Eperimental modeling: learning models from data

a user point of view

Mario Milanese

Dip. di Automatica e Informatica

Politecnico di Torino

The Logic of Modeling

Alta Scuola Politecnica

Milano, April 21, 2006

Outline

- Models as tools for making inferences from system data

  prediction, simulation, control, filtering, fault detection

- Model structures

  physical law based, input-output description, linear, nonlinear

- Model estimation

  statistical/parametric, set membership, structured

- Model quality evaluation (vs. model validation)

- Application examples
  - ✓ ***Prediction of atmospheric pollution***
  - ✓ ***Simulation of dam crest dynamics***
  - ✓ ***Identification of vehicles with controlled suspensions***

Regression form of system representation

- System S^o produces ***output signal y*** when driven by ***input signal u*** :

- Output y is related to input u by the regression function f^o :

$$y^{t+1} = f^o(w^t)$$

$$w^t = [y^t \cdots y^{t-n_y} u_1{}^t \cdots u_1{}^{t-n_{u1}} \, u_2{}^t \cdots u_2{}^{t-n_{u2}} \cdots]$$

Regression form of system representation

- Linear system ⇨ f^o is linear in w^t :

$$y^{t+1} = a_o y^t + a_1 y^{t-1} \cdots + a_{n_y} y^{t-n_y} + b_o u^t + b_1 u^{t-1} \cdots + b_{n_u} u^{t-n_u}$$

ARMA system

- If n_y=0 : MA (FIR) system
- If n_u=0 : AR system
- If f^o nonlinear : NARMA, NFIR, NAR systems

Making inferences from data

- It is desired to make an inference on system S^o :

  prediction, identification, simulation, control, filtering, fault detection

- The system S^o is unknown, but a finite number of noise corrupted measurements of y^t, w^t are available:

$$\tilde{y}^{t+1} = f^o(\tilde{w}^t) + d^t, \quad t = 1, \cdots, T$$

  d^t accounts for errors in data $\tilde{y}^t, \tilde{w}^t$

- The inference is described by the operator $I(f^o, w^T)$

  - *one-step prediction* ➡ $I(f^o, w^T) = f^o(w^T)$
  - *identification* ➡ $I(f^o, w^T) = f^o$

Making inferences from data

- Problems :
  - *for given estimates* $\hat{f} \simeq f^o, \hat{w}^T \simeq w^T$

    evaluate the inference error $\left\| I(f^o, w^T) - I(\hat{f}, \hat{w}^T) \right\|$
  - *find estimates* $\hat{f} \simeq f^o, \hat{w}^T \simeq w^T$

    "minimizing" the inference error

- The inference error cannot be exactly evaluated since f^o and w^T are not known

- Need of prior assumptions on f^o and d^t for deriving finite bounds on inference error

Model structures

- The model is described by:

$$\tilde{y}^{t+1} = f(\tilde{w}^t) + d^t$$

$$\tilde{w}^t = [\tilde{y}^t \cdots \tilde{y}^{t-n_y} \tilde{u}_1{}^t \cdots \tilde{u}_1{}^{t-n_{u1}} \tilde{u}_2{}^t \cdots \tilde{u}_2{}^{t-n_{u2}} \cdots]$$

- Model structure is defined by:
  - type of function ***f***
  - type of noise ***d***
  - which inputs $\boldsymbol{u_1, u_2, \ldots}$
  - lag values $\boldsymbol{n_y, n_{u1}, n_{u2}, \ldots}$

Statistical/parametric approach
Model structures

- Typical assumptions in literature:
  - on system: $f^o \in \mathcal{F}(\theta) = \left\{ f(w,\theta) = \sum_{i=1}^{r} \alpha_i \sigma_i(w,\beta_i) \right\}$
  - known lag values $n_y, n_{u1}, n_{u2}, \ldots$
  - on noise: iid stochastic noise
- Functional form of $F(\theta)$ required:
  - derived from physical laws
  - σ_i: "basis" function (polynomial, sigmoid,..)
- Parameters θ are estimated by optimizing Least Squares (LS) or Maximum Likelihood functionals

Statistical/parametric approach
Model structures

- If possible, physical laws are used to obtain the parametric representation of $f(w, \theta)$
- When the physical laws are not well known or too complex, input-output parameterizations are used

"Fixed" basis parametrization

Polinomial, trigonometric, etc.

"Tunable" basis parametrization

Neural networks, wawelets , etc.

often called black-box models

Statistical/parametric approach
Model structures: "fixed" basis

$$f(w,\theta) = \sum_{i=1}^{r} \alpha_i \sigma_i(w) \qquad \theta = [\alpha_1 \cdots \alpha_r]'$$

$\sigma_i(w)$: "Basis"

- **Problem:** Can σ_i's be found such that

$$f(w,\theta) \xrightarrow[r\to\infty]{} f^o(w) \qquad ?$$

Statistical/parametric approach
Model structures: "fixed" basis

- For continuous f^o, bounded $W \subset \Re^n$ and σ_i polynomial of degree i (Weierstrass):

$$\lim_{r \to \infty} \sup_{w \in W} \left| f^o(w) - f(w, \theta) \right| = 0$$

⬇

Polynomial NARX models

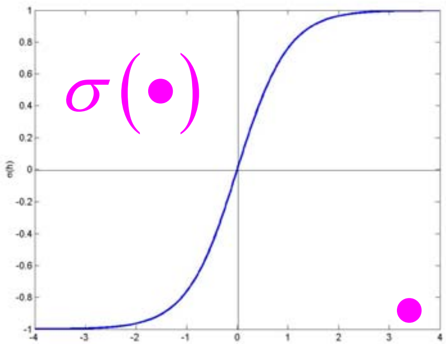
Statistical/parametric approach
Model structures: "tunable" basis

$$f(w,\theta) = \sum_{i=1}^{r} \alpha_i \sigma\left(w, \beta_i\right)$$

$$\theta = \left[\alpha_1 \cdots \alpha_r \ \beta_{11} \cdots \beta_{rq}\right]', \quad \beta_i \in \Re^q$$

- One of the most common "tunable" parameterization is the one-hidden layer sigmoidal neural network

$$\sigma\left(w, \beta_i\right) = \sigma\left(w^T a_i + b_i\right)$$



sigmoid

Statistical/parametric approach
Model estimation

$$f^o = f(w, \theta^o) = \sum_{i=1}^{r} \alpha_i^o \sigma\left(w, \beta_i^o\right)$$

- Given *T* noise-corrupted measurements of y^t, w^t:

$$\tilde{y}^2 = f(\tilde{w}^1, \theta^o) + d^1$$
$$\tilde{y}^3 = f(\tilde{w}^2, \theta^o) + d^2$$
$$\vdots$$
$$\tilde{y}^{T+1} = f(\tilde{w}^T, \theta^o) + d^T$$

$$\tilde{Y} = F\left(\theta^o\right) + D$$

Measured output ($\tilde{Y}$)

Known function (F)

Error (D)

Statistical/parametric approach
Model estimation

$$\tilde{Y} = F\left(\theta^o\right) + D$$

Gaussian pdf

Maximum Likelihood –

Least Squares estimate

$$\hat{\theta} = \arg\min_{\theta} R\left(\theta\right)$$

$$R(\theta) = \frac{1}{T} D'D = \frac{1}{T}\left[Y - F(\theta)\right]'\left[Y - F(\theta)\right]$$

- **Problem:** $R\left(\theta\right)$ is in general non-convex

Statistical/parametric approach
Model estimation

"Fixed" basis: $f(w,\theta)=\sum_{i=1}^{r}\alpha_i\sigma_i(w) \qquad \theta=[\alpha_1\cdots\alpha_r]'$

- Estimation of θ is a linear problem: $\tilde{Y}=L\theta^o+D$

$$L=\begin{bmatrix}\sigma_1(\tilde{w}_1) & \cdots & \sigma_r(\tilde{w}_1)\\ \vdots & \ddots & \vdots \\ \sigma_1(\tilde{w}_T) & \cdots & \sigma_r(\tilde{w}_T)\end{bmatrix} \qquad Y=\left[\tilde{y}^2\ \tilde{y}^3\cdots\tilde{y}^{T+1}\right]'$$

- If D is iid gaussian: $\hat{\theta}^{ML}=(L'L)^{-1}L'Y$

Statistical/parametric approach
Estimation accuracy

- For fixed basis and *D* iid gaussian:

$$\left|\vartheta_i^o - \hat{\theta}_i^{ML}\right| \leq 2\left[\left(L'L\right)^{-1}\right]_{ii} \sigma_i \quad w.p.\ 0.95$$

standard deviation of noise component d^i

- For tunable basis this results holds asymptotically ($T \rightarrow \infty$) with:

$$L = \left(\frac{\partial F}{\partial \vartheta}\right)_{\vartheta = \vartheta^o}$$

