GATNDOR topic: Collocation uncertainty in vertical profiles. Toward an unified approach for vertical profiles

Alessandro Fassò*

in collaboration with Fabio Madonna

* University of Bergamo - DIIMM - alessandro.fasso@unibg.it

Tokyo, **ICM-4**, 5-9 March 2012

RegioneLombardia

Scientific questions

We would like to answer the following questions:

1. is the collocation uncertainty related to environmental factors ?
2. is the collocation uncertainty related to the paired trajectories distance ?
3. Is the collocation uncertainty related to height ?
4. Are above point valid for all ECV ?
5. Is uncertainty a static or dynamic concept ?

Scientific questions

We would like to answer the following questions:

1. is the collocation uncertainty related to environmental factors ?
2. is the collocation uncertainty related to the paired trajectories distance ?
3. Is the collocation uncertainty related to height ?
4. Are above point valid for all ECV ?
5. Is uncertainty a static or dynamic concept ?

Scientific questions

We would like to answer the following questions:

1. is the collocation uncertainty related to environmental factors ?
2. is the collocation uncertainty related to the paired trajectories distance ?
3. Is the collocation uncertainty related to height ?
4. Are above point valid for all ECV ?
5. Is uncertainty a static or dynamic concept ?

Scientific questions

We would like to answer the following questions:

1. is the collocation uncertainty related to environmental factors ?
2. is the collocation uncertainty related to the paired trajectories distance ?
3. Is the collocation uncertainty related to height ?
4. Are above point valid for all ECV ?
5. Is uncertainty a static or dynamic concept ?

Scientific questions

We would like to answer the following questions:

1. is the collocation uncertainty related to environmental factors ?
2. is the collocation uncertainty related to the paired trajectories distance ?
3. Is the collocation uncertainty related to height ?
4. Are above point valid for all ECV ?
5. Is uncertainty a static or dynamic concept ?

Introduction

"... The primary goals of GRUAN are to **provide vertical profiles** of reference measurements suitable for reliably detecting changes in global and regional climate on decadal time scales" (GRUAN Manual V7).

$$\Downarrow$$

This implies the concept of profile uncertainty

Introduction

- We consider here an empirical approach to uncertainty analysis with no reference to a priori metadata but relying on data by means of a statistical modelling approach.
- The integration with technical information on sensors can be done at any stage.
- Automatic use of pre-identified models is useful in Quality monitoring and nearly automatic validation
- NRT & the operational GRUAN may benefit by nearly-automatic QA/QC
- Automatic model update is possible as new data become available

Introduction

- We consider here an empirical approach to uncertainty analysis with no reference to a priori metadata but relying on data by means of a statistical modelling approach.
- The integration with technical information on sensors can be done at any stage.
- Automatic use of pre-identified models is useful in Quality monitoring and nearly automatic validation
- NRT & the operational GRUAN may benefit by nearly-automatic QA/QC
- Automatic model update is possible as new data become available

Introduction

- We consider here an empirical approach to uncertainty analysis with no reference to a priori metadata but relying on data by means of a statistical modelling approach.
- The integration with technical information on sensors can be done at any stage.
- Automatic use of pre-identified models is useful in Quality monitoring and nearly automatic validation
- NRT & the operational GRUAN may benefit by nearly-automatic QA/QC
- Automatic model update is possible as new data become available

Introduction

- We consider here an empirical approach to uncertainty analysis with no reference to a priori metadata but relying on data by means of a statistical modelling approach.
- The integration with technical information on sensors can be done at any stage.
- Automatic use of pre-identified models is useful in Quality monitoring and nearly automatic validation
- NRT & the operational GRUAN may benefit by nearly-automatic QA/QC
- Automatic model update is possible as new data become available

Introduction

- We consider here an empirical approach to uncertainty analysis with no reference to a priori metadata but relying on data by means of a statistical modelling approach.
- The integration with technical information on sensors can be done at any stage.
- Automatic use of pre-identified models is useful in Quality monitoring and nearly automatic validation
- NRT & the operational GRUAN may benefit by nearly-automatic QA/QC
- Automatic model update is possible as new data become available

