2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

TIME SERIES FORECASTING WITH R:
from CLASSICAL to MODERN Methods

Suhartono

(B.Sc.-ITS; M.Sc.-UMIST,UK; Dr.-UGM; Postdoctoral-UTM)

Department of Statistics,
Institut Teknologi Sepuluh Nopember, Indonesia

Email: *suhartono@statistika.its.ac.id*, *gmsuhartono@gmail.com*

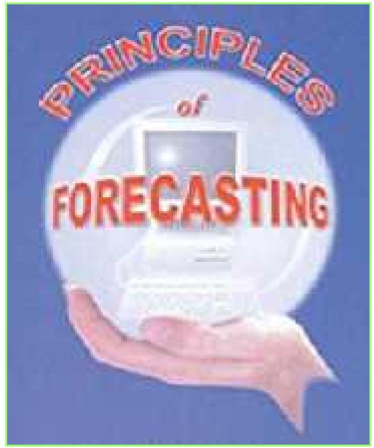
Department of Mathematics, Universitas Andalas, Padang

17-18 July 2017

2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

TIME SERIES FORECASTING WITH R: from CLASSICAL to MODERN Methods



Department of Mathematics, Universitas Andalas, Padang

17-18 July 2017

25 years of time series forecasting

De Gooijer & Hyndman (International Journal of Forecasting, 2006)

- 25 years of time series forecasting
  - Introduction
  - Exponential smoothing
    - Preamble
    - Variations
    - State space models
    - Method selection
    - Robustness
    - Prediction intervals
    - Parameter space and model properties
  - ARIMA models
    - Preamble
    - Univariate
    - Transfer function
    - Multivariate
  - Seasonality
  - State space and structural models and the Kalman filter

1980 ↓ 2005

- Nonlinear models
  - Preamble
  - Regime-switching models
  - Functional-coefficient model
  - Neural nets
  - Deterministic versus stochastic dynamics
  - Miscellaneous
- Long memory models
- ARCH/GARCH models
- Count data forecasting
- Forecast evaluation and accuracy measures
- Combining
- Prediction intervals and densities
- A look to the future
- Acknowledgments
- References

Motivation

☞ **The M3-Competition**: results, conclusions and implications

(1) Statistically sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones.

(2) The relative ranking of the performance of the various methods varies according to the accuracy measure being used.

(3) The accuracy when various methods are being combined outperforms, on average, the individual methods being combined and does very well in comparison to other methods.

(4) The accuracy of the various methods depends upon the length of the forecasting horizon involved.

Makridakis & Hibon (International Journal of Forecasting, 2000)

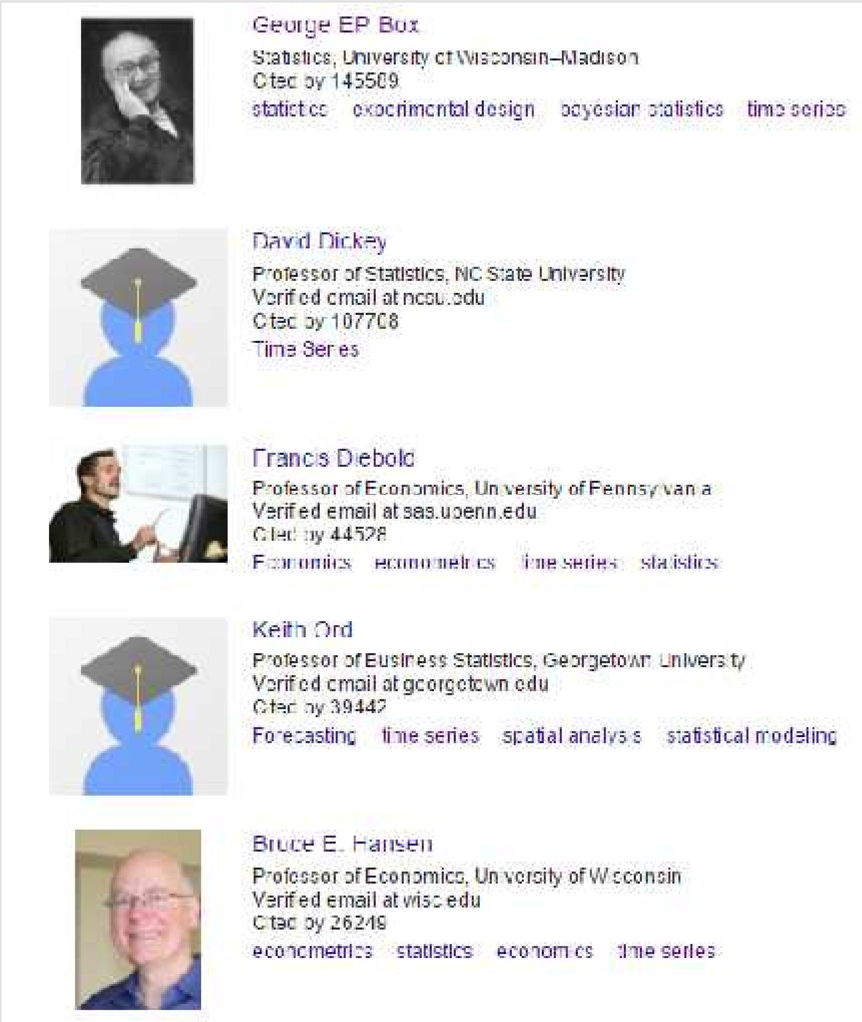
Material

1. Time Series Regression (TSR) & ARIMA model
   - ☞ Seasonal models: Multiplicative, Additive, Subset
   - ☞ Multiple Seasonal models.
2. ARIMAX & Multivariate Time Series Model
   - ☞ Intervention Model & Outlier Detection
   - ☞ Calendar Variation Model, Transfer Function Model.
3. Nonlinear Time Series (Modern) Models
   - ☞ Non-linearity test, Neural Networks.
4. Hybrid Models
   - ☞ TSR-NN, ARIMA-NN, ARIMAX-NN.

Reference

1. Armstrong, J.S. (2002)
   Principles of Forecasting: A Handbook for Researchers and Practicioners, Kluwer Academic Publisher.
2. Hanke, J.E. and Reitsch, A.G. (1995, 2005, 2008)
   Business Forecasting, 5th, 7th and 9th edition, Prentice Hall.
3. Bowerman, B.L. and O'Connell, R.T. (1993, 2004)
   Forecasting and Time Series: An Applied Approach,
   3rd and 4th edition, Duxbury Press: USA.
4. Makridakis, S., Wheelwright, S. C. and Hyndman, R. J. (1998)
   Forecasting: Method and Applications, New York: Wiley & Sons.
5. Wei, W.W.S. (1990, 2006)
   Time Series Analysis: Univariate and Multivariate Methods
   Addison-Wesley Publishing Co., USA.

10 Scholars in Time Series



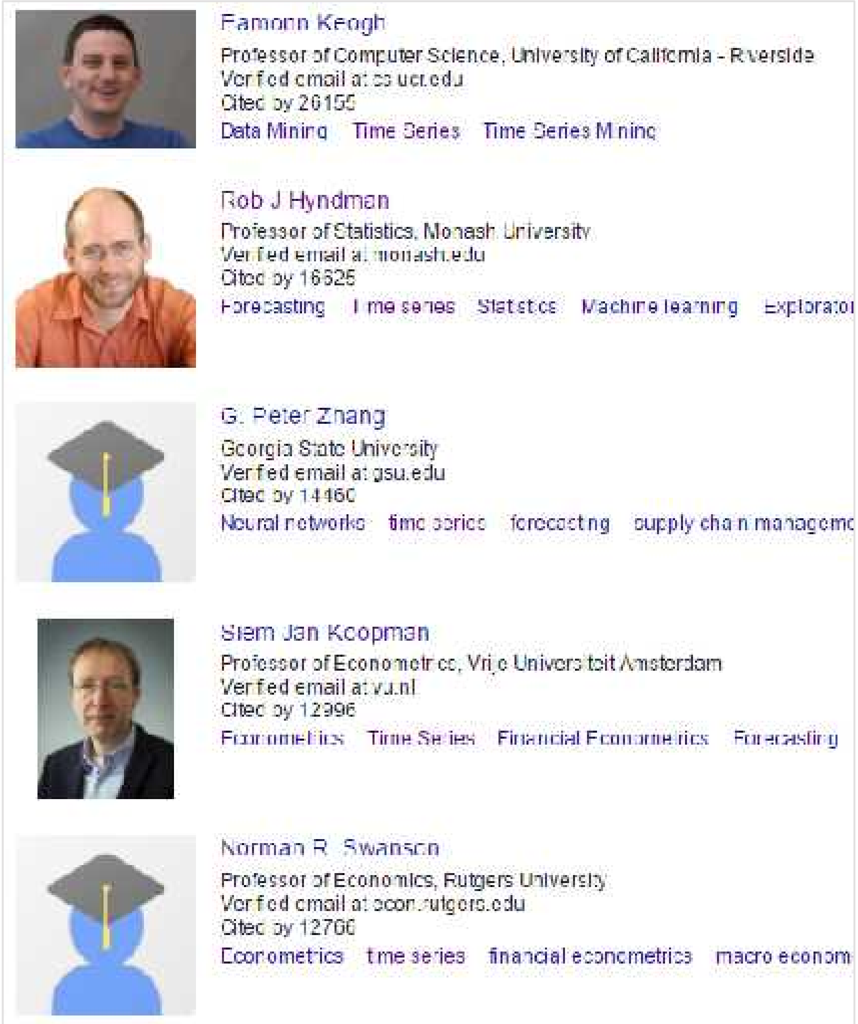
George EP Box
Statistics, University of Wisconsin–Madison
Cited by 145509
statistics · experimental design · bayesian statistics · time series

David Dickey
Professor of Statistics, NC State University
Verified email at ncsu.edu
Cited by 107708
Time Series

Francis Diebold
Professor of Economics, University of Pennsylvania
Verified email at sas.upenn.edu
Cited by 44528
Economics · econometrics · time series · statistics

Keith Ord
Professor of Business Statistics, Georgetown University
Verified email at georgetown.edu
Cited by 39442
Forecasting · time series · spatial analysis · statistical modeling

Bruce E. Hansen
Professor of Economics, University of Wisconsin
Verified email at wisc.edu
Cited by 26249
econometrics · statistics · economics · time series



Eamonn Keogh
Professor of Computer Science, University of California - Riverside
Verified email at cs.ucr.edu
Cited by 20155
Data Mining · Time Series · Time Series Mining

Rob J Hyndman
Professor of Statistics, Monash University
Verified email at monash.edu
Cited by 16626
Forecasting · Time series · Statistics · Machine learning · Explorato

G. Peter Zhang
Georgia State University
Verified email at gsu.edu
Cited by 14460
Neural networks · time series · forecasting · supply chain manageme

Siem Jan Koopman
Professor of Econometrics, Vrije Universiteit Amsterdam
Verified email at vu.nl
Cited by 12996
Econometrics · Time Series · Financial Econometrics · Forecasting

Norman R. Swanson
Professor of Economics, Rutgers University
Verified email at econ.rutgers.edu
Cited by 12766
Econometrics · time series · financial econometrics · macro econom

10 Scholars in Forecasting

j scott armstrong
The Wharton School, University of Pennsylvania
Verified email at wharton.upenn.edu
Cited by 21370
Forecasting advertising survey research marketing peer review

ROBIN MILES HOGARTH
Universitat Pompeu Fabra
Verified email at upf.edu
Cited by 19745
Decision making judgment forecasting

Keith Ord
Professor of Business Statistics, Georgetown University
Verified email at georgetown.edu
Cited by 19505
Forecasting time series spatial analysis statistical modeling

VS Subrahmanian
University of Maryland College Park
Verified email at cs.umd.edu
Cited by 14649
Artificial Intelligence Databases Social Networks Behavior Analytics Forecas

Luis G. Vargas
Professor of Business Analytics and Operations, University of Pittsburgh
Verified email at pitt.edu
Cited by 13141
Operations DecisionTheory Analytic Hierarchy Process Revenue Managemen

Ronald Lee
Professor of Graduate Studies in Demography and Economics
Verified email at demog.berkeley.edu
Cited by 12238
demography population aging economic demography inter

Michael Lewis-Beck
University of Iowa - Department of Political Science
Verified email at uiowa.edu
Cited by 12094
Elections Voting Behavior Forecasting Research Methodolo

Spyros Makridakis
Rector, Neapolis University of Pafos (NUP) and Emeritus Professor,
Verified email at insead.edu
Cited by 11651
Forecasting Decision Making Medical Research

Rob J Hyndman
Professor of Statistics, Monash University
Verified email at monash.edu
Cited by 10267
Forecasting Time series Statistics Machine learning Data

Siem Jan Koopman
Professor of Econometrics, VU University Amsterdam
Verified email at vu.nl
Cited by 9707
Econometrics Time Series Financial Econometrics Forecast

Indonesian Scholars in Time Series

sinta
Science and Technology Index

Search Area time series Submit

1 Page: 1 Items per page: 10

No	Area of Interest	Authors
1	Time Series	10
2	Time Series Analysis	7
3	Time Series Prediction	1
4	Time series forecasting	1
5	Spatial Time Series	1
6	Fuzzy Time Series	1

1

No		Peneliti
1		SUHARTONO Institut Teknologi Sepuluh Nopember
2		ANSARI SALEH AHMAR Universitas Negeri Makassar
3		SUPARMAN Universitas Ahmad Dahlan

UNAND Scholars in sinta

Science and Technology Index

« Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... | 143 | Next » Page: 2 ▼ Items per page: 10 ▼

15	Universitas Andalas *UNAND*	467	15676

No	Author Name	Google Scholar Indexed			Scopus Indexed		
		Citation	i10-index	H-index	Citation	i10-index	H-index
1	DAYAR ARBAIN *Universitas Andalas*	401	16	13	247	7	9
2	RAHMIANA ZEIN *Universitas Andalas*	354	8	9	271	7	8
3	DACHRIYANUS *Universitas Andalas*	259	7	8	172	4	7

UNAND Scholars in sinta

FERRA YANUAR

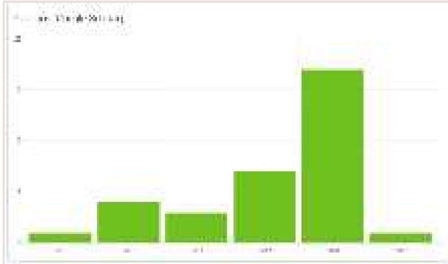
Universitas Andalas

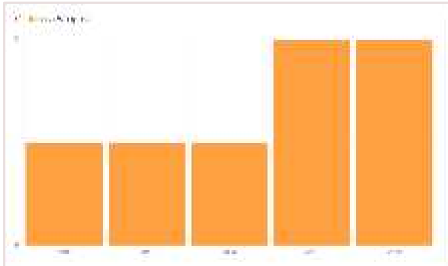
Mathematics

SINTA Score : 3

(SINTA Score is composed based on quantity, quality and impact of published documents)

	Articles	Citations	H-Index	i10-Index
Scopus	7	19	2	1
Google Scholar	22	33	3	2
Inasti	0	0	0	0





Hybrid Model in Time Series Forecasting

Google

Time series forecasting using a hybrid

Cendekia

Sekitar 51.600 hasil (0,07 dtk)

Artikel

Koleksiku

Kapan saja

Sejak 2017

Sejak 2016

Sejak 2013

Rentang khusus...

2000 —

Telusuri

Urutkan menurut relevansi

Time series forecasting using a **hybrid** ARIMA and neural network model
GP Zhang - Neurocomputing, 2003 - Elsevier
Autoregressive integrated moving average (ARIMA) is one of the popular linear models in **time series forecasting** during the past three decades. Recent research activities in **forecasting** with artificial neural networks (ANNs) suggest that ANNs can be a promising
Dirujuk 1417 kali Artikel terkait 10 versi Kutip Tersimpan

Financial **time series forecasting using** support vector machines
K Kim - Neurocomputing, 2003 - Elsevier
... [15] showed the applicability of SVM to **time-series forecasting**. Recently, Tay and Cao [18] examined the predictability of financial **time-series** including five **time series** data with **...** Since we attempt to **forecast** the direction of daily price change in the stock price index, technical **...**
Dirujuk 1025 kali Artikel terkait 20 versi Kutip Simpan

Time-series forecasting using flexible neural tree model
Y Chen, B Yang, J Dong, A Abraham - Information sciences, 2005 - Elsevier
... A **hybrid** learning algorithm for evolving the neural tree models is given in Section 3. Section 4 presents some simulation results for two **time-series forecasting** problems **...** reason for choosing the representation is that the tree can be created and evolved **using** the existing **...**
Dirujuk 274 kali Artikel terkait 65 versi Kutip Simpan

Hybrid Model in Time Series Forecasting

Google | Time series forecasting using a hybrid

Cendekia

Artikel

Koleksiku

Kapan saja

Sejak 2017

Sejak 2016

Sejak 2013

Rentang khusus...

Urutkan menurut relevansi

Urutkan menurut tanggal

Hybrid methodology for tuberculosis incidence **time-series forecasting** based on ARIMA and a NAR neural network

KW Wang, C Deng, JP Li, YY Zhang, XY Li... - Epidemiology & ..., 2017 - cambridge.org

... Page 11. modelled by the NAR model. Moreover, this study compares the results obtained from the **hybrid** model with the **forecast** results from the single ARIMA model. **...** Comparison of ARIMA, neural networks and **hybrid** models in **time series**: tourist arrival **forecasting**. **...**

Artikel terkait 5 versi Kutip Simpan

Hybrid ARIMA-BPNN model for **time series** prediction of the Chinese stock market

L Xiong, Y Lu - ... (ICIM), 2017 3rd International Conference on, 2017 - ieeexplore.ieee.org

... 35, pp. 670-680, October 2015. [3] CH Su and CH Cheng, "A **hybrid** fuzzy **time series** model based on ANFIS and integrated nonlinear feature selection method for **forecasting** stock," Neurocompting, vol. 205, pp. **...** [5] XF Lu, DF Q and GX Cao, "Volatility **Forecast** Based on **...**

Kutip Simpan

Hybrid DARIMA-NARX model for **forecasting** long-term daily inflow to Dez reservoir **using** the North Atlantic Oscillation (NAO) and rainfall data

ME Banihabib, A Ahmadian, FS Jamali - GeoResJ, 2017 - Elsevier

... Autoregressive integrated moving average (ARIMA) models (classified as **time series** models) and artificial neural **...** index (IIFFE) to assess mean absolute relative error (MARE), **time** to peak **...** **forecaster** has the most prominent influence on the increasing **forecasting** accuracy, while **...**

Artikel terkait Kutip Simpan

“Padang” in Time Series Forecasting research

Google | peramalan di padang, indonesia

PERAMALAN PASOKAN BAHAN BAKU DAN PENJUALAN SIR 20 DI PT. PERKEBUNAN NUSANTARA VII UNIT **PADANG** PELAWI KEC. SUKARAJA KAB. SELUMA

EM Manihuruk, MM Romdhon - JURNAL AGRISEP, 2016 - ejournal.unib.ac.id

... Dengan adanya perhitungan **peramalan** menggunakan software e-views dengan data pasokan ... Perkebunan VII Unit **Padang** Pelawi diketahui bahwa metode Arima memiliki nilai ... yang menjadi masalah bagi perusahaan dalam memproduksi Standard **Indonesia** Rubber (SIR). ...

Artikel terkait 2 versi Kutip Simpan

Peramalan Kebutuhan Energi Jual pada PT Perusahaan Listrik Negara (PLN) Cabang Bukittinggi dengan Menggunakan Metode Dekomposisi Sensus Ii

S Wahyuni, H Helma, N Amalita - UNP Journal of Mathematics, 2014 - ejournal.unp.ac.id

Page 1. **Peramalan** Kebutuhan Energi Jual pada PT Perusahaan Listrik Negara (PLN) Cabang Bukittinggi dengan Menggunakan Metode Dekomposisi Sensus Ii Sujantri Wahyuni1, Helma2, Nonong Amalita3 1 Mathematics Department State University of **Padang**, **Indonesia** ...

Kutip Simpan

"Padang" in Time Series Forecasting research

Google | peramalan di padang, indonesia

Peramalan Harga Ayam Broiler di Lima Kota di Sumatera Barat
A Amri - 2009 - repository.ipb.ac.id
... nilai MAD terkecil untuk meramalkan harga ayam broiler di Kota **Padang** dan Payakumbuh ... Sedangkan model **peramalan** untuk Kota Solok dan Kabupaten Tanah Datar adalah winter aditif lag ... Raya Darmaga Kampus IPB Darmaga Bogor 16680 West Java **Indonesia** All rights ...
Kutip Simpan Lainnya

[DOC] Teknik baru statistika dalam **peramalan** curah hujan ekstrim untuk penentuan musim tanam produk–produk pertanian
M Irfan, A Santoso - 2011 - repository.ipb.ac.id
... iklim ditengarai menyebabkan sekitar 400 hektare sawah di Kabupaten **Padang** Pariaman tidak ... penggunaan metode alternative statistika ini akan mampu melakukan **peramalan** lebih akurat ... sebelumnya dengan melihat nilai RMSE dari masing – masing hasil **peramalannya**. ...
Kutip Simpan Lainnya

PROGRAM KREATIVITAS MAHASISWA

"Padang" in Time Series Forecasting research

Google: peramalan di padang, indonesia

... pendukung keputusan efisiensi penggunaan sumber daya di Rumah Sakit berdasarkan prediksi kunjungan pasien:: Studi kasus di Rumah Sakit Semen **Padang**

R AFYENNI, S Hartati - 2005 - etd.repository.ugm.ac.id

... Sistem pendukung keputusan efisiensi penggunaan sumber daya di Rumah Sakit berdasarkan prediksi kunjungan pasien :: Studi kasus di Rumah Sakit Semen **Padang**. Penulis. ... Kata kunci : Sistem Pendukung Keputusan, Model, **Peramalan**. ... Bahasa, **Indonesia**. Jenis, Tesis. ...

Kutip Simpan Lainnya

Sistem pendukung keputusan efisiensi penggunaan sumber daya di Rumah Sakit berdasarkan prediksi kunjungan pasien :: Studi kasus di Rumah Sakit Semen Padang

Penulis

Afyenni, Rita

Pembimbing: Dra. Sri Hartati, M.Sc.,Ph.D

"Padang" in Time Series Forecasting research

Google | peramalan di padang, indonesia

Peramalan Kebutuhan Manajemen Logistik Pada Usaha Depot Air Minum Isi Ulang Al-Fitrah
H Yulius - Jurnal EDik Informatika, 2017 - ejournal.stkip-pgri-sumbar.ac.id
... ISSN : 2407-0491 E-ISSN : 2541-3716 **Peramalan** Kebutuhan Manajemen Logistik Pada Usaha Depot Air Minum Isi Ulang Al-Fitrah Henny Yulius 1 , Islami Yetti 2 Universitas Putra **Indonesia** „YPTK" **Padang** henny_yulius27@yahoo.com **...**
Kutip Simpan

Pengendalian Perencanaan Produksi Premium Dan Harga Pesan Crude Oil Ekonomis Menggunakan Metode **Peramalan** Dan Economic Order Quantity (Studi Kasus ...
H Yulius - Jurnal EDik Informatika, 2017 - ejournal.stkip-pgri-sumbar.ac.id
... Raya Lubuk Begalung **Padang** – Sumatera Barat E-mail : henny_yulius27@yahoo.com, daviddeska_p@yahoo.com **...** hasil perbandingan tabel tersebut diatas dengan nilai SEE yang terkecil maka dapat diketahui **peramalan** perencanaan produksi **...** Chevron pasific **Indonesia** (PT. **...**
Artikel terkait Kutip Simpan

"Padang" in Time Series Forecasting research

Google | peramalan di padang, indonesia

[PDF] PENENTUAN RESIKO INVESTASI DENGAN MODEL GARCH PADA INDEKS HARGA SAHAM PT. INDOFOOD SUKSES MAKMUR TBK.
L Mahlindiani, H Yozza - Jurnal Matematika UNAND, 2017 - jmua.fmipa.unand.ac.id
... dan Ilmu Pengetahuan Alam, Universitas Andalas, Kampus UNAND Limau Manis **Padang**, **Indonesia**, email : laramahlindiani@yahoo ... Perusahaan ini juga ter- gabung di Bursa Efek **Indonesia**. ... Berikutnya akan dihitung nilai **peramalan** return, **peramalan** varian dan volatili- tas. ...
Kutip Simpan Lainnya

[PDF] PEMODELAN DAN **PERAMALAN** DATA DERET WAKTU DENGAN METODE SEASONAL ARIMA
A ul Ukhra - Jurnal Matematika UNAND, 2014 - jmua.fmipa.unand.ac.id
... Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas, Kampus UNAND Limau Manis **Padang**, **Indonesia**, annisaulukhra25@gmail.com ... dengan model **peramalannya** adalah ... Pemodelan dan **Peramalan** Data Deret Waktu dengan Metode Seasonal ARIMA 65 ...
Artikel terkait Kutip Simpan Lainnya

Jurnal Matematika UNAND
Vol. VI No. 1 Hal. 25 – 32

TIME SERIES FORECASTING WITH R: from CLASSICAL to MODERN Methods

JURUSAN MATEMATIKA | FAKULTAS MIPA

UNIVERSITAS ANDALAS

E-JURUSAN

LABORATORIUM

Matematika Kombinatorika
Analisis dan Geometri
MATH LAB
Aljabar
Matematika Terapan
Statistika dan Teori Peluang

JURNAL MATEMATIKA

Jurnal Matematika
Universitas Andalas

AKREDITASI

AKREDITASI
BAN-PT
A
Nomor 2515/SK/BAN-PT/Akred/S/X/2016

2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

DOUBLE SEASONAL ARIMA MODEL WITH R, MINITAB AND SAS

Suhartono

(B.Sc.-ITS; M.Sc.-UMIST,UK; Dr.-UGM; Postdoctoral-UTM)

Department of Statistics,
Institut Teknologi Sepuluh Nopember, Indonesia

Email: *suhartono@statistika.its.ac.id*, *gmsuhartono@gmail.com*

Department of Mathematics, Universitas Andalas, Padang

17-18 July 2017

Motivation

- Develop the best forecast ARIMA model for Short-term Electricity Load Data
- **MINITAB**: Descriptive evaluation about the pattern of DOUBLE Seasonal Time Series Data
- **R**: Theoretical ACF and PACF of DOUBLE Seasonal ARIMA model
- **MINITAB**: The Data: identification of stationary & tentative order of DOUBLE Seasonal ARIMA
- **SAS**: Estimation & Diagnostic check.
- Discussion

Review: DSARIMA - TSARIMA

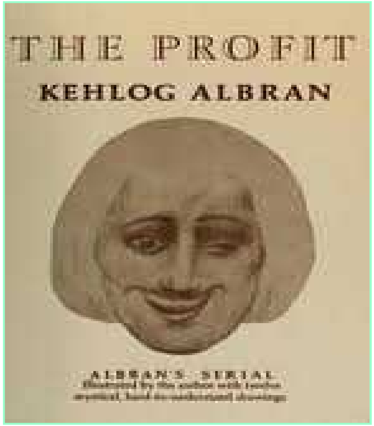
- **Multiple Seasonal ARIMA models:**
  Double Seasonal ARIMA, Triple Seasonal ARIMA model.

Website: http://users.ox.ac.uk/~mast0315/

Introduction

☞ **Kehlog Albran**

"*The Profit* (1973)"

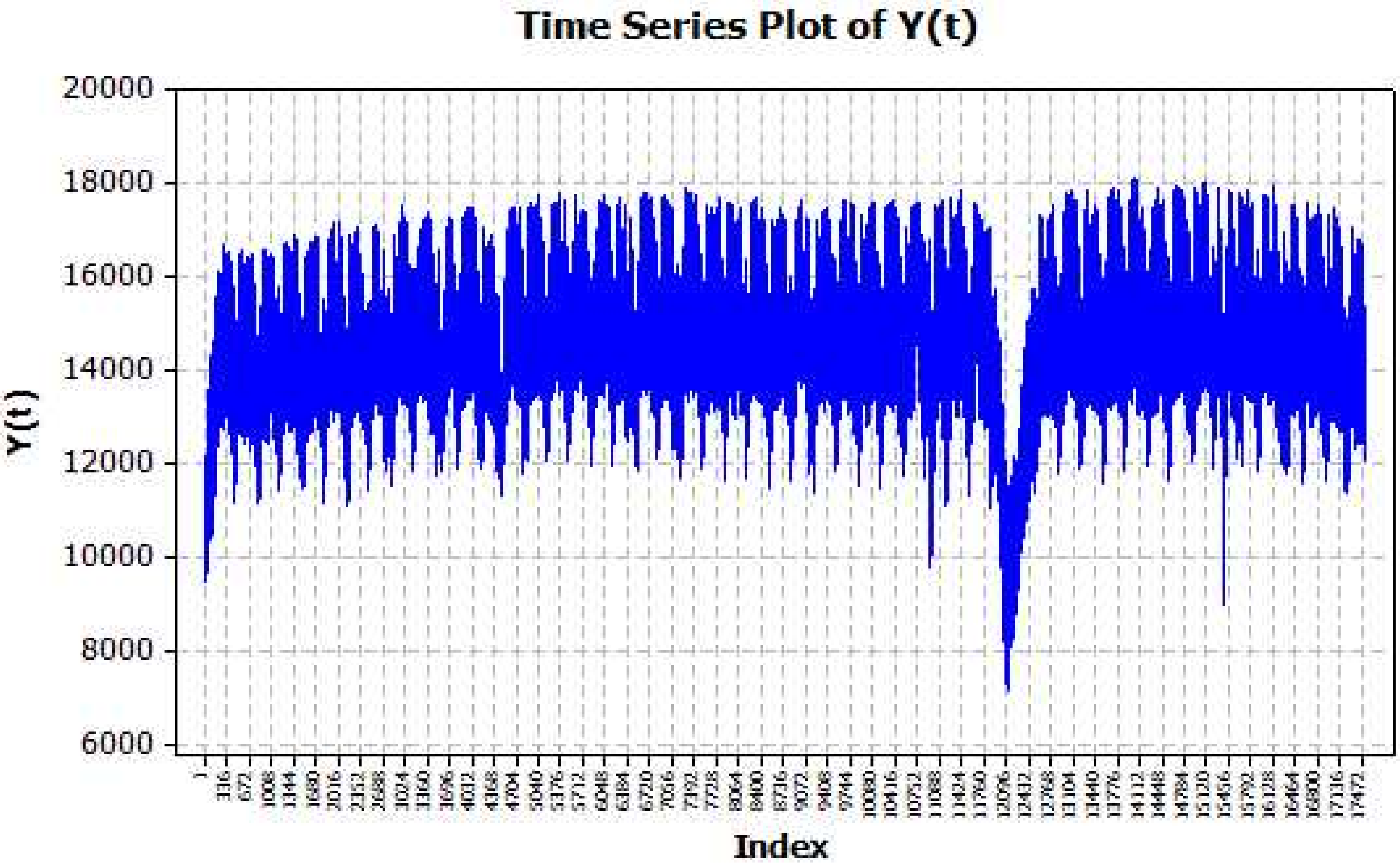


I have seen the future

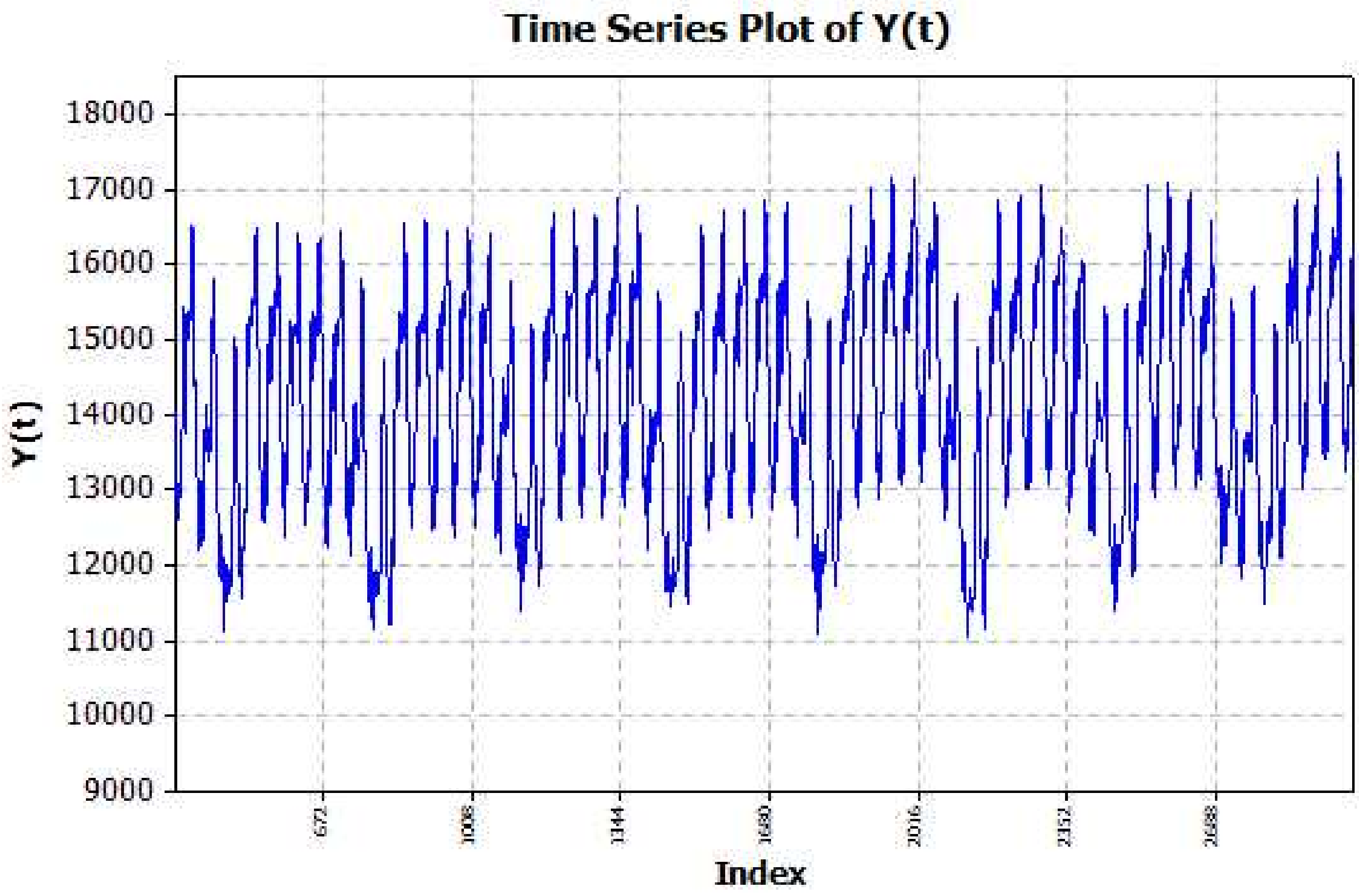
and it is just like the present,

only longer.

