FORECASTING OF TRAFFIC FLOWS AND JAMS ON HIGH-WAYS

IGOR GRABEC

Univerza v Ljubljani & Amanova d.o.o., Tehnološki park Ljubljana

Optimiranje ruta vozila korištenjem stvarno-vremenskih prometnih podataka,
30. lipnja 2014, ZUK Borongaj, Fakultet prometnih znanosti u Zagrebu

How a child copes with a complex traffic? Based upon learning!

Objective of the lecture

- **Problem:** Traffic flows are very complex and therefore hardly modeled and forecast analytically. Our aim is to show how this problem can be solved statistically.

- **General solution:** Non-parametric regression expressed in terms of measured data.

- **Examples:** Forecasting of traffic flows, road slipperiness and jams.

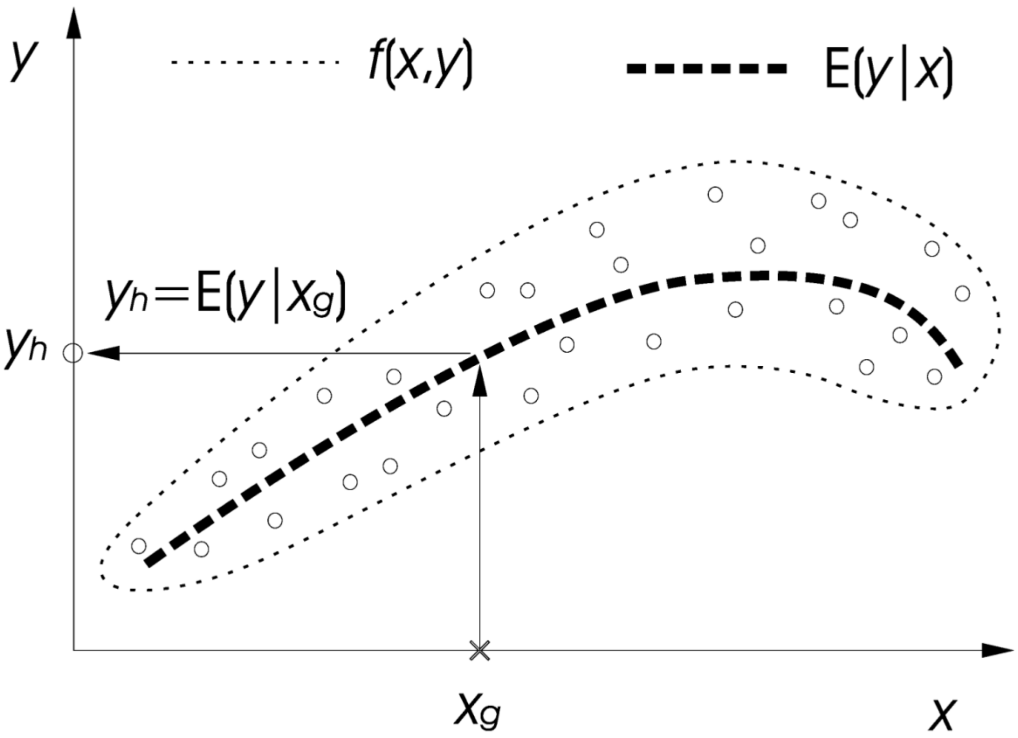
Statistical treatment

- **Basis of description:** Experimentally estimated probability density function - PDF. The kernel of the estimator is the instrument scattering function.

- **Estimation of relations between measured variables:** Non-parametric regression determined by the conditional mean estimator.

Statistical estimation of functions from measured data by the conditional mean

o	measured data
f(*x*,*y*)	probability density estimated from data
E(*y*|*x*)	conditional mean defined by f(x,y)
------	estimated function $y(x)=E(y|x)$
x_g	given datum
y_h	hidden datum estimated by E(*y*|*x*)



Basic properties of data in the two-dimensional case

- A vector variable $Z = (X,Y)$ is considered
- z_i is the instrument output in a continuous joint sample space $Sz = Sx \otimes Sy$ of size $2L \otimes 2L$
- N measured joint data $\{z_1, \ldots, z_N\}$ are given
- Calibration by a unit $w = u \otimes v$ yields the gaussian instrument scattering function:

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

Scattering width σ is equal for both components

Joint probability density *f*(*x*,*y*)

- From data samples $\{(x,y)_{n\ ;\ n=1 \ldots N}\}$ the probability density is estimated by:

$$f(x,y) = \frac{1}{N} \sum_{n=1}^{N} g\left(x, x_n; \sigma\right) g\left(y, y_n; \sigma\right)$$

- The same model is applicable in a multi-dimensional case, just the number of components is increased: $z = (x,y,\ldots)$

Extraction of a law from PDF

An optimal MSE predictor of a law *Y(x)*
is the conditional average CA:

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

Expressed by data it gets the form:

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

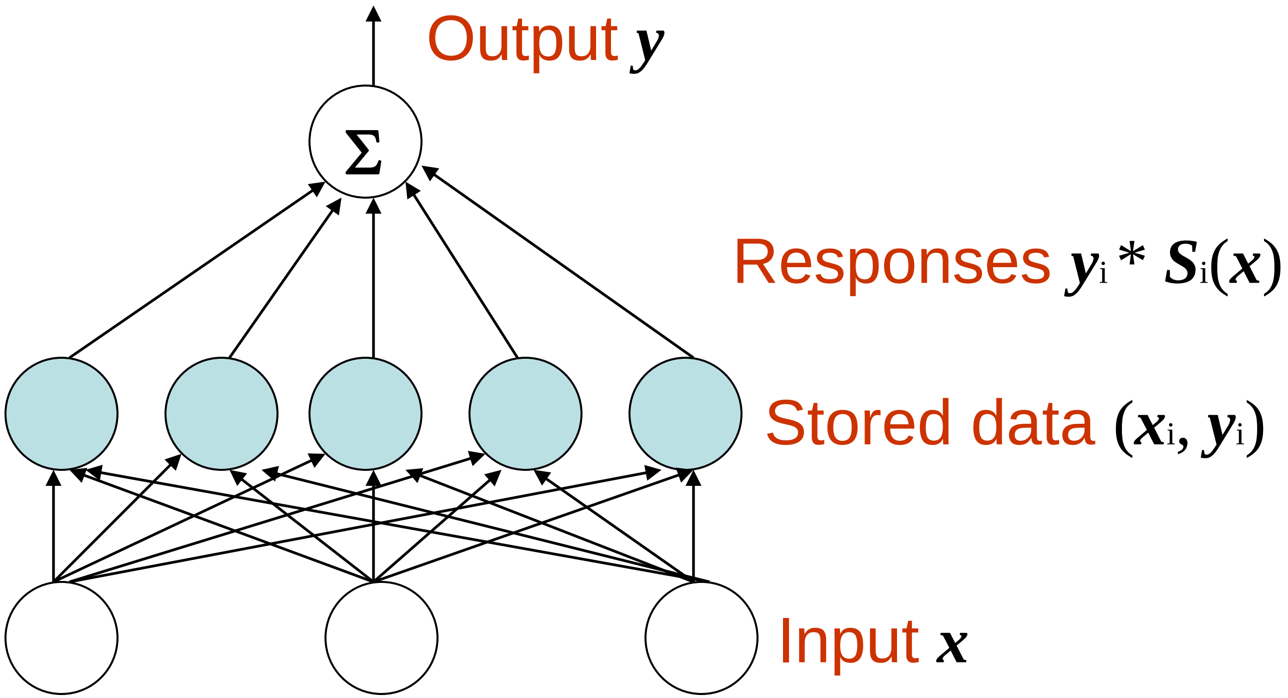
Properties of similarity measure $S_i(x)$

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

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

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

Scheme of the predictor resembles a radial basis function neural network



Predictor quality

$$Q = 1 - \frac{\mathrm{E}\left[(Y_p - Y)^2\right]}{\mathrm{Var}(Y_p) + \mathrm{Var}(Y)}$$

$$= \frac{2\,\mathrm{Cov}(Y_p, Y)}{\mathrm{Var}(Y_p) + \mathrm{Var}(Y)} - \frac{(\mathrm{E}[Y_p] - \mathrm{E}[Y])^2}{\mathrm{Var}(Y_p) + \mathrm{Var}(Y)}$$

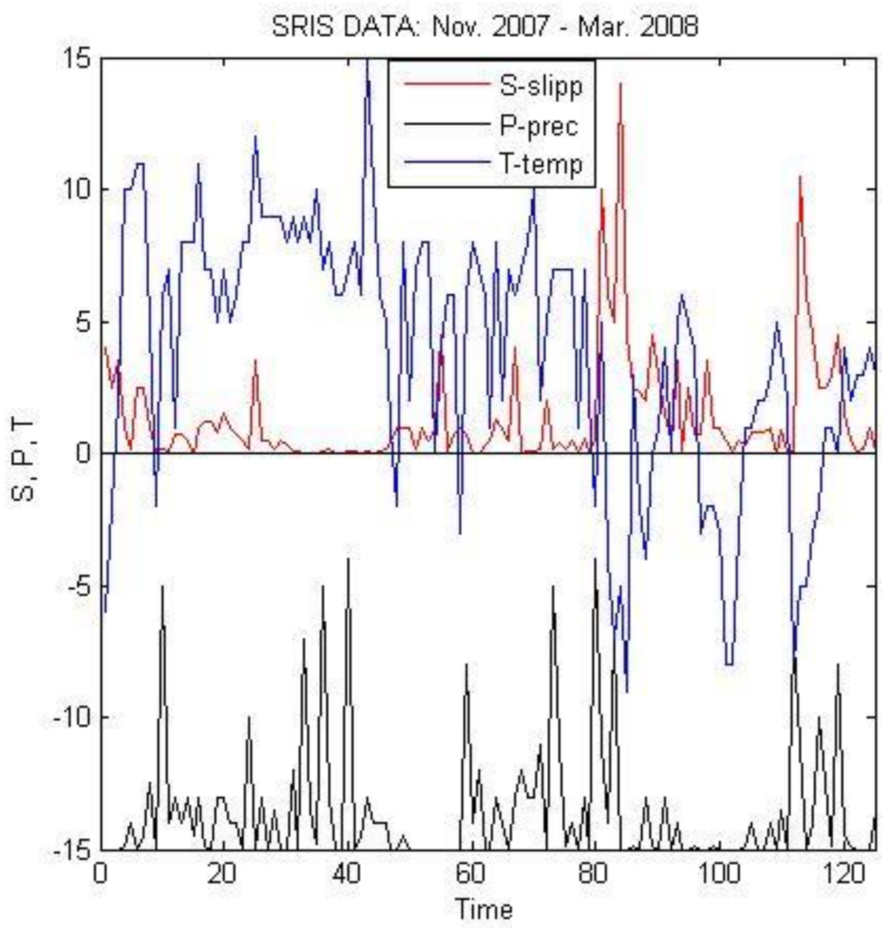
Q=1 for an exact prediction: $Y_p = Y$,
Q=0 for statistically independent Y and Y_p

Q is approximately equal to correlation coefficient r

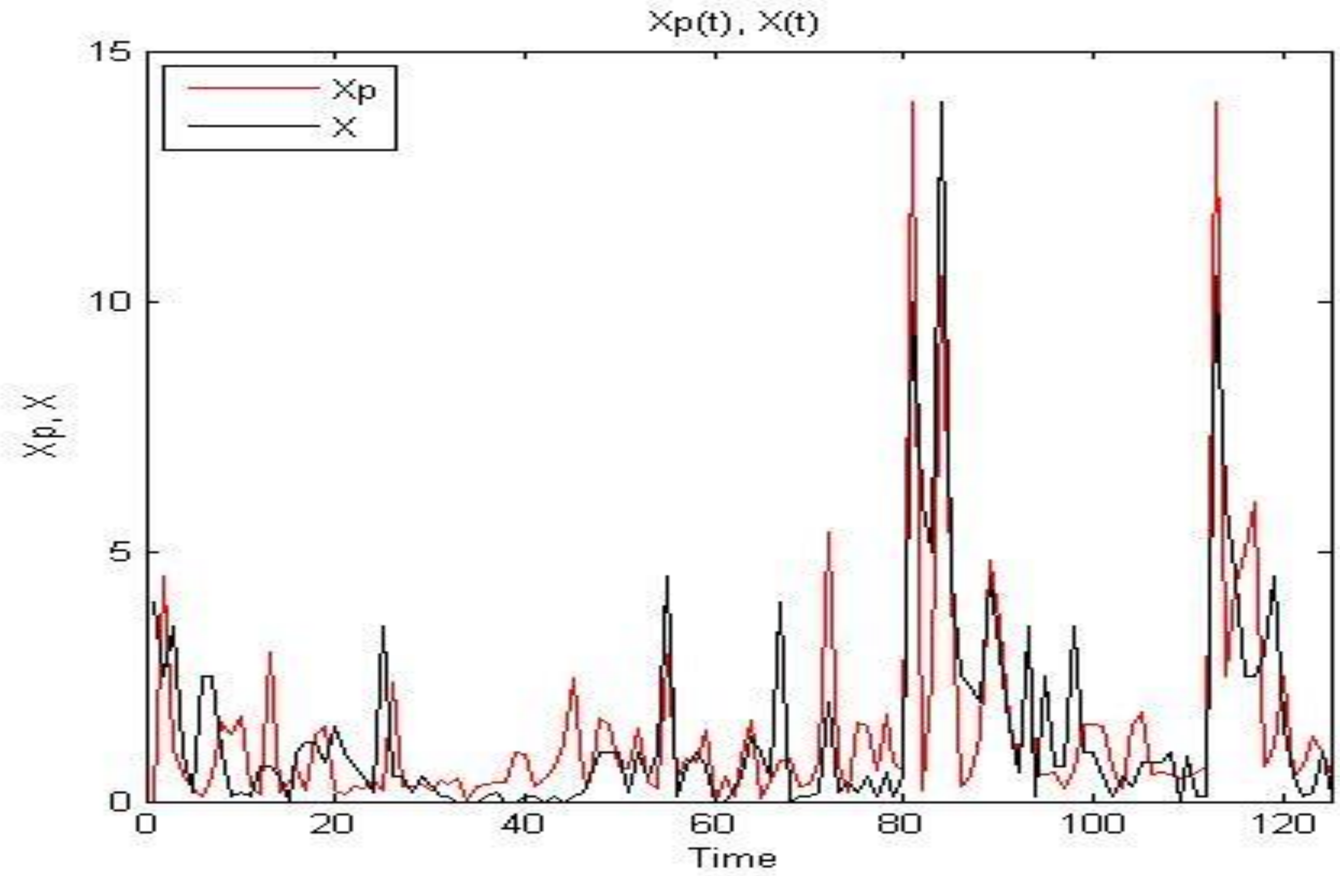
Prediction of road slipperiness from weather forecast in Sweden

