European Network on New Sensing Technologies for Air Pollution Control and Environmental Sustainability - *EuNetAir*

COST Action TD1105

WGs Meeting, Belgrade, 13 - 14 October 2015

organized by VINCA Institute and co-organized by Public Health Institute of Belgrade

hosted by Faculty of Mechanical Engineering, University of Belgrade

Action Start date: 16/05/2012 - Action End date: 30/04/2016

Year 4: 1 July 2015 - 30 April 2016 (*Ongoing Action*)

COMPARATIVE STATISTICAL ANALYSIS OF MULTIVARIATE LINEAR REGRESSION MODELS FOR CALIBRATION OF LOW COST SENSORS FOR AIR POLLUTION MONITORING

Dušan Topalović, Miloš Davidović, Milena Jovašević-Stojanović

Vinča Institute of Nuclear Sciences, University of Belgrade, Serbia

dusan.topalovic@vin.bg.ac.rs

COST EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

European Network on New Sensing Technologies for Air-Pollution Control and Environmental Sustainability

COST is supported by the EU Framework Programme

ESF provides the COST Office through a European Commission contract

Outline:

1) INTRODUCTION
2) DNET EB700 SYSTEM FOR AIR POLLUTION MONITORING
3) RESULTS
4) MULTIVARIATE LINEAR REGRESSION MODELS FOR CALIBRATION OF LOW-COST SENSORS FOR AIR-POLLUTION MONITORING
5) CONCLUSION

INTRODUCTION

- 50% of people live in cities
- The UN estimates-in 2050. 70% of the worlds population will live in cities
- Accommodation of a growing trend-"smart city" and ICT sector
- Citizens expectations:
  - Better quality of life
  - Detailed information about the citys environmental conditions
- Cities occupy 3% of the worlds geography-80% of CO_2 emission

INTRODUCTION

- Maintenance of environmental quality-"green cities "
- Attracting tourists, investors, business and to provide additional information about the environmental conditions
- Optimization of industrial production:
  - Remote monitoring of work processes
  - Monitoring hazardous gases
  - Personal exposure to pollutants
  - Eco-friendly solutions
- Trend of lowering tolerance to air pollution
- Improving of monitoring services and putting in place polices-not following the regulations impose adequate penalties

INTRODUCTION

- Monitoring of air pollution- networks of static measurement stations under the control of public authorities
- Fixed Stations:
  - Reliable
  - Large
  - Expensive, with large cash costs
  - Limited number of installations
- The concentration of pollutants is spatially and temporally dependent

New possibilities for air pollution monitoring

- Solid-state gas sensors have started to be used for measuring the pollutants in the atmosphere
- Electrochemical reactions when exposed to a specific gas
- Measuring the output voltage or resistance of the tin oxide layer
- They are typically used in the industry
- *Advantages*:
  - Small
  - Cheap
  - Provides higher spatial resolution
- *Disadvantages*:
  - Limited measurement accuracy
  - Instability
  - Poor selectivity
  - ***Sensitivity to weather conditions*** and other gases

DNET EB700 SYSTEM FOR AIR POLLUTION MONITORING

- The DNET system is designed in such a way to provide a complete end-to-end solution for the environmental monitoring
- Low cost sensors for meteorological parameters (T,rH,p) and gases have been produced by Alphasense (UK)
- Particulate matter monitor (less then PM2.5 and PM10) has been produced by DYLOS (USA)
- Testing system performances and calibration in the field performing by Vinča institute (Serbia) in cooperation with SEPA (Serbian Environmental Protection Agency)
- Activities of system validation and data presentation and analysing has been performed in cooperation of Institute Vinča, QUT and DunavNet

DNET EB700 SYSTEM FOR AIR POLLUTION MONITORING

- The DNET system is designed in such a way to provide a complete end-to-end solution for the environmental monitoring
- Low cost sensors for meteorological parameters (T,rH,p) and gases have been produced by Alphasense (UK)
- Particulate matter monitor (less then PM2.5 and PM10) has been produced by DYLOS (USA)
- Testing system performances and calibration in the field performing by Vinča institute (Serbia) in cooperation with SEPA (Serbian Environmental Protection Agency)
- Activities of system validation and data presentation and analysing has been performed in cooperation of Institute Vinča, QUT and DunavNet

