CASPA 2019

Session 2 : Capteurs - Données

METHODICAL EVALUATION OF IN SITU CALIBRATION STRATEGIES FOR ENVIRONMENTAL SENSOR NETWORKS

FLORENTIN DELAINE

Methodical Evaluation of In Situ Calibration Strategies for Environmental Sensor Networks

Florentin Delaine [1, 2, 3], Bérengère Lebental [1, 2, 3], Hervé Rivano [4]

[1] Efficacity, France

[2] Université Paris-Est IFSTTAR, COSYS/LISIS, Marne-la-Vallée, France

[3] LPICM, CNRS, École Polytechnique, Université Paris-Saclay, F-91128 Palaiseau, France[1]

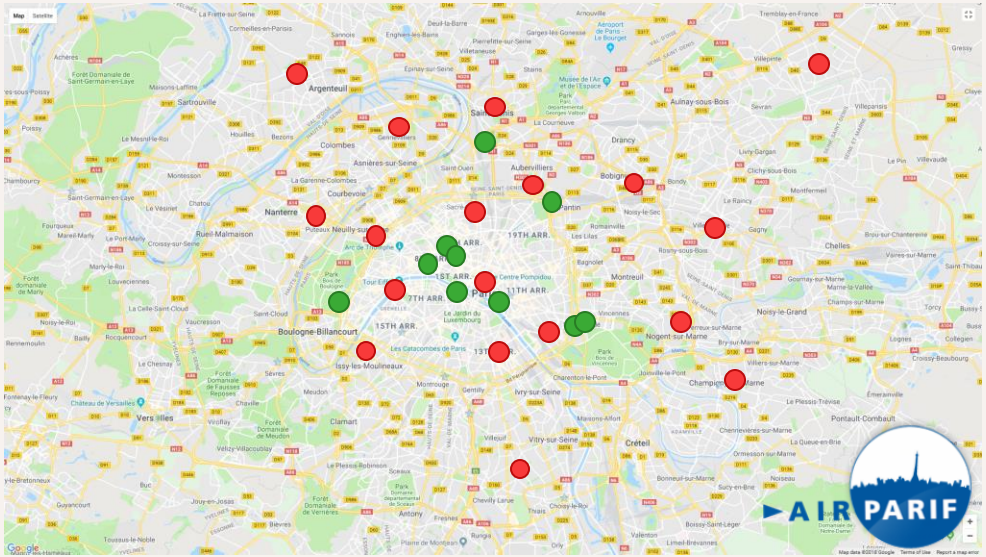
[4] Université de Lyon, INSA Lyon, Inria, CITI, F-69621 Villeurbanne, France

Colloque National Capteurs et Sciences Participatives

1-4 avril 2019, Paris

Reference measuring instruments cannot easily ensure monitoring at a high spatial scale

Map of air pollution monitoring stations of Airparif for Paris and its inner suburbs (December 2018)



Permanent urban station Permanent traffic station

AC32M (Environnement SA)

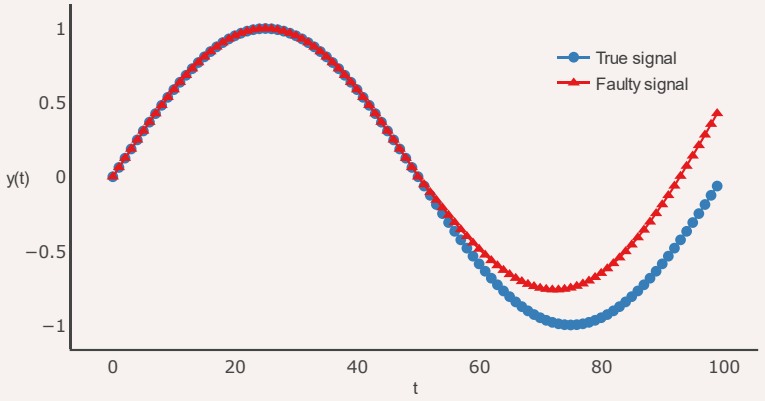
Example of reference system for NOx

Cumbersome (3U rack)
Heavy (13 kg)
Expensive (> 10 000 €)

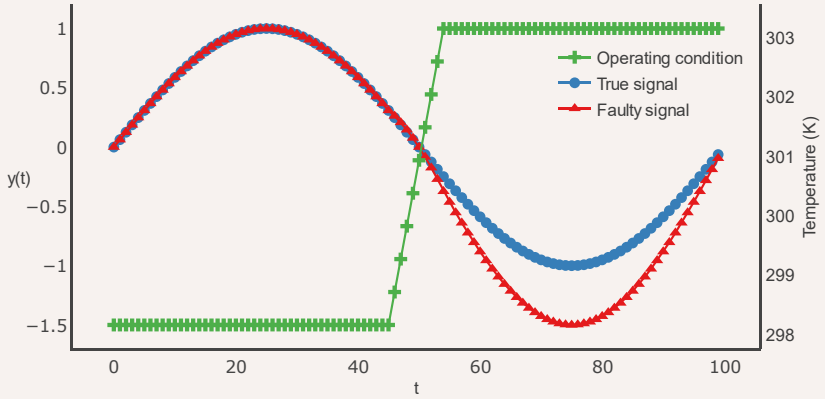
Low-cost sensors becomes interesting to allow deployments at a reasonable cost [Kumar 2015]

Performances of low cost sensors are subject to drifts…

… due to aging



… or influenced by operating conditions



(True) quantity value

Measuring instrument

Sensor
(Sensitive element, transducer…)

Processing(s)

…

Indication(s)

Quantity conversion

Measuring chain

Measured value

Other values

Indications

Standards

Calibration

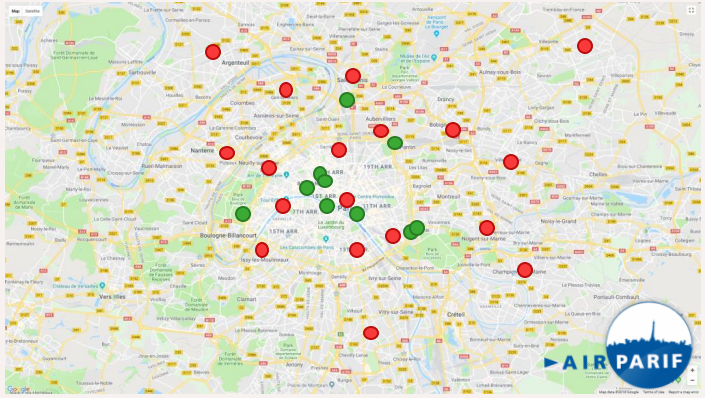
Calibration relationship

Recalibrating each instrument in laboratory is expensive

Recalibration procedure

1. Take it out of service
2. Bring it to a calibration facility
3. Calibrate
4. Bring it back to the field
5. Put it in service