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

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



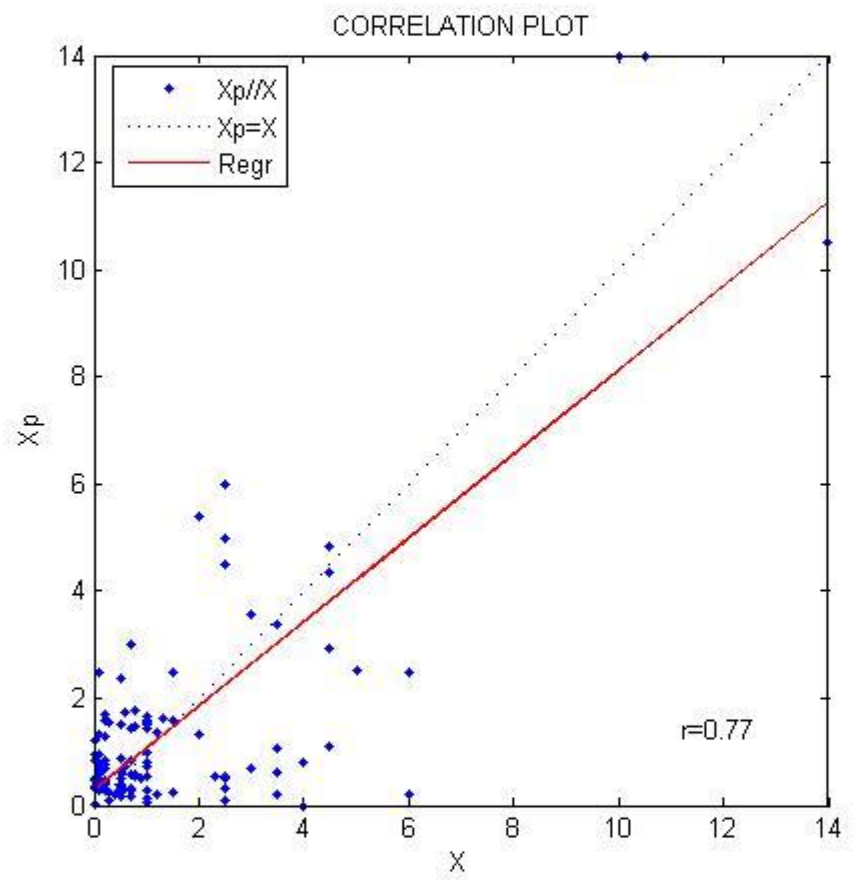
Predicted and original slipperiness



1 day ahead prediction

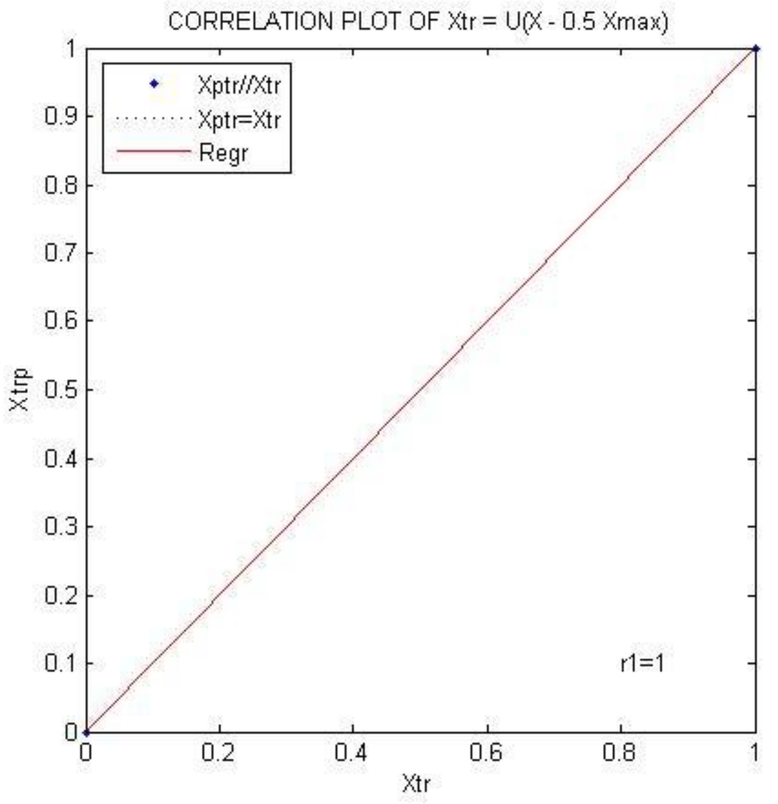
Accounting of past data improves the accuracy of prediction

Correlation plot of predicted and observed slipperiness



r – correlation coefficient of X_p and X

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



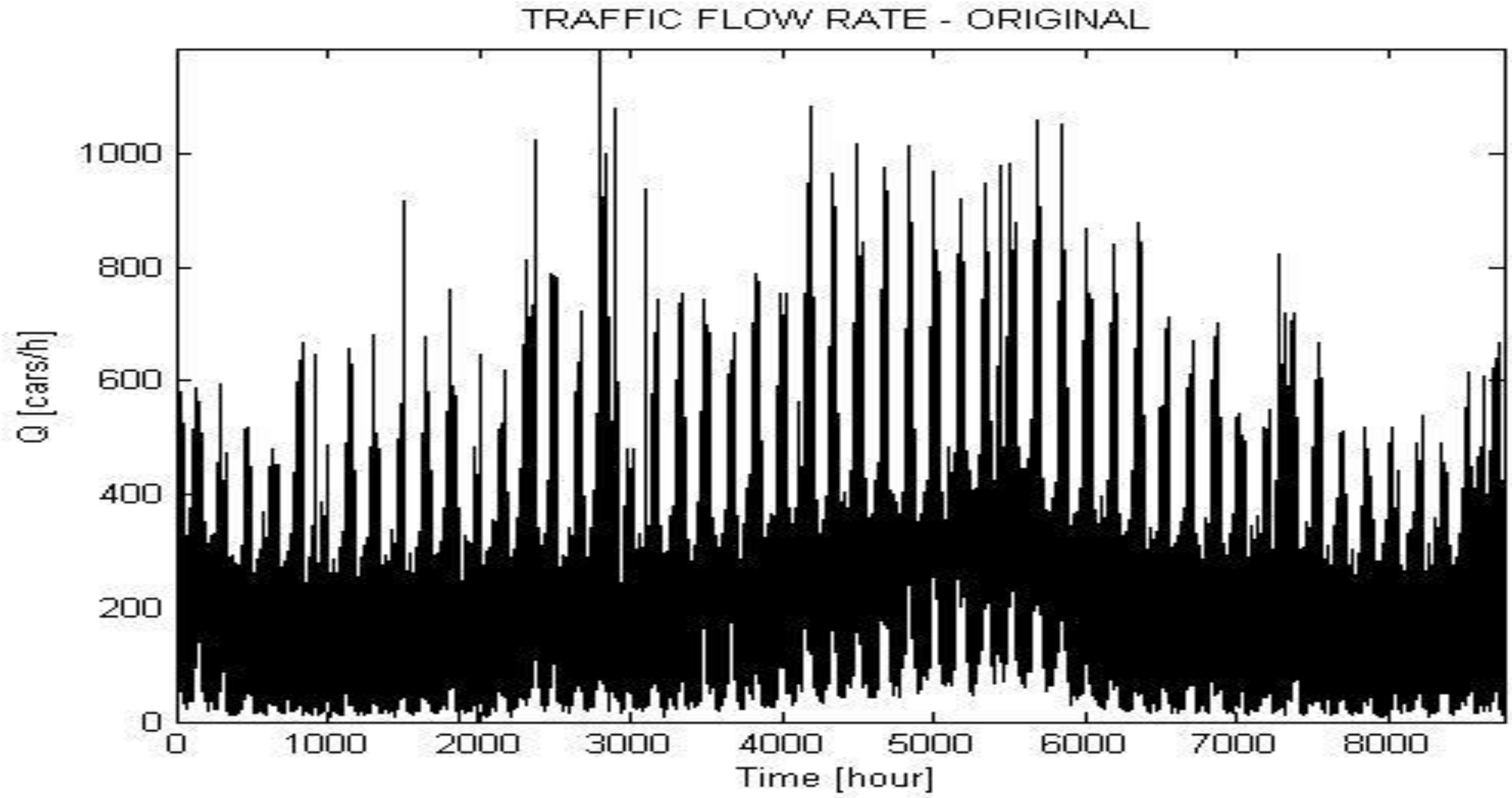
r – correlation coefficient of X_{trp} and X_{tr}

Example: modeling and forecasting the time series of traffic flow

- Codes of day *D(t)* and hour *H(t)* are joined with traffic flow rate *Q(t)* in the state vector:

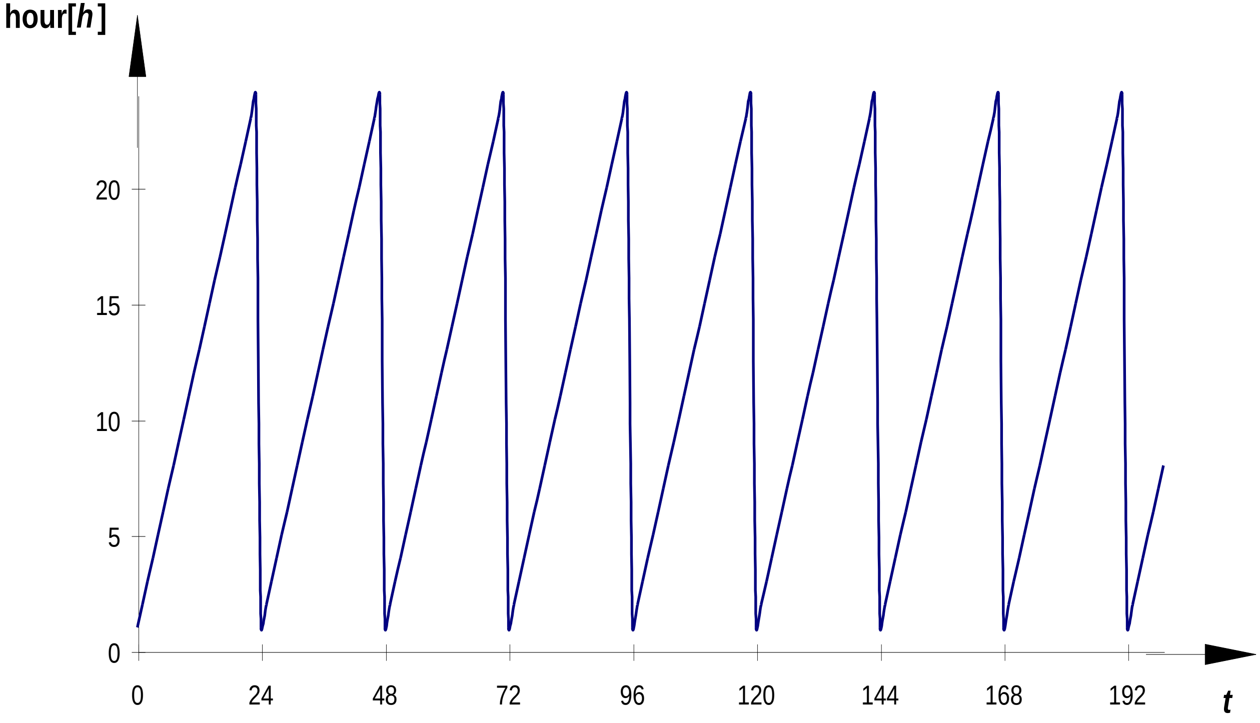
  $Z(t)=(D(t), H(t), Q(t), Q(t-1),\ldots)$
- Samples from the past time series are used to estimate the probability distribution of the state vector *Z*.
- From given codes and past flow rate, the future flow rate is optimally predicted by the conditional mean estimator: $Q_e(t) = \mathrm{E}[Q \mid D(t), H(t), Q(t-1),\ldots]$
- An optimal combination of condition variables is found by analysis of correlation between predicted and observed data.
- Weather data can be included into condition.

