UNIVERSITÀ DEGLI STUDI DI BERGAMO

Radiosonde Interpolation Uncertainty

Application to Temperature and Humidity profiles

Virtual meeting ICM-13

November 19, 2021

Author: Pietro Colombo
pietro.colombo@unibg.it
Author: Alessandro Fassò
alessandro.fasso@unibg.it

Contents

Introduction

- **High reliability of the data mandatorily includes**:
  - Uncertainty assessment
    1. Collocation mismatch, e.g. (Fassò et al., 2014)
    2. Uncertainty between satellites and radiosondes data, e.g. (Finazzi et al., 2019)
    3. ...
    4. Interpolation Uncertainty, (Fassò et al., 2020)
- **Why Interpolation uncertainty assessment?**
  - For various reasons, data gaps may appear along the profiles of virtually all radiosonde types, including Vaisala RS41.
  - Solution: Interpolation of missing values
  - Therefore, the interpolation uncertainty must be added to the total uncertainty budget.

Vaisala RS41 Data

Temperature profiles

Site	Country	Profiles
Beltsville	USA	15
Lauder	NZ	32
Lindenberg	DF	45
Ny-Alesund	DE/FR	35
Payerne	CH	30
Lamont	USA	16
Sodankyl	FI	4
Total		177

- 1 sec resolution.
- Gaps < 5 sec
- Profiles' lengths 3500-6500 (approx)
- 2014-2017

Humidity profiles

Site	Country	Profiles
Beltsville	USA	15
Lauder	NZ	32
Lindenberg	DF	45
Ny-Alesund	DE/FR	35
Payerne	CH	30
Lamont	USA	16
Sodankyl	FI	1
Total		174

- 1 sec resolution.
- Gaps < 2 sec (1 exception).
- Profiles' lengths 3500-6500 (approx).
- 2014-2017.

Strategy

BB-CV

Block-bootstrap Cross-validation Scheme allows generating a random number of gaps of different lengths across each fully observed profile

Interpolation of the missing values

Interpolation of the simulated missings using:

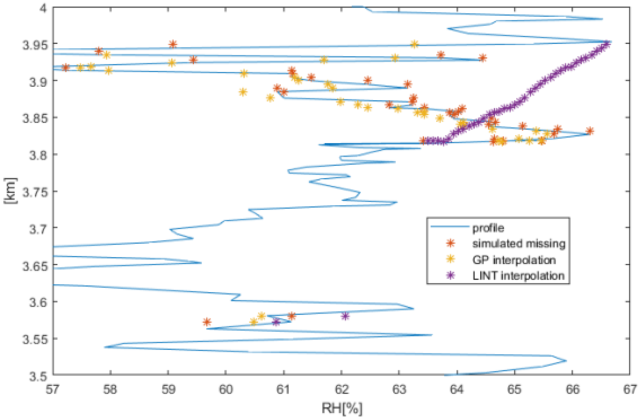
- Gaussian process
- Linear Interpolation

Uncertainty assessment

The uncertainty is evaluated in relevant dimensions such as Altitude, Interpolation distance, Launch site.

Example

Detail of RS41 humidity profile at Lamont site on 2014-06-06, near 3.5 km altitude.



Block Bootstrap Cross-Validation

Each fully observed profile $Y = (y(1), ..., y(T))$ is taken and partitioned as follows: $[Y^L, Y^*]$

- Y^L: Learning set
- Y^*: Testing set

The Y^* values are chosen as follows:

1. For each profile, $n_G = T\frac{f}{\mu_g}$ gap sequences are generated.
   - f Sampling fraction
   - μ_g Average gap size
   - T Profile length
2. The extraction process is repeated B times.
3. The procedure is replicated for $\mu_g = [4, 10, 30, 60]$.

Interpolators

Temperature Interpolators

- $m_i(t^*)$ Linear interpolator
- $m_i(t^*)$ Gaussian process interpolator
  - Squared Exponential covariance function

Humidity Interpolators

- Linear interpolator
- Gaussian process interpolator
  - Exponential covariance function
- Multidimentional interpolator $m_i(x1, x2, x3)$

Model	Predictors
Model 1	Time
Model 2	Alt
Model 3	Time,Temp
Model 4	Alt, Temp
Model 5	Lat, Lon, Alt
Model 6	Time, Alt, Temp

Uncertainty

Temperature Uncertainty

- **Gaussian Process based approach**
  - U_{GP} is based on a local Gaussian Process approximation of the specific profile.
  - It might underestimate the IU
  - It considers the individual profile autocorrelation structure.
- **Bootstrap correction approach (I.A.)**
  - $MSE_B = mean(e^2)$
  - $U_B = MSE_B - avg(U_{GP}^2)$
  - $IU_{Total}^2 = U_{GP}^2 + U_B^2$

