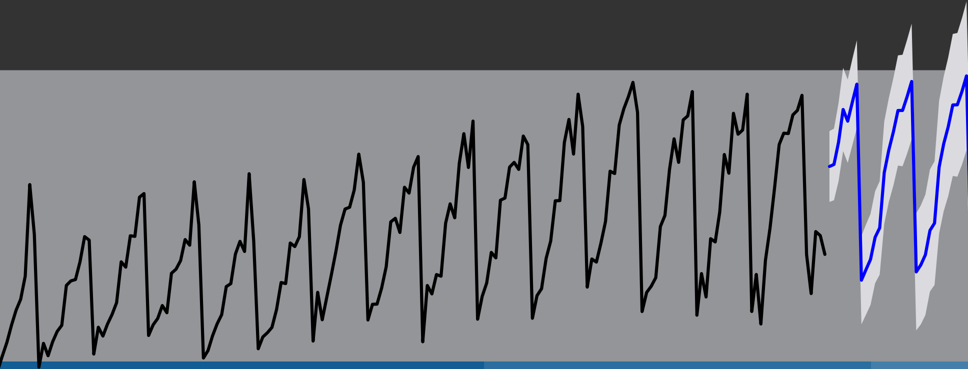
MONASH University

Rob J Hyndman

Automatic time series forecasting



Outline

Motivation

Australian Government

Department of Health and Ageing

Motivation

FOXTEL digital

Australian Government

Department of Health and Ageing

Motivation

FOXTEL digital

Incitec Pivot

Australian Government

Department of Health and Ageing

Motivation

FOXTEL digital

CAT

Incitec Pivot

Australian Government

Department of Health and Ageing

Motivation

FOXTEL digital

CAT

hp

Incitec Pivot

Australian Government

Department of Health and Ageing

Motivation

1. Common in business to have over 1000 products that need forecasting at least monthly.
2. Forecasts are often required by people who are untrained in time series analysis.

Specifications

Automatic forecasting algorithms must:

- ➥ determine an appropriate time series model;
- ➥ estimate the parameters;
- ➥ compute the forecasts with prediction intervals.

Motivation

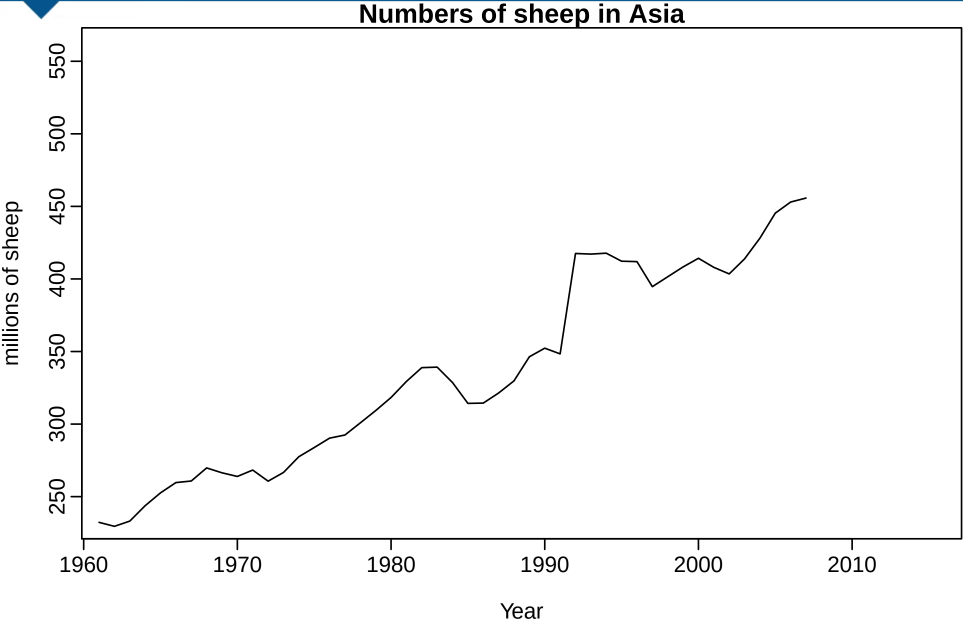
1. Common in business to have over 1000 products that need forecasting at least monthly.
2. Forecasts are often required by people who are untrained in time series analysis.

Specifications

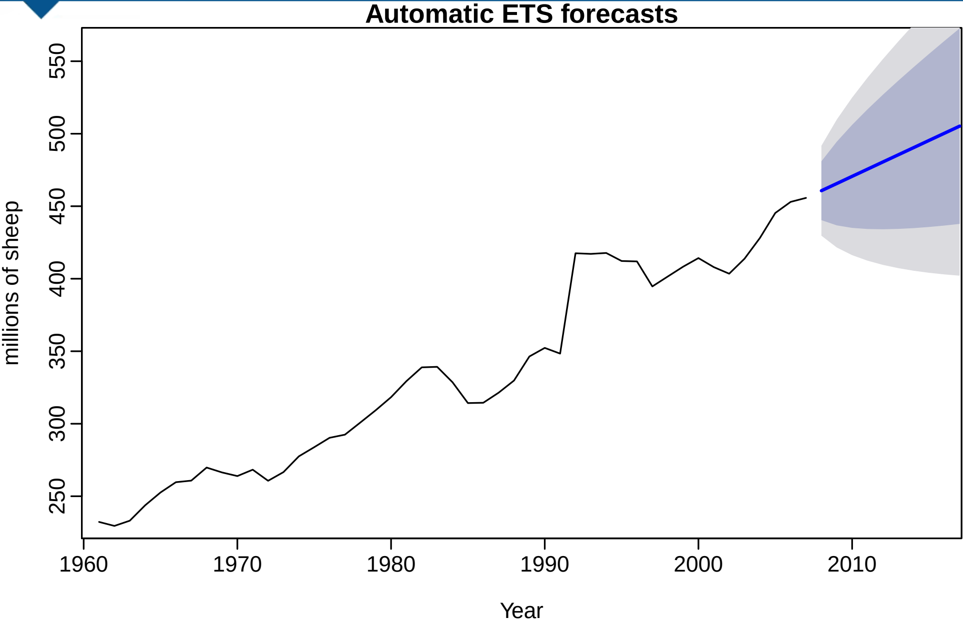
Automatic forecasting algorithms must:

- determine an appropriate time series model;
- estimate the parameters;
- compute the forecasts with prediction intervals.

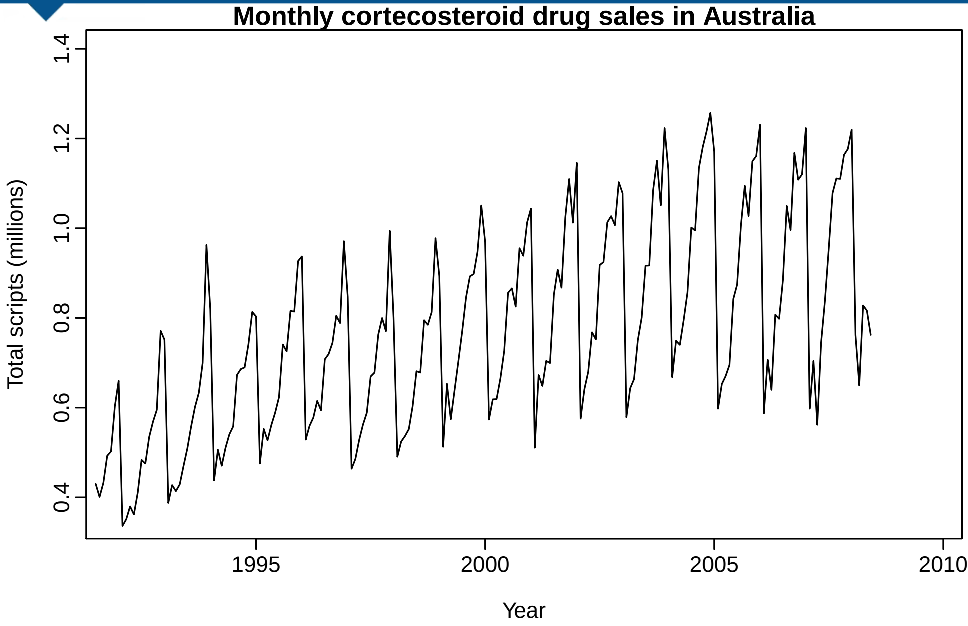
Example: Asian sheep



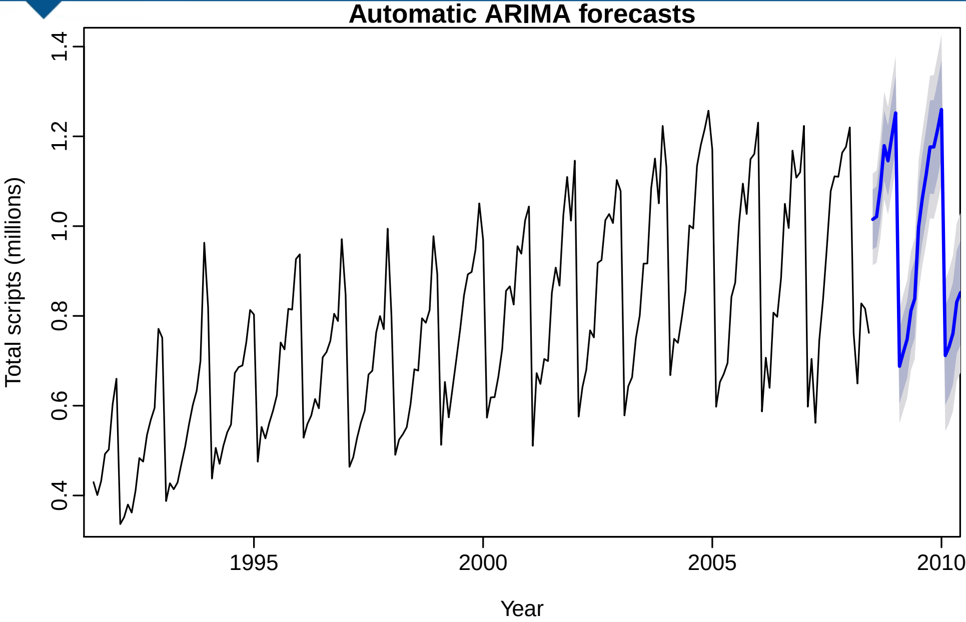
Example: Asian sheep



Example: Cortecosteroid sales



Example: Cortecosteroid sales



Outline

Makridakis and Hibon (1979)

J. R. Statist. Soc. A (1979),
142, *Part* 2, *pp*. 97–145

Accuracy of Forecasting: An Empirical Investigation

By Spyros Makridakis and Michèle Hibon

INSEAD—The European Institute of Business Administration

[Read before the Royal Statistical Society on Wednesday, December 13th, 1978, the President, Sir Claus Moser in the Chair]

Summary

In this study, the authors used 111 time series to examine the accuracy of various forecasting methods, particularly time-series methods. The study shows, at least for time series, why some methods achieve greater accuracy than others for different types of data. The authors offer some explanation of the seemingly conflicting conclusions of past empirical research on the accuracy of forecasting. One novel contribution of the paper is the development of regression equations expressing accuracy as a function of factors such as randomness, seasonality, trend-cycle and the number of data points describing the series. Surprisingly, the study shows that for these 111 series simpler methods perform well in comparison to the more complex and statistically sophisticated ARMA models.

