Prediction of severe driving conditions in winter

Franc Švegl
and Igor Grabec

Amanova Ltd,
Technology Park Ljubljana, Slovenia

SIRWEC Quebec, 5. 2. 2010

ROADIDEA

AMANOVA d.o.o.
Inteligentni sistemi, senzorji in novi materiali
Intelligent systems, sensors and new materials
Član Tehnološkega parka Ljubljana
Member of Technology Park Ljubljana
Tehnološki park 18, Ljubljana

Objectives

- **Problem:** Statistical estimation of relation between measured environmental variables and driving conditions in winter.
- **Basis for treatment:** Experimentally estimated probability density function - PDF.
- **Extraction of a relation:** General regression expressed by the conditional average estimator - CA.
- **Goal:** To provide data for optimization of winter roads service by intelligent control.

Experimental basis

The picture can't be displayed.

- A vector variable $Z = (X, Y)$ is utilized to join data about environment - X, and driving conditions - Y.

- Calibration by units u and v yields the joint instrument scattering function

$$w(z,w;\sigma) = w(x,u;\sigma)\,w(y,v;\sigma)$$

σ is assumed to be equal for all components

Statistical basis
the joint probability density - PDF

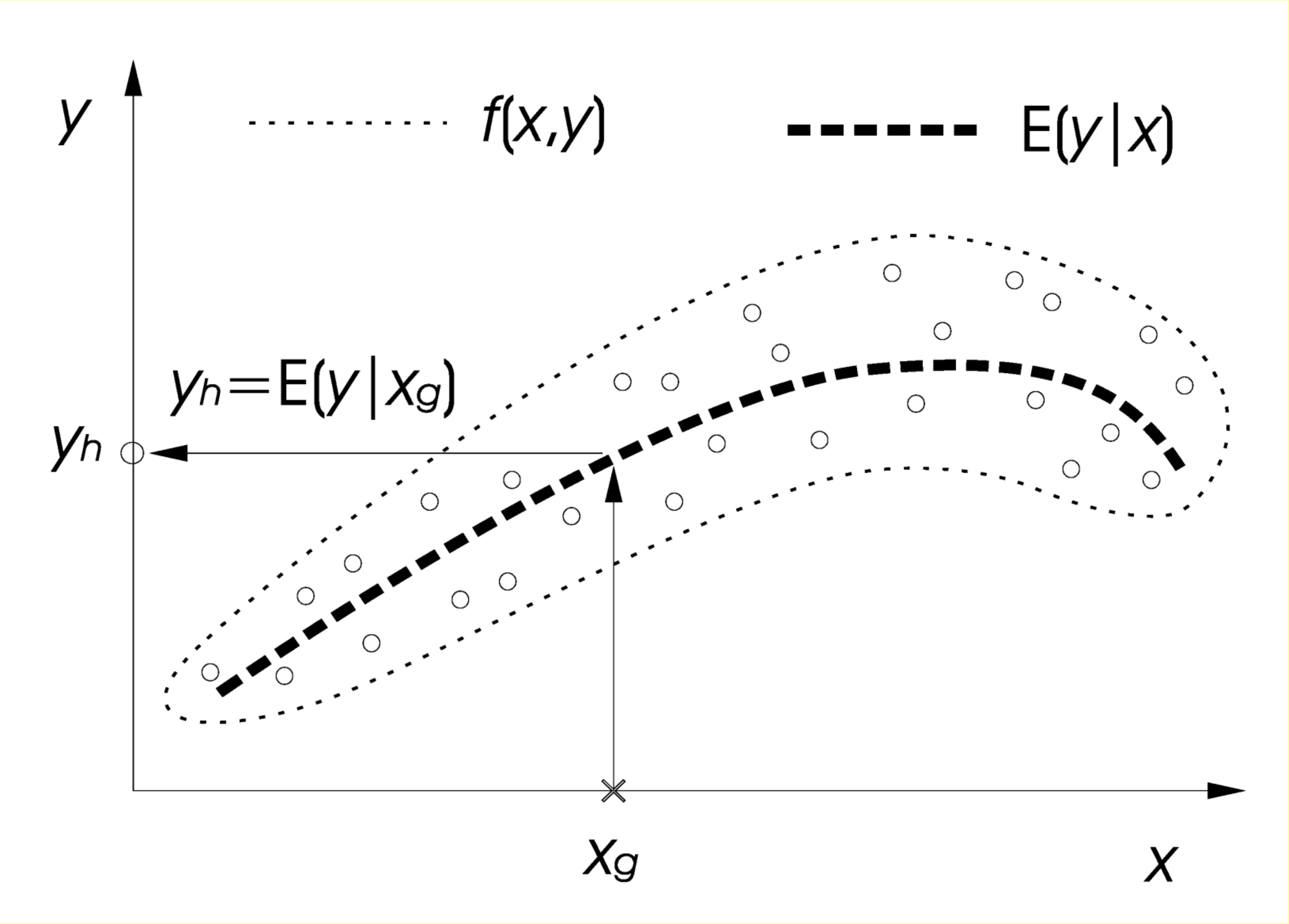
N measured joint data $z_1, \ldots, z_N$ are given

