforms the background: the quality of the performance decreases rather sharply with the increase in the precipitation intensity, nevertheless the RR1-Rad performances decreases slowly than the background thus indicating the benefit of using in-situ observations. In the summer

season we get better scores than in winter and the difference is particularly significant for the background. In the case of light precipitation (less than $1 \, \text{mm} \cdot \text{h}^{-1}$), the BIAS is close to 1 indicating that the frequency of "yes" events is better predicted than by using radar-derived data only and the POD value around 0.8 together with FAR of 0.2 indicates skill of RR1-Rad in distinguishing between precipitation/no-precipitation events. In the case of intense precipitation (greater than $10 \, \text{mm} \cdot \text{h}^{-1}$) the BIAS is between 0.2 and 0.4, which means that the predicted frequency of "yes" events is far less than the observed one. This result, which is also valid for the analysis, indicates that we need to reduce the uncertainty of the representativity error component (which is part of the observation error) in our OI scheme: that's probably due to the fact that our assumptions in OI, such as both observation and background are unbiased estimates for the truth, are not valid for intense precipitation. On the other hand, we should investigate more in detail the procedure used for the temporal disaggregation of daily to hourly precipitation, which may introduce significant uncertainty especially for intense precipitation. As a last remark, it's quite interesting the increase in FAR in winter, which is not occurring during summer.

RR-Rad for daily precipitation the evaluation shows much better results than for the hourly case. There are many reason for that: we should expect a larger spatial coherence for the RR1-Rad field than for RR1-Rad, furthermore the station network for daily precipitation is denser and we don't have to take into account the effect of the disaggregation procedure. It's useful to remark that for RR-Rad the background (Yb) we're considering in Figs. 12, 13 and 14 it's not a raw radar-derived product but it's the sum of the 24 RR1-Rad fields (see Eq.(11)), while the analysis (Ya) refers to the subsequent OI of Eqs. (12) and (13). It looks like the analysis is obviously improving the performances at station locations but for grid points (CV-Ya) is not really capable of adding value to the background (i.e. the CV-analysis curves present worse results compared to the background curves), as a consequence for grid points may be better to use the sum of RR1-Rad as daily analysis value. This result might indicate that we should use a different formulation for the error covariance matrices in the RR-Rad OI than the one presented in Eq (16), perhaps without including the radar attenuation coefficient, which is yet considered for RR1-Rad production. The RR-Rad performances are still dependent on precipitation intensity, nonetheless the

14

ETS shows a reasonable skill also for intense precipitation. It seems that the performances in summer and winter are not significantly different, perhaps the RR-Rad products show better performances in winter but that's not really clear.

7 Conclusions

In the process of establishing an high-resolution climate gridded dataset for hourly and daily precipitation on the Norwegian mainland (*KliNoGrid: RR1-Rad* and *RR-Rad* products), the first experiments on the combination of radar-derived fields and in-situ observations are described in this report.

The daily observed precipitation, which is measured by a dense network of stations, is disaggregated to hourly precipitation in a way that is consistent with the subsequent spatial interpolation of hourly values, as a consequence the in-situ observation network used to correct the radar background is denser than the one we would have by considering hourly measurements only.

The spatial interpolation method implemented is based on Optimal Interpolation (OI). For locations which are far from any radar installation, the predictions are equivalent to an OI based on raingauge observations only. The evaluation presented shows the benefit of the combined products compared to the raw radar-derived product. The OI scheme described here should be regarded as proof of concept and the presented results show that the OI has practical potential. In addition, efforts have been devoted to the development of an automatic quality control procedure, which must be further developed and thoroughly tested in the near future.

The OI scheme itself may be improved: we may considered to include a transformation of the precipitation values before they enter the OI (in order to better fulfill the requirement of Gaussian errors); the evaluation shows the worst performances for intense hourly precipitation, which might be improved through the use of a better radar background (soon available) or with more realistic OI assumptions on the radar error.

8 Figures

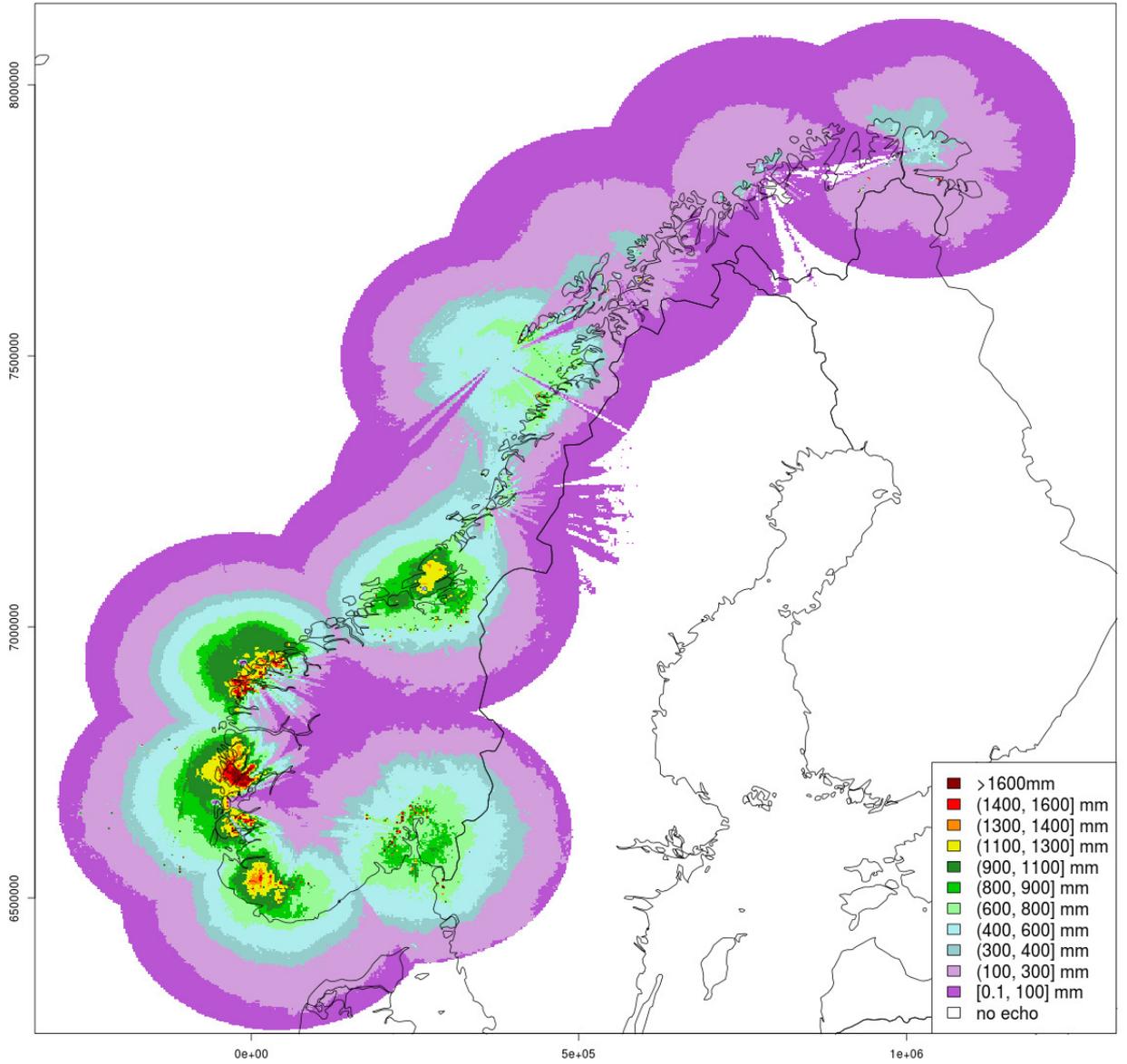**accumulated precipitation from 2014.07 to 2015.08 (input: 1h radar-derived prec fields)**

Figure 1: 1-year (from July 2014 to August 2015) accumulated precipitation field obtained from hourly radar-derived fields.

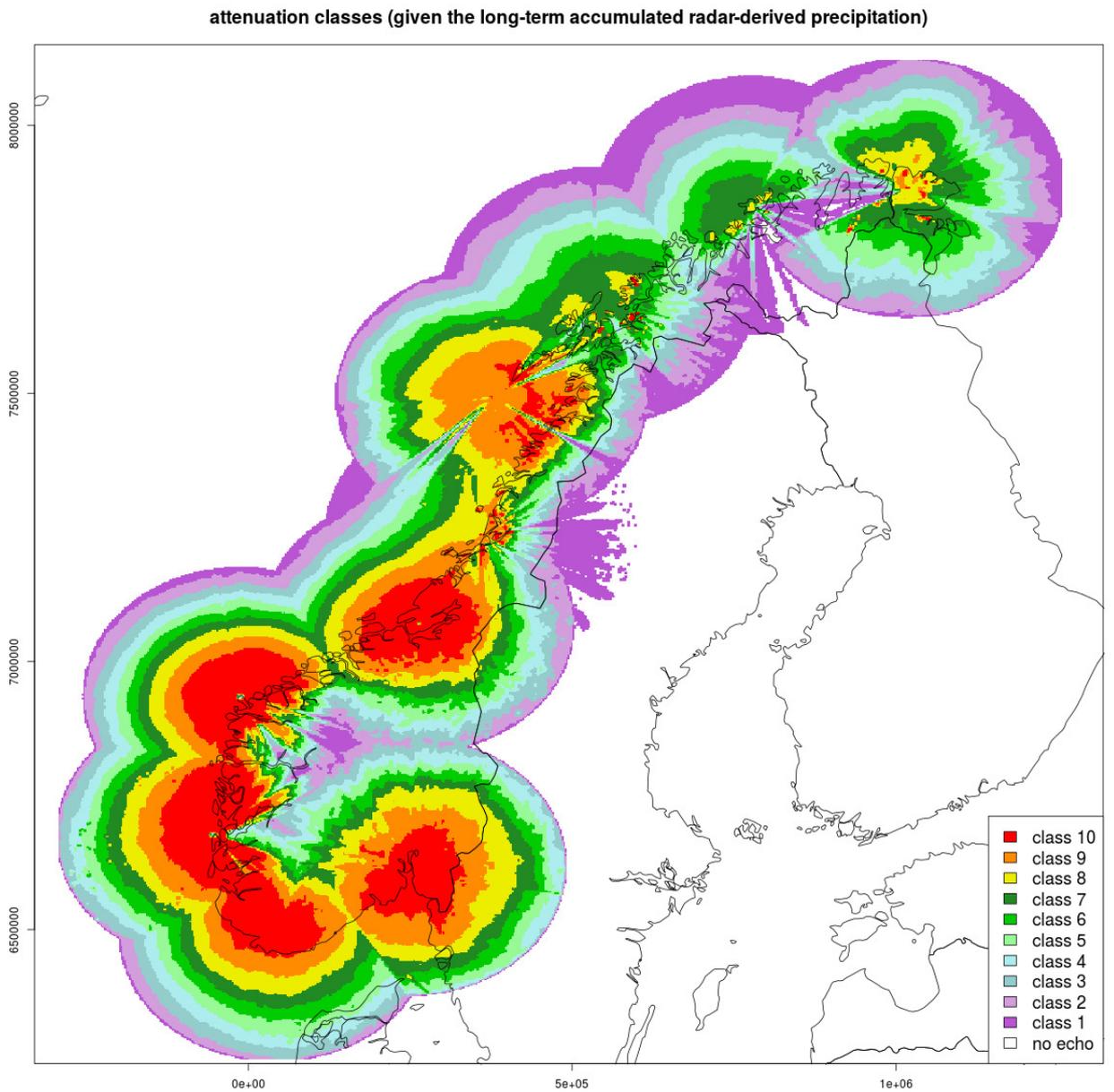**attenuation classes (given the long-term accumulated radar-derived precipitation)**