Statistical/parametric approach
Model structures: properties

- Model structure choice:
  - “basis” type
  - Number r of “basis”
  - Number n of regressors

- **Problem:** “curse of dimensionality”

  The number r of basis needed to obtain “accurate” approximation of f^o may grow exponentially with the dimension n of regressor space

More relevant in the case of “fixed” basis

Statistical/parametric approach
Model structures: properties

Using tunable basis:

- Under suitable regularity conditions on the function to approximate, the number of parameters r required to obtain "accurate" models grows linearly with n

- Estimation of θ requires to solve a non-convex minimization problem

⇓

Trapping in local minima

Statistical/parametric approach
Modeling errors

- Basic to the statistical/parametric approach is the assumption of no modeling error

↓

$$\exists\, \vartheta^{o} : f^{o} = f(w, \vartheta^{o})$$

↓

$$d^{t} = \tilde{y}^{t} - f(w, \vartheta^{o})$$

stochastic variable

independent of input u

Statistical/parametric approach
Modeling errors

- Searches for the functional form of unknown f^o are time consuming and lead to approximate model structures

⇓

- d^t is no more a stochastic variable independent of u

- Statistical estimation in presence of modeling errors is a hard problem

⇓

Set Membership approach:

- no assumption on the functional form of f^o
- no statistical assumption on d^t

Set Membership approach

- SM assumptions:
  - on system: $f^o \in \mathcal{F}(\gamma) = \left\{ f \in C^1 : \left\| f'(w) \right\|_2 \leq \gamma, \forall w \in W \right\}$

    bounded set $\in \boldsymbol{R}^n$
  - on noise: $\left| d^t \right| \leq \varepsilon^t + \gamma \delta^t, t = 1, .., T$

- Significant improvements obtained by:
  - **use of "local" bound** $\left\| f'(w) \right\|_2 \leq \gamma(w)$
  - **scaling of regressors *w* to adapt to data**

Set Membership approach

- All information (prior and data) are summarized in the Feasible Systems Set:

$$FSS^T \doteq \left\{ f \in \mathcal{F}(\gamma) : |\tilde{y}^t - f(\tilde{w}^t)| \leq \varepsilon^t + \gamma\delta^t, \quad t = 1, \cdots, T \right\}$$

- FSS^T is the set of all systems $\in \mathcal{F}(\gamma)$ that could have generated the data

- Inference algorithm Φ maps all information into estimated inference:

$$\hat{I} = \Phi(FSS^T) \simeq I(f^o, w^T)$$

Set Membership approach
Prior assumptions validation

- Prior assumptions are invalidated by data
  if FSS^T is empty

- Prior assumptions are considered validated
  if $FSS^T \neq \emptyset$

- The fact that the priors are validated by using the present data does not exclude that they may be invalidated by future data

  (Popper, "Conjectures and Refutations: the Growth of Scientific Knowledge", 1969)

Set Membership approach
Prior assumptions validation

- Define: $$\overline{f}(w) = \min_{t=1,..,T-1} (\overline{h}^t + \gamma \| w - \tilde{w}^t \|_2)$$

$$\underline{f}(w) = \max_{t=1,..,T-1} (\underline{h}^t - \gamma \| w - \tilde{w}^t \|_2)$$

$$\overline{h}^t = \tilde{y}^{t+1} + \varepsilon^t + \gamma\delta^t, \ \underline{h}^t = \tilde{y}^{t+1} - \varepsilon^t - \gamma\delta^t$$

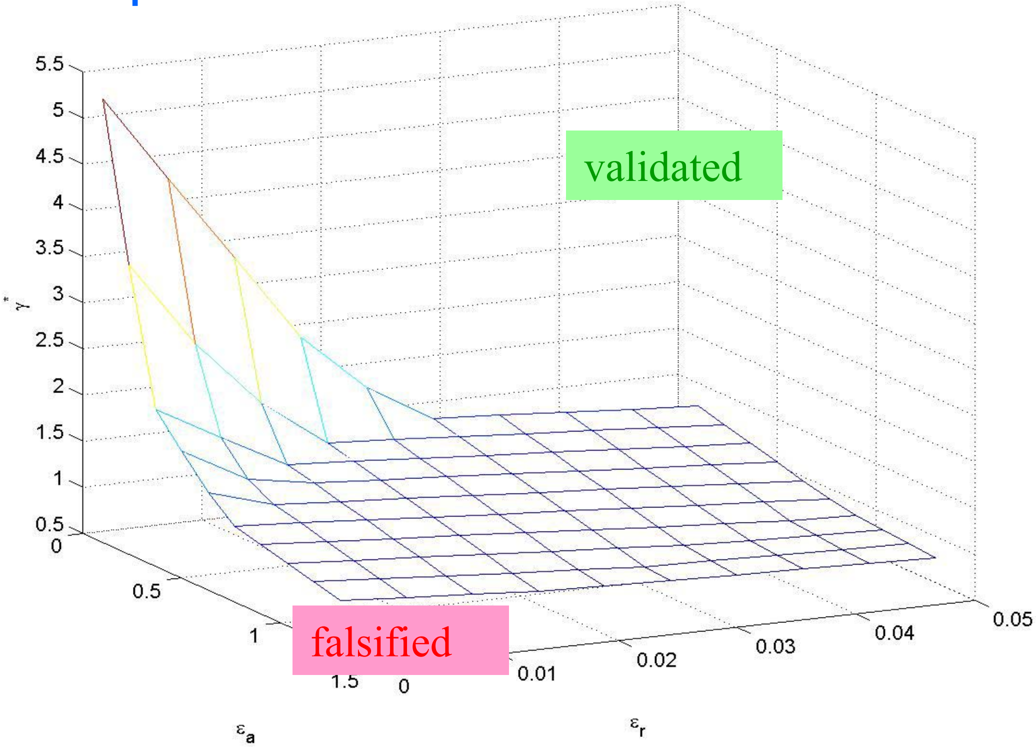
Theorem:

Conditions for assumptions to be validated are:

- necessary: $\overline{f}(\tilde{w}^t) \geq \underline{h}^t, t = 1,..,T$
- sufficient: $\overline{f}(\tilde{w}^t) > \underline{h}^t, t = 1,..,T$

Set Membership approach
Prior assumptions validation

- In space (γ,ε) the surface $\gamma^*(\varepsilon) = \inf_{FSS^T \neq \emptyset} \gamma$ separates falsified values from validated ones



Used for the choice of γ,ε values

Set Membership approach
Error and optimality concepts

- (Local) Inference error:

$$E(\hat{I}) = E[\Phi(FSS^T)] = \sup_{f \in FSS^T} \sup_{|w^T - \tilde{w}^T| \le \varepsilon^T + \gamma\delta^T} \| \Phi(FSS^T) - I(f\ ,w^T) \|$$

- An algorithm Φ^* is optimal if:

$$E[\Phi^*(FSS^T)] = \inf_{\Phi} E[\Phi(FSS^T)] = r \quad \forall FSS^T$$

  - r: (local) radius of information

- An algorithm Φ^α is α-optimal if:

$$E[\Phi^\alpha(FSS^T)] \le \alpha \inf_{\Phi} E[\Phi(FSS^T)] \quad \forall FSS^T$$

Set Membership approach

Inference ⇨ **Identification: $I(f,w^T)=f$**

- Let $\|I(f,w^T)\|=\|f\|_p= [\int_W |f(w)|^p \mathrm{d}w]^{1/p}$
- Define $f^c(w) = \frac{1}{2}[\underline{f}(w)+\overline{f}(w)]$

Theorem:

i) The identification algorithm $\Phi^c(FSS^T)=f^c$

is optimal for any L_p norm, $1 \leq p \leq \infty$

ii) The radius of information r is:

$$E[f^c] = r = \frac{1}{2}\|\overline{f}-\underline{f}\|_p$$

Set Membership approach

Inference ⇨ **Prediction:** $\mathrm{I}(f, w^T) = f(w^T)$

- Let: $\| \mathrm{I}(f, w^T) \| = | f(w^T) |$
- Assume: $|d^t| \le \varepsilon^t + \gamma \delta^t$
- Let: $B_\delta(\tilde{w}^t) = \left\{ w \in W : \|w - \tilde{w}^t\|_2 \le \delta^t \right\}$

Set Membership approach

Inference ➡ **Prediction:** $I(f, w^T) = f(w^T)$

Theorem:

***i*)** The prediction algorithm $\Phi^c(FSS^T) = f^c(\tilde{w}^T)$

is 2-optimal, with prediction error bounded by:

$$E\left[\Phi^c\left(FSS^T\right)\right] \le \frac{1}{2}\left[\overline{f}(\tilde{w}^T) - \underline{f}(\tilde{w}^T)\right] + \gamma\delta^T$$

***ii*)** If $B_\delta(\tilde{w}^T) \subset \underline{C}^T \cap \overline{C}^T$, then prediction $\hat{y}^{T+1} = f^c(\tilde{w}^T)$

is optimal and the radius of information is:

$$E\left[\Phi^c\right] = r = \frac{1}{2}\left[\overline{f}(\tilde{w}^T) - \underline{f}(\tilde{w}^T)\right] + \gamma\delta^T$$

Structured identification

- In the case of large dimension of regressor space it is often very hard to obtain satisfactory modeling accuracy.

