A Constructive-Fuzzy System Modeling for Time Series Forecasting

I. Luna[1], R. Ballini[2] and S. Soares[1]

{iluna,dino}@cose.fee.unicamp.br, ballini@eco.unicamp.br

[1]Department of Systems, School of Electrical and Computer Engineering, University of Campinas - Brazil

[2]Institute of Economy, University of Campinas - Brazil

Summary

- Introduction;
- Time series analysis: data pre-processing;
- Input selection;
- General structure of a fuzzy rule-based model;
- Constructive learning;
- Case study: NN3 competition;
- Conclusions and future works.

Time series modeling

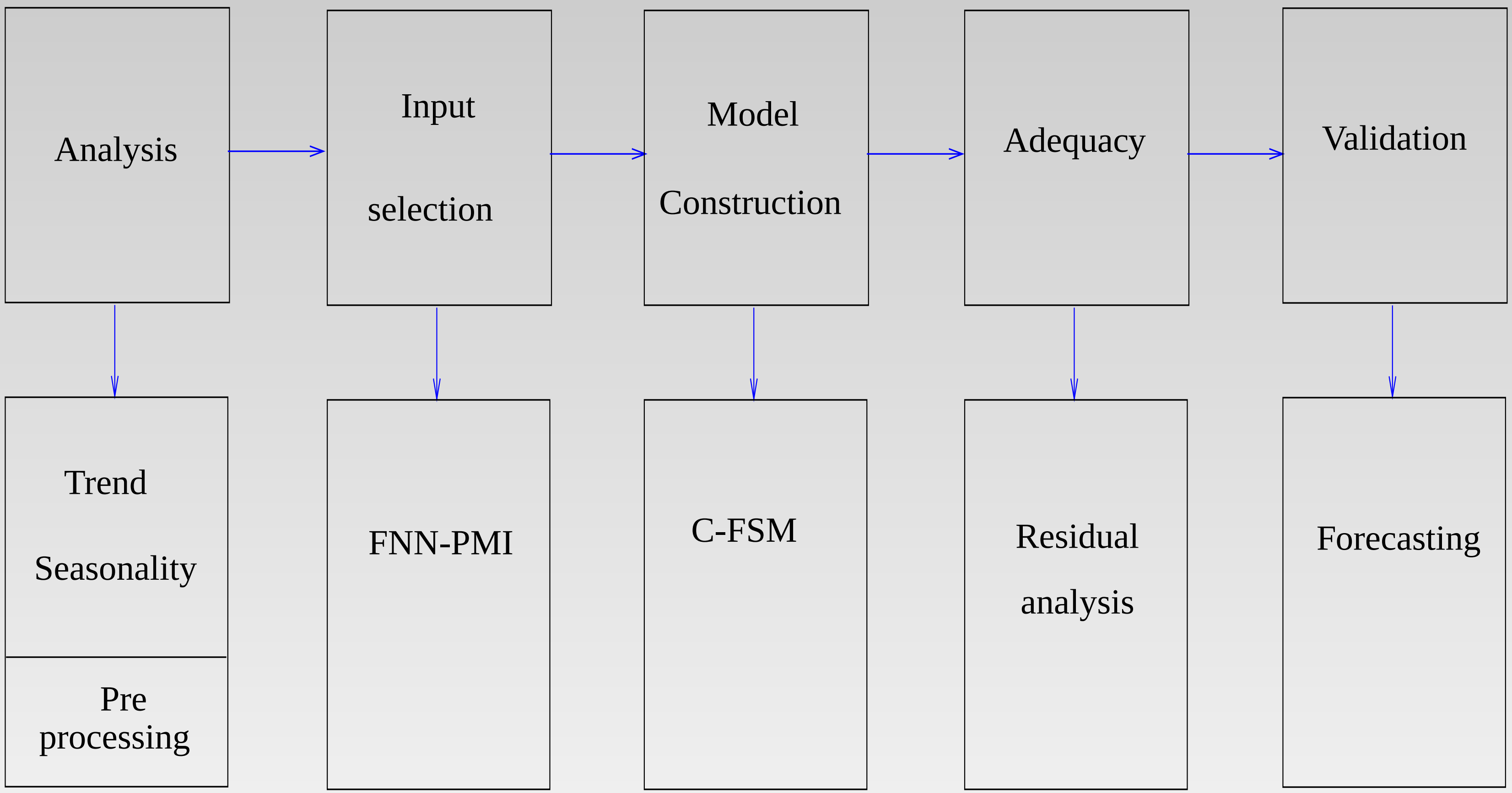


Figure 1: Time series modeling.

Analysis and pre-processing

- Reduced data set of the NN3 competition;
- Stationarity: required for input selection;
- Seasonal and trend components:

Econometric Views 2.0 - Quantitative Software, 1995.

- Trend component: series 1, 5 and 9;
- All series with no trend were transformed according to:

$$z^k(m) = \frac{y^k(m) - \mu(m)}{\sigma(m)}$$

z^k: stationary version of the time series;
$y^k,\ k = 1, \ldots$: $k-$th observation;
$\mu(m)$ is the monthly average value and $\sigma(m)$ is the monthly standard deviation.