Figure 2: radar attenuation coefficient field obtained from the accumulated precipitation field shown in Fig. (1). The coefficients are presented here as "classes" (the coefficient is equal to the class divided by 10).
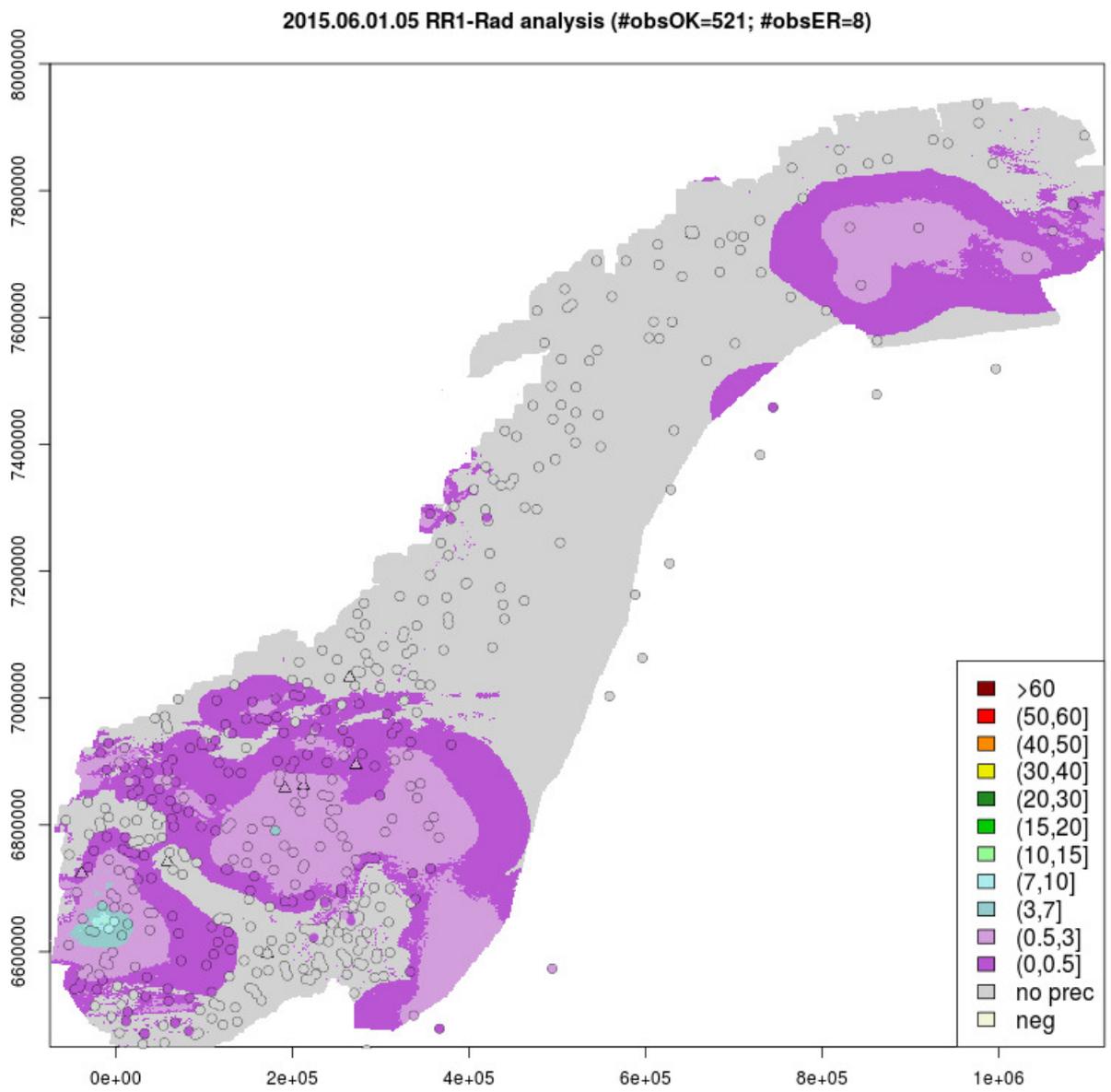
Figure 3: RR1-Rad analysis, 2015.06.01 05:00 UTC.
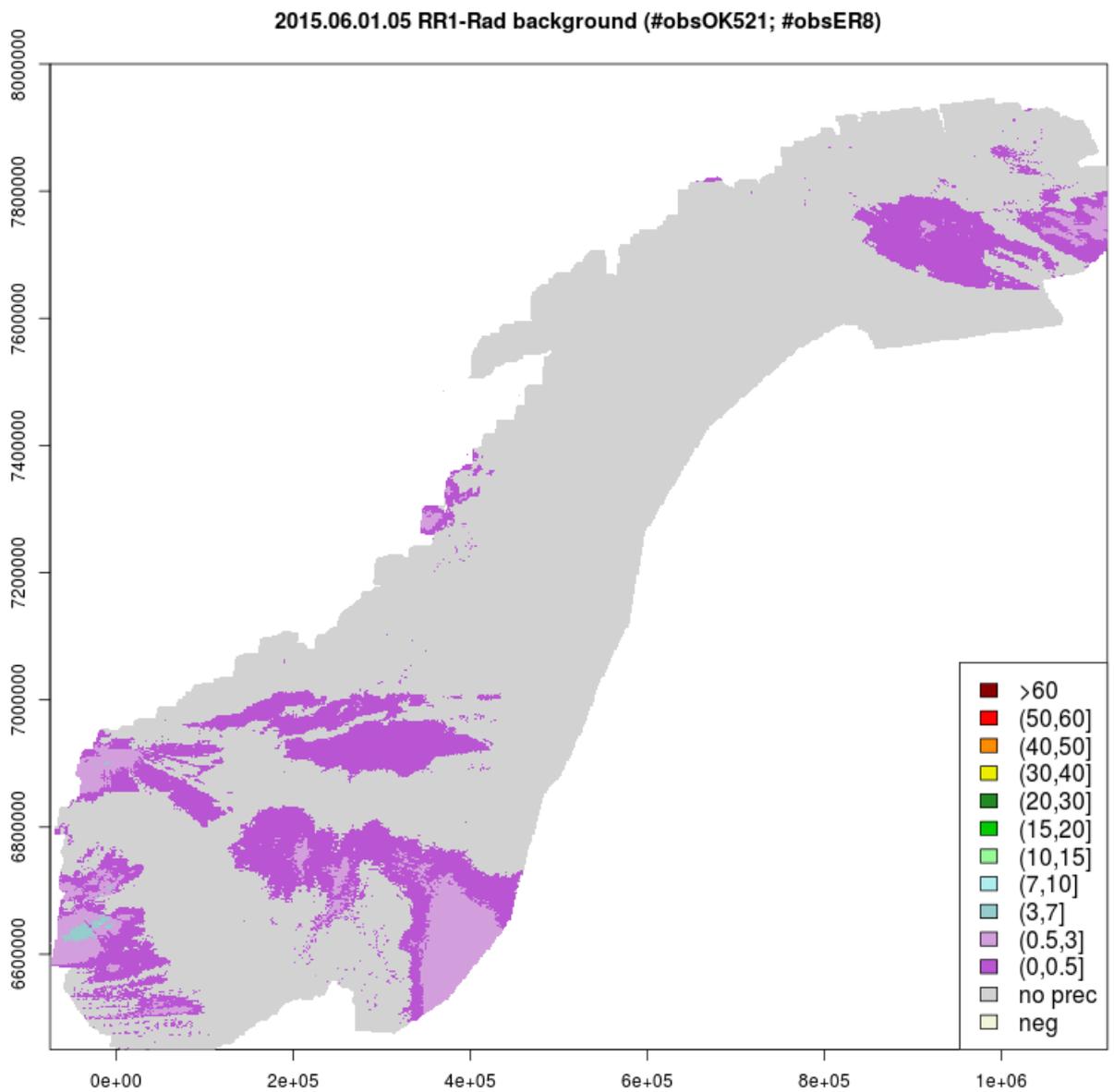
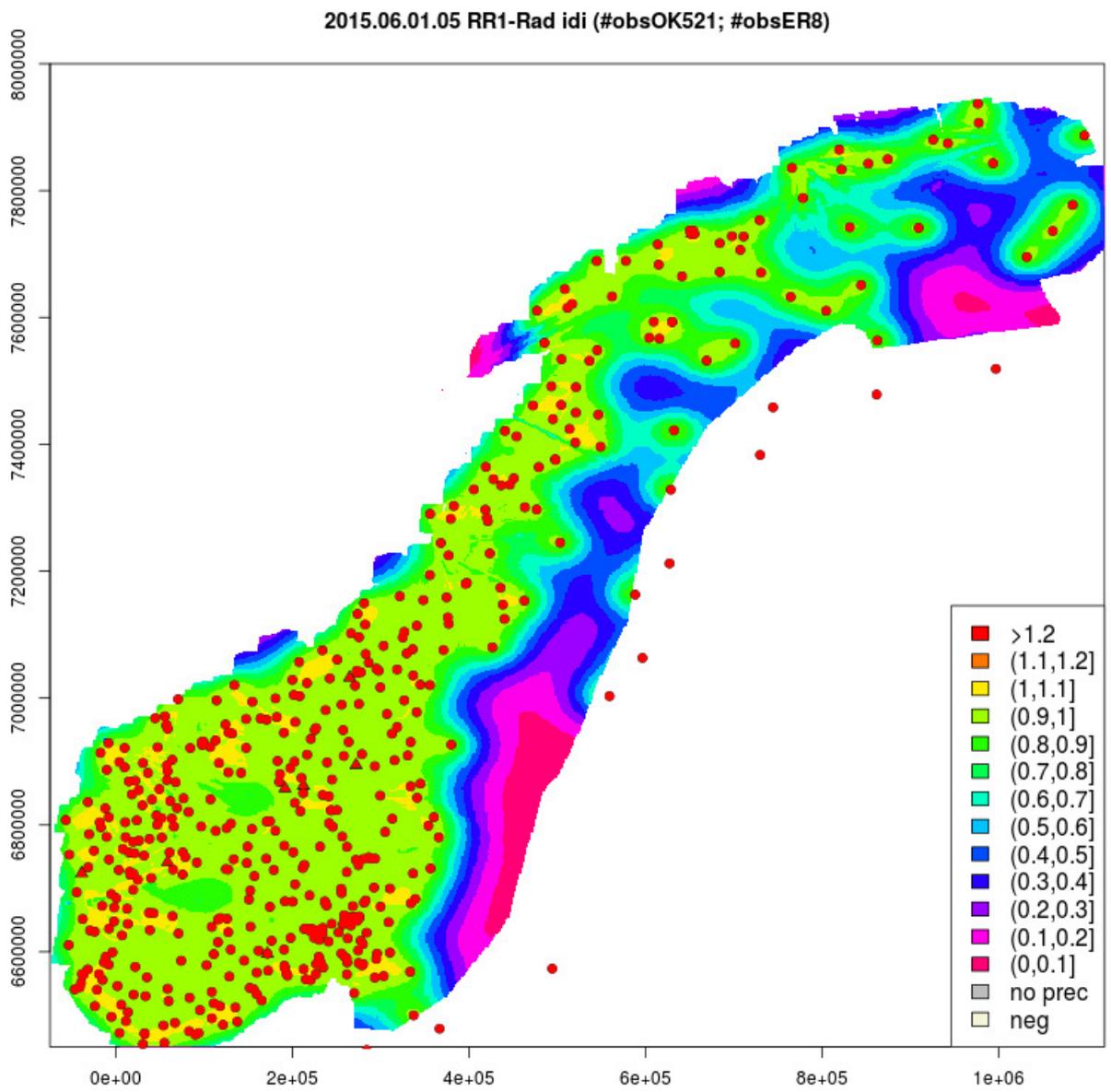Figure 4: RR1-Rad background (i.e. radar-derived precipitation estimate), 2015.06.01 05:00 UTC.