$$f(x,y) = \frac{1}{N} \sum_{i=1}^{N} w(x, x_i; \sigma)\, w(y, y_i; \sigma)$$

w denotes normal probability distribution

σ is the mean distance between data points

Statistical estimation of hidden driving conditions Y_h from given weather data X_g



Extraction of relation *Y*(*x*) from PDF

Optimal predictor is the conditional average:

The picture can't be displayed.

$$Y_p(x) = \mathrm{E}\left[\, y \,|\, x \,\right] = \int y\, f(y|x)\, dy$$

Expressed by data it gets the form:

The picture can't be displayed.

$$Y_p(x) = \frac{\sum_{i=1}^{N} y_i\, w(x - x_i;\sigma)}{\sum_{j=1}^{N} w(x - x_j;\sigma)} = \sum_{i=1}^{N} y_i\, S_i(x)$$

Properties of $S_i(x)$

$$S_i(x) = \frac{w(x - x_i; \sigma)}{\sum_{j=1}^{N} w(x - x_j; \sigma)}$$

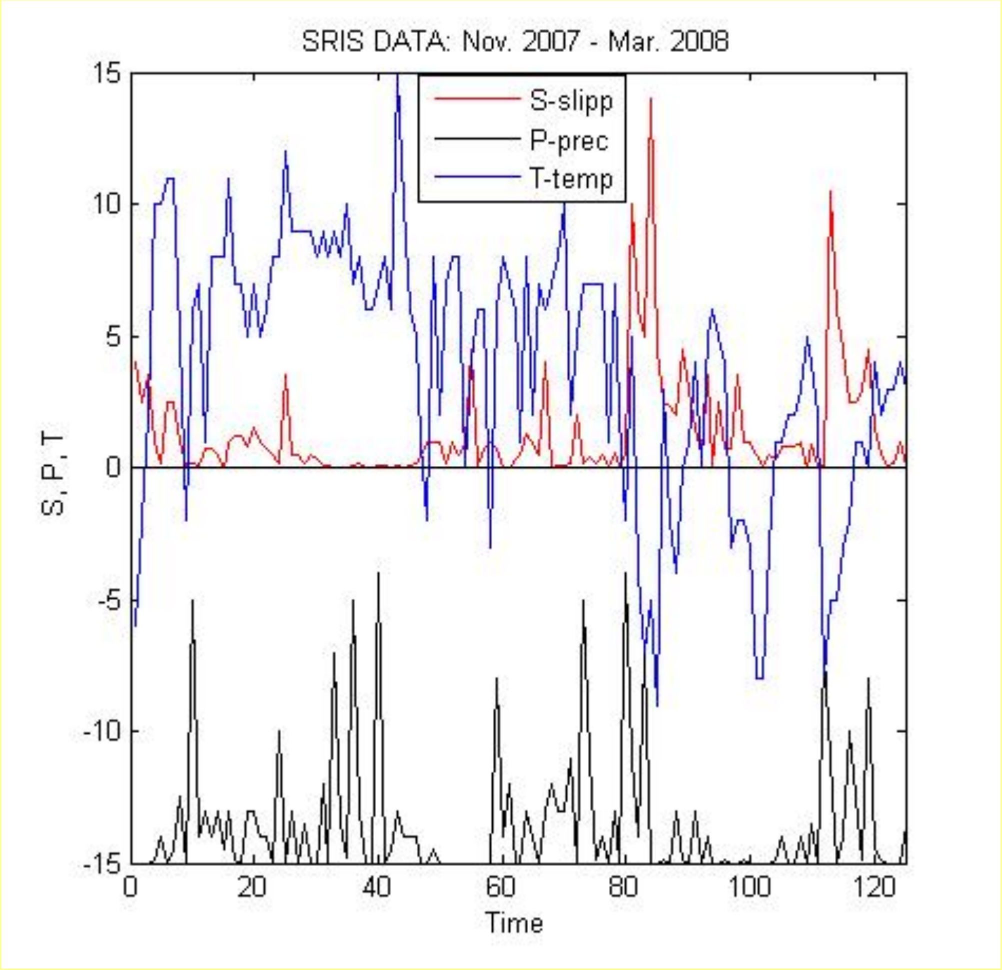
S_i is a normalized measure of similarity between given x and stored x_i