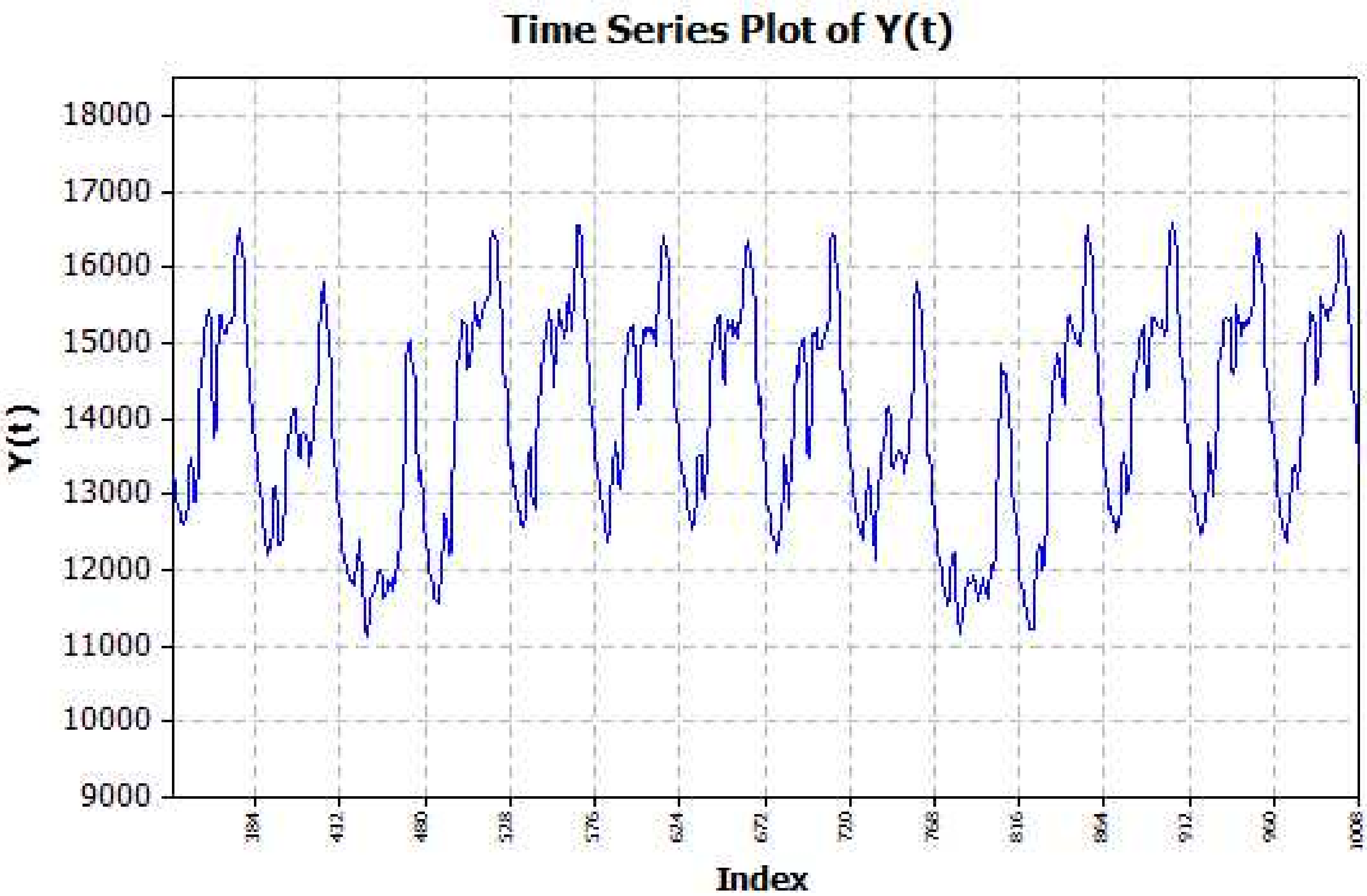
Problem: Prediction of half hourly load data



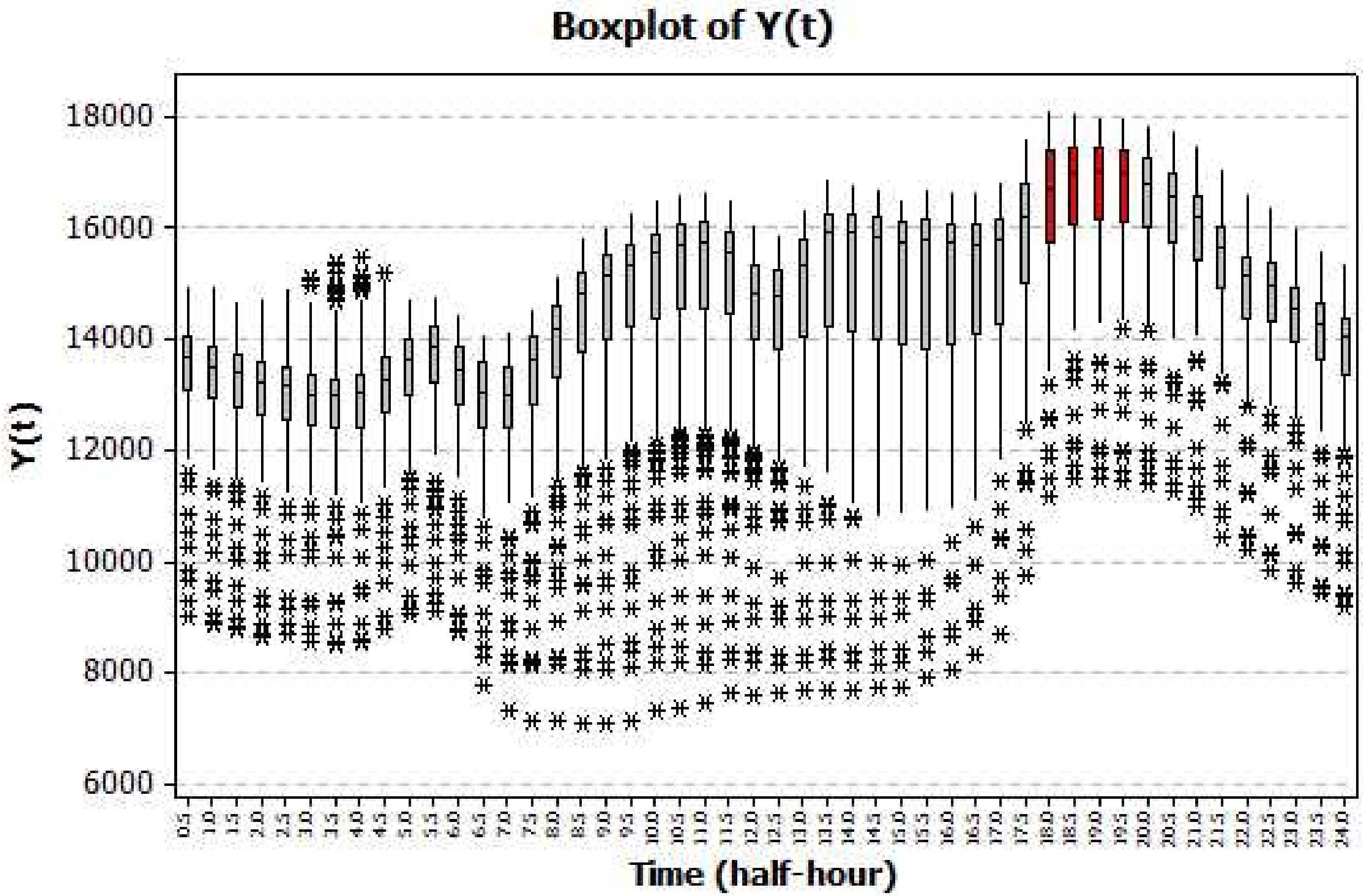
MINITAB Descriptive: half hourly load data



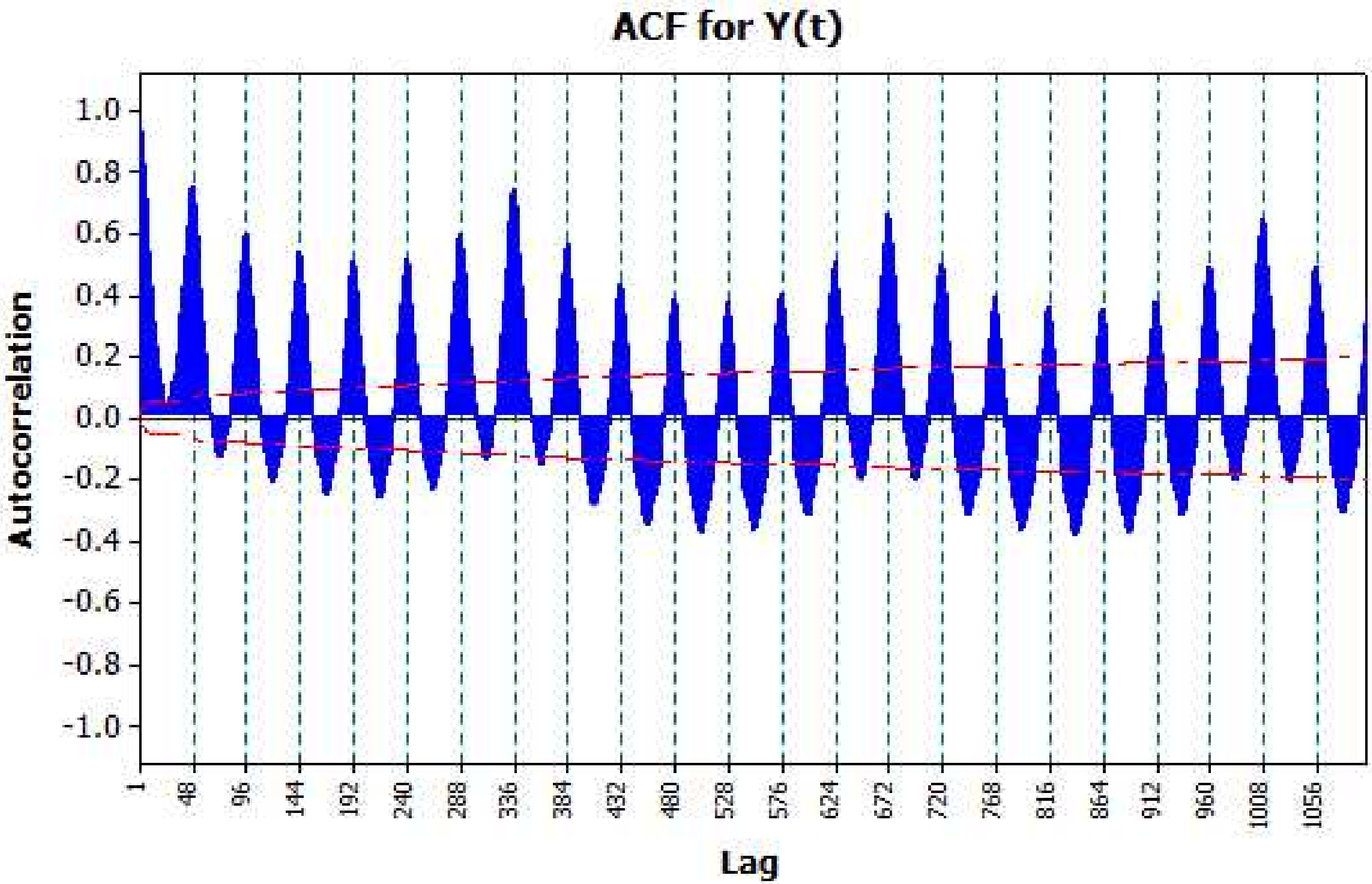
MINITAB Descriptive: half hourly load data



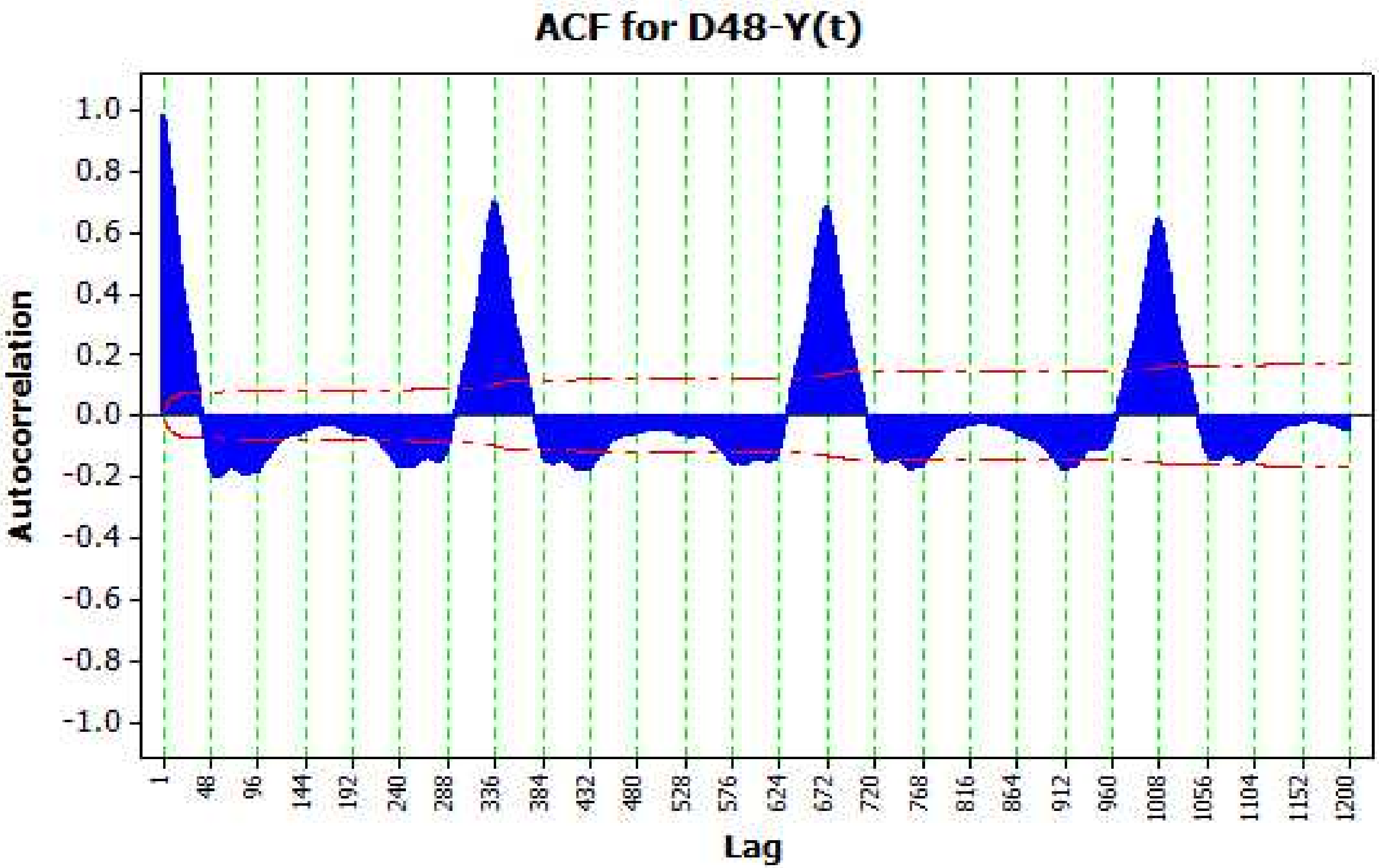
MINITAB Descriptive: half hourly load data



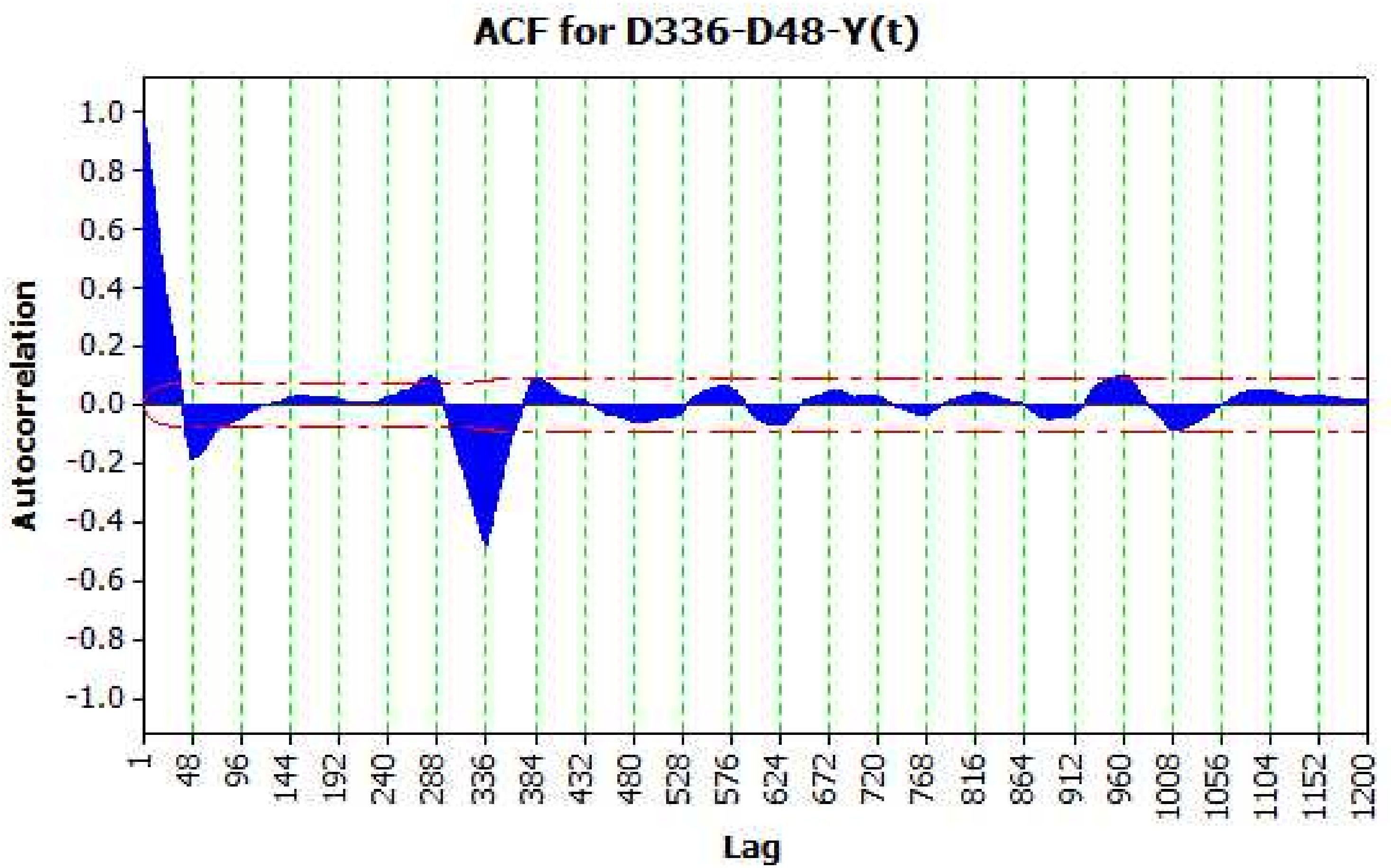
MINITAB Identification: half hourly load data



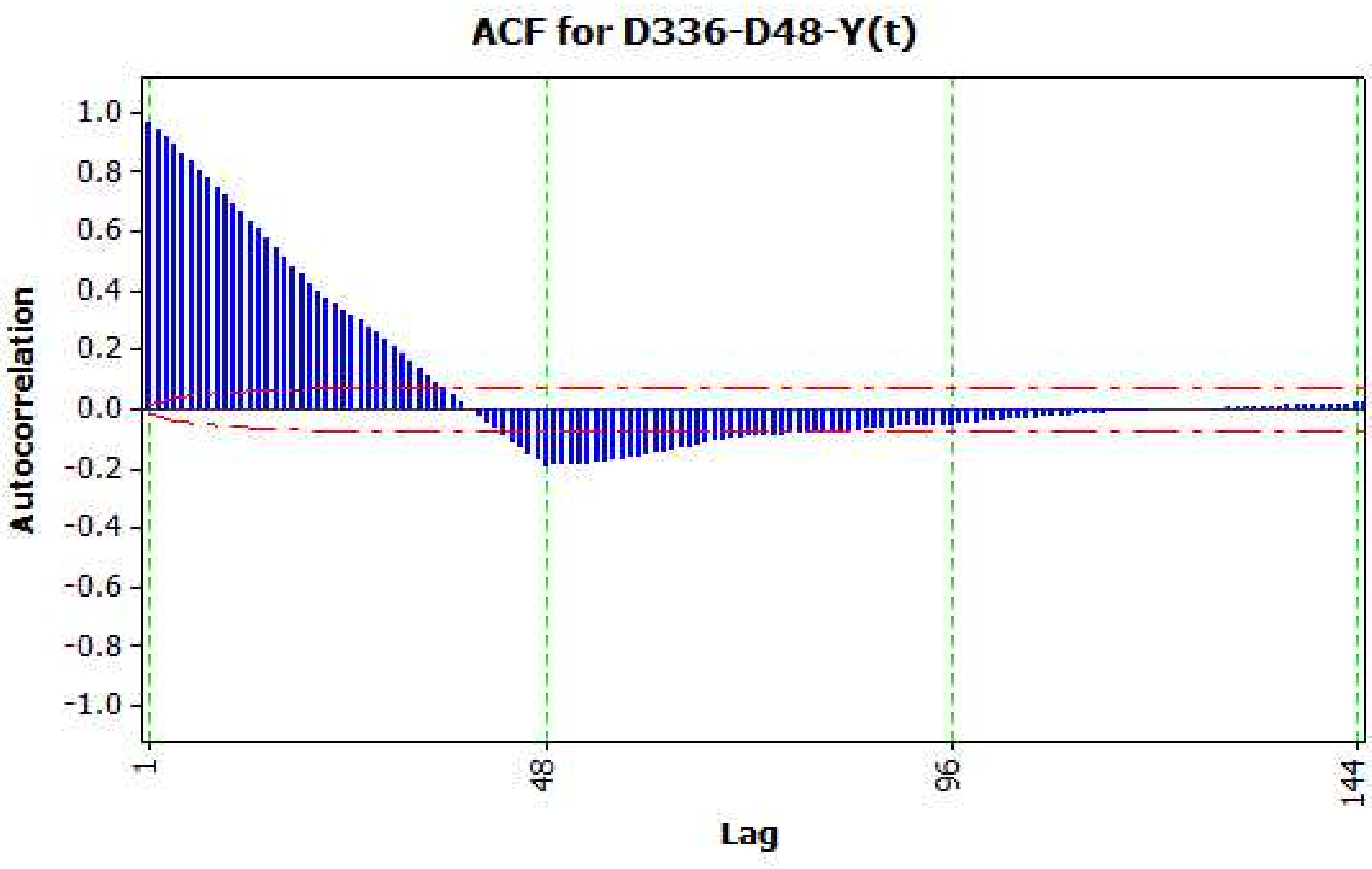
MINITAB Identification: half hourly load data



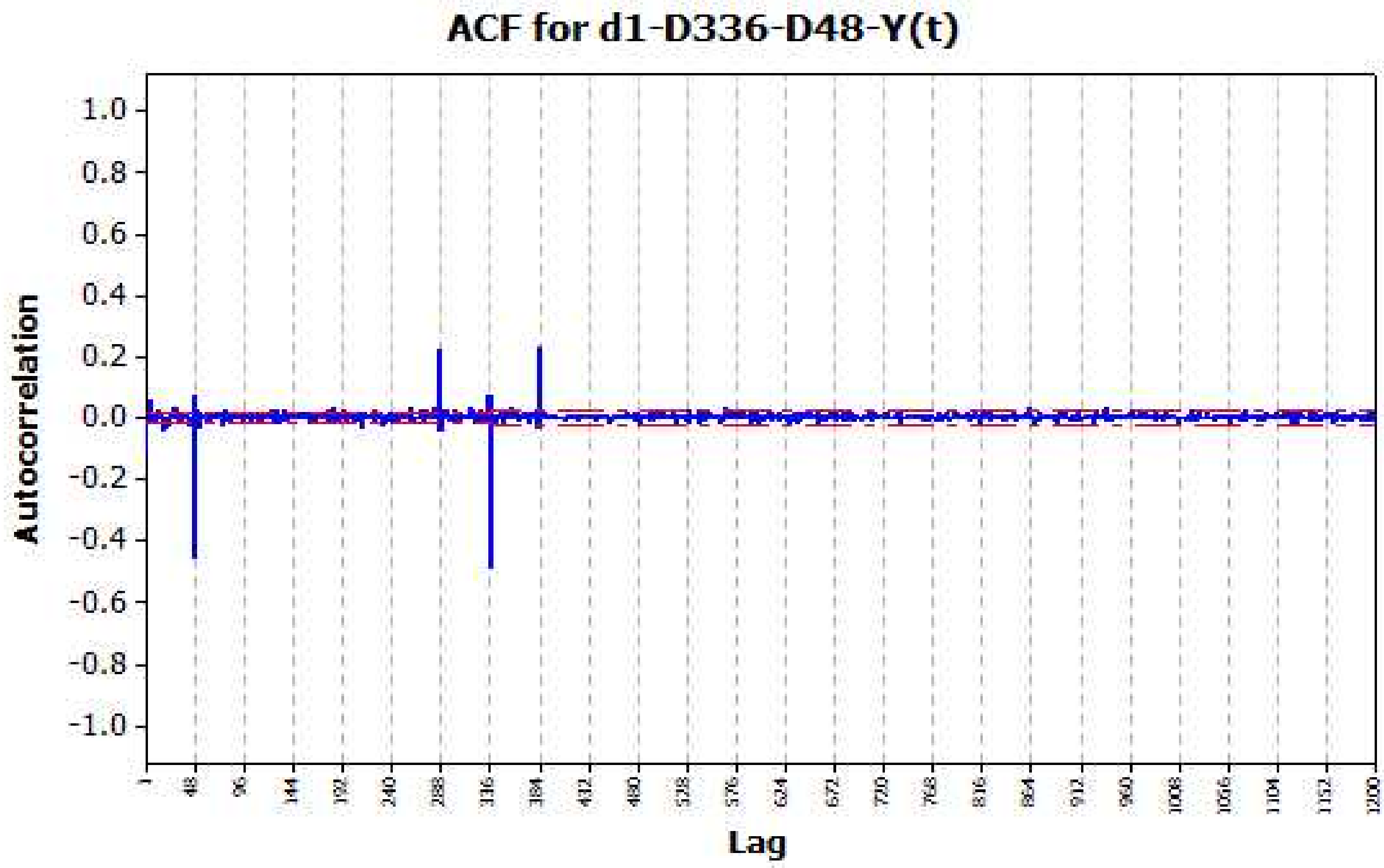
MINITAB Identification: half hourly load data