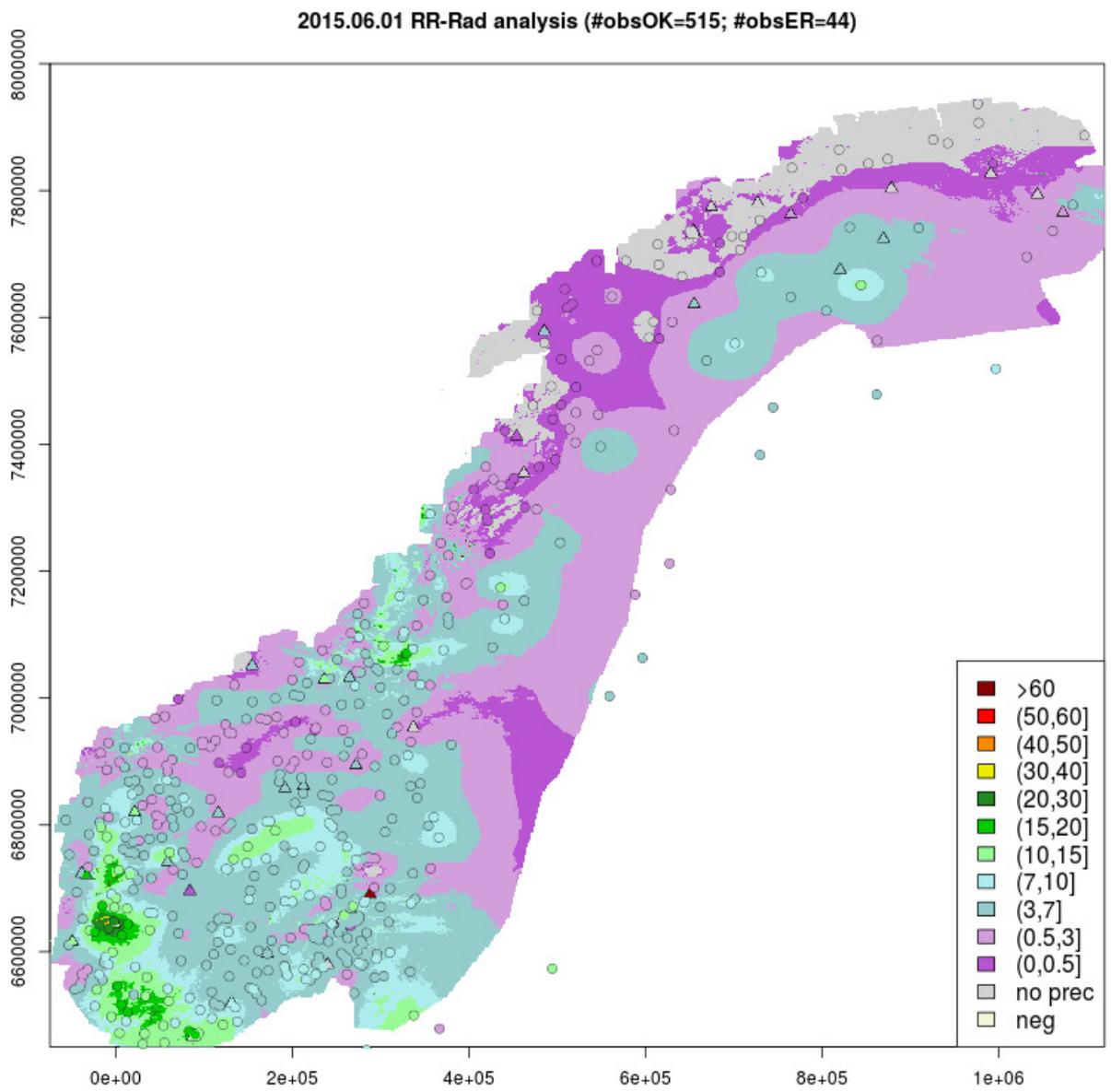Figure 5: RR1-Rad Integral Data Influence, 2015.06.01 05:00 UTC.
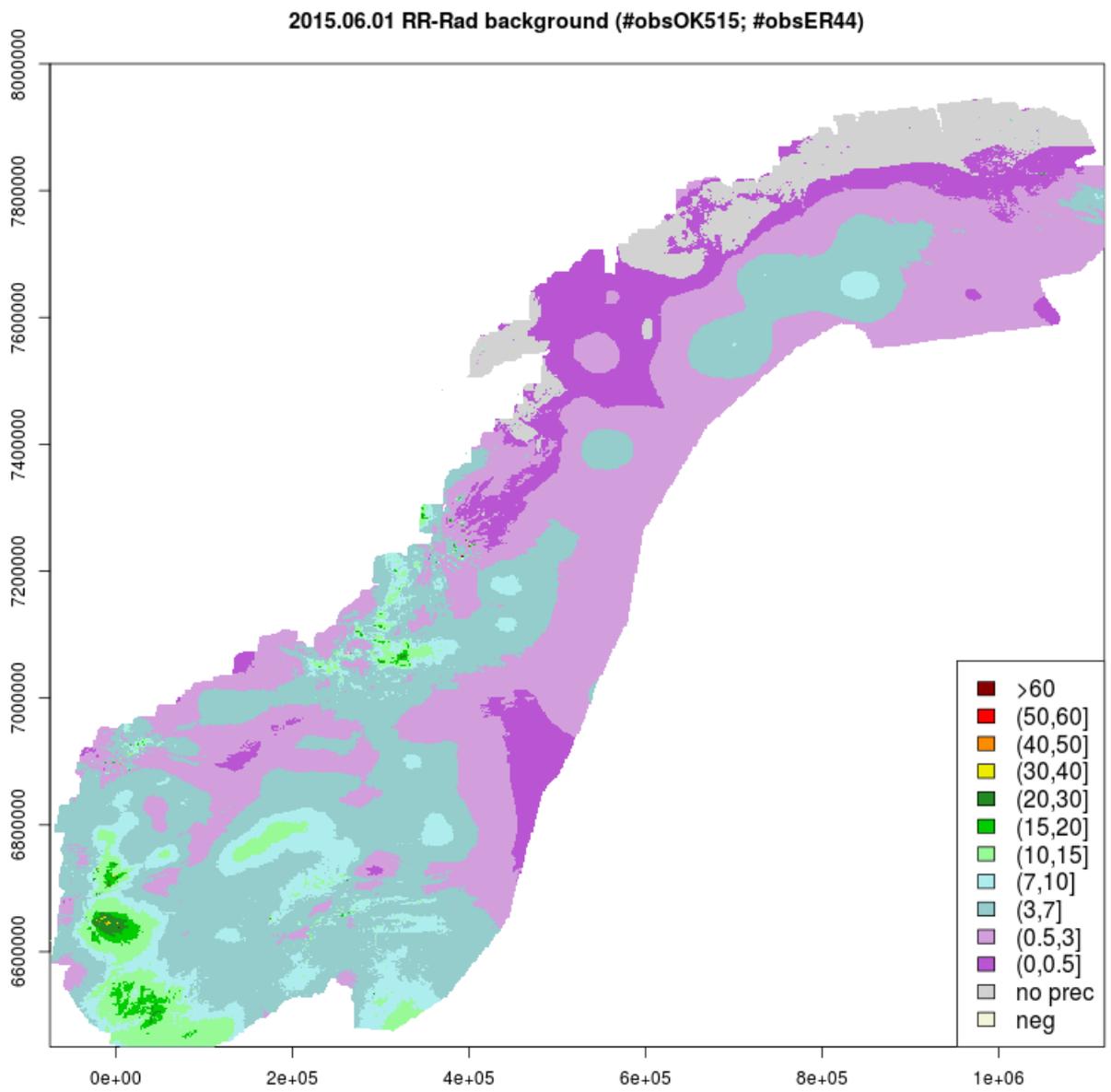
Figure 6: RR-Rad analysis, 2015.06.01