The more instruments are deployed, the more calibration is challenging



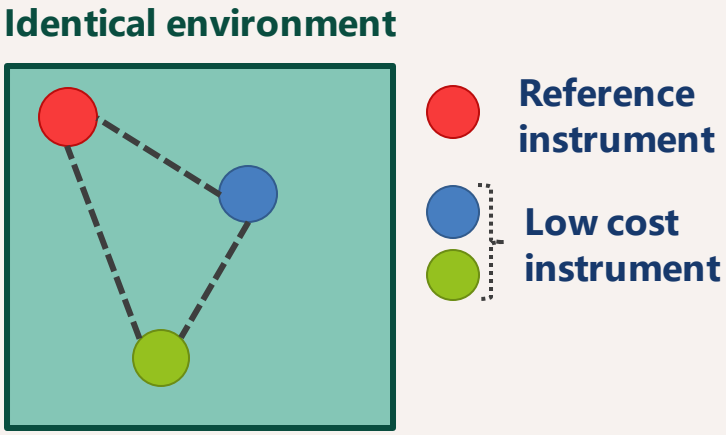
Permanent urban station

Permanent traffic station

The behaviour of instruments may also be different under uncontrolled conditions [Ramanathan 2006]

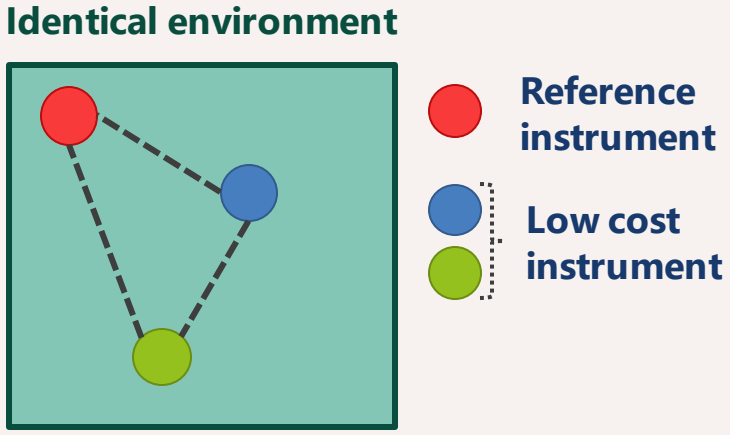
Solution

- **Idea:** perform an *in situ* calibration
- **Definition:** Calibration of measuring instruments while leaving them in the field, preferably without any physical intervention
- **Synonyms:** field, in place, remote, online, self calibration



The use of measurement results from multiple instruments may be a solution

- **Idea:** perform an *in situ* calibration
- **Necessary elements for calibration [BIPM 2012]:**
  - **Availability of standard values**
  - **Known operating conditions**
  - **Ability to quantify the resulting uncertainty**



Classification of *in situ* calibration strategies [Delaine 2019 (accepted, IEEE Sensors)]

- Use of reference instruments
- Mobility of the instruments
- Grouping strategies
- Calibration relationships

Similar classification and observations made in [Barcelo-Ordinas 2019]

Problem statement

Publication	P_a	P_b	P_c
Algorithm	A_a	A_b	A_c
Case study	C_a	C_b	C_c
Conclusion	A_a works on C_a	$A_b > A_a$ on C_b	$A_c > A_b$ on C_c
Not proved	-	$A_b > A_a$ on C_a	$A_c > A_b$ on $C_{a,b}$ $A_c > A_a$ on $C_{a,b,c}$

Reimplementations or codes available

Simulations or experiments

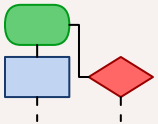
No commonly accepted case study like in ML [Dua 2017]

Phenomena to monitor, sensor network…

Major limitation to be able to compare strategies

Problem statement

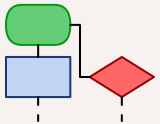
- **Challenge**: Be able to compare performances of



One strategy

One network

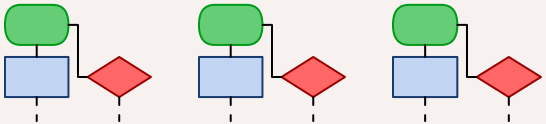
Multiple environments



One strategy

Multiple networks

One environment



Multiple strategies

One network

One environment

By which mean should comparison be performed?