Modeling and forecasting of traffic flow on a high-way

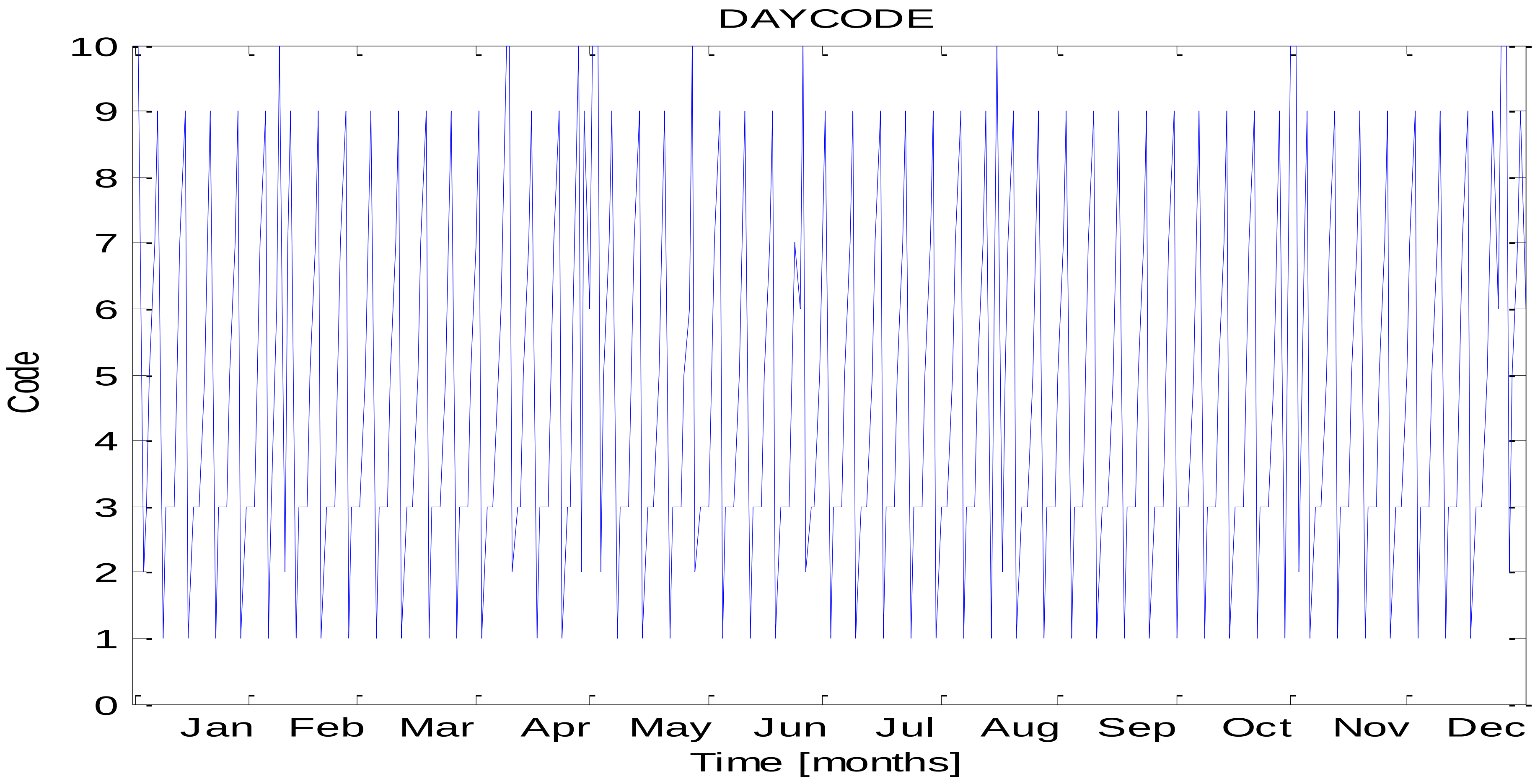


A record of traffic flow rate over a year

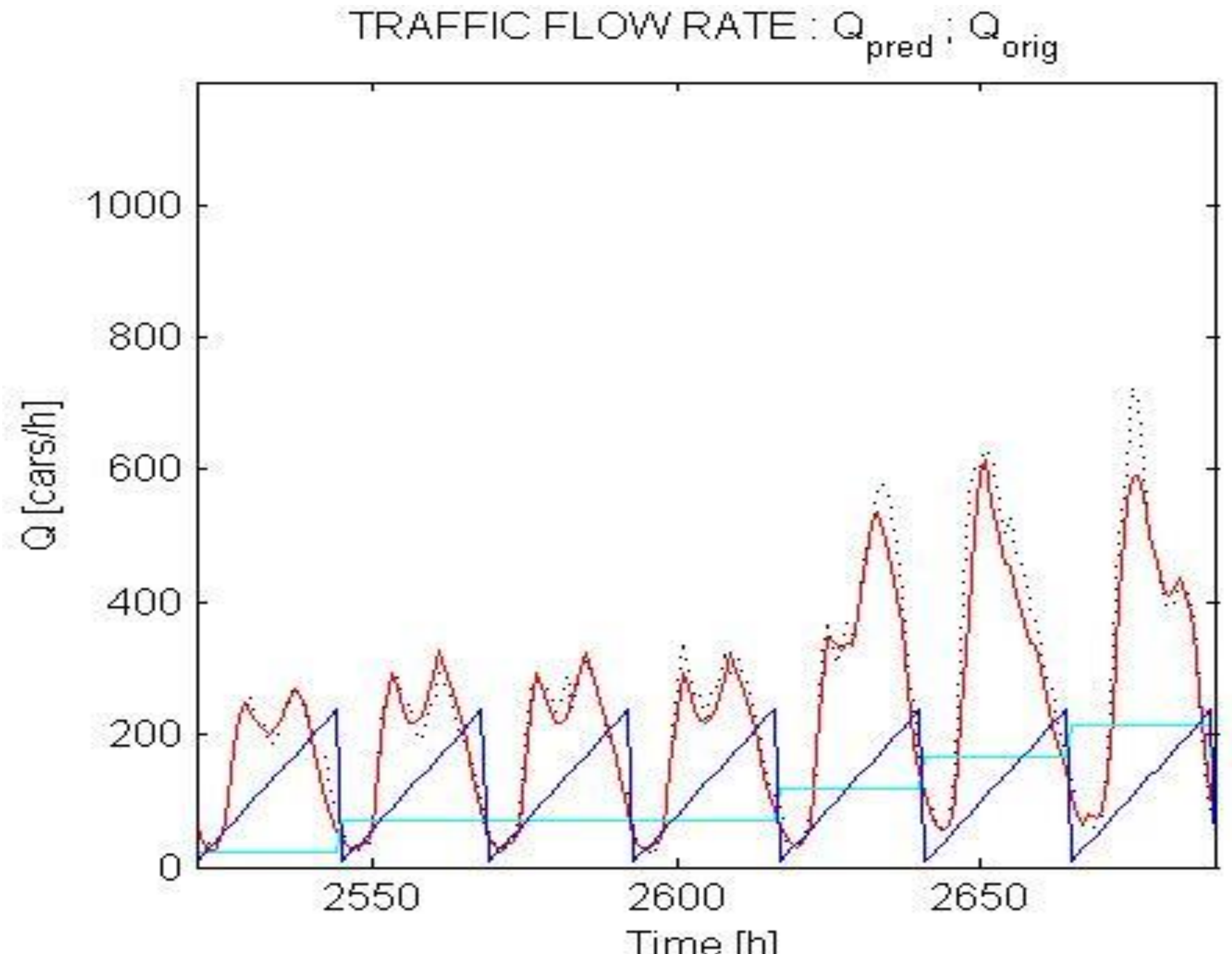
Hour-variable C_h used in modeling of traffic flow dynamics



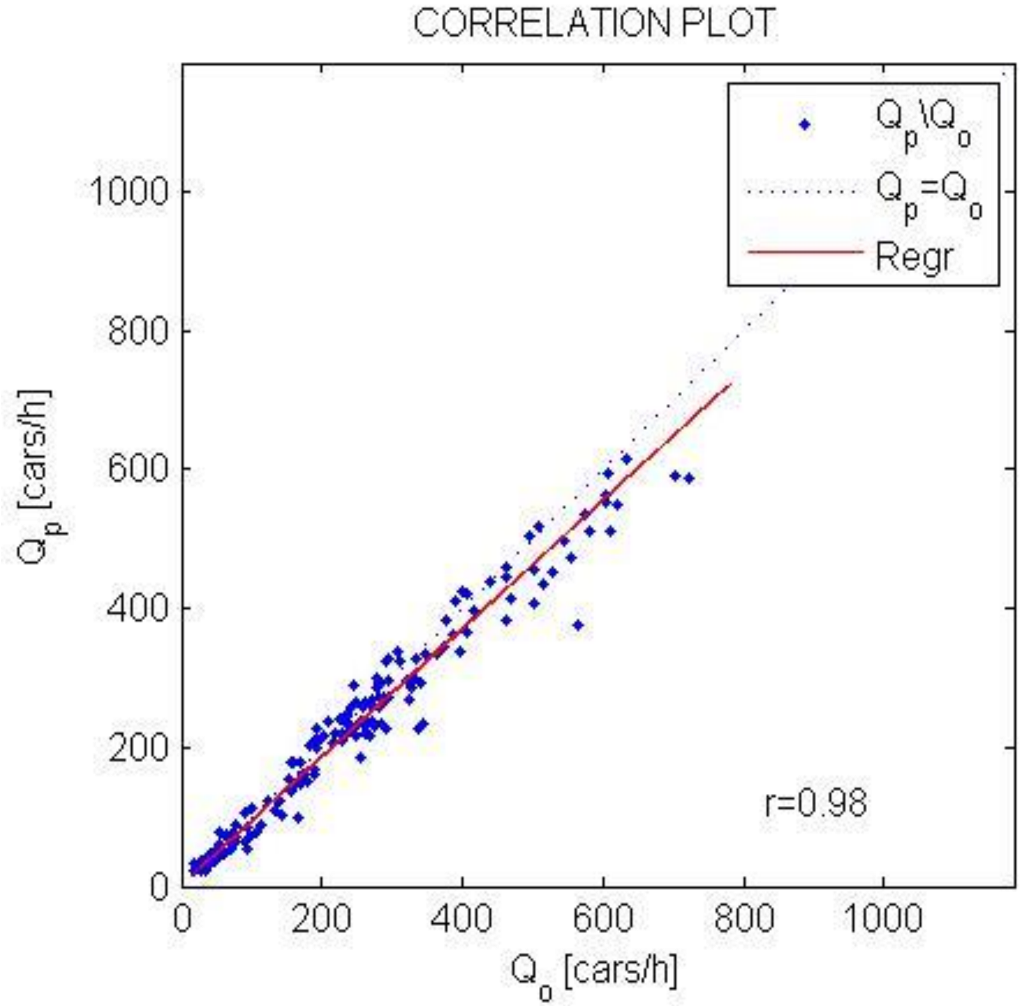
Encoding of days provides additional information for analysis