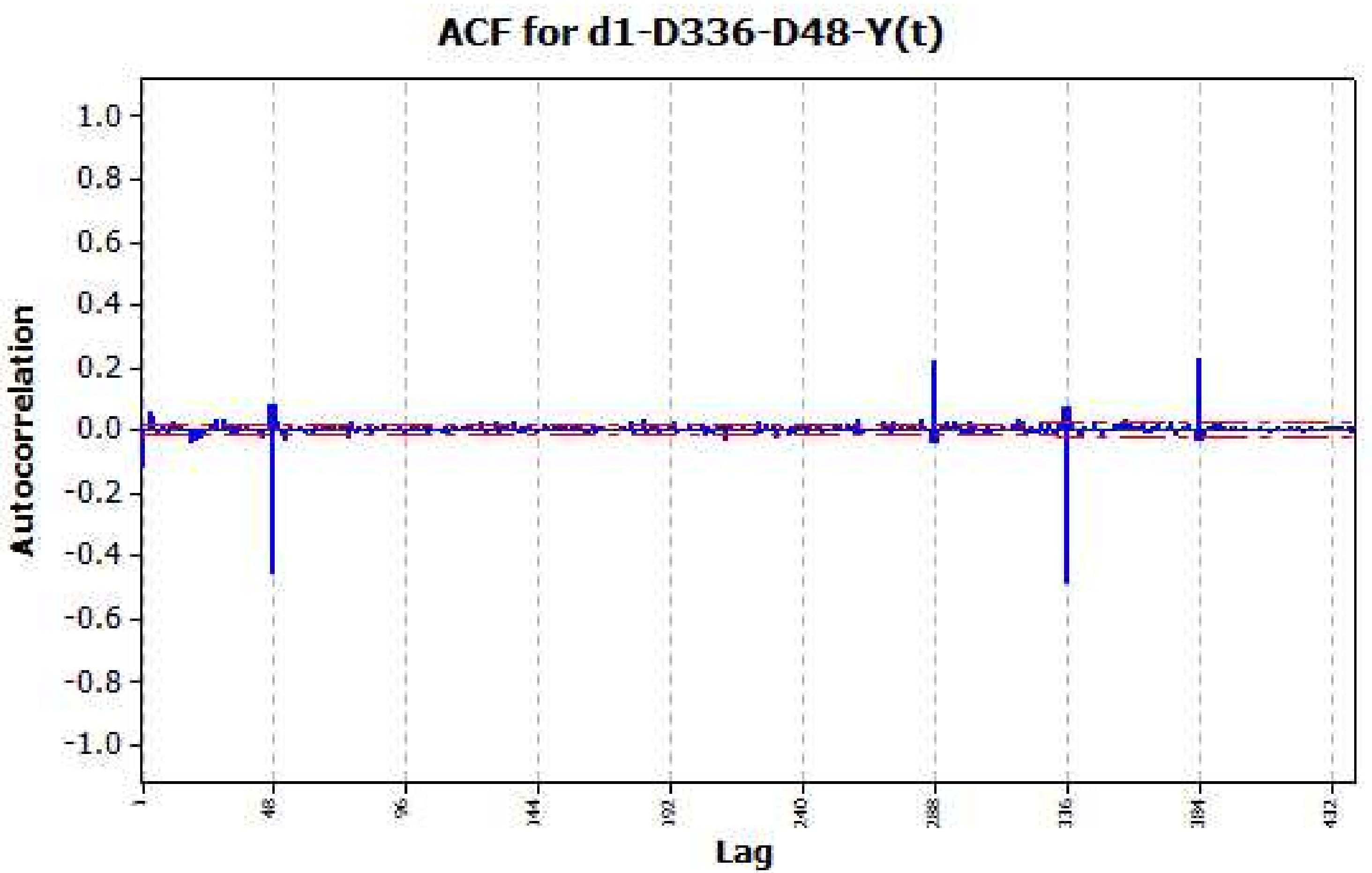
MINITAB Identification: half hourly load data



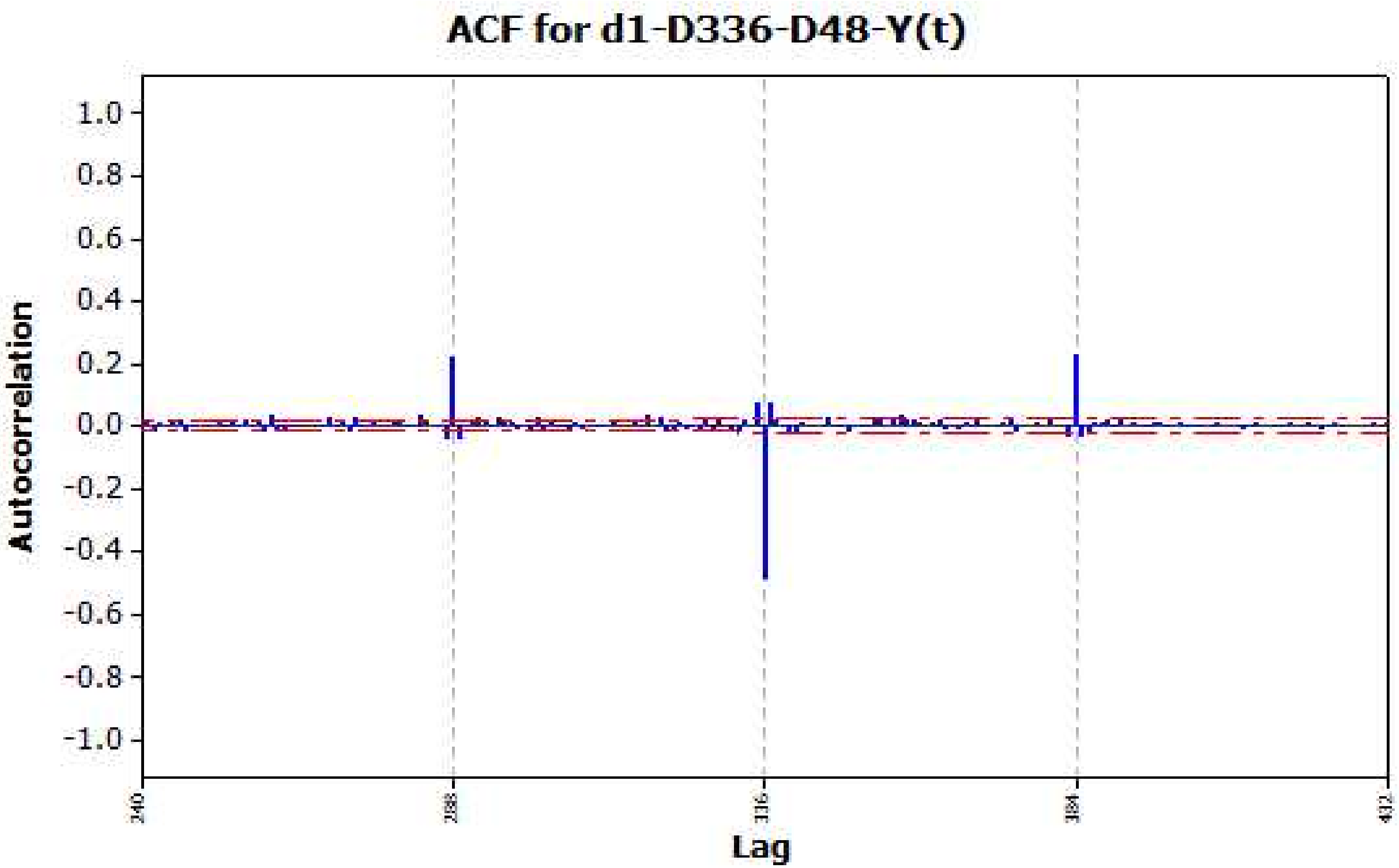
MINITAB Identification: half hourly load data



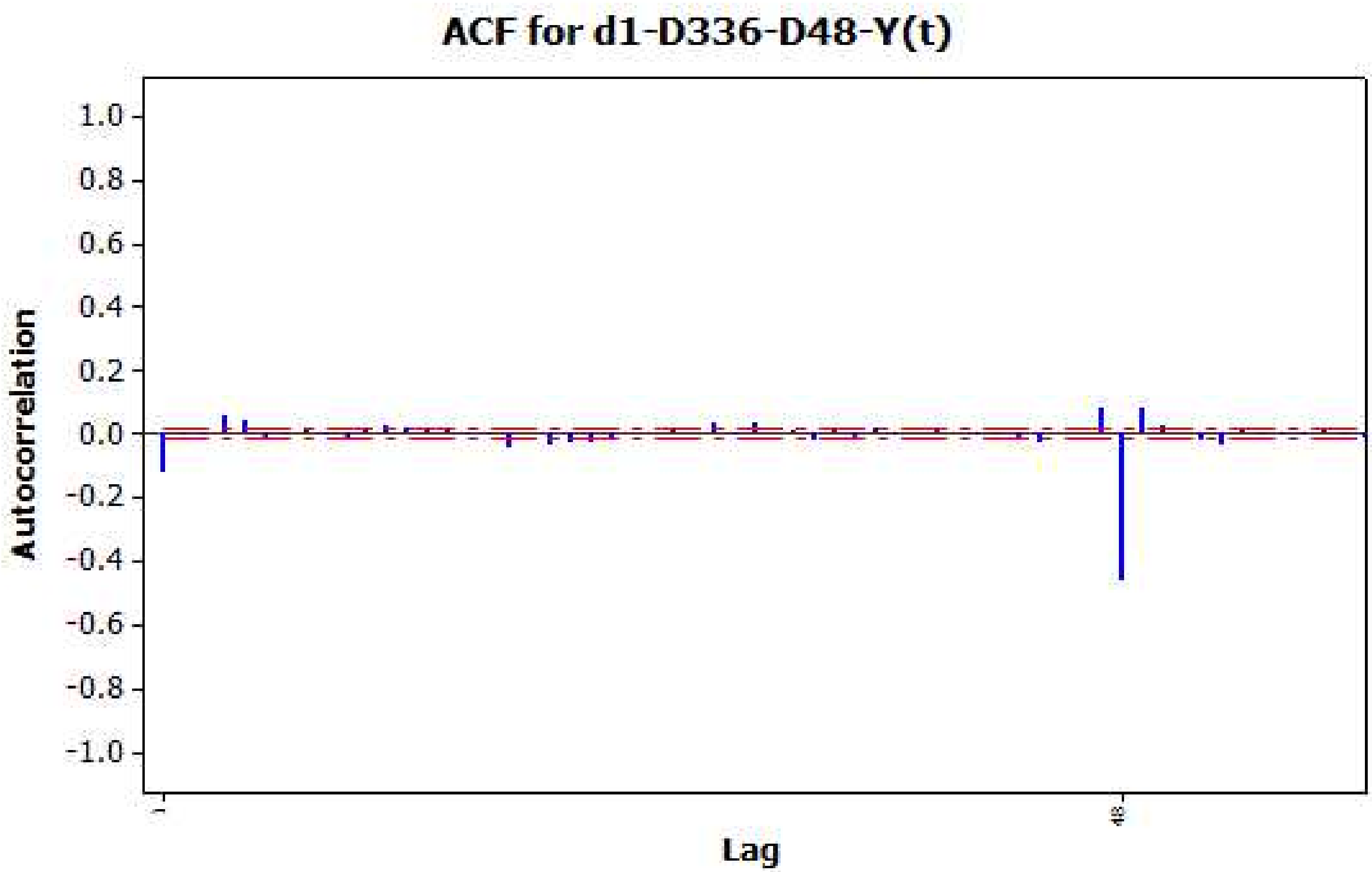
MINITAB Identification: half hourly load data



MINITAB Identification: half hourly load data



MINITAB Identification: half hourly load data



ARIMA, SARIMA, DSARIMA model

- ARIMA model

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t$$

- SARIMA model

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D \dot{Z}_t = \theta_q(B)\Theta_Q(B^s)a_t$$

- DSARIMA model

$$\phi_p(B)\Phi_{P_1}(B^{s_1})\Phi_{P_2}(B^{s_2})(1-B)^d(1-B^{s_1})^{D_1}(1-B^{s_2})^{D_2} Z_t = \theta_q(B)\Theta_{Q_1}(B^{s_1})\Theta_{Q_2}(B^{s_2})a_t$$

ACF of SARIMA model

- SARIMA$(0,0,1)(0,0,1)^{24}$ model

$$\rho_k = \begin{cases} 1, & k = 0, \\ \dfrac{-\theta_1}{(1+\theta_1^2)}, & k = 1, \\ \dfrac{\theta_1 \Theta_{24}}{(1+\theta_1^2)(1+\Theta_{24}^2)}, & k = 23 \text{ and } 25, \\ \dfrac{-\Theta_{24}}{(1+\Theta_{24}^2)}, & k = 24, \\ 0, & k \text{ others.} \end{cases}$$

ACF of DSARIMA model

O DSARIMA(0,0,1)(0,0,1)^{24}(0,0,1)^{168}

$$
{}^{\circ}{}_{k} = \begin{cases}
1, & k = 0, \\
\dfrac{-{}_{1}}{(1+{}_{1}^{2})}, & k = 1, \\
\dfrac{{}_{1}\ {}_{24}}{(1+{}_{1}^{2})(1+{}_{24}^{2})}, & k = 23 \text{ and } 25, \\
\dfrac{-{}_{24}}{(1+{}_{24}^{2})}, & k = 24, \\
\dfrac{-{}_{1}\ {}_{24}\ {}_{168}}{(1+{}_{1}^{2})(1+{}_{24}^{2})(1+{}_{168}^{2})}, & k = 143, 145, 191 \text{ and } 193, \\
\dfrac{{}_{24}\ {}_{168}}{(1+{}_{1}^{2})(1+{}_{24}^{2})(1+{}_{168}^{2})}, & k = 144, \\
\dfrac{{}_{1}\ {}_{168}}{(1+{}_{1}^{2})(1+{}_{168}^{2})}, & k = 167 \text{ and } 169, \\
\dfrac{-{}_{168}(1+{}_{24}^{2}+{}_{1}^{2}{}_{24}^{2})}{(1+{}_{1}^{2})(1+{}_{24}^{2})(1+{}_{168}^{2})}, & k = 168, \\
\dfrac{{}_{24}\ {}_{168}}{(1+{}_{24}^{2})(1+{}_{168}^{2})}, & k = 192, \\
0, & k \text{ others}.
\end{cases}
$$

R script: ACF of SARIMA

- $SARIMA(0,0,1)(0,0,1)^{24}$ model

```
# Program to calculate ACF and PACF theoretically
theta = c(-0.6,rep(0,22),-0.5,0.3)
acf.arma = ARMAacf(ar=0, ma=theta, 168)
pacf.arma = ARMAacf(ar=0, ma=theta, 168, pacf=T)
acf.arma = acf.arma[2:169]
c1 = acf.arma
c2 = pacf.arma
arma = cbind(c1, c2)
arma   # ACF and PACF theoretically
par(mfrow=c(1,2))
plot(acf.arma, type="h", xlab="lag", ylim=c(-1,1))
abline(h=0)
plot(pacf.arma, type="h", xlab="lag", ylim=c(-1,1))
abline(h=0)
```

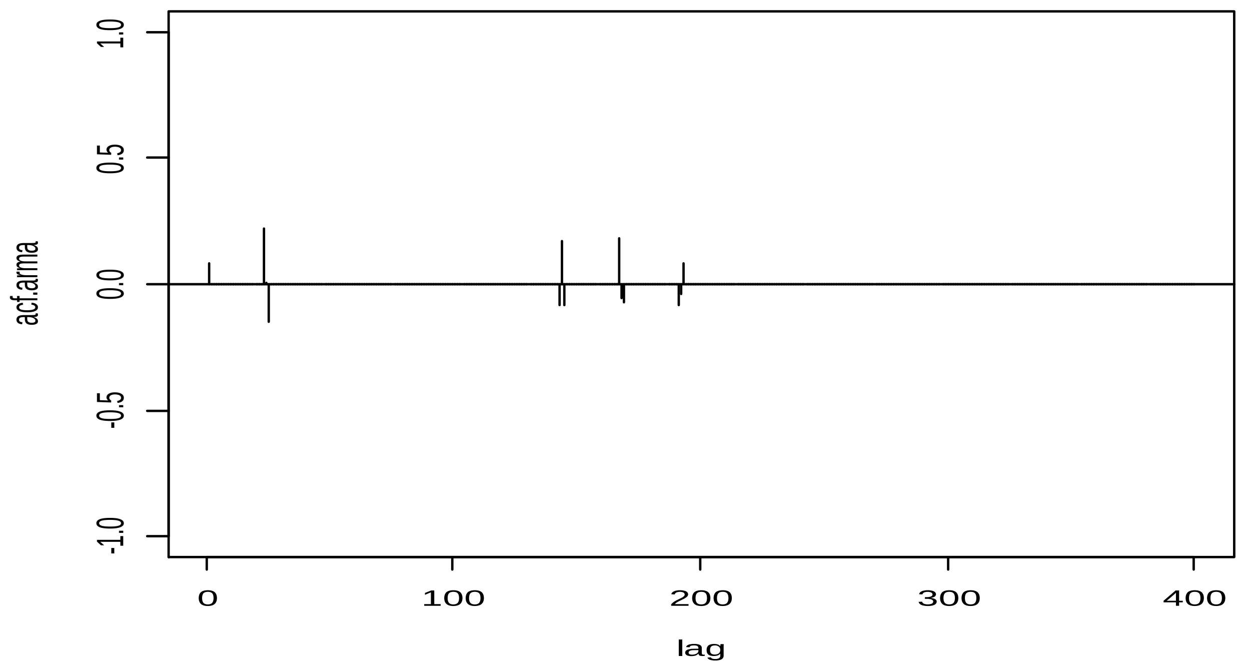
R script: ACF of SARIMA

- $SARIMA(0,0,1)(0,0,1)^{24}(0,0,1)^{168}$ model

```
# Program to calculate ACF and PACF theoretically
theta = c(-0.6,rep(0,22),-0.5,0.3,rep(0,143),-0.4,0.24,rep(0,12),0.2,-0.12)
acf.arma = ARMAacf(ar=0, ma=theta, 200)
pacf.arma = ARMAacf(ar=0, ma=theta, 200, pacf=T)
acf.arma = acf.arma[2:201]
c1 = acf.arma
c2 = pacf.arma
arma = cbind(c1, c2)
arma   # ACF and PACF theoretically
par(mfrow=c(1,2))
plot(acf.arma, type="h", xlab="lag", ylim=c(-1,1))
abline(h=0)
plot(pacf.arma, type="h", xlab="lag", ylim=c(-1,1))
abline(h=0)
```

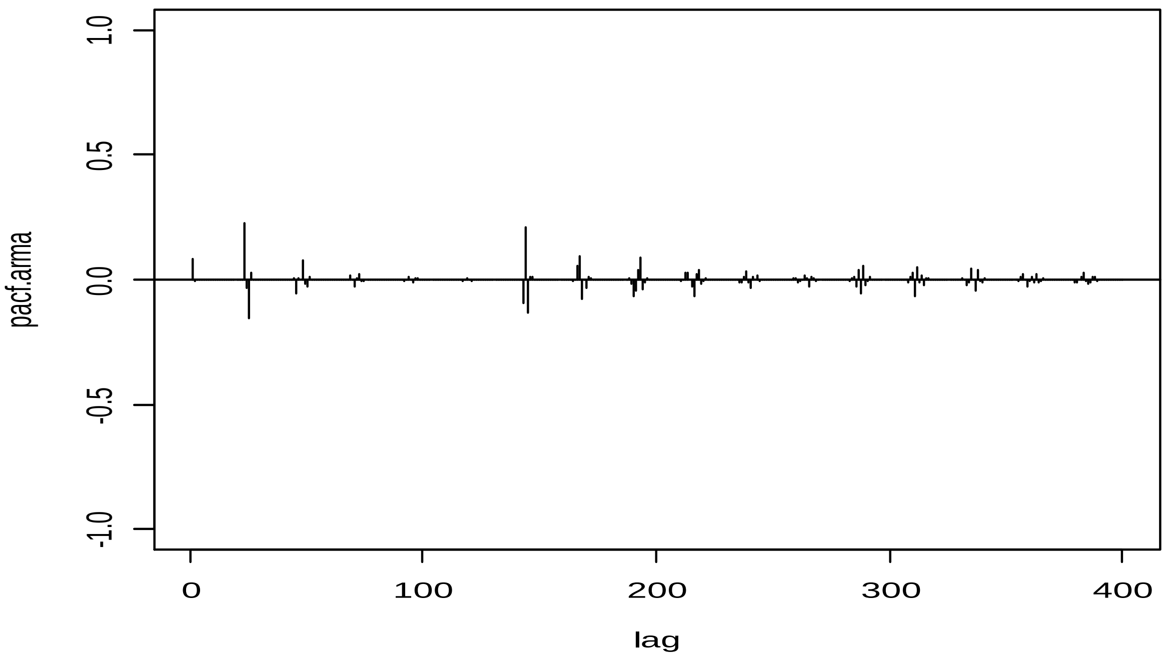
R: the result – ACF of DSARIMA

- SARIMA$(0,0,1)(0,0,1)^{24}(0,0,1)^{168}$ model with $_{1}$=0.8, $_{24}$=0.65, and $_{168}$=0.45



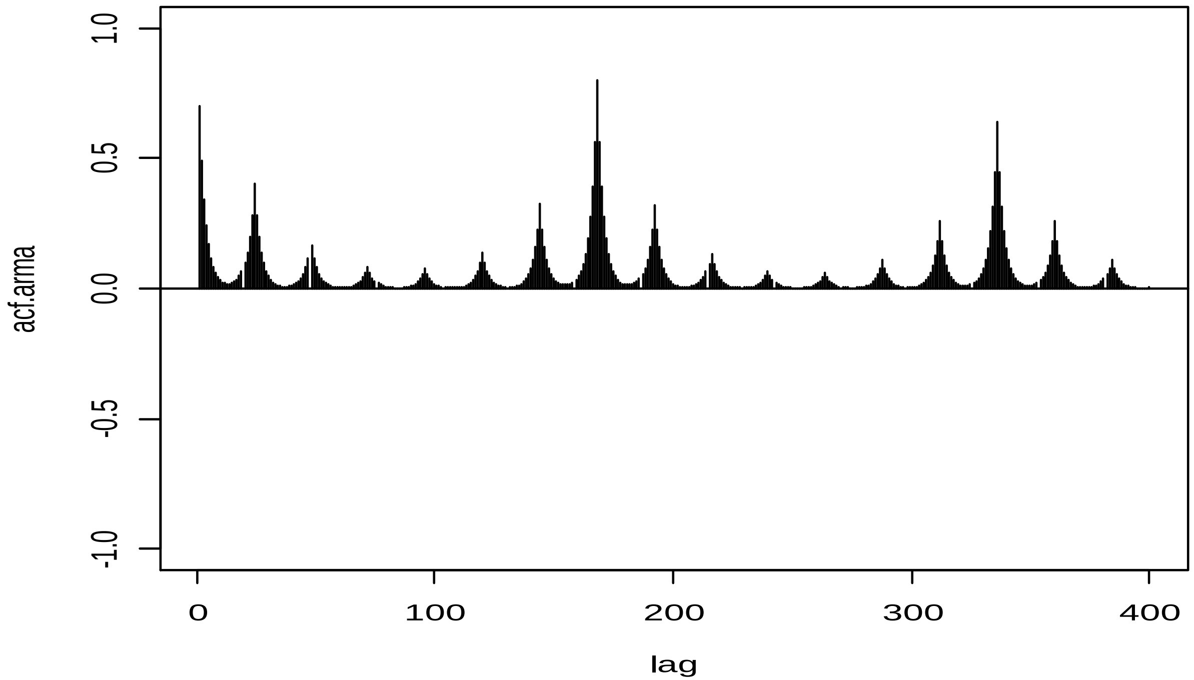
R: the result – PACF of DSARIMA

- SARIMA$(0,0,1)(0,0,1)^{24}(0,0,1)^{168}$ model with $_{1}=0.8$, $_{24}=0.65$, and $_{168}=0.45$



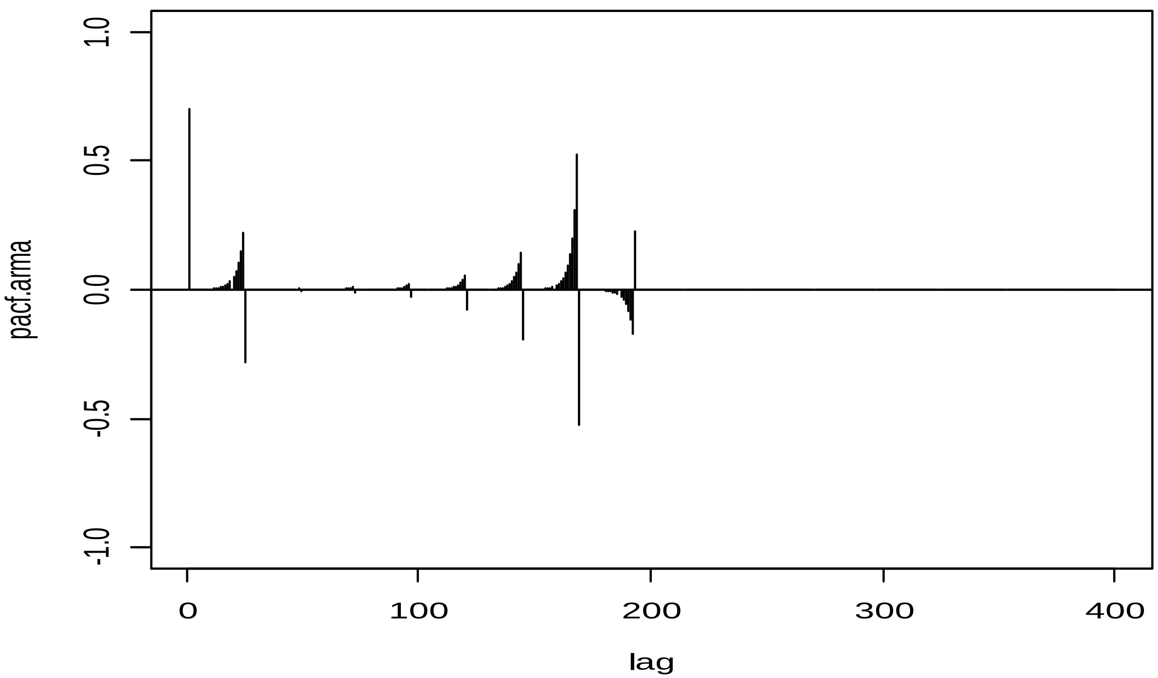
R: the result – ACF of DSARIMA

- SARIMA$(1,0,0)(1,0,0)^{24}(1,0,0)^{168}$ model with $_{1}=0.7$, $_{24}=0.4$, and $_{168}=0.8$



R: the result – PACF of DSARIMA

- SARIMA$(1,0,0)(1,0,0)^{24}(1,0,0)^{168}$ model with
  $_1 = 0.7$, $_{24} = 0.4$, and $_{168} = 0.8$



SAS Program: Estimation DSARIMA

```
data listrik;
  input y;
cards;
12123.6
…
11947.8
;
proc arima data=listrik out=b1;
    /*** IDENTIFICATION Step ***/
    identify var=y(1,48,336) nlag=12;
    run;
    /*** PARAMETER ESTIMATION & DIAGNOSTIC CHECK Step ***/
    estimate p=(5,8) q=(1)(24)(168) noconstant;
    run;
    /*** FORECASTING Step ***/
    forecast lead=336 out=b2 noprint;
    run;
```

SAS Program: Estimation DSARIMA

```
proc arima data=listrik out=b1;
    /*** IDENTIFICATION Step ***/
    identify var=y(1,48,336) nlag=12;
    run;
    /*** PARAMETER ESTIMATION & DIAGNOSTIC CHECK Step ***/
    estimate p=(5,8) q=(1)(24)(168) noconstant;
    run;
    /*** FORECASTING Step ***/
    forecast lead=336 out=b2 noprint;
    run;

proc export data= work.b2
          outfile= "D:\results1.xls"
          dbms=excel2000 replace;
    sheet="om_41";
    run;
```

Conclusion

☞ This paper shows that R, MINITAB and SAS must be used comprehensively for model building of DSARIMA from certain time series data.

- R: To calculate the theoretical ACF and PACF from DOUBLE Seasonal ARIMA models
- MINITAB: Descriptive evaluation & Identification step.
- SAS: Parameter Estimation, Diagnostic check, and Forecasting steps.

2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

Two Levels Regression Modeling of Trading Day and Holiday Effects for Forecasting Retail Data

Suhartono
(B.Sc.-ITS; M.Sc.-UMIST,UK; Dr.-UGM; Postdoctoral-UTM)

Department of Statistics,
Institut Teknologi Sepuluh Nopember, Indonesia

Email: *suhartono@statistika.its.ac.id*, *gmsuhartono@gmail.com*

Department of Mathematics, Universitas Andalas, Padang
17-18 July 2017

Outline

- **Introduction**: *General time series "pattern"*
- **The aims of this paper**: *Develop two levels calendar variation model*
- **Data**: *Monthly men's jeans and women's trousers sales in a retail company*
- **Modeling method**: *Based on time series regression*
- **Results, analysis and evaluation**: *forecast accuracy*
- **Conclusion and future works**

Introduction

- ✓ General time series "**PATTERN**"
  - ✍ Stationer
  - ✍ Trend: ***linear*** & ***nonlinear***
  - ✍ Seasonal: ***additive*** & ***multiplicative***
  - ✍ Cyclic
  - ✍ **Calendar Variation**

Introduction

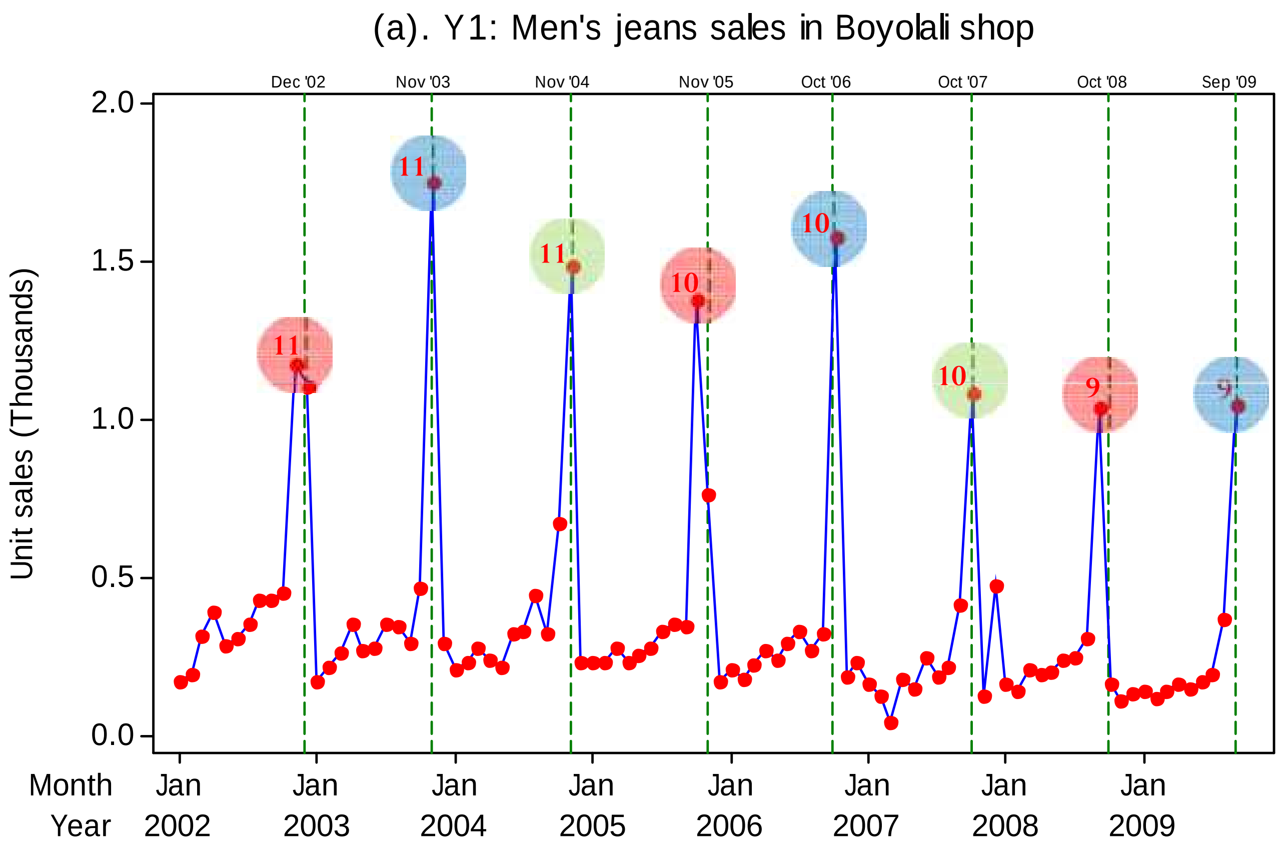
- Two kinds of calendar variation effects:
  1. Trading day effects

     The levels of economics or business activities may change depending on the day of the week. The composition of days of the week varies from month to month and year to year.
  2. Holiday (traditional festivals) effects

     Some traditional festivals or holidays, such as Eid ul-Fitr, Easter, Chinese New Year, and Jewish Passover are set according to lunar calendars and the dates of such holidays may vary between two adjacent months in the Gregorian calendar from year to year.