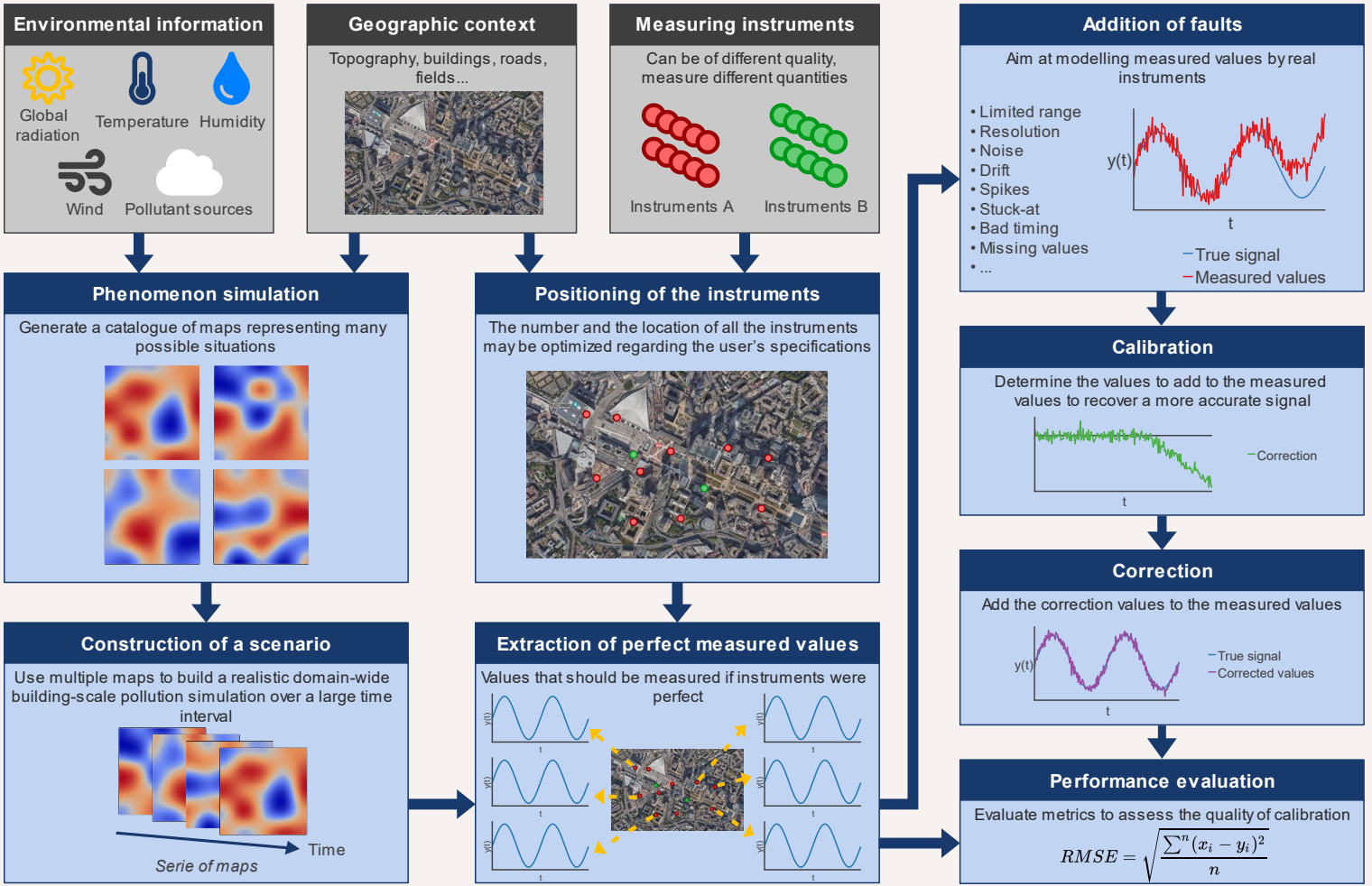
Experiment

- **Pros:**
  - Real conditions
  - Real instruments
- **Cons:**
  - Instrument dependent
  - Measurand dependent
  - Necessity of high quality instruments

Simulation

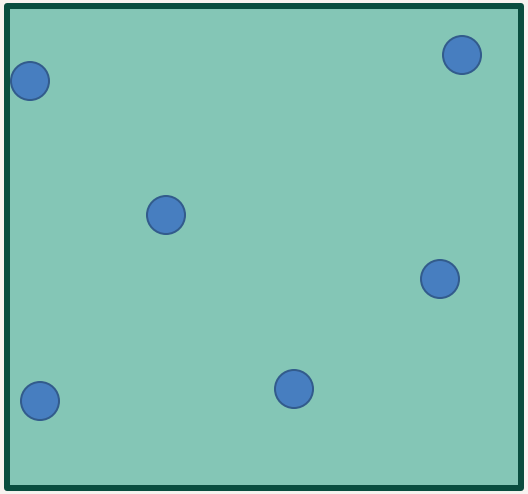
- **Pros:**
  - Any measurand possible
  - Unlimited number of instruments
  - Accessible “true values”
- **Cons:**
  - Results as good as the simulation is

Proposed methodology

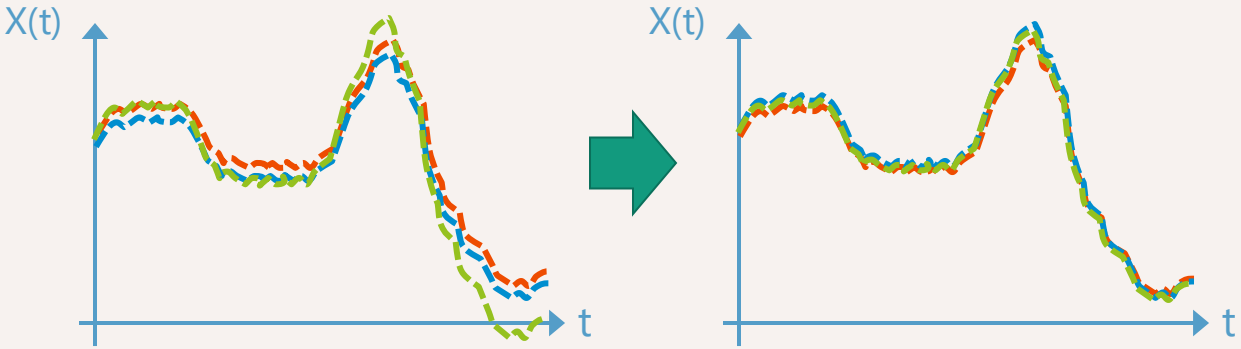


Case study

- **Goal:** Compare two existing strategies: [Balzano 2007] and [Wang 2015]
- Both based on the subspace model



Consider N measured values for each of M instruments, measuring a quantity with a high spatial density



Values are corrected supposing the true ones lies in a lower sub-space of the space formed by the measured values

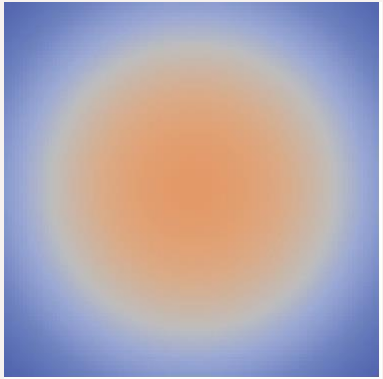
Modelling of the phenomenon

- **Goal: one year, hourly time step**
- Environment considered: flat field 1000 x 1000m, discretized with a step of 10m
- Concentration of pollutant modelled as

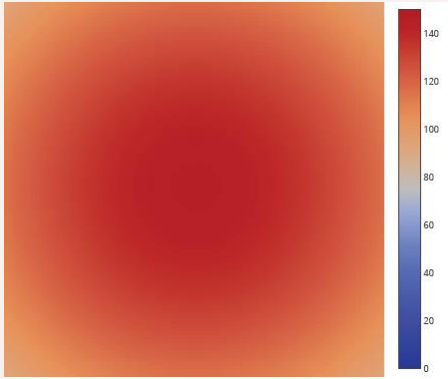
$$C(x,y,t) = A(t)\exp\left(-\frac{x^2+y^2}{\sigma(t)^2}\right)$$

Standard deviation expressed in terms of full width at half maximum (FWHM)