Good prediction of Q from the condition $\{D, H\}$ is obtained for normal days

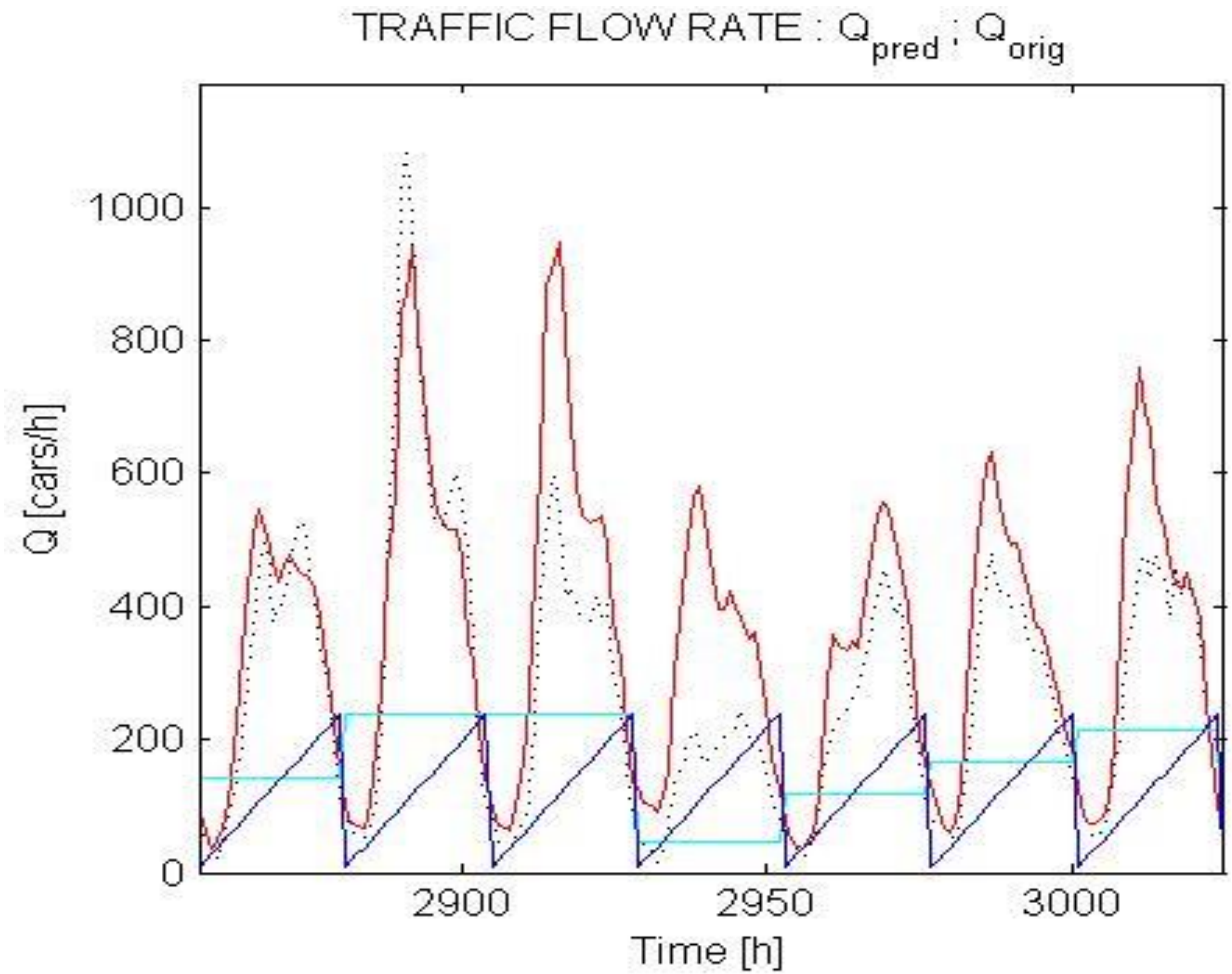


Comparison of predicted and observed flow rate in normal days

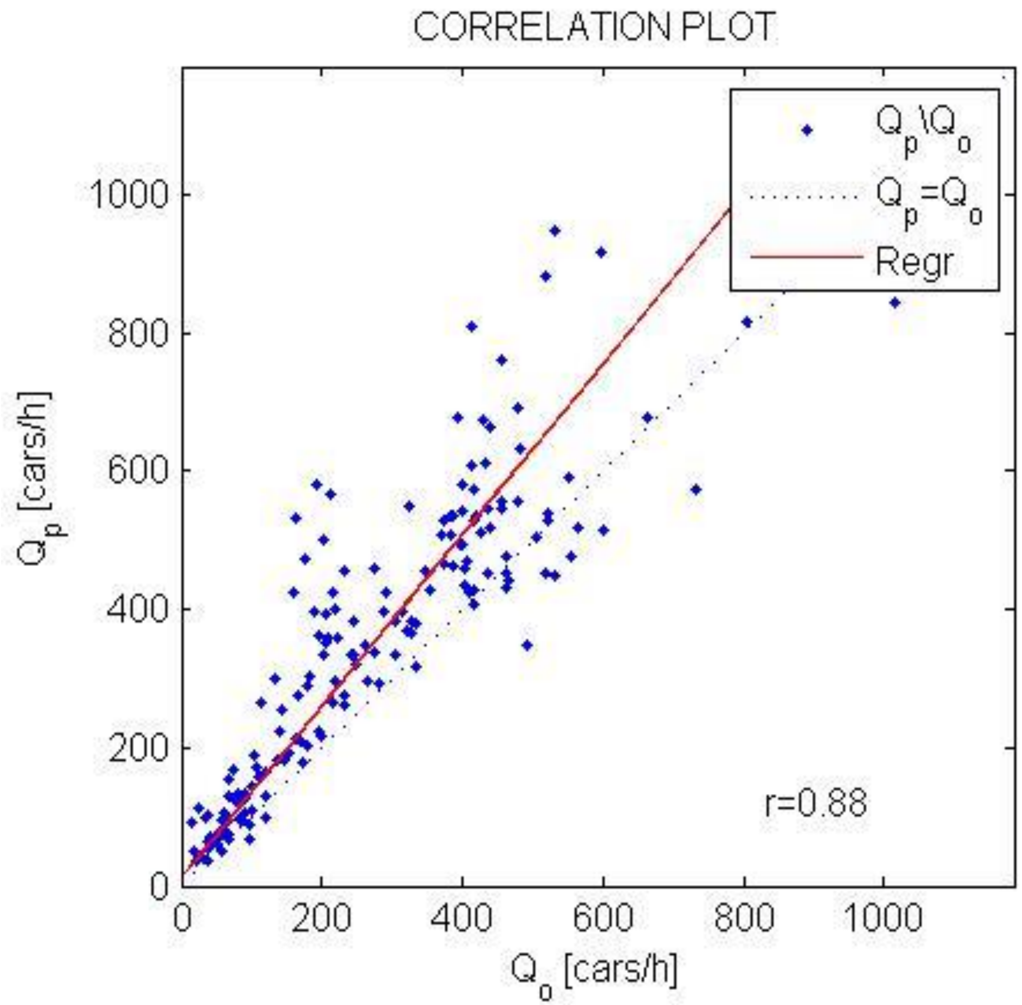


r = 0.98

Worse prediction from the condition {*D*, *H* } is obtained for holidays

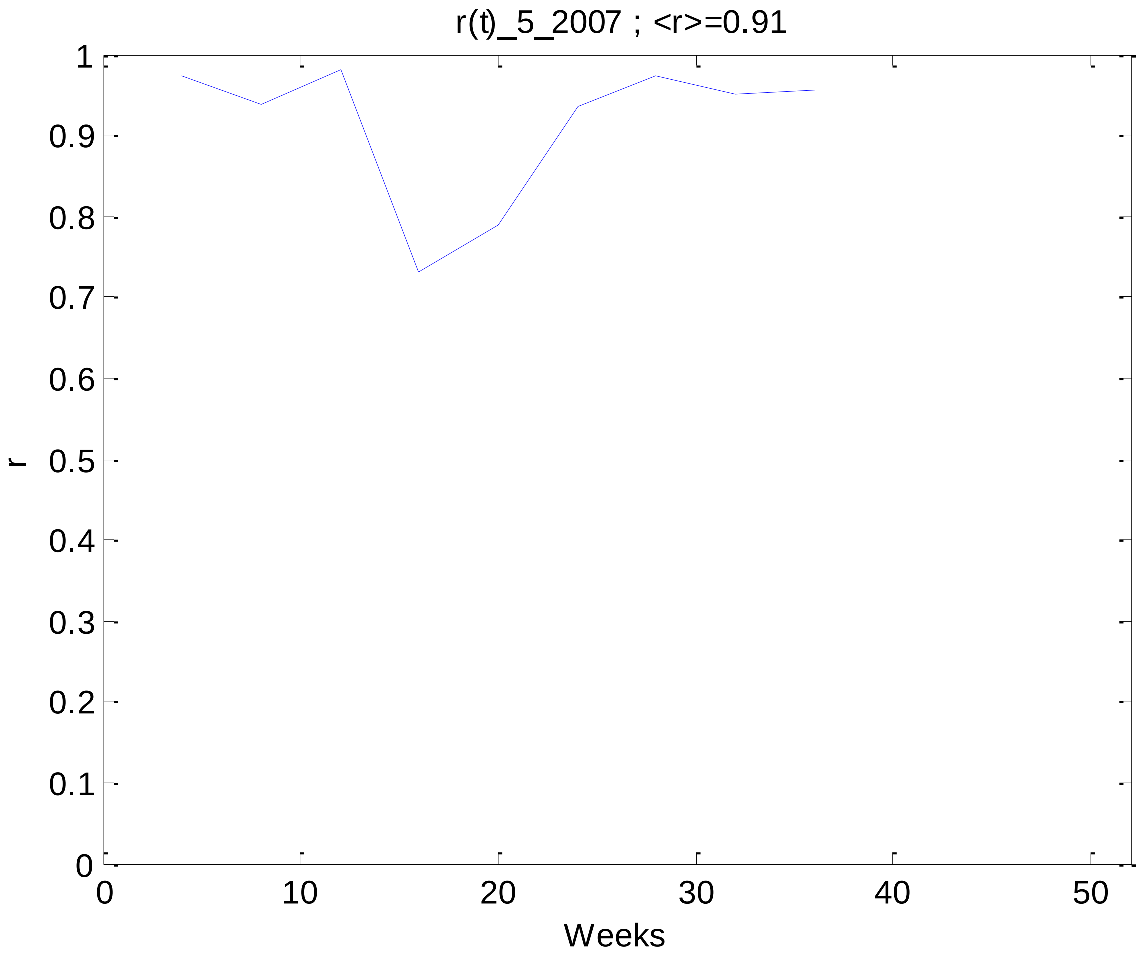


Comparison of predicted and observed flow rate in holidays

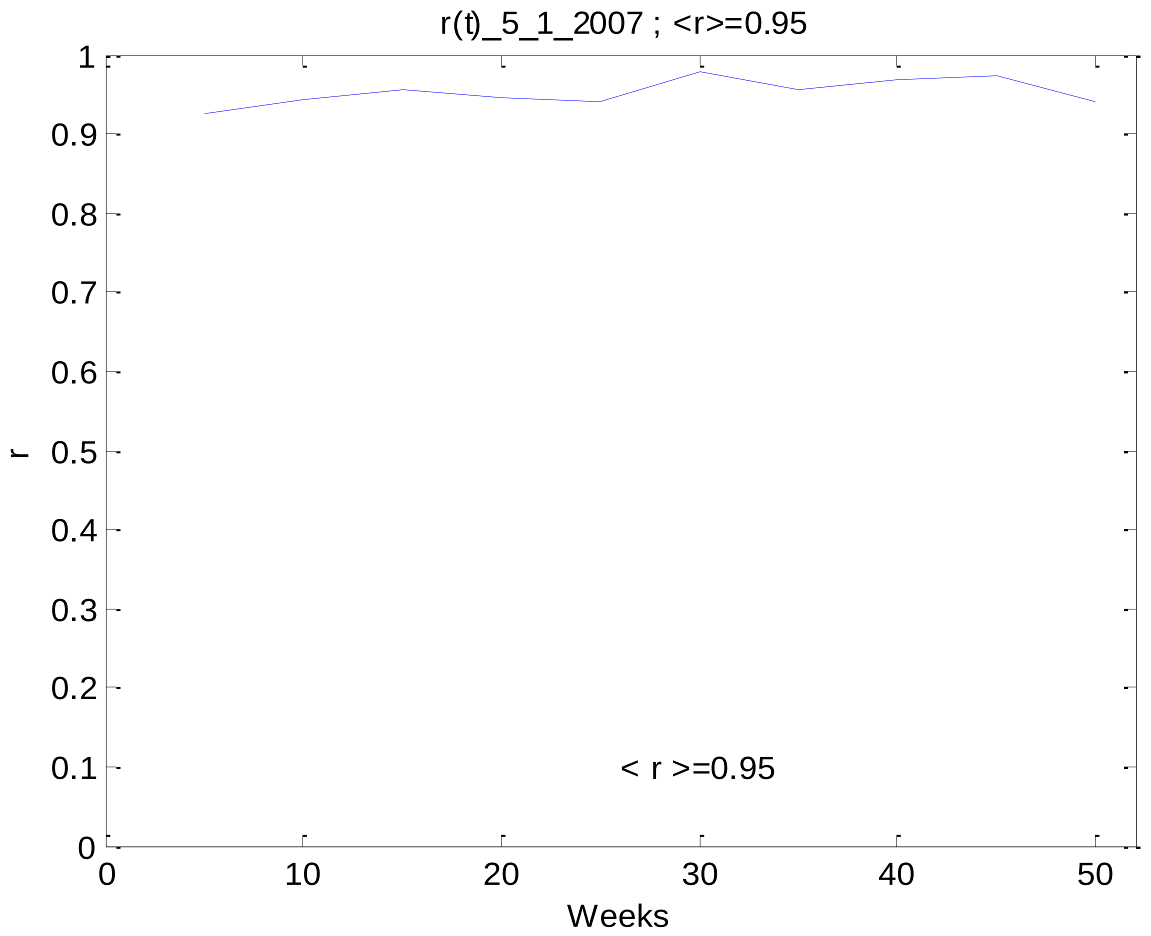


r = 0.88

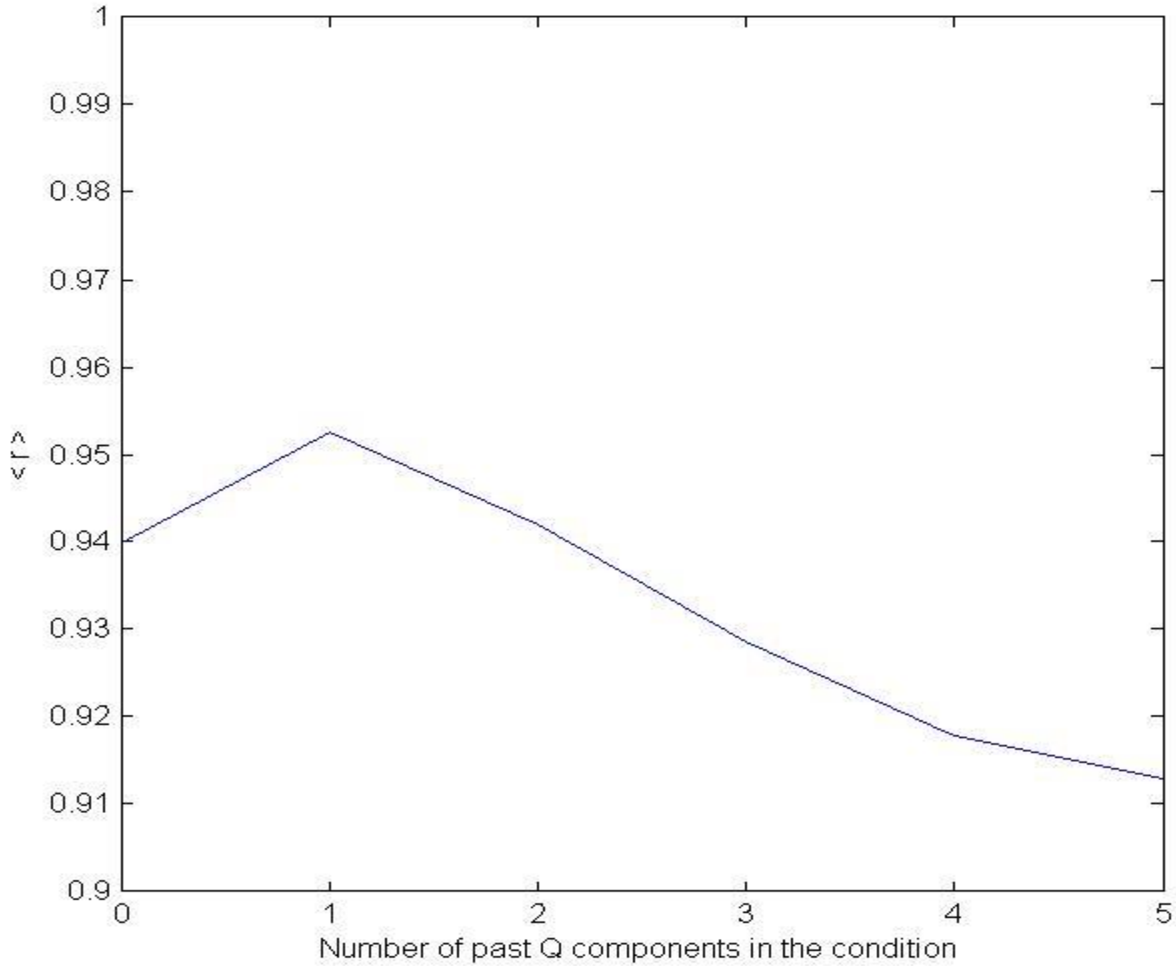
Correlation coefficient



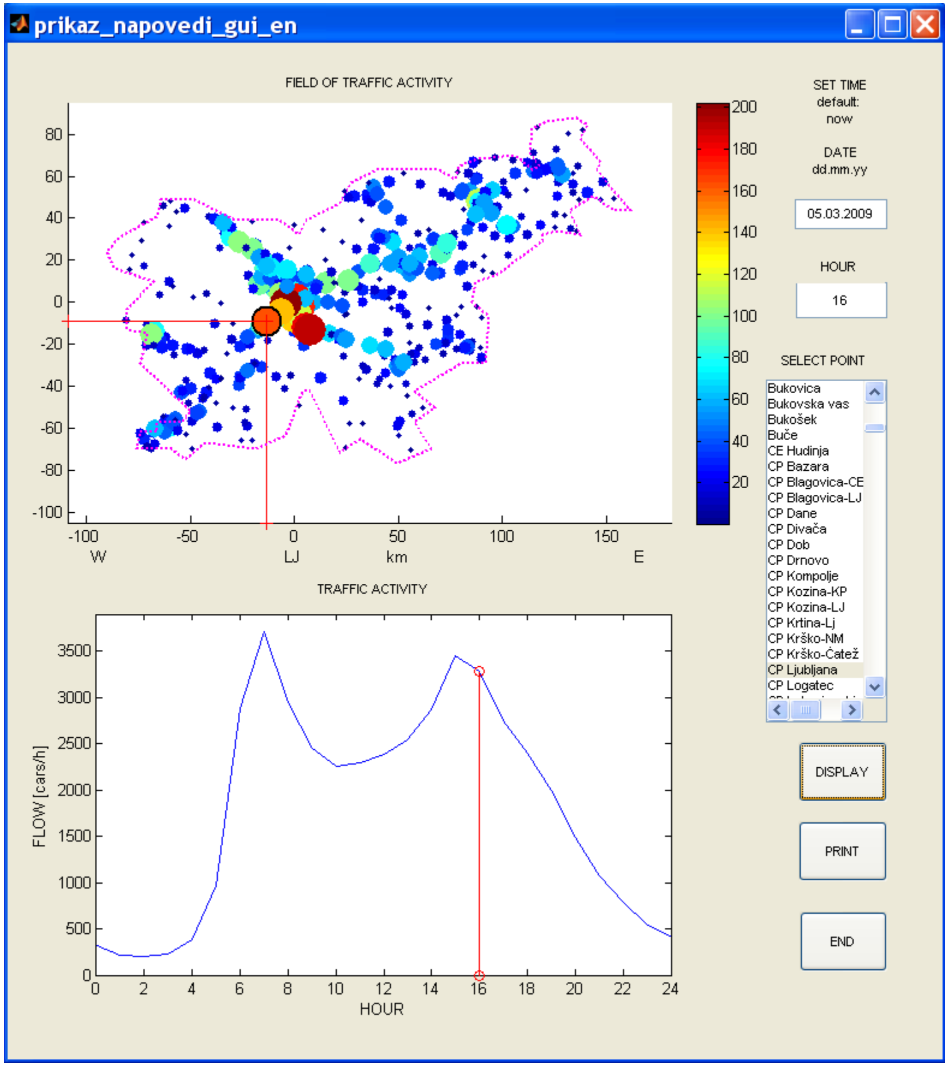
Merging of condition {*D(t), H(t)*} with past *Q* improves the prediction



Dependence of < *r* > on the number of past *Q* components in the condition



Graphic user interface for prediction of traffic flow



- User sets: the day, hour, and point of prediction.
- The field of traffic activity is displayed in the top graph.
- The predicted time series of traffic flow rate is displayed in the bottom graph.
- The selected place and hour of prediction are marked in graphs

- The next problem is to map the predicted flow to parameters of jams evolving due to various disturbances at critical regions!

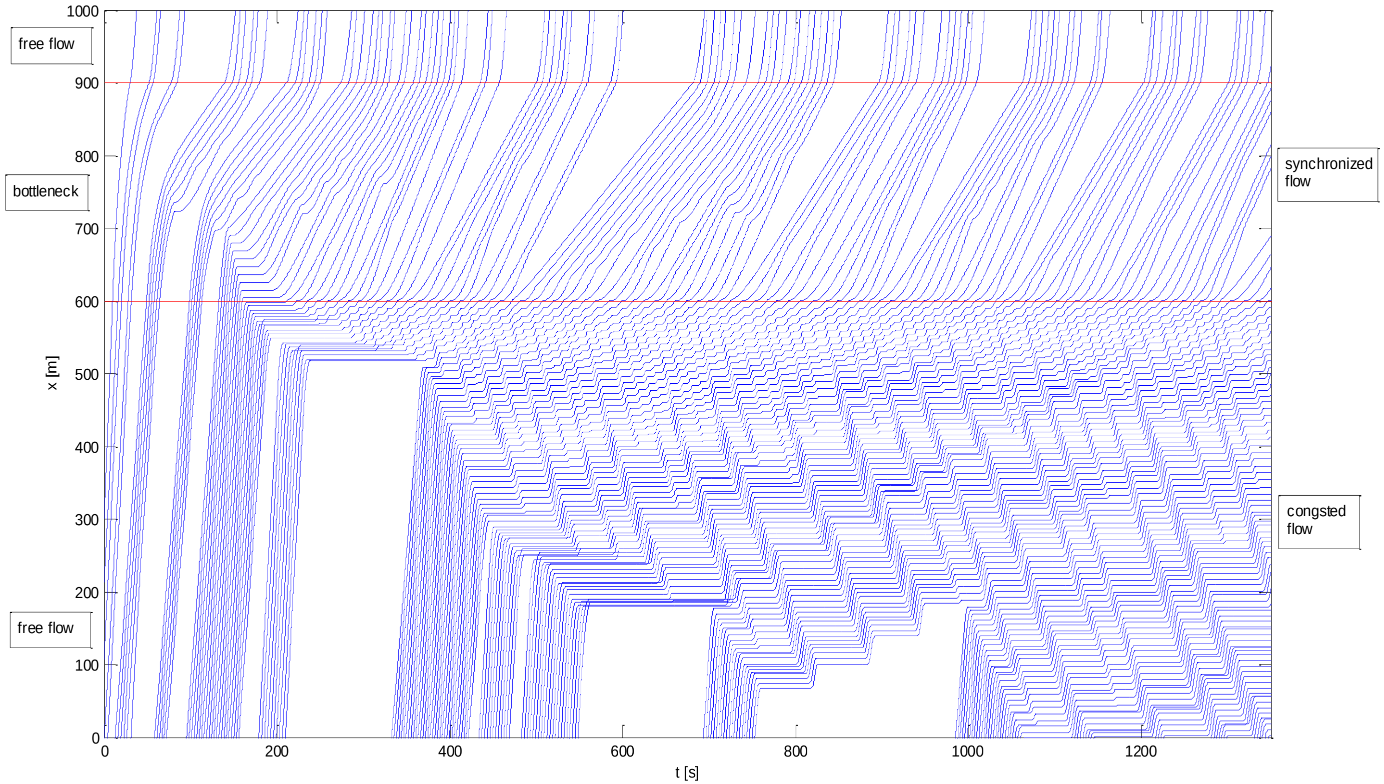
Forecasting of traffic jam

- **Problem:** Traffic jams on high-ways are developing due to various disturbances that decrease the road capacity. Our next aim is to describe a forecasting method.
- **Basic information:** Properties of the disturbance and the statistical data about the traffic flow rate.
- **Mathematical tool:** Statistical predictor of traffic flow rate and road capacity.
- **Goal:** To develop a program for transformation of predicted traffic flow rate to variables characterizing jam properties.