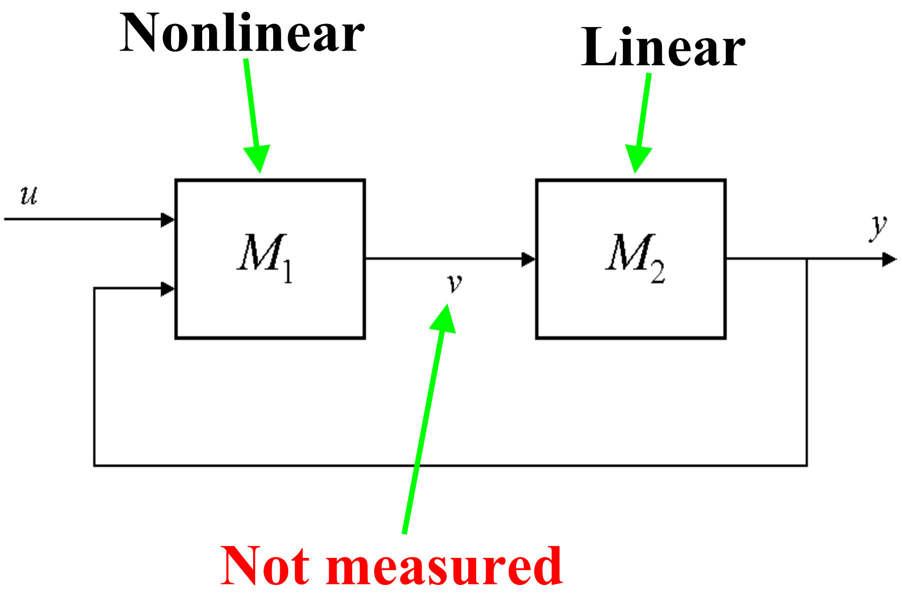
↓

Structured (block-oriented) identification

↓

- The high-dimensional problem is reduced to the identification of lower dimensional subsystems and to the estimation of their interactions

Structured identification



- Typical cases: Wiener, Hammerstein and Lur’e systems

Structured identification

Iterative identification algorithm:

- Initialisation: get an initial guess $M_2^{(0)}$ of M_2
- Step k:

1) Compute $v^{(k)}$ such that $M_2^{(k-1)}[v^{(k)}]=y$
2) Identify $M_1^{(k)}$ using u and y as inputs, $v^{(k)}$ as output
3) Identify $M_2^{(k)}$ using $v^{(k)} = M_2^{(k)}[u,y]$ as input, y as output and return to step 1)

Key feature:

The identification error is non-increasing for increasing iteration.

Model quality evaluation

- The usual approach is to look for model validity
- Model invalidity only can be surely asserted, when the model does not explain the measured data

$$|\tilde{y}^t - y_M{}^t| > \quad \textit{expected noise size}$$

- Infinitely many not-invalidated models can be derived
- Even more, infinitely many models exactly explaining the data can be derived

"overfitting" danger

Model quality evaluation

- Finding models exactly explaining the data

choose #r of basis functions = #T of measured data

$$L = \begin{bmatrix} \sigma_1(\tilde{w}_1) & \cdots & \sigma_T(\tilde{w}_1) \\ \vdots & \ddots & \vdots \\ \sigma_1(\tilde{w}_T) & \cdots & \sigma_T(\tilde{w}_T) \end{bmatrix}$$

← ***invertible***

$$\hat{\vartheta} = (L'L)^{-1} L'\tilde{Y} \quad \Rightarrow \quad Y_M = L\hat{\vartheta} = L(L'L)^{-1}L'\tilde{Y} = \tilde{Y}$$

Model quality evaluation

Example:

$\tilde{u}^1 = -2 \quad \tilde{u}^2 = 0.5 \quad \tilde{u}^3 = 0.8 \quad \tilde{u}^4 = -0.5$ ← *input*

$\tilde{y}^1 = 0 \quad \tilde{y}^2 = 1 \quad \tilde{y}^3 = -8 \quad \tilde{y}^4 = 0.125$ ← *output*

$$M_1(\vartheta) \Rightarrow \quad {y_{M1}}^{t+1} = \vartheta_1 u^t + \vartheta_2\, u^{t-1}$$

$$M_2(\vartheta) \Rightarrow \quad {y_{M2}}^{t+1} = \vartheta_1 u^t + \vartheta_2 (u^{t-1})^2$$

$$M_3(\vartheta) \Rightarrow \quad {y_{M3}}^{t+1} = \vartheta_1 u^t + \vartheta_2 (u^{t-1})^3$$

← *candidate model structures*

Model quality evaluation

Estimation of M_1, M_2, M_3

$$Y = L\,\theta$$

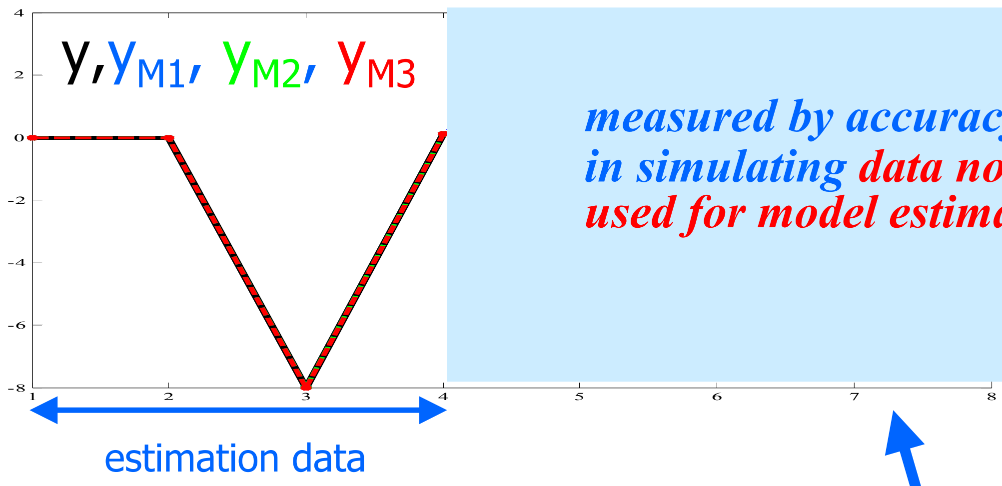
$$M_1(\vartheta) \Rightarrow \begin{matrix} t=2\rightarrow \\ t=3\rightarrow \end{matrix} \begin{bmatrix} -8 \\ 0.125 \end{bmatrix} = \begin{bmatrix} 0.5 & -2 \\ 0.8 & 0.5 \end{bmatrix} \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{\vartheta}_1 \\ \hat{\vartheta}_2 \end{bmatrix} = L^{-1}Y = \begin{bmatrix} -2.03 \\ 3.49 \end{bmatrix}$$

$$M_2(\vartheta) \Rightarrow \begin{matrix} t=2\rightarrow \\ t=3\rightarrow \end{matrix} \begin{bmatrix} -8 \\ 0.125 \end{bmatrix} = \begin{bmatrix} 0.5 & -4 \\ 0.8 & 0.25 \end{bmatrix} \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{\vartheta}_1 \\ \hat{\vartheta}_2 \end{bmatrix} = L^{-1}Y = \begin{bmatrix} 0.81 \\ -2.10 \end{bmatrix}$$

$$M_3(\vartheta) \Rightarrow \begin{matrix} t=2\rightarrow \\ t=3\rightarrow \end{matrix} \begin{bmatrix} -8 \\ 0.125 \end{bmatrix} = \begin{bmatrix} 0.5 & -8 \\ 0.8 & 0.125 \end{bmatrix} \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{\vartheta}_1 \\ \hat{\vartheta}_2 \end{bmatrix} = L^{-1}Y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Model quality evaluation

- All models M_1, M_2, M_3 explain exactly the given data y



measured by accuracy in simulating data not used for model estimation

- How to choose among them ?

choose the one with the best “predictive ability”