Input selection

1. **FNN** (*False Nearest Neighbors*): determines the minimum number of lags necessary to represent each pattern or *state* of the time series;

2. **PMI** (*Partial Mutual Information*): measure of information that each new variable $\mathbf{x}$ provides, taking into account an existing set of inputs $\mathbf{Z}$. Given variables X e Y, **PMI** score between X and Y is defined by:

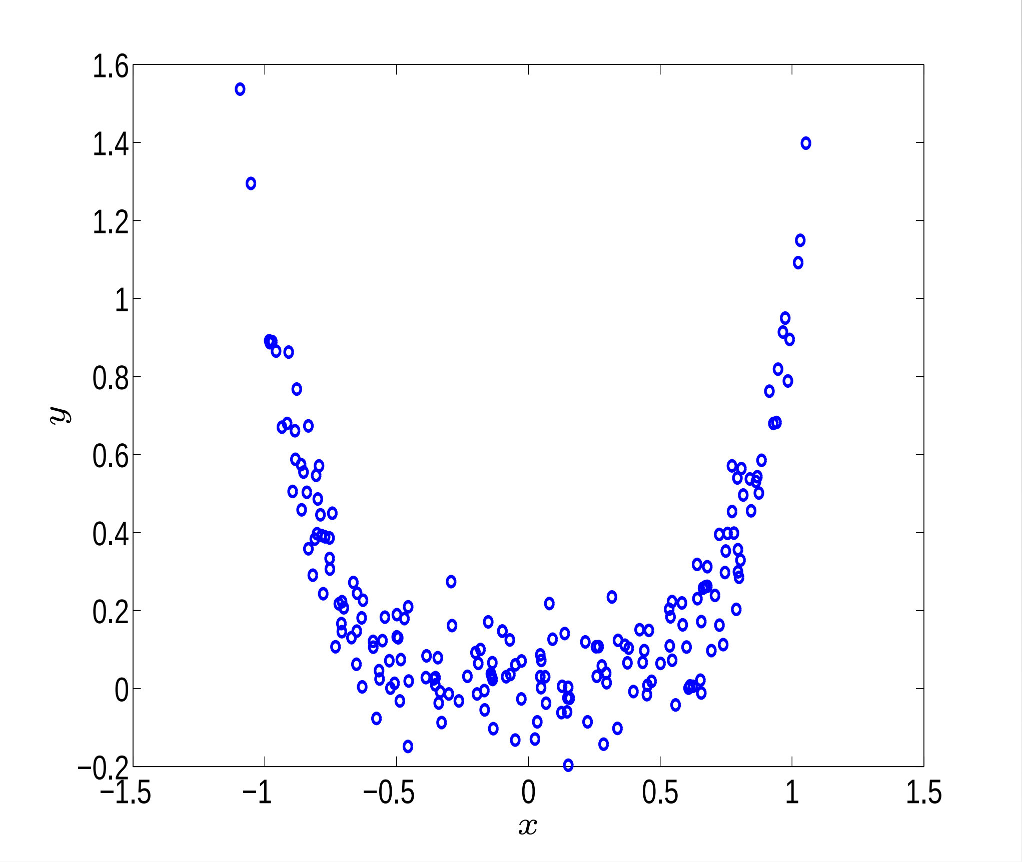
$$\mathrm{PMI} = \frac{1}{N}\sum_{i=1}^{N} \log_e \left[\frac{f_{X',Y'}(x'_i, y'_i)}{f_{X'}(x'_i) f_{Y'}(y'_i)} \right] \tag{1}$$

where:

$$x'_i = x_i - E(x_i|\mathbf{Z}) \quad \text{e} \quad y'_i = y_i - E(y_i|\mathbf{Z})$$

$\mathbf{Z}$ is the st of inputs already chosen. $E(\cdot|\mathbf{Z})$ is the conditional expected value; N is the number of input-output patterns.

Input selection



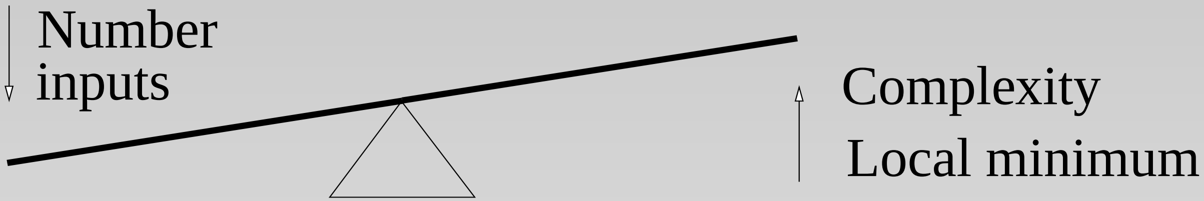
$$y = x^4 + e_1$$

$$x = \text{sen}(2\pi t/T) + e_2$$

$T = 20$; $t = 1, \ldots, 200$; e_1 and e_2 are noisy signals with normal distribution, $\mu = 0$ and $\sigma = 0, 1$.

Measure	Value
MI	0.4199
Correlation	0.0032

Input selection



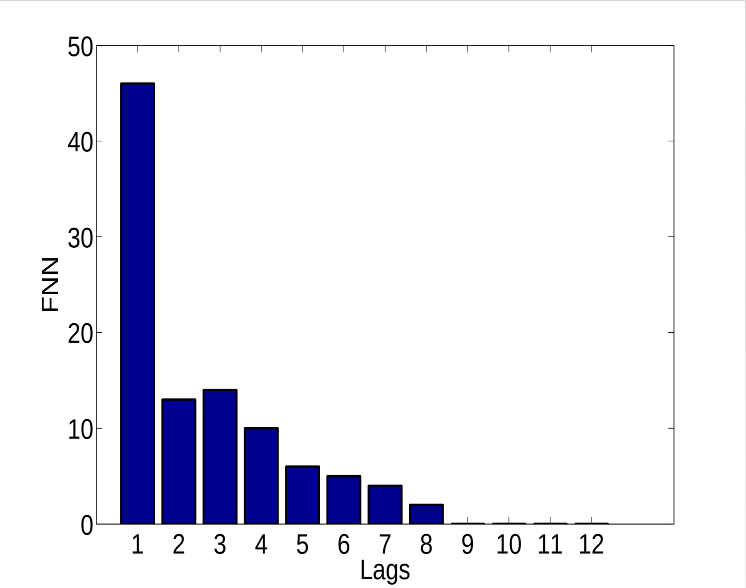


Figure 2: FNN.

