Prediction of sunspot activity

- Prediction of sunspot activity is important in telecommunication, meteorology, agriculture, etc.
- The set of data 1770-1892 has been chosen because used as a **benchmark** to test prediction methods:
  - 1770-1869: Estimation data
  - 1870-1892: Testing data
- A comparison of the Nonlinear Set Membership (NSM) almost optimal predictors is made with the following predictors:
  - AR: Linear AR model — Box-Jenkins, 1976
  - BL: Bilinear model — Granger-Andersen, 1978
  - TAR: Piece-wise AR model — Tong-Lim, 1980
  - GMDH: Polynomial model — Ivakhnenko, 1970
  - NN: Neural Network (sigmoidal)

Sunspot activity, 1-step ahead prediction:
NSMG1 predictor

- NSMG1 is obtained considering a model of the form:

$$y_{t+1} = f\left(w_t\right)$$
$$w_t = \left[y_t \; y_{t-1} \; y_{t-2} \; u_t\right]^T$$

$$u_t = 0, \;\forall t$$
$$\varepsilon_t = 0, \;\forall t$$
$$\delta_t = \delta, \;\forall t$$

  where u_t is a noise acting on the system.

- The values of $\delta = 5$ and $\gamma = 5.5$ are chosen on the base of the trade-off curve $\gamma^*\left(\delta\right)$ and of a rough evaluation of $\gamma = 3 \div 4$ derived from a neural network approximation of f^o.

- The NSMG1 prediction is:

$$\widehat{y}_{T+1} = f^c_G(\widetilde{w}_T) = \phi^c_G(FSS_T)$$

Sunspot activity, 1-step ahead prediction:
NSML1 predictor

- NSML1 is obtained by considering the series of residuals:

$$\Delta y_{t+1} = \widetilde{y}_{t+1} - f_G^c(\widetilde{w}_t)$$

  and by using a model for such series of the form:

$$\Delta y_{t+1} = \Delta f\,(w_t)$$
$$w_t = \left[y_t\; y_{t-1}\; y_{t-2}\; u_t\right]^T$$

$$u_t = 0, \quad \varepsilon_t = 0, \quad \delta_t = \delta,\ \forall t$$

  where u_t is a noise acting on the system.

- A bound on the weighted norm of grad$\Delta f(w)$:

$$\|\mathrm{grad}\Delta f(w)\|_2^\nu \leq \gamma_r$$

  with $\nu = [0.9 \quad 0.8 \quad 0.6 \quad 0]$, is taken.

- Based on validation analysis, the following values are chosen:

$$\delta = 1, \quad \gamma_r = 0.2$$

- The NSML1 prediction is:

$$\widehat{y}_{T+1} = f_G^c(\widetilde{w}_T) + \phi_\Delta^c\,(FSS_T)$$

Sunspot activity: 1-step ahead prediction performances

- The following 1-step ahead prediction performances have been evaluated on the testing data set:
  - $RMSE_1$: Root Mean Square Error
  - $MAXE_1$: Maximum Prediction Error

Predictor	$RMSE_1$	$MAXE_1$
NSMG1	14.6	28
NSML1	13.8	27
NN	16.2	41
AR	18.0	47
BL	16.6	46
SETAR	16.1	44
GMDH	14.7	42

Sunspot activity, one-step ahead prediction: NSMG11 predictor

- NSMG11 is obtained considering a model of the form:

$$y_{t+11} = f\,(w_t)$$
$$w_t = [y_t \; y_{t-2} \cdots y_{t-12} \; u_t]^T$$

$$u_t = 0, \;\forall t$$
$$\varepsilon_t = 0, \;\forall t$$
$$\delta_t = \delta, \;\forall t$$

where u_t is a noise acting on the system.

- The values of $\delta = 5$ and $\gamma = 7$ are chosen on the base of the trade-off curve $\gamma^*(\delta)$.
- The NSMG11 prediction is:

$$\widehat{y}_{T+11} = f_G^c(\widetilde{w}_T) = \phi_G^c(FSS_T)$$

Sunspot activity, 11-step ahead prediction: NSML11 predictor

- NSML11 is obtained by considering the series of residuals:

$$\Delta y_{t+11} = \widetilde{y}_{t+11} - f_G^c(\widetilde{w}_t)$$

and by using a model for such series of the form:

$$\Delta y_{t+11} = \Delta f\,(w_t)$$
$$w_t = [y_t\; y_{t-1} \cdots\; y_{t-5}\; u_t]^T$$

$$u_t = 0, \quad \varepsilon_t = 0, \quad \delta_t = \delta,\ \forall t$$

where u_t is a noise acting on the system.

- A bound on the weighted norm of $\text{grad}\Delta f(w)$:

$$\|\text{grad}\Delta f(w)\|_2^\nu \leq \gamma_r$$

with $\nu = [0.8 \quad 1.2 \quad 1.2 \quad 1 \quad 1 \quad 1.4 \quad 1 \quad 0]$, is taken.