$$\sigma = \frac{FWHM}{2\ln(2)}$$



A = 100, FWHM = 1000

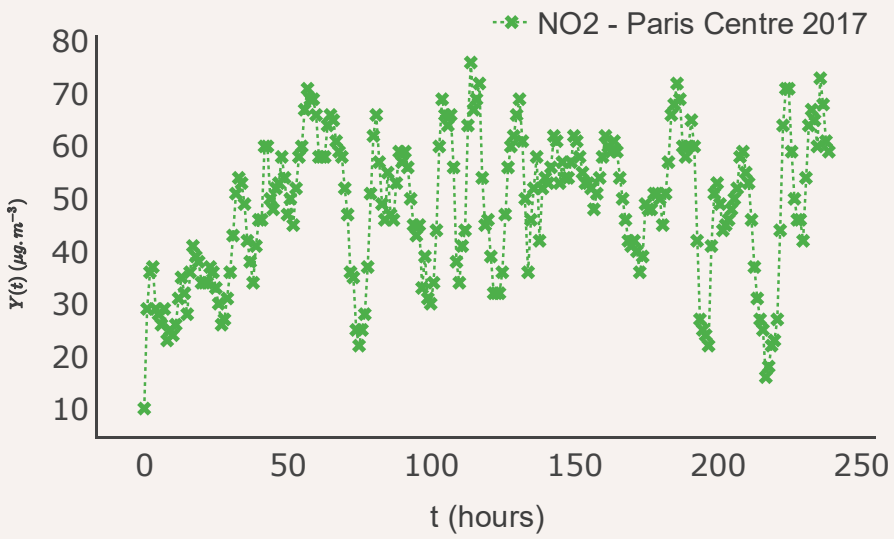


A = 145, FWHM = 1500

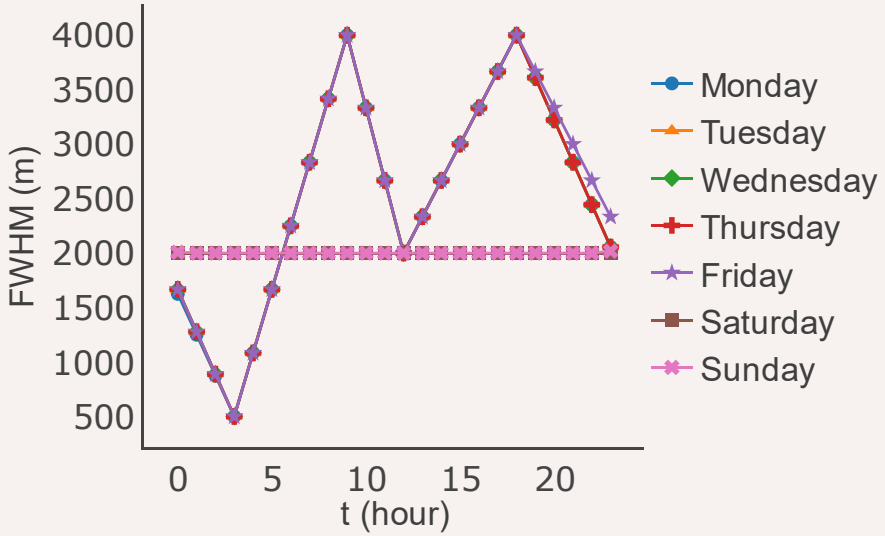
- Number of maps simulated defined arbitrary
  - Amplitude: from 0 to 150 µg/m3, step 5 µg/m3
  - FWHM: from 0 to 4000 m, step 500 m

Modelling of the phenomenon

- **Amplitude modelling**: based on NO2 measured values of Airparif



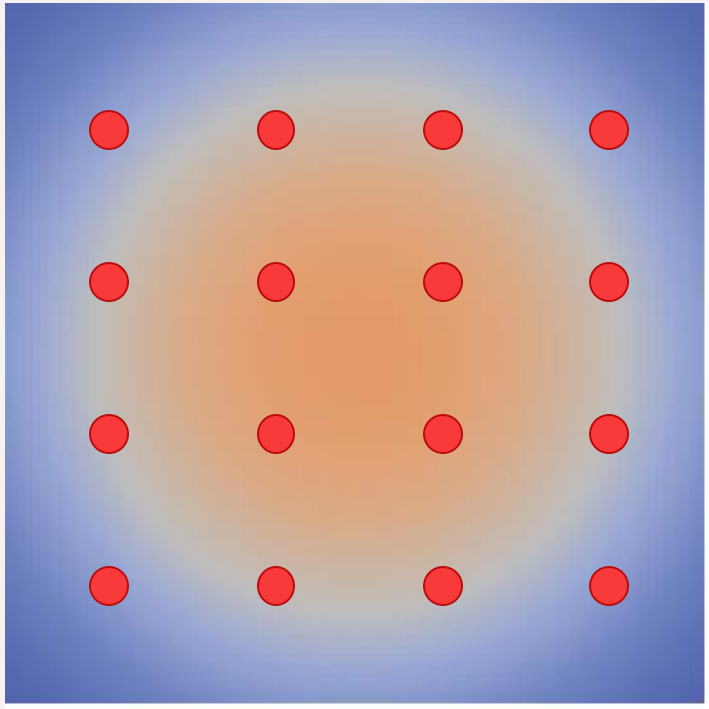
- **Full width at half maximum:** profiles for week-end and week days



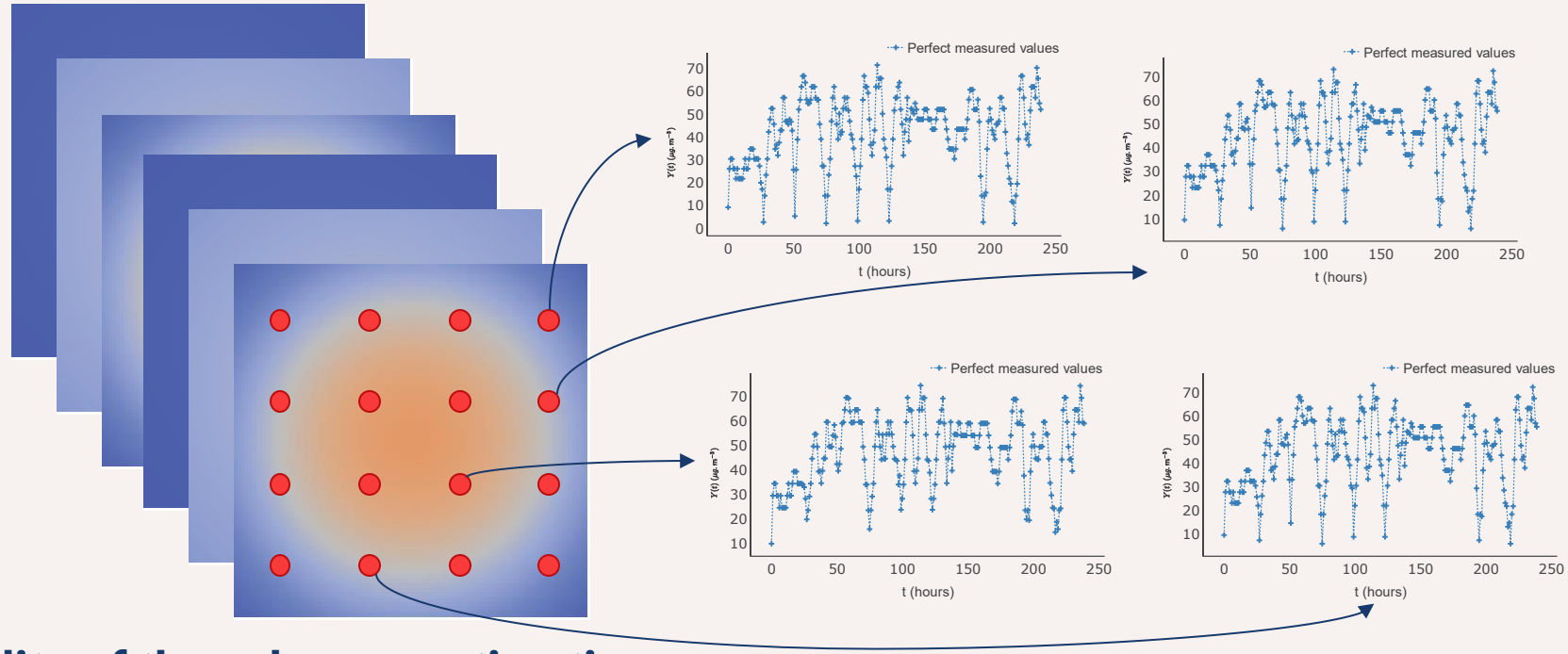
- **Result:** A time series of 8760 maps (one year, hourly time step)