Humidity Uncertainty

- **Data approach (BB-CV)**
  - $e = \hat{y}(t^*|s, l, d, Alt, B) - y(t^*|s, l, d, Alt)$
  - $MSE_B = mean(e^2)$
  - $MSE_B = SE^2 + bias^2$.
  - This approach can be made operational as lookup tables
  - The individual profile contribution to the uncertainty is not considered

Results

Temperature results

		$\mu_G = 4''$		$\mu_G = 10''$		$\mu_G = 30''$		$\mu_G = 60''$	
Site	Profiles	GP	Linear	GP	Linear	GP	Linear	GP	Linear
BEL	15	0.084	0.088	0.159	0.160	0.338	0.363	0.590	0.604
LAU	32	0.106	0.107	0.180	0.184	0.370	0.389	0.599	0.612
LIN	45	0.073	0.074	0.145	0.145	0.314	0.324	0.548	0.542
NYA	35	0.072	0.073	0.127	0.130	0.269	0.269	0.463	0.460
PAY	30	0.098	0.098	0.180	0.181	0.370	0.391	0.659	0.658
SGP	16	0.107	0.109	0.189	0.187	0.401	0.420	0.703	0.698
SOD	4	0.074	0.076	0.137	0.138	0.281	0.363	0.426	0.478
	177	0.087	0.088	0.159	0.160	0.334	0.349	0.574	0.576

Humidity results

		Average gap size				
Model	Interpolator	4s	10s	30s	60s	Mean
Model 1	GP	0.559	1.126	2.243	3.268	1.799
	LINT*	0.556	1.122	2.237	3.260	1.794*
Model 2	GP	0.563	1.127	2.244	3.268	1.801
	LINT	0.559	1.122	2.237	3.260	1.795
Model 3	GP	0.551	1.070	2.125	3.109	1.714
	LINT	0.975	1.580	3.021	4.090	2.416
Model 4	GP	0.553	1.069	2.124	3.097	1.711
	LINT	0.792	1.385	2.653	3.742	2.143
Model 5	GP	0.583	1.149	2.294	3.377	1.851
	LINT	1.075	1.935	7.643	19.227	7.470
Model 6	GP*	0.531	1.038	2.078	3.047	1.674*
	LINT	1.182	2.326	6.093	7.863	4.366

Uncertainty

Temperature results

- Larger IU Lamont Payern and Lauder
- IU: 0.1-0.6 K
- GP and LINT equivalent.

Humidity results

- Larger IU Lamont, Payerne, Beltsville
- IU: $0.3 - 8\% RH$ units .
- GP lower IU: ($0.2\% RH$ units or $6\% RMSE$ units).
- For large gaps at lower altitudes $1 - 1.5\%$.

Humidity Graphs

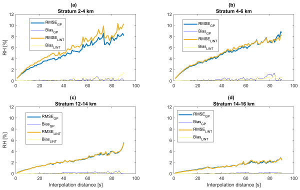
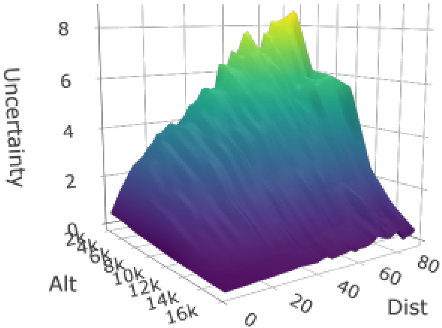


Figure 5.1: The figure depicts the interpolation uncertainty of GP and LINT for different atmosphere strata.

Uncertainty Surface and Lookup tables



References I

Fassò, A., Ignaccolo, R., Madonna, F., Demoz, B. B., and Franco-Villoria, M. (2014). Statistical modelling of collocation uncertainty in atmospheric thermodynamic profiles. *Atmospheric Measurement Techniques*, 7(6):1803–1816.

Fassò, A., Sommer, M., and von Rohden, C. (2020). Interpolation uncertainty of atmospheric temperature profiles. *Atmospheric Measurement Techniques*, 13(12):6445–6458.

Finazzi, F., Fassò, A., Madonna, F., Negri, I., Sun, B., and Rosoldi, M. (2019). Statistical harmonization and uncertainty assessment in the comparison of satellite and radiosonde climate variables. *Environmetrics*, 30(2):e2528.

Thank You.