Micro-dynamic modeling of jam evolution at a bottleneck

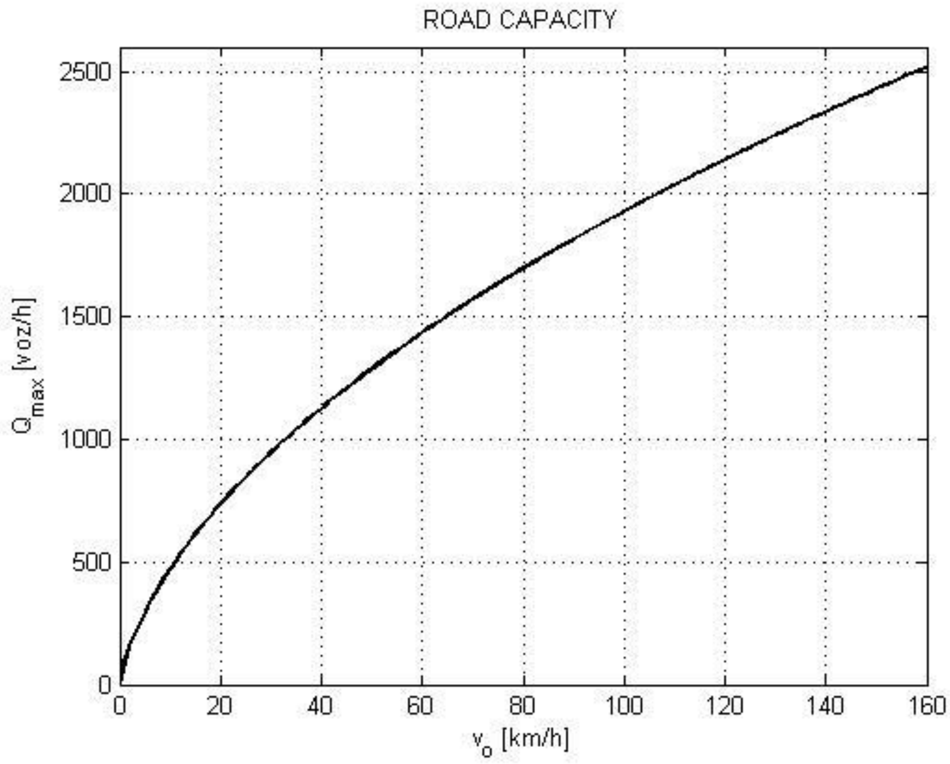
Micro-dynamic model is based upon driving rules and time series of traffic flow rate.

Micro-modeling is not convenient for application due to many trajectories.



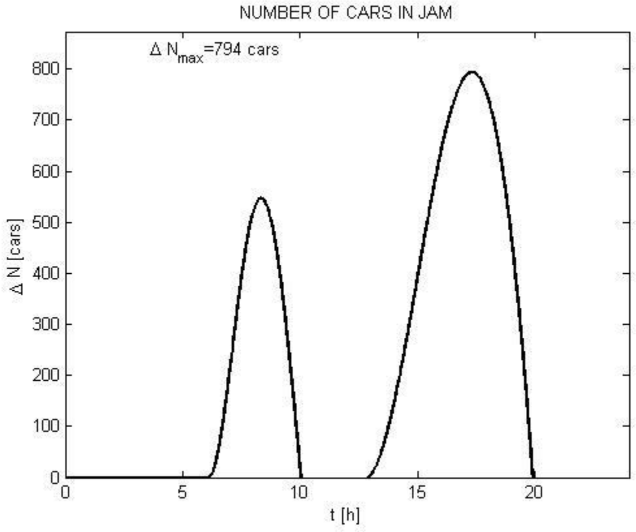
Hint: Apply macro-modeling by continuity equation in which a boundary condition is determined by the predicted traffic flow

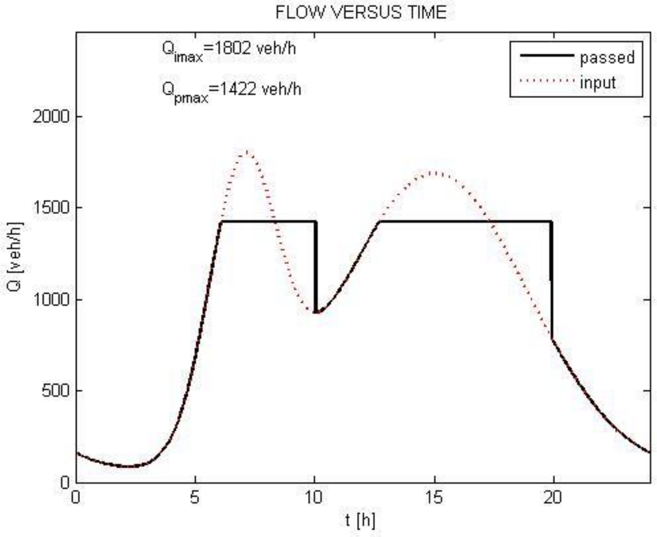
Road capacity Q_{max}



Dependence of road capacity Q_{max} on the speed limit v_o.

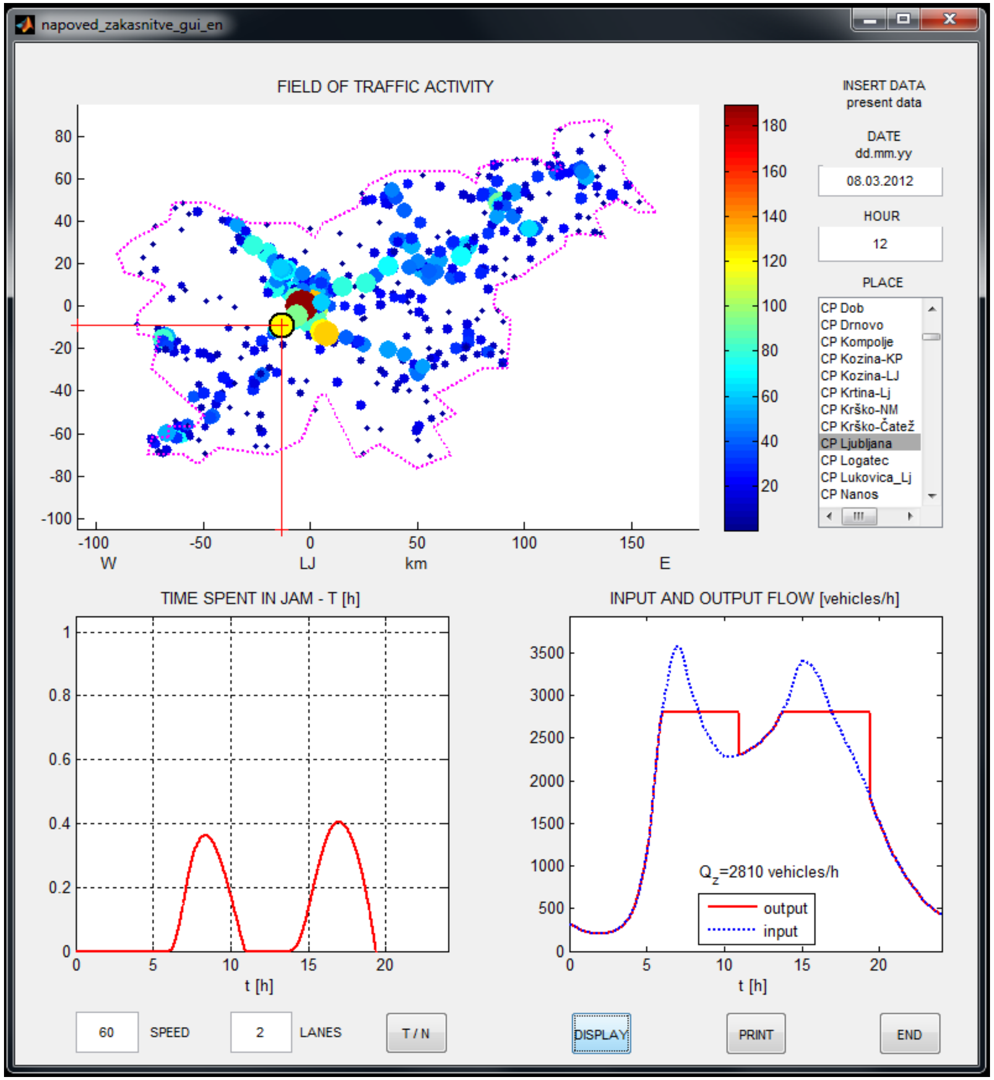
Example of jam estimation





- Jam properties are estimated from the given road capacity Q_{max} and the predicted input flow Q_{in} (right).
- When the input flow surpasses the road capacity: $Q_{in} > Q_{max}$, a jam starts to evolve - shown on the right.
- The number N of cars in a jam (left) is estimated by integrating the difference: Q_{in} - Q_{max} .

GUI for estimation of traffic jam properties



- User sets: the day, hour, and point of prediction.
- The field of traffic activity is displayed in the top graph.
- The predicted time series of traffic flow rate is displayed in the bottom right graph.
- User also sets: proper speed, number of lanes, and selects display of T or N in jam.
- The input and passed traffic flow rate are indicated in the right bottom diagram.
- The forecast evolution of jam is shown in the left diagram.

More advanced description of disturbance, traffic flow, and jam evolution

- From the properties of the disturbance a proper value of the speed limit can be estimated.
- The speed limit provides for the description of the equilibrium traffic by two fundamental laws.
- To describe a variable traffic at a disturbance one has to solve partial differential equations of traffic field.
- Homogenization of congested flow is possible by an optimal control of speed limit.
- More general problem of an optimal traffic control can be treated by the methods of intelligent control.

Properties of statistical modeling

- Demonstrated prediction of driving conditions needs no analytical model, but just measured data.

- The method provides information support for drivers and winter roads service.

Conclusions

- Demonstrated forecasting of driving conditions and traffic flows needs no analytical model, but just measured data.
- Statistical modeling by CA yields rather accurate prediction of traffic flow rate on a high-way
- Based upon the predicted flow rate and decreased road capacity the properties of traffic jams at disturbances on a high-ways can be forecast.

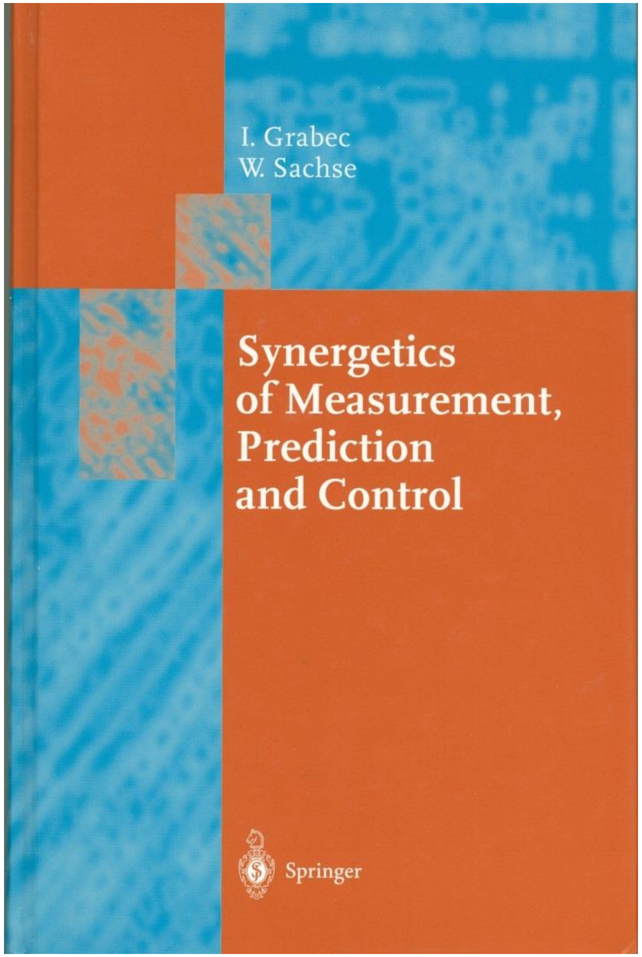
References

1. I. Grabec, W. Sachse, *Synergetics of Mesurement, Prediction nad Control*, Springer, Berlin, 1997.
2. I. Grabec, K. Kalcher, F. Švegl, *Modeling and Forecasting of Traffic Flow*, Nonlinear Phenomena in Complex Systems, **13**(1), 53-63, 2010.
3. I. Grabec, F. Švegl, *Statistical forecasting of high-way traffic jam at a bottleneck*, Advances in Methodology and Statistics, **9**(1), 81-93, 2012.
4. I. Grabec, F. Švegl, *Foerecasting of traffic jams on high-ways caused by adverse weather,* 16th International Road Weather Conference - SIRWEC , Helsinki, FI, May 23-25, 2012, ID:27
5. I. Grabec, F. Švegl, *Stabilization of traffic by speed limit variation*, ISEP 2014, Ljubljana, SI, March 24-25, 2014, R7.

Suggested reference for further work

Aimed at those interested in:

- adaptive and autonomous statistical modeling of natural laws from sensory signals,
- sensory-neural networks,
- intelligent and self-control,
- synergetics and informatics.



Stopping distance and friction coefficient

- For constant μ_o : $x_{st} = x_{react} + x_{break} = \tau v + v^2 / 2\,\mu_o\,g$;
- *τ – reaction time, g – acceleration of gravity.*
- Generally μ depends on velocity v as: $\mu(v) = \mu_o \exp(-v/c)$
- c – decay velocity: $c \cong 85$km/h
- Due to decay of $\mu(v)$ the stopping distance is increased:
- $x_{st} \cong \tau v + \exp(0.7v/c)\; v^2/2\,\mu_o\,g$
- A proper speed limit on slippery road is obtained by equalizing breaking distances at normal and adverse conditions: $x_{st1} = x_{st2}$
- A proper speed limit at decreased visibility is obtained by equalizing stopping and visibility distance: $x_{st2} = x_{vis}$

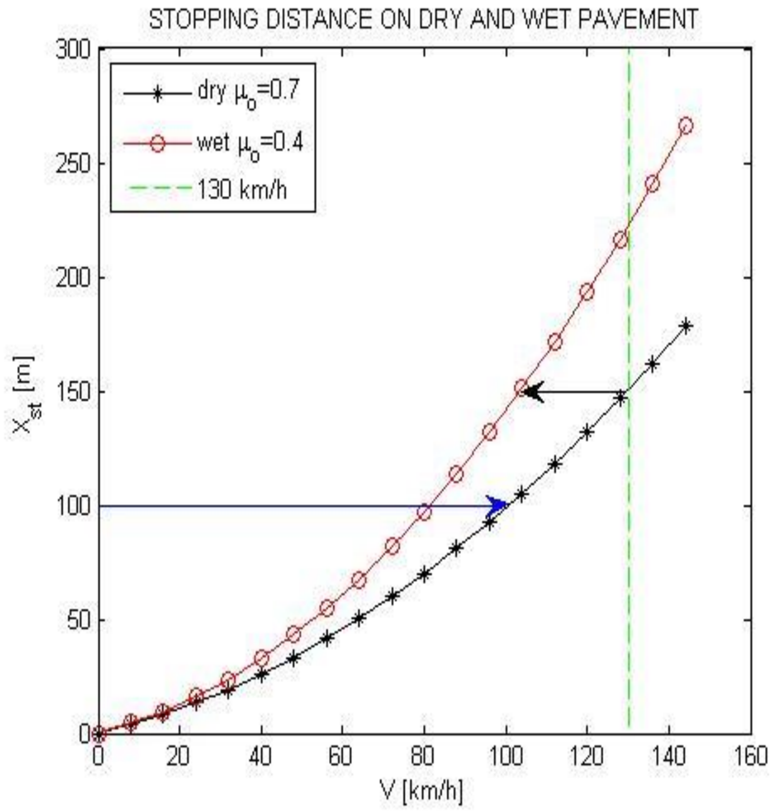
Estimation of speed limit from stopping distance characteristics

- A proper speed limit on a wet road is obtained by equalizing stopping distances on dry and wet pavement:

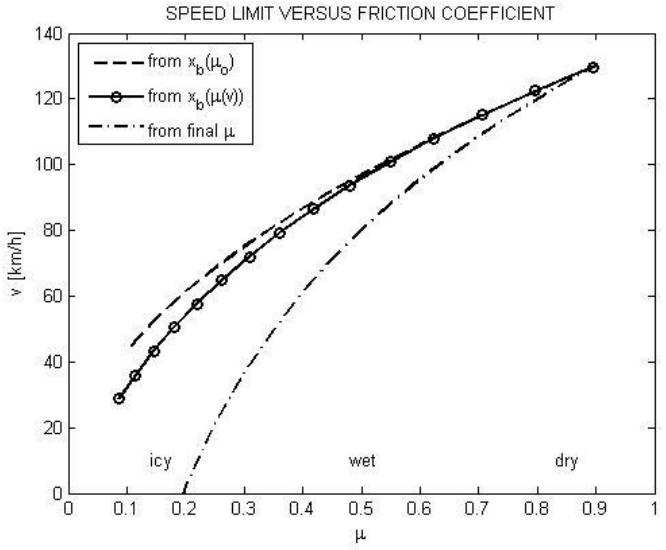
  $x_{st1} = x_{st2}$ - black arrow

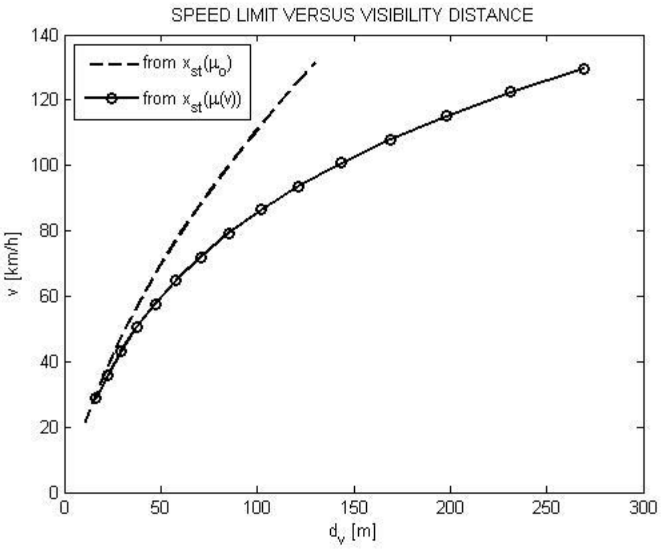
- A proper speed limit at decreased visibility is obtained by equalizing stopping and visibility distance:

  $x_{st2} = x_{vis}$ - blue arrow



Characteristics of proper speed limit





Dependence of proper speed limit on friction coefficient (left) obtained from various assumptions about breaking distance.

Dependence of proper speed limit on visibility distance (right) obtained from various assumptions about breaking distance.

From the speed limit the road capacity Q_{max} can be estimated.

Variables needed for macro-modeling of traffic dynamics

- Basic variables:
  - mean distance between cars: r
  - density of cars: $\rho = 1/r$
  - equilibrium velocity of cars: v_e
  - equilibrium flow rate: $Q_e = \rho v_e$
- Parameters and reference variables:
  - car length: $\lambda \cong 5\text{m}$, reaction time: $\tau \cong 1\text{s}$
  - speed limit: v_0 , speed reference: $u = \lambda / \tau$
  - clear spacing: $r - \lambda$, proper velocity: $w = (r - \lambda)/ \tau$

Equilibrium velocity

- Supposition: Equilibrium velocity v_e is smaller than the speed limit: $v_e \leq v_o$ and the proper velocity: $v_e \leq w$.
- Joint constraint: $1/v_e = 1/v_o + A/w$
- Observations yield the weight: $A \cong 3u/w$

  and the first fundamental law of traffic:

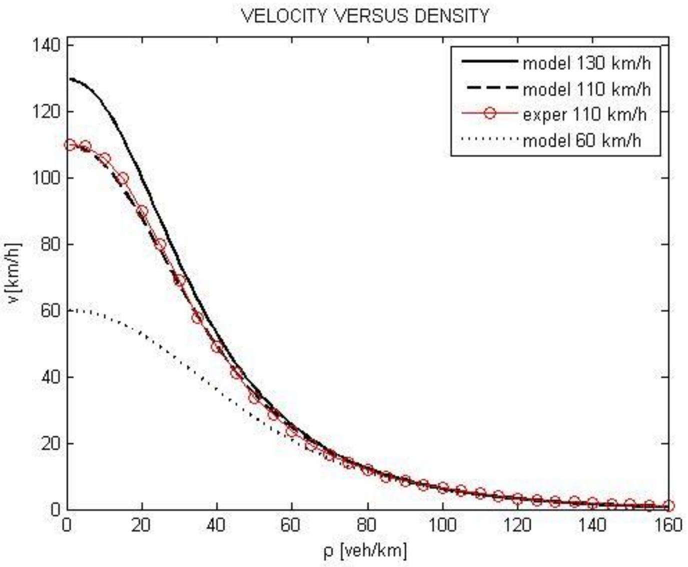
$$v_e = \frac{v_o}{1 + \frac{uv_o}{w^2}}$$

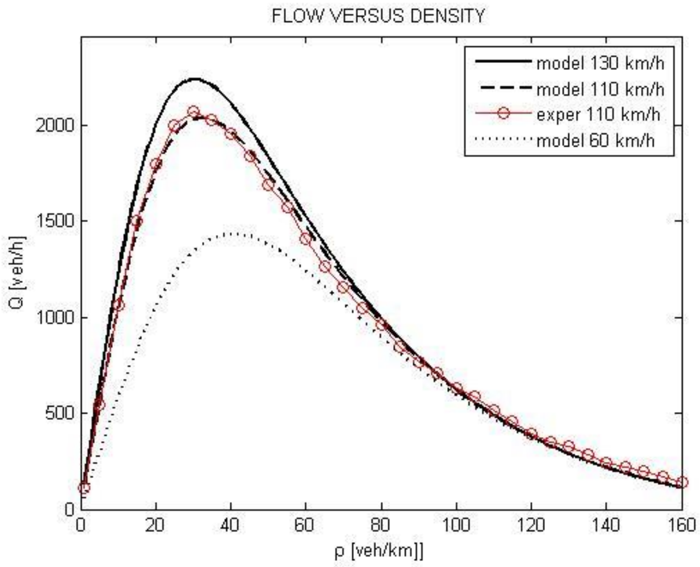
- From $u(\rho)$ and $w(\rho)$ we obtain: $v_e = v_e(\rho)$ and $Q_e(\rho) = \rho v(\rho)$

Fundamental diagrams of equilibrium traffic

$v_e(\rho)$

$Q_e(\rho)$





Equations of traffic field: $v(x,t), \rho(x,t)$

- Velocity adaptation law:

$$\frac{dv}{dt} = \frac{v_e - v}{T} \quad ; \qquad \text{relaxation time: } T \sim 3\tau$$

- Continuity equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = I$$

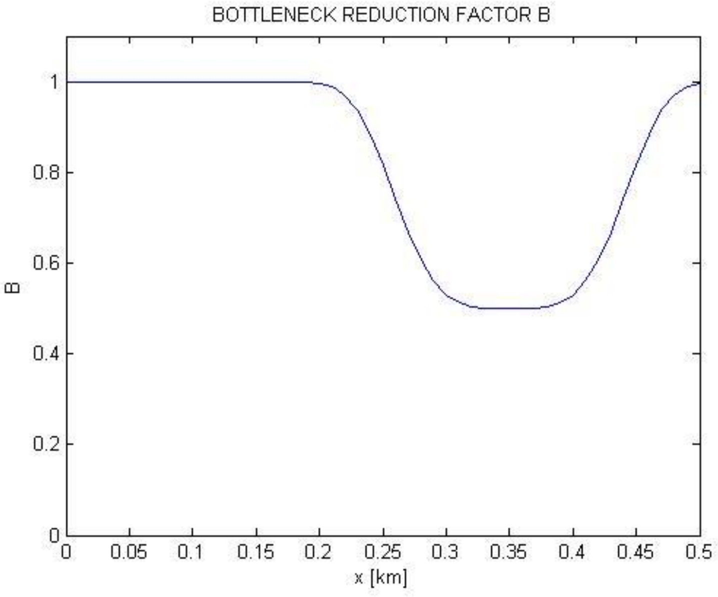
- Traffic source term:
- $I(x,t) = Q(t)\,\delta(x - x_o)$; $Q(t)$ is forecast

Numerical treatment

- Cell dimensions: $\Delta x = \lambda, \ \Delta t = 0.1\tau$
- Intervals: $0 < x < 0.5$km ; $0 < t < 1$h
- Initial and boundary conditions: $\rho=0$; $v=0$
- Source term specified by the forecast flow rate Q centered at rush hour: $t = 0.5$h.
- Transition to non-dimensional variables:
- $t/\tau \rightarrow \mathbf{T}$; $x/\lambda \rightarrow \mathbf{X}$; $v\,\tau/\lambda \rightarrow \mathbf{V}$; $\rho\,\lambda \rightarrow \boldsymbol{\rho}$; $Q\,\tau \rightarrow \mathbf{Q}$

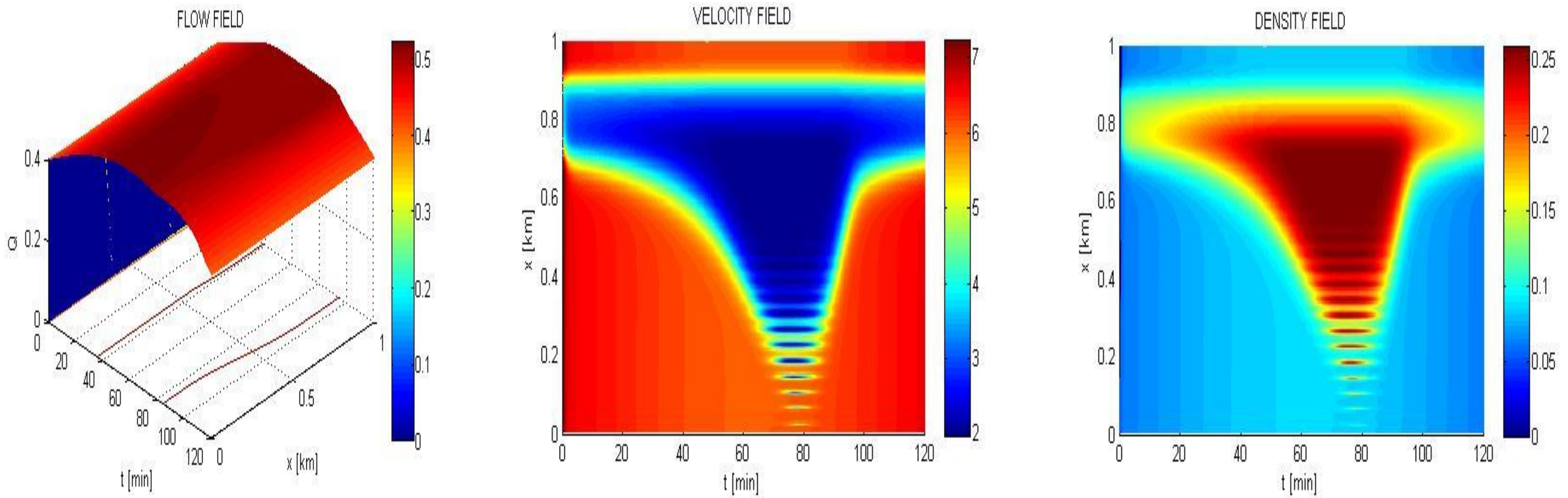
Specification of a bottleneck

- Position: $0.2\text{km} < x < 0.4\text{km}$
- Reduced speed: $0.5\ v_o$



Dependence of the velocity reduction factor *B* on *x.*

Field distributions



Parameters: v_o=130 km/h ; Q_{max}=1875 veh/h

Application of jam forecasting

Observation: The evolution of traffic jam at the bottleneck critically depends on the input flow. Forecasting of its properties is possible based upon the predicted flow.

Advice: Before installing a bottleneck one can estimate its influence and diminish traffic disturbance by a proper adaptation of the bottleneck structure.