Model quality evaluation

- Several indexes have been proposed for estimating the predictive ability of models:
  - $FPE = R(\hat{\vartheta})\frac{T+n}{T-n}$
  - $AIC = \ln R(\hat{\vartheta}) + \frac{2n}{T}$
  - $BIC = \ln R(\hat{\vartheta}) + \frac{n \ln T}{T}$

T : number of data

n : number of parameters ϑ

$$R(\theta) = \frac{1}{T}[Y - L\vartheta]'[Y - L\vartheta]$$

- They provide quite crude approximations, especially for nonlinear systems

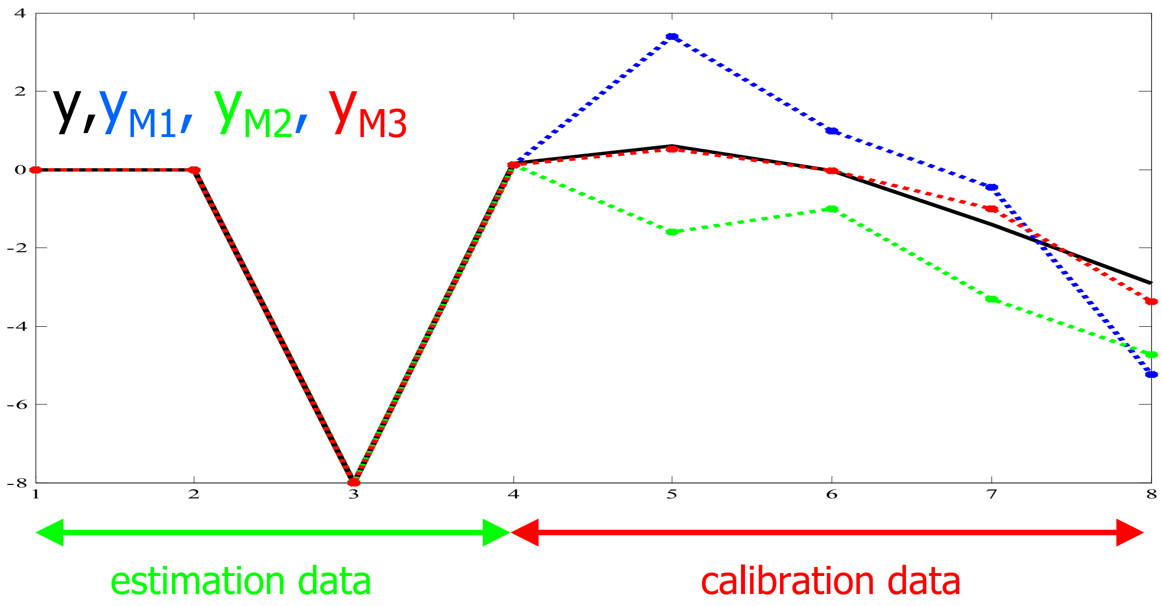
- A simple but effective approach: splitting of data
  - ***estimation data:*** *estimate candidate models* $\boldsymbol{M_i}$, $\boldsymbol{i=1,..,m}$
  - ***calibration data:*** *choose the best one among* $\boldsymbol{M_i}$

Model quality evaluation

- Best model among candidate ones M_i

minimum simulation error on the "calibration" data

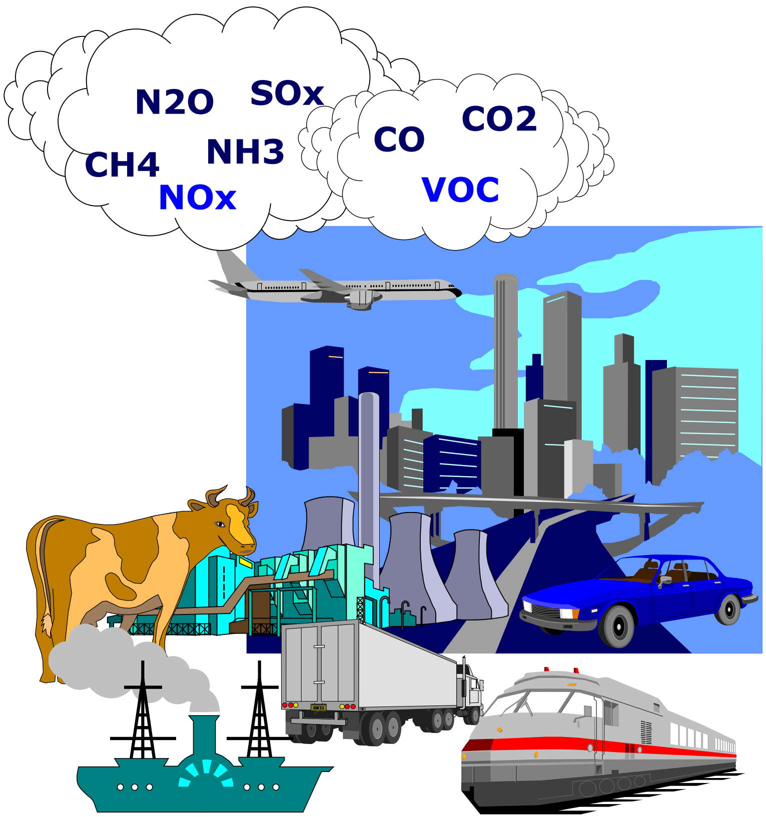
- Example: M_3 is the best one among M_1, M_2, M_3

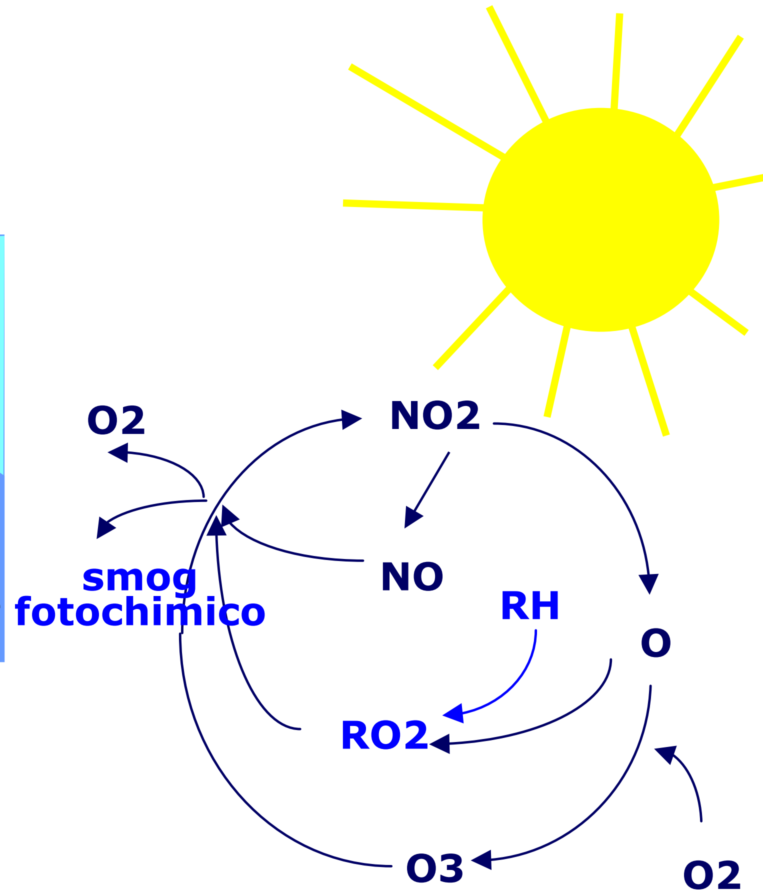


Applications

- Prediction of atmospheric pollution
- Simulation of dam crest dynamics
- Identification of vehicles with controlled suspensions