Introduction (continued)

Temporal stochastic structure of vertical profiles is important from the statistical point of view not only for facing the collocation problem but may be useful in various applications such as:

- network design
- redundancy
- BESTA
- multiresolution data fusion
- Spatial trend analysis
- Temporal trend analysis

Outline

1. Stochastic spatio temporal models
2. Spatio temporal functional data analysis FDA
3. Collocation model
   1. general case
   2. linear case
4. Application to pressure and temperature from Beltsville-Sterling radiosonde
5. Toolbox scheme

Thanks

1. Fabio Madonna for introducing me into GRUAN data and sensors in general, identifying the relevant dataset, and fruitful discussions and collocation and else ...
2. Belay Demoz for providing the data
3. Lombardy Region's Project EN17-FA2009, *'Methods for the integration of different renewable energy sources and impact monitoring with satellite data'.*

Stochastic Modeling of Vertical Profiles

- Let vertical profile measurements (T,RH,pressure ..) be given by

$$y(s,t,h), \quad s \in \mathcal{S}, \quad t \geq t_0, \quad h \geq h_0$$

where t_0 and h_0 are launch time and height.

- We may use two useful approaches for describing the variability and uncertainty of y

Stochastic Modeling of Vertical Profiles

- Let vertical profile measurements (T,RH,pressure ..) be given by
$$y\left(s,t,h\right), \quad s\in\mathcal{S}, \quad t\geq t_0, \quad h\geq h_0$$
where t_0 and h_0 are launch time and height.
- We may use two useful approaches for describing the variability and uncertainty of y

Spatio Temporal Stochastic Models

We suppose that data are generated by a mix of fixed and random factors

$$y = \mu + \tau\varepsilon$$

- All components depend on (s, t, h)
- μ is a Deterministic/Random trend component
- τ^2 *is* a Deterministic/Random uncertainty component
- ε is a standardized (**Gaussian !?**) ideal instrumental error.
- y may be a scalar or a vector: mutivariable (MIMO) modelling is not considered here but can be covered in this approach contributing to BESTA, redundancy analysis and multiresolution collocation.

Spatio Temporal Stochastic Models

We suppose that data are generated by a mix of fixed and random factors

$$y = \mu + \tau\varepsilon$$

- All components depend on (s, t, h)
- μ is a Deterministic/Random trend component
- τ^2 *is* a Deterministic/Random uncertainty component
- ε is a standardized (**Gaussian !?**) ideal instrumental error.
- y may be a scalar or a vector: mutivariable (MIMO) modelling is not considered here but can be covered in this approach contributing to BESTA, redundancy analysis and multiresolution collocation.

Spatio Temporal Stochastic Models

We suppose that data are generated by a mix of fixed and random factors

$$y = \mu + \tau\varepsilon$$

- All components depend on (s, t, h)
- μ is a Deterministic/Random trend component
- τ^2 *is* a Deterministic/Random uncertainty component
- ε is a standardized (**Gaussian !?**) ideal instrumental error.
- y may be a scalar or a vector: mutivariable (MIMO) modelling is not considered here but can be covered in this approach contributing to BESTA, redundancy analysis and multiresolution collocation.

Spatio Temporal Stochastic Models

We suppose that data are generated by a mix of fixed and random factors

$$y = \mu + \tau\varepsilon$$

- All components depend on (s, t, h)
- μ is a Deterministic/Random trend component
- τ^2 *is* a Deterministic/Random uncertainty component
- ε is a standardized (**Gaussian !?**) ideal instrumental error.
- y may be a scalar or a vector: mutivariable (MIMO) modelling is not considered here but can be covered in this approach contributing to BESTA, redundancy analysis and multiresolution collocation.