NN3_102

$\Longrightarrow$

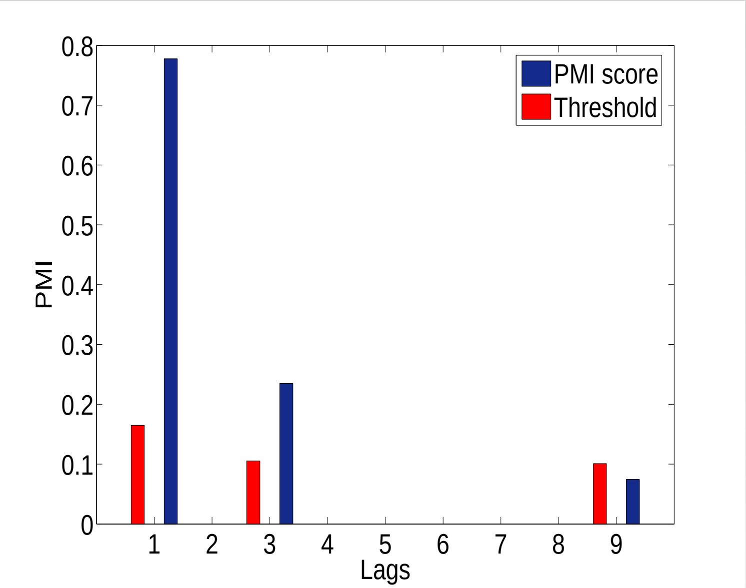


Figure 3: PMI.

A general structure

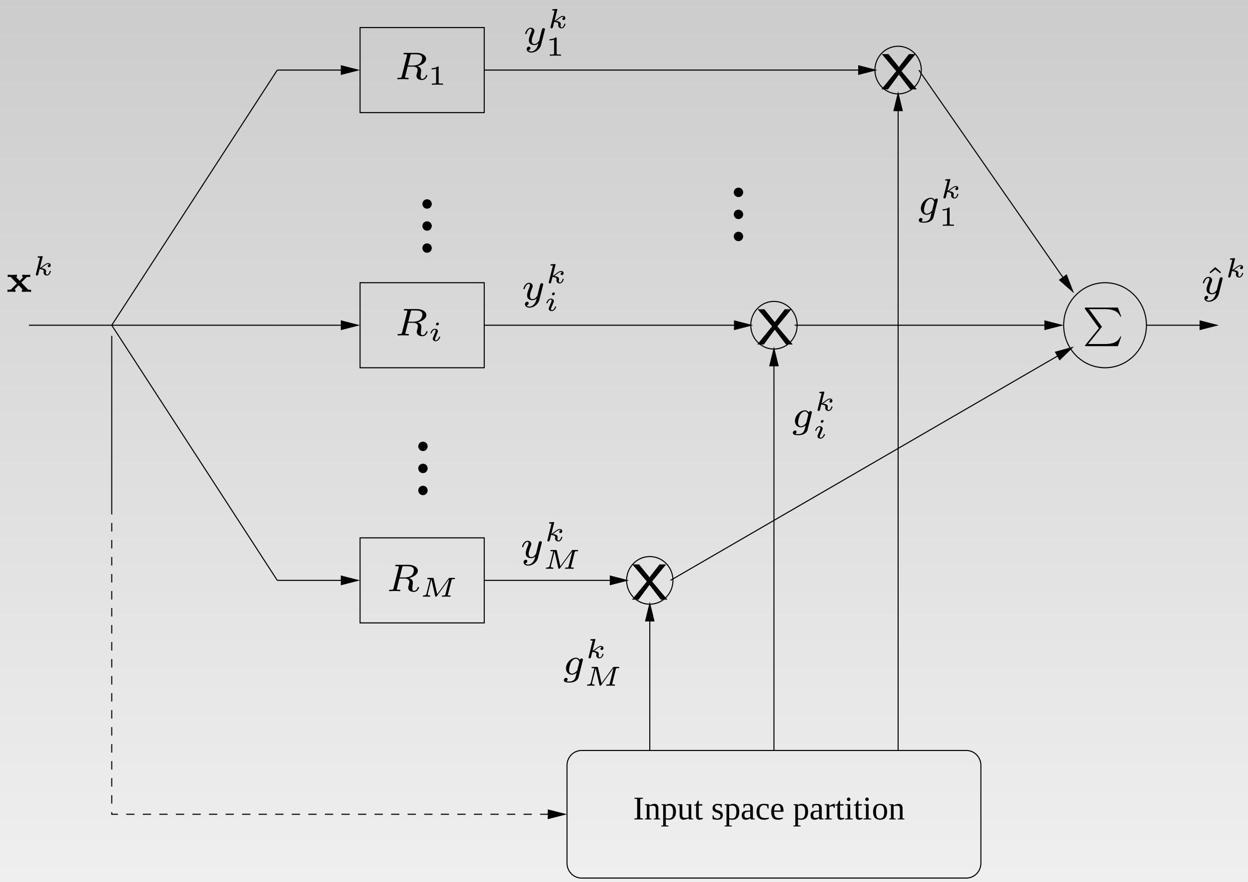


Figure 4: A general structure of a fuzzy-rule based system, composed by a total of M fuzzy rules.

A general structure

- $\mathbf{x}^k = [x_1^k,\ x_2^k,\ \ldots,\ x_p^k] \in \mathbb{R}^p$ is the input vector at instant k, $k \in \mathbb{Z}_0^+$;
- $\hat{y}^k \in \mathbb{R}$ is the estimate output;
- Given centers $\mathbf{c}_i \in \mathbb{R}^p$ and covariance matrices $\mathbf{V}_i$ $i = 1, \ldots, M$, membership degrees $g_i(\mathbf{x}^k)$ are defined as:

$$g_i(\mathbf{x}^k) = g_i^k = \frac{\alpha_i \cdot P[\, i \mid \mathbf{x}^k \,]}{\displaystyle\sum_{q=1}^{M} \alpha_q \cdot P[\, q \mid \mathbf{x}^k \,]} \tag{2}$$

with $\alpha_i \geq 0$, $\sum_{i=1}^{M} \alpha_i = 1$ and:

$$\begin{aligned} P[\, i \mid \mathbf{x}^k \,] \quad &= \quad \frac{1}{(2\pi)^{p/2} \det(\mathbf{V}_i)^{1/2}} \times \\ &\times \exp\left\{ -\frac{1}{2} (\mathbf{x}^k - \mathbf{c}_i) \mathbf{V}_i^{-1} (\mathbf{x}^k - \mathbf{c}_i)^T \right\} \end{aligned} \tag{3}$$