Prediction of urban ozone peaks





Prediction of urban ozone peaks

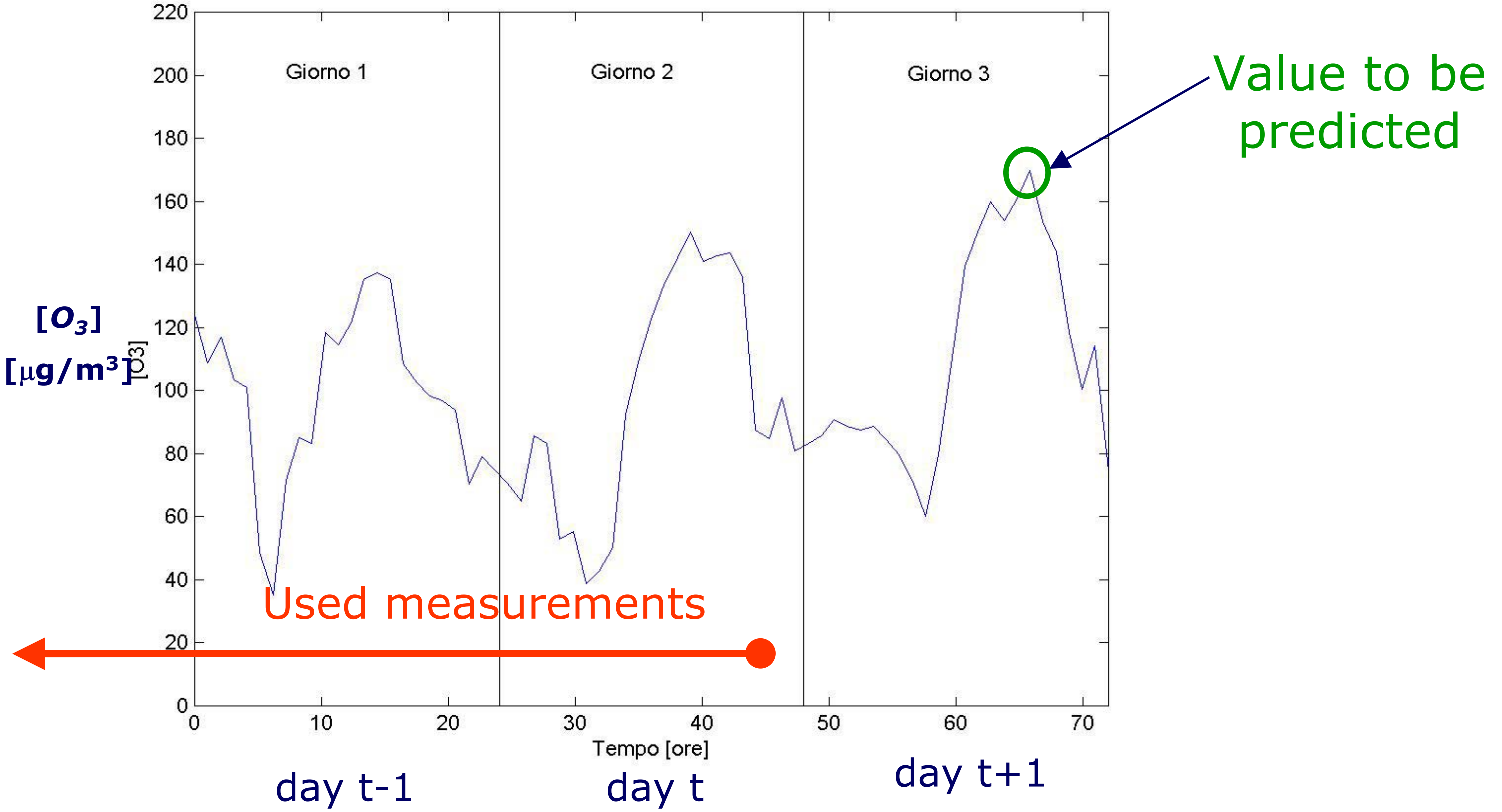
- Combustion processes and high solar radiation cause high tropospheric ozone concentrations
- Prediction of ozone concentrations is important for authorities in charge of pollution control and prevention
- Studies in the literature show that physical models are not able to reliably forecast the links between precursor emissions (No_x, VOC), methereological conditions and ozone concentrations

➢Sillman "The relation between ozone, No_x and hydrocarbons", Atmos. Environ., 1999

➢Jenkin-Clemitshaw "Ozone and other photochemical pollutants: chemical processes governing their formation", Atmos. Environ., 1999

Prediction of urban ozone peaks

typical data at Broletto (Bs)



Prediction of urban ozone peaks

- Structure of used models:

$$y^{t+1} = f^{o}(w^{t})$$

$$w^{t} = [y^{t}\, u_1^{\,t}\, u_2^{\,t}\, u_3^{\,t}\, u_4^{\,t}]$$

- y^{t}: **max O_3 concentration at day t**
- $u_1^{\,t}$: **mean NO_2 concentration at 4-8 pm of day t**
- $u_2^{\,t}$: **mean O_3 concentration at 4-8 pm of day t**
- $u_3^{\,t}$: **max temperature at day t**
- $u_4^{\,t}$: **forecast of max temperature at day t+1**

Prediction of urban ozone peaks

- Prediction methods tested:
  - **PERS:** $y^{t+1} = y^{t}$
  - **ARCX:** periodic ARX
  - **NN:** sigmoidal neural net
  - **NF:** neuro-fuzzy
  - **NSM:** nonlinear set membership

- Hourly data measured at Brescia center:
  - **1995-1998**: estimation data set
  - **1999:** calibration data set
  - **2000-2001:** testing data set

Prediction of urban ozone peaks

Indexes measuring the ability to predict concentrations exceeding a given threshold:

	observed		total
predicted	yes	no	
yes	a	f – a	f
no	m - a	N + a – m – f	N – f
total	m	N - m	N

- ✓ fraction of Correct Predictions: **CP=(a/m)%**
- ✓ fraction of False Alarms: **FA=(1-a/f)%**
- ✓ Success index: **SI=[(a/m)+((N+a-m-f)/(N-m))-1]%**

European Environmental Agency, Tech. Report 9, 1998

Prediction of urban ozone peaks

Calibration data set: m=63 exceeded thresholds

	PERS	ARCX	NN	NF	NSM
CP	65.1	61.9	69.8	63.5	71
FA	33.9	25	27.9	25.9	27.4
SI	47.6	51.1	55.7	51.8	51.2

Testing data set: m=39 exceeded thresholds

	PERS	ARCX	NN	NF	NSM
CP	41.5	35.9	53.8	66.7	71.8
FA	57.5	51.7	40	44.7	44
SI	34.4	31.3	49.6	60.2	63.5

Model of Schlegeis Arch Dam

- Model to simulate the crest displacement of the dam as function of:
  - **water level**
  - **concrete temperature**
  - **air temperature**

- Daily data available in period 1992-2000

- Difficulties in deriving reliable physical models

- Models tested: ARX, NN, NSM

Model of Schlegeis Arch Dam

- Structure of used models:

$$y^{t+1} = f^{o}(w^{t})$$

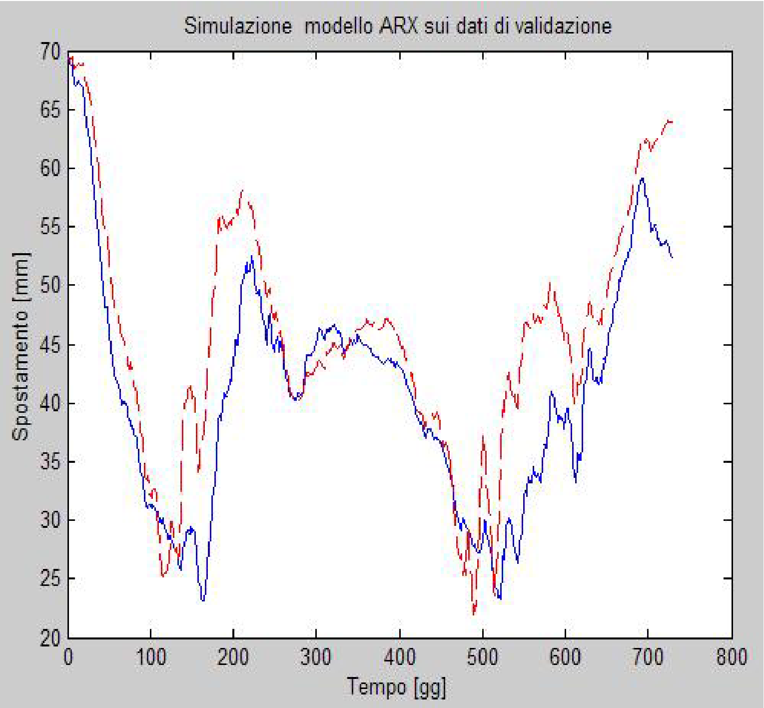
$$w^{t} = [y^{t}\ y^{t-1}\ u_1^{t+1}\ u_1^{t}\ u_1^{t-1}\ u_2^{t+1}\ u_2^{t}\ u_3^{t+1}\ u_3^{t}]$$

- y^{t}: **crest displacement at day t**
- u_1^{t}: **water level at day t**
- u_2^{t}: **concrete temperature at day t**
- u_3^{t}: **mean air temperature at day t**