Keywords: FORECASTING; TIME SERIES; FORECASTING ACCURACY

0. Introduction

The ultimate test of any forecast is whether or not it is capable of predicting future events

Makridakis and Hibon (1979)

J. R. Statist. Soc. A (1979),
142, *Part* 2, *pp.* 97–145

Accuracy of Forecasting: An Empirical Investigation

By SPYROS MAKRIDAKIS and MICHÈLE HIBON

INSEAD—The European Institute of Business Administration

ATISTICAL SOCIETY on Wednesday,
dent, SIR CLAUS MOSER in the Ch

SUMMARY

ed 111 time series to examine
arly time-series methods. The
s achieve greater accuracy than
me explanation of the seemingl
the accuracy of forecasting. O
of regression equations expres
mness, seasonality, trend-cycle
Surprisingly, the study shows
ll in comparison to the more

Keywords: FORECASTING; TIME SERIES; FORECASTING ACCURACY

0. INTRODUCTION

THE ultimate test of any forecast is whether or not it is capable of predicting future events

Makridakis and Hibon (1979)

This was the first large-scale empirical evaluation of time series forecasting methods.

- Highly controversial at the time.
- Difficulties:
  - ✗ How to measure forecast accuracy?
  - ✗ How to apply methods consistently and objectively?
  - ✗ How to explain unexpected results?
- Common thinking was that the more sophisticated mathematical models (ARIMA models at the time) were necessarily better.
- If results showed ARIMA models not best, it must be because analyst was unskilled.

Makridakis and Hibon (1979)

I do not believe that it is very fruitful to attempt to classify series according to which forecasting techniques perform “best”. The performance of any particular technique when applied to a particular series depends essentially on (a) the model which the series obeys; (b) our ability to identify and fit this model correctly and (c) the criterion chosen to measure the forecasting accuracy.

— *M.B. Priestley*

. . . the paper suggests the application of normal scientific experimental design to forecasting, with measures of unbiased testing of forecasts against subsequent reality, for success or failure. A long overdue reform.

— *F.H. Hansford-Miller*

Makridakis and Hibon (1979)

I do not believe that it is very fruitful to attempt to classify series according to which forecasting techniques perform “best”. The performance of any particular technique when applied to a particular series depends essentially on (a) the model which the series obeys; (b) our ability to identify and fit this model correctly and (c) the criterion chosen to measure the forecasting accuracy.

— M.B. Priestley

. . . the paper suggests the application of normal scientific experimental design to forecasting, with measures of unbiased testing of forecasts against subsequent reality, for success or failure. A long overdue reform.

— F.H. Hansford-Miller

Makridakis and Hibon (1979)

Modern man is fascinated with the subject of forecasting — *W.G. Gilchrist*

It is amazing to me, however, that after all this exercise in identifying models, transforming and so on, that the autoregressive moving averages come out so badly. I wonder whether it might be partly due to the authors not using the backwards forecasting approach to obtain the initial errors.

— *W.G. Gilchrist*

Makridakis and Hibon (1979)

Modern man is fascinated with the subject of forecasting — *W.G. Gilchrist*

It is amazing to me, however, that after all this exercise in identifying models, transforming and so on, that the autoregressive moving averages come out so badly. I wonder whether it might be partly due to the authors not using the backwards forecasting approach to obtain the initial errors.

— *W.G. Gilchrist*

Makridakis and Hibon (1979)

I find it hard to believe that Box-Jenkins, if properly applied, can actually be worse than so many of the simple methods — *C. Chatfield*

Why do empirical studies sometimes give different answers? It may depend on the selected sample of time series, but I suspect it is more likely to depend on the skill of the analyst and on their individual interpretations of what is meant by Method *X*.

— *C. Chatfield*

. . . these authors are more at home with simple procedures than with Box-Jenkins. — *C. Chatfield*

Makridakis and Hibon (1979)

I find it hard to believe that Box-Jenkins, if properly applied, can actually be worse than so many of the simple methods — *C. Chatfield*

Why do empirical studies sometimes give different answers? It may depend on the selected sample of time series, but I suspect it is more likely to depend on the skill of the analyst and on their individual interpretations of what is meant by Method *X*.

— *C. Chatfield*

… these authors are more at home with simple procedures than with Box-Jenkins. — *C. Chatfield*

Makridakis and Hibon (1979)

I find it hard to believe that Box-Jenkins, if properly applied, can actually be worse than so many of the simple methods — *C. Chatfield*

Why do empirical studies sometimes give different answers? It may depend on the selected sample of time series, but I suspect it is more likely to depend on the skill of the analyst and on their individual interpretations of what is meant by Method *X*.

— *C. Chatfield*

. . . these authors are more at home with simple procedures than with Box-Jenkins. — *C. Chatfield*

Makridakis and Hibon (1979)

> There is a fact that Professor Priestley must accept: empirical evidence is in *disagreement* with his theoretical arguments. — *S. Makridakis & M. Hibon*

> Dr Chatfield expresses some personal views about the first author . . . It might be useful for Dr Chatfield to read some of the psychological literature quoted in the main paper, and he can then learn a little more about biases and how they affect prior probabilities. — *S. Makridakis & M. Hibon*

Makridakis and Hibon (1979)

There is a fact that Professor Priestley must accept: empirical evidence is in *disagreement* with his theoretical arguments. — *S. Makridakis & M. Hibon*

Dr Chatfield expresses some personal views about the first author . . . It might be useful for Dr Chatfield to read some of the psychological literature quoted in the main paper, and he can then learn a little more about biases and how they affect prior probabilities. — *S. Makridakis & M. Hibon*

Consequences of M&H (1979)

As a result of this paper, researchers started to:

- ➥ consider how to automate forecasting methods;
- ➥ study what methods give the best forecasts;
- ➥ be aware of the dangers of over-fitting;
- ➥ treat forecasting as a different problem from time series analysis.

Makridakis & Hibon followed up with a new competition in 1982:

- 1001 series
- Anyone could submit forecasts (avoiding the charge of incompetence)
- Multiple forecast measures used.

Consequences of M&H (1979)

As a result of this paper, researchers started to:

- ➥ consider how to automate forecasting methods;
- ➥ study what methods give the best forecasts;
- ➥ be aware of the dangers of over-fitting;
- ➥ treat forecasting as a different problem from time series analysis.

Makridakis & Hibon followed up with a new competition in 1982:

- 1001 series
- Anyone could submit forecasts (avoiding the charge of incompetence)
- Multiple forecast measures used.

M-competition

Journal of Forecasting, *Vol. 1*, *111–153 (1982)*

The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition

S. MAKRIDAKIS
INSEAD, Fontainebleau, France

A. ANDERSEN
University of Sydney, Australia

R. CARBONE
Université Laval, Quebec, Canada

R. FILDES
Manchester Business School, Manchester, England

M. HIBON
INSEAD, Fontainebleau, France

R. LEWANDOWSKI
Marketing Systems, Essen, Germany

J. NEWTON
E. PARZEN
Texas A & M University, Texas, U.S.A.

R. WINKLER
Indiana University, Bloomington, U.S.A.

ABSTRACT

In the last few decades many methods have become available for forecasting. As always, when alternatives exist, choices need to be made so that an appropriate forecasting method can be selected and used for the specific situation being considered. This paper reports the results of a forecasting competition that provides information to facilitate such choice. Seven experts in each of the 24 methods forecasted up to 1001 series for six up to eighteen time horizons. The results of the competition are presented in this paper whose purpose is to provide empirical evidence about *differences* found to exist among the various extrapolative (time series) methods used in the competition.

M-competition

Main findings (taken from Makridakis & Hibon, 2000)

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.

M3 competition

ELSEVIER

International Journal of Forecasting 16 (2000) 451–476

international journal of forecasting

www.elsevier.com/locate/ijforecast

The M3-Competition: results, conclusions and implications

Spyros Makridakis, Michèle Hibon*

INSEAD, Boulevard de Constance, 77305 Fontainebleau, France

Abstract

This paper describes the M3-Competition, the latest of the M-Competitions. It explains the reasons for conducting the competition and summarizes its results and conclusions. In addition, the paper compares such results/conclusions with those of the previous two M-Competitions as well as with those of other major empirical studies. Finally, the implications of these results and conclusions are considered, their consequences for both the theory and practice of forecasting are explored and directions for future research are contemplated.

Keywords: Comparative methods — time series: univariate; Forecasting competitions; M-Competition; Forecasting methods, Forecasting accuracy

Makridakis and Hibon (2000)

"The M3-Competition is a final attempt by the authors to settle the accuracy issue of various time series methods. . . The extension involves the inclusion of more methods/ researchers (in particular in the areas of neural networks and expert systems) and more series."

- 3003 series
- All data from business, demography, finance and economics.
- Series length between 14 and 126.
- Either non-seasonal, monthly or quarterly.
- All time series positive.
- M&H claimed that the M3-competition supported the findings of their earlier work.
- However, best performing methods far from "simple".

Makridakis and Hibon (2000)

Best methods:

Theta

- A very confusing explanation.
- Shown by Hyndman and Billah (2003) to be average of linear regression and simple exponential smoothing with drift, applied to seasonally adjusted data.
- Later, the original authors claimed that their explanation was incorrect.

Forecast Pro

- A commercial software package with an unknown algorithm.
- Known to fit either exponential smoothing or ARIMA models using BIC.