$$\sum_{i=1}^{N} S_i = 1 \qquad\qquad 0 \le S_i \le 1$$

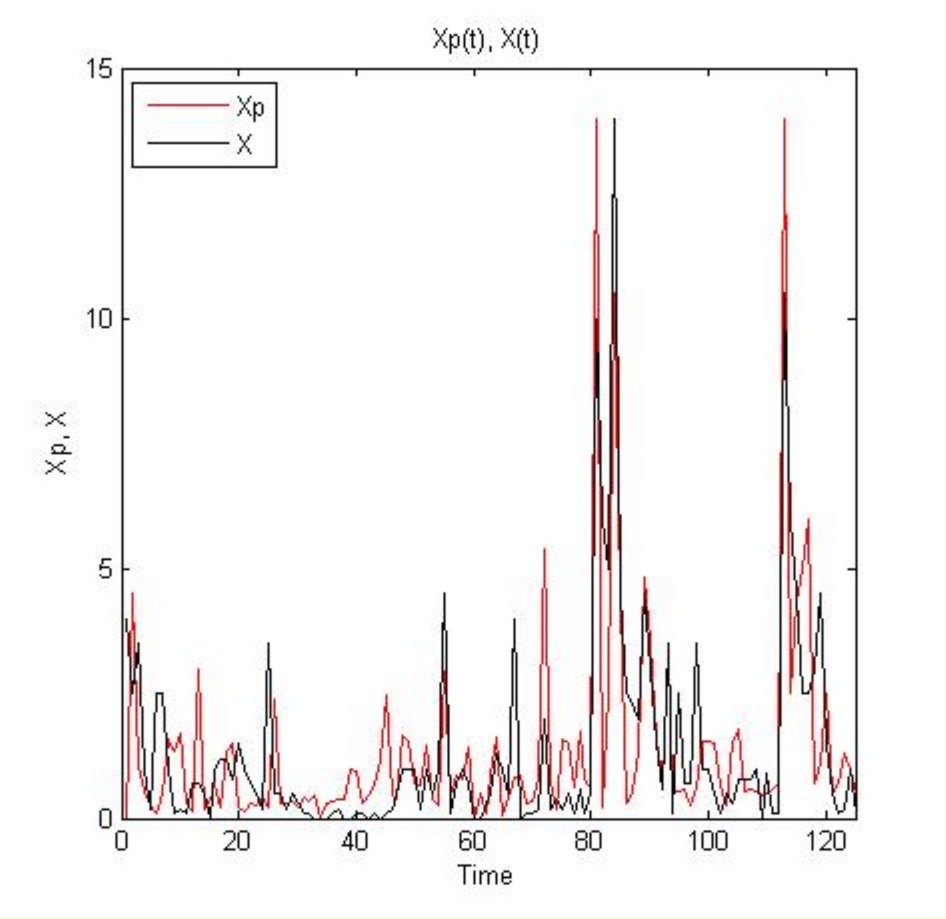
Prediction of road slipperiness from weather forecasts in Sweden