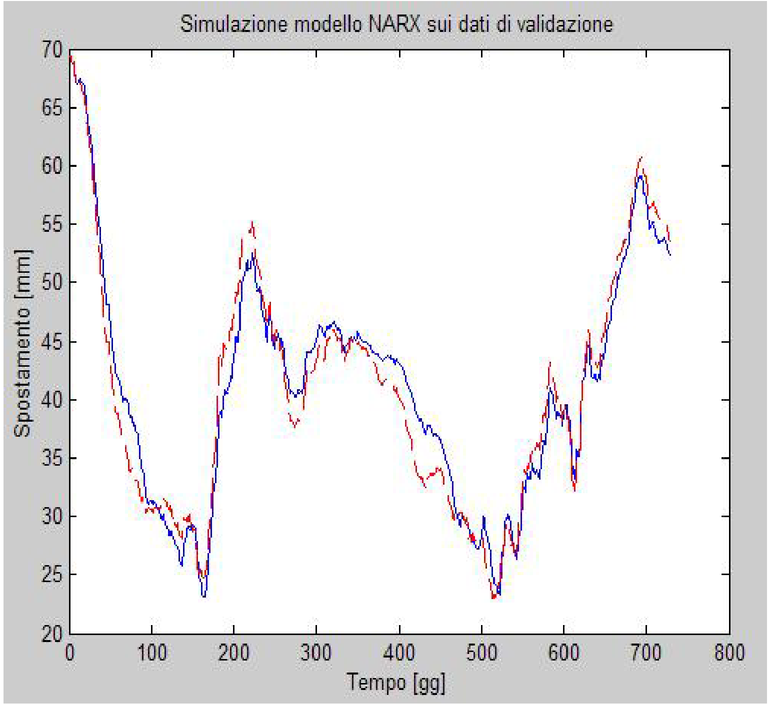
- Daily data:
  - **1992-1996**: estimation data set
  - **1997-1998:** calibration data set
  - **1999-2000:** testing data set

Model of Schlegeis Arch Dam

- Simulation results on the testing data set:



ARX model



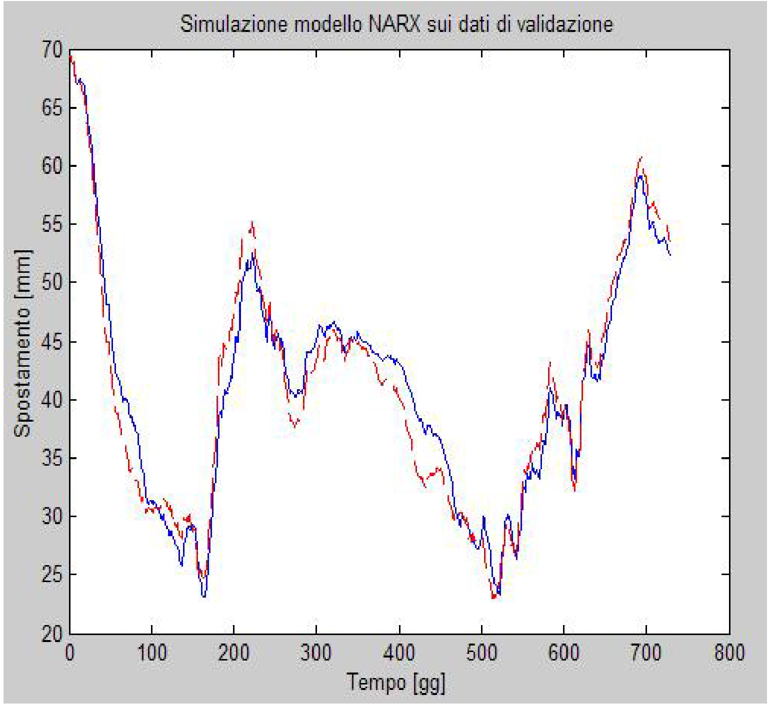
NN model

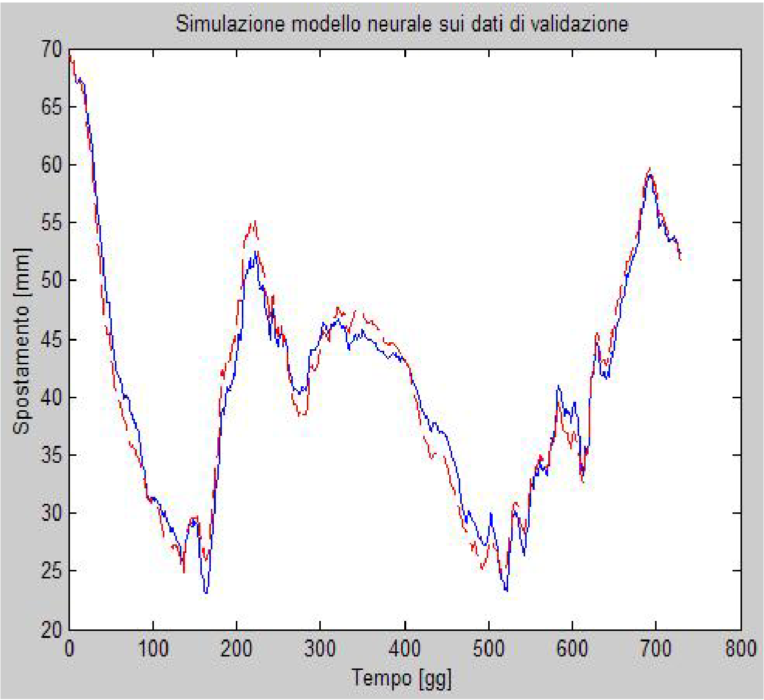
experimental data

model

Model of Schlegeis Arch Dam

- Simulation results on the testing data set:





NN model

experimental data

model

NSM model

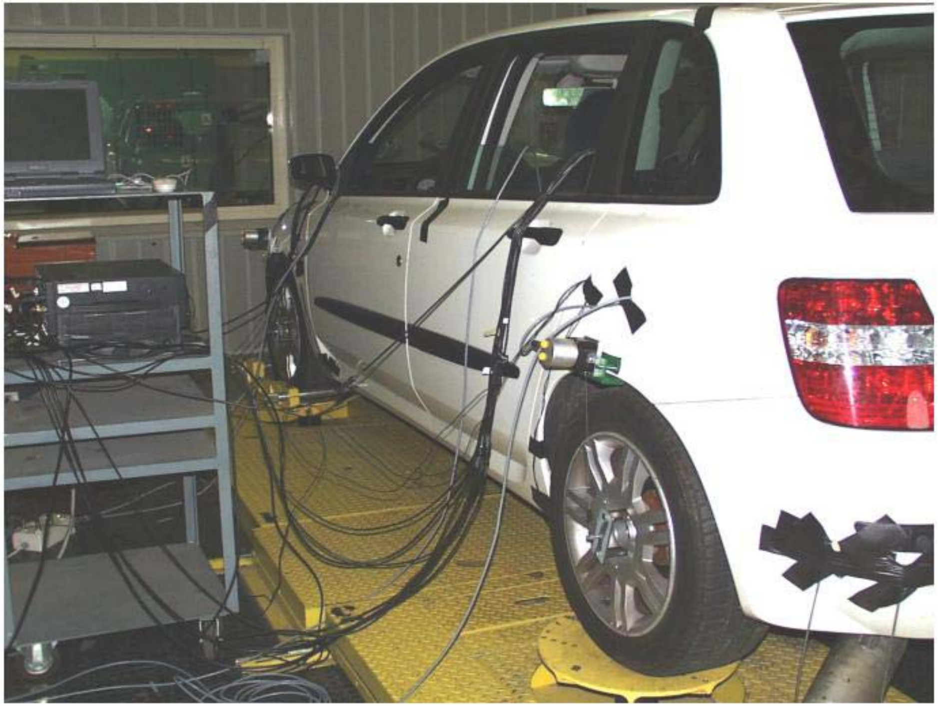
Identification of vehicles with controlled suspensions

GOAL: Derive a model for simulation of chassis and wheels accelerations as function of road profile and damper control

USE: Virtual design and tuning of Continuous Damping Control systems

Experimental setting

- C-segment prototype vehicle with controlled dampers and CDC-Skyhook (Continuous Damping Control system).



- Measurements are performed on a four-poster test bench of FIAT-Elasis Research Center.

Experimental setting

Road profiles:

- Random: random road.
- English Track: road with irregularly spaced holes and bumps.
- Short Back: impulse road.
- Motorway: level road.
- Pavé track: road with small amplitude irregularities.
- Drain well: negative impulse road.

Note: The road profiles are symmetric (left=right).