Pilot outdoor measurements in the municipalities of New Belgrade

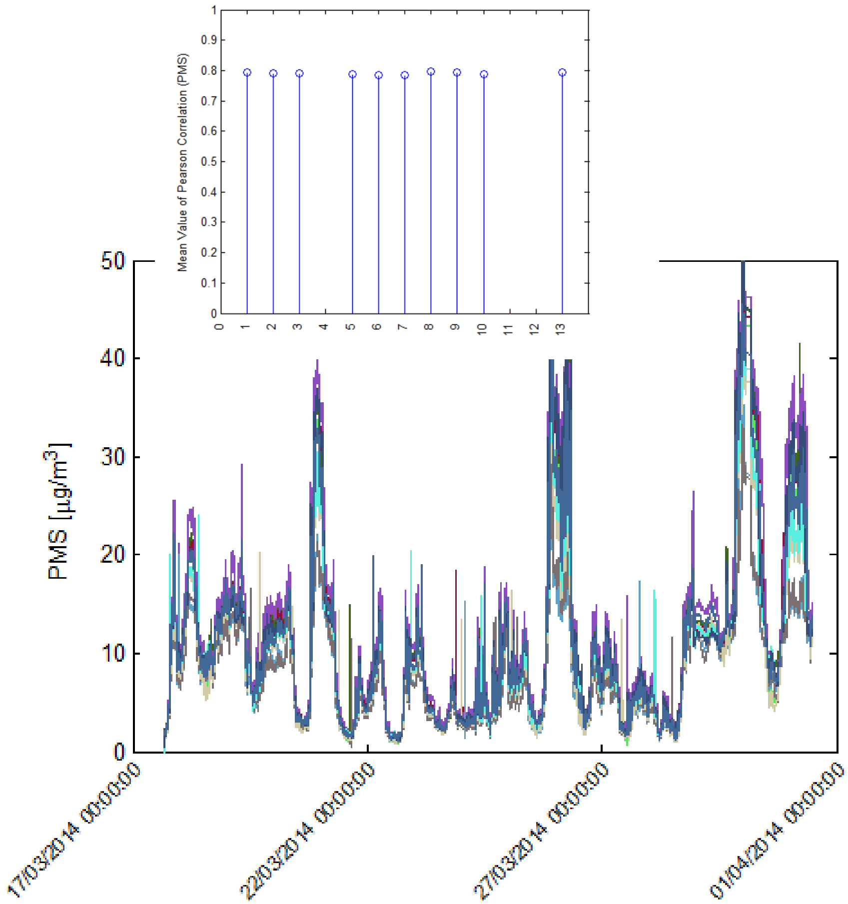
- Based on experience from studies that have used cheap sensors for air pollution monitoring, it was recommended that when low-cost sensors are used in real time measurements it is necessary to perform an evaluation of their characteristics through a calibration procedure
- Pods were deployed within the SEPA stations for access to reference measurements (**calibration**)
- 13 pods at the same location, close to each other
- Correlations between the same sensor from different platforms
- Correlation between the platforms and the reference monitor (important for calibration)
- Time synchronization of the sensors and reference signal by interpolation
- Filtering of the signal with the aim of noise reduction