Data: S – Slipperiness
P – Precipitation,
T – Temperature,

Data provided by: Slippery road information system – SRIS - www.sris.nu



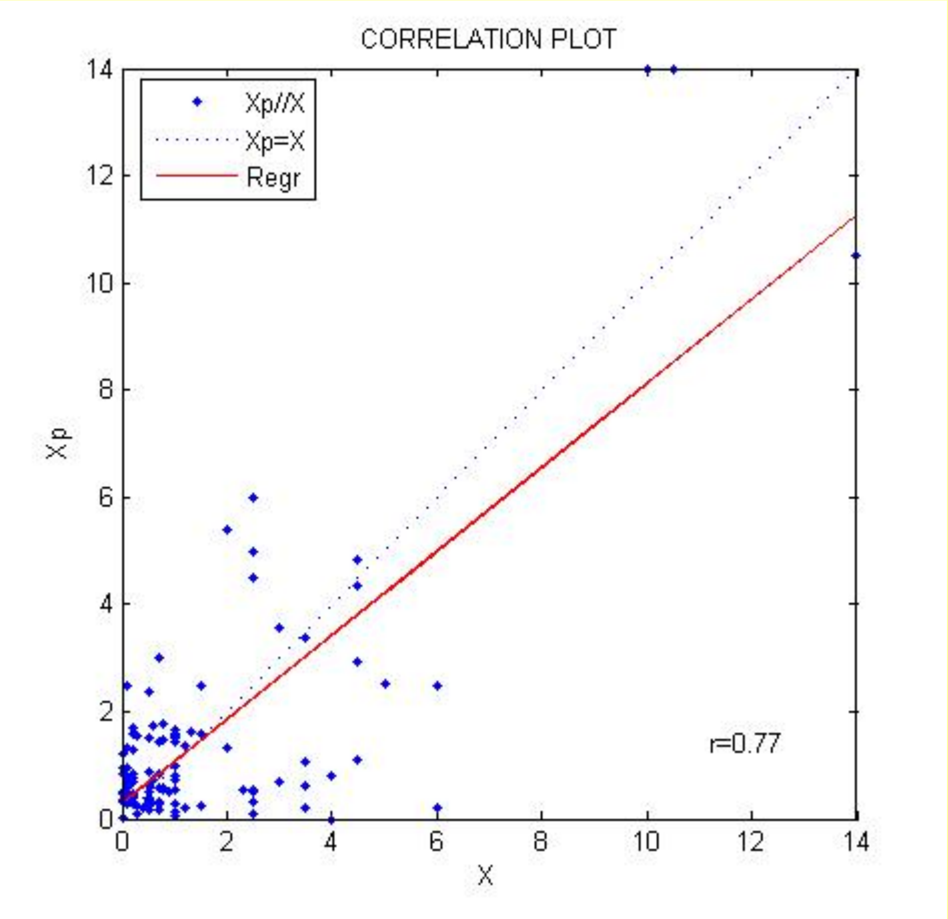
Predicted and original slipperiness



1 day ahead prediction

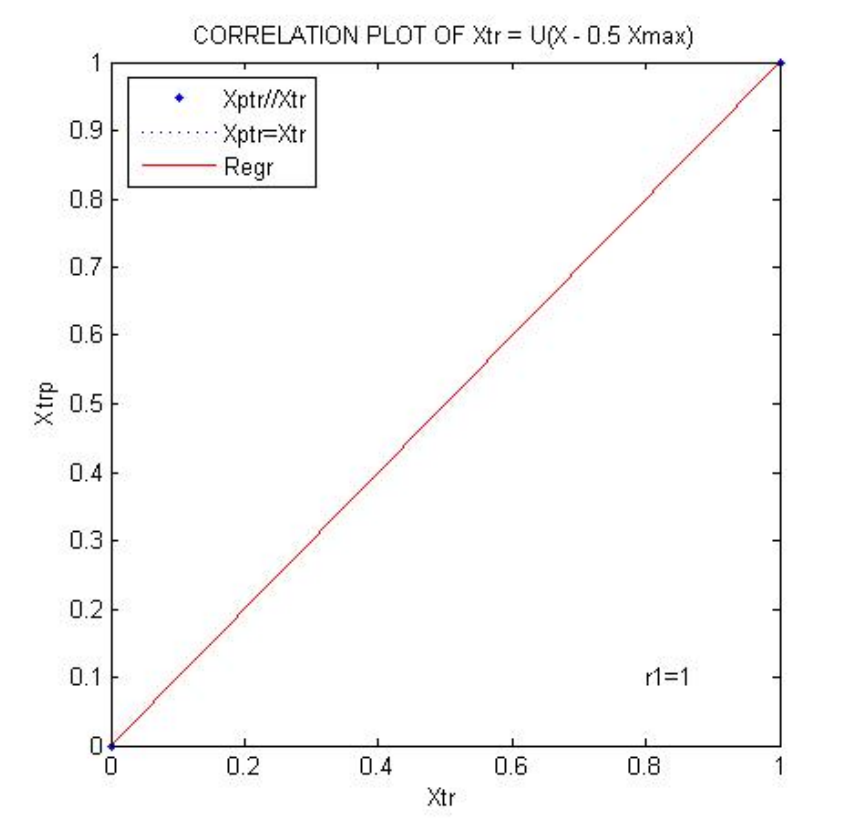
Accounting of past data improves accuracy of prediction