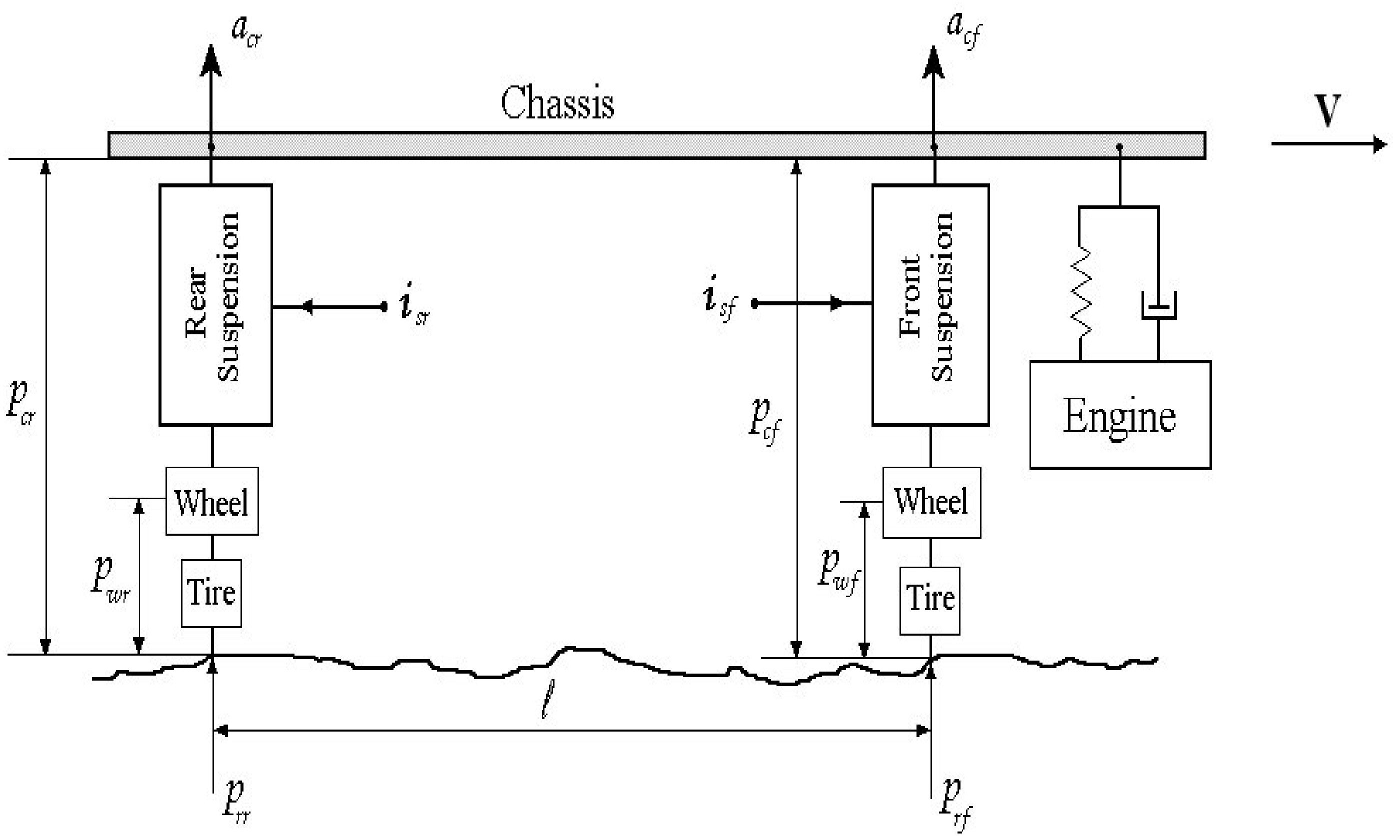
Experimental setting

Data set: 93184 data, collected with a sampling frequency of 512 Hz, partitioned as follows:

- Estimation data set: 0-5 seconds of each acquisition.
- Calibration data set: 5-7 seconds of each acquisition.
- Testing set: 7-14 seconds of each acquisition.

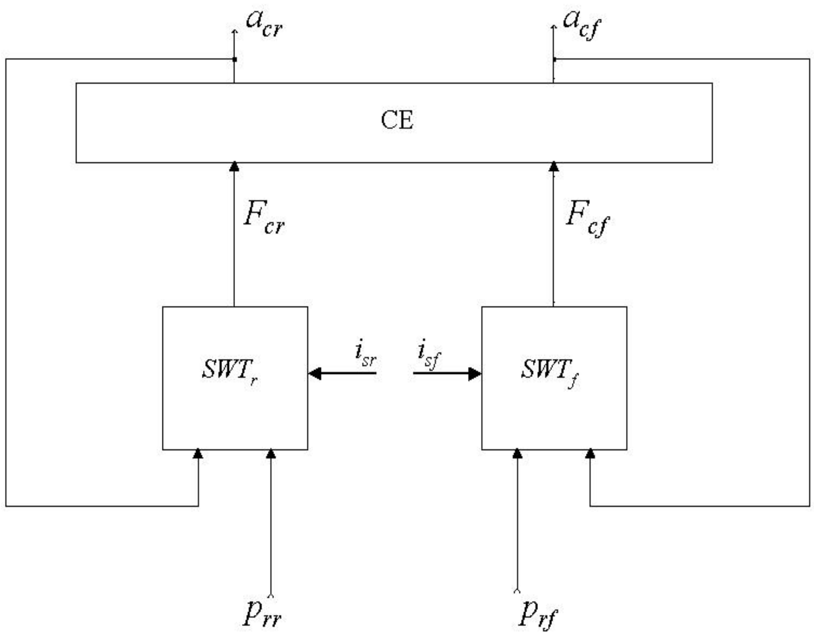
Structure of vehicles vertical dynamics

Since the road profiles are symmetric,
a Half-car model has been considered:



Structured Identification of vehicles vertical dynamics

Structure decomposition:



- CE: chassis + engine
- SWT: suspension + wheel + tire

Measured variables:

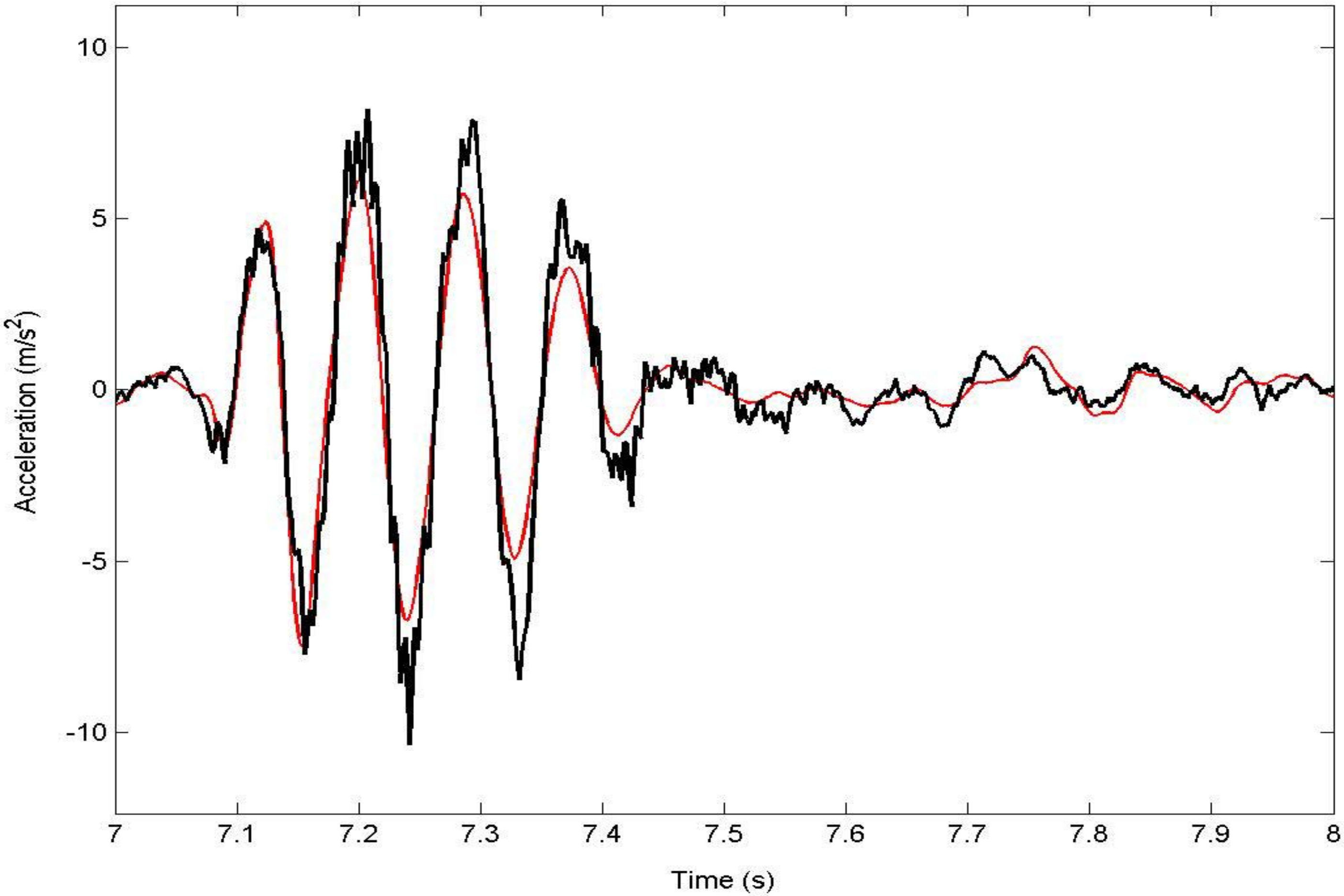
- p_{rf} and p_{rr}: front and rear road profiles.
- i_{sf} and i_{sr}: control currents of front and rear suspensions.
- a_{cf} and a_{cr}: front and rear chassis vertical accelerations.