Network definition

- n = 16 nodes, uniformly deployed in the environment
- Static
- Blind



Extraction of perfect measured values



Validity of the subspace estimation

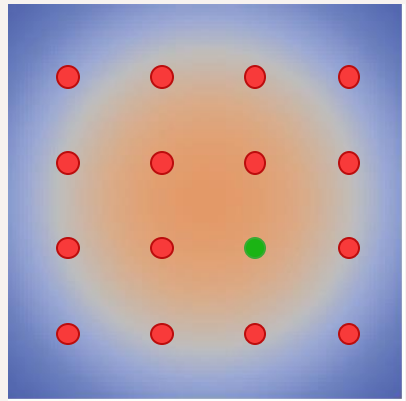
- Principal component analysis on the matrix formed by the concatenation of the 16 time series obtained
- **Result: r = 4 < n**

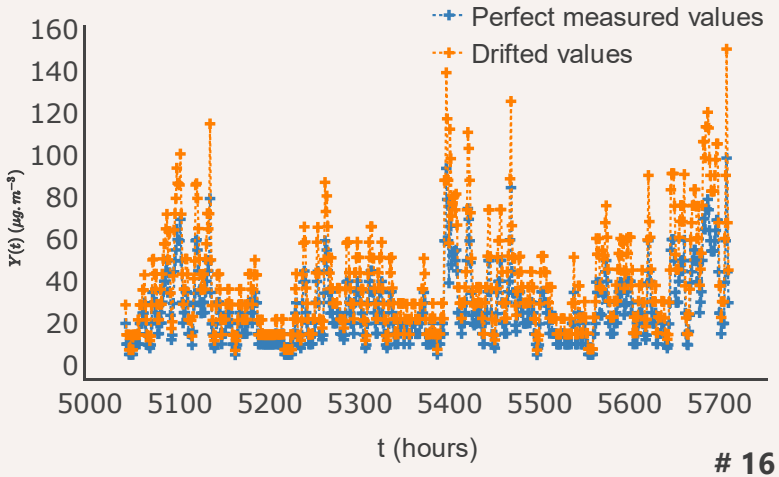
Fault addition

- Sensors assumed as initially calibrated and faultless for 8 eight weeks
- One sensor drifting
- Linear drift of its gain, up to 5% weekly

$$G_0 = 1 + 0.05 d_0$$
$$G_i = G_{i-1} + 0.05 d_i$$
$$d_i \sim \mathcal{U}(0, 1)$$

- Start at the first week of drift (week 12)