Spatio Temporal Stochastic Models

We suppose that data are generated by a mix of fixed and random factors

$$y = \mu + \tau\varepsilon$$

- All components depend on (s, t, h)
- μ is a Deterministic/Random trend component
- τ^2*is* a Deterministic/Random uncertainty component
- ε is a standardized (**Gaussian !?**) ideal instrumental error.
- y may be a scalar or a vector: mutivariable (MIMO) modelling is not considered here but can be covered in this approach contributing to BESTA, redundancy analysis and multiresolution collocation.

Trend component

The measurement trend is given by

$$\mu = b(x) + m(x) + \omega$$

where

- All component depend on (s, t, h)
- x is a set of environmental factors related to point (s, t, h)
- b is a determinstic (or stochastic) instrumental bias
  e.g.: $b(s, t, h) = b_i$ constant bias for instrument type "i"
- m is the "true" profile
  For example seasonal local linear component $m(x) = \beta(h)_t' x$

Trend component

The measurement trend is given by

$$\mu = b(x) + m(x) + \omega$$

where

- All component depend on (s, t, h)
- x is a set of environmental factors related to point (s, t, h)
- b is a determinstic (or stochastic) instrumental bias
  e.g.: $b(s, t, h) = b_i$ constant bias for instrument type "i"
- m is the "true" profile
  For example seasonal local linear component $m(x) = \beta(h)_t' x$

Trend component

The measurement trend is given by

$$\mu = b(x) + m(x) + \omega$$

where

- All component depend on (s, t, h)
- x is a set of environmental factors related to point (s, t, h)
- b is a determinstic (or stochastic) instrumental bias
  e.g.: $b(s, t, h) = b_i$ constant bias for instrument type "i"
- m is the "true" profile
  For example seasonal local linear component $m(x) = \beta(h)_t' x$

Trend component

The measurement trend is given by

$$\mu = b(x) + m(x) + \omega$$

where

- All component depend on (s, t, h)
- x is a set of environmental factors related to point (s, t, h)
- b is a determinstic (or stochastic) instrumental bias
  e.g.: $b(s, t, h) = b_i$ constant bias for instrument type "i"
- m is the "true" profile
  For example seasonal local linear component $m(x) = \beta(h)_t' x$

Trend component (continued)

...

$$\mu = b(x) + m(x) + \omega$$

- ω is a zero mean (Gaussian) spatially, temporally and vertically correlated random factor, independent on x
- e.g.

$$\omega(s,t,h) = \omega_S(s) + \omega_T(t) + \omega_H(h)$$

  with
  - ω_S a geostatistical component
  - ω_T a markovian component over time
  - ω_H a markovian component over the vertical profile

Trend component (continued)

...

$$\mu = b(x) + m(x) + \omega$$

- ω is a zero mean (Gaussian) spatially, temporally and vertically correlated random factor, independent on x
- e.g.

$$\omega(s,t,h) = \omega_S(s) + \omega_T(t) + \omega_H(h)$$

with
- ω_S a geostatistical component
- ω_T a markovian component over time
- ω_H a markovian component over the vertical profile

Skedastic component

- The skedastic component τ^2 has a structure which is similar to μ but
  - without bias component and
  - with possibly different environmental factors

$$\tau^2 = \sigma^2(x) + \zeta > 0$$

- ζ is a zero mean small white noise
- $\sigma^2 > 0$ is the skedastic component
  e.g. $\sigma^2(x) = \alpha_t' x$

Skedastic component

- The skedastic component τ^2 has a structure which is similar to μ but
  - without bias component and
  - with possibly different environmental factors

$$\tau^2 = \sigma^2(x) + \zeta > 0$$

- ζ is a zero mean small white noise
- $\sigma^2 > 0$ is the skedastic component
  e.g. $\sigma^2(x) = \alpha_t' x$

Skedastic component

- The skedastic component τ^2 has a structure which is similar to μ but
  - without bias component and
  - with possibly different environmental factors

$$\tau^2 = \sigma^2(x) + \zeta > 0$$

- ζ is a zero mean small white noise
- $\sigma^2 > 0$ is the skedastic component
  e.g. $\sigma^2(x) = \alpha_t' x$