Introduction

cont'



Introduction

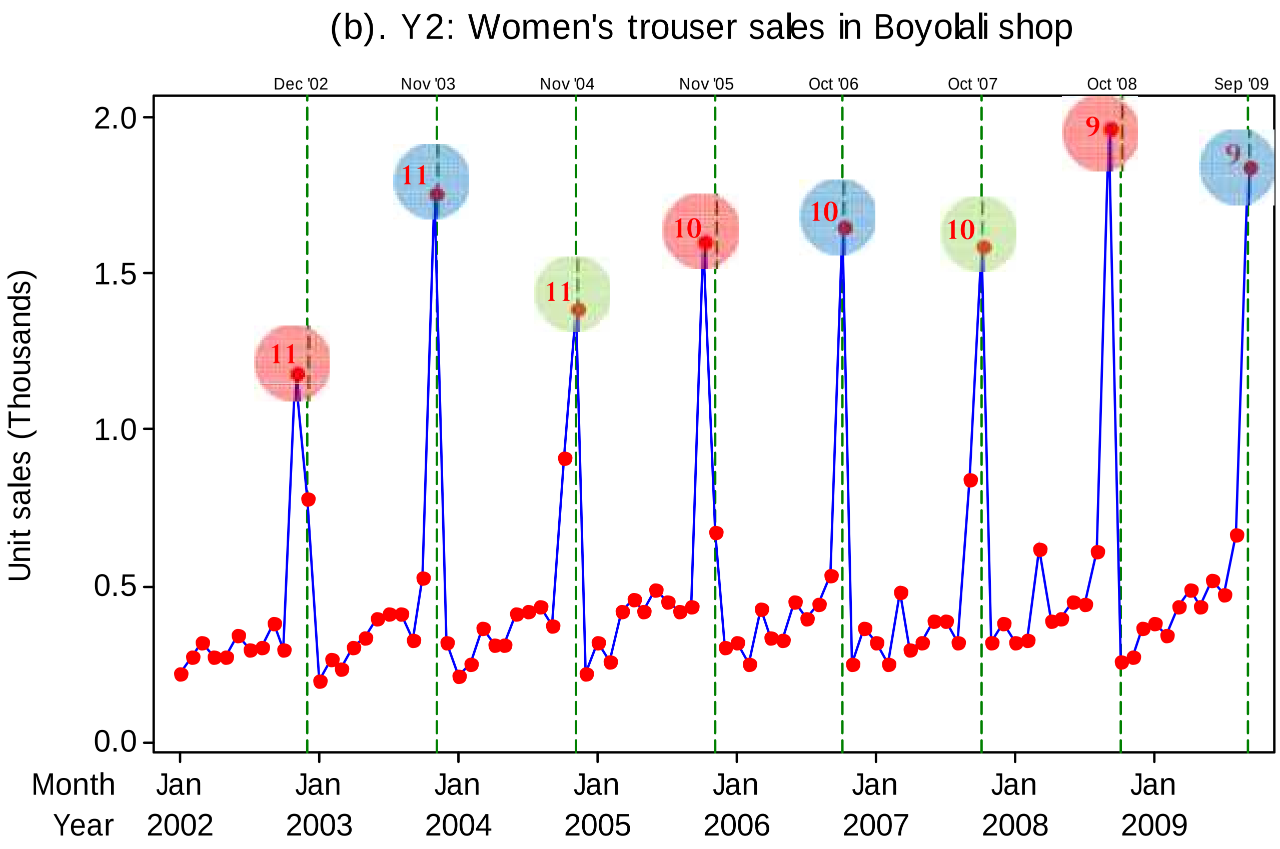
cont'

Eid holidays for the period 2002 to 2011

Year	Date	Explanation
2002	06-07 December	There are 5 days before Eid in December
2003	25-26 November	There are 24 days before Eid in November
2004	14-15 November	There are 13 days before Eid in November
2005	03-04 November	There are 2 days before Eid in November
2006	23-24 October	There are 22 days before Eid in October
2007	12-13 October	There are 11 days before Eid in October
2008	01-02 October	There is 0 day before Eid in October
2009	21-22 September	There are 20 days before Eid in September
2010	10-11 September	There are 9 days before Eid in September
2011	30-31 August	There are 29 days before Eid in August

Introduction

cont'



Introduction

cont'

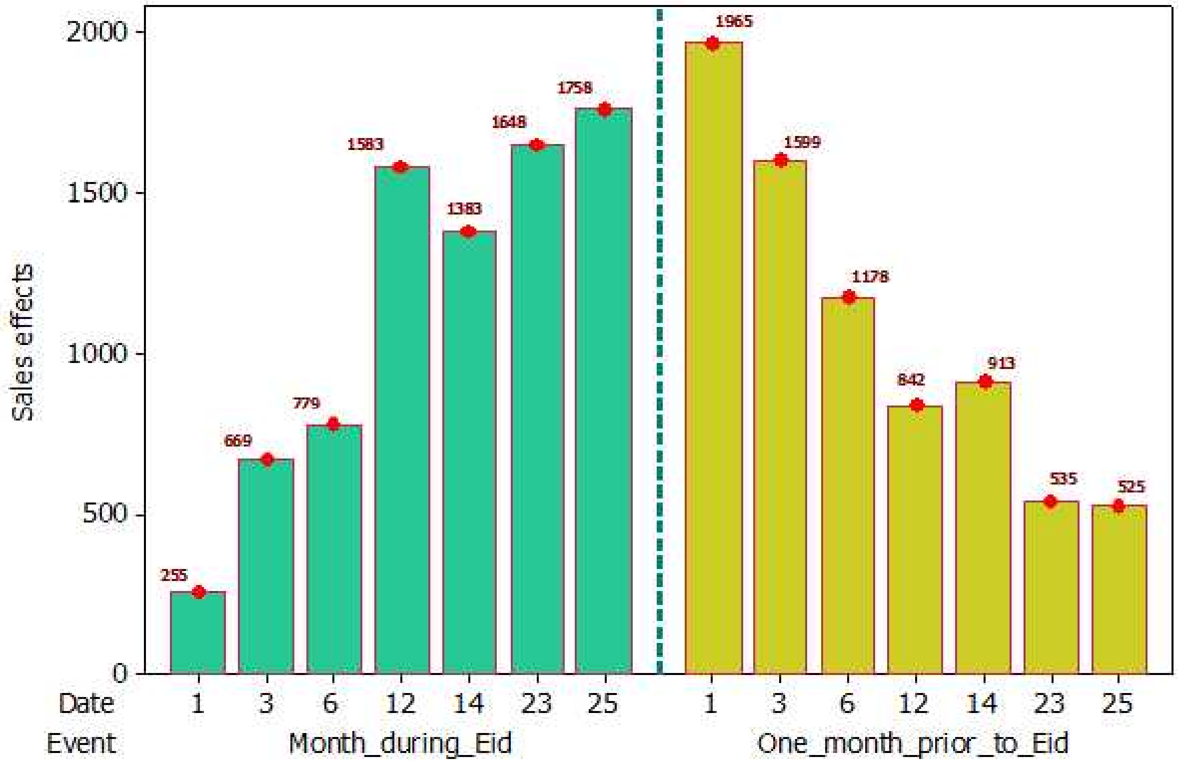


Fig. 2. Bar chart of Eid effects on the women's trouser sales in the month during and one month prior to the Eid celebration in Boyolali shop.

The aims

- To develop ***two levels calendar variation model*** based on **time series regression method** for forecasting sales data with the ***Eid ul-Fitr*** effects.

- To compare the **forecast accuracy** with other forecasting methods, i.e.
  - ARIMA model
  - Feed-forward Neural Networks (FFNN)

Modeling method

- Model for linear trend:

$$y_t = \beta_0 + \beta_1 t + w_t \qquad \ldots (1)$$

- Regression with dummy variable for seasonal pattern:

$$y_t = \beta_0 + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + w_t \qquad \ldots (2)$$

- Regression for calendar effects:

$$y_t = \beta_0 + \beta_1 V_{1,t} + \beta_2 V_{2,t} + \cdots + \beta_p V_{p,t} + w_t \qquad \ldots (3)$$

The Proposed Model

- Model at the first level :

$$Y_t = \delta_1 t + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} +$$
$$\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \,.$$

- Model at the second level :

1. Linear model

$\hat{\alpha}_j = \upsilon_0 + \upsilon_1 j$

$\hat{\gamma}_j = \omega_0 + \omega_1 j$

2. Exponential model

$\hat{\alpha}_j = \upsilon_0 \, e^{\upsilon_1 j}$

$\hat{\gamma}_j = \ln(\omega_0 + \omega_1 j)$

Background of Two Levels

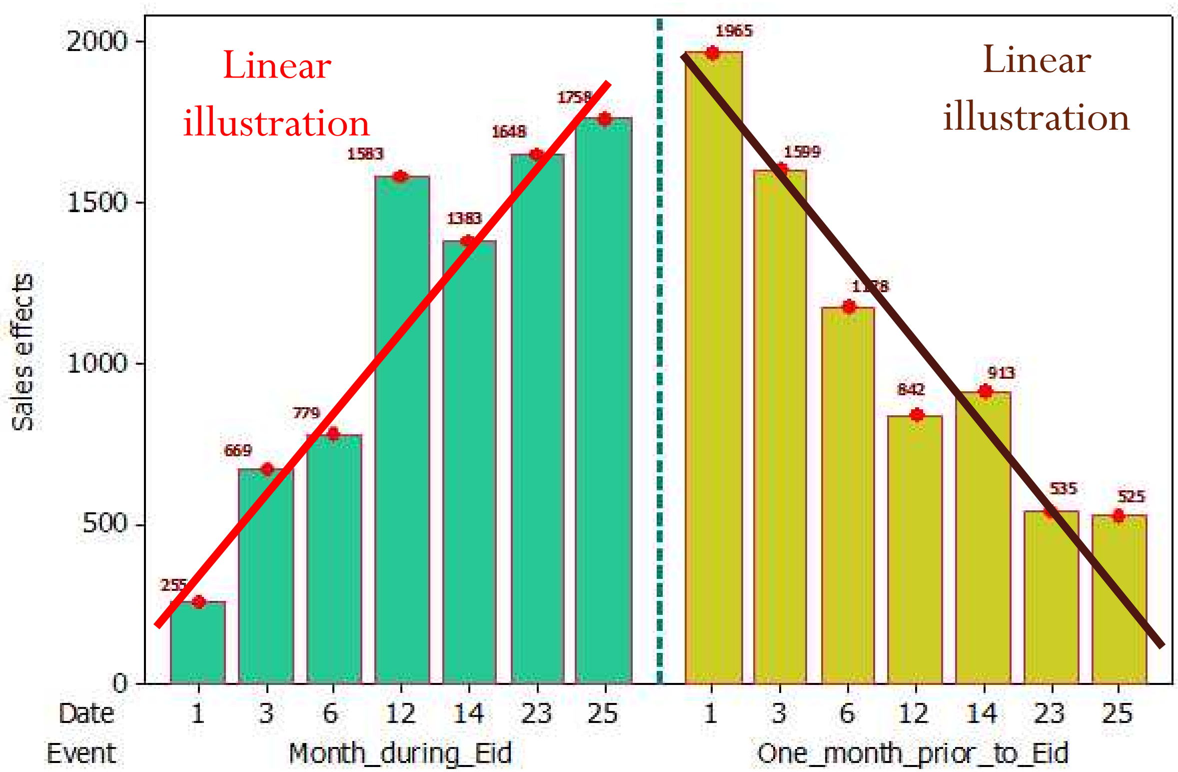


Fig. 2. Bar chart of Eid effects on the women's trouser sales in the month during and one month prior to the Eid celebration in Boyolali shop.

Dummy at Two Levels

Year	Date	Explanation
2002	06-07 December	$D_{5,t}$ = December, and $D_{5,t-1}$ = November
2003	25-26 November	$D_{24,t}$ = November, and $D_{24,t-1}$ = October
2004	14-15 November	$D_{13,t}$ = November, and $D_{13,t-1}$ = October
2005	03-04 November	$D_{2,t}$ = November, and $D_{2,t-1}$ = October
2006	23-24 October	$D_{22,t}$ = October, and $D_{22,t-1}$ = September
2007	12-13 October	$D_{11,t}$ = October, and $D_{11,t-1}$ = September
2008	01-02 October	$D_{0,t}$ = October, and $D_{0,t-1}$ = September
2009	21-22 September	$D_{20,t}$ = September, and $D_{20,t-1}$ = August
2010	10-11 September	$D_{9,t}$ = September, and $D_{9,t-1}$ = August
2011	30-31 August	$D_{29,t}$ = September, and $D_{20,t-1}$ = August

The proposed procedure

Step 1: Determination of dummy variable for calendar variation period.

Step 2: Determination of deterministic trend and seasonal model.

Step 3: Simultaneous estimation of calendar effects and other patterns.

Step 4: Diagnostic checks on error model. If error is not white noise, add significant lags (autoregressive order) based on ACF and PACF plots of error model.

Step 5: Re-estimate calendar effect, other pattern (trend, seasonal), and appropriate lags (autoregressive order) simultaneously for the first level model.

Step 6: Estimate the second level model to predict the effects of calendar variation in every possibility number of days before Eid ul-Fitr celebration.

Step 1

- Based on the time series plot, two dummy variables are used for evaluating calendar variation effect, i.e.
  - The months prior to Eid ul Fitr,

    $D_{j,t-1}$ = dummy variable for ONE month prior to Eid ul-Fitr celebration.
  - During the month of Eid ul-Fitr celebration,

    $D_{j,t}$ = dummy variable for during the month of Eid ul-Fitr celebration.
- o j = number of days before Eid ul-Fitr celebration

Step 2 -3

- Model for linear trend:

$$y_t = \beta_0 + \beta_1 t + w_t$$

- Regression with dummy variable for seasonal pattern:

$$y_t = \beta_0 + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + w_t$$

- Regression for calendar effects and other patterns:

$$Y_t = \delta_1 t + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + N_t \,.$$

Step 4-6

- Model at the first level :

$$Y_t = \delta_1 t + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} +$$

$$\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \,.$$

- Model at the second level :

1. *Linear model*

$\Rightarrow$ $\hat{\alpha}_j = \upsilon_0 + \upsilon_1 j$

$\Rightarrow$ $\hat{\gamma}_j = \omega_0 + \omega_1 j$

2. *Exponential model*

$\Rightarrow$ $\hat{\alpha}_j = \upsilon_0 \, e^{\upsilon_1 j}$

$\Rightarrow$ $\hat{\gamma}_j = \ln(\omega_0 + \omega_1 j)$

Results: monthly sales of men's jeans

a. The first level model

$$Y_{1,t} = 0.197M_{1,t} + 0.201M_{2,t} + 0.238M_{3,t} + 0.281M_{4,t} + 0.240M_{5,t} + 0.291M_{6,t} + 0.318M_{7,t} + 0.346M_{8,t} + 0.359M_{9,t} + 0.462M_{10,t} + 0.162M_{11,t} + 0.283M_{12,t} + 0.605D_{2,t} + 0.825D_{5,t} + 0.627D_{11,t} + 1.32D_{13,t} + 1.12D_{22,t} + 1.60D_{24,t} + 0.917D_{2,t-1} + 1.02D_{5,t-1} + 0.215D_{13,t-1} + \varepsilon_t . \quad (16)$$

b. The second level model

b.1. Linear form

$$\hat{\alpha}_j = 0.408 + 0.0473\,j , \quad (17a)$$

$$\hat{\gamma}_j = 0.946 - 0.0458\,j . \quad (17b)$$

b.2. Exponential form

$$\hat{\alpha}_j = \ln(1.614 + 0.098\,j) , \quad (18a)$$

$$\hat{\gamma}_j = 2.103\,e^{-0.038\,j} - 1 . \quad (18b)$$

Results: monthly sales of women's trouser

a. The first level model

$$Y_{2,t} = 0.00126t + 0.225M_{1,t} + 0.217M_{2,t} + 0.331M_{3,t} + 0.284M_{4,t} + 0.285M_{5,t} + 0.366M_{6,t} + 0.346M_{7,t} + 0.338M_{8,t} + 0.345M_{9,t} + 0.280M_{10,t} + 0.202M_{11,t} + 0.257M_{12,t} + 0.408D_{2,t} + 0.507D_{5,t} + 1.21D_{11,t} + 1.14D_{13,t} + 1.29D_{22,t} + 1.53D_{24,t} + 1.26D_{2,t-1} + 0.963D_{5,t-1} + 0.410D_{11,t-1} + 0.590D_{13,t-1} + 0.118D_{22,t-1} + 0.217D_{24,t-1} + \varepsilon_t .$$

b. The second level model

b.1. Linear form

$$\hat{\alpha}_j = 0.552 + 0.0361j ,$$

$$\hat{\gamma}_j = 1.20 - 0.0469j .$$

b.2. Exponential form

$$\hat{\alpha}_j = \ln(1.203 + 0.138j) ,$$

$$\hat{\gamma}_j = 1.519e^{-0.094j} .$$

Results: monthly sales of women's trouser

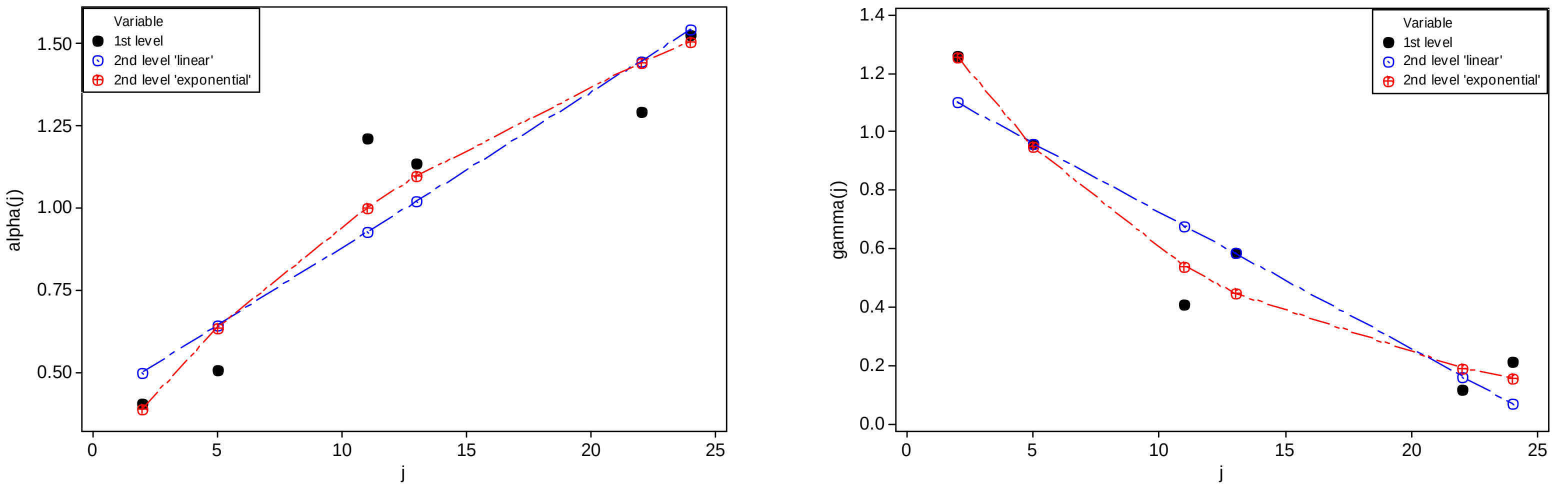
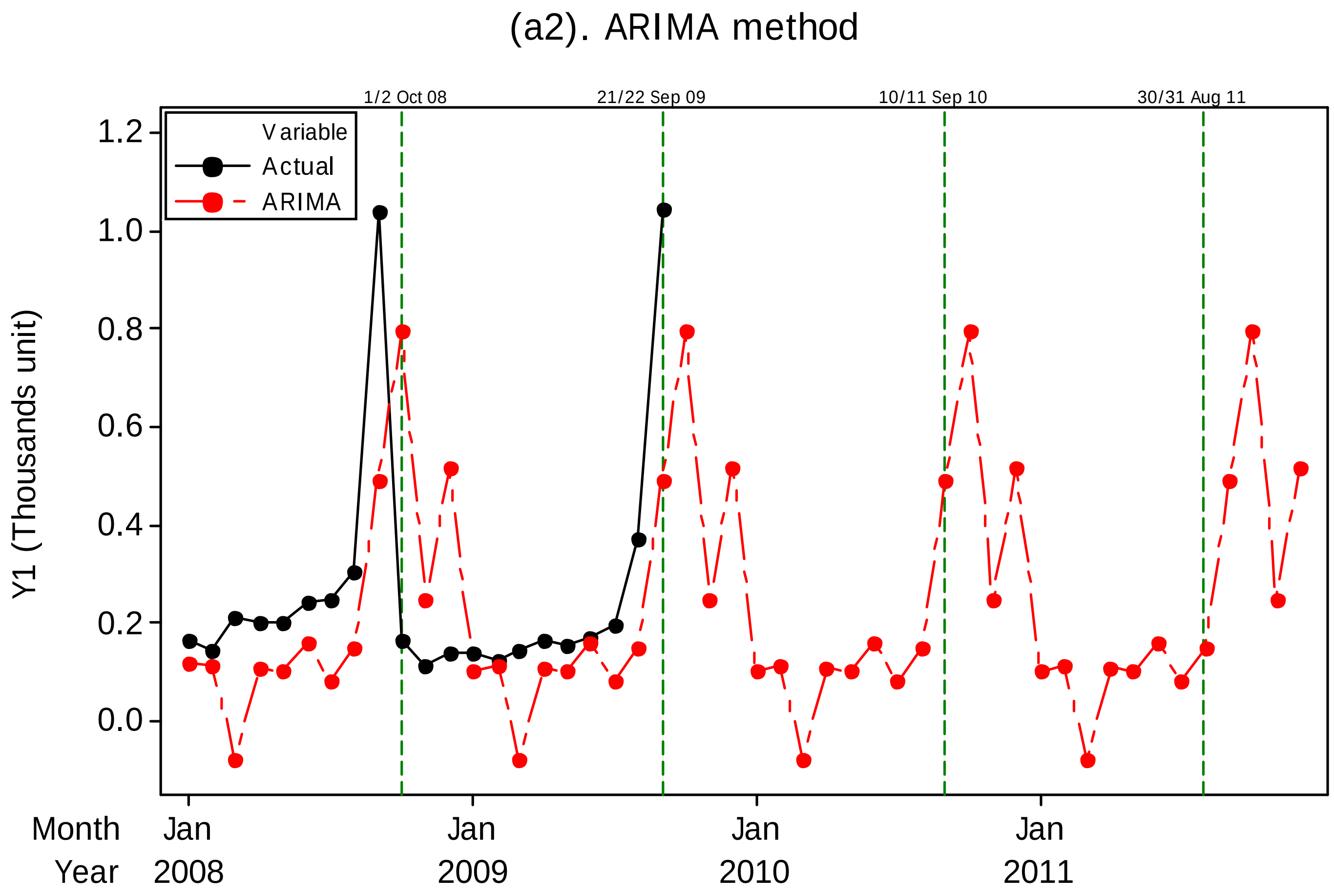


Fig. 3. The fitted line of the second level model regression in Eq. (13a)-(14b) for monthly sales of women's trouser data

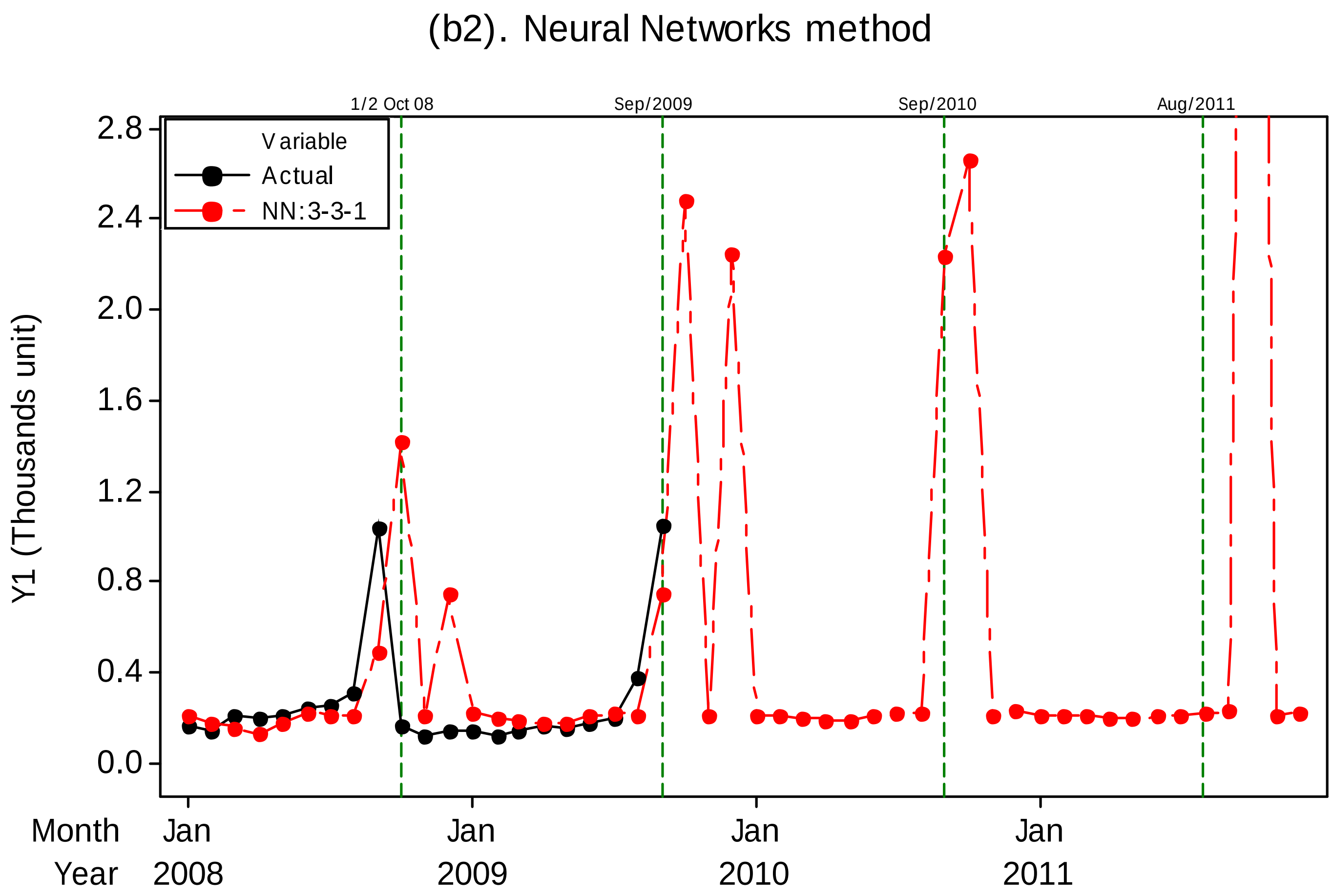
Results

Method	$Y_{1,t}$ = men's jeans		$Y_{2,t}$ = women trouser	
	in-sample	out-sample	in-sample	out-sample
ARIMA	0.1408	0.2634	0.1685	0.4235
FFNN: no skip layer				
3-1-1	0.1188	0.3847	0.0845	0.3290
3-2-1	0.0809	4.3466	0.0741	0.3844
3-3-1	0.0786	0.3375	0.0657	0.2789
…	…	…	…	…
3-9-1	0.0709	11.7676	0.0551	1.7997
3-10-1	0.0894	5.6064	0.0598	10.6219
FFNN: with skip layer				
3-1-1	0.1148	0.4159	0.0889	0.3273
3-2-1	0.0809	0.5659	0.0710	0.3383
3-3-1	0.0708	0.6290	0.0663	0.2855
…	…	…	…	…
3-9-1	0.1245	2.6E+01	0.0616	2.6E+01
3-10-1	0.1087	1.9E+07	0.0561	8.2E+01
Two levels regression	**0.0686**		**0.0510**	
2^{nd}: linear model		0.2434		0.1508
2^{nd}: exponential model		**0.2424**		**0.0929**

Graphical Results

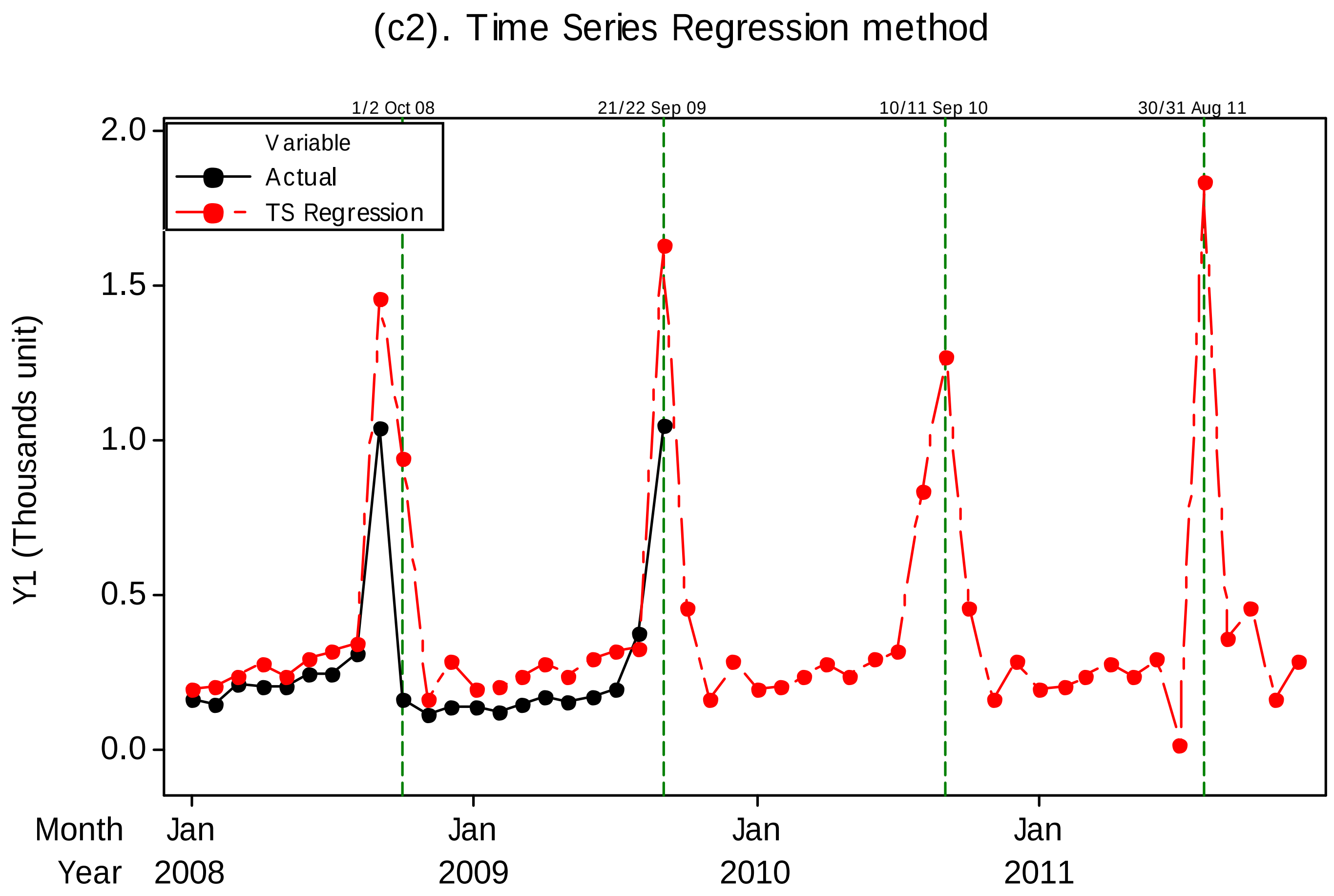


Graphical Results



*Inspiring **Creative** & **Innovative** Minds* – ***Dept. of Mathematics, UNAND***

Graphical Results



Conclusion

- The proposed two levels calendar variation model based on time series regression yield better prediction for out-sample data, compared to those of ARIMA model and neural networks.
- The application of ARIMA model usually yield spurious results, particularly about seasonal pattern and the presence of outliers.
- Whereas, neural networks perform well only for in-sample data.