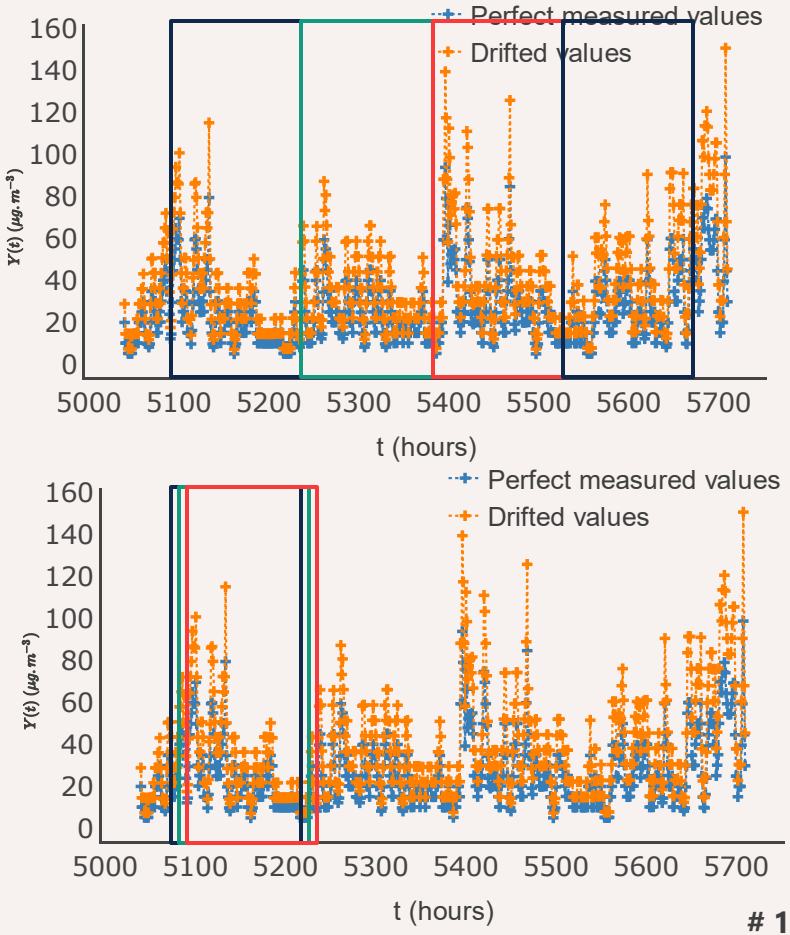




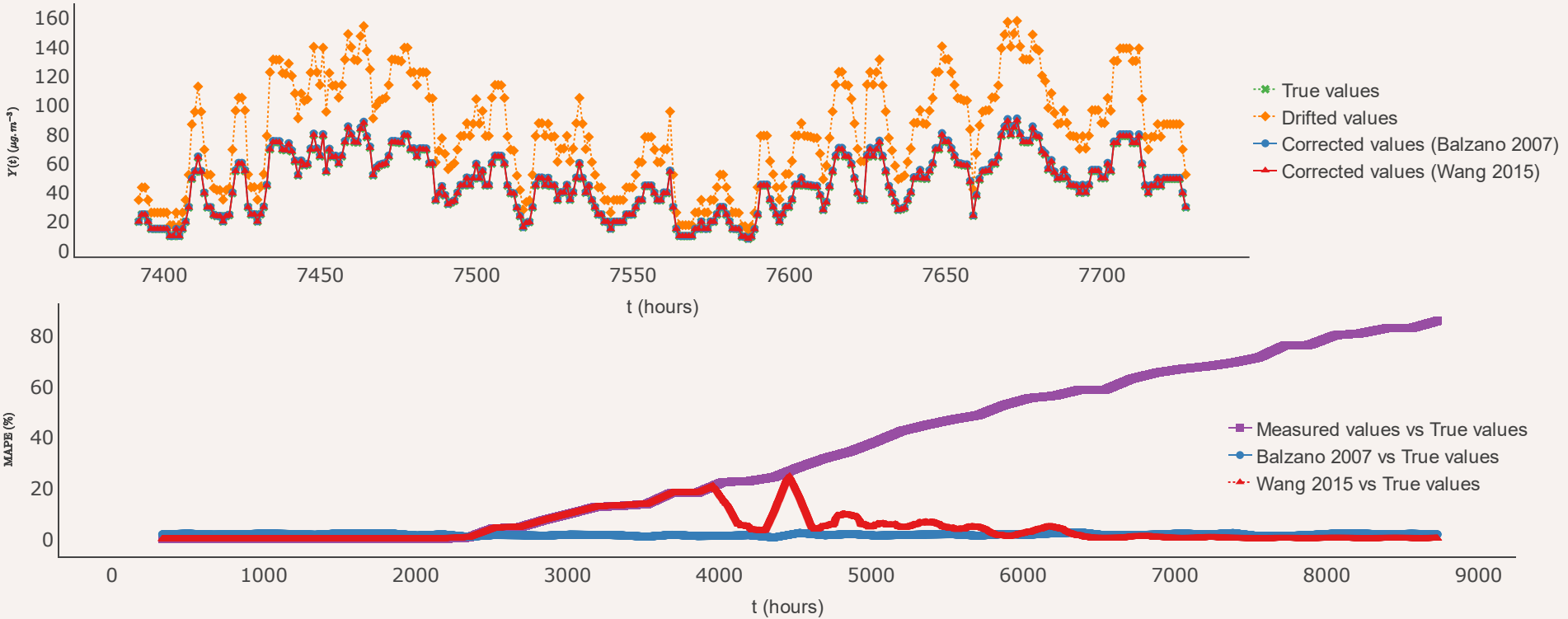
Parametrization of calibration strategies

- Balzano et al.
  - Correction: gain
  - Principle: linear system solving
  - Strategy applied each w = 7 days

- Wang et al.
  - Correction: variable offset
  - Principle: Kalman filter
  - Applied at each time step (1 hour)
  - Starts at t = w = 7 days = 168 hours



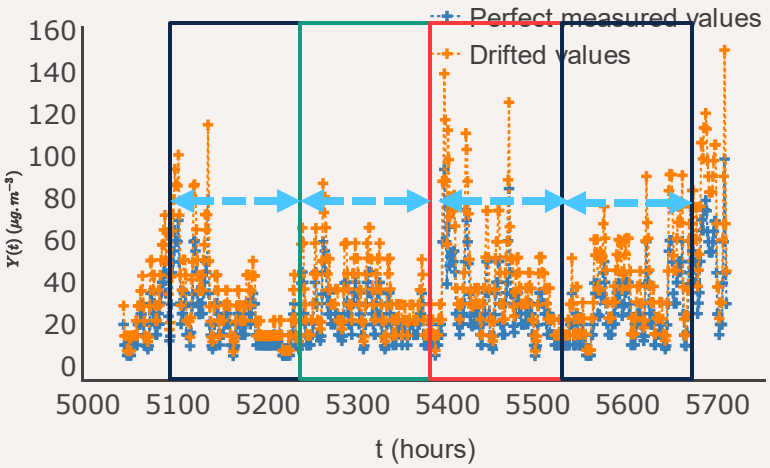
Results



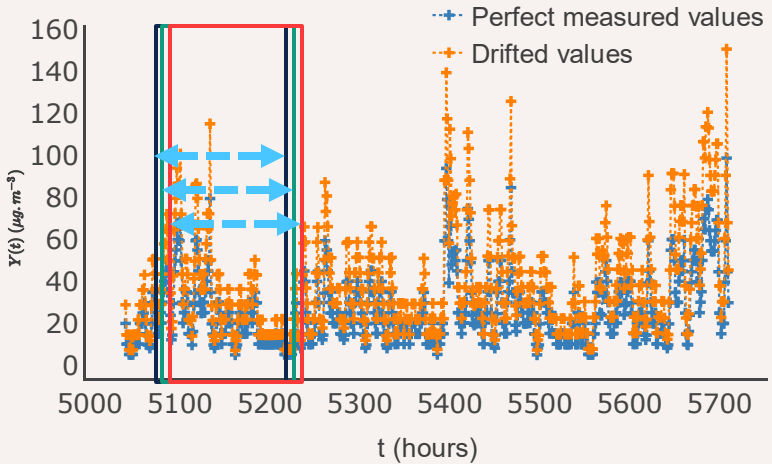
Metric	**[Balzano 2007]**	**[Wang 2015]**
Mean absolute percentage error (rolling, 7 days)	Mean: 1.7% Max:16.8%	Mean: 3.0% Max: 19.5%

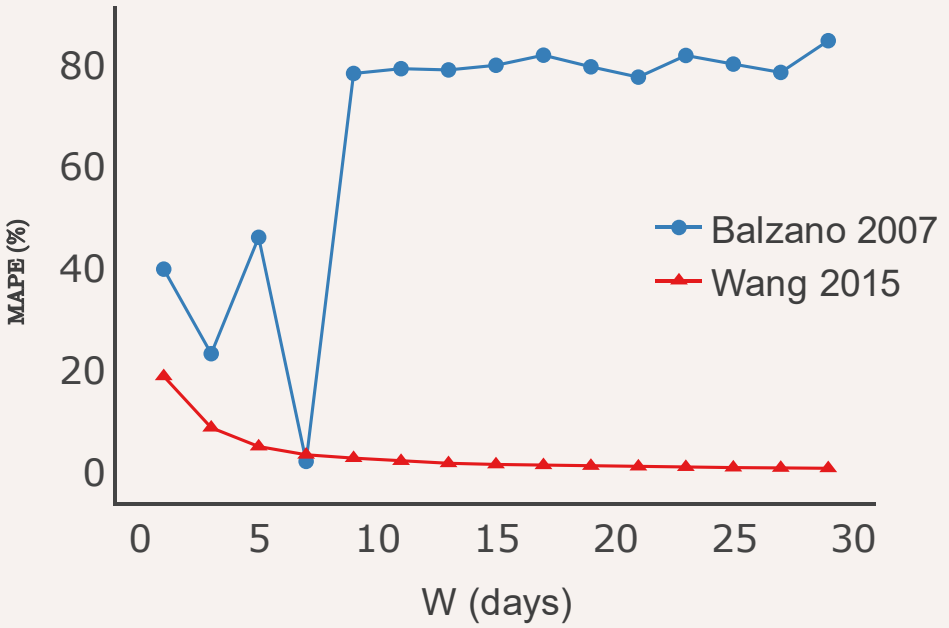
With varying window width

[Balzano 2007]