Conditional uncertainty profile

- Consider the space-time vertical trajectory

$$h \rightarrow (s_h, t_h) \text{ for } h_0 \leq h \leq h_1$$

- $$U(h|x) = E\left((y-\mu)^2 |x\right) = b(x)^2 + \sigma_\omega^2 + \sigma^2(h)$$

- The second term is the colored random component

$$\sigma_\omega^2 = Var(\omega(s,t,h))$$

which describes the unaccounted environmental factors

Conditional uncertainty profile

- Consider the space-time vertical trajectory

$$h \longrightarrow (s_h, t_h) \text{ for } h_0 \leq h \leq h_1$$

- $$U(h|x) = E\left((y-\mu)^2 |x\right) = b(x)^2 + \sigma_\omega^2 + \sigma^2(h)$$

- The second term is the colored random component

$$\sigma_\omega^2 = Var(\omega(s,t,h))$$

which describes the unaccounted environmental factors

Conditional uncertainty profile

- Consider the space-time vertical trajectory

$$h \longrightarrow (s_h, t_h) \text{ for } h_0 \leq h \leq h_1$$

- $$U(h|x) = E\left((y-\mu)^2 |x\right) = b(x)^2 + \sigma_\omega^2 + \sigma^2(h)$$

- The second term is the colored random component

$$\sigma_\omega^2 = Var(\omega(s,t,h))$$

which describes the unaccounted environmental factors

Conditional total uncertainty

We take for semplicity $b = 0$. Then the total uncertainty is

$$U(x) = \frac{1}{\Delta h} \int_{h_0}^{h_1} U(h|x)\, dh = \sigma_\omega^2 + \frac{1}{\Delta h} \int_{h_0}^{h_1} \sigma^2(h|x)\, dh$$

Global total uncertainty

When enough information about x is available, we may compute the total marginal uncertainty

$$U = E_x\left(U\left(x\right)\right) = \sigma^2_\omega + \frac{1}{\Delta h}\int_{h_0}^{h_1} E_x\left(\sigma^2\left(h\right)|x\right)dh$$

Functional data analysis

We consider now a vertical profile as a single object (smooth function):

$$\mu(\cdot) = \mu(h) = \mu\left(s'_h, t'_h, h\right), \quad h \geq h_0$$

so observation labelled by launch place and time (s, t) is given by a random function

$$y_{s,t} = \mu_{s,t}(\cdot) + \varepsilon_{s,t}(\cdot)$$

where $\mu(\cdot)$ is the "true" profile and $\varepsilon(\cdot)$ is the zero mean functional error.

Spatio temporal FDA and regression FDA

- If $\mu_{s,t}(\cdot)$ or $\varepsilon_{s,t}(\cdot)$ are correlated over space (s) and/or time (t), we have spatio temporal functional data models.
- If (functional) environmental factors $x(\cdot)$ are in force

$$\mu(\cdot) = \beta(\cdot)' x(\cdot) + \omega(\cdot)$$

Spatio temporal FDA and regression FDA

- If $\mu_{s,t}(\cdot)$ or $\varepsilon_{s,t}(\cdot)$ are correlated over space (s) and/or time (t), we have spatio temporal functional data models.
- If (functional) environmental factors $x(\cdot)$ are in force

$$\mu(\cdot) = \beta(\cdot)' x(\cdot) + \omega(\cdot)$$

Conditional uncertainty profile

- Assuming no bias, $b = 0$, the conditional uncertainty profile is given by
$$U(\cdot|x) = \sigma_\omega^2(\cdot) + \sigma_\varepsilon^2(\cdot)$$
- The first term is the colored random component function
$$\sigma_\omega^2(h) = Var(\omega(h))$$
which describes the unaccounted functional environmental factors

Conditional uncertainty profile

- Assuming no bias, $b = 0$, the conditional uncertainty profile is given by
$$U(\cdot|x) = \sigma_\omega^2(\cdot) + \sigma_\varepsilon^2(\cdot)$$
- The first term is the colored random component function
$$\sigma_\omega^2(h) = Var(\omega(h))$$
which describes the unaccounted functional environmental factors