References

1. Alagidede, P. (2008a). Day of the week seasonality in African stock markets. *Applied Financial Economics Letters*, 4, 115-120.
2. Alagidede, P. (2008b). Month-of-the-year and pre-holiday seasonality in African stock markets. *Stirling Economics Discussion Paper* 2008-23, Department of Economics, University of Stirling.
3. Al-Khazali, O.M., Koumanakos, E.P., & Pyun, C.S. (2008). Calendar anomaly in the Greek stock market: Stochastic dominance analysis. *International Review of Financial Analysis*, 17, 461–474.
4. Ariel, R.A. (1990). High Stock Returns before Holidays: Existence and Evidence on Possible Causes. *The Journal of Finance*, 45(5), 1611-1626.
5. Balaban, E. (1995). Day of the week effects: new evidence from an emerging stock market. *Applied Economics Letters*, 2, 139-143.
6. Bell, W.R., & Hillmer, S.C. (1983). Modeling Time Series with Calendar Variation. *Journal of the American Statistical Association*, 78, 526-534.
7. Brockman, P., & Michayluk, D. (1998). The Persistent Holiday Effect: Additional Evidence. *Applied Economics Letters*, 5, 205-209.
8. Brooks, C., & Persand, G. (2001). Seasonality in Southeast Asian stock markets: some new evidence on day-of-the week effects. *Applied Economics Letters*, 8, 155-158.

References

9. Cleveland, W.S. & Devlin, S.J. (1980). Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods. *Journal of the American Statistical Association*, 75(371), 487-496.
10. Cleveland, W.S. & Devlin, S.J. (1982). Calendar Effects in Monthly Time Series: Modeling and Adjustment. *Journal of the American Statistical Association*, 77(379), 520-528.
11. Evans, K.P., & Speight, A.E.H. (2010). Intraday periodicity, calendar and announcement effects in Euro exchange rate volatility. *Research in International Business and Finance,* 24, 82-101.
12. Faraway, J., & Chatfield, C. (1998). Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics*, 47, 231-250.
13. Gao, L., & Kling, G. (2005). Calendar Effects in Chinese Stock Market. *Annals of Economics and Finance*, 6, 75-88.
14. Hillmer, S.C. (1982). Forecasting Time Series with Trading Day Variation. *Journal of Forecasting*, 1, 385-395.
15. Holden, K., Thompson, J., & Ruangrit, Y. (2005). The Asian crisis and calendar effects on stock returns in Thailand. *European Journal of Operational Research*, 163, 242–252.
16. Liu, L.M. (1980). Analysis of Time Series with Calendar Effects. *Management Science*, 2, 106-112.

2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

Two Levels ARIMAX and Regression Models for Forecasting Time Series Data with Calendar Variation Effects

Suhartono

(B.Sc.-ITS; M.Sc.-UMIST,UK; Dr.-UGM; Postdoctoral-UTM)

Department of Statistics,
Institut Teknologi Sepuluh Nopember, Indonesia

Email*: suhartono@statistika.its.ac.id*, *gmsuhartono@gmail.com*

Department of Mathematics, Universitas Andalas, Padang

17-18 July 2017

Outline

- **Introduction**: *General time series "pattern"*
- **The aims of this paper**: ***Develop two levels calendar variation model***
- **Data**: ***Monthly men's jeans*** *and* ***women's trousers*** *sales in a retail company*
- **Modeling method**: *Based on* ***ARIMAX*** *and* ***Regression***
- **Results, analysis and evaluation**: *forecast accuracy*
- **Conclusion and future works**

Introduction

- ✓ General time series "**PATTERN**"
  - ✍ Stationer
  - ✍ Trend: ***linear*** & ***nonlinear***
  - ✍ Seasonal: ***additive*** & ***multiplicative***
  - ✍ Cyclic
  - ✍ **Calendar Variation**

Introduction cont'

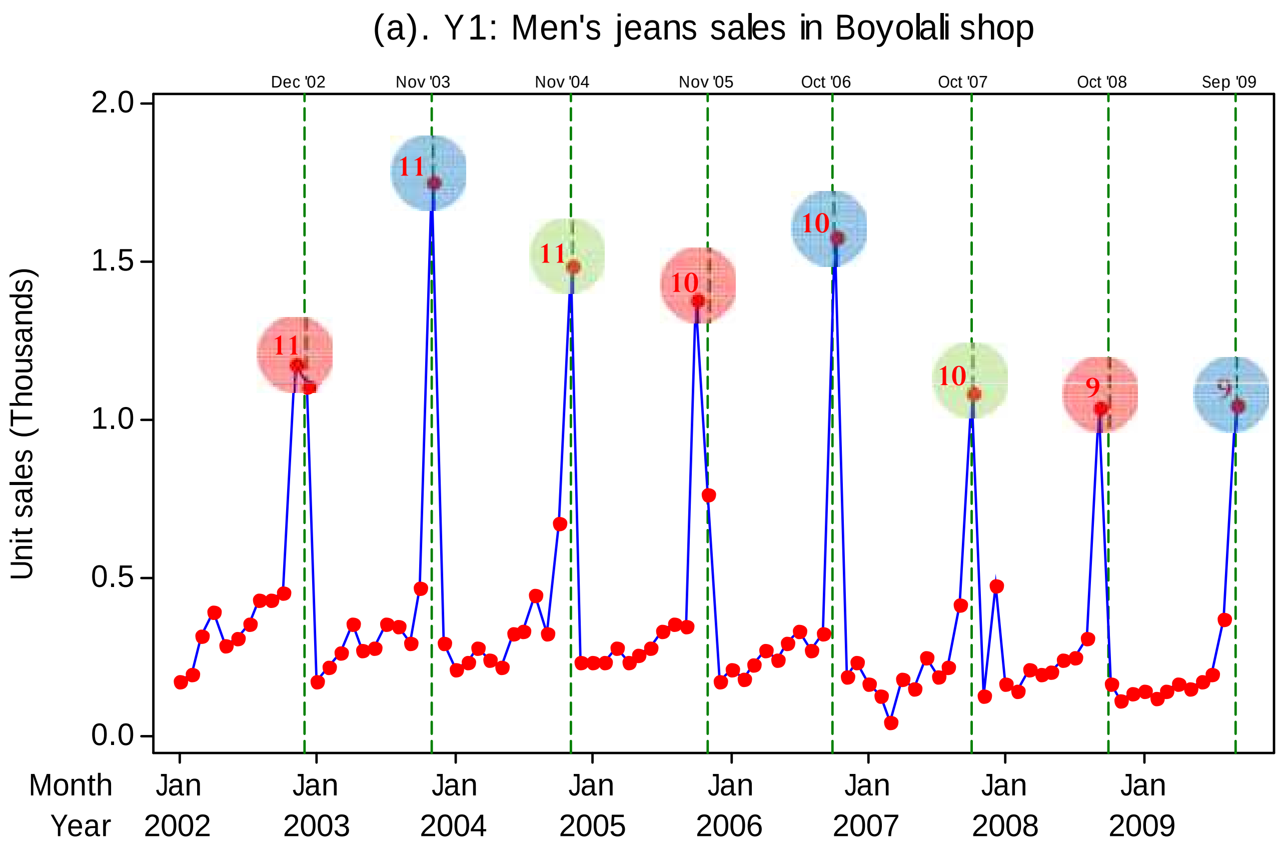
- Two kinds of calendar variation effects:
  1. **Trading day effects**

     The levels of economics or business activities may change depending on the day of the week. The composition of days of the week varies from month to month and year to year.
  2. **Holiday (traditional festivals) effects**

     Some traditional festivals or holidays, such as Eid ul-Fitr, Easter, Chinese New Year, and Jewish Passover are set according to lunar calendars and the dates of such holidays may vary between two adjacent months in the Gregorian calendar from year to year.

Introduction

cont'



Introduction

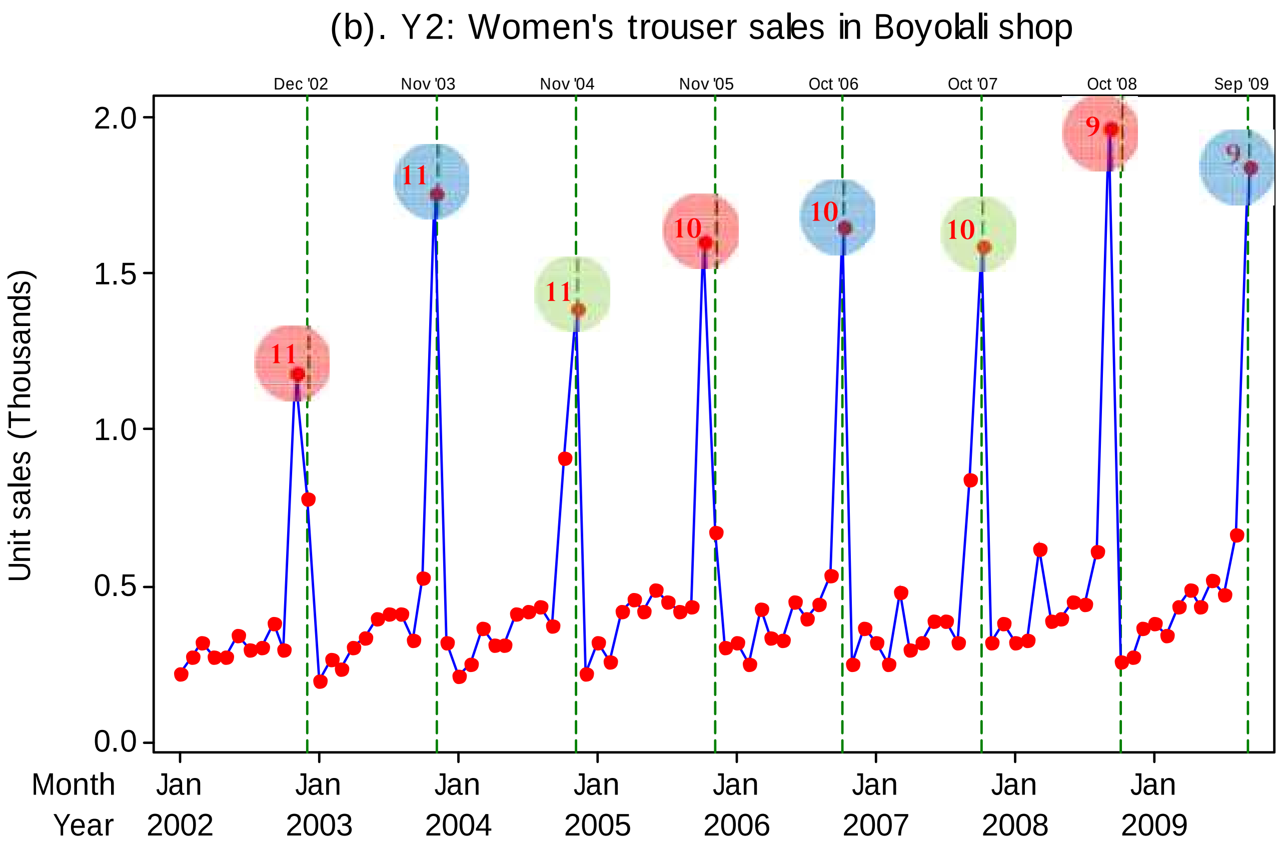
cont'

Eid holidays for the period 2002 to 2011

Year	Date	Explanation
2002	06-07 December	There are 5 days before Eid in December
2003	25-26 November	There are 24 days before Eid in November
2004	14-15 November	There are 13 days before Eid in November
2005	03-04 November	There are 2 days before Eid in November
2006	23-24 October	There are 22 days before Eid in October
2007	12-13 October	There are 11 days before Eid in October
2008	01-02 October	There is 0 day before Eid in October
2009	21-22 September	There are 20 days before Eid in September
2010	10-11 September	There are 9 days before Eid in September
2011	30-31 August	There are 29 days before Eid in August

Introduction

cont'



Introduction

cont'

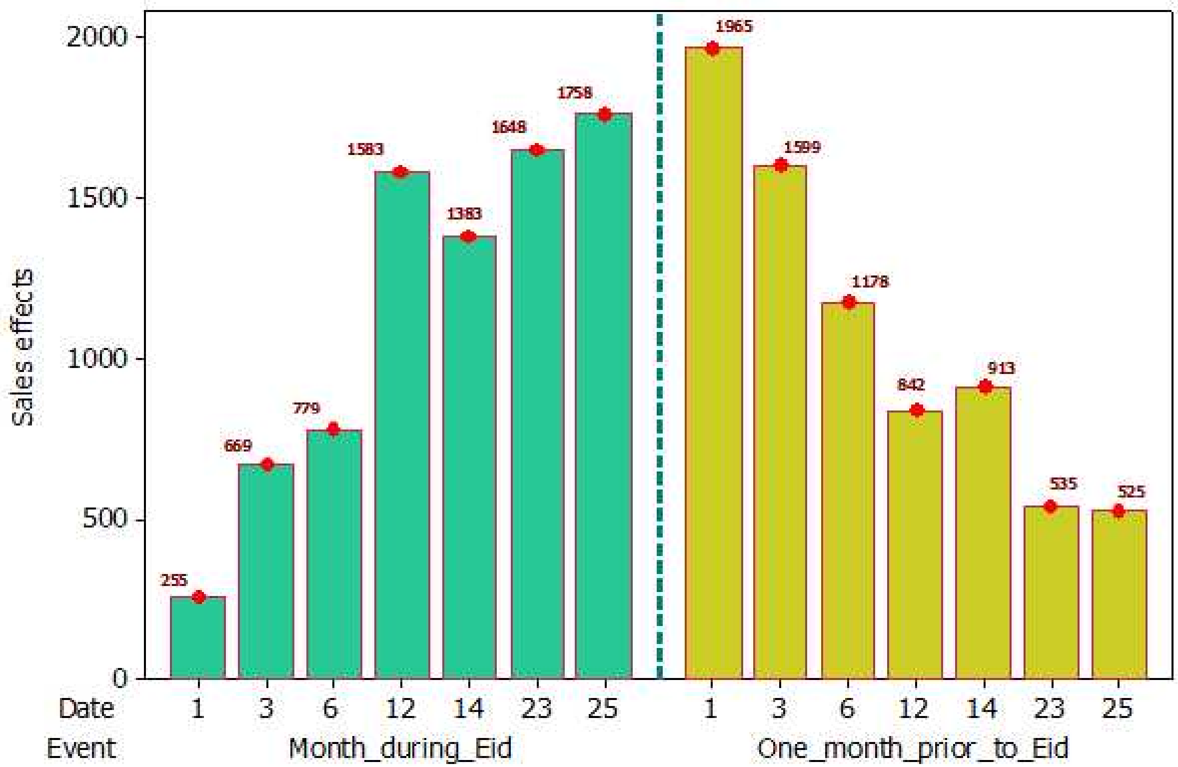


Fig. 2. Bar chart of Eid effects on the women's trouser sales in the month during and one month prior to the Eid celebration in Boyolali shop.

The aims

- To develop ***two levels calendar variation model*** based on **ARIMAX** and **regression** method for forecasting sales data with the ***Eid ul-Fitr*** effects.

- To compare the **forecast accuracy** with other forecasting methods, i.e.
  - **ARIMA** model
  - Feed-forward **Neural Networks** (**FFNN**)
  - **Two levels Time Series Regression**

Modeling method

- Model for **linear trend**:

$$y_t = \beta_0 + \beta_1 t + w_t \qquad \ldots (1)$$

- Regression with dummy variable for **seasonal pattern**:

$$y_t = \beta_0 + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + w_t \qquad \ldots (2)$$

- Regression for **calendar effects**:

$$y_t = \beta_0 + \beta_1 V_{1,t} + \beta_2 V_{2,t} + \cdots + \beta_p V_{p,t} + w_t \qquad \ldots (3)$$

The Two Levels Regression Model

- Model at the **first level** :

$$Y_t = \delta_1 t + \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_s S_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} +$$

$$\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \, .$$

- Model at the **second level** :

1. *Linear model*

$\Rightarrow$ $\hat{\alpha}_j = \upsilon_0 + \upsilon_1 j$

$\Rightarrow$ $\hat{\gamma}_j = \omega_0 + \omega_1 j$

2. *Exponential model*

$\Rightarrow$ $\hat{\alpha}_j = \upsilon_0 \, e^{\upsilon_1 j}$

$\Rightarrow$ $\hat{\gamma}_j = \ln(\omega_0 + \omega_1 j)$

The PROPOSED Model

- Model at the **first level** → **ARIMAX-1**: **stochastic** TREND-SEASONAL

$$Y_t = \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t$$

- Model at the **first level** → **ARIMAX-2**: **deterministic** TREND-SEASONAL

$$Y_t = \delta_1\, t + \beta_1 M_{1,t} + \beta_2 M_{2,t} + \cdots + \beta_s M_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + \frac{\theta_q(B)}{\phi_p(B)}\varepsilon_t$$

- Model at the **second level** :
  - ✓ *Linear model*

$$\boxed{\hat{\alpha}_j = \upsilon_0 + \upsilon_1 j}$$

$$\boxed{\hat{\gamma}_j = \omega_0 + \omega_1 j}$$

Background of Two Levels

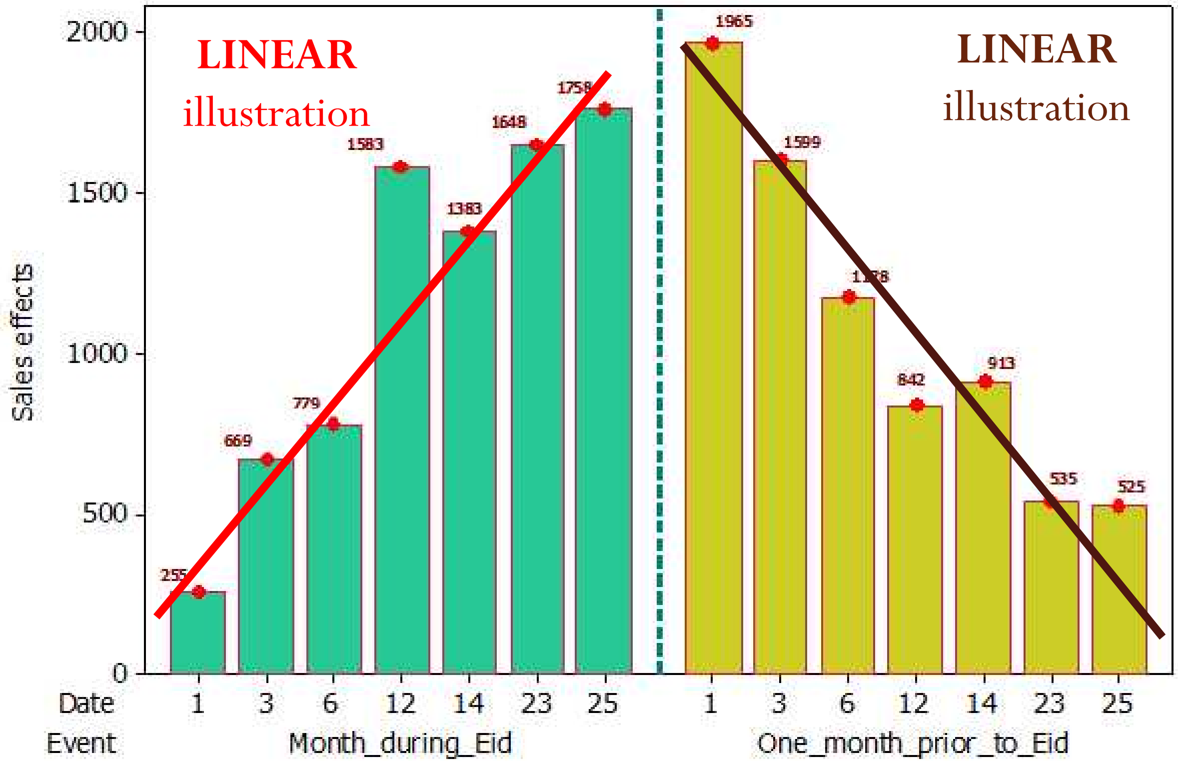


Fig. 2. Bar chart of Eid effects on the women's trouser sales in the month during and one month prior to the Eid celebration in Boyolali shop.

Dummy at Two Levels

Year	Date	Explanation
2002	06-07 December	$D_{5,t}$ = December, and $D_{5,t-1}$ = November
2003	25-26 November	$D_{24,t}$ = November, and $D_{24,t-1}$ = October
2004	14-15 November	$D_{13,t}$ = November, and $D_{13,t-1}$ = October
2005	03-04 November	$D_{2,t}$ = November, and $D_{2,t-1}$ = October
2006	23-24 October	$D_{22,t}$ = October, and $D_{22,t-1}$ = September
2007	12-13 October	$D_{11,t}$ = October, and $D_{11,t-1}$ = September
2008	01-02 October	$D_{0,t}$ = October, and $D_{0,t-1}$ = September
2009	21-22 September	$D_{20,t}$ = September, and $D_{20,t-1}$ = August
2010	10-11 September	$D_{9,t}$ = September, and $D_{9,t-1}$ = August
2011	30-31 August	$D_{29,t}$ = September, and $D_{29,t-1}$ = August

The Proposed Procedure

Step 1: Determination of dummy variable for calendar variation period.

Step 2: Remove the calendar variation effect form the response by fitting

$$Y_t = \beta_0 + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + N_t$$

for model with stochastic trend and seasonal model, or fitting

$$Y_t = \delta_1\, t + \beta_1 M_{1,t} + \beta_2 M_{2,t} + \cdots + \beta_s M_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + N_t$$

simultaneuously for model with deterministic trend and seasonal, to obtain the error, N_t.

Step 3: Find the best ARIMA model of N_t using Box-Jenkins procedure.

Step 4: Simultaneously fit the model from step 2 and 3. This model is the first level of calendar variation model based on ARIMAX method.

Step 5: Test the significance of parameter and perform diagnostic check.

Step 6: Estimate the second level model to predict the effects of calendar variation in every possibility number of days before Eid ul-Fitr.

Step 1

- Based on the time series plot, **TWO** DUMMY VARIABLES are used for evaluating calendar variation effect, i.e.
  - **The months prior to Eid ul Fitr**,

    $D_{j,t-1}$ = dummy variable for **ONE** month prior to Eid ul-Fitr celebration.
  - **During the month of Eid ul-Fitr celebration**,

    $D_{j,t}$ = dummy variable for **during** the month of Eid ul-Fitr celebration.
  - j = **number of days before Eid ul-Fitr celebration**

Step 2 - 6

- Model at the **first level** → **ARIMAX-1** : **stochastic** TREND-SEASONAL

$$Y_t = \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t$$

- Model at the **first level** → **ARIMAX-2**: **deterministic** TREND-SEASONAL

$$Y_t = \delta_1 t + \beta_1 M_{1,t} + \beta_2 M_{2,t} + \cdots + \beta_s M_{s,t} + \sum_j \alpha_j D_{j,t} + \sum_j \gamma_j D_{j,t-1} + \frac{\theta_q(B)}{\phi_p(B)}\varepsilon_t$$

- Model at the **second level** :
  - ✓ *Linear model*

➡ $\hat{\alpha}_j = \upsilon_0 + \upsilon_1 j$ ➡ $\hat{\gamma}_j = \omega_0 + \omega_1 j$

Results: monthly sales of men's jeans

a. ARIMAX-1 method

a.1. The first level model

$$Y_{1,t} = 0.109871 + 0.51153D_{2,t} + 0.79284D_{5,t} + 0.89457D_{11,t} + 1.21885D_{13,t} + 1.31881D_{22,t} + 1.44439D_{24,t} + 1.07963D_{2,t-1} + 0.80625D_{5,t-1} + 0.18174D_{11,t-1} + 0.38253D_{13,t-1} + 0.12689D_{24,t-1} + \frac{(1+0.57B^{12})}{(1-0.60498B)}\varepsilon_t .$$

a.2. The second level model

$$\hat{\alpha}_j = 0.537 + 0.0385\,j ,$$

$$\hat{\gamma}_j = 0.983 - 0.0431\,j .$$

Results: monthly sales of men's jeans

b. ARIMAX-2 method

b.1. The first level model

$$
\begin{aligned}
Y_{1,t} = {} & 0.21334M_{1,t} + 0.21110M_{2,t} + 0.24403M_{3,t} + 0.28412M_{4,t} + 0.24168M_{5,t} + \\
& 0.29194M_{6,t} + 0.31880M_{7,t} + 0.34575M_{8,t} + 0.33287M_{9,t} + 0.44423M_{10,t} + \\
& 0.12521M_{11,t} + 0.27299M_{12,t} + 0.69022D_{2,t} + 0.85376D_{5,t} + 0.67169D_{11,t} + \\
& 1.38182D_{13,t} + 1.10961D_{22,t} + 1.61269D_{24,t} + 0.94901D_{2,t-1} + 1.05741D_{5,t-1} + \\
& 0.5262D_{11,t-1} + 0.24495D_{13,t-1} + \frac{1}{(1-0.58642B)}\varepsilon_t .
\end{aligned}
$$

b.2. The second level model

$$\hat{\alpha}_j = 0.626 + 0.0333j ,$$

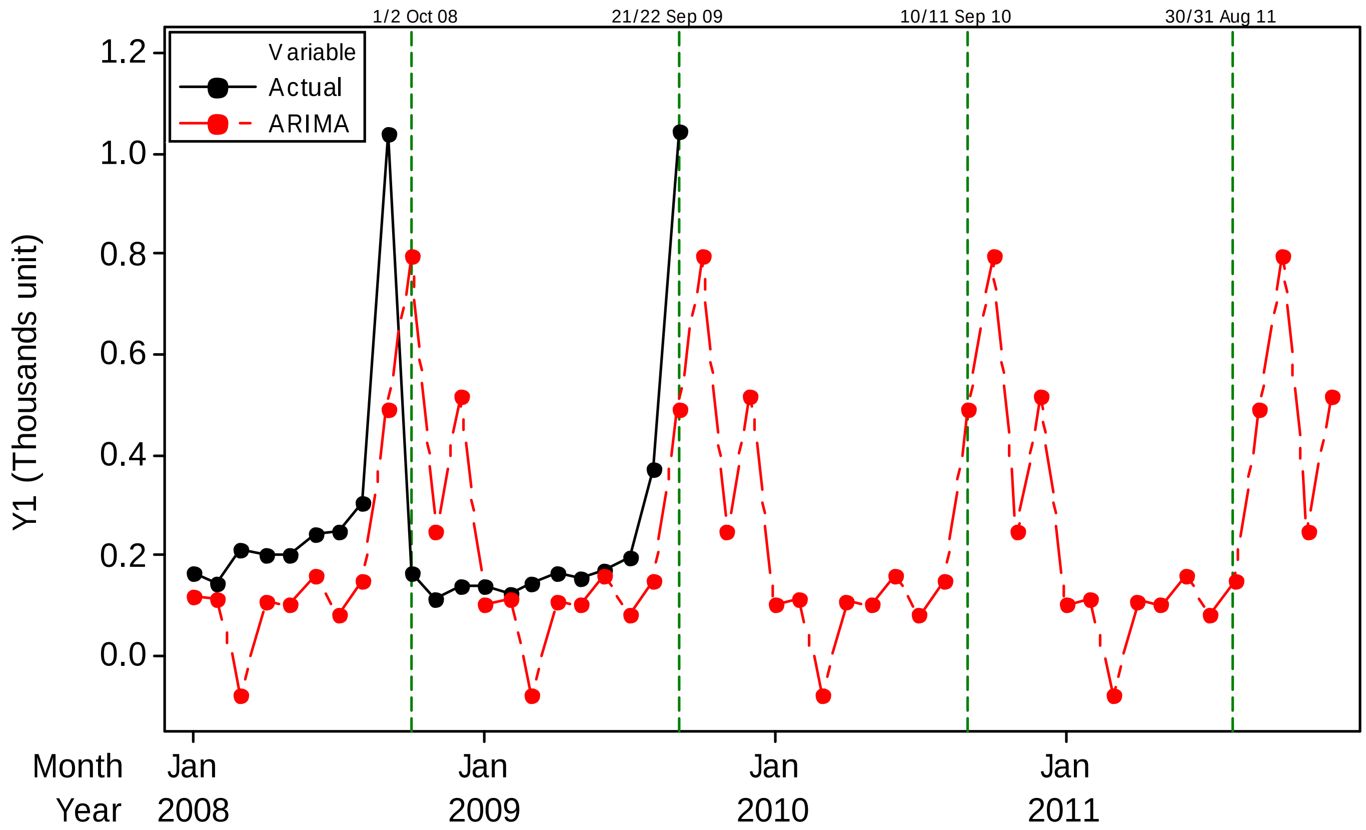
$$\hat{\gamma}_j = 1.018 - 0.0481j .$$

Results

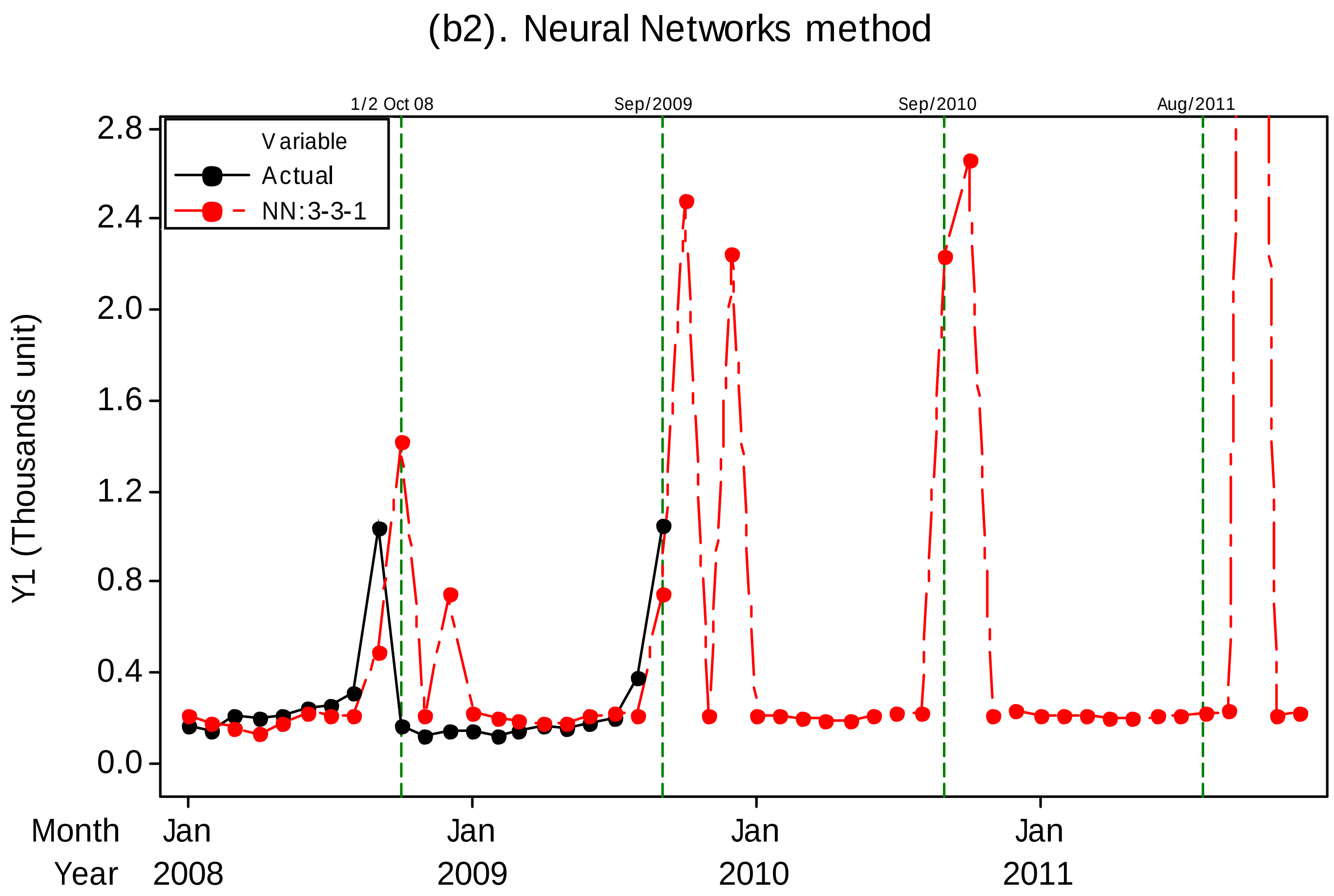
Method	$Y_{1,t}$		$Y_{2,t}$	
	in-sample	out-sample	in-sample	out-sample
ARIMA	**0.1408**	**0.2634**	**0.1685**	**0.4235**
FFNN: no skip layer				
3-1-1	0.1188	**0.3847**	0.0845	0.3290
3-2-1	0.0809	4.3466	0.0741	0.3844
3-3-1	0.0786	0.3375	0.0657	**0.2789**
...	...	...	...	...
3-10-1	0.0894	5.6064	0.0598	10.6219
FFNN: with skip layer				
3-1-1	0.1148	**0.4159**	0.0889	0.3273
3-2-1	0.0809	0.5659	0.0710	0.3383
3-3-1	**0.0708**	0.6290	0.0663	**0.2855**
...	...	...	...	...
3-10-1	0.1087	1.9E+07	0.0561	0.8200
Two levels regression	**0.0686**	**0.2434**	**0.0510**	**0.1508**
Two levels ARIMAX-1	**0.0671**	**0.2169**	**0.0742**	**0.1757**
Two levels ARIMAX-2	**0.0606**	**0.2599**	**0.0483**	**0.1688**