A general structure

- Each local model y_i^k, $i = 1, \ldots, M$ is estimated by a linear one:

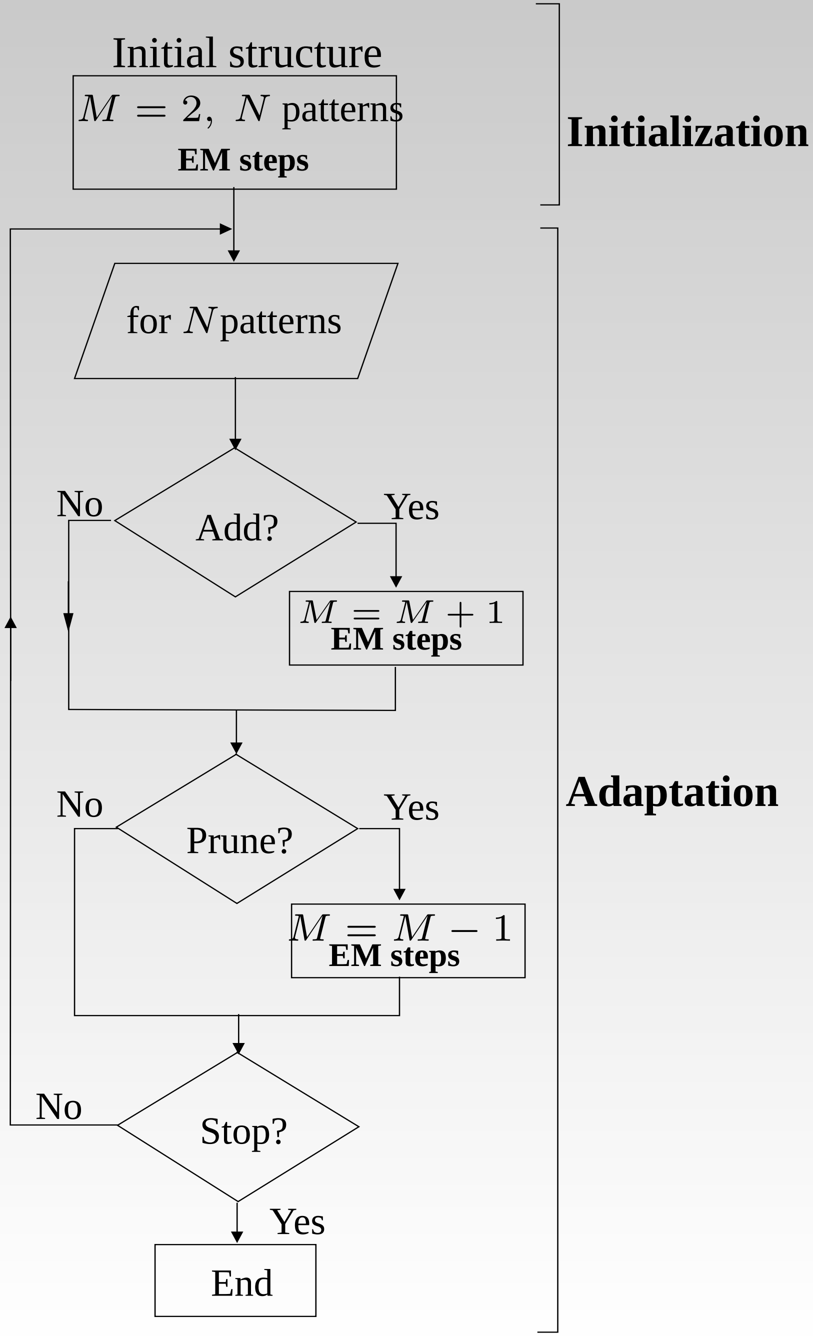
$$y_i^k = \phi^k \times {\theta_i}^T \tag{4}$$

where $\phi^k = [1 \; x_1^k \; x_2^k \; \ldots x_p^k]$ and $\theta_{\mathbf{i}} = [\theta_{i0} \; \theta_{i1} \; \ldots \; \theta_{ip}]$.

- The output model $\hat{y}^k$ is computed as:

$$\hat{y}^k = \sum_{i=1}^{M} g_i(\mathbf{x}^k) \; y_i^k \tag{5}$$

Constructive learning



- **E step**: g_i^k is estimated given $\mathbf{x}^k$ and $y^k \Rightarrow$ *posterior* estimate h_i^k;

$$h_i^k = \frac{\alpha_i P(i \mid \mathbf{x}^k) P(y^k \mid \mathbf{x}^k,\ \theta_i)}{\sum_{q=1}^{M} \alpha_q P(q \mid \mathbf{x}^k) P(y^k \mid \mathbf{x}^k,\ \theta_q)}$$

- **M step**:
  - Model parameters are adjusted;
  - Adding and pruning conditions are verified.

$$\alpha_i = \frac{1}{N} \sum_{k=1}^{N} h_i^k \qquad (6)$$

Second phase: Adaptation

- **Adding a new rule**: Assuming a normal input data distribution, with a confidence level equal to $\gamma\%$:

Ω is an i/o data set so that

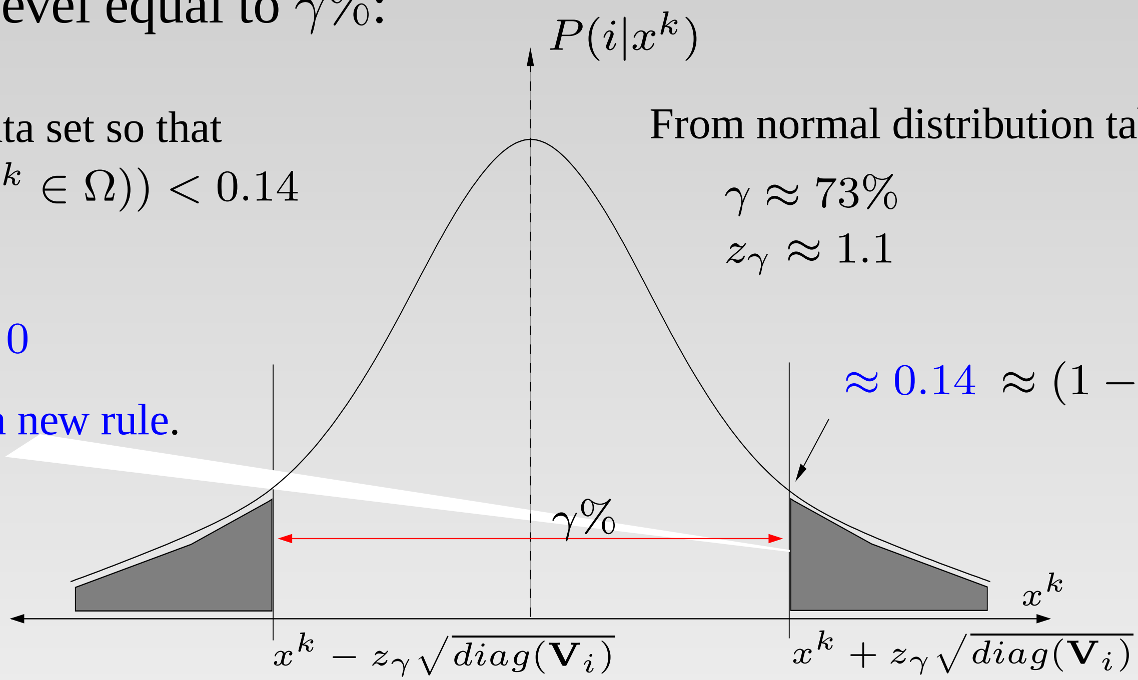
$$\max_{i=1,\ldots,M}(P(i|\mathbf{x}^k \in \Omega)) < 0.14$$

$\Rightarrow$ If $N_\Omega > 0$

Then create a new rule.

From normal distribution table:

$$\gamma \approx 73\%$$
$$z_\gamma \approx 1.1$$



- **Pruning a new rule**: α_i is proportional to the sum of all h_i^k. Thus, the more times the rule is strongly activated, the higher its α_i will be.

$\Rightarrow$ If $\alpha_i < \alpha_{min}$, then the $i-$th rule will be pruned.

Case study: NN3 competition

- Reduced data set;

Table 1: Global prediction errors for series NN3_102 and NN3_104.

Series	In sample 1 step ahead		Out of sample 1 step ahead		Out of sample 1 to 18 steps ahead	
NN3_102	$k = 4, \ldots, 108$		$k = 109, \ldots, 126$		$k = 109, \ldots, 126$	
NN3_104	$k = 4, \ldots, 97$		$k = 98, \ldots, 115$		$k = 98, \ldots, 115$	
Series	M	**sMAPE** (%)	**MAE** (u)	**sMAPE** (%)	**MAE** (u)	**sMAPE** (%)	**MAE** (u)
NN3_102	3	3.41	179.36	4.72	287.98	11.40	658.05
NN3_104	3	10.98	438.27	6.75	334.93	12.32	612.32