Correlation plot of X_p and actual X



r – correlation coefficient of X_p and X

Correlation plot of transformed critical variable: $X_{tr} = U(X - 0.5\ X_{max})$



r – correlation coefficient of X_{trp} and X_{tr}

Prediction of air pollution ARPV data about *PM10*

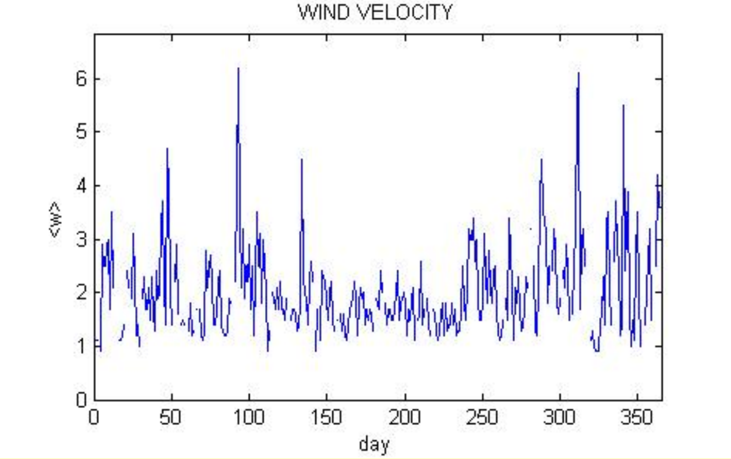
Microsoft Excel - Arpv_data

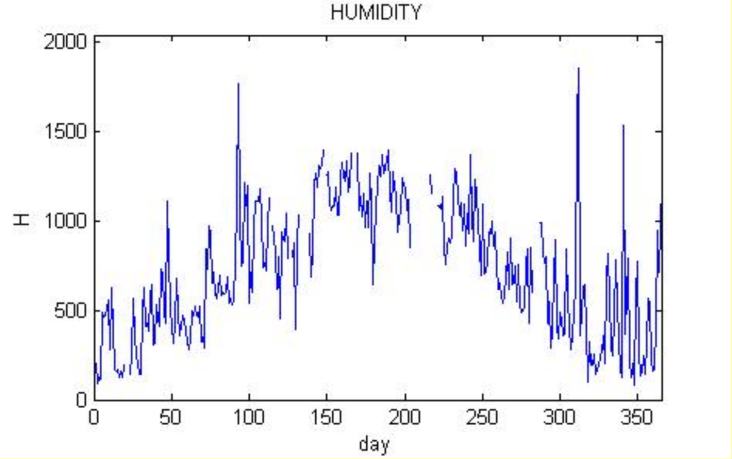
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1		Percentage of wind calm <0.5 m/s	Average wind speed	Average wind speed (without values<0.5 m/s)	Average wind direction	Hmix rural	Hmix urban	Stanford Index	Average radiation	T minimum	Average T	Total precipitation	Equavalent potetial temperature	Concentration PM10 Air Quality Station in Arcella	Concentration PM10 in Air Quality Station in Mandia	Concentration PM10 in Air Quality Station. Average Mandria and Arcella				
2	date	% wind calm	vv_m	vvc_m	dv_m	zi_r	zi_u	stan_u	rmed	tmin	tmed	ptot	thte	PM10_arc	PM10_man	PM10_tipo		aa	mm	gg
3	1.1.2003	17	1	1 1	289,1	156	206	0,54	81	277	280	0	296	63	62	63,0		2003	1	1
4	2.1.2003	12	1	1 1	352,9	46	90	0,69	7	278	278	0,2	294	55	54	55,0		2003	1	2
5	3.1.2003	25	1		85,2	74	126	1,37	17	277	278	0	293	53	46	50,0		2003	1	3
6	4.1.2003	17	0,8	0 9	75,1	93	117	2,74	50	274	276	0,2	289	62	46	54,0		2003	1	4
7	5.1.2003	8	2,7	2 9	299,5	266	491	0,39	31	276	278	5,8	293	44	40	42,0		2003	1	5
8	6.1.2003	8	2,3	2 5	326	262	464	0,05	42	274	276	8	288	48	36	42,0		2003	1	6
9	7.1.2003	0	2,5	2 5	338,2	245	471	0	19	274	275	5	287	33	25	29,0		2003	1	7
10	8.1.2003	0	2,9	2 9	336,4	271	521	0	12	274	275	0,2	285		41	41,0		2003	1	8
11	9.1.2003	0	3	3	334,6	285	553	0,02	28	273	274		301		32	32,0		2003	1	9
12	10.1.2003	12	1,5	1 7	337,2	188	284	0	51	273	274	0	289					2003	1	10
13	11.1.2003	0	3,5	3 5	323,2	347	623	0,03	87	272	274	0	298					2003	1	11
14	12.1.2003	4	2,1	2 1	357,9	257	391	0,36	100	269	272	0	302					2003	1	12
15	13.1.2003	37	0,7		57,1	136	178	2,75	83	267	271	0	311		130	130,0		2003	1	13
16	14.1.2003	62	0,5		41	119	156	1,81	84	269	273	0,2	314		152	152,0		2003	1	14
17	15.1.2003	62	0,5		14,4	125	166	5,1	80	270	273	0,2	315		188	188,0		2003	1	15
18	16.1.2003	8	1	1 1	12	91	126	4,98	46	270	274	0,2	305		196	196,0		2003	1	16
19	17.1.2003	17	0,9	1 1	41,7	120	163	3,54	72	272	276	0	300		127	127,0		2003	1	17
20	18.1.2003	17	1,1	1 2	34,8	84	124	1,43	39	272	277	0	291	128	103	116,0		2003	1	18
21	19.1.2003	0	1,4	1 4	68,7	156	198	4,08	105	271	274	0,2	318	76	67	72,0		2003	1	19
22	20.1.2003	33	0,7		78,4				93	271	275	0,4	306	101	97	99,0		2003	1	20
23	21.1.2003	4	2,3	2 4	312,2				17	276	278	15	293	83	87	85,0		2003	1	21
24	22.1.2003	17	1,8	2 1	39,8				31	275	279	2	294	49	50	50,0		2003	1	22
25	23.1.2003	37	0,8		70	112	141	2,39	56	273	277	0,2	296	84	73	79,0		2003	1	23
26	24.1.2003	0	1,9	1 9	339,7	170	258	2,17	107	274	279	0	291	67	52	60,0		2003	1	24
27	25.1.2003	0	3,1	3 1	326,3	331	566	1,22	108	277	279	0,2	290	49	37	43,0		2003	1	25
28	26.1.2003	0	2,3	2 3	335,9	268	365	2,23	111	275	278	0	288		47	47,0		2003	1	26
29	27.1.2003	4	1,1	1 2	66,4	192	245	3,29	96	273	277	0	294		86	86,0		2003	1	27
30	28.1.2003	12	1,2	1 4	0,4	141	200	2,63	87	272	276	0	303	148	134	141,0		2003	1	28