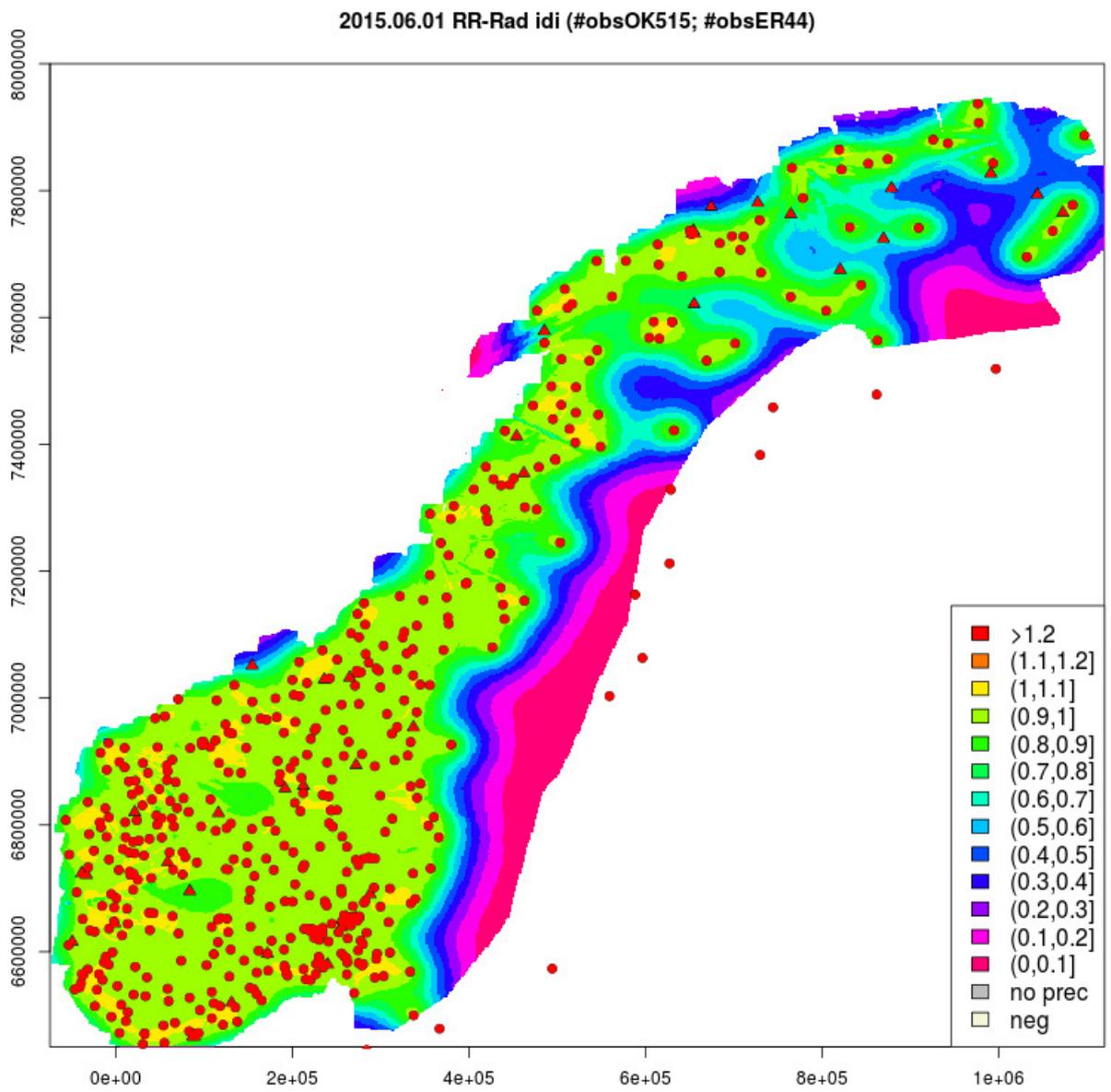Figure 7: RR-Rad background, 2015.06.01.

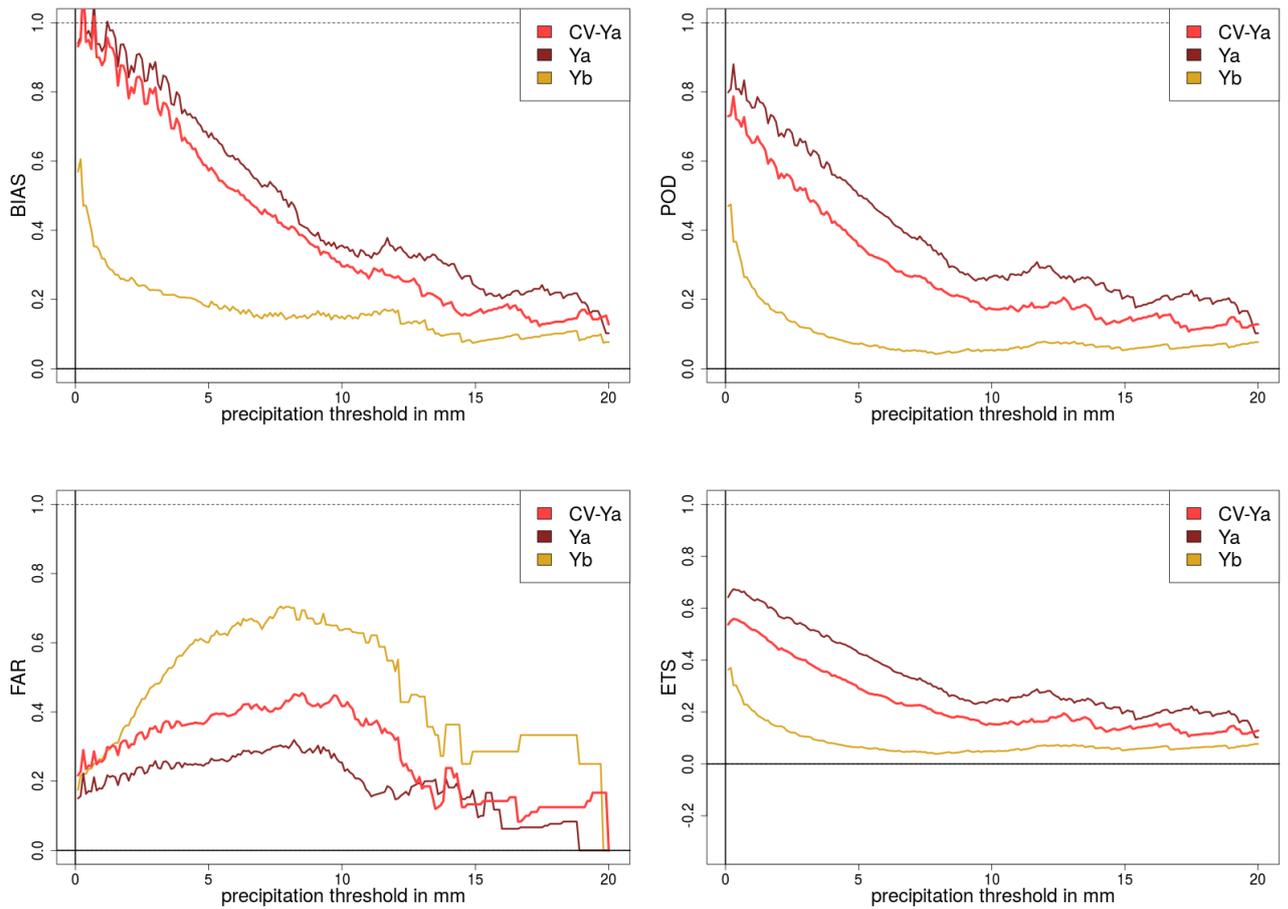Figure 8: RR-Rad Integral Data Influence, 2015.06.01.