Case study: NN3 competition

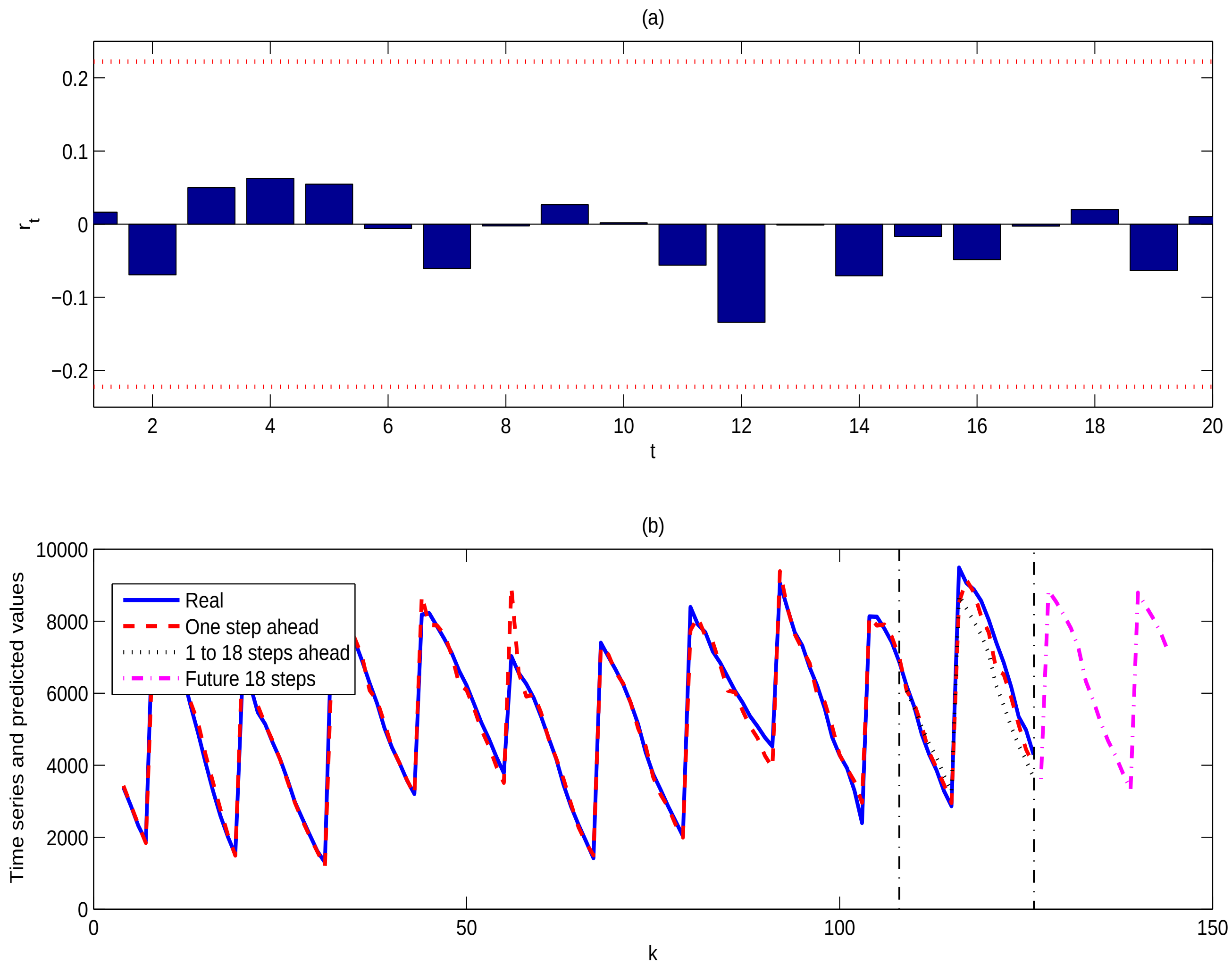


Figure 5: Multiple steps ahead: (a) autocorrelation coefficients estimates, (b) predictions for series NN3_102.

Case study: NN3 competition

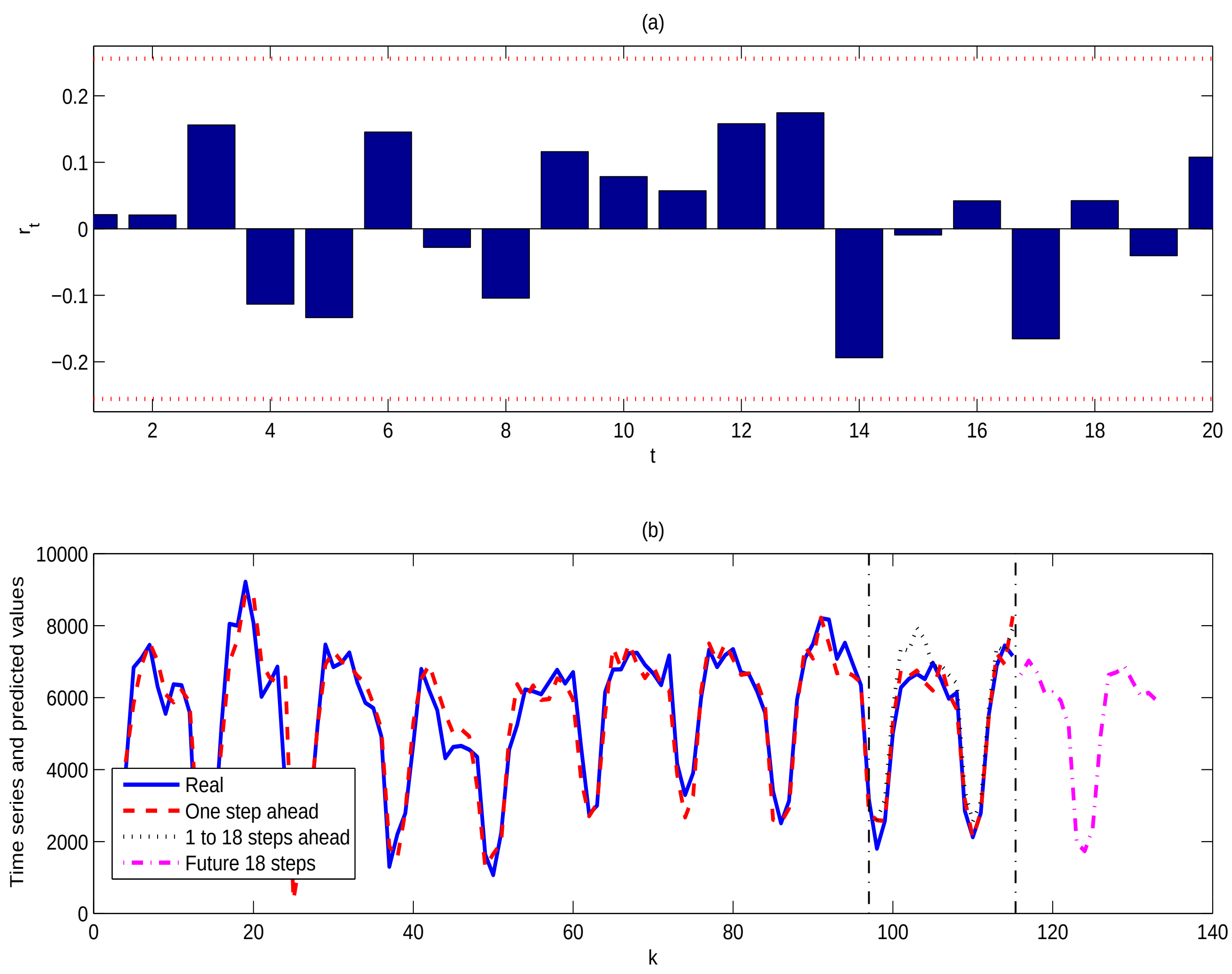


Figure 6: Multiple steps ahead: (a) autocorrelation coefficients estimates, (b) predictions for series NN3_104.

Case study: NN3 competition

Table 2: Some characteristics of input selection and model construction.