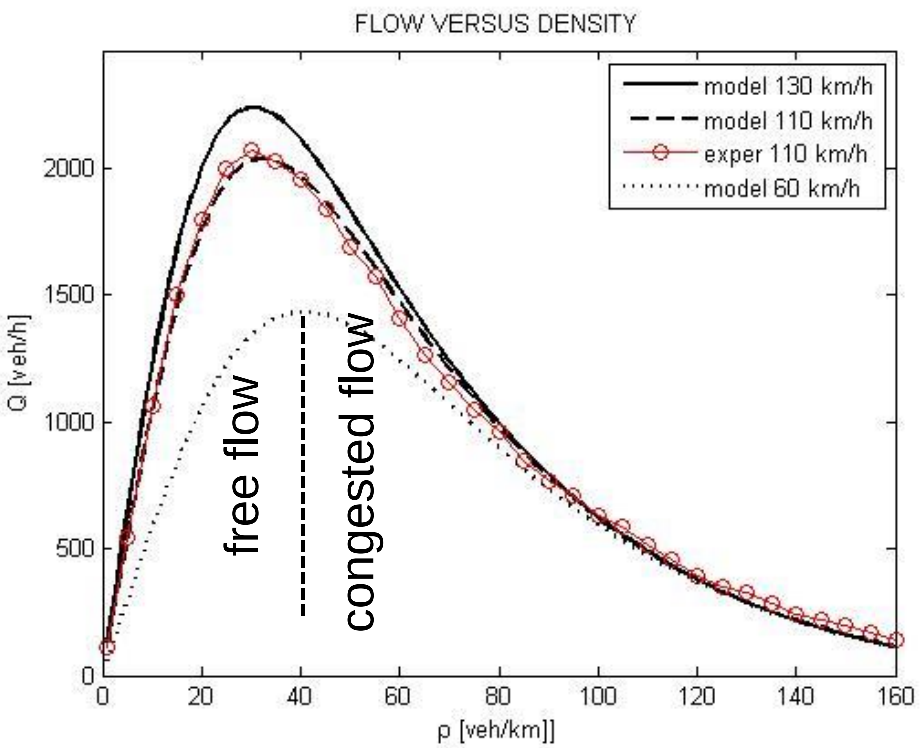
STABILIZATION OF TRAFFIC FLOW BY THE SPEED LIMIT CONTROL

Basic properties of congested traffic

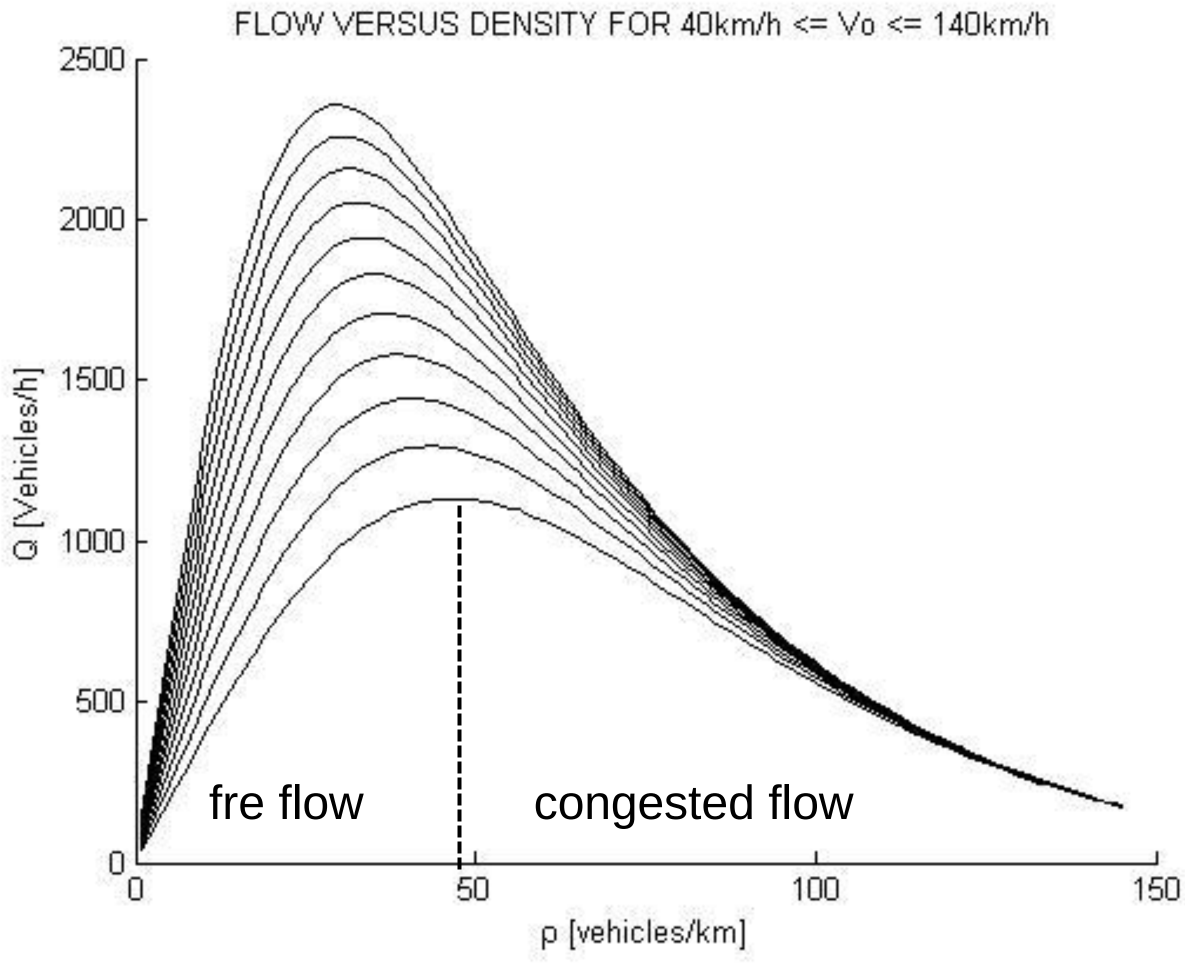
- The maximum of traffic flow Q_{max} takes place at some optimal value of density ρ_{max}.
- The congested traffic at $\rho > \rho_{max}$ is subject to dynamic instability that causes evolution of moving jams.
- These jams diminish the road capacity and lead to detrimental economic consequences.
- Consequently, the basic problem of an optimal control of congested traffic is to avoid the instability and so to assure homogeneous state.

An optimal control of speed limit

At each density ρ it is reasonable to set the value of speed limit v_0 so that the traffic state corresponds to the maximum of flow Q. In this case the state is stable and moving jams do not develop.



Family of characteristics $Q(\rho, v_o)$

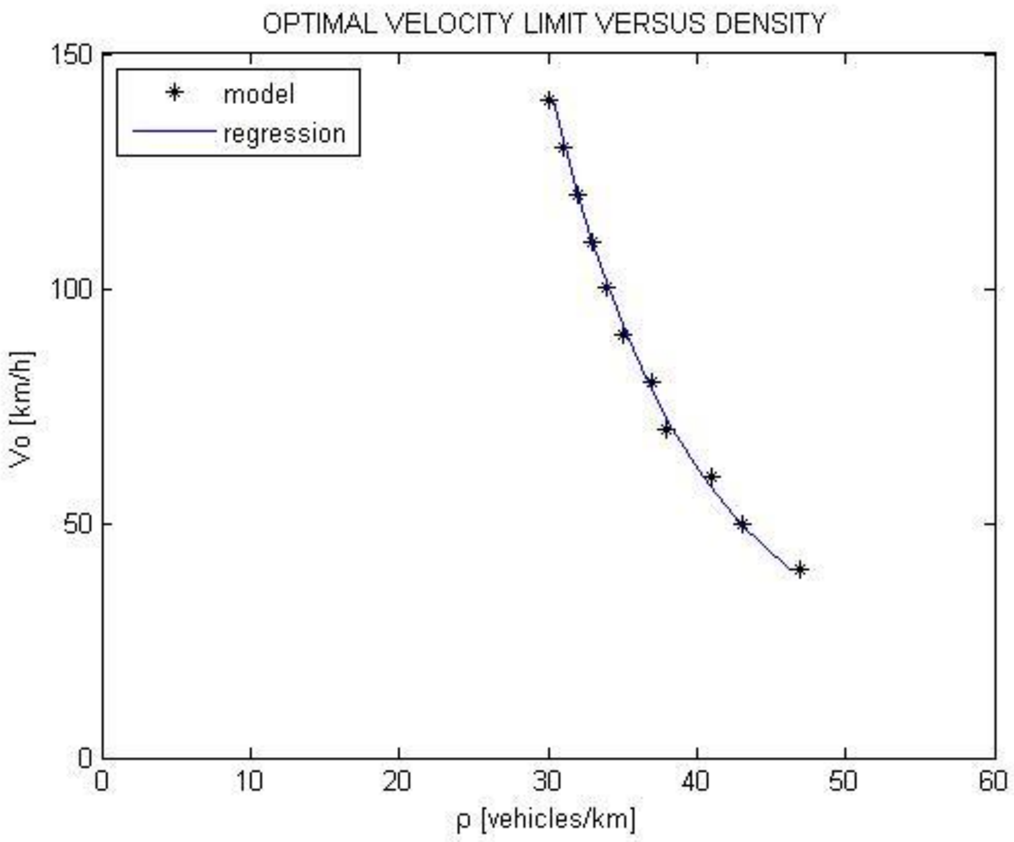


Characteristic of the optimal speed limit

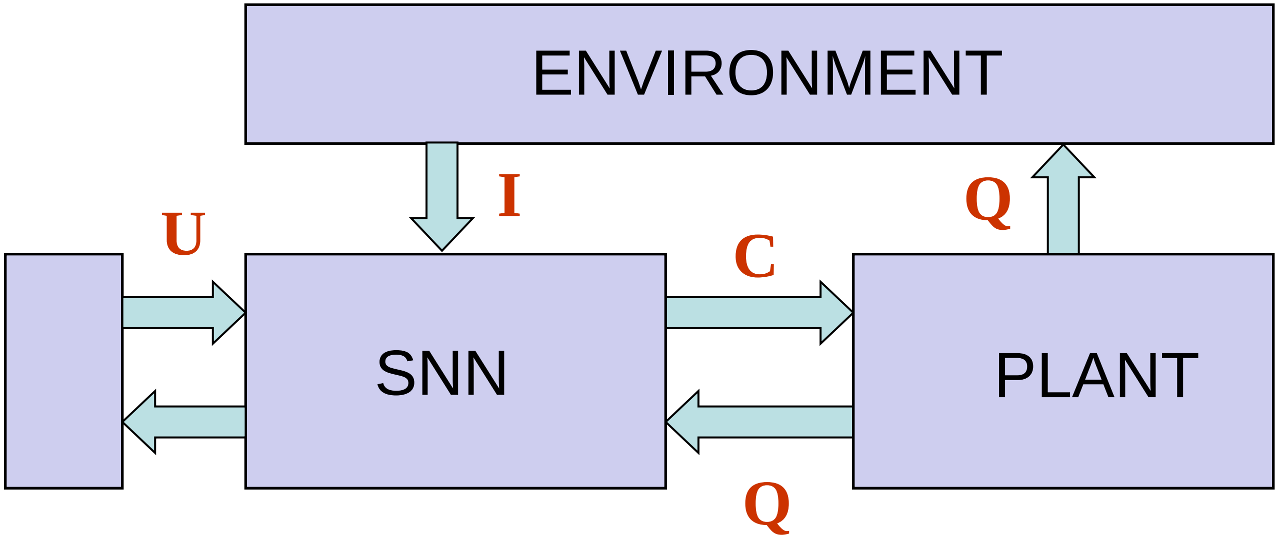
From the family of characteristics $Q(\rho, v_0)$ we obtain the relation between density and the optimal speed limit. (right)

A paradox !?!

With an increasing density ρ the value of the optimal speed limit v_0 must be properly decreased if we want to keep homogeneous traffic state without moving jams!



More general case in changing environment and self-controlled system



I – input, **U** – utility, **C** – control, **Q** – plant state

Basic problem of intelligent control

- Dynamics of the system is determined by:

  $dQ/dt=F(Q,C)$, but not known !!!
- Problem: Find the control $C=C(I,Q)$

  that optimizes the system utility: $U=U(I,Q,C)$
- Hint:
- 1) Apply given examples and use general regression for modeling of functions: F, C, U
- 2) Find the optimal utility by reinforcement learning

Conclusions about the traffic control

- Homogenization of congested flow is possible based upon an optimal control of speed limit.

- More general problem of an optimal traffic control can be treated by the methods of intelligent control.

Prediction of air pollution
ARPV data about PM10

Microsoft Excel - Arpv_data

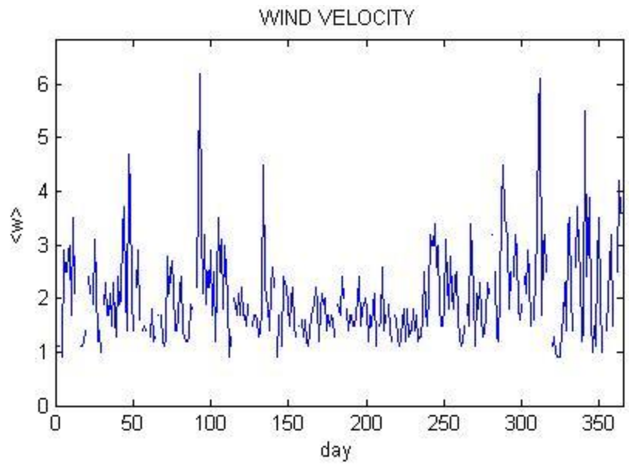
File Edit View Insert Format Tools Data Window Help

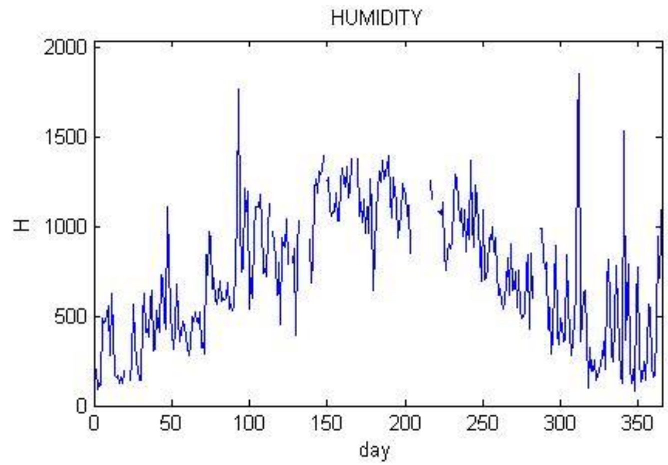
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1		Percentage of wind calm <0.5 m/s	Average wind speed	Average wind speed (without values<0.5 m/s)	Average wind direction	Hmix rural	Hmix urban	Stanford Index	Average radiation	T minimum	Average T	Total precipitation	Equavalent potetial temperature	Concentration 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

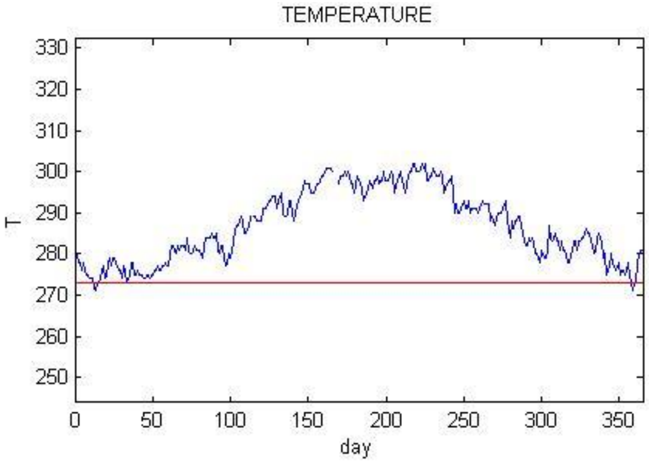
Selection of variables used in modeling predictor of PM10

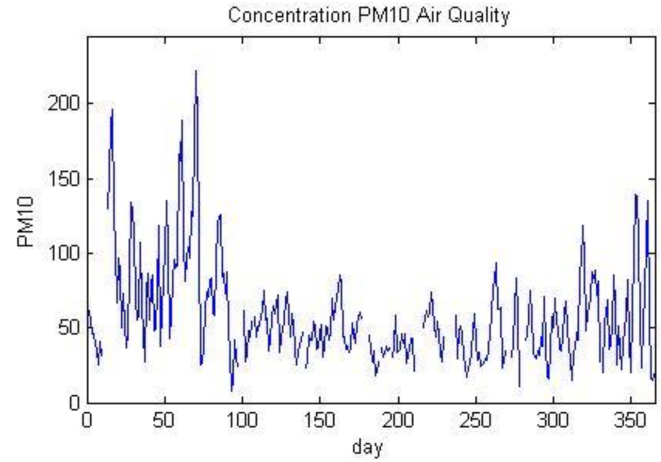
- As given variables we consider: the average wind velocity – *W*, humidity – *H* and average temperature – *T.*
- As hidden variable we consider concentration *P=PM10.*
- Using sample vectors $\mathbf{Z}_n = (W,H,T,P)_n$ *from* the recorded data base we create statistical model of the relation *P=G(W,H,T)* 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 *n.*
- Agreement between predicted and measured data is described by the correlation coefficient *r* and shown in correlation diagram.