Figure 9: RR1-Rad. dataset: from 2014.07.01 to 2016.03.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).
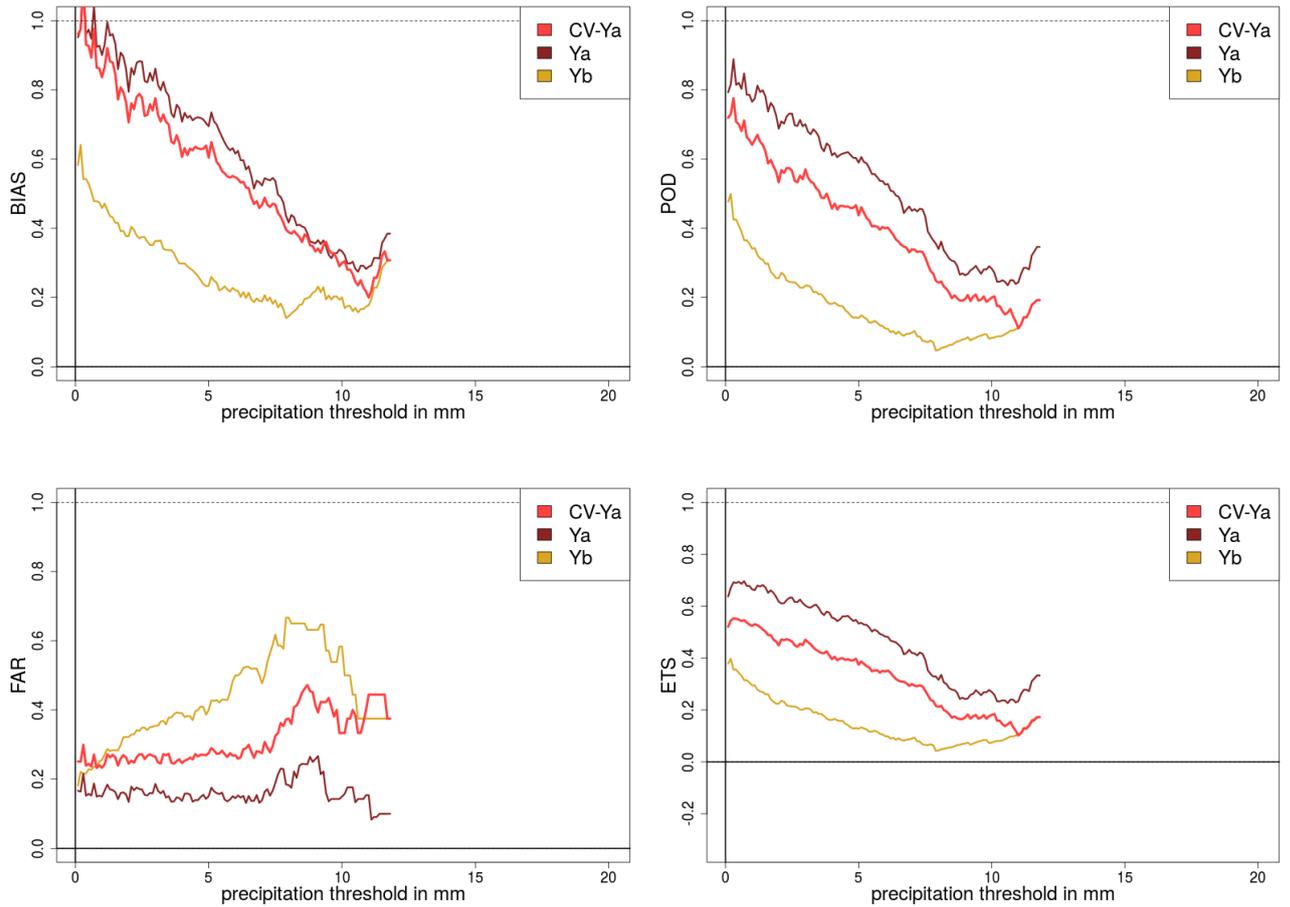
Figure 10: RR1-Rad. *JJA* dataset: from 2015.06.01 to 2015.09.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).

Figure 11: RR1-Rad. *DJF* dataset: from 2015.12.01 to 2016.03.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).

Figure 12: RR-Rad. dataset: from 2014.07.01 to 2016.03.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).

Figure 13: RR-Rad. *JJA* dataset: from 2015.06.01 to 2015.09.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).
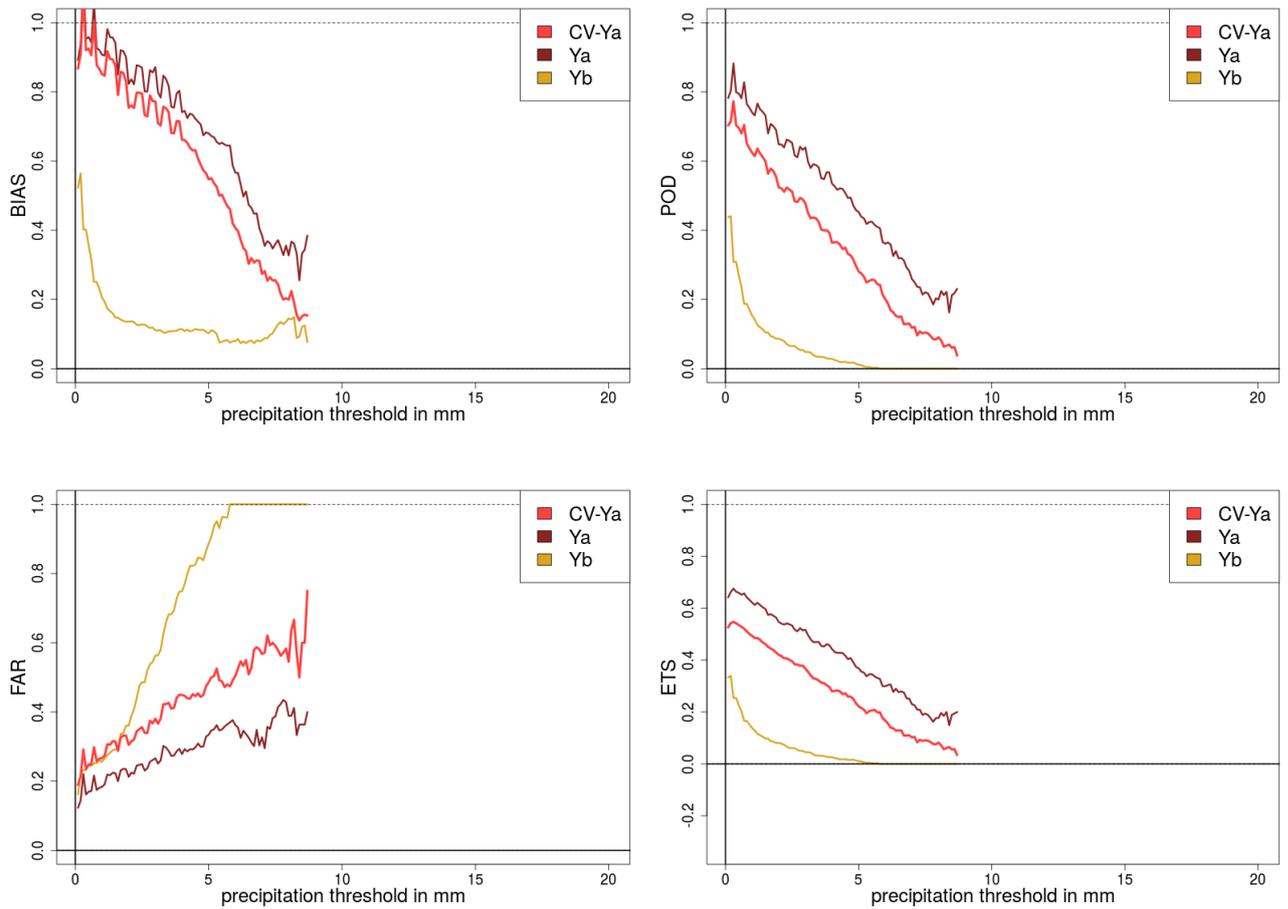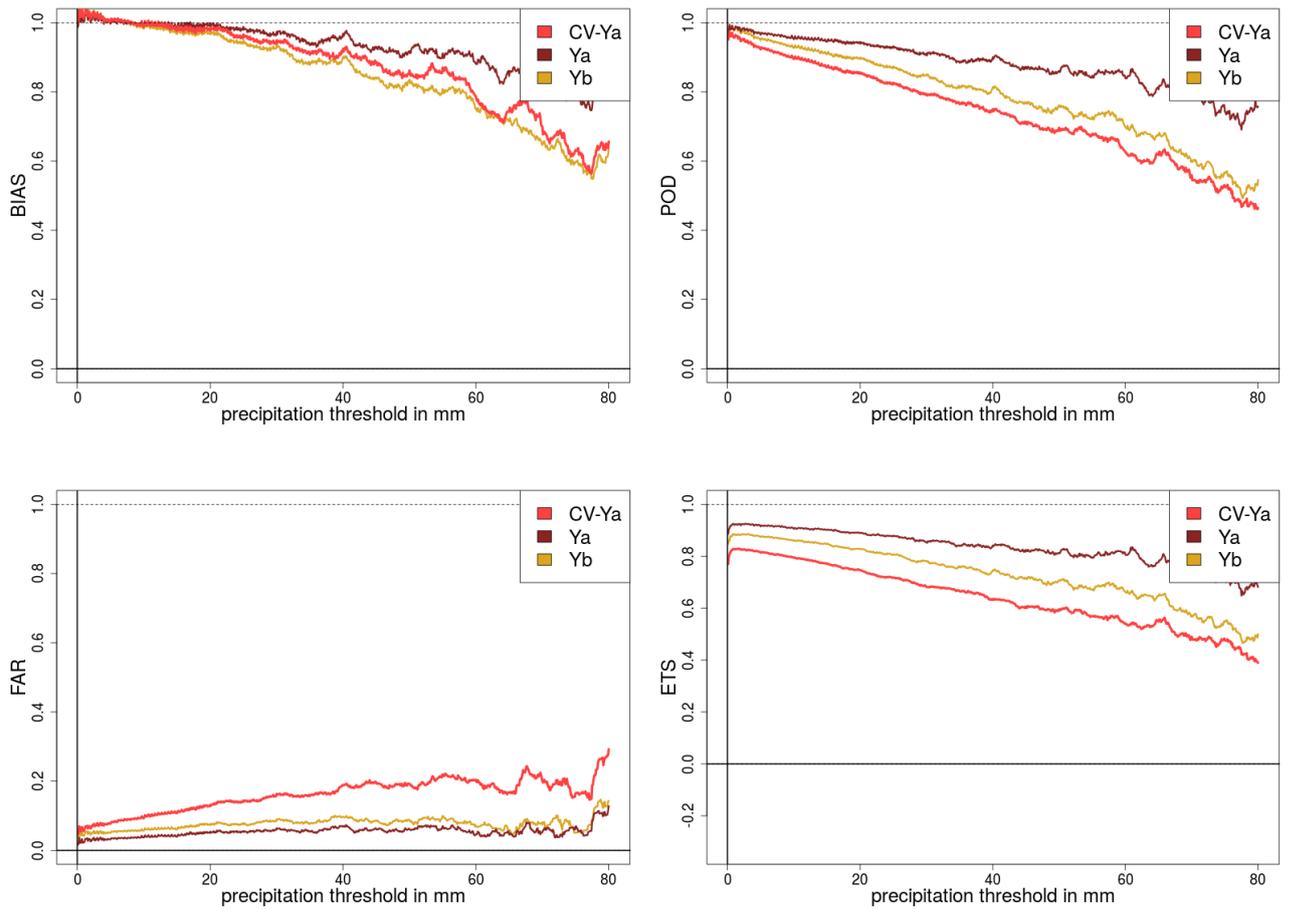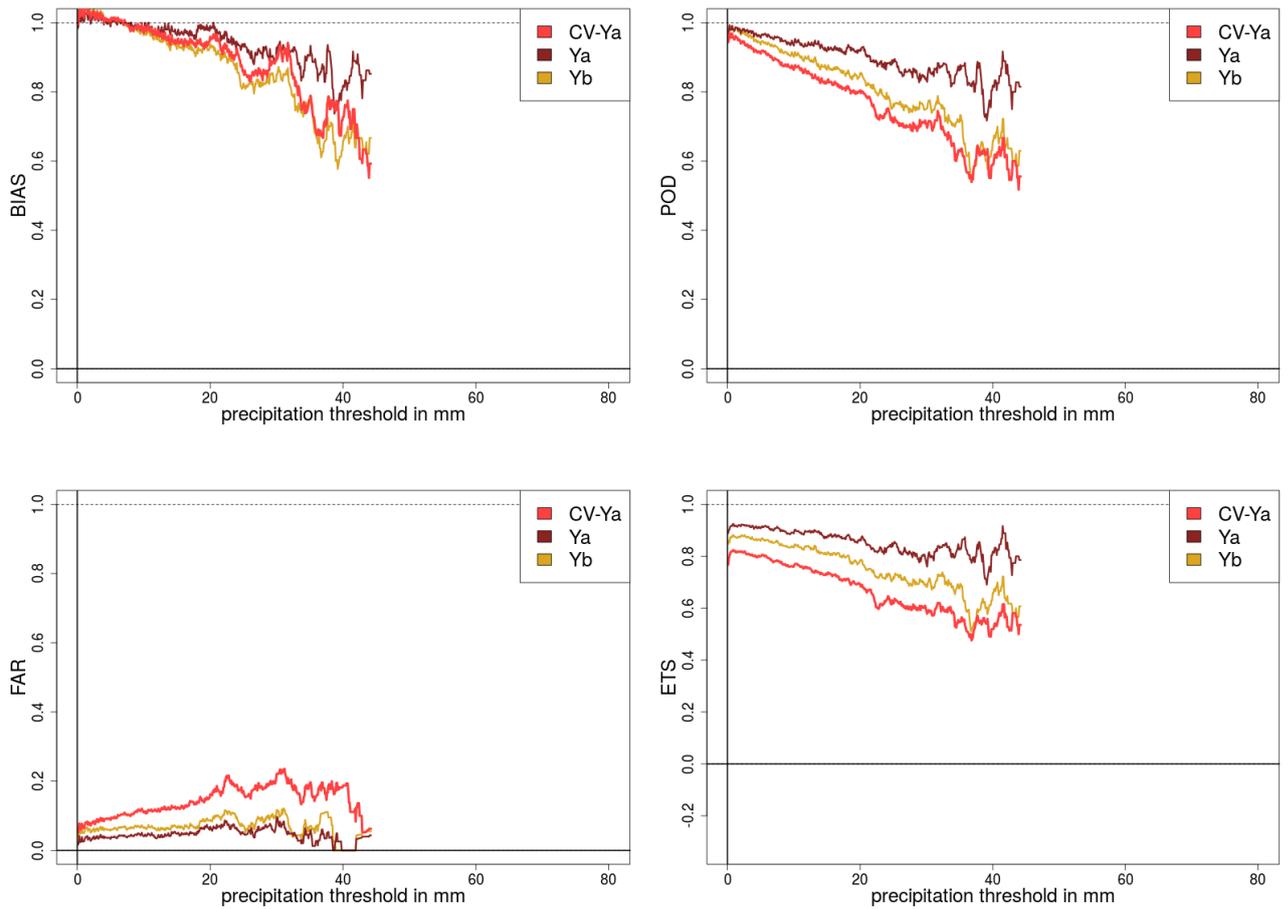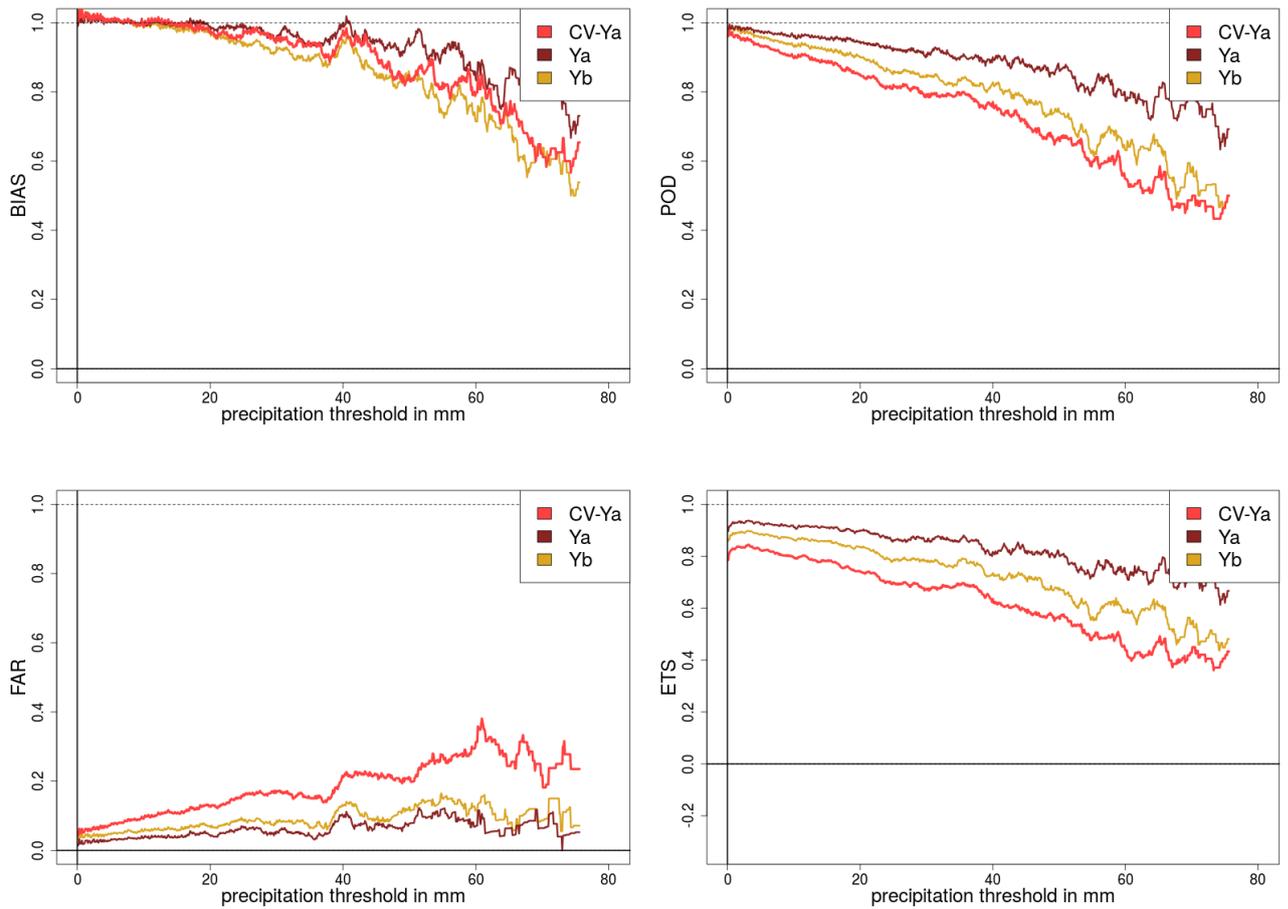
Figure 14: RR-Rad. *DJF* dataset: from 2015.12.01 to 2016.03.01. Categorical statistics computed from the contingency table: BIAS (top-left); POD (top-right); FAR (bottom-left); ETS (bottom-right).