Time series	Difference	Num. inputs	Inputs (lags)	M
1	1	3	1, 2, 4	3
2	0	2	1, 3	3
3	0	2	1, 10	3
4	0	3	1, 2, 3	8
5	1	2	1, 2	6
6	0	3	2, 3, 4	3
7	0	1	1	2
8	0	1	2	2
9	1	2	2, 3	12
10	0	2	1, 6	5
11	0	1	1	2

Case study: NN3 competition

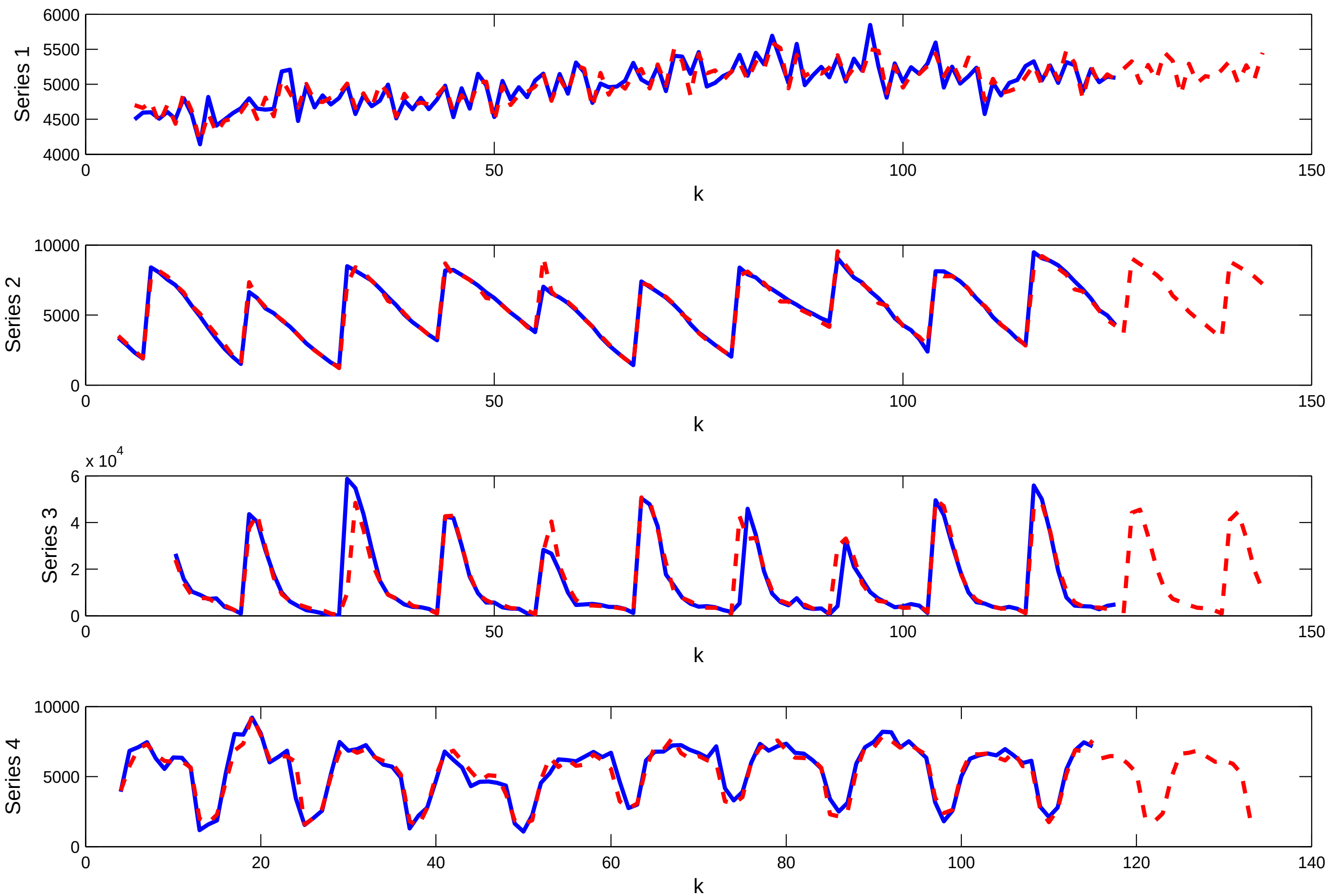


Figure 7: One and multi-step ahead forecasting for time series NN3_101 to NN3_104.