Graphical Results

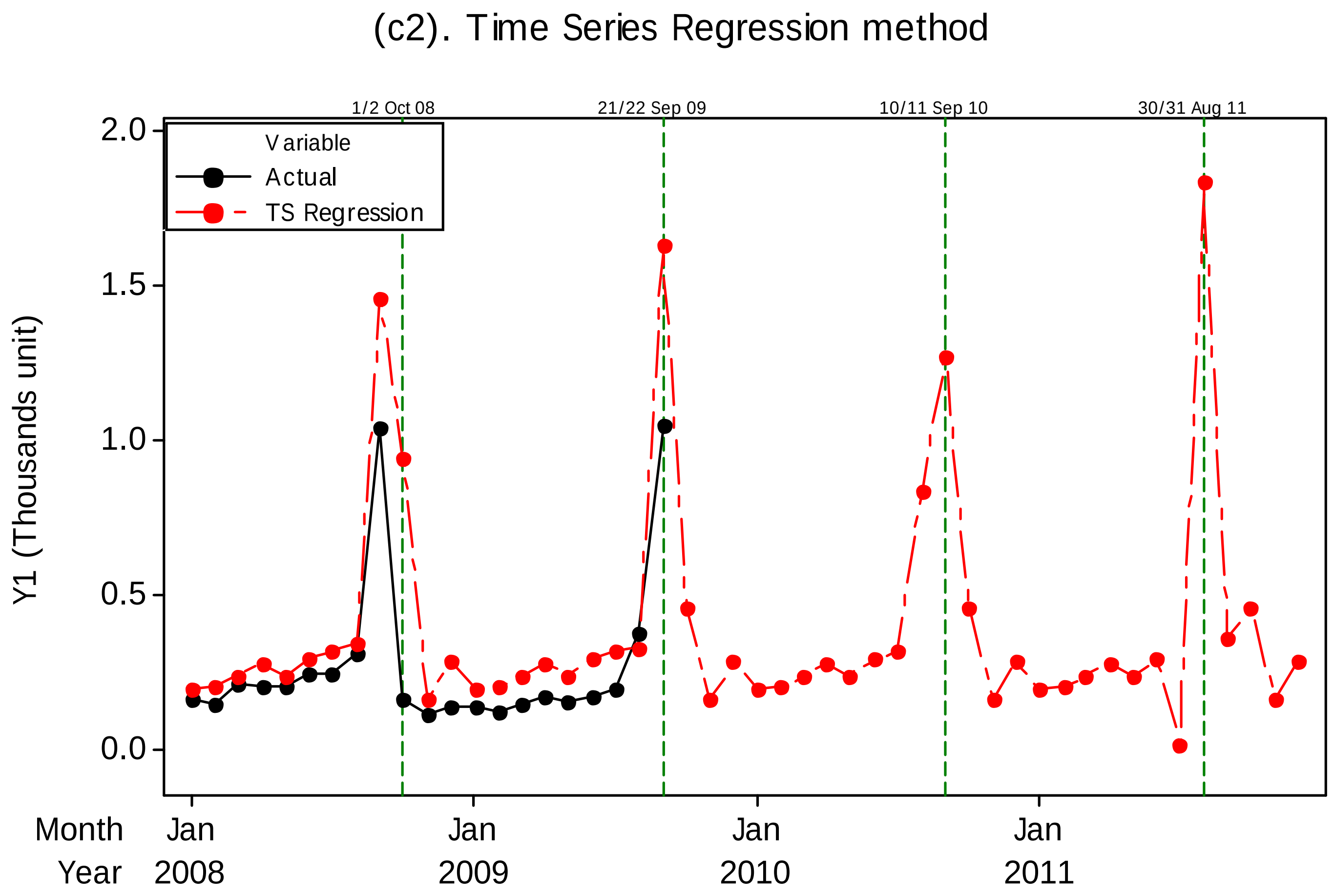
(a2). ARIMA method



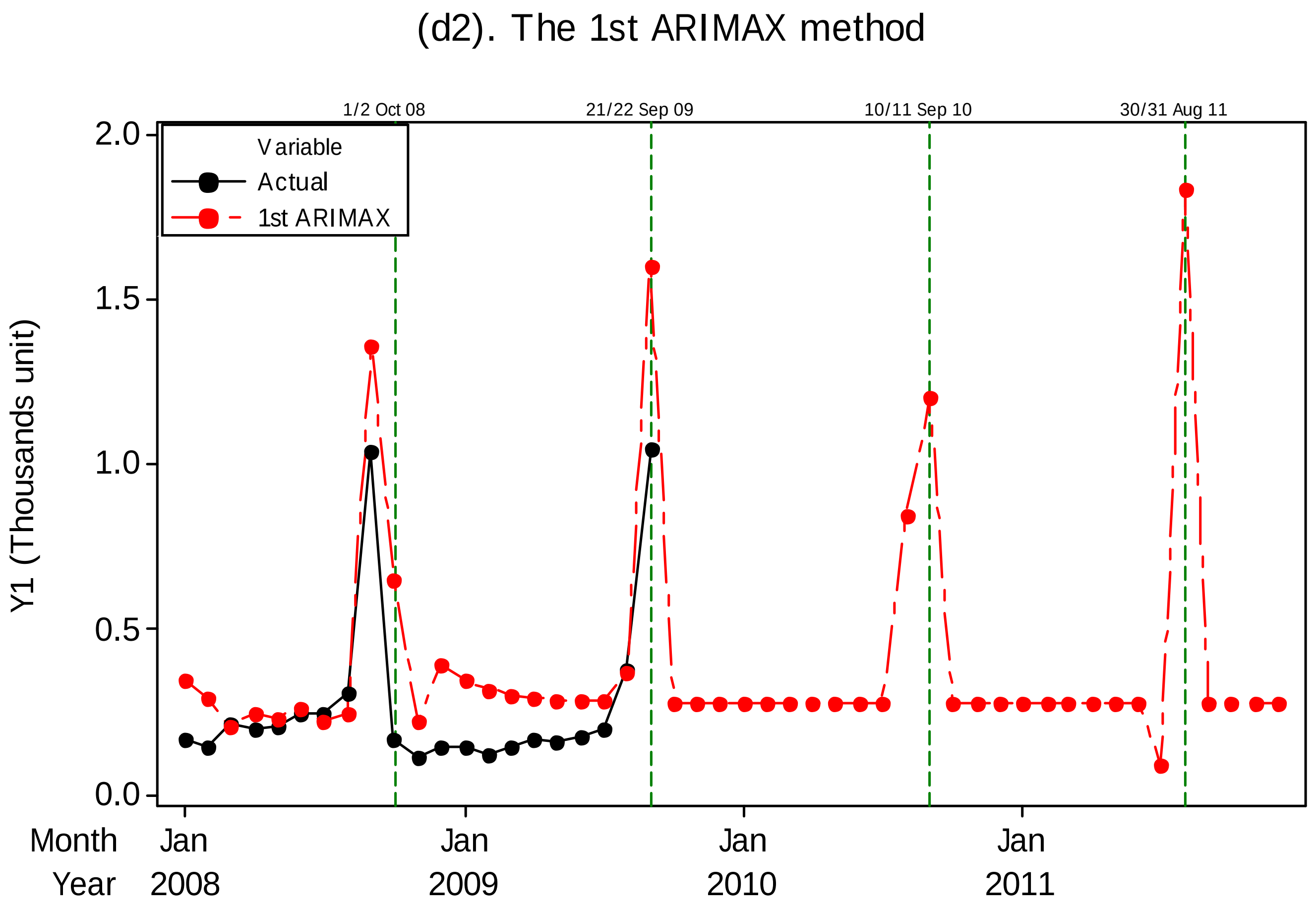
Graphical Results



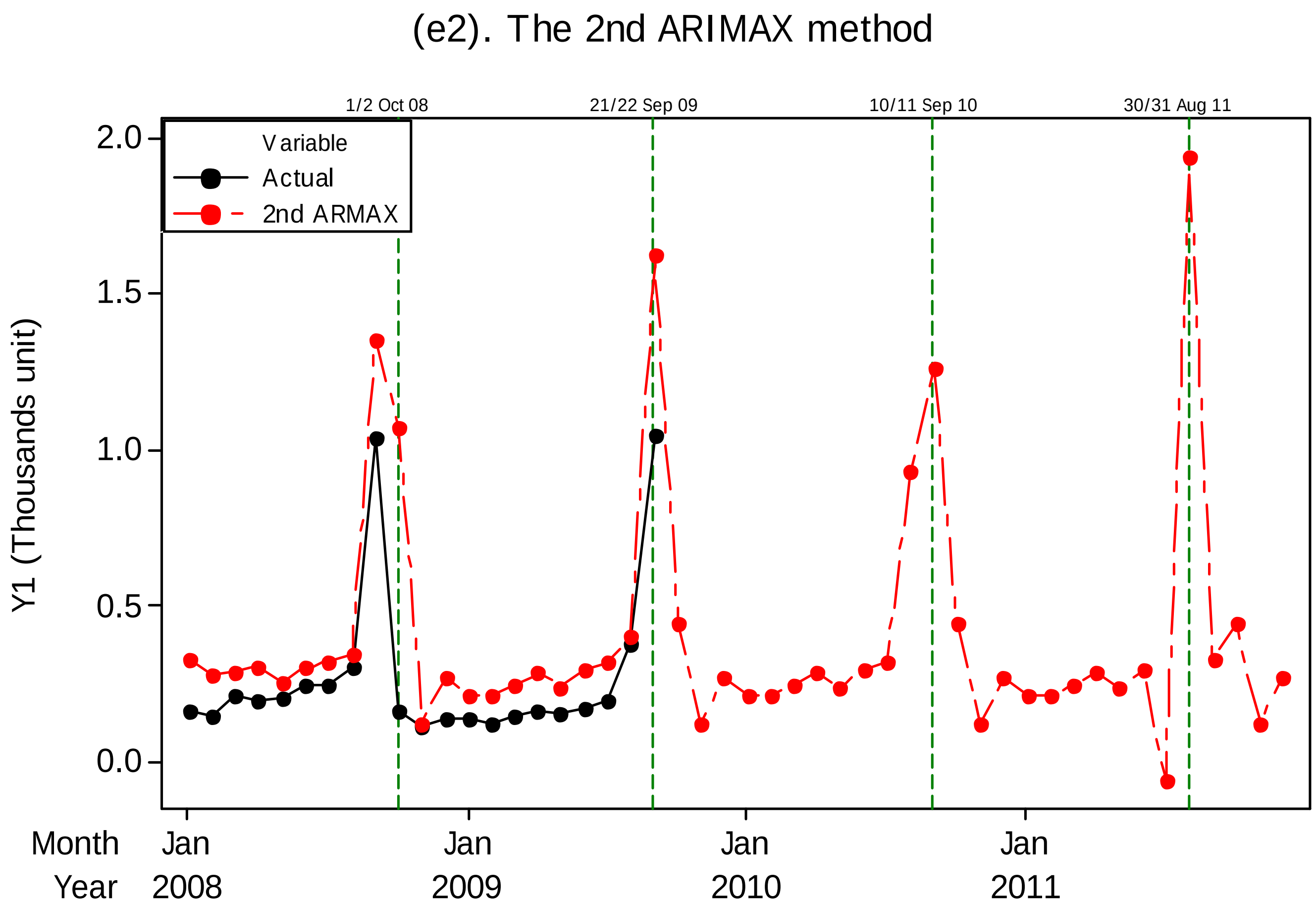
Graphical Results



Graphical Results



Graphical Results



Conclusion

- The **proposed two levels calendar variation model** based on **ARIMAX** and **Regression** method **yield better prediction for out-sample data**, compared to those of **ARIMA** model and **neural networks**.
- The application of **ARIMA** model usually yield **spurious results**, particularly about **seasonal pattern** and the presence of **outliers**.
- Whereas, **Neural Networks perform well only** for **in-sample** data.

References

1. L.M. Liu, Analysis of Time Series with Calendar Effects, Management Science, **2**, 106-112 (1980).
2. L.M. Liu, Identification of Time Series Models in the Presence of Calendar Variation, International Journal of Forecasting, **2**, 357-372 (1986).
3. T.C. Mills and J.A. Coutts, Calendar effects in the London Stock Exchange FT-SE indices, The European Journal of Finance, **1**, 79-93 (1995).
4. L.Gao and G. Kling, Calendar Effects in Chinese Stock Market, Annals of Economics and Finance, **6**, 75-88 (2005).
5. R. Sullivan, A. Timmermann and H. White, Dangers of data mining: The case of calendar effects in stock returns, Journal of Econometrics, **105**, 249-286 (2001).
6. F.J. Seyyed, A. Abraham and M. Al-Hajji, Seasonality in stock returns and volatility: The Ramadan effect, Research in International Business and Finance, **19**, 374-383 (2005).
7. Suhartono and M.H. Lee, Two levels regression modeling of trading day and holiday effects for forecasting retail data, Proceeding The 7th IMT-GT International Conference on Mathematics, Statistics and its Applications (ICMSA 2011), Bangkok, Thailand, pp. 150-164.
8. W.S. Cleveland and S.J. Devlin, Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods, Journal of the American Statistical Association, **75**(371), 487-496 (1980).

References

9. C. Brooks and G. Persand, Seasonality in Southeast Asian stock markets: some new evidence on day-of-the week effects, Applied Economics Letters, **8**, 155-158 (2001).

10. K. Holden, J. Thompson and Y. Ruangrit, The Asian crisis and calendar effects on stock returns in Thailand, European Journal of Operational Research, **163**, 242–252 (2005).

11. W.W.S. Wei, Time Series Analysis: Univariate and Multivariate Methods, 2nd edition, California: Addison-Wesley Publishing Company, Inc. (2006).

12. G.E.P. Box, G.M. Jenkins and G.C. Reinsel, Time Series Analysis, Forecasting and Control, 4th edition, New Jersey: John Wiley & Sons, Inc. (2008).

13. J.D. Cryer and K.S. Chan, Time Series Analysis: with Application in R, 2nd edition, New York: Springer-Verlag. (2008).

14. Suhartono, Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive, Journal of Mathematics and Statistics, **7** (1), 20-27, (2011).

15. G.E.P. Box and G.M. Jenkins, Time Series Analysis: Forecasting and Control, San Fransisco: Holden-Day, Revised ed. (1976).

16. Suhartono and M.H. Lee, A Hybrid Approach based on Winter's Model and Weighted Fuzzy Time Series for Forecasting Trend and Seasonal Data. Journal of Mathematics and Statistics, 7 (3), 177-183, (2011).

2 Days Workshop

Faculty of MIPA, UNIVERSITAS ANDALAS PADANG
& INSTITUT TEKNOLOGI SEPULUH NOPEMBER

Applying Data Analytics Using Neural Networks

Suhartono

(B.Sc.-ITS; M.Sc.-UMIST,UK; Dr.-UGM; Postdoctoral-UTM)

Department of Statistics,
Institut Teknologi Sepuluh Nopember, Indonesia

Email: *suhartono@statistika.its.ac.id*, *gmsuhartono@gmail.com*

Department of Mathematics, Universitas Andalas, Padang

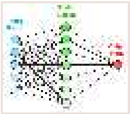
17-18 July 2017

Outline

- **Introduction**: *Background*, *Motivation*, *Jargons*, *Goals*.
- **Architecture of Neural Networks:** *Supervised* & *Unsupervised networks*
- **Model selection in Neural Networks**: *Inputs*, *Number of hidden neurons*, *Activation function*, *Preprocessing method*.
- **Application and Development:** *Forecasting and Classification problems*.

Neural Networks – NN

- **Sven F. Crone:**
  http://www.sven-crone.com/presentations.htm
  http://www.neural-forecasting.com/
- **Halbert L. White:**
  http://weber.ucsd.edu/~hwhite/
- **Warren S. Sarle:**
  ftp://ftp.sas.com/pub/neural/FAQ.html



General Background

- During the last few decades,
  - ❶ modeling to explain nonlinear relationship between variables, and
  - ❷ some procedures to detect this nonlinear relationship

  have grown in a spectacular way and received a great deal of attention.
  - Granger, C.W.J. and Terasvirta,T., (1993)
  - Terasvirta, T., Tjostheim, D. and Granger, C.W.J., (1994)

- Due to computational advances and increased computational power, nonparametric models that do not make assumptions about the parametric form of the functional relationship between the variables to be modelled have become more easily applicable.

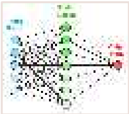
Motivation of NN Research

- Today's research is largely motivated by the possibility of using NN model as an instrument to solve a wide variety of application problems such as:

  pattern recognition (classification), signal processing, process control, and forecasting.

- The use of the NN model in applied work is generally motivated by a mathematical result stating that under mild regularity conditions, a relatively simple NN model is capable of approximating any Borel-measureable function to any given degree of accuracy.

  (see e.g. Hornik, Stichombe and White (1989, 1990), White (1990); Cybenko (1989))



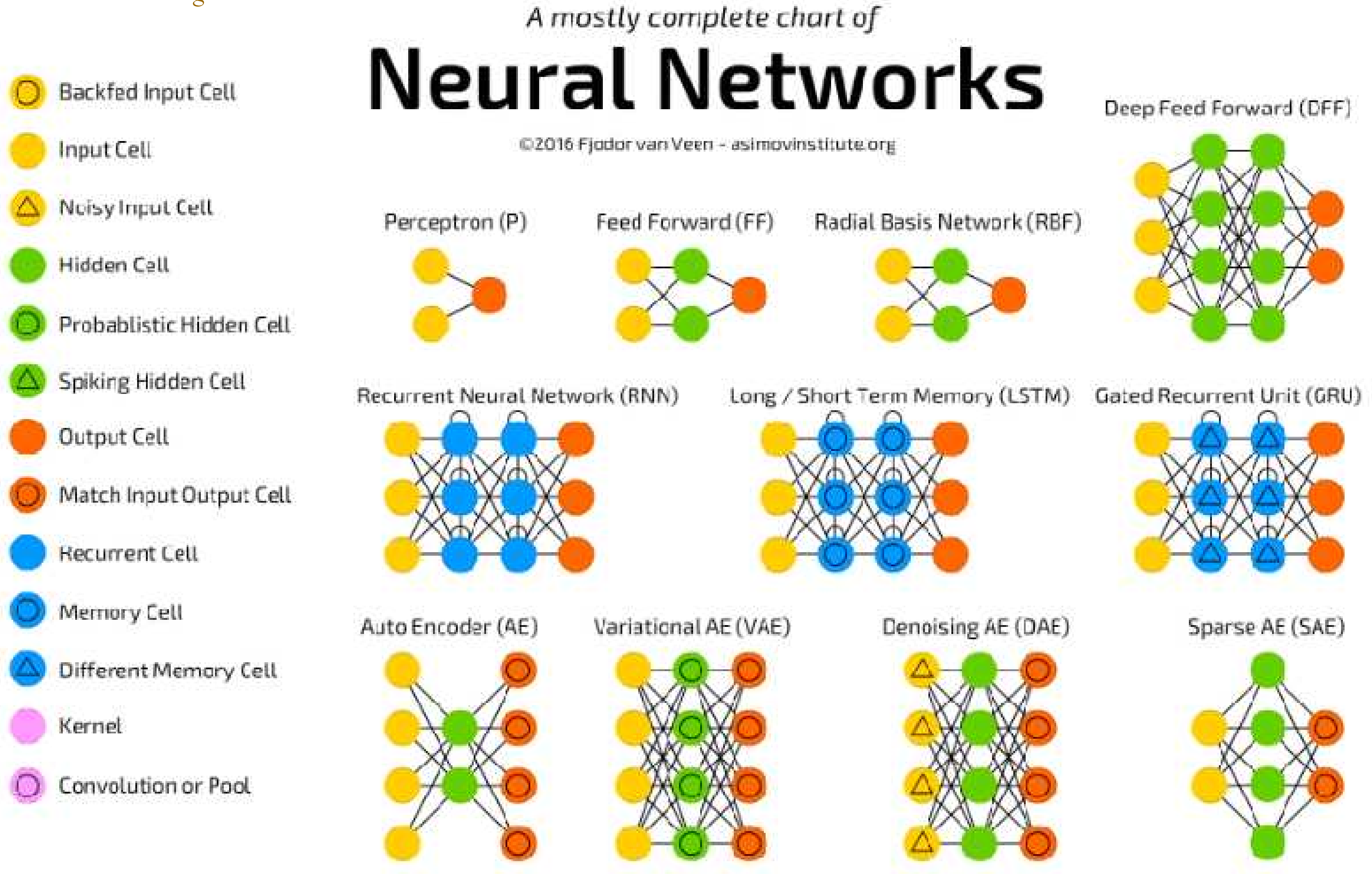
The use of NN ... [Sarle, 1994]

1. as models of biological nervous systems and "intelligence",
2. as real-time adaptive signal processors or controllers implemented in hardware for applications such as robots,
3. as data analytic methods.

➲ This paper is concerned with NN for **DATA ANALYSIS**.

Chart of Neural Networks

http://www.asimovinstitute.org/neural-network-zoo/



Supervised vs Unsupervised networks

	Supervised Learning	Unsupervised Learning
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

Dependence Methods

Interdependence Methods

Multivariate Data Analysis

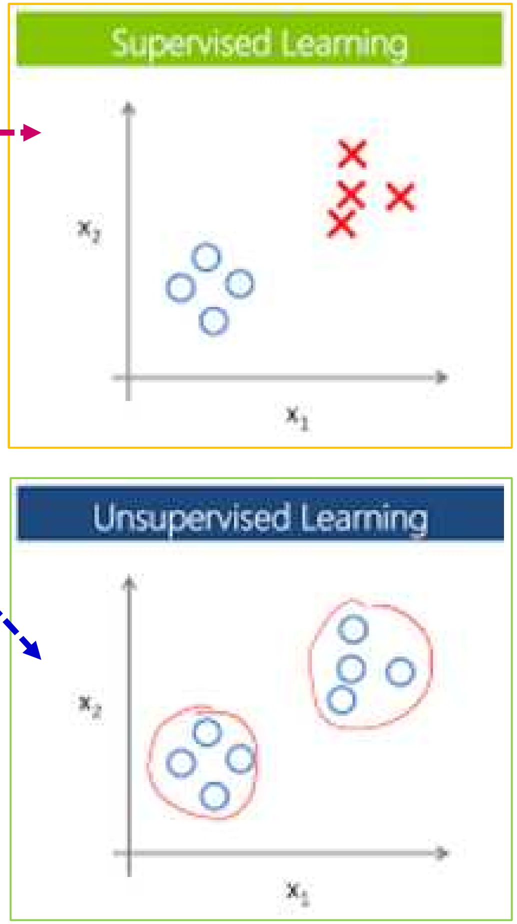


Chart of Neural Networks

http://www.asimovinstitute.org/neural-network-zoo/

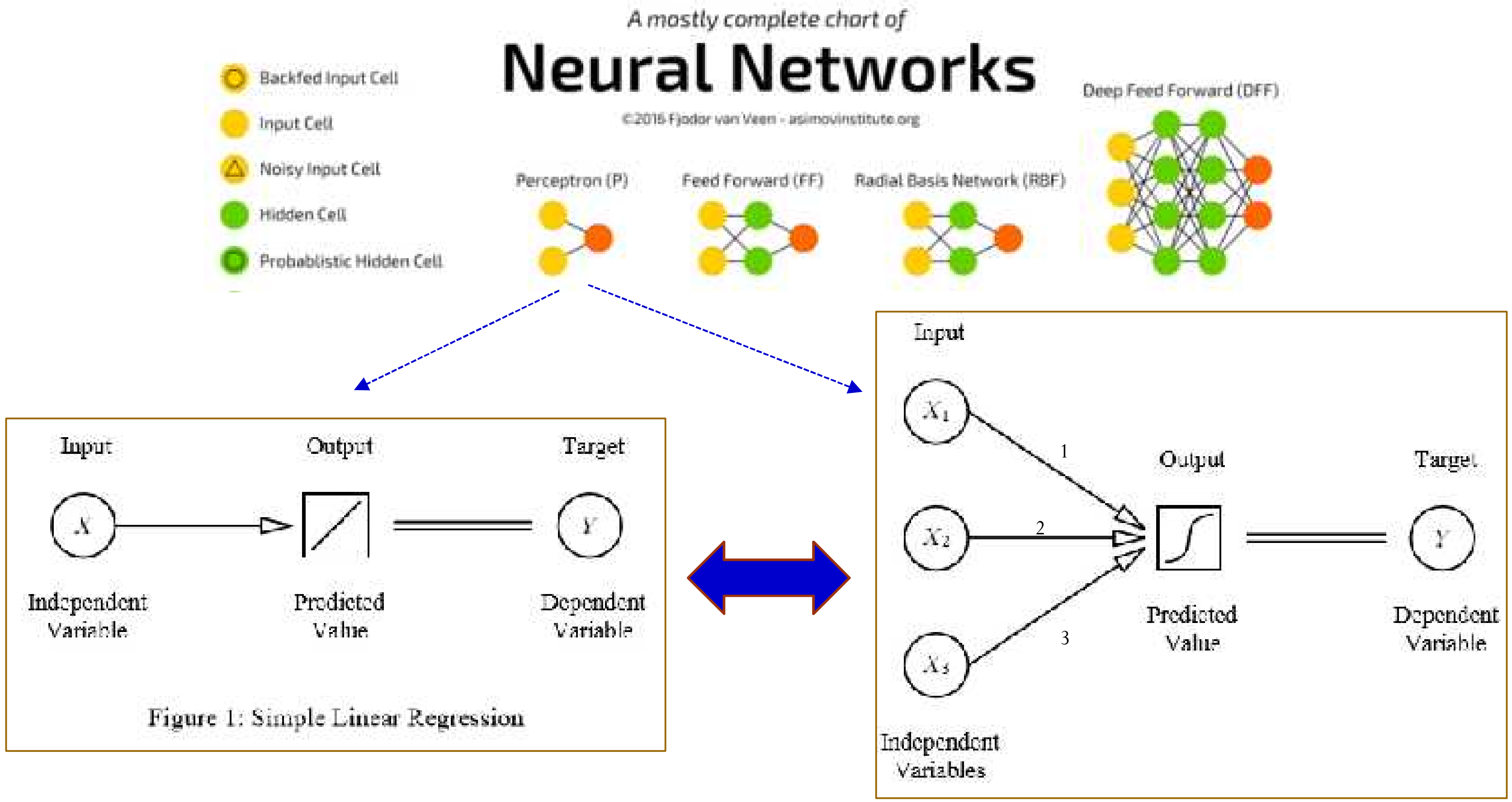
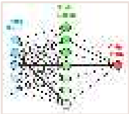


Figure 2: Simple Nonlinear Perceptron = Logistic Regression

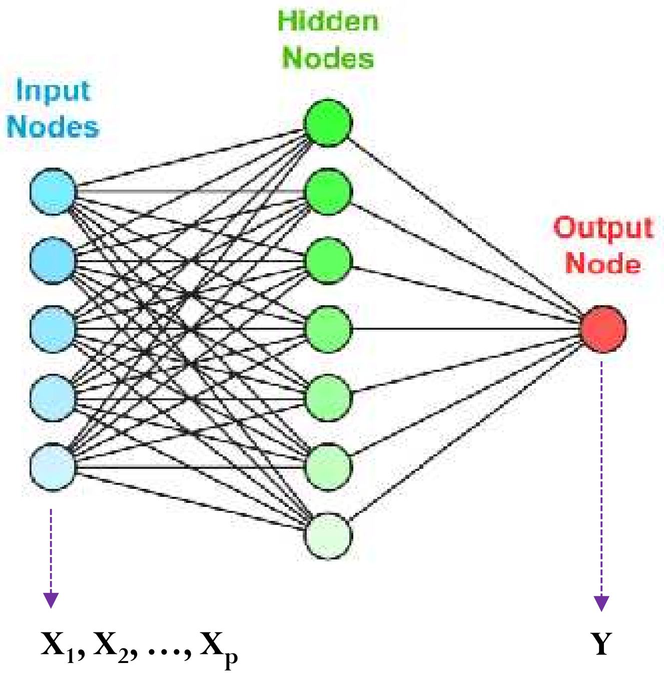
Source: Sarle (1994)



Feed Forward Neural Networks

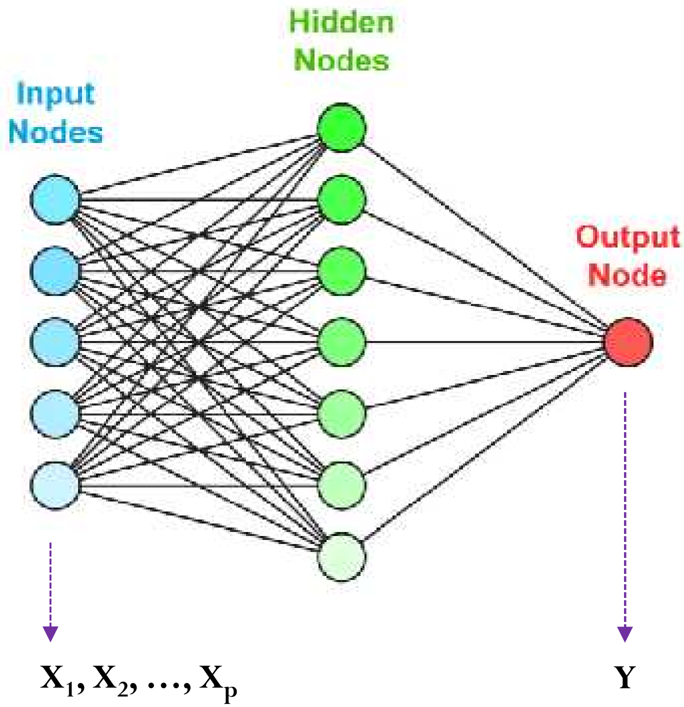
- **Multilayer perceptron** (**MLP**), also known as **feedforward neural networks** (**FFNN**), is probably the **most commonly** used NN architecture in engineering application.
- Typically, applications of NN for regression, time series modeling and classification (discriminant analysis) are **based on the FFNN architecture**.