M3 results (recalculated)

Method	MAPE	sMAPE	MASE
Theta	17.42	12.76	1.39
ForecastPro	18.00	13.06	1.47
ForecastX	17.35	13.09	1.42
Automatic ANN	17.18	13.98	1.53
B-J automatic	19.13	13.72	1.54

M3 results (recalculated)

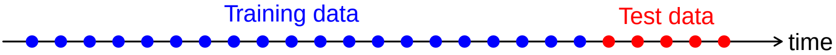
Method	MAPE	sMAPE	MASE
Theta	17.42	12.76	1.39
ForecastPro	18.00	13.06	1.47
ForecastX	17.35	13.09	1.42

- Calculations do not match published paper.
- Some contestants apparently submitted multiple entries but only best ones published.

Outline

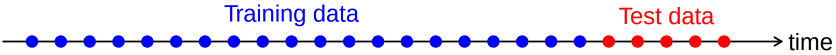
Cross-validation

Traditional evaluation

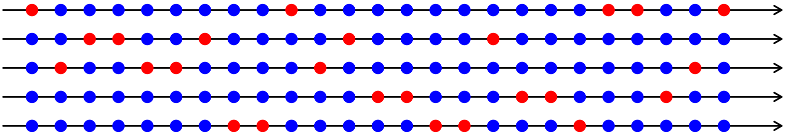


Cross-validation

Traditional evaluation

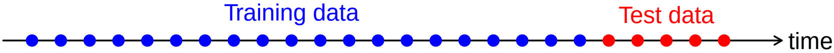


Standard cross-validation

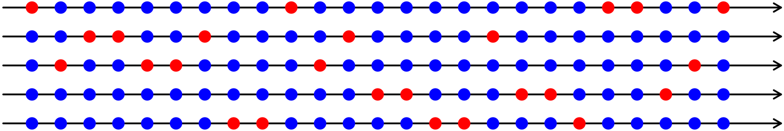


Cross-validation

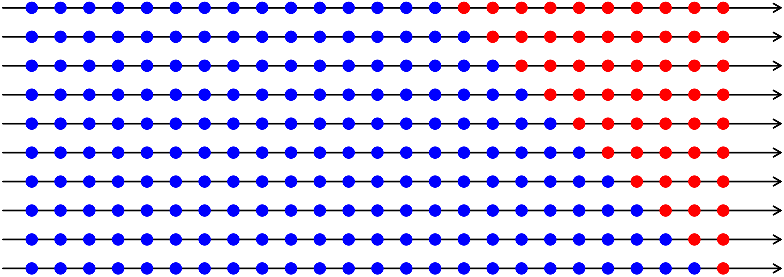
Traditional evaluation



Standard cross-validation

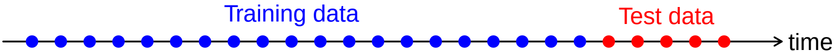


Time series cross-validation

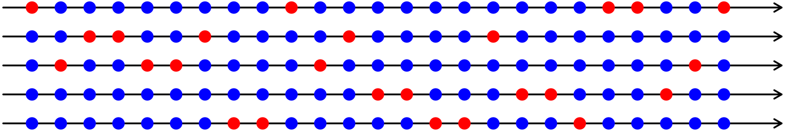


Cross-validation

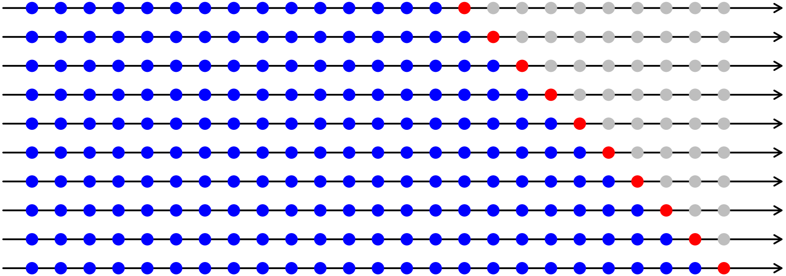
Traditional evaluation



Standard cross-validation

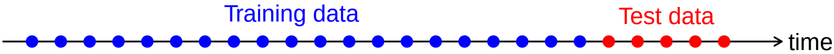


Time series cross-validation

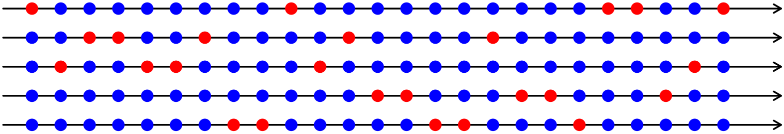


Cross-validation

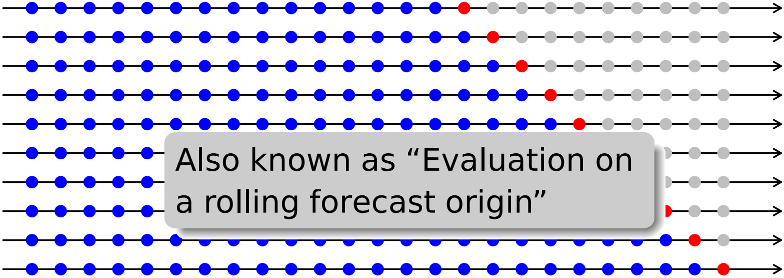
Traditional evaluation



Standard cross-validation



Time series cross-validation



Also known as “Evaluation on a rolling forecast origin”

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

where L is the likelihood and k is the number of estimated parameters in the model.

- This is a *penalized likelihood* approach.
- If L is Gaussian, then $\text{AIC} \approx c + T \log \text{MSE} + 2k$ where c is a constant, MSE is from one-step forecasts on **training set**, and T is the length of the series.

Minimizing the Gaussian AIC is asymptotically equivalent (as $T \to \infty$) to minimizing MSE from one-step forecasts on **test set** via time series cross-validation.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

where L is the likelihood and k is the number of estimated parameters in the model.

- This is a *penalized likelihood* approach.
- If L is Gaussian, then $\text{AIC} \approx c + T\log\text{MSE} + 2k$ where c is a constant, MSE is from one-step forecasts on **training set**, and T is the length of the series.

Minimizing the Gaussian AIC is asymptotically equivalent (as $T \to \infty$) to minimizing MSE from one-step forecasts on **test set** via time series cross-validation.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

where L is the likelihood and k is the number of estimated parameters in the model.

- This is a *penalized likelihood* approach.
- If L is Gaussian, then $\text{AIC} \approx c + T \log \text{MSE} + 2k$ where c is a constant, MSE is from one-step forecasts on **training set**, and T is the length of the series.

Minimizing the Gaussian AIC is asymptotically equivalent (as $T \to \infty$) to minimizing MSE from one-step forecasts on **test set** via time series cross-validation.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

where L is the likelihood and k is the number of estimated parameters in the model.

- This is a *penalized likelihood* approach.
- If L is Gaussian, then $\text{AIC} \approx c + T\log \text{MSE} + 2k$ where c is a constant, MSE is from one-step forecasts on **training set**, and T is the length of the series.

Minimizing the Gaussian AIC is asymptotically equivalent (as $T \to \infty$) to minimizing MSE from one-step forecasts on **test set** via time series cross-validation.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

where L is the likelihood and k is the number of estimated parameters in the model.

- This is a *penalized likelihood* approach.
- If L is Gaussian, then $\text{AIC} \approx c + T \log \text{MSE} + 2k$ where c is a constant, MSE is from one-step forecasts on **training set**, and T is the length of the series.

Minimizing the Gaussian AIC is asymptotically equivalent (as $T \to \infty$) to minimizing MSE from one-step forecasts on **test set** via time series cross-validation.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

Corrected AIC

For small T, AIC tends to over-fit. Bias-corrected version:

$$\text{AIC}_{\text{C}} = \text{AIC} + \frac{2(k+1)(k+2)}{T-k}$$

Bayesian Information Criterion

$$\text{BIC} = \text{AIC} + k[\log(T) - 2]$$

- BIC penalizes terms more heavily than AIC
- Minimizing BIC is consistent **if there is a true model**.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

Corrected AIC

For small T, AIC tends to over-fit. Bias-corrected version:

$$\text{AIC}_\text{C} = \text{AIC} + \frac{2(k+1)(k+2)}{T-k}$$

Bayesian Information Criterion

$$\text{BIC} = \text{AIC} + k[\log(T) - 2]$$

- BIC penalizes terms more heavily than AIC
- Minimizing BIC is consistent **if there is a true model**.

Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k$$

Corrected AIC

For small T, AIC tends to over-fit. Bias-corrected version:

$$\text{AIC}_{\text{C}} = \text{AIC} + \frac{2(k+1)(k+2)}{T-k}$$

Bayesian Information Criterion

$$\text{BIC} = \text{AIC} + k[\log(T) - 2]$$

- BIC penalizes terms more heavily than AIC
- Minimizing BIC is consistent **if there is a true model**.

What to use?

Choice: AIC, AICc, BIC, CV-MSE

- CV-MSE too time consuming for most automatic forecasting purposes. Also requires large T.
- As $T \to \infty$, BIC selects *true* model if there is one. But that is never true!
- AICc focuses on forecasting performance, can be used on small samples and is very fast to compute.
- Empirical studies in forecasting show AIC is better than BIC for forecast accuracy.

What to use?

Choice: AIC, AICc, BIC, CV-MSE

- CV-MSE too time consuming for most automatic forecasting purposes. Also requires large T.
- As $T \to \infty$, BIC selects *true* model if there is one. But that is never true!
- AICc focuses on forecasting performance, can be used on small samples and is very fast to compute.
- Empirical studies in forecasting show AIC is better than BIC for forecast accuracy.

What to use?

Choice: AIC, AICc, BIC, CV-MSE

- CV-MSE too time consuming for most automatic forecasting purposes. Also requires large T.
- As $T \to \infty$, BIC selects *true* model if there is one. But that is never true!
- AICc focuses on forecasting performance, can be used on small samples and is very fast to compute.
- Empirical studies in forecasting show AIC is better than BIC for forecast accuracy.

What to use?

Choice: AIC, AICc, BIC, CV-MSE

- CV-MSE too time consuming for most automatic forecasting purposes. Also requires large T.
- As $T \to \infty$, BIC selects *true* model if there is one. But that is never true!
- AICc focuses on forecasting performance, can be used on small samples and is very fast to compute.
- Empirical studies in forecasting show AIC is better than BIC for forecast accuracy.

What to use?