Case study: NN3 competition

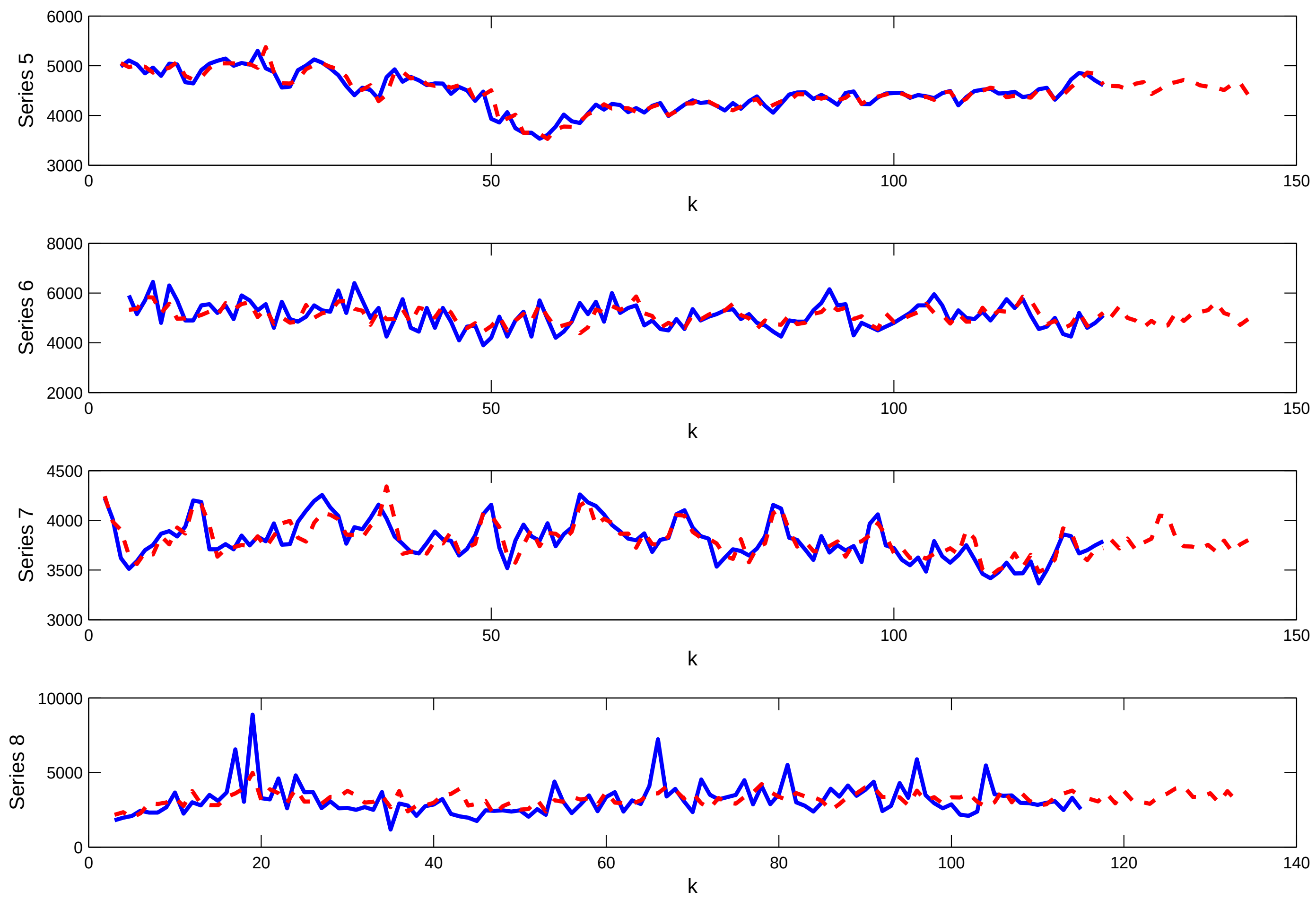


Figure 8: One and multi-step ahead forecasting for time series NN3_105 to NN3_108.

Case study: NN3 competition

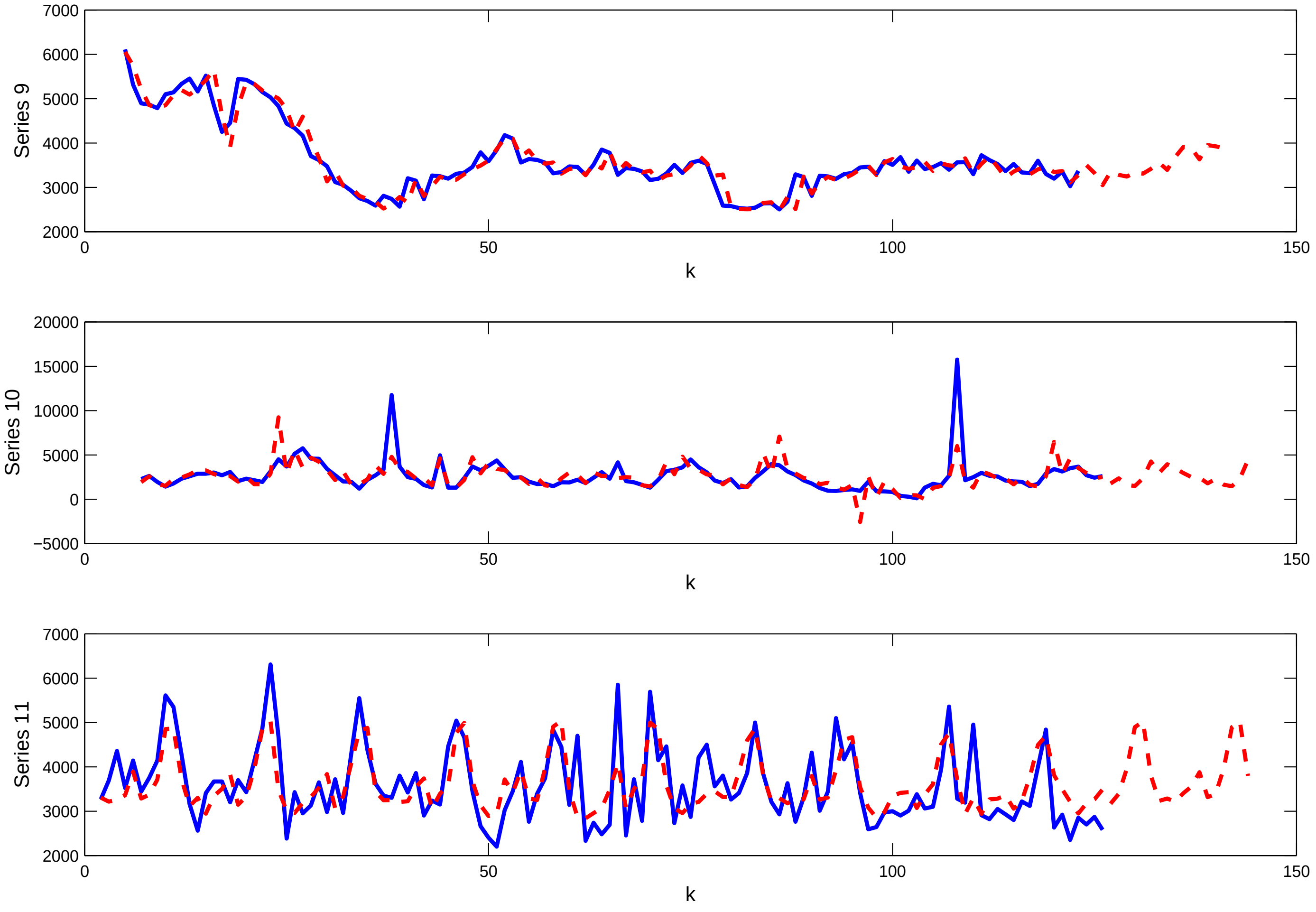


Figure 9: One and multi-step ahead forecasting for time series NN3_109 to NN3_111.

Conclusions and Future works

- This work presents a methodology for time series modeling.
- Statistical tools combined with novel methodologies provide adequate models.
- Objectives achieved:
  - The study of the different tasks that compose the methodology: from data pre-processing to model validation.
  - The automatic selection of a suitable model structure;
- What needs to be improved:
  - Initialization phase;
  - Adding and pruning conditions.

Thanks for your attention.

Ivette Luna

iluna@cose.fee.unicamp.br