Appendix

A Output grid specifications

- Coordinate Reference system. proj4 string="+proj=utm +zone=33 +datum=WGS84 +units=m +no_defs +ellps=WGS84 +towgs84=0,0,0"

- dimensions (number of grid points): easting = 1195 ; northing = 1550 ;

- grid spacing: easting = 1000 meter ; northing = 1000 meter.

B Spatial Consistency Test (SCT) and disaggregation

of daily to hourly precipitation

This Section describes the simultaneous application of SCT and disaggregation of daily to hourly precipitation. The notation $\mathbf{a} circ \mathbf{b}$ denotes the component-wise multiplications between two vectors

- *loop1*.

  - temporal disaggregation of daily precipitation to hourly precipitation (see Sec. B.1)

  - *loop2*. For every timestep $t = 1, \ldots, 24$

    * *loop3*.

      · $\mathbf{y}^{\mathrm{a}}(t) = \mathbf{y}^{\mathrm{b}}(t) + \mathbf{W}\left[\mathbf{y}^{\mathrm{o}}(t) - \mathbf{y}^{\mathrm{b}}(t)\right]$

      · if $\max\left\{\left[\mathbf{y}^{\mathrm{o}}(t) - \mathbf{y}^{\mathrm{a}}(t)\right] \circ \left[\mathbf{y}^{\mathrm{o}}(t) - \check{\mathbf{y}}^{\mathrm{a}}(t)\right]\right\} > \mathrm{T}^{h}$ then

        1. identify the i-th observation such that $\left[y_i^{\mathrm{o}}(t) - y_i^{\mathrm{a}}(t)\right]\left[y_i^{\mathrm{o}}(t) - \check{y}_i^{\mathrm{a}}(t)\right] = \max\left\{\left[\mathbf{y}^{\mathrm{o}}(t) - \mathbf{y}^{\mathrm{a}}(t)\right] \circ \left[\mathbf{y}^{\mathrm{o}}(t) - \check{\mathbf{y}}^{\mathrm{a}}(t)\right]\right\}$

        2. flag $y_i^{\mathrm{o}}(t)$ as a suspect observation

        3. if $y_i^{\mathrm{o}}(t)$ is obtained by temporal disaggregation of a daily observation, then flag as suspect y_i^{o} and all the derived hourly observations $y_i^{\mathrm{o}}(t)$, $t = 1, \ldots 24$

        4. restart *loop 1*

      · else: no more suspect observations to flag, exit *loop3*.

- ∗ end *loop3.*

- – end *loop2.*

- – $\mathbf{y}^{\text{SCTb}} = \sum_{t=1}^{24} \mathbf{y}^{\text{a}}(t)$

- – $\mathbf{y}^{\text{SCTa}} = \mathbf{y}^{\text{SCTb}} + \mathbf{W}\left[\mathbf{y}^{\text{o}} - \mathbf{y}^{\text{SCTb}}\right]$

- – if $\max\left\{\left[\mathbf{y}^{\text{o}} - \mathbf{y}^{\text{SCTa}}\right] \circ \left[\mathbf{y}^{\text{o}} - \check{\mathbf{y}}^{\text{SCTa}}\right]\right\} > \text{T}^d$ then

  - ∗ identify the *j*-th observation such that $\left[y_j^{\text{o}} - y_j^{\text{SCTa}}\right]\left[y_j^{\text{o}} - \check{y}_j^{\text{SCTa}}\right] = \max\left\{\left[\mathbf{y}^{\text{o}} - \mathbf{y}^{\text{SCTa}}\right] \circ \left[\mathbf{y}^{\text{o}} - \right.\right.$

  - ∗ flag as suspect y_j^{o} and the associated hourly observations $y_j^{\text{o}}(t)$, $t = 1, \ldots 24$
    , if any.