Note: F_{cf} and F_{cr} are not measured.

Results on testing set of NSM model

Front wheel acceleration: english track road

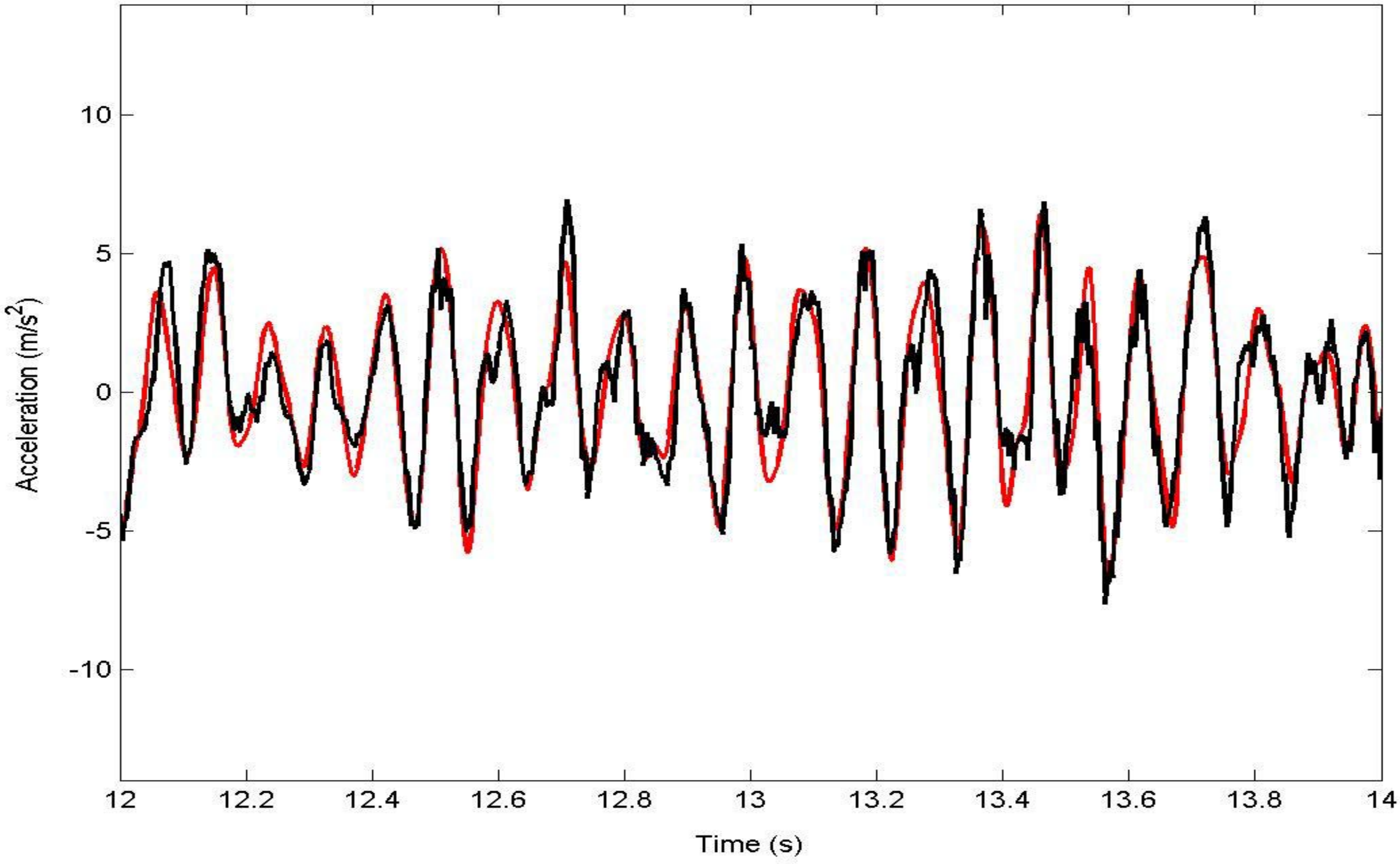
measurements, NSM model



Results on testing set of NSM model

Chassis front accelerations: random road

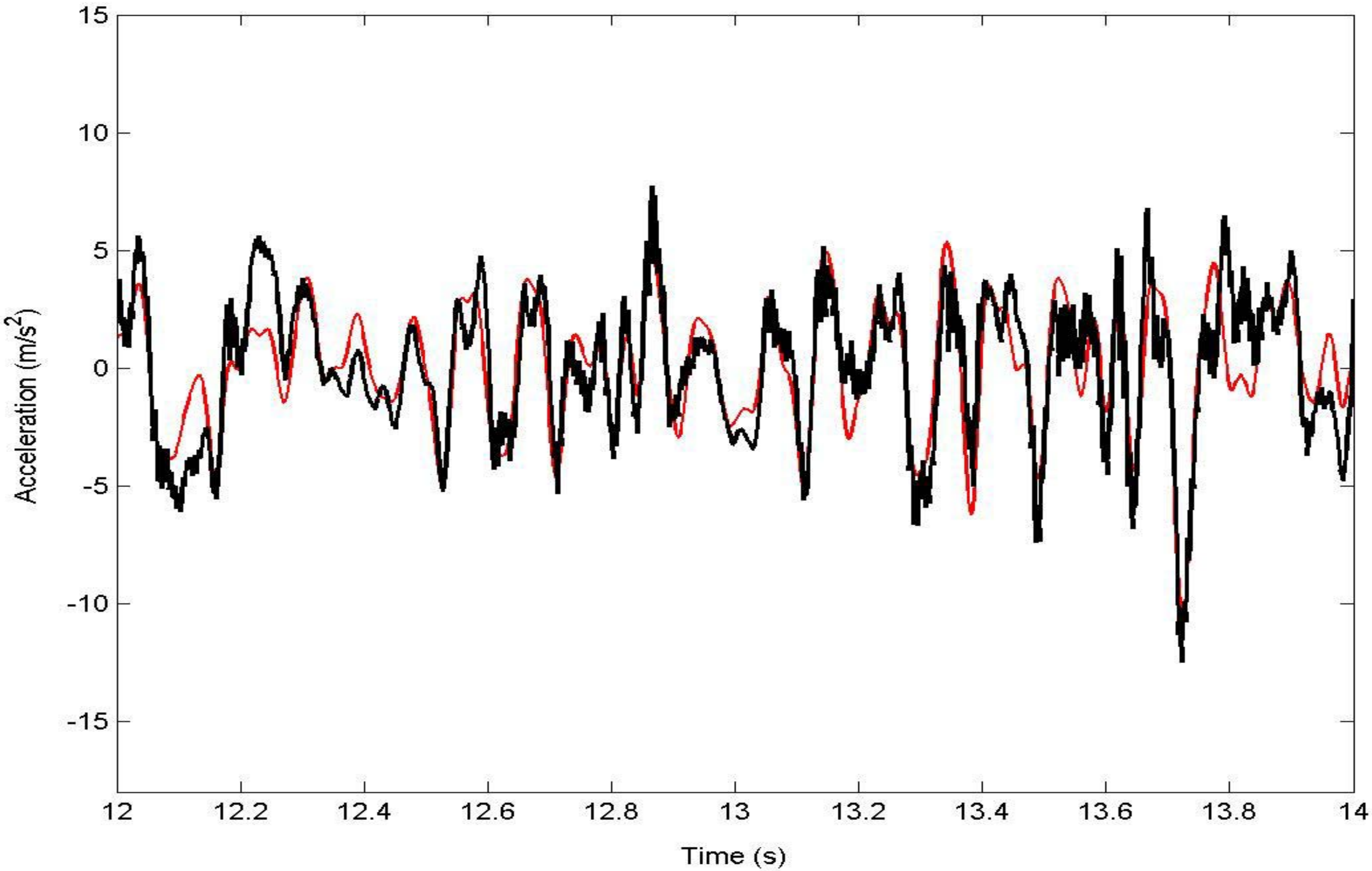
measurements, NSM model



Results on testing set of NSM model

Chassis rear accelerations: random road

measurements, NSM model.



Comparison with physical model

Chassis front accelerations: random road

measurements, NSM model, physical model