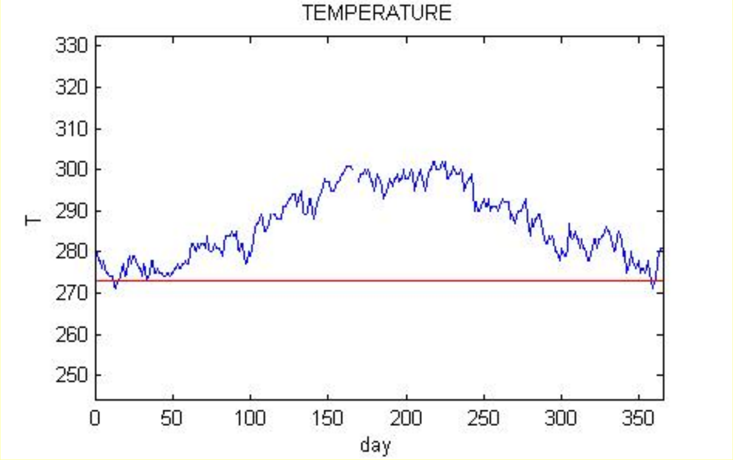
Selection of variables used in modelling predictor of *PM10*

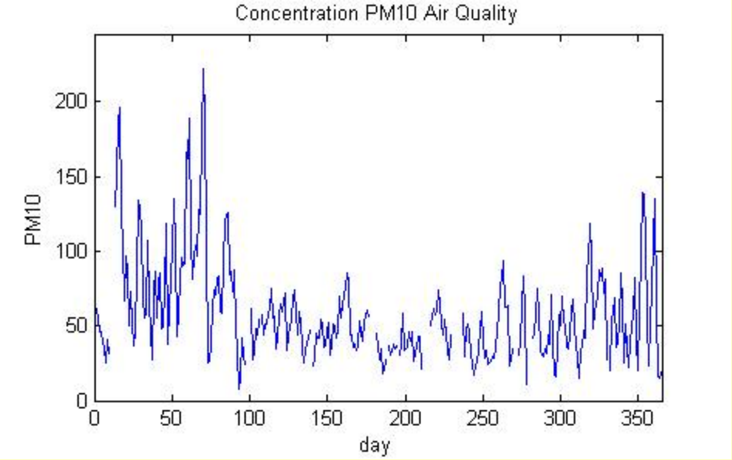
- As given variables we consider: the average wind velocity - *W*, humidity – *H* and average temperature – *T*.
- As hidden variables we consider concentration *P=PM10*.
- Using sample vectors $\mathbf{Z}_i = (W,H,T,P)$ *from* the data base we create statistical model of the relation *P=G(W,H,T)* by the CA estimator.
- By using the model we predict hidden *P* from some given data *W,H,T*.
- Here the time is used as sample index *i*.
- Agreement between predicted and really measured data is described by the correlation coefficient *r* and shown in a correlation diagram.