- – else: no more suspect observations to flag, exit *loop1.*

- end *loop1.*

B.1 temporal disaggregation of daily precipitation to hourly precipitation

Note: $\mathbf{W}'$ is defined as $\mathbf{W}$ in Eq. (19) except that here we set $D_H = 50\,\text{Km}$.

- *loop.* For every timestep $t = 1, \ldots, 24$

  - – $\widetilde{\mathbf{y}}^{\text{a}}(t) = \mathbf{y}^{\text{b}}(t) + mathbfW'\left[\widetilde{\mathbf{y}}^{\text{o}}(t) - \mathbf{y}^{\text{b}}(t)\right]$.

  - – if $\left(\widetilde{\mathbf{y}}^{\text{a}}(t) < \mathbf{y}^{\text{b}}(t)\right)$ then $\widetilde{\mathbf{y}}^{\text{a}}(t) = \mathbf{y}^{\text{b}}(t)$

  - – if $\left(\widetilde{\mathbf{y}}^{\text{a}}(t) < 0\right)$ then $\widetilde{\mathbf{y}}^{\text{a}}(t) = 0$

- end *loop.*

- *loop* over the daily observations to be disaggregated $m = 1, \ldots, M$

  - – if for some reason $\widetilde{y}_m^{\text{a}}(t)$ is not defined for one or more t, then set the location
    as not available for spatial interpolation

  - – $\widetilde{y}_m^{\text{a}} = \sum_{t=1}^{24} \widetilde{y}_m^{\text{a}}(t)$

  - – in case $\widetilde{y}_m^{\text{a}}$ does measure precipitation and:

    - ∗ y_m^{o} does measure precipitation, then $c_m(t) = \widetilde{y}_m^{\text{a}}(t)/\widetilde{y}_m^{\text{a}}$, $t = 1, \ldots, 24$

    - ∗ y_m^{o} does not measure precipitation, then $c_m(t) = 0, \forall t$

  - – in case $\widetilde{y}_m^{\text{a}}$ does not measure precipitation and:

    - ∗ y_m^{o} does measure precipitation, then $c_m(t) = 1/24, \forall t$

32

* y_m^o does not measure precipitation, then $c_m(t) = 0, \forall t$

- end *loop*.

- Finally, the disaggregated observations are obtained by using the coefficients:

$$\mathbf{y}^o(t) = \mathbf{c}(t) \circ \widetilde{\mathbf{y}}^o(t)$$

C SCT formulation for one single observation

This Section is mostly based on *Lorenc* (1984) and *Lorenc and Hammon* (1988). As in the previous Sections, SCT stands for Spatial Consistency Test. The object is to check for the presence of gross measurement errors in our observations.

Consider an observation and a prior estimate of the unknown true value (i.e. background) at one fixed location in space, we define the events:

$$
\begin{aligned}
T &= \text{(true value is between } t \text{ and } t+dt\text{)} \\
O &= \text{(observed value is between } o \text{ and } o+do\text{)} \\
B &= \text{(background value is between } b \text{ and } b+db\text{)} \\
G &= \text{(observation is affected by gross error)}
\end{aligned}
$$

Notation: discrete events are indicated as capital letters; their probabilities are denoted by $P(\ldots)$; $P(X|Y)$ is the a posteriori probability of X, given that Y has occurred; the continuous variables are in small letters; prior knowledge of b is implicit in all of our probabilities; PDF stands for probability density function. The SCT we are describing is based on the following approximation:

$$
\begin{aligned}
P(G|O) > \frac{1}{2} &\Rightarrow P(G|O) = 1 \\
P(G|O) \leq \frac{1}{2} &\Rightarrow P(G|O) = 0
\end{aligned}
\tag{24}
$$

If given the observed value o (and first guess value b) we have a probability of gross measurements error greater than $\frac{1}{2}$, then we assume that the observation is affected by gross measurement error. As a consequence, the observation should be flagged as suspect observation and not used in the interpolation procedure.

To implement the test of Eq. (24), $P(G|O)$ must be specified.

By using the Bayes' theorem we can write:

$$P(G|O) = P(O|G)\frac{P(G)}{P(O)} \tag{25}$$

If a gross measurements error is actually present, the corresponding observed value is not informative of the true state of the atmosphere. In other words, a uniform PDF is assumed for the $P(O|G)$ term of Eq. (25):

$$P(O|G) = k \, do \tag{26}$$

where k is a normalization constant related to the range (width $1/k$) of plausible values, such that $\int_{-\infty}^{+\infty} k \, do = 1$.

The term $P(G)$ in Eq. (25) represents the a-priori probability of having a gross measurements error and it may be regarded as a property of the measurements network. In general, $P(G)$ can be estimated by using the history of the measurements network.

The last term we have to specify in Eq. (25) is $P(O)$, which we define taking into account the possibility of having a gross measurement error:

$$P(O) = P(O|G)\,P(G) + P(O|\overline{G})\,P(\overline{G}) \tag{27}$$

If the observation is not affected by gross error, we assume a Gaussian PDF for the observation error $N(x, \sigma^2)$:

$$P(O|\overline{G}) \;\; = \;\; N(o - t, \sigma_o^2) do$$

As a consequence:

$$P(O) = \left[kP(G) + N(o - t, \sigma_o^2)P(\overline{G}) \right] do \tag{28}$$

By substituting Eqs. (26)-(28) in Eq. (25) and by assuming know the reliability $P(G)$, the test condition of Eq. (24) can be written as:

$$(o - b)^2 > T^2(\sigma_o^2 + \sigma_b^2) \tag{29}$$

with:

$$T^2 = 2ln\frac{P(\overline{G})}{P(G)} - ln\left\{(\sigma_o^2 + \sigma_b^2)k^2\right\} - ln(2\pi)$$

According to *Lorenc* (1986): "we have thus derived an expression for the checking tolerance which is an empirically tuned parameter in many existing quality control scheme."

Proof: By Substituting Eqs. (26)-(28) in Eq. (25) we obtain:

$$P(G|O) = \frac{kP(G)}{kP(G) + N(o - b, \sigma_o^2 + \sigma_b^2)P(\overline{G})}$$

Then we can rewrite inequality $P(G|O) > \frac{1}{2}$ as:

$$\frac{kP(G)}{kP(G) + N(o - b, \sigma_o^2 + \sigma_b^2)P(\overline{G})} > \frac{1}{2}$$

$$\frac{1}{1 + N(o - b, \sigma_o^2 + \sigma_b^2)\frac{P(\overline{G})}{kP(G)}} > \frac{1}{2}$$

$$1 + N(o - b, \sigma_o^2 + \sigma_b^2)\frac{P(\overline{G})}{kP(G)} < 2$$

$$\frac{1}{\sqrt{2\pi\left(\sigma_o^2 + \sigma_b^2\right)}}e^{-\frac{1}{2}\left[\frac{(o-b)^2}{\sigma_o^2 + \sigma_b^2}\right]}\frac{P(\overline{G})}{P(G)} < k$$

$$\frac{(o - b)^2}{\sigma_o^2 + \sigma_b^2} > 2ln\left[\frac{P(\overline{G})}{P(G)}\right] - 2ln(k) - ln(2\pi) - ln(\sigma_o^2 + \sigma_b^2)$$

End Proof.

Note1 If the occurrence of gross errors is allowed as a possibility and explicitly treated in the theory, then the analysis PDF becomes distinctly bimodal. In contrast, in the standard OI the analysis PDF is always Gaussian (see Figs.(15) and (16)). Since operationally we need a single best analysis, then we have the problem of picking the best value: *the local mode nearest the mean is the best*.