Choice: AIC, AICc, BIC, CV-MSE

- CV-MSE too time consuming for most automatic forecasting purposes. Also requires large T.
- As $T \to \infty$, BIC selects *true* model if there is one. But that is never true!
- AICc focuses on forecasting performance, can be used on small samples and is very fast to compute.
- Empirical studies in forecasting show AIC is better than BIC for forecast accuracy.

Outline

Exponential smoothing

https://www.otexts.org/fpp/7

OTexts
ONLINE, OPEN-ACCESS TEXTBOOKS

Home Books Authors About Donation

Search

7 Exponential smoothing

Exponential smoothing was proposed in the late 1950s (Brown 1959, Holt 1957 and Winters 1960 are key pioneering works) and has motivated some of the most successful forecasting methods. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older. In other words, the more recent the observation the higher the associated weight. This framework generates reliable forecasts quickly and for a wide spectrum of time series which is a great advantage and of major importance to applications in industry.

This chapter is divided into two parts. In the first part we present in detail the mechanics of all exponential smoothing methods and their application in forecasting time series with various characteristics. This is key in understanding the intuition behind these methods. In this setting, selecting and using a forecasting method may appear to be somewhat ad-hoc. The

Book information

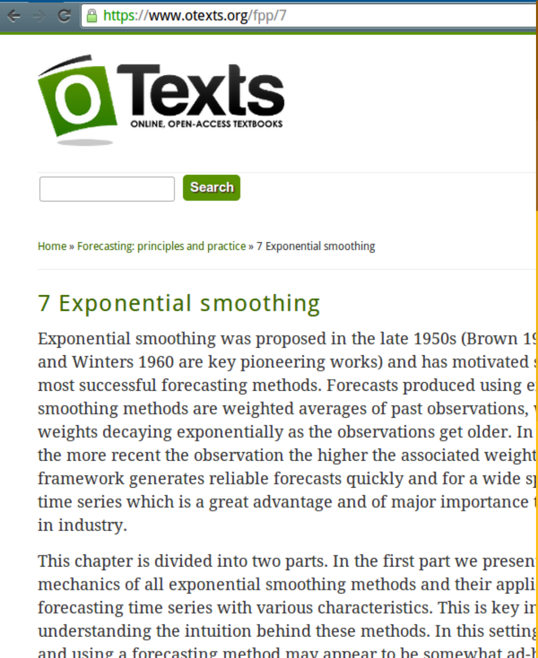
About this book

Feedback on this book

Rob J Hyndman
George Athanasopoulos

Forecasting: principles and practice

Exponential smoothing


https://www.otexts.org/fpp/7

OTexts ONLINE, OPEN-ACCESS TEXTBOOKS

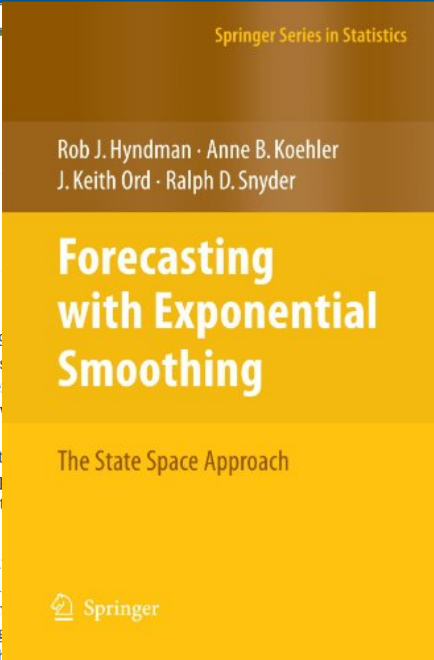
Search

Home » Forecasting: principles and practice » 7 Exponential smoothing

7 Exponential smoothing

Exponential smoothing was proposed in the late 1950s (Brown 1
and Winters 1960 are key pioneering works) and has motivated
most successful forecasting methods. Forecasts produced using e
smoothing methods are weighted averages of past observations,
weights decaying exponentially as the observations get older. In
the more recent the observation the higher the associated weight
framework generates reliable forecasts quickly and for a wide s
time series which is a great advantage and of major importance
in industry.

This chapter is divided into two parts. In the first part we presen
mechanics of all exponential smoothing methods and their appli
forecasting time series with various characteristics. This is key i
understanding the intuition behind these methods. In this setting
and using a forecasting method may appear to be somewhat ad-h



Exponential smoothing methods

		Seasonal Component		
Trend Component		N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing

A,N: Holt's linear method

Exponential smoothing methods

		Seasonal Component		
Trend Component		N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing
A,N: Holt's linear method
A_d,N: Additive damped trend method

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing
A,N: Holt's linear method
A_d,N: Additive damped trend method
M,N: Exponential trend method

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing
A,N: Holt's linear method
A_d,N: Additive damped trend method
M,N: Exponential trend method
M_d,N: Multiplicative damped trend method

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing
A,N: Holt's linear method
A_d,N: Additive damped trend method
M,N: Exponential trend method
M_d,N: Multiplicative damped trend method
A,A: Additive Holt-Winters' method

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

N,N: Simple exponential smoothing
A,N: Holt's linear method
A_d,N: Additive damped trend method
M,N: Exponential trend method
M_d,N: Multiplicative damped trend method
A,A: Additive Holt-Winters' method
A,M: Multiplicative Holt-Winters' method

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

- There are 15 separate exponential smoothing methods.

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

- There are 15 separate exponential smoothing methods.
- Each can have an additive or multiplicative error, giving 30 separate models.

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

- There are 15 separate exponential smoothing methods.
- Each can have an additive or multiplicative error, giving 30 separate models.
- Only 19 models are numerically stable.

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

Examples:

A,N,N: Simple exponential smoothing with additive errors

A,A,N: Holt's linear method with additive errors

M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	Seasonal Component A (Additive)	Seasonal Component M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

↑

Trend

Examples:

A,N,N: Simple exponential smoothing with additive errors

A,A,N: Holt's linear method with additive errors

M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Exponential smoothing methods

		Seasonal Component		
Trend Component		N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

$\uparrow$ $\nwarrow$

Trend **S**easonal

Examples:

A,N,N: Simple exponential smoothing with additive errors

A,A,N: Holt's linear method with additive errors

M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

Error **T**rend **S**easonal

Examples:

A,N,N: Simple exponential smoothing with additive errors

A,A,N: Holt's linear method with additive errors

M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A_d	(Additive damped)	A_d,N	A_d,A	A_d,M
M	(Multiplicative)	M,N	M,A	M,M
M_d	(Multiplicative damped)	M_d,N	M_d,A	M_d,M

General notation E T S : **E**xponen**T**ial **S**moothing

Error **T**rend **S**easonal

Examples:

A,N,N: Simple exponential smoothing with additive errors

A,A,N: Holt's linear method with additive errors

M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Exponential smoothing methods

Innovations state space models

- All ETS models can be written in innovations state space form (IJF, 2002).
- Additive and multiplicative versions give the same point forecasts but different prediction intervals.

General notation E T S : **E**xponen**T**ial **S**moothing

Error **T**rend **S**easonal

Examples:

A,N,N: Simple exponential smoothing with additive errors
A,A,N: Holt's linear method with additive errors
M,A,M: Multiplicative Holt-Winters' method with multiplicative errors

Innovations state space models

Let $\boldsymbol{x}_t = (\ell_t, b_t, s_t, s_{t-1}, \dots, s_{t-m+1})$ and $\varepsilon_t \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma^2)$.

$$y_t = \underbrace{h(\boldsymbol{x}_{t-1})}_{\mu_t} + \underbrace{k(\boldsymbol{x}_{t-1})\varepsilon_t}_{e_t} \qquad \text{Observation equation}$$

$$\boldsymbol{x}_t = f(\boldsymbol{x}_{t-1}) + g(\boldsymbol{x}_{t-1})\varepsilon_t \qquad \text{State equation}$$

Additive errors:

$$k(\boldsymbol{x}_{t-1}) = 1. \qquad y_t = \mu_t + \varepsilon_t.$$

Multiplicative errors:

$$k(\boldsymbol{x}_{t-1}) = \mu_t. \qquad y_t = \mu_t(1 + \varepsilon_t).$$

$\varepsilon_t = (y_t - \mu_t)/\mu_t$ is relative error.

Innovations state space models

- All models can be written in state space form.
- Additive and multiplicative versions give same point forecasts but different prediction intervals.

Estimation

$$L^*(\theta, \boldsymbol{x}_0) = n \log\left(\sum_{t=1}^{n} \varepsilon_t^2 / k^2(\boldsymbol{x}_{t-1})\right) + 2\sum_{t=1}^{n} \log|k(\boldsymbol{x}_{t-1})|$$

$$= -2\log(\text{Likelihood}) + \text{constant}$$

Innovations state space models

- All models can be written in state space form.
- Additive and multiplicative versions give same point forecasts but different prediction intervals.

Estimation

$$
\begin{aligned}
L^*(\theta, \boldsymbol{x}_0) &= n \log\left(\sum_{t=1}^{n} \varepsilon_t^2 / k^2(\boldsymbol{x}_{t-1})\right) + 2\sum_{t=1}^{n} \log|k(\boldsymbol{x}_{t-1})| \\
&= -2\log(\text{Likelihood}) + \text{constant}
\end{aligned}
$$

- Minimize wrt $\theta = (\alpha, \beta, \gamma, \phi)$ and initial states $\boldsymbol{x}_0 = (\ell_0, b_0, s_0, s_{-1}, \dots, s_{-m+1})$.

Innovations state space models

- All models can be written in state space form.
- Additive and multiplicative versions give same point forecasts but different prediction intervals.

Estimation

$$L^*(\theta, \boldsymbol{x}_0) = n \log\left(\sum_{t=1}^{n} \varepsilon_t^2 / k^2(\boldsymbol{x}_{t-1})\right) + 2\sum_{t=1}^{n} \log|k(\boldsymbol{x}_{t-1})|$$
$$= -2\log(\text{Likelihood}) + \text{constant}$$

- Minimize wrt $\theta = (\alpha, \beta, \gamma, \phi)$ and initial states $\boldsymbol{x}_0 = (\ell_0, b_0, s_0, s_{-1}, \dots, s_{-m+1})$.

Innovations state space models

- All models can be written in state space form.
- Additive and multiplicative versions give same point forecasts but different prediction intervals.