[Wang 2015]



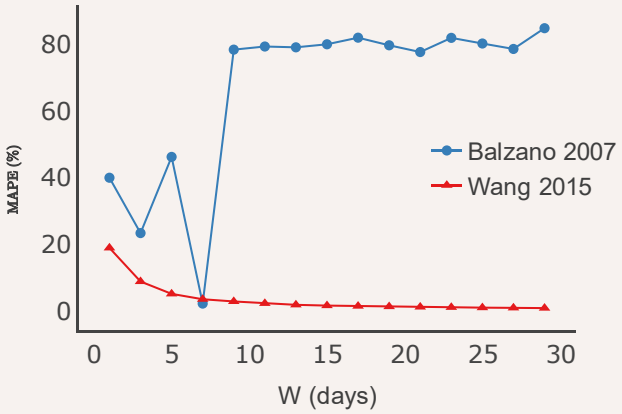


Conclusions of the case study

- A unique case study may not be conclusive – results may be case specific

Publication	P_a	P_b	P_c
Algorithm	A_a	A_b	A_c
Case study	C_a	C_b	C_c
Conclusion	A_a works on C_a	$A_b > A_a$ on C_b	$A_c > A_b$ on C_c
Not proved	-	$A_b > A_a$ on C_a	$A_c > A_b$ on $C_{a,b}$ $A_c > A_a$ on $C_{a,b,c}$

- The parametrization of the existing methods should be investigated



Summary

- A protocol for a methodical **evaluation of in situ calibration strategies** was introduced

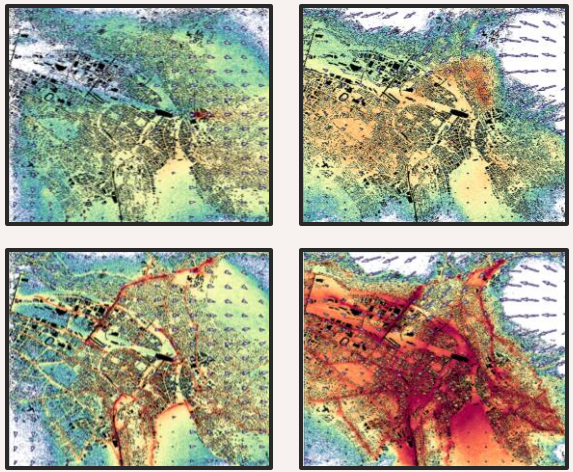
- A **simple case study** for the comparison of two existing strategies was developed

- First results show an influence of the design of the case studies on performance results. It justifies the necessity of **extended studies** toward the **definition of systematic procedures** for the evaluation calibration strategies

Future work

Quantities simulation

- Realistic quantities
- Real geometry
- Multiple quantities

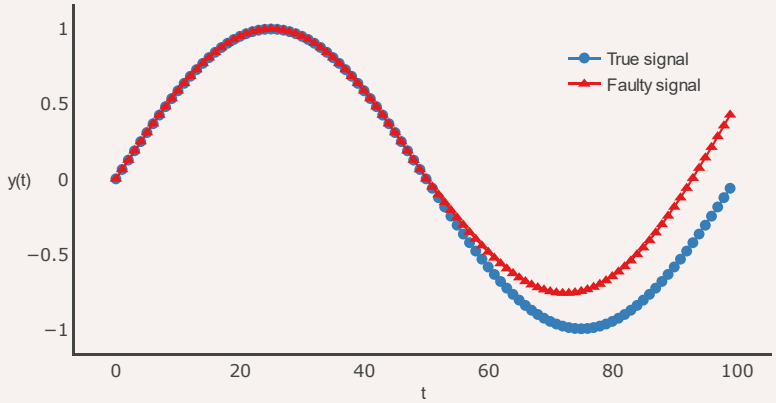


Catalogue based simulation [Berchet 2017]

Sensor network definition

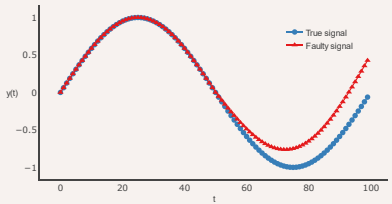
- Mobility
- Optimization of positioning [Boubrima 2017]

Fault addition

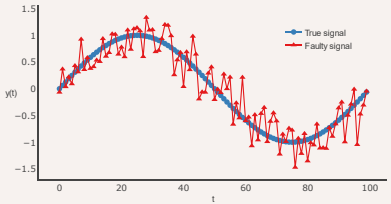


Drift

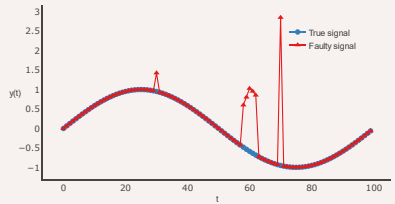
Fault addition



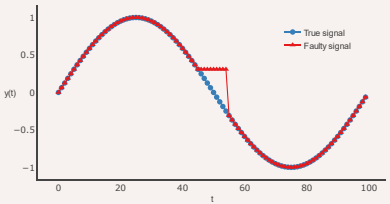
Drift



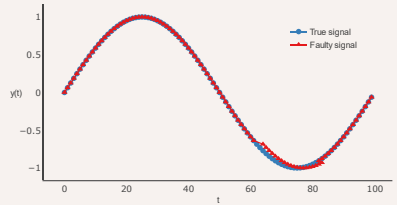
Noise



Spike

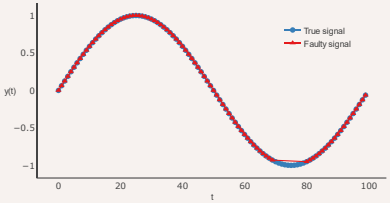


Stuck-at

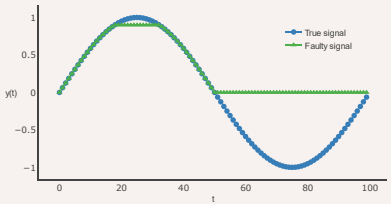


Bad timing

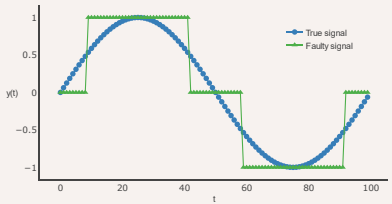
Fault taxonomy (based on [Ni 2009])