Conditional total uncertainty

The average/total uncertainty is given by

$$
\begin{aligned}
U(x) &= \frac{1}{\Delta h} \int_{h_0}^{h_1} U(h|x)\, dh \\
&= \frac{1}{\Delta h} \int_{h_0}^{h_1} \sigma_{\omega}^{2}(h)\, dh + \frac{1}{\Delta h} \int_{h_0}^{h_1} \sigma_{\varepsilon}^{2}(h)\, dh
\end{aligned}
$$

Global total uncertainty

When enough information about x is available we can compute the total marginal uncertainty

$$U = \frac{1}{\Delta h} \int_{h_0}^{h_1} \beta^2 (h)\, \sigma_x^2 (h)\, dh + E_x (U (x))$$

Collocation model

Suppose we are comparing two instruments with the same resolution.
According to the stochastic model approach we have

$$\Delta y = y\left(s, t, h\right) - y'\left(s', t', h\right) = \Delta\mu + \Delta b + \Delta\left(\sigma\varepsilon\right)$$

- $\Delta\mu$ is the collocation drift
- Δb is the collocation instrumental bias
- $\Delta\left(\sigma\varepsilon\right)$ is the collocation conditional uncertainty

Collocation model

Suppose we are comparing two instruments with the same resolution.
According to the stochastic model approach we have

$$\Delta y = y(s,t,h) - y'(s',t',h) = \Delta\mu + \Delta b + \Delta(\sigma\varepsilon)$$

- $\Delta\mu$ is the collocation drift
- Δb is the collocation instrumental bias
- $\Delta(\sigma\varepsilon)$ is the collocation conditional uncertainty

Collocation model

Suppose we are comparing two instruments with the same resolution.
According to the stochastic model approach we have

$$\Delta y = y(s,t,h) - y'(s',t',h) = \Delta\mu + \Delta b + \Delta(\sigma\varepsilon)$$

- $\Delta\mu$ is the collocation drift
- Δb is the collocation instrumental bias
- $\Delta(\sigma\varepsilon)$ is the collocation conditional uncertainty

1. Parametric approach (Gaussian) to conditional, profile and total uncertainty

$$\begin{aligned} Total &: E\left(\left(\Delta y\right)^2 |x\right), \\ Collocation\ drift &: E\left(\left(\Delta \mu\right)^2 |x\right), \\ Instrumental\ \ Bias &: E\left(\left(\Delta b\right)^2 |x\right), \\ Conditional\ \ error &: E\left(\left(\Delta\left(\sigma\varepsilon\right)\right)^2 |x\right) \end{aligned}$$

2. Quantile profile (e.g. 95%) approach

$$\zeta_p = \zeta_p\left(h\right) : P\left(\left|\Delta y\right| > \zeta_p |x\right) = p$$

1. Parametric approach (Gaussian) to conditional, profile and total uncertainty

$$Total : E\left(\left(\Delta y\right)^2 |x\right),$$
$$Collocation\ drift : E\left(\left(\Delta \mu\right)^2 |x\right),$$
$$Instrumental\ \ Bias : E\left(\left(\Delta b\right)^2 |x\right),$$
$$Conditional\ \ error : E\left(\left(\Delta\left(\sigma\varepsilon\right)\right)^2 |x\right)$$

2. Quantile profile (e.g. 95%) approach

$$\zeta_p = \zeta_p\left(h\right) : P\left(|\Delta y| > \zeta_p|x\right) = p$$

Uncertainty by collocation risk

- Total conditional profile collocation risk, for given λ

$$P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global profile collocation risk

$$P\left(max_t|\Delta y| > \lambda|h\right) = E_{x|h}P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global collocation risk

$$P\left(max_t|\Delta y| > \lambda\right) = E_h P\left(max_t|\Delta y| > \lambda|h\right)$$