Estimation

$$L^*(\boldsymbol{\theta},\boldsymbol{x}_0) = n\log\left(\sum_{t=1}^{n}\varepsilon_t^2/k^2(\boldsymbol{x}_{t-1})\right) + 2\sum_{t=1}^{n}\log|k(\boldsymbol{x}_{t-1})|$$
$$= -2\log(\text{Likelihood}) + \text{constant}$$

- Minimize wrt $\boldsymbol{\theta} = (\alpha,\beta,\gamma,\phi)$ and initial states $\boldsymbol{x}_0 = (\ell_0, b_0, s_0, s_{-1},\dots,s_{-m+1})$.

Innovations state space models

- All models can be written in state space form.
- Additive and multiplicative versions give same point forecasts but different prediction intervals.

Estimation

$$L^*(\boldsymbol{\theta}, \boldsymbol{x}_0) = n \log\left(\sum_{t=1}^{n} \varepsilon_t^2 / k^2(\boldsymbol{x}_{t-1}) \right) + 2 \sum_{t=1}^{n} \log |k(\boldsymbol{x}_{t-1})|$$
$$= -2 \log(\text{Likelihood}) + \text{constant}$$

- Minimize wrt $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \phi)$ and initial states $\boldsymbol{x}_0 = (\ell_0, b_0, s_0, s_{-1}, \dots, s_{-m+1})$.

ets algorithm in R

Based on Hyndman & Khandakar (IJF 2008):

- Apply each of 19 models that are appropriate to the data. Optimize parameters and initial values using MLE.
- Select best method using AICc.
- Produce forecasts using best method.
- Obtain prediction intervals using underlying state space model.

ets algorithm in R

Based on Hyndman & Khandakar (IJF 2008):

- Apply each of 19 models that are appropriate to the data. Optimize parameters and initial values using MLE.
- Select best method using AICc.
- Produce forecasts using best method.
- Obtain prediction intervals using underlying state space model.

ets algorithm in R

Based on Hyndman & Khandakar (IJF 2008):

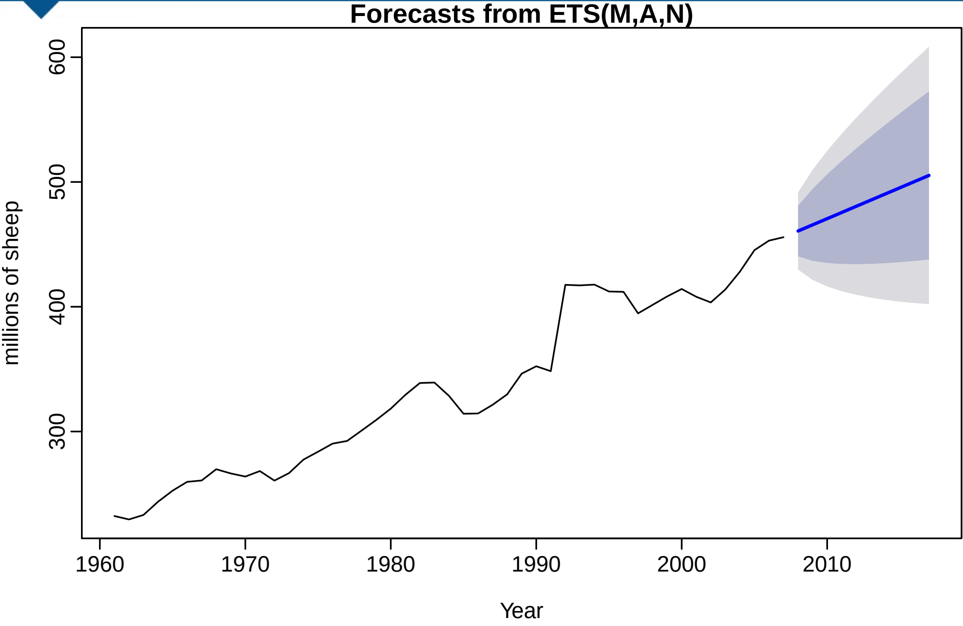
- Apply each of 19 models that are appropriate to the data. Optimize parameters and initial values using MLE.
- Select best method using AICc.
- Produce forecasts using best method.
- Obtain prediction intervals using underlying state space model.

ets algorithm in R

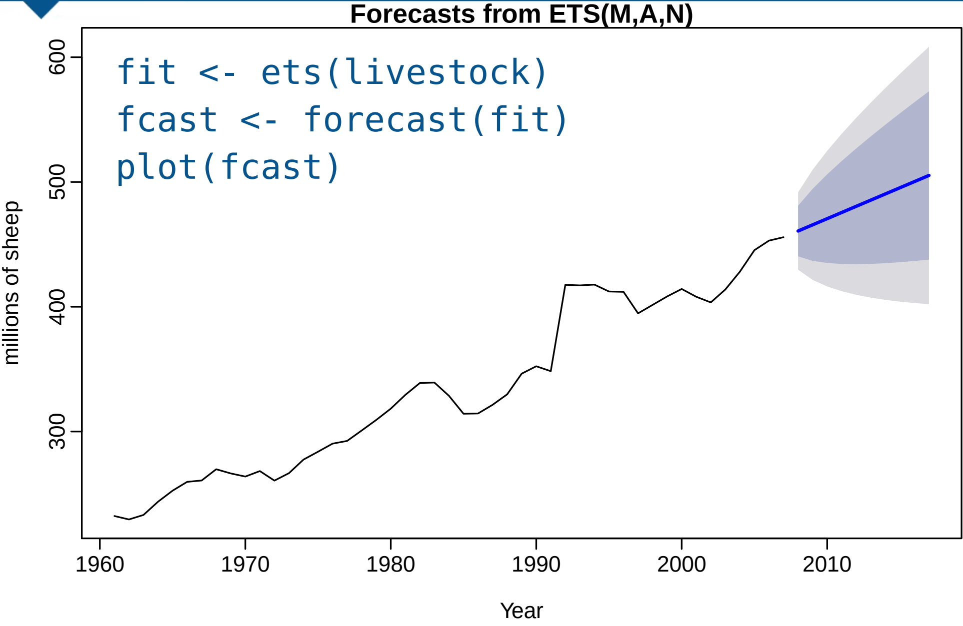
Based on Hyndman & Khandakar (IJF 2008):

- Apply each of 19 models that are appropriate to the data. Optimize parameters and initial values using MLE.
- Select best method using AICc.
- Produce forecasts using best method.
- Obtain prediction intervals using underlying state space model.

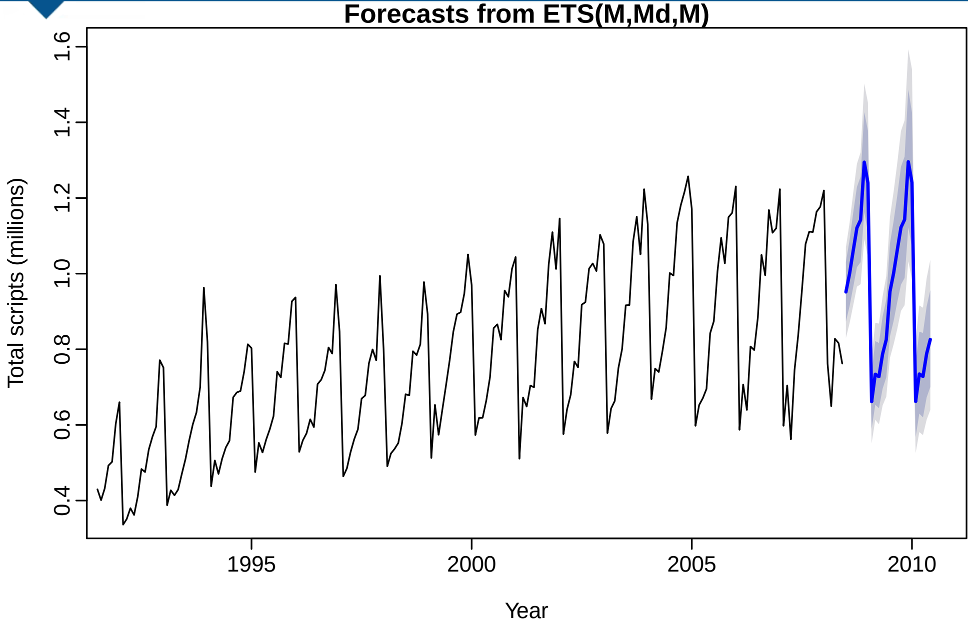
Exponential smoothing



Exponential smoothing

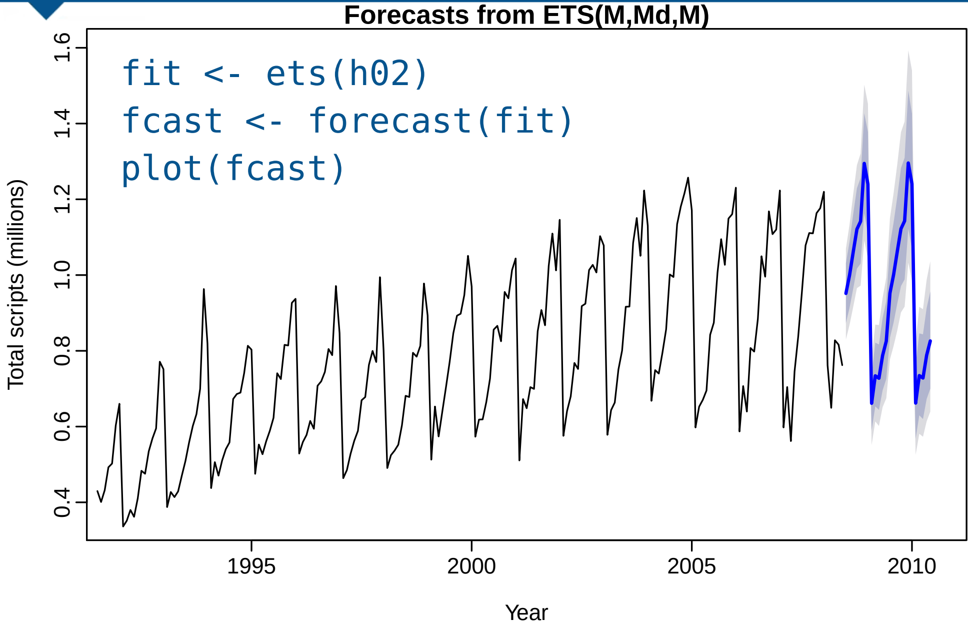


Exponential smoothing



Exponential smoothing

```r
fit <- ets(h02)
fcast <- forecast(fit)
plot(fcast)
```



M3 comparisons

Method	MAPE	sMAPE	MASE
Theta	17.42	12.76	1.39
ForecastPro	18.00	13.06	1.47
ForecastX	17.35	13.09	1.42
Automatic ANN	17.18	13.98	1.53
B-J automatic	19.13	13.72	1.54
ETS	18.06	13.38	1.52

Outline

ARIMA modelling

A non-seasonal ARIMA process

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t$$

B is backshift operator and ε_t is (Gaussian) iid noise.