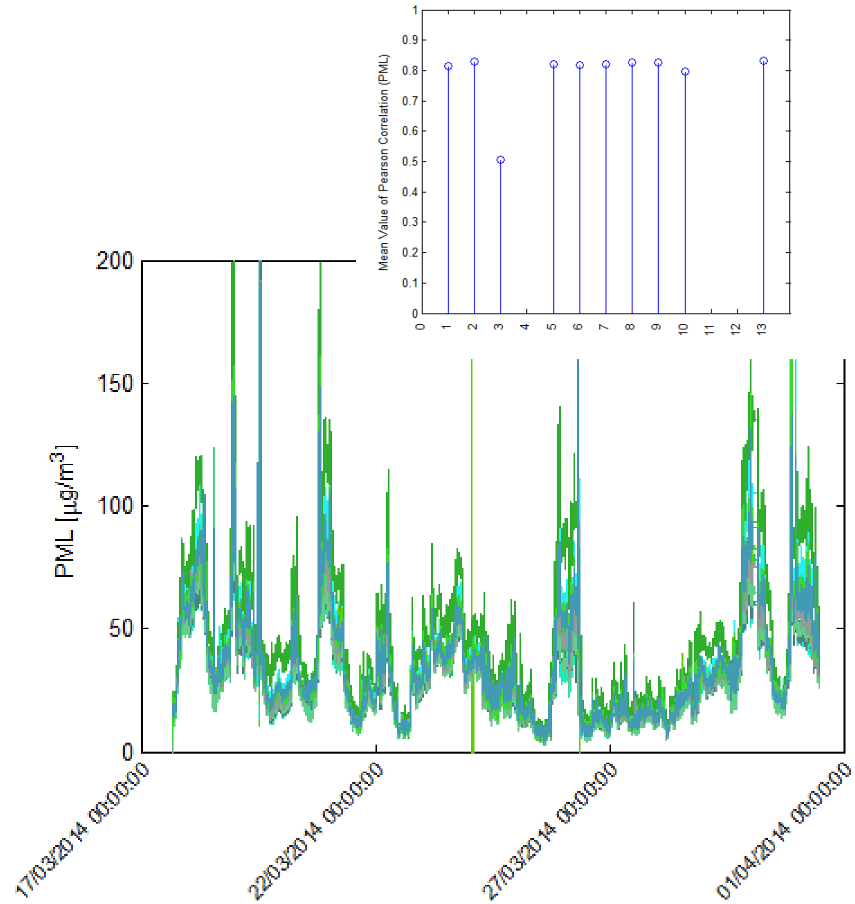
RESULTS

PMS (NB)	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.000	0.993	0.996		0.991	0.963	0.963	0.991	0.989	0.993			0.995
2	0.993	1.000	0.992		0.985	0.954	0.954	0.983	0.989	0.996			0.989
3	0.996	0.992	1.000		0.994	0.955	0.957	0.985	0.987	0.995			0.993
4													
5	0.991	0.985	0.994		1.000	0.951	0.953	0.982	0.979	0.987			0.992
6	0.963	0.954	0.955		0.951	1.000	0.990	0.979	0.976	0.947			0.972
7	0.963	0.954	0.957		0.953	0.990	1.000	0.977	0.975	0.948			0.972
8	0.991	0.983	0.985		0.982	0.979	0.977	1.000	0.987	0.979			0.992
9	0.989	0.989	0.987		0.979	0.976	0.975	0.987	1.000	0.988			0.990
10	0.993	0.996	0.995		0.987	0.947	0.948	0.979	0.988	1.000			0.988
11													
12													
13	0.995	0.989	0.993		0.992	0.972	0.972	0.992	0.990	0.988			1.000

PML (NB)	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.000	0.952	0.522		0.950	0.961	0.958	0.974	0.947	0.864			0.972
2	0.952	1.000	0.520		0.972	0.941	0.953	0.962	0.985	0.962			0.982
3	0.522	0.520	1.000		0.516	0.518	0.519	0.523	0.519	0.493			0.524
4													
5	0.950	0.972	0.516		1.000	0.924	0.927	0.958	0.958	0.921			0.978
6	0.961	0.941	0.518		0.924	1.000	0.985	0.982	0.954	0.891			0.963
7	0.958	0.953	0.519		0.927	0.985	1.000	0.977	0.964	0.912			0.969
8	0.974	0.962	0.523		0.958	0.982	0.977	1.000	0.962	0.904			0.978
9	0.947	0.985	0.519		0.958	0.954	0.964	0.962	1.000	0.953			0.977
10	0.864	0.962	0.493		0.921	0.891	0.912	0.904	0.953	1.000			0.930
11													
12													
13	0.972	0.982	0.524		0.978	0.963	0.969	0.978	0.977	0.930			1.000

RESULTS


Mean Value of Pearson Correlation (PMS)
PMS [µg/m3]
17/03/2014 00:00:00
22/03/2014 00:00:00
27/03/2014 00:00:00
01/04/2014 00:00:00


Mean Value of Pearson Correlation (PML)
PML [µg/m3]
17/03/2014 00:00:00
22/03/2014 00:00:00
27/03/2014 00:00:00
01/04/2014 00:00:00

RESULTS

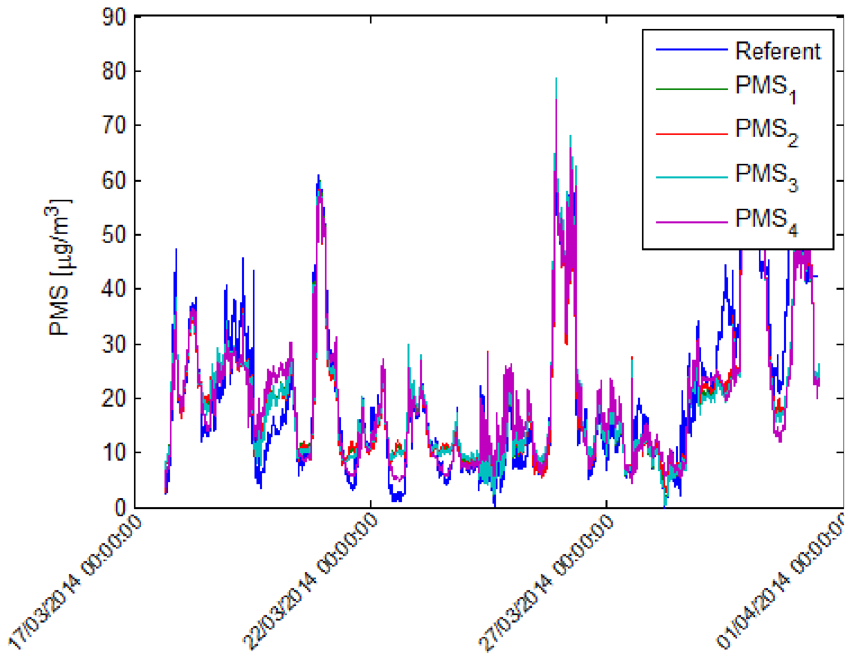
(NB)	PMS/Ref
1	0.910595
2	0.907096
3	0.914725
4	
5	0.922014
6	0.913865
7	0.918648
8	0.912448
9	0.918137
10	0.90797
11	
12	
13	0.92564