- Similar definitions apply for drift, bias and conditional error

Uncertainty by collocation risk

- Total conditional profile collocation risk, for given λ

$$P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global profile collocation risk

$$P\left(max_t|\Delta y| > \lambda|h\right) = E_{x|h}P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global collocation risk

$$P\left(max_t|\Delta y| > \lambda\right) = E_h P\left(max_t|\Delta y| > \lambda|h\right)$$

- Similar definitions apply for drift, bias and conditional error

Uncertainty by collocation risk

- Total conditional profile collocation risk, for given λ

$$P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global profile collocation risk

$$P\left(max_t|\Delta y| > \lambda|h\right) = E_{x|h}P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global collocation risk

$$P\left(max_t|\Delta y| > \lambda\right) = E_h P\left(max_t|\Delta y| > \lambda|h\right)$$

- Similar definitions apply for drift, bias and conditional error

Uncertainty by collocation risk

- Total conditional profile collocation risk, for given λ

$$P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global profile collocation risk

$$P\left(max_t|\Delta y| > \lambda|h\right) = E_{x|h}P\left(max_t|\Delta y| > \lambda|x\right)$$

- Total global collocation risk

$$P\left(max_t|\Delta y| > \lambda\right) = E_h P\left(max_t|\Delta y| > \lambda|h\right)$$

- Similar definitions apply for drift, bias and conditional error

Multiresolution collocation

- When comparing radiosonde with satellite data, we have that the two vertical profiles may be referred to two different spatial resolutions.
- In this case we prefer a joint modelling by the multivariate dynamic coregionalization model or other similar COSP approaches.

Multiresolution collocation

- When comparing radiosonde with satellite data, we have that the two vertical profiles may be referred to two different spatial resolutions.
- In this case we prefer a joint modelling by the multivariate dynamic coregionalization model or other similar COSP approaches.

Linear co-calibrated collocation

The level of modelling complexity must related to data availability (and human-time resources).
In the radiosonde collocation exercise below, paired baloons are supposed to have similar instruments with comparable calibration.
Then we have

$$\Delta b = 0$$

Linear collocation drift

We use a simple linear seasonal model for the **collocation conditional drift** $\Delta\mu$, namely

$$\Delta\mu = \beta' x$$

In the rest of the talk, with abuse of notation, we will use μ instead of $\Delta\mu$.

Linear collocation variance

Similarly, for the **collocation conditional variance** $\Delta(\tau\varepsilon)$

$$\Delta(\tau\varepsilon)^2 = \sigma^2(h|x) = \alpha' x$$

where among the candidate terms for x, we have also the distance among paired ballons:

$$\Delta s = s - s'$$

Total conditional Uncertainty

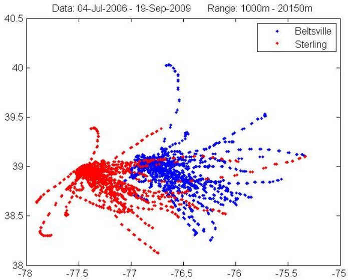
$$
\begin{aligned}
U(x) &= \frac{1}{\Delta h} \int_{h_0}^{h_1} \sigma^2 (h|x)\, dh \\
&= \frac{1}{\Delta h} \int_{h_0}^{h_1} \alpha' x(h)\, dh \cong \hat{\alpha}' \sum_{j=}^{n} x(h_j) \frac{\Delta h_j}{\Delta h}
\end{aligned}
$$

Global Uncertainty

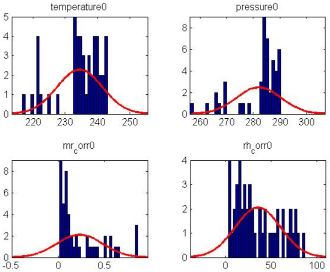
$$U(h) \cong \hat{E}_x\left(\hat{\mu}(h|x)^2\right) + \hat{E}_x\left(\hat{\sigma}^2(h|x)\right)$$