- Based on validation analysis, the following values are chosen:

$$\delta = 1, \quad \gamma_r = 1$$

- The NSML11 prediction is:

$$\widehat{y}_{T+11} = f_G^c(\widetilde{w}_T) + \phi_\Delta^c\,(FSS_T)$$

Sunspot activity:
11-step ahead prediction performances

- The following 11-step ahead prediction performances have been evaluated on the testing data set:
  - $RMSE_{11}$: Root Mean Square Error
  - $MAXE_{11}$: Maximum Prediction Error

Predictor	$RMSE_{11}$	$MAXE_{11}$
NSMG11	19.9	45
NSML11	17.7	45
NN	23.4	63
AR	32.6	81
BL	32.6	81
SETAR	35.0	76
GMDH	29.4	81

Prediction of river flow

- Data set: 13500 measurements of the mean daily discharges of the Dora Baltea river (Valle d'Aosta-Piemonte) from January 1, 1941 to December 31, 1977:
  - 1-13000: Estimation data
  - 13200-13500: Testing data
- This data set has been used in:
  -Porporato-Ridolfi, "Nonlinear analysis of river flow time sequences", Water Resour. Res., 1977.
  - Porporato-Ridolfi, "Clues to existence of deterministic chaos in river flow", Int. J. of Modern Physics B, 1996.
- Evidence of a component of deterministic chaos has been found.

Prediction of river flow

- A comparison of 1-step ahead prediction performances is made among the following predictors:
  - NSMG: Nonlinear SM predictor with global information on grad$f$
  - NSML: Nonlinear SM predictor with local information on grad$f$
  - NN: Neural Network (sigmoidal)
  - JIT: Just In Time predictor (Porporato-Ridolfi, 1997)

River flow: NSMG predictor

- NSMG is obtained considering a model of the form:

$$y_{t+1} = f\left(w_t\right)$$
$$w_t = \left[y_t \; y_{t-1} \; y_{t-2} \; y_{t-3} \; u_t\right]^T$$

$$u_t = 0, \; \forall t$$
$$\varepsilon_t = \varepsilon \left|y_t\right|, \; \forall t$$
$$\delta_t = \delta, \; \forall t$$

where u_t is a noise acting on the system.

- A bound on the weighted norm of grad$f^o(w)$:

$$\|\text{grad} f^o(w)\|_2^\nu \leq \gamma$$

with $\nu = [0.6 \quad 0.8 \quad 1.2 \quad 1 \quad 0]$, is taken.

- Comparing the measured time series with the filtered one, an estimate of $\varepsilon = 0.3$ has been obtained. The values of $\delta = 10$ and $\gamma = 50$ are chosen on the base of the trade-off curve $\gamma^*\left(\delta\right)$ and of a rough evaluation of $\gamma = 30 \div 40$ derived from a neural network approximation of f.

- The NSMG prediction is:

$$\widehat{y}_{T+1} = f_G^c(\widetilde{w}_T)$$

River flow: NSML predictor

- NSML is obtained by considering the series of residuals:

$$\Delta y_{t+1} = \widetilde{y}_{t+1} - f_G^c(\widetilde{w}_t)$$

  and by using a model for such series of the form:

$$\Delta y_{t+1} = \Delta f\,(w_t)$$
$$w_t = [y_t\; y_{t-1}\; \cdots\; y_{t-5}\; u_t]^T$$

$$u_t = 0, \quad \varepsilon_t = \varepsilon\,|y_t|\,, \quad \delta_t = \delta,\ \forall t$$

  where u_t is a noise acting on the system.

- A bound on the weighted norm of grad$\Delta f(w)$:

$$\|\text{grad}\Delta f(w)\|_2^{\nu} \leq \gamma_r$$

  with $\nu = [0.6 \quad 0.8 \quad 1.2 \quad 1 \quad 0 \quad 0 \quad 0]$, is taken.

- Based on validation analysis, the following values are chosen:

$$\varepsilon = 0.6, \quad \delta = 10, \quad \gamma_r = 22$$

- The NSML prediction is:

$$\widehat{y}_{T+1} = f_G^c(\widetilde{w}_T) + \phi_\Delta^c\,(FSS_T)$$

River flow: prediction performances

- The following 1-step ahead prediction performances have been evaluated on the testing data set:
  - $RMSE$: Root Mean Square Error
  - $MAXE$: Maximum Prediction Error

Predictor	$RMSE$	$MAXE$
NSMG	86.8	688
NSML	85.1	731
NN	86.5	861
JIT	90.2	872

Is SM approach too conservative ?

- The SM approach makes use of quite weak assumptions on function f^o and noises e and e'.
- This feature is basic for good performances in applications where **reliable information** on the functional form of f^o and on noise properties (uncorrelation, pdf,...) **is not available**.
- In case that reliable information is available, it may be guessed that the SM approach gives largely conservative results in comparison with methods able to account for such information.
- A **most adverse situation** for SM approach is simulated:
  - data are generated by a linear AR model driven by iid gaussian noise
  - using such information, the optimal predictor, minimizing the expected 1-step prediction error is computed.
  - the prediction performances are compared with the ones of NSM predictors

Is SM approach too conservative ?

- 200 time series of 150 data was generated by the equation:

$$y_{t+1} = 14.7 + 1.425 y_t - 0.731 y_{t-1} + d_t$$

where d_t is an iid gaussian noise of variance $\sigma_d^2 = 50$.

- The time series have been divided into an estimation set of 100 data and a testing set of 50 data.
- The root mean square errors $\overline{RMSE}_k$ and the maximum prediction errors in absolute value $\overline{MAXE}_k$ for the k-step ahead prediction, with $k = 1, 11$ are evaluated on the testing data and averaged over the 200 realizations of the time series.

Predictor	$\overline{RMSE}_1$	$\overline{MAXE}_1$	$\overline{RMSE}_{11}$	$\overline{MAXE}_{11}$
NSMG1	8.9	23	19.2	46
NSMG11	-	-	19.4	45
NSML1	8.5	22	19.5	47
NSML11	-	-	18.8	43
AR	7.2	18	17.9	43