A seasonal ARIMA process (period m)

$$\Phi(B^m)\phi(B)(1-B)^d(1-B^m)^D y_t = c + \Theta(B^m)\theta(B)\varepsilon_t$$

$-2\times$ Log likelihood function

$$L^*(\beta) = f(\beta) + T\log\sigma^2 + \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t - \hat{y}_{t|t-1})^2/r_t(\beta)$$

where β contains all parameters to be estimated and σ^2 is variance of $\{\varepsilon_t\}$.

ARIMA modelling

ELSEVIER

International Journal of Forecasting 16 (2000) 497–508

international journal of forecasting

www.elsevier.com/locate/ijforecast

Automatic ARIMA modeling including interventions, using time series expert software

G. Mélard*, J.-M. Pasteels

ISRO CP 210 (bldg NO room 2.O.9.300), Campus Plaine, Université Libre de Bruxelles, Bd du Triomphe, B-1050 Bruxelles, Belgium

Abstract

This article has three objectives: (a) to describe the method of automatic ARIMA modeling (AAM), with and without intervention analysis, that has been used in the analysis; (b) to comment on the results; and (c) to comment on the M3 Competition in general. Starting with a computer program for fitting an ARIMA model and a methodology for building univariate ARIMA models, an expert system has been built, while trying to avoid the pitfalls of most existing software packages. A software package called Time Series Expert TSE-AX is used to build a univariate ARIMA model with or without an intervention analysis. The characteristics of TSE-AX are summarized and, more especially, its automatic ARIMA modeling method. The motivation to take part in the M3-Competition is also outlined. The methodology is described mainly

ARIMA modelling

A Course in Time Series Analysis
Edited by Daniel Peña, George C. Tiao and Ruey S. Tsay

CHAPTER 7

Automatic Modeling Methods for Univariate Series

Víctor Gómez
Ministerio de Hacienda

Agustín Maravall
Banco de España

ARIMA modelling

Journal of Statistical Software

July 2008, Volume 26, Issue 3. http://www.jstatsoft.org/

Automatic Time Series Forecasting: The forecast Package for R

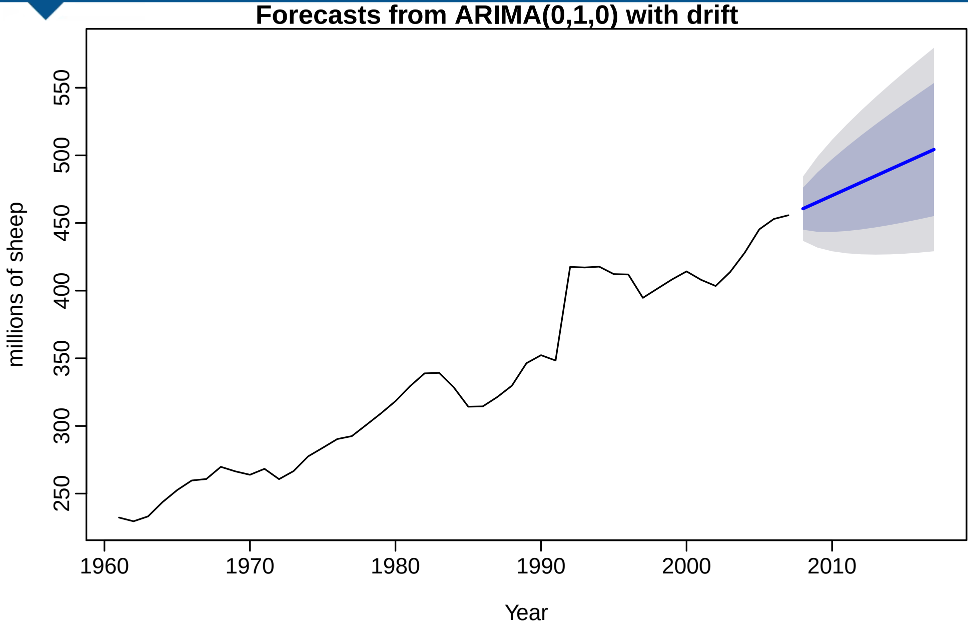
Rob J. Hyndman
Monash University

Yeasmin Khandakar
Monash University

Abstract

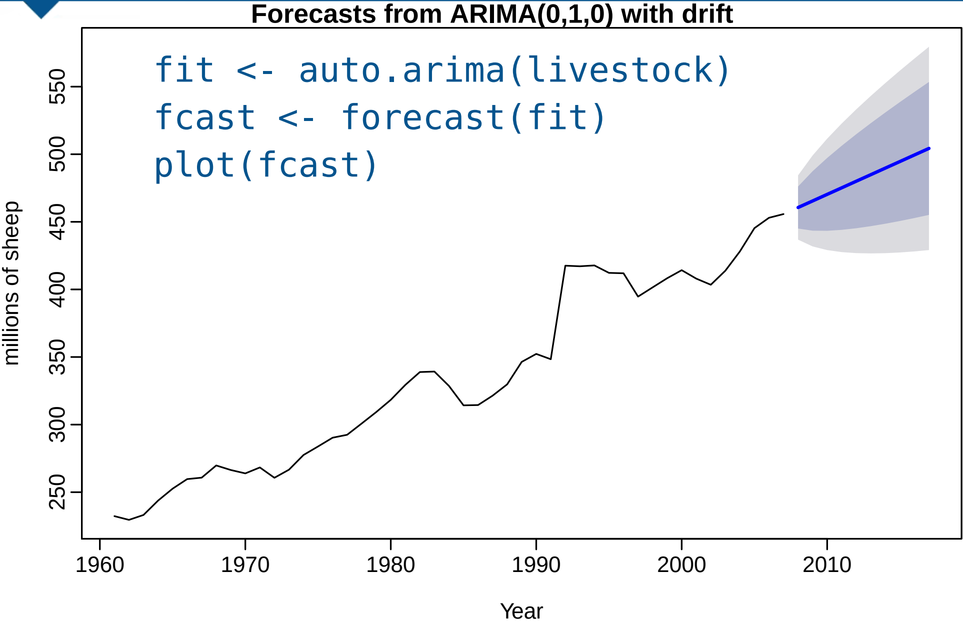
Automatic forecasts of large numbers of univariate time series are often needed in

Auto ARIMA

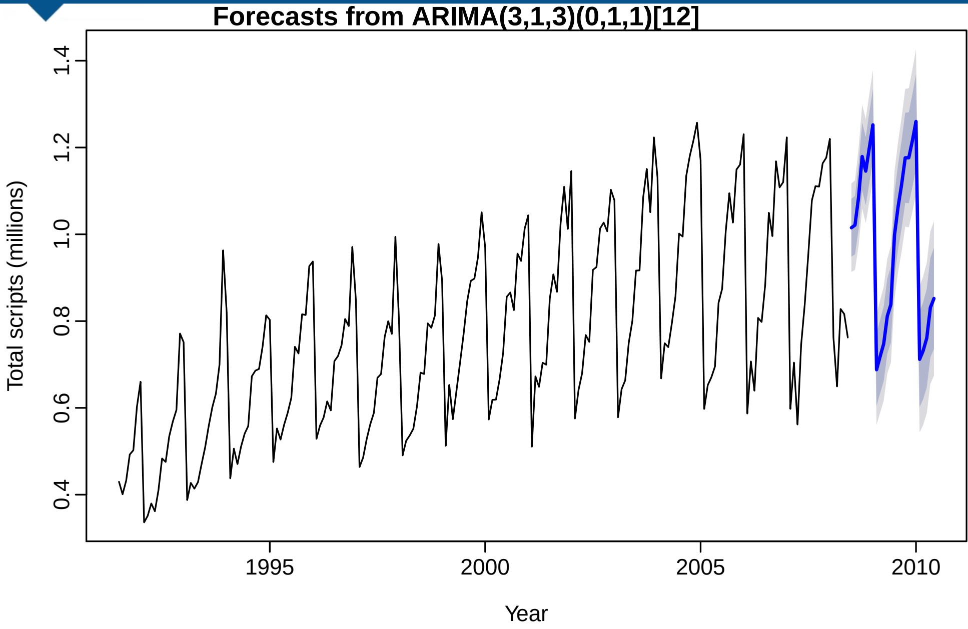

Forecasts from ARIMA(0,1,0) with drift
millions of sheep
250
300
350
400
450
500
550
1960
1970
1980
1990
2000
2010
Year

Auto ARIMA

```
fit <- auto.arima(livestock)
fcast <- forecast(fit)
plot(fcast)
```



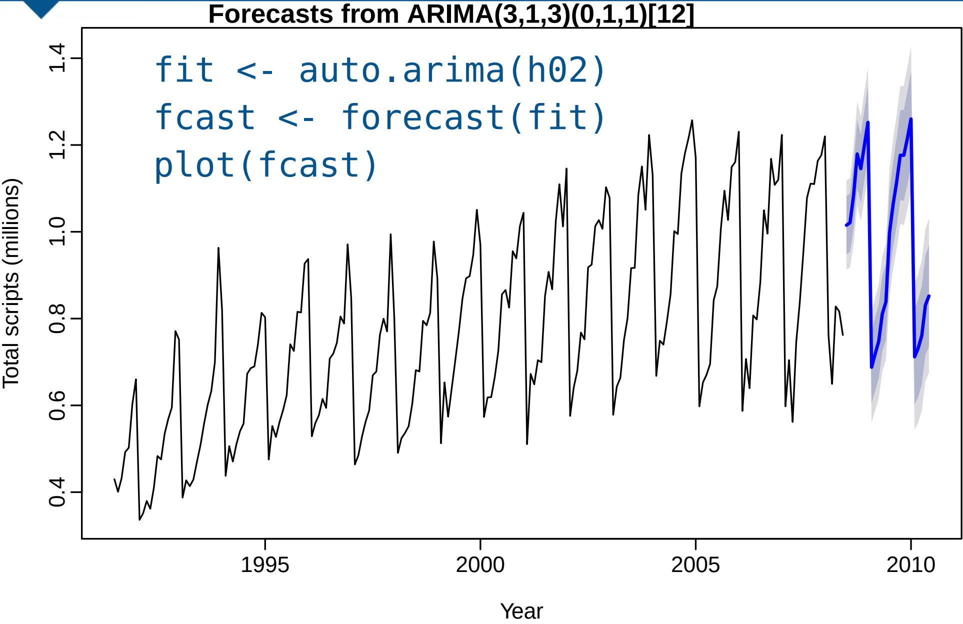
Auto ARIMA



Forecasts from ARIMA(3,1,3)(0,1,1)[12]
Total scripts (millions)
0.4
0.6
0.8
1.0
1.2
1.4
1995
2000
2005
2010
Year

Auto ARIMA

```
fit <- auto.arima(h02)
fcast <- forecast(fit)
plot(fcast)
```



How does auto.arima() work?