Variables used in modelling

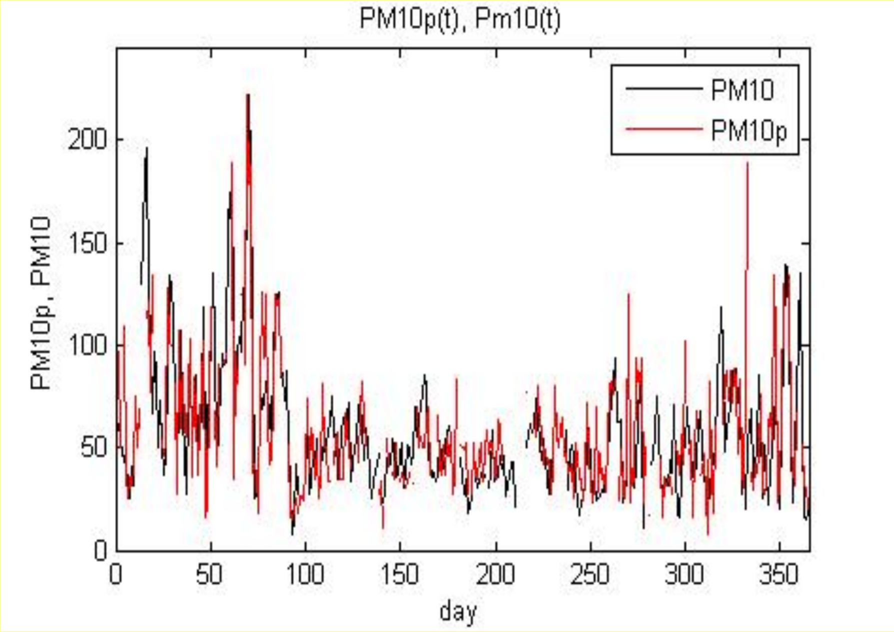


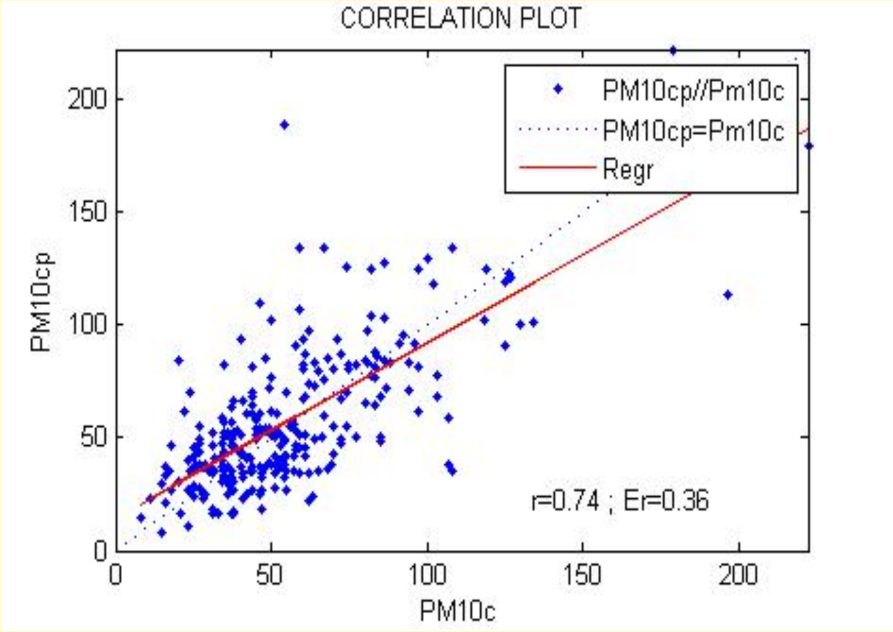




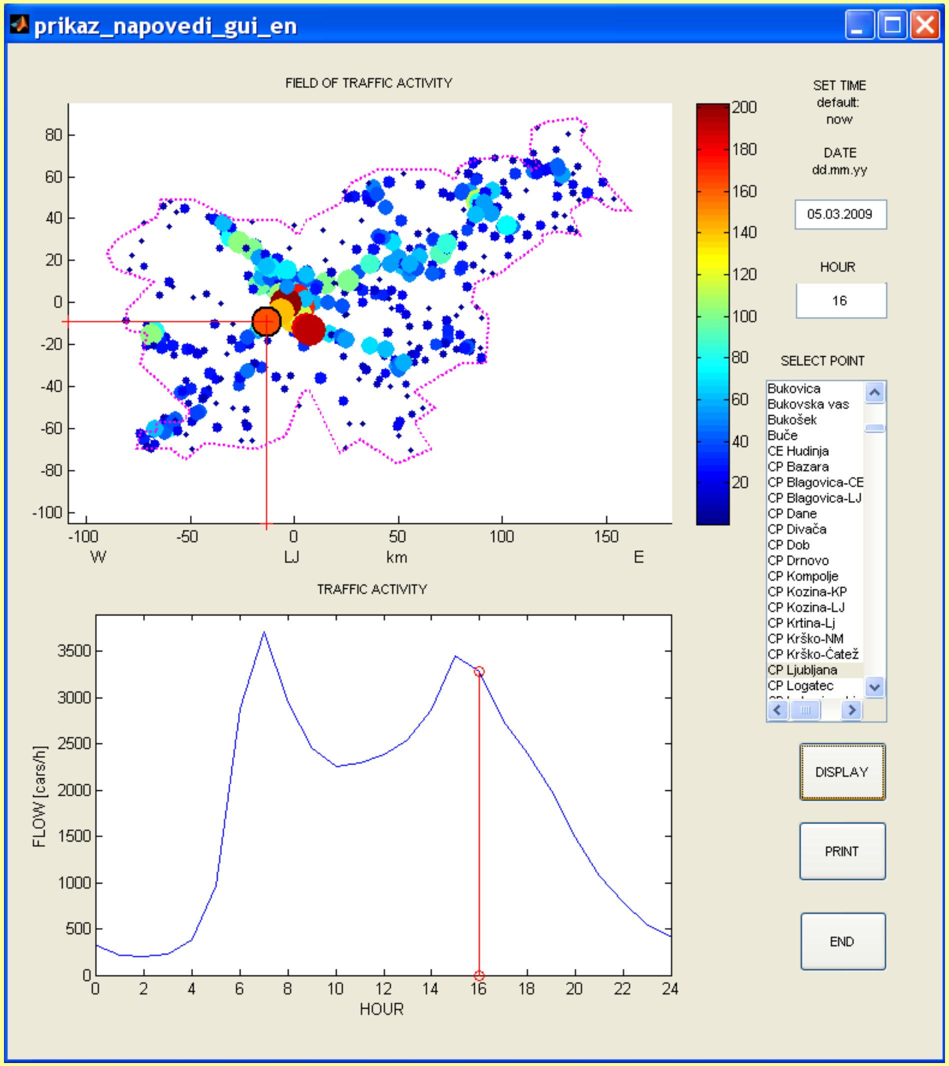


Results of prediction





Graphic user interface for display of predicted traffic activity in Slovenia



- User sets time – day and hour - of prediction in the interval from now to the year 2015.
- From the pop-up menu an observation point is selected.
- GUI displays the field of traffic activity from the start of the selected day to the selected hour in the top graph.
- GUI displays the distribution of predicted traffic flow over the selected day in the bottom graph.
- The selected place and hour of prediction are marked in graphs.
- The prediction can be repeated with varied time and place.
- The display can be printed.

Coclusions

- Our approach to prediction of driving conditions needs no analytical model, but extracts it from experimental data.
- The same approach was successfully applied also to forecasting of traffic flows.
- The method provides information support for planning of winter roads service.
- The next step is an approch to intelligent control of winter roads service.

References

- I. Grabec, W. Sachse, *Synergetics of Measurement, Prediction and Control*, Springer-Verlag, Berlin, 1997
- I. Grabec, K. Kalcher, F. Švegl, *Modelling and Forecasting of Traffic Flow on Slovenian High-Ways*, Transport Research Arena Europe 2008", Ljubljana, SI, April 21-24, 2008
- I. Grabec, K. Kalcher, F. Švegl, *Statistical Forecasting of Traffic Flow Rate*, Proc. of the conf.: SIRWEC – 14th International Road Weather Conference, Prague, CZR, May 14-16, 2008
- Igor Grabec, Kurt Kalcher, Franc Švegl, *Statistical forecasting of traffic activity*, 9. Slovenian Congress on Traffic, Portorož, October 22-24, 2008