Records of variables used in modeling

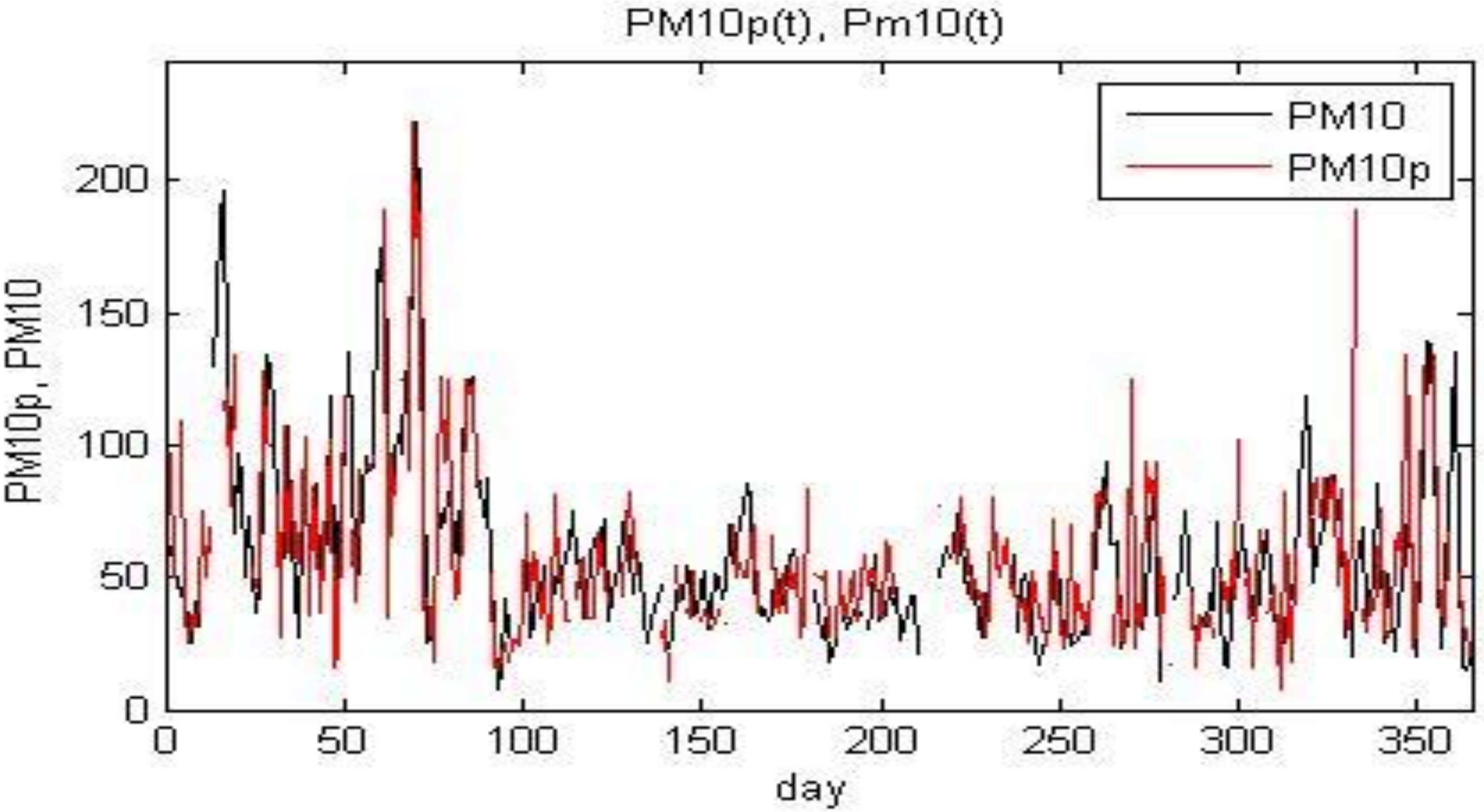




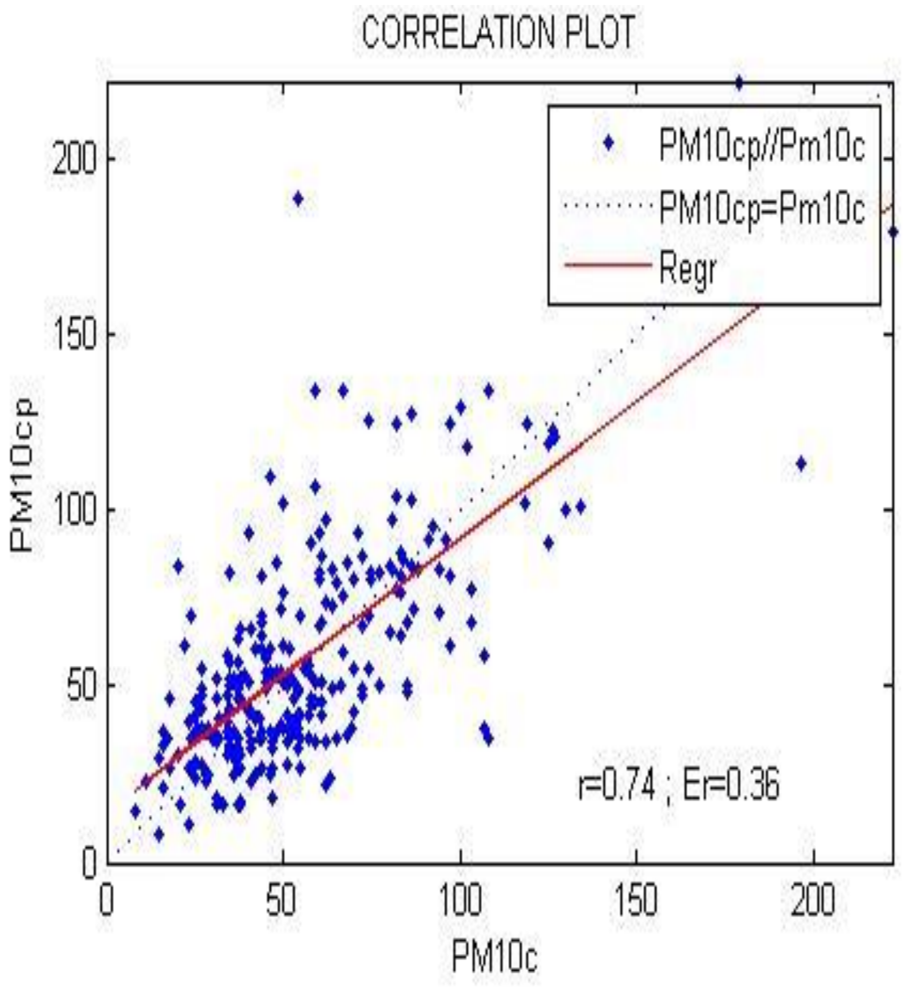




Predicted and observed PM10



Correlation plot of predicted and observed PM10

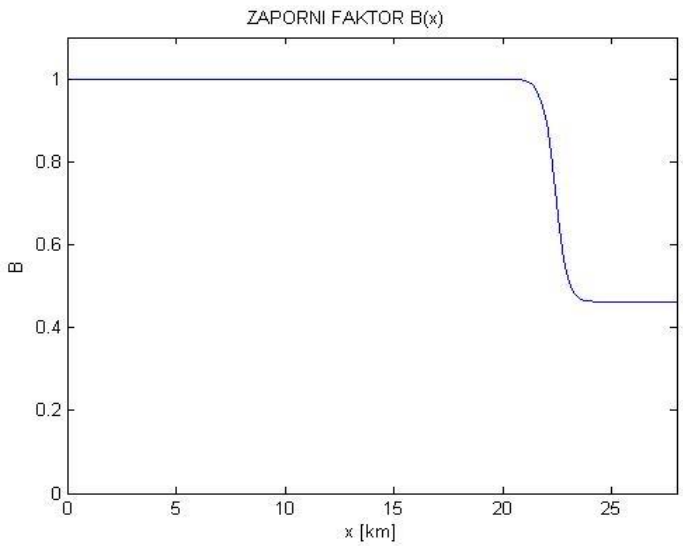


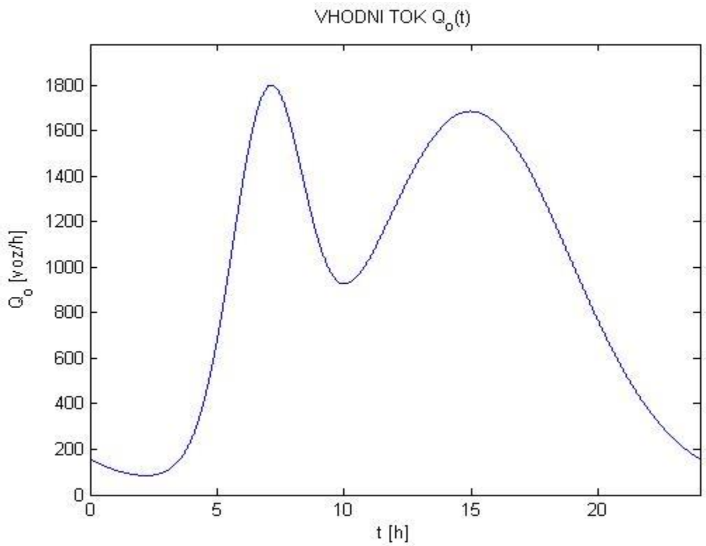
Properties of statistical modeling

- Demonstrated prediction of driving conditions needs no analytical model, but just measured data.

- The method provides information support for drivers and winter roads service.

Disturbance on a road sector



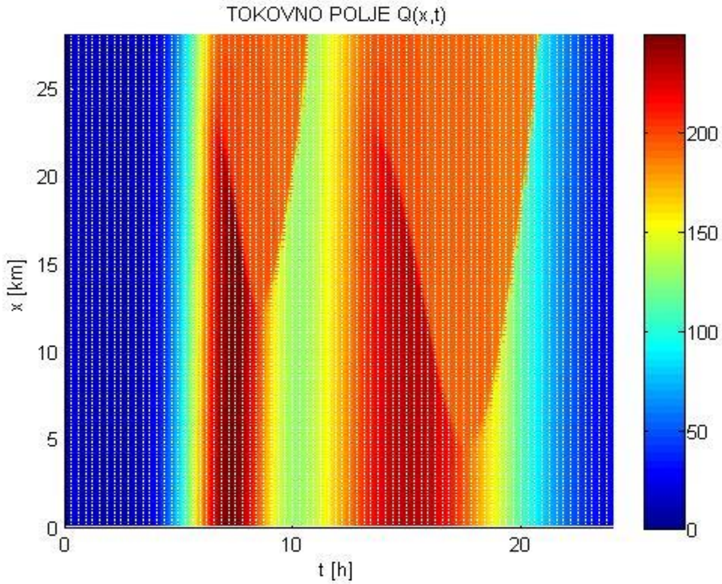


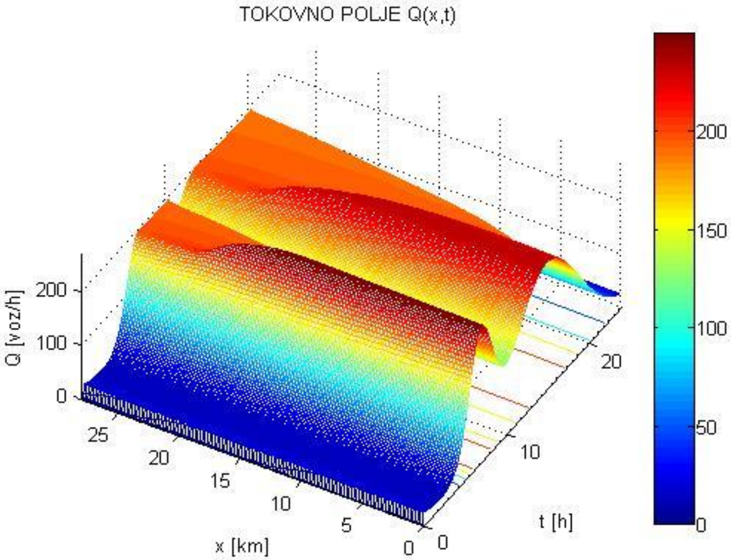
Dependence of the velocity reduction factor *B* on *x*

Dependence of input flow rate *Q* on time

Parameters: v_o =130km/h , Q_{max}=1875veh/h

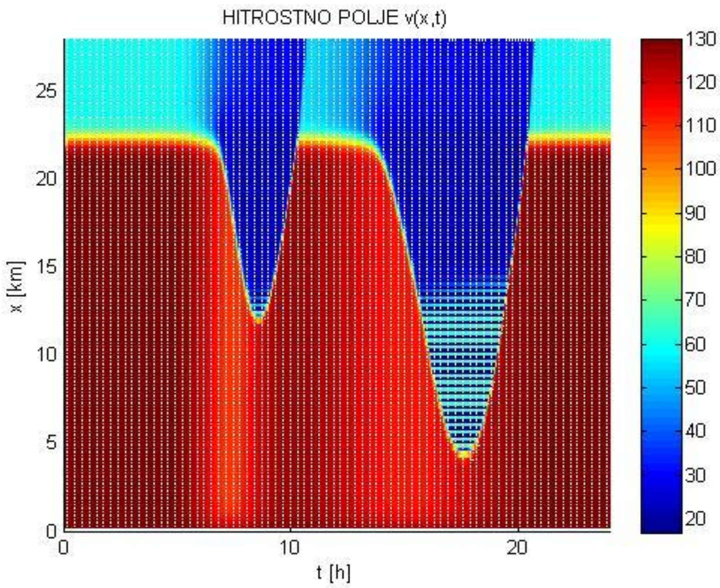
Traffic flow field

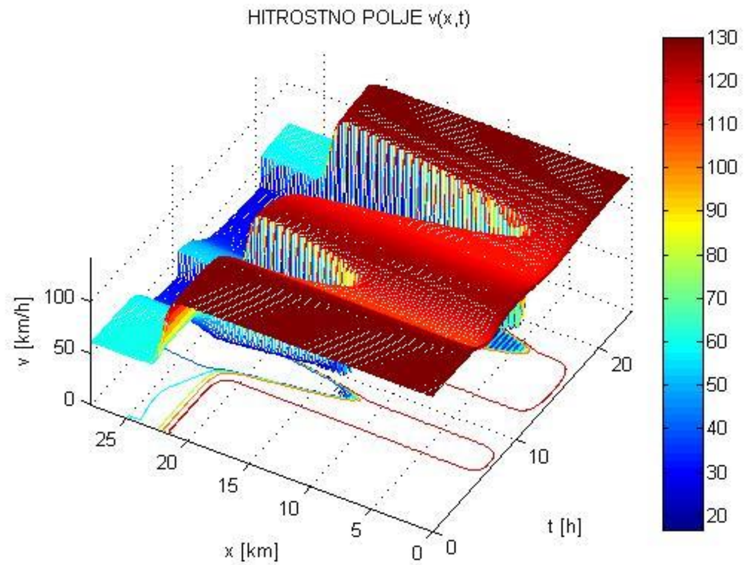




Distribution of traffic flow field
left – ground plan, right – side view.

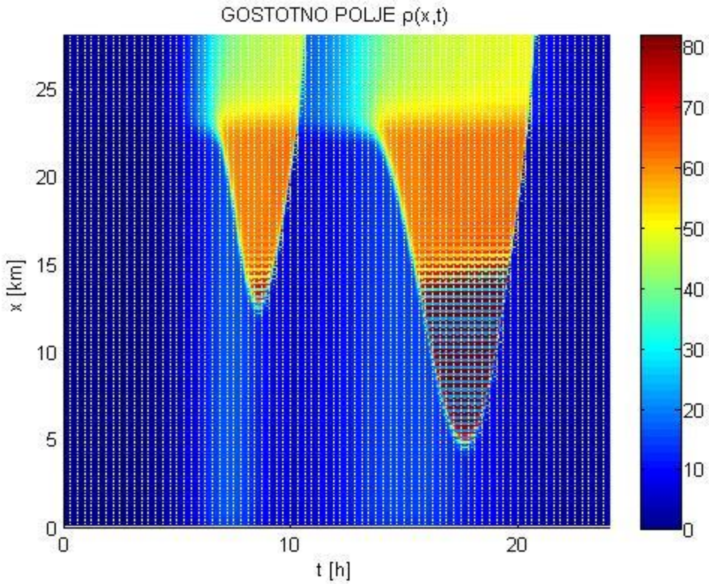
Traffic velocity field

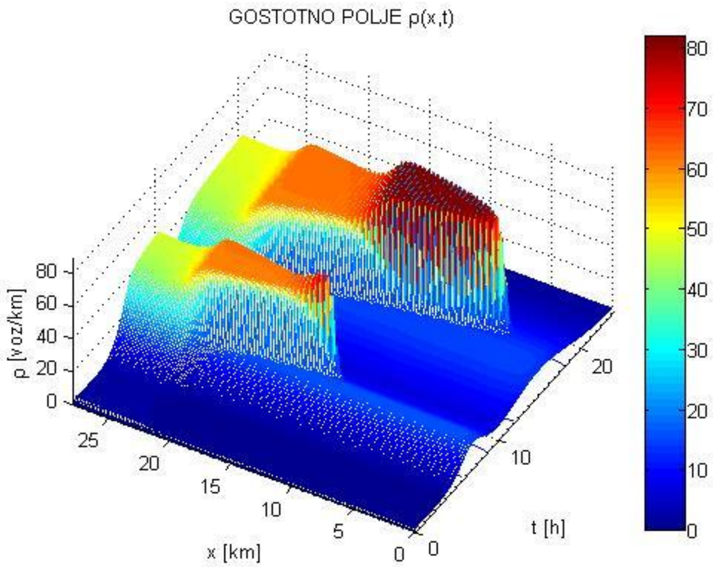




Distribution of traffic velocity field
left – ground plan, right – side view.

Traffic density field





Distribution of traffic density field
left – ground plan, right – side view.