The Beltsville case study

Profiles averaged over 150m thick vertical resolution for the four ECV below



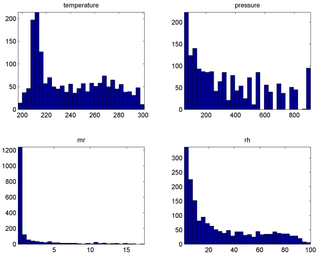
n=49 trajectories



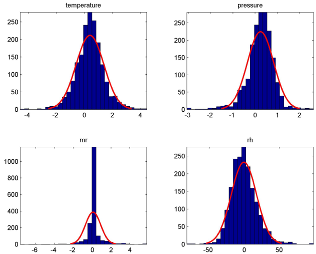
Histograms at 10,000m

The Data (continued)

Levels and collocation errors for $1,804$ records

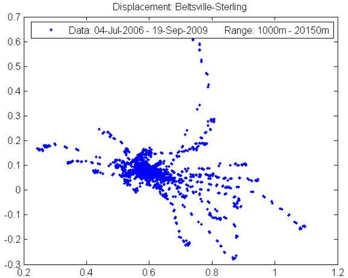


ECV's levels

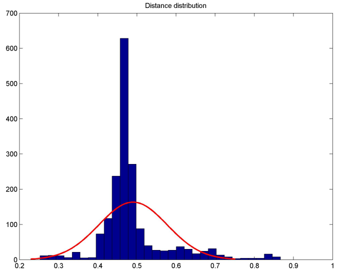


ECV's collocation errors

The Collocation distance



Mutual departure trajectories



Collocation distance at all hights

Modelling of pressure

- The response variable used here is pressure and all the data from the other radiosonda and all but pressure from the "collocated" radiosonda are used as explanatory factors.
- The resulting collocation error analysis corresponds to forecasting the single sensor rather than all radiosonda ECV's.
- Information on location, distance and ECV's have been used as regressors for both μ and σ^2 and selected using suitable statistical model selection techniques.

Modelling of pressure

- The response variable used here is pressure and all the data from the other radiosonda and all but pressure from the "collocated" radiosonda are used as explanatory factors.
- The resulting collocation error analysis corresponds to forecasting the single sensor rather than all radiosonda ECV's.
- Information on location, distance and ECV's have been used as regressors for both μ and σ^2 and selected using suitable statistical model selection techniques.

Modelling of pressure

- The response variable used here is pressure and all the data from the other radiosonda and all but pressure from the "collocated" radiosonda are used as explanatory factors.
- The resulting collocation error analysis corresponds to forecasting the single sensor rather than all radiosonda ECV's.
- Information on location, distance and ECV's have been used as regressors for both μ and σ^2 and selected using suitable statistical model selection techniques.

Modelling (continued)

- The collocation errors and the residuals are white noise in time but show vertical autocorrelation, which has been modelled as a vertical AR(1) model:

$$\varepsilon(h) = \rho\varepsilon(h - 150) + \xi(h)$$

- This autcorrelation structure is similar for μ and σ^2. It has been estimated using the empirical *GLS* approach.

Modelling (continued)

- The collocation errors and the residuals are white noise in time but show vertical autocorrelation, which has been modelled as a vertical $\mathrm{AR}(1)$ model:

$$\varepsilon(h) = \rho\varepsilon(h - 150) + \xi(h)$$

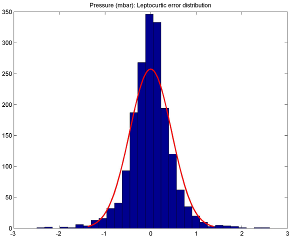
- This autcorrelation structure is similar for μ and σ^2.
  It has been estimated using the empirical *GLS* approach.