Neural Networks & Statistical Jargon



Neural Networks	Statistics
▪ **features**	▪ **variables**
▪ **inputs**	▪ **independent variables**
▪ **outputs**	▪ **predicted values**
▪ **targets or training values**	▪ **dependent variables**
▪ errors	▪ residuals
▪ training, learning, adaptation	▪ estimation
▪ patterns or training pairs	▪ observations
▪ **weights**	▪ **parameter estimates**
▪ **supervised learning**	▪ **regression & discriminant**
▪ unsupervised learning	▪ data reduction
▪ adaptive vector quantization	▪ cluster analysis
▪ generalization	▪ interpolation & extrapolation

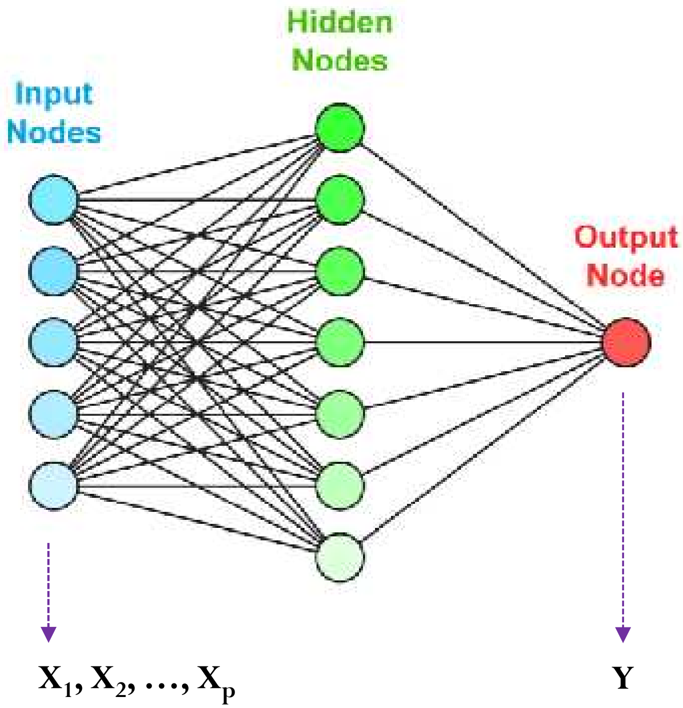
FFNN as Nonlinear regression



- FFNN includes estimated weights between the inputs and the hidden layer, and the hidden layer uses nonlinear activation functions such as the logistic function, the FFNN becomes genuinely nonlinear model, i.e., nonlinear in the parameters.

- In this case, FFNN can be seen as nonlinear regression. FFNN can have multiple inputs and outputs (This figure is multiple inputs with single output), and this architecture is similar to **multiple nonlinear regression**.

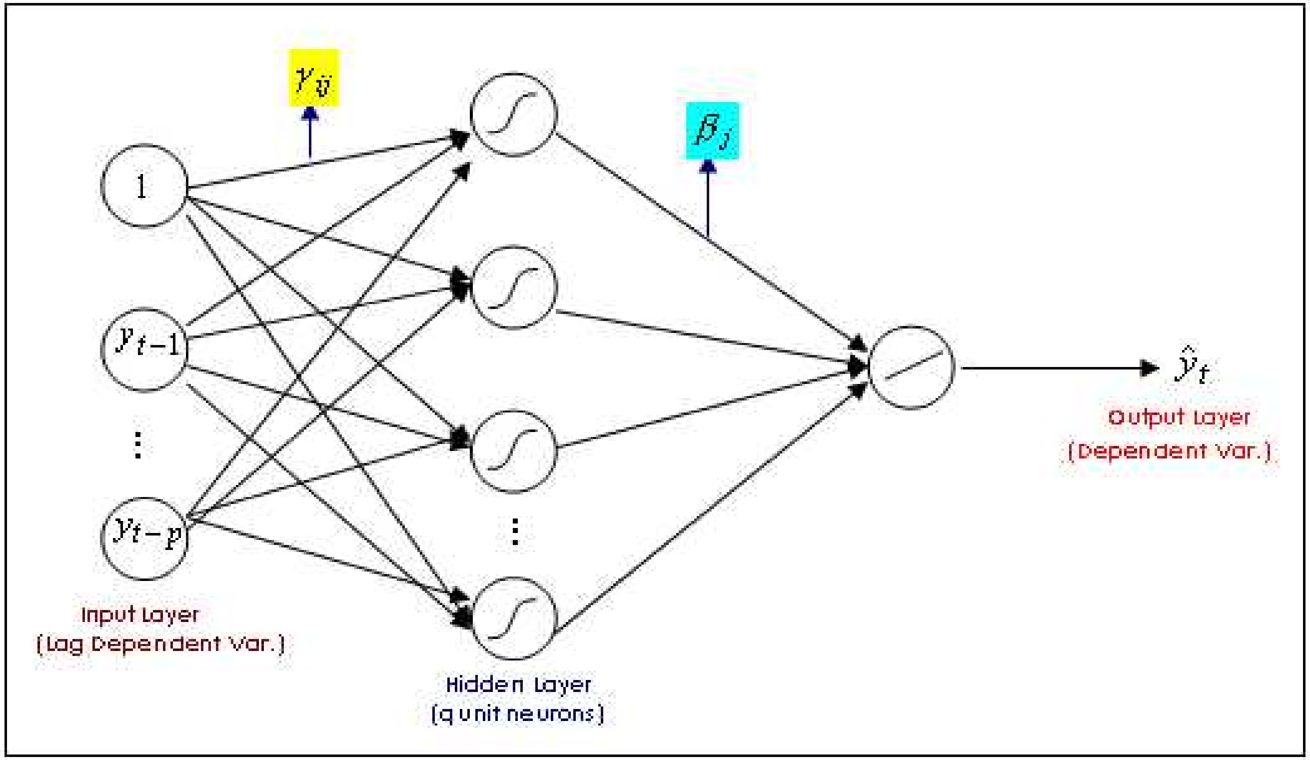
FFNN as Logistic Regression and Discriminant Analysis



- FFNN with nonmetric data (dichotomous/polycothomus) in target values is identical to **logistic regression** and **nonlinear discriminant analysis**.

- In this case, FFNN often use a multiple logistic function to estimate the conditional probabilities of each class. A multiple logistic function is called a ***softmax*** activation function in the NN literature.

FFNN as Nonlinear AR(p) model



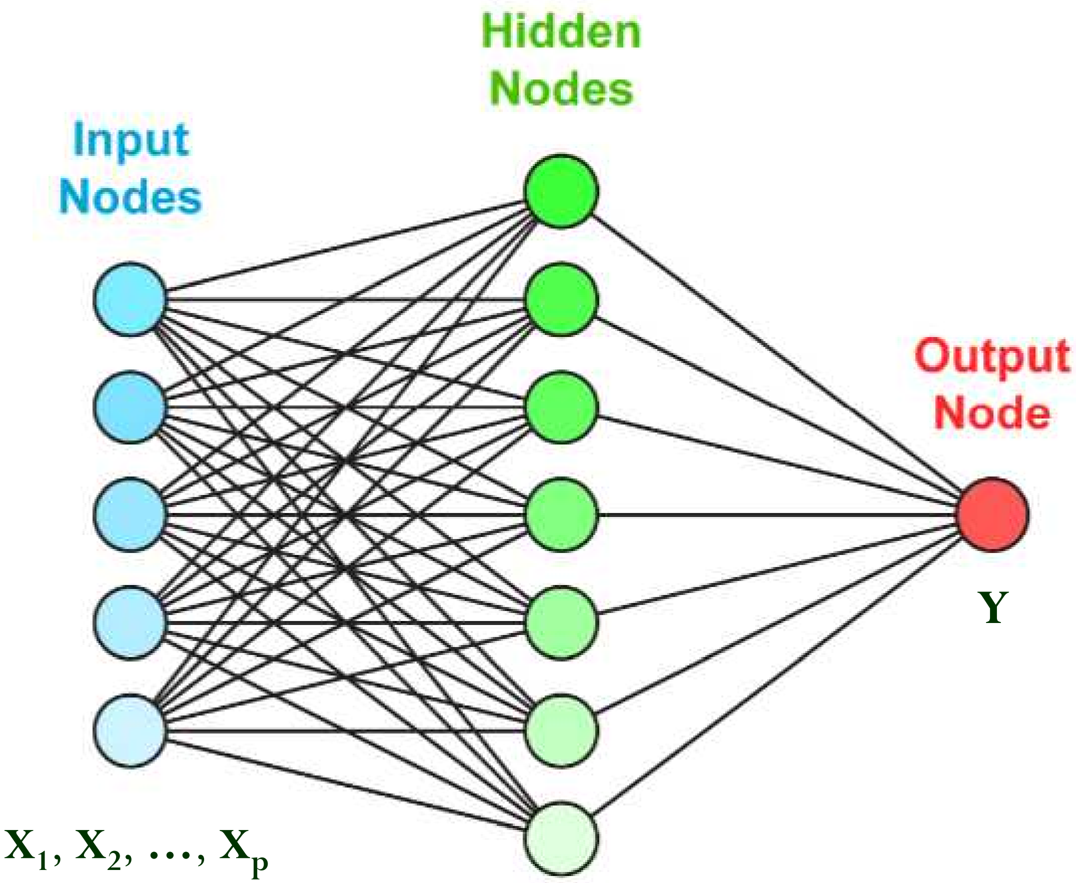
$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j f\left(\sum_{i=1}^{p} \gamma_{ij} y_{t-i} + \gamma_{oj} \right) + \varepsilon_t$$

FFNN as Nonlinear AR(p) model

- Model building strategy that proposed by Terasvirta *et al.* (1994)

  1. Test Y_t for linearity, using linearity test (neglected nonlinearity).
  2. If linearity is rejected, consider a small number of alternative parametric models and/or nonparametric models.
  3. These models should be estimated in-sample and compared out-of-sample.

FFNN: the main problems !!!

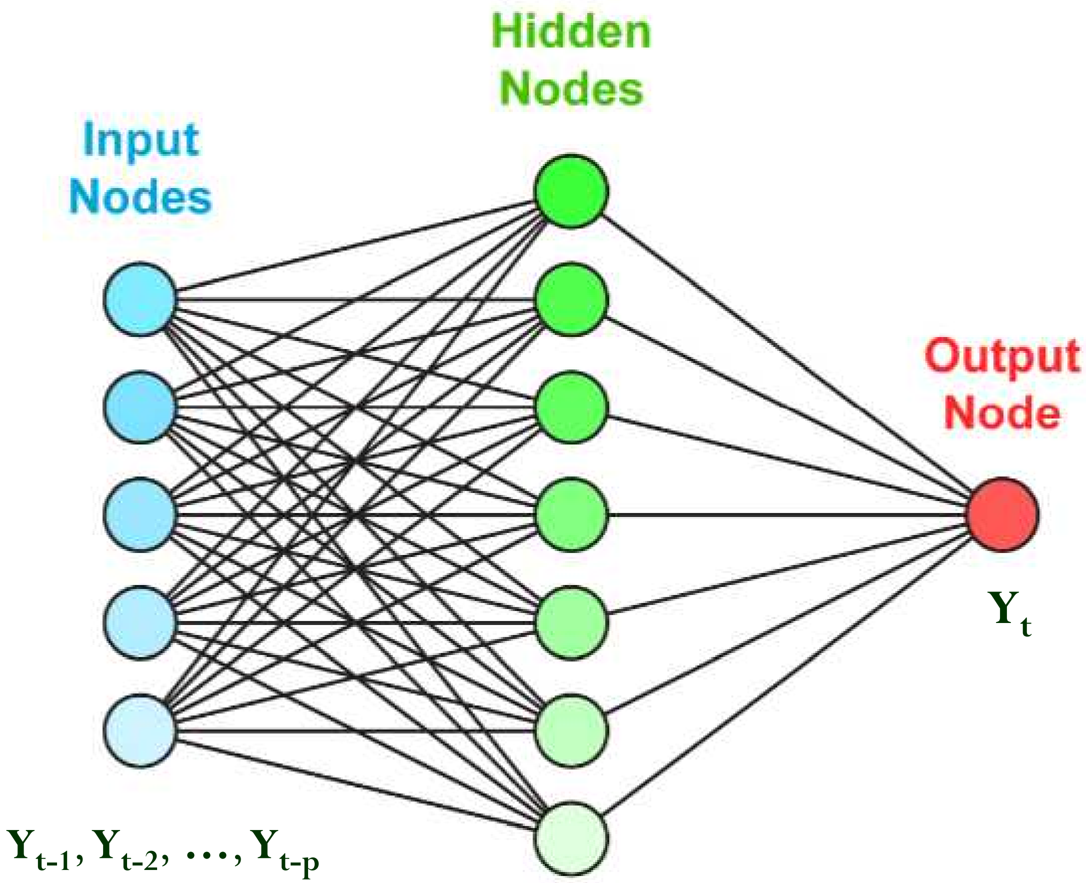


In Classification & Regression

1. How many **nodes** (**neurons**) in hidden layer?
2. What is the best **inputs** (**features selection**)?
3. What is the best **pre-processing** method?
4. What is the best **activation function** in **hidden** and **output** layer?

Model selection in Neural Networks

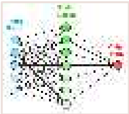
FFNN: the main problems !!!



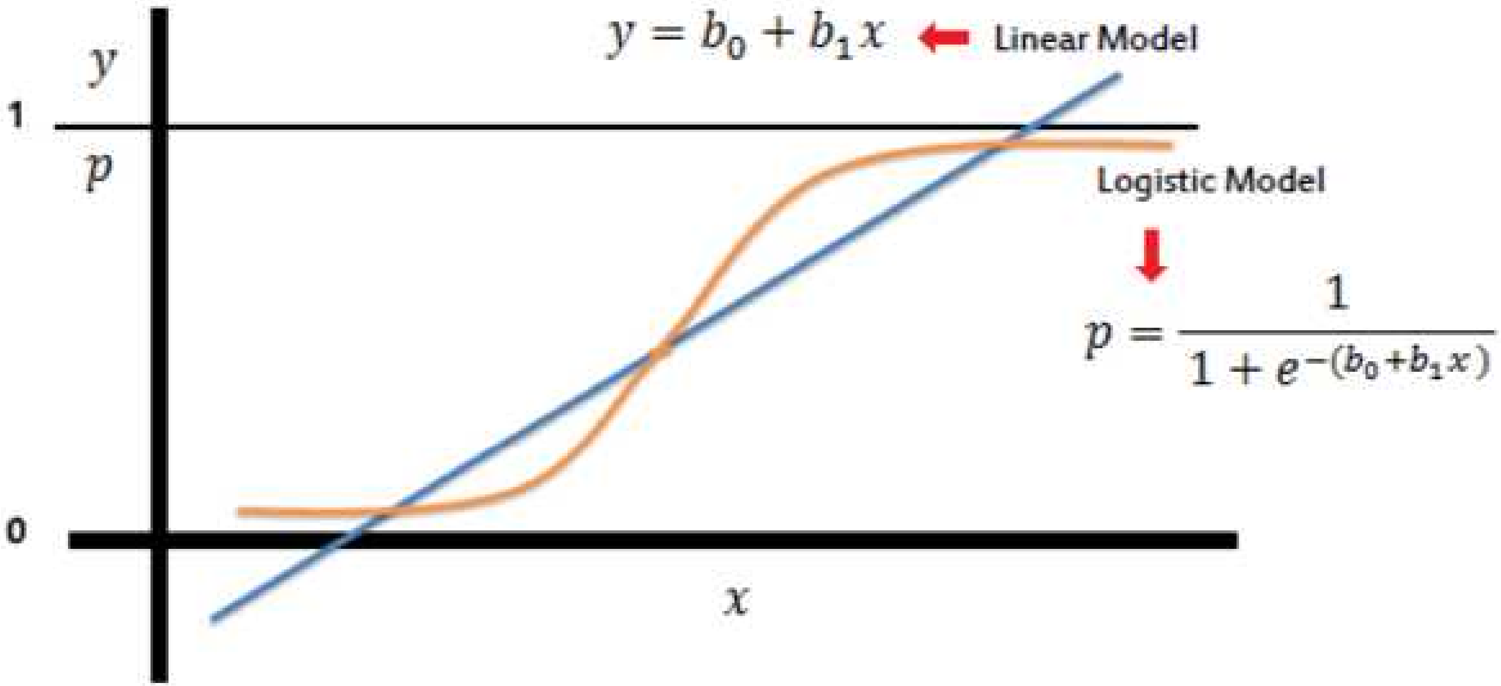
In **Time Series Forecasting**

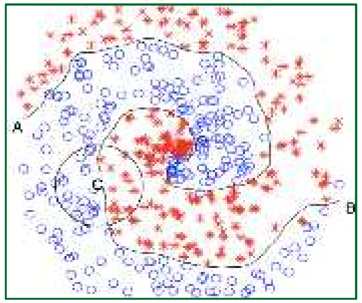
1. What is the best **inputs** (**features selection**)?
2. How many **nodes** (**neurons**) in hidden layer?
3. What is the best **pre-processing** method?
4. What is the best **activation function** in **hidden** and **output** layer?

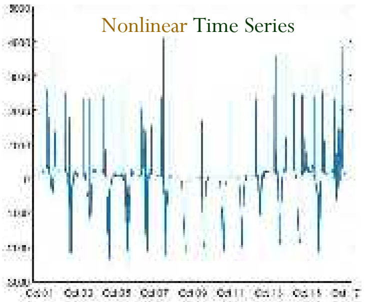
Model selection in Neural Networks

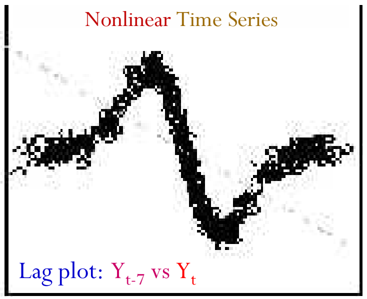


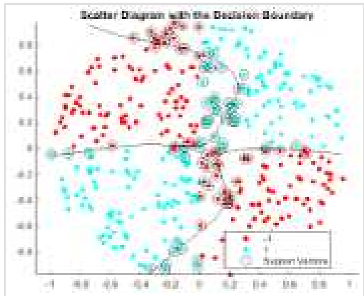
Nonlinear relationship Concept







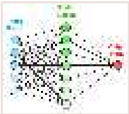




Model Selection in Neural Network

- In general, there are **two procedures** usually used to find **the best FFNN model** or **the optimal architecture**, those are "general-to-specific" or "top-down" and "specific-to-general" or "bottom-up" procedures.

➲ "Top-down" procedure is started from complex model and then applies an algorithm to reduce number of parameters (number of input variables and unit nodes in hidden layer) by using some stopping criteria, whereas "bottom-up" procedure works from a simple model.



Neural Networks "software"

R: Training of neural networks ▾ Find in Topic

neuralnet {neuralnet} R Documentation

Training of neural networks

Description

`neuralnet` is used to train neural networks using backpropagation, re
(Riedmiller, 1994) or without weight backtracking (Riedmiller and Braun
version (GRPROP) by Anastasiadis et al. (2005). The function allows fl
error and activation function. Furthermore the calculation of generalized
is implemented.

Usage

```
neuralnet(formula, data, hidden = 1, threshold = 0.
          stepmax = 1e+05, rep = 1, startweights =
          learningrate.limit = NULL,
          learningrate.factor = list(minus = 0.5, p
          learningrate=NULL, lifesign = "none",
          lifesign.step = 1000, algorithm = "rprop+
          err.fct = "sse", act.fct = "logistic",
          linear.output = TRUE, exclude = NULL,
          constant.weights = NULL, likelihood = FAL
```

R: Fit Neural Networks ▾ Find in Topic

nnet {nnet}

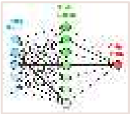
Fit Neural Networks

Description

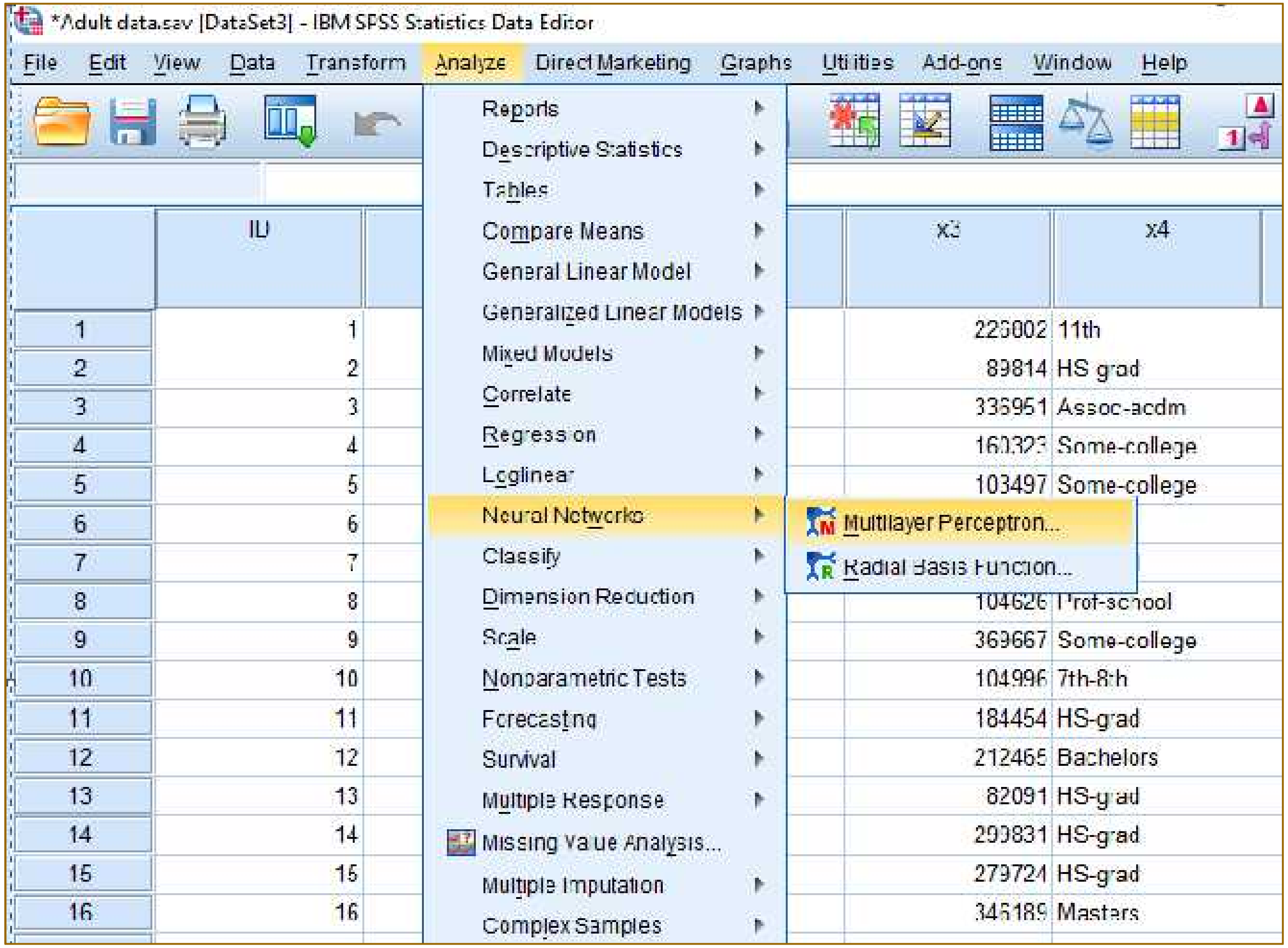
Fit single-hidden-layer neural network, possibly with skip-layer connections.

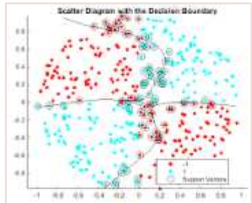
Usage

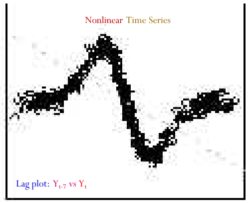
```
nnet(x, ...)

## S3 method for class 'formula'
nnet(formula, data, weights, ...,
     subset, na.action, contrasts = NULL)

## Default S3 method:
nnet(x, y, weights, size, Wts, mask,
     linout = FALSE, entropy = FALSE, softmax = FALSE,
     censored = FALSE, skip = FALSE, rang = 0.7, decay = 0,
     maxit = 100, Hess = FALSE, trace = TRUE, MaxNWts = 1000,
     abstol = 1.0e-4, reltol = 1.0e-8, ...)
```

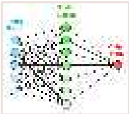


Neural Networks "software"

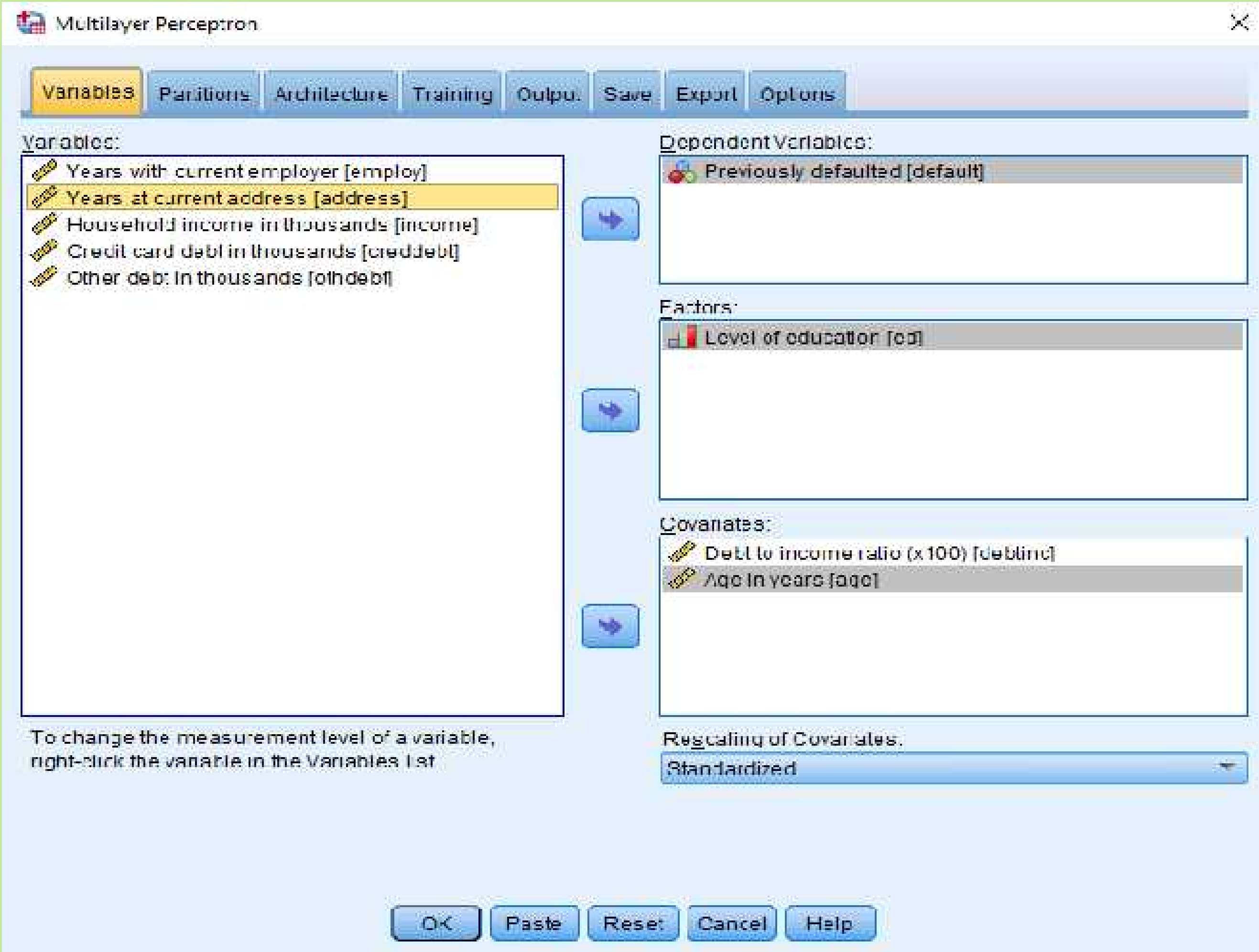


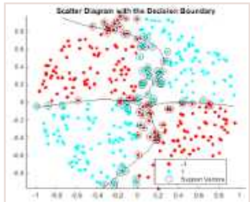


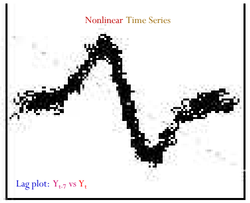




Neural Networks “software”







Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Variable	Notation
Y	default (Yes = 1 , No = 0)
X_1	age
X_2	ed (categorical)
X_3	employ
X_4	address
X_5	income
X_6	debtinc
X_7	creddebt
X_8	othdebt

Name	Type	Width	Decimals	Label
age	Numeric	4	0	Age in years
ed	Numeric	4	0	Level of education
employ	Numeric	4	0	Years with current employer
address	Numeric	4	0	Years at current address
income	Numeric	8	2	Household income in thousands
debtinc	Numeric	8	2	Debt to income ratio (x100)
creddebt	Numeric	8	2	Credit card debt in thousands
othdebt	Numeric	8	2	Other debt in thousands
default	Numeric	4	0	Previously defaulted

Level of education →

1 = "Did not complete high school"
2 = "High school degree"
3 = "Some college"
4 = "College degree"
5 = "Post-undergraduate degree"

Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Input variables	All Input	$X_1, X_2, \ldots, X_8$: age, ed, …, othdebt
	Best Input	employ, address, debtinc, and creddebt
Number of neurons	1 – 25	
Activation Function	Logistic Sigmoid vs Tangent Hyperbolic	
Preprocessing Method	None, Standardized, Normalized, Adjusted Normalized	

1. What is the best **inputs** (**features selection**)?
2. How many **nodes** (**neurons**) in hidden layer?
3. What is the best **activation function** in **hidden** and **output** layer?
4. What is the best **pre-processing** method?

Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Input variables	All Input	$X_1, X_2, \ldots, X_8$: age, ed, …, othdebt
	Best Input	employ, address, debtinc, and creddebt

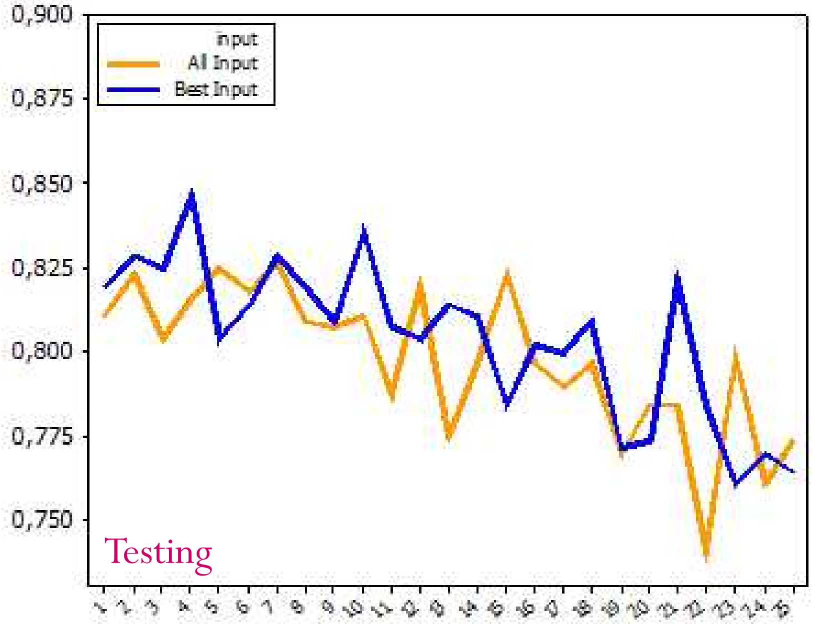
Stepwise Discriminant			Logistic Regression		
Variable	F	Wilk's Lambda	**Variable**	Wald	p-value
Debtinc	30,531	0,747	**Employ**	63,360	0,000
Employ	73,671	0,798	**Address**	15,621	0,000
Creddebt	43,584	0,762	**Debtinc**	20,129	0,000
Address	9,560	0,721	**Creddebt**	35,799	0,000

Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Percentage **correct** of classification

Number of **neurons** in hidden layer



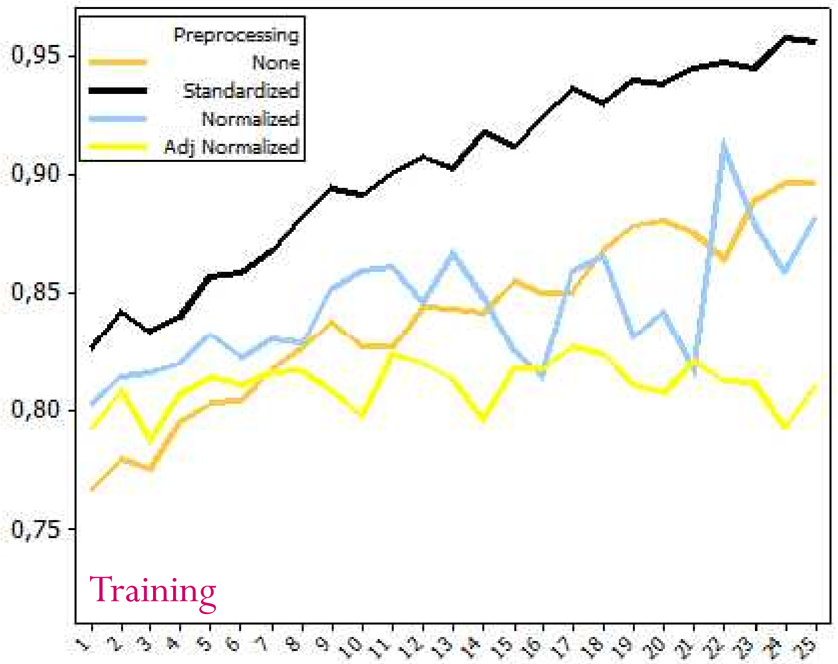
Number of **neurons** in hidden layer

➲ The effect of INPUTS and number of NEURONS in hidden layer

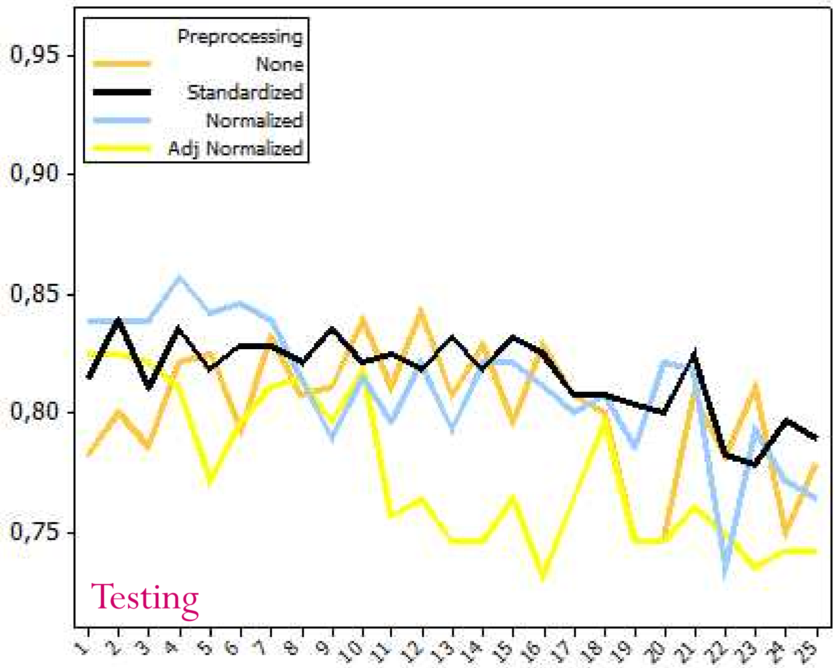
Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Percentage **correct** of classification



Number of **neurons** in hidden layer



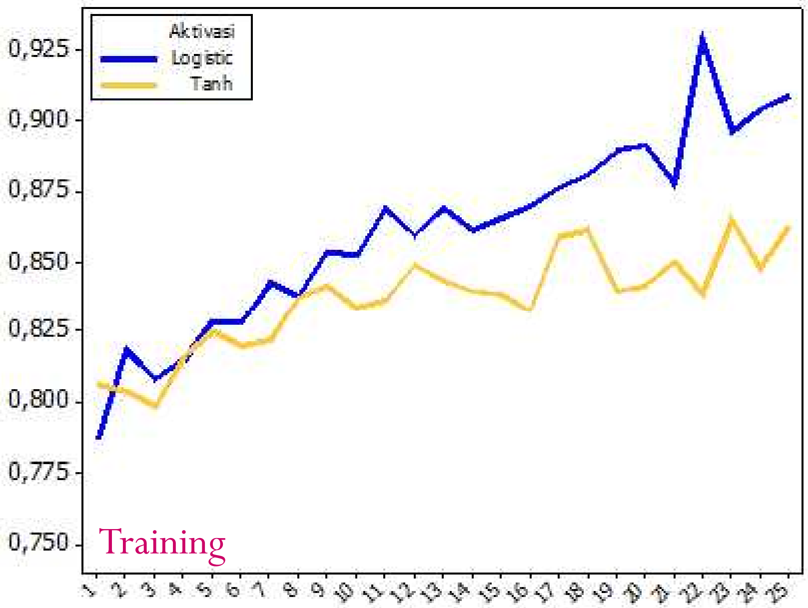
Number of **neurons** in hidden layer

➲ The effect of PREPROCESSING method & number of NEURONS

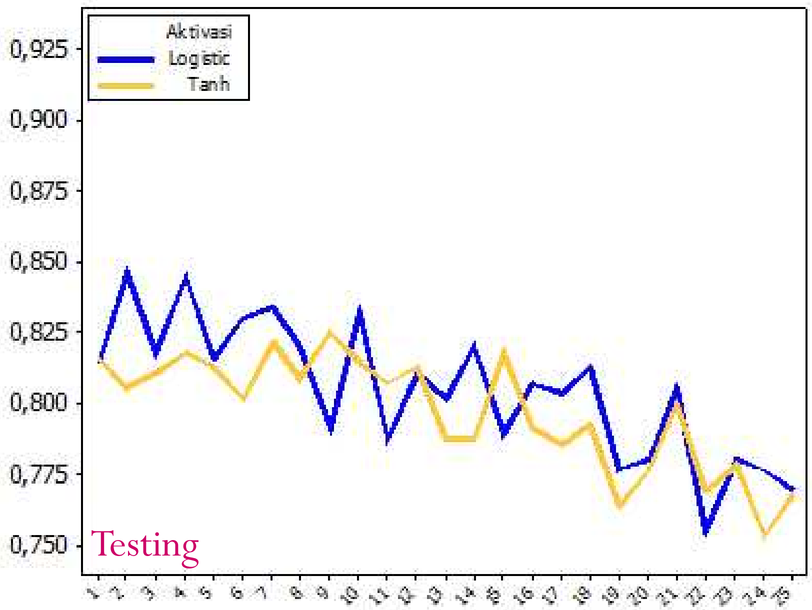
Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

Percentage **correct** of classification



Number of **neurons** in hidden layer



Number of **neurons** in hidden layer

➲ The effect of ACTIVATION function method & number of NEURONS

Application: NN for Classification

- Source: **bankloan.sav** from **SPSS**

☞ ***Summary*** of the results:

(1) More sophisticated or complex methods do not necessarily provide more accurate classification than simpler ones, particularly at testing dataset.

(2) The performance of the various NN methods for classification problem depends upon :

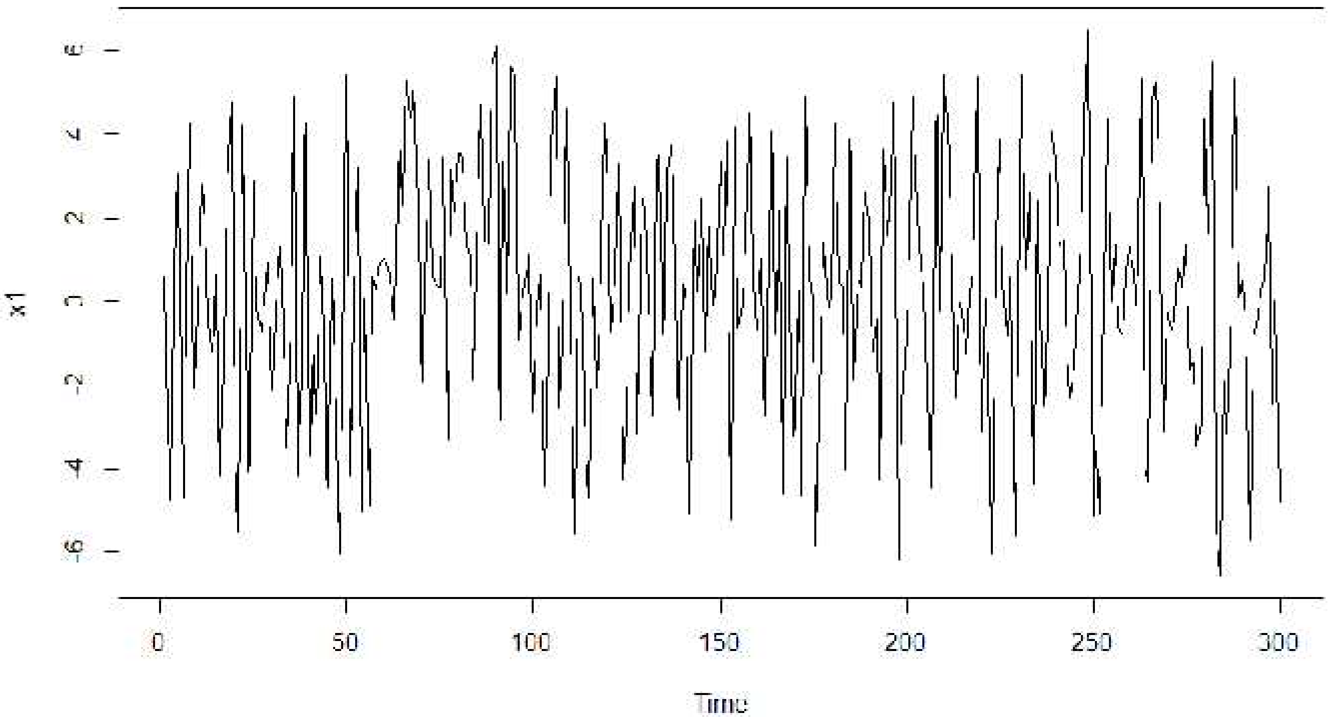
Inputs,
Number of **neurons** in hidden layer,
Pre-processing method, and
Activation function.

.

Application: NN for Time Series Forecasting

- Source: **simulation study** using ESTAR(1)[7] model

$$x_t = 6.5x_{t-7}\exp(-0.25x_{t-7}^2) + e_t \qquad , e_t \sim N(0, 0.5)$$



Application: NN for Time Series Forecasting

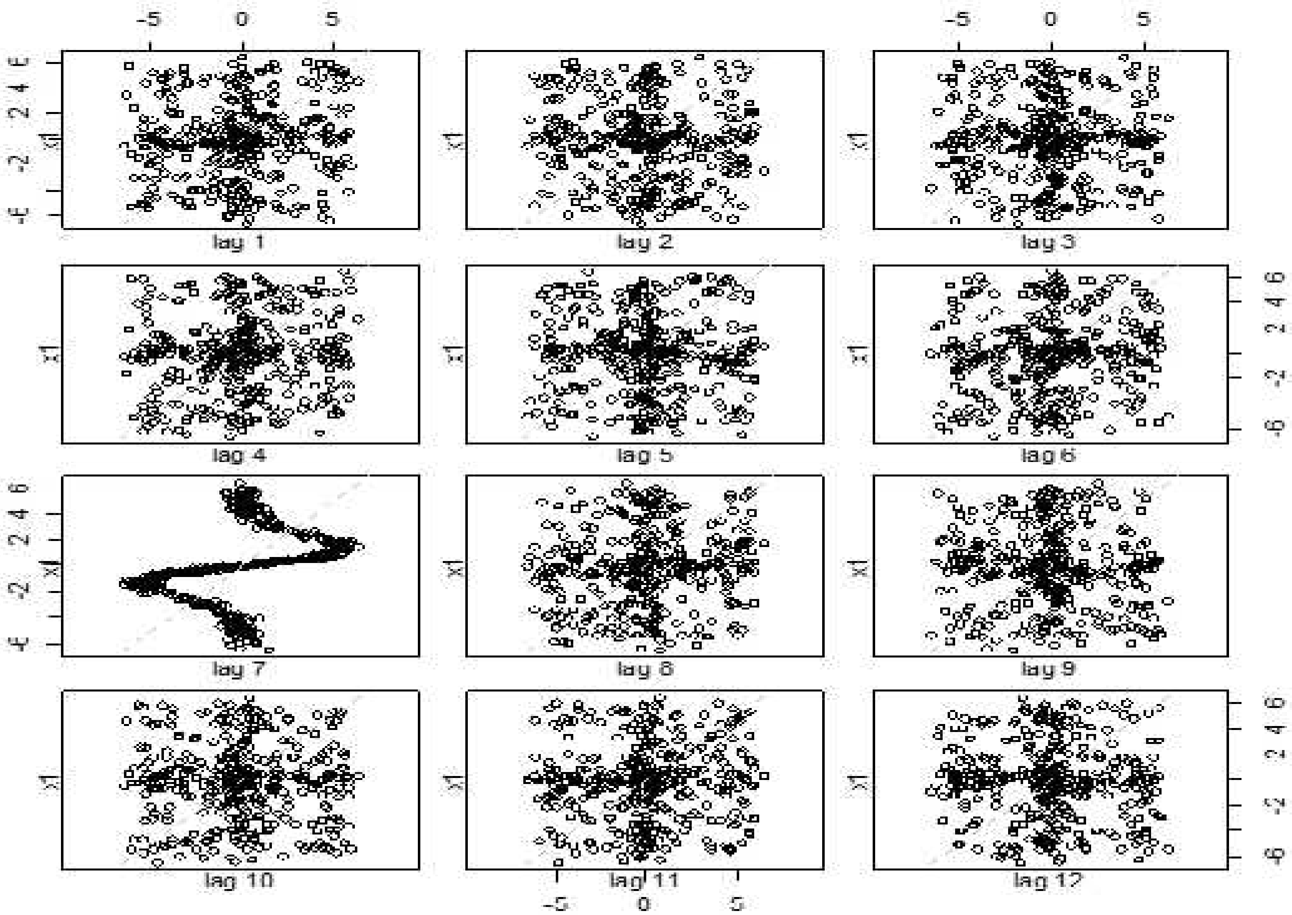
- Source: **simulation study** using ESTAR(1)[7] model

Input variables	Many Inputs	include lag 7 (X_{t-7}) and without lag 7
	Best Input	only using lag 7 or X_{t-7}
Number of neurons	1,2,3,4,5,10,15	
Activation Function	Logistic Sigmoid vs Tangent Hyperbolic	
Preprocessing Method	None, Standardized, Normalized, Adjusted Normalized	

1. What is the best **inputs** (**features selection**)?
2. How many **nodes** (**neurons**) in hidden layer?
3. What is the best **activation function** in **hidden** and **output** layer?
4. What is the best **pre-processing** method?

Application: NN for Time Series Forecasting

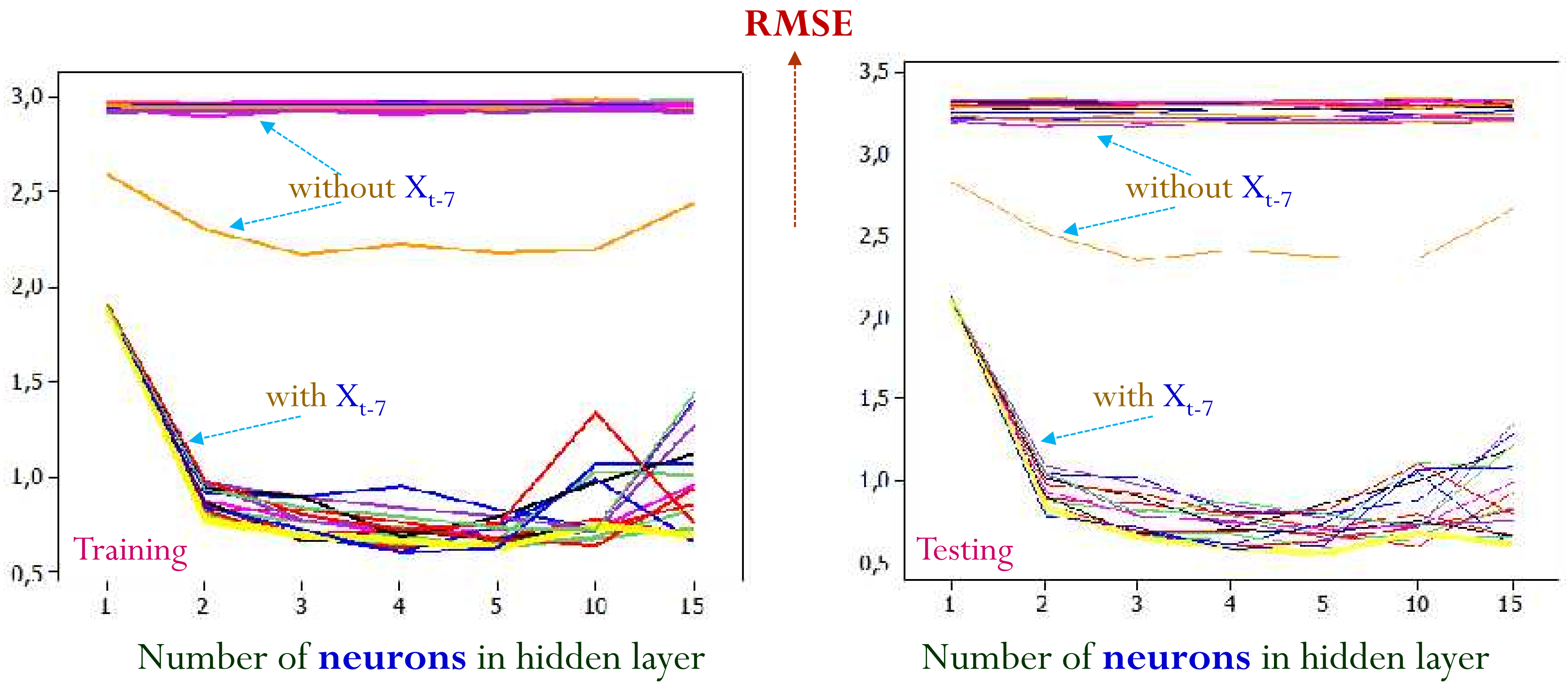
- Source: **simulation study** using ESTAR(1)[7] model



➲ Identification the appropriate lag inputs: use **LAG PLOT** in ***R***

Application: NN for Time Series Forecasting

- Source: **simulation study** using ESTAR(1)[7] model

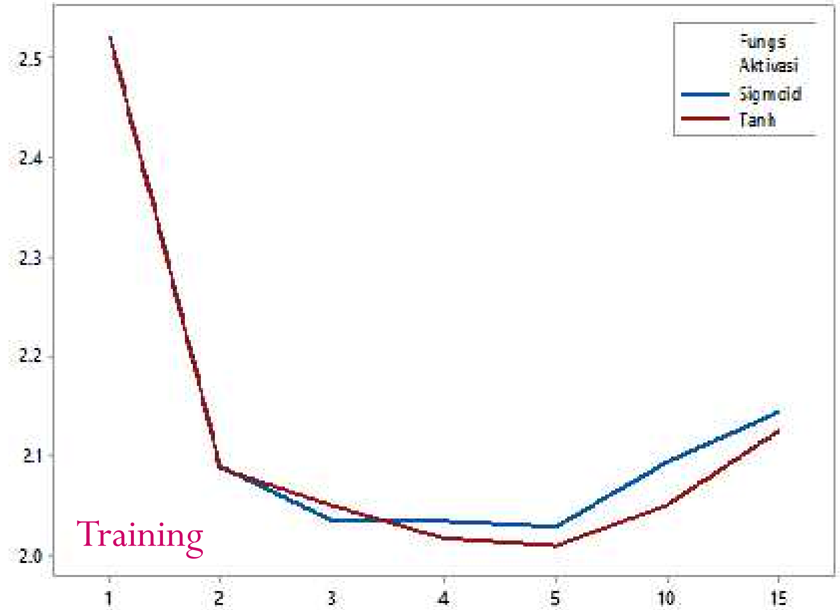


➲ The effect of INPUTS and number of NEURONS in hidden layer

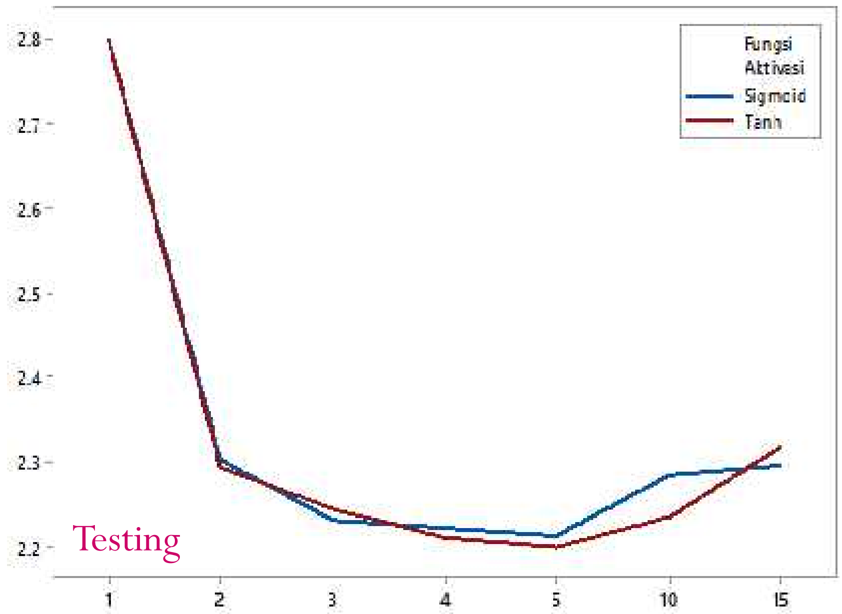
Application: NN for Time Series Forecasting

- Source: **simulation study** using ESTAR(1)[7] model

RMSE



Number of **neurons** in hidden layer



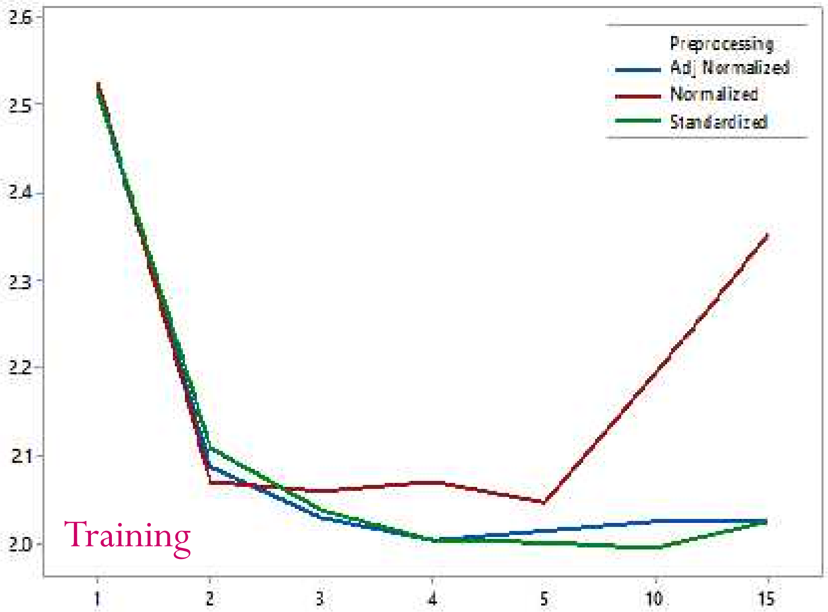
Number of **neurons** in hidden layer

➲ The effect of ACTIVATION function and number of NEURONS

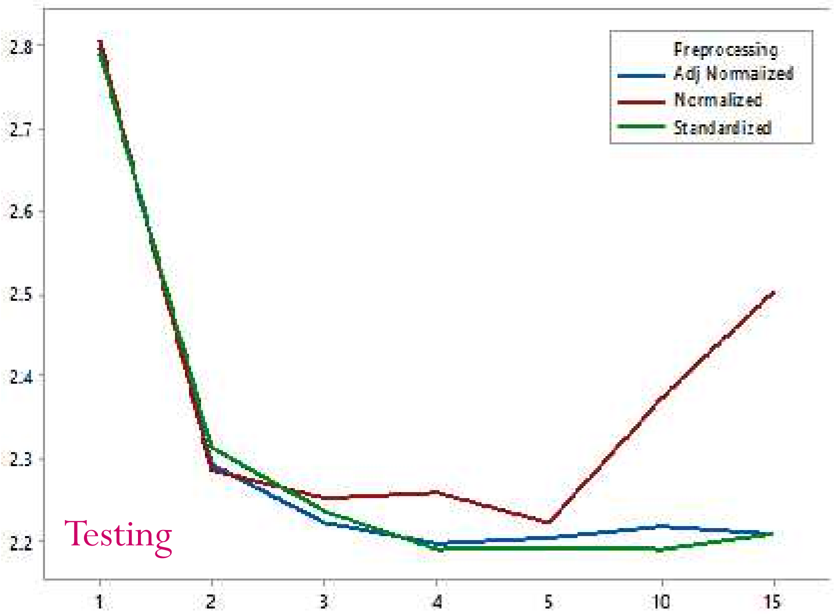
Application: NN for Time Series Forecasting

- Source: **simulation study** using ESTAR(1)[7] model

RMSE



Number of **neurons** in hidden layer



Number of **neurons** in hidden layer

➲ The effect of PREPROCESSING method & number of NEURONS

Application: NN for Time Series Forecasting

- Source: **simulation study** using ESTAR(1)[7] model

☞ ***Summary*** of the results:

(1) More sophisticated or complex methods do not necessarily provide more accurate forecast than simpler ones.

(2) The performance of the various NN methods for time series forecasting problem depends upon :

⇒ **Inputs** or **lag variables**,
Number of **neurons** in hidden layer,
Pre-processing method.

25 years of time series forecasting

De Gooijer & Hyndman (International Journal of Forecasting, 2006)

- 25 years of time series forecasting
  - Introduction
  - Exponential smoothing
    - Preamble
    - Variations
    - State space models
    - Method selection
    - Robustness
    - Prediction intervals
    - Parameter space and model properties
  - ARIMA models
    - Preamble
    - Univariate
    - Transfer function
    - Multivariate
  - Seasonality
  - State space and structural models and the Kalman filter

1980 → 2005

- Nonlinear models
  - Preamble
  - Regime-switching models
  - Functional-coefficient model
  - Neural nets
  - Deterministic versus stochastic dynamics
  - Miscellaneous
- Long memory models
- ARCH/GARCH models
- Count data forecasting
- Forecast evaluation and accuracy measures
- Combining
- Prediction intervals and densities
- A look to the future
- Acknowledgments
- References

Research Motivation

☞ **The M3-Competition**: results, conclusions and implications

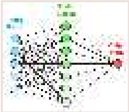
(1) Statistically sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones.

(2) The relative ranking of the performance of the various methods varies according to the accuracy measure being used.

(3) The accuracy when various methods are being combined outperforms, on average, the individual methods being combined and does very well in comparison to other methods.

(4) The accuracy of the various methods depends upon the length of the forecasting horizon involved.

Makridakis & Hibon (International Journal of Forecasting, 2000)



Recent development of NN for forecasting

☞ Hybrid Model – Combined – Ensemble

Model Selection in Neural Networks by Using Inference of R2 Incremental, PCA, and SIC Criteria for Time Series Forecasting [PDF] from ugm.ac.id

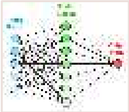
Authors	Suhartono, Subanar, S Guritno
Publication date	2006
Journal	JOURNAL OF QUANTITATIVE METHODS: Journal Devoted to The Mathematical and Statistical Application in Various Fields

New Procedures for Model Selection in Feedforward Neural Networks for Time Series Forecasting

Authors	Suhartono Suhartono
Publication date	2008/7/4

THE EFFECT OF DECOMPOSITION METHOD AS DATA PREPROCESSING ON NEURAL NETWORKS MODEL FOR FORECASTING TREND AND SEASONAL TIME SERIES [PDF] from petra.ac.id

Authors	Subanar Subanar, Suhartono Suhartono
Publication date	2007/2/1
Journal	Jurnal Teknik Industri



Recent development of NN for forecasting

☞ *Hybrid Model – Combined – Ensemble*

Two-level seasonal model based on hybrid ARIMA-ANFIS for forecasting short-term electricity load in Indonesia

4 Author(s) Suhartono ; Indah Puspitasari ; M. Sjahid Akbar ; Muhammad Hisyam Lee

Design of Experiment to Optimize the Architecture of Wavelet Neural Network for Forecasting the Tourist Arrivals in Indonesia

Authors	Bambang W Otok, Suhartono, Brodjol SS Ulama, Alfonsus J Endharta
Publication date	2011/11/14
Conference	International Conference on Informatics Engineering and Information Science

Seasonal Time Series Data Forecasting by Using Neural Networks Multiscale Autoregressive Model

Suhartono, B.S.S. Ulama and A.J. Endharta
Department of Statistics, Faculty of Mathematics and Natural Sciences,
Institute Technology Sepuluh Nopember, Surabaya 60111, Indonesia

Recent development of NN for forecasting

☞ Hybrid Model – Combined – Ensemble

Forecasting currency circulation data of Bank Indonesia by using hybrid ARIMAX-ANN model

I. Gede Surya Adi Prayoga, Suhartono, and Santi Puteri Rahayu

Citation: AIP Conference Proceedings **1842**, 030029 (2017); doi: 10.1063/1.4982867

Forecasting electricity load demand using hybrid exponential smoothing-artificial neural network model

Winita Sulandari, Subanar Subanar, Suhartono Suhartono, Herni Utami

Quality & Quantity

November 2015, Volume 49, Issue 6, pp 2633–2647

Artificial neural networks and fuzzy time series forecasting: an application to air quality

Authors | Authors and affiliations

Nur Haizum Abd Rahman, Muhammad Hisyam Lee, Suhartono, Mohd Talib Latif

Conclusion

- **Statistical models** and **NN** are not competing methodologies for data analysis. There is considerable many similarities between the two models.
- NN include several models, such as FFNN, that are useful for statistical applications.
- Statistical methodology is directly applicable to NN in a variety of ways, including estimation criteria, optimization algorithm, testing hypothesis and diagnostic check.

Main References

Neural Networks and Statistical Models

Proceedings of the Nineteenth Annual SAS Users Group International Conference, April, 1994

Warren S. Sarle, SAS Institute Inc., Cary, NC, USA

30 CONTRIBUTED RESEARCH ARTICLES

neuralnet: Training of Neural Networks

by Frauke Günther and Stefan Fritsch

provides functions to visualize the results or in gen-

Model Selection in Neural Networks

Ulrich Anders, Olaf Korn

Centre for European Economic Research (ZEW), Mannheim