When observation and background are close to each other, the best analysis is close to the observation. If the difference between observation and background is gradually increasing, we reach a point where the best analysis lies close to the background. The turning point is $P(G|O) = 1/2$.
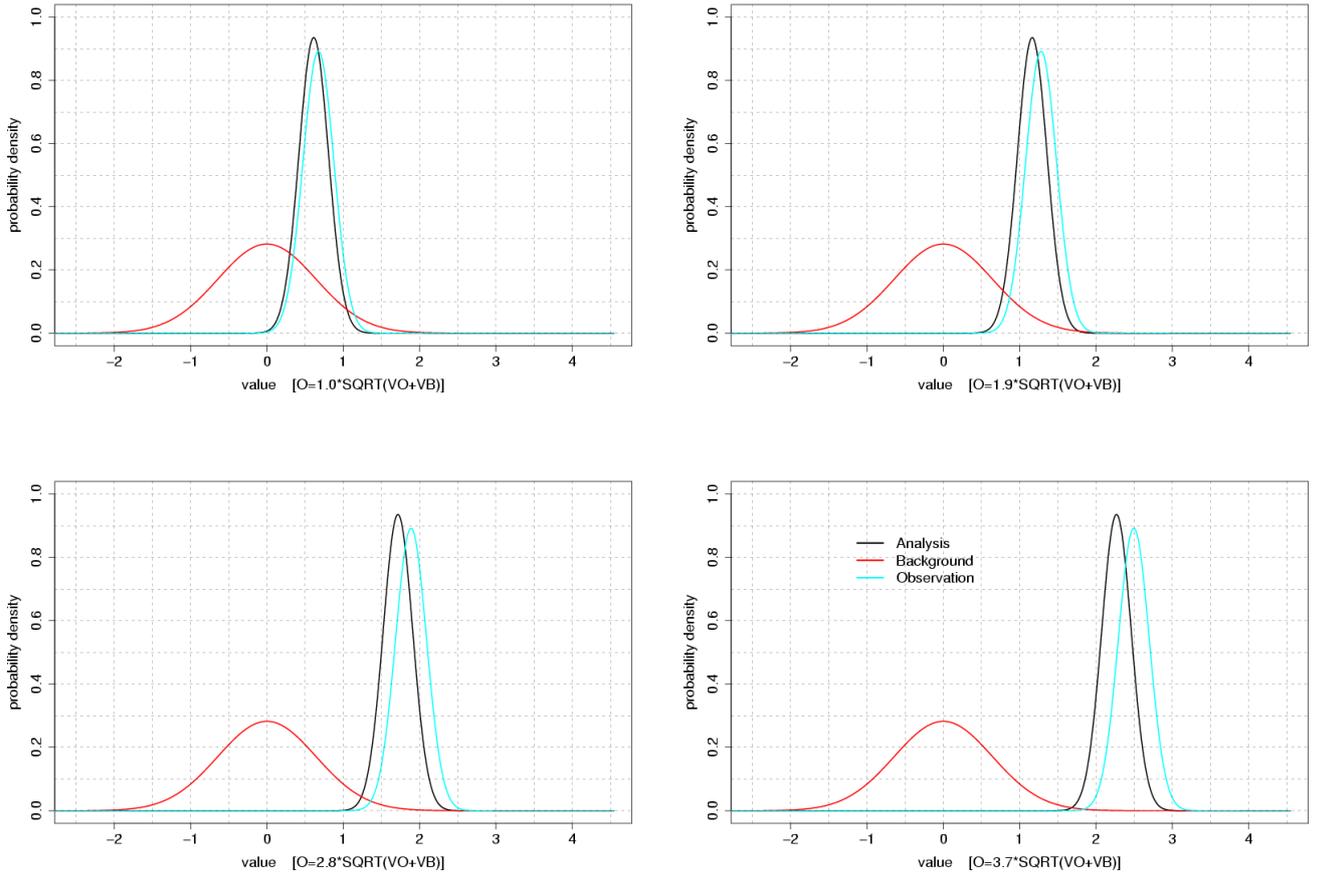
Figure 15: probability density functions for background, observation and analysis in 4 idealized cases. Gaussian PDF are assumed for error distributions with $[\sigma_o^2 = 0.1; \sigma_b^2 = 1]$ (taken from **?**).
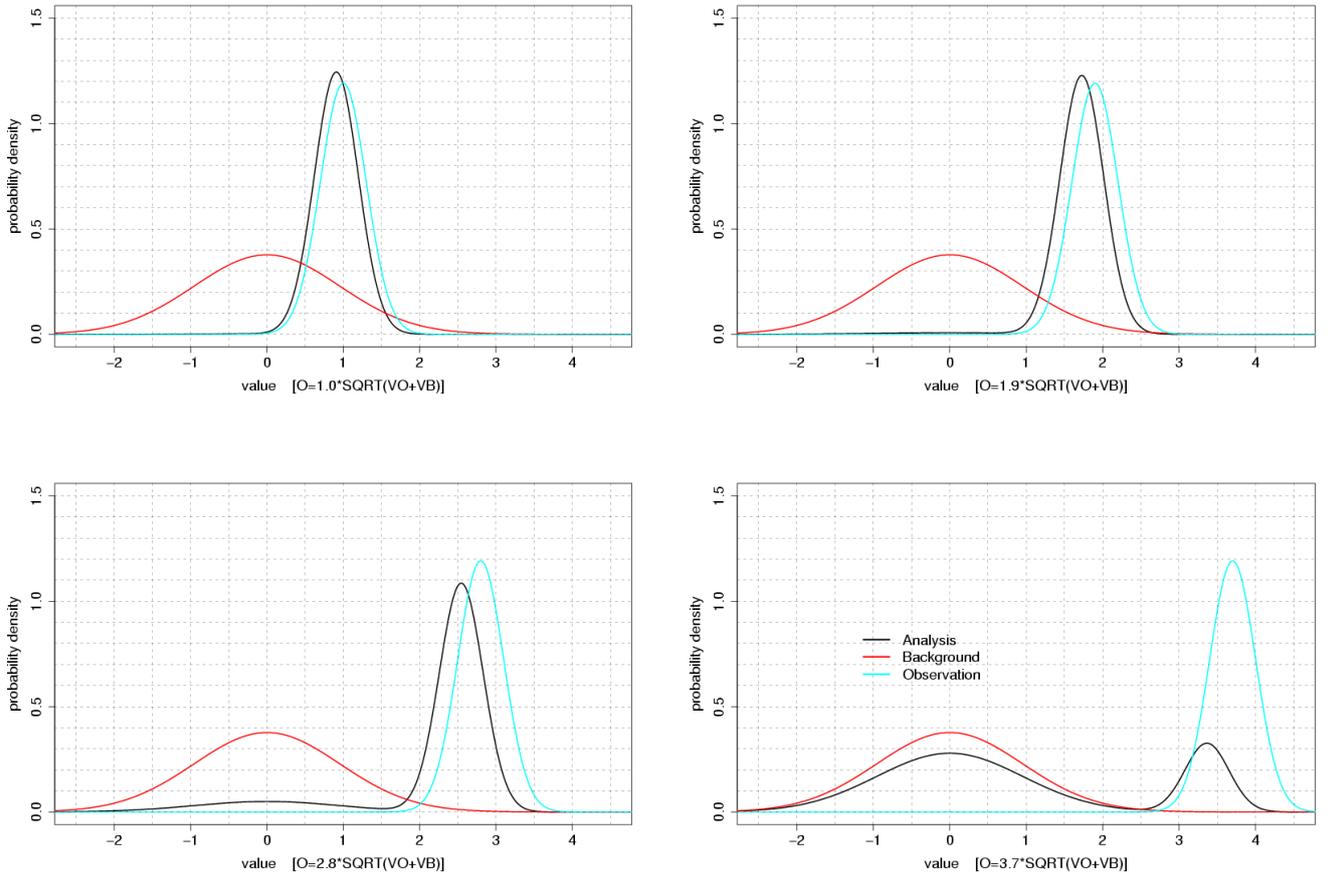
Figure 16: probability density functions for background, observation and analysis in 4 idealized cases. The error distributions are Gaussian but we introduce the possibility of gross errors in the observation as in Eq. (27) [$\sigma_o^2 = 0.1$; $\sigma_b^2 = 1$; $k = 0.017$; $P(G) = 0.06$] (taken from *Lorenc and Hammon* (1988)).

Note2 In the case of more than one observation, we deal with similar equations but we need to use $P(G_i | O_1 \cap O_2 \cap \ldots \cap O_N)$ instead of $P(O|G)$ and the final test formulation is described in *Lussana et al.* (2010).

D Verification scores

for reader's convenience we report the score definitions taken from *WWRP/WGNE Joint Working Group on Forecast Verification Research* http://www.cawcr.gov.au/projects/verification in Appendix

- Probability of detection (POD) or hit rate

    - Answers the question: What fraction of the observed "yes" events were correctly forecast?

    - Range: 0 to 1. Perfect score: 1.

    - Characteristics: Sensitive to hits, but ignores false alarms. Very sensitive to the climatological frequency of the event. Good for rare events. Can be artificially improved by issuing more "yes" forecasts to increase the number of hits. Should be used in conjunction with the false alarm ratio.

- False alarm ratio (FAR)

    - Answers the question: What fraction of the predicted "yes" events actually did not occur (i.e., were false alarms)?

    - Range: 0 to 1. Perfect score: 0.

    - Characteristics: Sensitive to false alarms, but ignores misses. Very sensitive to the climatological frequency of the event. Should be used in conjunction with the probability of detection.

- bias score (BIAS):

    - Answers the question: How did the forecast frequency of "yes" events compare to the observed frequency of "yes" events?

    - Range: 0 to ∞. Perfect score: 1.

– Characteristics: Measures the ratio of the frequency of forecast events to the frequency of observed events. Indicates whether the forecast system has a tendency to underforecast (BIAS<1) or overforecast (BIAS>1) events. Does not measure how well the forecast corresponds to the observations, only measures relative frequencies.

- Equitable threat score (ETS)

  – Answers the question: How well did the forecast "yes" events correspond to the observed "yes" events (accounting for hits due to chance)?

  – Range: -1/3 to 1, 0 indicates no skill. Perfect score: 1.

  – Characteristics: Measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance (for example, it is easier to correctly forecast rain occurrence in a wet climate than in a dry climate). The ETS is often used in the verification of rainfall in NWP models because its "equitability" allows scores to be compared more fairly across different regimes. Sensitive to hits. Because it penalises both misses and false alarms in the same way, it does not distinguish the source of forecast error.

References

American Meteorological Society (2016), Marshall–palmer relation. glossary of meteorology, `http://glossary.ametsoc.org/wiki/Marshall-palmer_relation`, [Online; accessed 28-December-2016].

Daley, R. (1991), *Atmospheric Data Analysis*, 457 pp., Cambridge University Press.

Gandin, L. S., and R. Hardin (1965), *Objective analysis of meteorological fields*, vol. 242, Israel program for scientific translations Jerusalem.

Jazwinski, A. H. (2007), *Stochastic processes and filtering theory*, Courier Dover Publications.

Lorenc, A. (1984), Analysis methods for the quality control of observations, in *proceedings of ECMWF workshop on the use and quality control of meteorological observations for numerical weather prediction*, pp. 397–428, 6-9 November 1984, Reading, UK.

Lorenc, A. (1986), Analysis methods for numerical weather prediction, *Quart. J. Roy. Meteorol. Soc.*, *112*, 1177–1194.

Lorenc, A., and O. Hammon (1988), Objective quality control of observations using bayesian methods. theory, and a practical implementation, *Quarterly Journal of the Royal Meteorological Society*, *114*(480), 515–543.

Lussana, C., F. Uboldi, and M. R. Salvati (2010), A spatial consistency test for surface observations from mesoscale meteorological networks, *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 1075–1088.

Mahfouf, J.-F., B. Brasnett, and S. Gagnon (2007), A canadian precipitation analysis (capa) project: Description and preliminary results, *Atmosphere-Ocean*, *45*(1), 1–17.

Marshall, J. S., and W. M. K. Palmer (1948), The distribution of raindrops with size, *Journal of meteorology*, *5*(4), 165–166.

Soci, C., E. Bazile, F. Besson, and T. Landelius (2016), High-resolution precipitation re-analysis system for climatological purposes, *Tellus A*, *68*.

Uboldi, F., C. Lussana, and M. Salvati (2008), Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks, *Meteorological Applications*, *15*(3), 331–345.