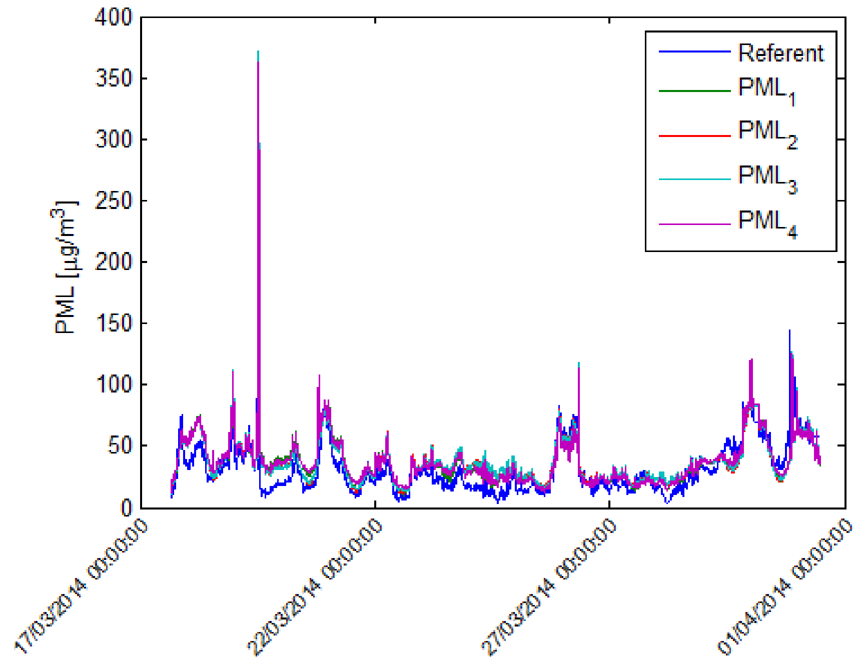
(NB)	PML/Ref
1	0.855904
2	0.8479
3	0.808534
4	
5	0.858145
6	0.791292
7	0.796262
8	0.834348
9	0.84524
10	0.794618
11	
12	
13	0.86068

MULTIVARIATE LINEAR REGRESSION MODELS FOR CALIBRATION OF LOW-COST SENSORS FOR AIR-POLLUTION MONITORING

Model	Parameters	R^2	RMSE	AARE
MLRpms-1(NB)	*PMS, T, H, p*	0.922487	5.878108	3.981672
MLRpms-2(NB)	*PMS, T*	0.921241	5.923232	4.039527
MLRpms-3(NB)	*PMS, H*	0.921908	5.899118	4.039738
MLRpms-4(NB)	*PMS, p*	0.911240	6.271679	4.236747
MLRpml-1 (NB)	*PML, T, H, p*	0.861990	10.3595	7.216037
MLRpml-2 (NB)	*PML, T*	0.857081	10.52824	7.247795
MLRpml-3 (NB)	*PML, H*	0.857213	10.52376	7.231833
MLRpml-4 (NB)	*PML, p*	0.860658	10.40566	7.287425

MULTIVARIATE LINEAR REGRESSION MODELS FOR CALIBRATION OF LOW-COST SENSORS FOR AIR-POLLUTION MONITORING





CONCLUSION

- From the results it was evident that R^2 parameter that determines the degree of correlation between the output and target data has a higher value for the PMS group in comparison to the PML fraction
- Also, based on the values of statistical tests for different models it was concluded that the use of the maximum number of input parameters in the model gives the best results
- Differences in the results are not high and it was noted that humidity (pressure) has the greatest impact in the linear regression
- When one applying regression methods, for simplicity, it can also be used a model that contains only the humidity(pressure) as an input parameter