A non-seasonal ARIMA process

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t$$

Need to select appropriate orders p, q, d, and whether to include c.

Algorithm choices driven by forecast accuracy.

How does auto.arima() work?

A non-seasonal ARIMA process

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t$$

Need to select appropriate orders p, q, d, and whether to include c.

Hyndman & Khandakar (JSS, 2008) algorithm:

- Select no. differences d via KPSS unit root test.
- Select p, q, c by minimising AICc.
- Use stepwise search to traverse model space, starting with a simple model and considering nearby variants.

Algorithm choices driven by forecast accuracy.

How does auto.arima() work?

A non-seasonal ARIMA process

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t$$

Need to select appropriate orders p, q, d, and whether to include c.

Hyndman & Khandakar (JSS, 2008) algorithm:

- Select no. differences d via KPSS unit root test.
- Select p, q, c by minimising AICc.
- Use stepwise search to traverse model space, starting with a simple model and considering nearby variants.

Algorithm choices driven by forecast accuracy.

How does auto.arima() work?

A seasonal ARIMA process

$$\Phi(B^m)\phi(B)(1-B)^d(1-B^m)^D y_t = c + \Theta(B^m)\theta(B)\varepsilon_t$$

Need to select appropriate orders p, q, d, P, Q, D, and whether to include c.

Hyndman & Khandakar (JSS, 2008) algorithm:

- Select no. differences d via KPSS unit root test.
- Select D using OCSB unit root test.
- Select p, q, P, Q, c by minimising AIC.
- Use stepwise search to traverse model space, starting with a simple model and considering nearby variants.

M3 comparisons

Method	MAPE	sMAPE	MASE
Theta	17.42	12.76	1.39
ForecastPro	18.00	13.06	1.47
B-J automatic	19.13	13.72	1.54
ETS	18.06	13.38	1.52
AutoARIMA	19.04	13.86	1.47

M3 comparisons

Method	MAPE	sMAPE	MASE
Theta	17.42	12.76	1.39
ForecastPro	18.00	13.06	1.47
B-J automatic	19.13	13.72	1.54
ETS	18.06	13.38	1.52
AutoARIMA	19.04	13.86	1.47
ETS-ARIMA	17.92	13.02	1.44

M3 conclusions

MYTHS

- Simple methods do better.
- Exponential smoothing is better than ARIMA.

FACTS

- The best methods are hybrid approaches.

M3 conclusions

MYTHS

- Simple methods do better.
- Exponential smoothing is better than ARIMA.

FACTS

- The best methods are hybrid approaches.
- ETS-ARIMA (the simple average of ETS-selected and AutoARIMA) is the only fully documented method that is comparable to the M3 competition winners.

M3 conclusions

MYTHS

- Simple methods do better.
- Exponential smoothing is better than ARIMA.

FACTS

- The best methods are hybrid approaches.
- ETS-ARIMA (the simple average of ETS-additive and AutoARIMA) is the only fully documented method that is comparable to the M3 competition winners.

M3 conclusions

MYTHS

- Simple methods do better.
- Exponential smoothing is better than ARIMA.

FACTS

- The best methods are hybrid approaches.
- ETS-ARIMA (the simple average of ETS-additive and AutoARIMA) is the only fully documented method that is comparable to the M3 competition winners.

M3 conclusions

MYTHS

- Simple methods do better.
- Exponential smoothing is better than ARIMA.

FACTS

- The best methods are hybrid approaches.
- ETS-ARIMA (the simple average of ETS-additive and AutoARIMA) is the only fully documented method that is comparable to the M3 competition winners.

Outline

Automatic nonlinear forecasting

- Automatic ANN in M3 competition did poorly.
- Linear methods did best in the NN3 competition!
- Very few machine learning methods get published in the IJF because authors cannot demonstrate their methods give better forecasts than linear benchmark methods, even on supposedly nonlinear data.
- Some good recent work by Kourentzes and Crone (*Neurocomputing*, 2010) on automated ANN for time series.
- **Watch this space!**

Automatic nonlinear forecasting

- Automatic ANN in M3 competition did poorly.
- Linear methods did best in the NN3 competition!
- Very few machine learning methods get published in the IJF because authors cannot demonstrate their methods give better forecasts than linear benchmark methods, even on supposedly nonlinear data.
- Some good recent work by Kourentzes and Crone (*Neurocomputing*, 2010) on automated ANN for time series.
- **Watch this space!**

Automatic nonlinear forecasting

- Automatic ANN in M3 competition did poorly.
- Linear methods did best in the NN3 competition!
- Very few machine learning methods get published in the IJF because authors cannot demonstrate their methods give better forecasts than linear benchmark methods, even on supposedly nonlinear data.
- Some good recent work by Kourentzes and Crone (*Neurocomputing*, 2010) on automated ANN for time series.
- **Watch this space!**

Automatic nonlinear forecasting

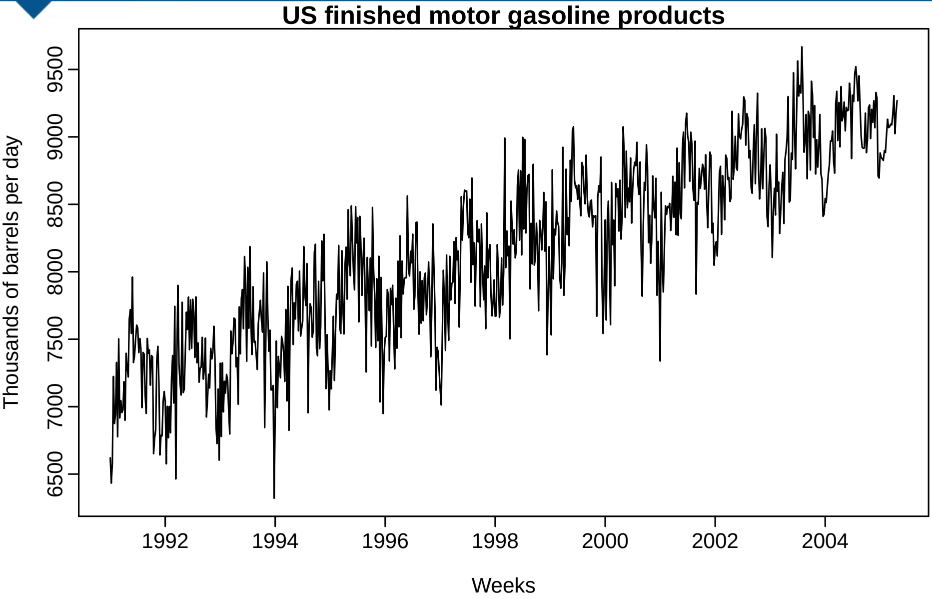
- Automatic ANN in M3 competition did poorly.
- Linear methods did best in the NN3 competition!
- Very few machine learning methods get published in the IJF because authors cannot demonstrate their methods give better forecasts than linear benchmark methods, even on supposedly nonlinear data.
- Some good recent work by Kourentzes and Crone (*Neurocomputing*, 2010) on automated ANN for time series.
- **Watch this space!**

Automatic nonlinear forecasting

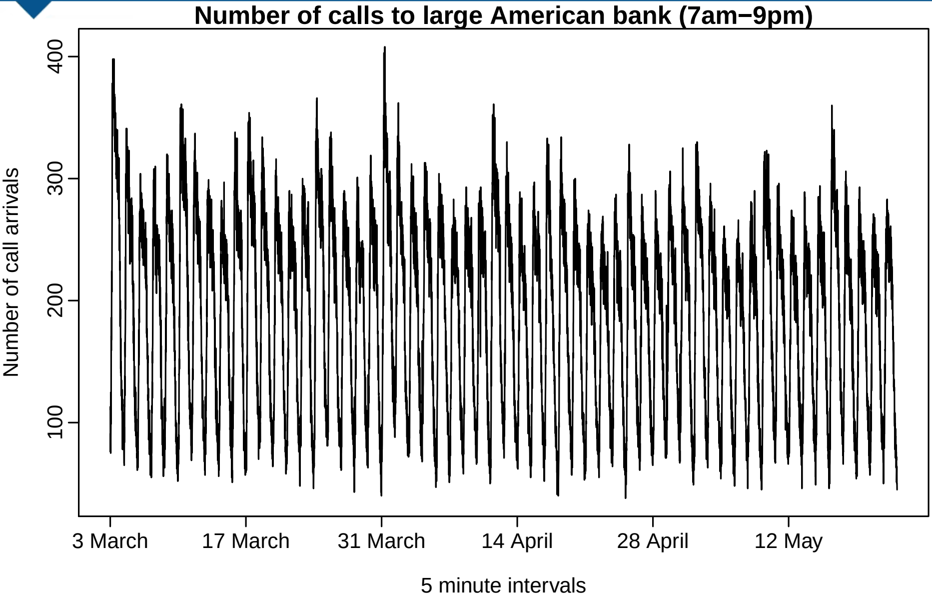
- Automatic ANN in M3 competition did poorly.
- Linear methods did best in the NN3 competition!
- Very few machine learning methods get published in the IJF because authors cannot demonstrate their methods give better forecasts than linear benchmark methods, even on supposedly nonlinear data.
- Some good recent work by Kourentzes and Crone (*Neurocomputing*, 2010) on automated ANN for time series.
- **Watch this space!**