Drift estimate

mhu estimation						
variable	beta_GLS	se(beta)		variable	beta_GLS	se(beta)
global component				summer_delta		
height0	-0.56047	0.088835		time0	0.58975	0.15434
lon0	0.18078	0.030264		pressure0	0.23904	0.11063
time0	-0.39836	0.16266		vWind0	0.22472	0.023815
pressure0	-0.55718	0.11791		wspd0	0.084676	0.031538
temperatu	-0.18204	0.077572				
uWind0	-0.19729	0.027014		d_time	0.13997	0.01798
wdir0	0.055457	0.015768		d_vWind	-0.1354	0.037574
datetime0	0.20175	0.028563		d_c_seas	10.215	0.096932
				c_glob	-0.60925	0.084831
d_tempera	0.068282	0.013193				
d_rh_corr	0.026491	0.015512		rho	0.69657	
d_vWind	0.073847	0.034				
d_lon	-0.19945	0.030326				

sigma^2 estimation		
variable	beta_GLS	se(beta)
global component		
pressure0	0.31239	0.03188
uWind0	0.071846	0.017112
vWind0	0.027944	0.015949
d_time	0.1694	0.043564
d_uWind	-0.022168	0.012609
summer_delta		
height0	0.17796	0.035344
d_time	-0.20659	0.047104
d_mr_corr	0.013247	0.01314
c_glob	0.17918	0.016529
rho	0.54796	

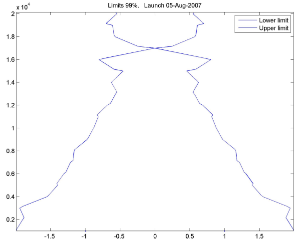
Non Gaussian errors



The estimated errors are described by a rescaled Student's t distribution with 3.63 estimated degrees of freedom:

$$\frac{\hat{\varepsilon} - 0.002}{0.32} \equiv t_{3.63}$$

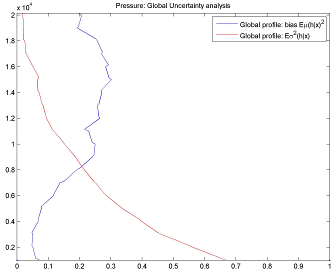
Conditional uncertainty profile

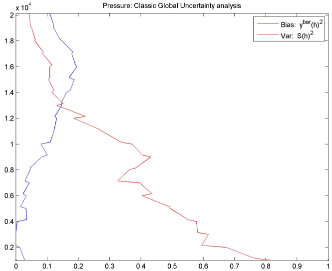


Based on Student's $t_{3.63}$ 99^{th} percentile and

$$\hat{\sigma}^2\left(h|x_{5/8/12}\right) = \hat{\alpha}' x_{5/8/12}\left(h\right)$$

Global uncertainty profile





Total uncertainty table			
	μ^2	σ^2	TOT
Conditional (5 Aug 2007)	-	0.18	-
Global	0.19	0.20	0.39
Classical	0.09	0.31	0.40

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
   1. Minimal technique: use LSE + stepwise or penalty driven model choice
   2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
   3. Consider using alternative models covering for missing data
   4. If big outliers are present use robust methods
   5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
   1. Use the errors of step1
   2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
    1. Minimal technique: use LSE + stepwise or penalty driven model choice
    2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
    3. Consider using alternative models covering for missing data
    4. If big outliers are present use robust methods
    5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
    1. Use the errors of step1
    2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Toolbox scheme

1. Estimate an appropriate drift function $\mu(h|x)$
    1. Minimal technique: use LSE + stepwise or penalty driven model choice
    2. If temporal and/or vertical autocorrelations are present use at least *GLSE* estimate
    3. Consider using alternative models covering for missing data
    4. If big outliers are present use robust methods
    5. The choice of x should include as much information as possible
2. Estimate an appropriate skedastic function $\sigma(h|x)$
    1. Use the errors of step1
    2. Minimal technique: as in 1.1) but Issue 1.2) applies
3. If errors are not Gaussian consider approriate distribution for percentiles, confidence intervals and robust estimation.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Conclusions

At the beginning we asked the following questions:

- is the collocation uncertainty related to measured factors ?
- is the collocation uncertainty related to distance ?
- Is the collocation uncertainty related to height ?
- Are above point valid for all ECV ?
- Is uncertainty a static or dynamic concept ?
- Moreover the scheme of a toolbox for implemeting a basic methodology in similar situations is available.
- Nevertheless, a more general approach as described in the first part is reccommended especially for large datasets.

Thanks for yur attention