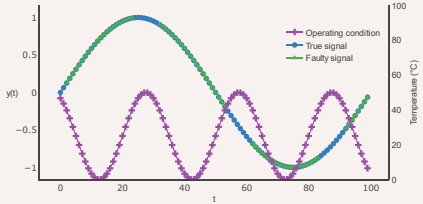
Missing values



Range

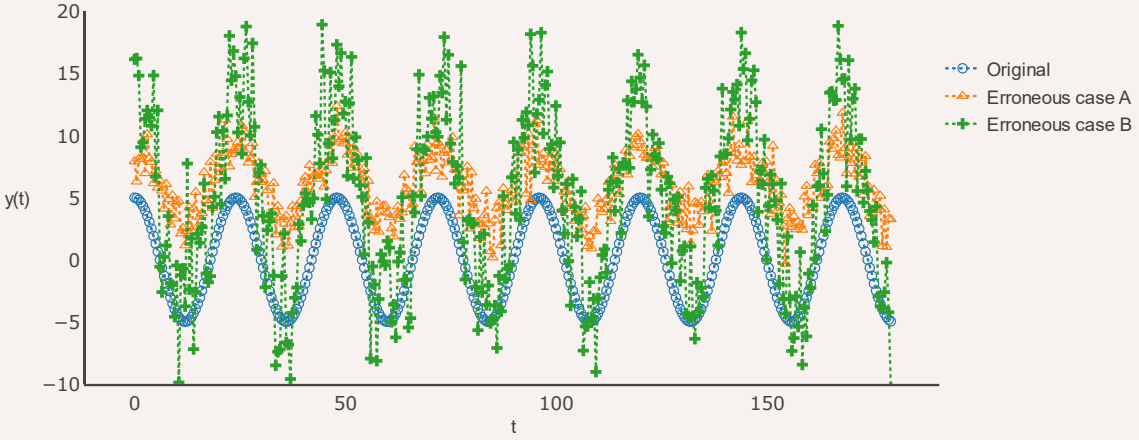


Resolution

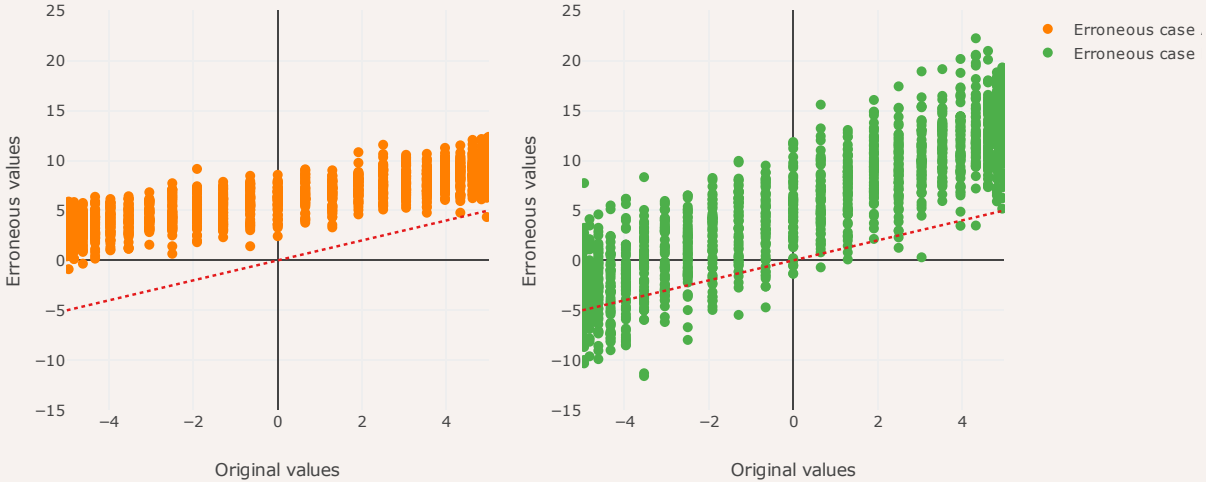


Op. cond. out-of-range

Performance evaluation



Metric	A	B
Bias	6.1	5.1
MSE	40.7	43.5
Lin. corr.	0.85	0.88
Coef. det.	0.72	0.77
MSE SS	-2.26	-2.46



Interdependence
Underdetermination
Incompleteness
Linear error assumption
[Tian 2016]

References

[Airparif] Airparif. Airparif website. https://www.airparif. asso.fr

[Kumar 2015] P. Kumar et al.,"The rise of low-cost sensing for managing air pollution in cities," Environment International, vol. 75, pp. 199–205, 2015

[Ramanathan 2006] N. Ramanathan et al., "Rapid Deployment with Confidence: Calibration and Fault Detection in Environmental Sensor Networks", UCLA, Tech. Report, 2006

[Delaine 2019] F. Delaine, B. Lebental and H. Rivano, "In Situ Calibration Algorithms for Environmental Sensor Networks: a Review", IEEE Sensors, 2019 [Accepted Paper]

[Barcelo-Ordinas 2019] J. M. Barcelo-Ordinas et al., "Self-Calibration Methods for Uncontrolled Environments in Sensor Networks: A Reference Survey," Ad Hoc Networks, vol. 88, pp. 142–159, 2019

[BIPM 2012] BIPM, International vocabulary of metrology - Basic and general concepts and associated terms (VIM), 3rd ed. JCGM 200:2012, 2012

[Dua 2017] D. Dua and K. T. Efi, "UCI Machine Learning Repository", University of California, Irvine, School of Information and Computer Science

[Balzano 2007] L. Balzano and R. Nowak, "Blind Calibration of Sensor Networks," 2007 6th International Symposium on Information Processing in Sensor Networks, pp. 79–88, 2007

[Wang 2015] [15] Y. Wang, A. Yang, Z. Li, P. Wang, and H. Yang, "Blind drift calibration of sensor networks using signal space projection and Kalman filter," in 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015, pp. 1–6.

[Berchet 2017] A. Berchet et al., "A cost-effective method for simulating city-wide air flow and pollutant dispersion at building resolving scale," Atmospheric Environment, vol. 158, pp. 181–196, 2017.

[Boubrima 2017] A. Boubrima, W. Bechkit, and H. Rivano, "Optimal WSN Deployment Models for Air Pollution Monitoring," IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 2723–2735, 2017

[Ni 2009] K. Ni et al., "Sensor network data fault types," ACM Transactions on Sensor Networks, vol. 5, no. 3, pp. 1–29, 2009

[Tian 2016] Y. Tian et al., "Performance Metrics, Error Modeling, and Unvertainty Quantification," Monthly Weather Review, vol. 144, no. 2, pp. 607–613, 2016

IFSTTAR efficacity Inria

Questions?

université PARIS-SACLAY ÉCOLE POLYTECHNIQUE LPICM cnrs