Outline

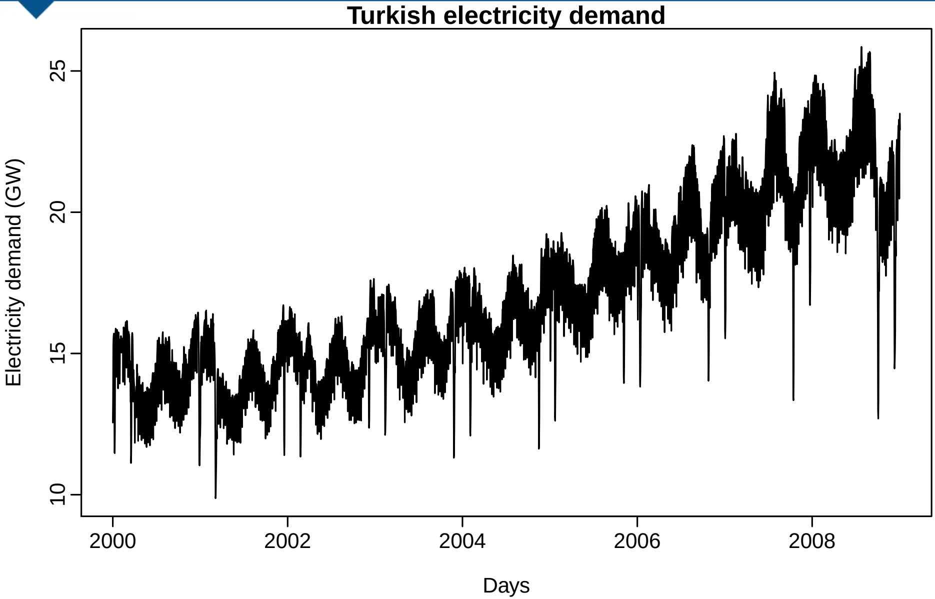
Examples


US finished motor gasoline products
Thousands of barrels per day
6500
7000
7500
8000
8500
9000
9500
1992
1994
1996
1998
2000
2002
2004
Weeks

Examples


Number of calls to large American bank (7am−9pm)
Number of call arrivals
100
200
300
400
3 March
17 March
31 March
14 April
28 April
12 May
5 minute intervals

Examples


Turkish electricity demand
Electricity demand (GW)
25
20
15
10
2000
2002
2004
2006
2008
Days

TBATS model

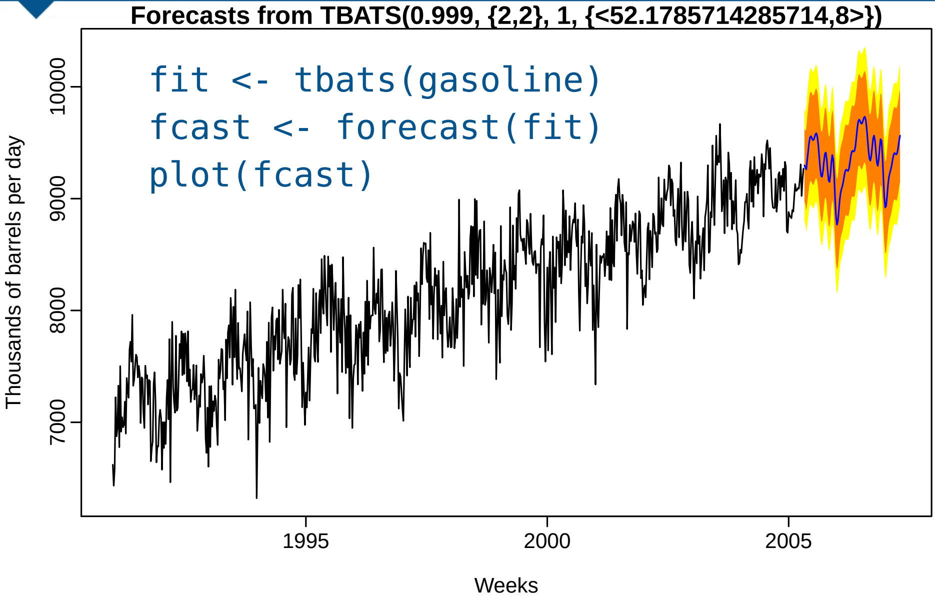
TBATS

Trigonometric terms for seasonality
Box-Cox transformations for heterogeneity
ARMA errors for short-term dynamics
Trend (possibly damped)
Seasonal (including multiple and non-integer periods)

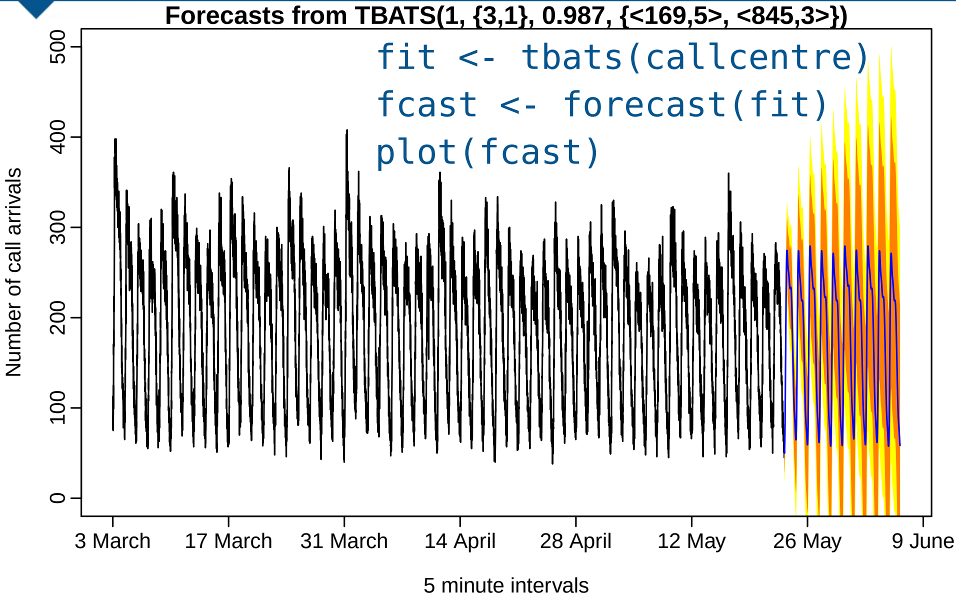
Automatic algorithm described in AM De Livera, RJ Hyndman, and RD Snyder (2011). “Forecasting time series with complex seasonal patterns using exponential smoothing”. *Journal of the American Statistical Association* **106**(496), 1513–1527.

Examples

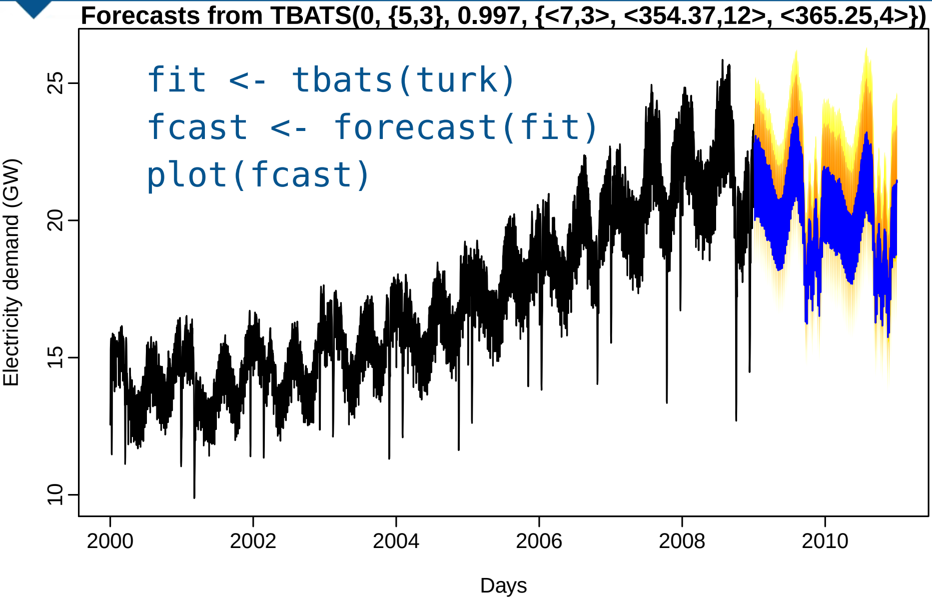
```
fit <- tbats(gasoline)
fcast <- forecast(fit)
plot(fcast)
```



Examples


Forecasts from TBATS(1, {3,1}, 0.987, {<169,5>, <845,3>})
fit <- tbats(callcentre)
fcast <- forecast(fit)
plot(fcast)
Number of call arrivals
0
100
200
300
400
500
3 March
17 March
31 March
14 April
28 April
12 May
26 May
9 June
5 minute intervals

Examples


Forecasts from TBATS(0, {5,3}, 0.997, {<7,3>, <354.37,12>, <365.25,4>})
fit <- tbats(turk)
fcast <- forecast(fit)
plot(fcast)
Electricity demand (GW)
Days
10
15
20
25
2000
2002
2004
2006
2008
2010

Outline

Forecasts about forecasting

1. Automatic algorithms will become more general — handling a wide variety of time series.
2. Model selection methods will take account of multi-step forecast accuracy as well as one-step forecast accuracy.
3. Automatic forecasting algorithms for multivariate time series will be developed.
4. Automatic forecasting algorithms that include covariate information will be developed.

Forecasts about forecasting

1. Automatic algorithms will become more general — handling a wide variety of time series.
2. Model selection methods will take account of multi-step forecast accuracy as well as one-step forecast accuracy.
3. Automatic forecasting algorithms for multivariate time series will be developed.
4. Automatic forecasting algorithms that include covariate information will be developed.

Forecasts about forecasting

1. Automatic algorithms will become more general — handling a wide variety of time series.
2. Model selection methods will take account of multi-step forecast accuracy as well as one-step forecast accuracy.
3. Automatic forecasting algorithms for multivariate time series will be developed.
4. Automatic forecasting algorithms that include covariate information will be developed.

Forecasts about forecasting

1. Automatic algorithms will become more general — handling a wide variety of time series.
2. Model selection methods will take account of multi-step forecast accuracy as well as one-step forecast accuracy.
3. Automatic forecasting algorithms for multivariate time series will be developed.
4. Automatic forecasting algorithms that include covariate information will be developed.

For further information

robjhyndman.com

- Slides and references for this talk.
- Links to all papers and books.
- Links to R packages.
- A blog about forecasting research.