MONASH University

MONASH BUSINESS SCHOOL

Feature-based forecasting algorithms for large collections of time series

Rob J Hyndman

25 January 2019

Outline

Outline

1. Makridakis forecasting competitions
2. Time series features
3. Feature-based forecasting algorithms

M competition: 1982

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: 1982

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
Univ

R. C
Univ

R. F
Man

M. H
INSE

R. L
Mark

J. N
E. P
Texa

R. W
India

ABS

In th … recasting. As a … that an appr … e 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: 1982

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

ABS…

In th… …ecasting. As a… …that an appr… …e 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

- 1001 series from demography, industry, economics.
- Annual, quarterly, monthly data.
- Anyone could submit forecasts.
- Multiple forecast measures used.

M3 competition: 2000

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

M3 competition: 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.

M4 competition: 2018

International Journal of Forecasting 34 (2018) 802–808

ELSEVIER

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

The M4 Competition: Results, findings, conclusion and way forward

Check for updates

Spyros Makridakis [a,b,*], Evangelos Spiliotis [c], Vassilios Assimakopoulos [c]

[a] *University of Nicosia, Nicosia, Cyprus*
[b] *Institute For the Future (IFF), Nicosia, Cyprus*
[c] *Forecasting and Strategy Unit, School of Electrical and Computer Engineering, National Technical University of Athens, Zografou, Greece*

ARTICLE INFO

Keywords:
Forecasting competitions
M Competitions

ABSTRACT

The M4 competition is the continuation of three previous competitions started more than 45 years ago whose purpose was to learn how to improve forecasting accuracy, and

M4 competition: 2018

- January – May 2018
- 100,000 time series: yearly, quarterly, monthly, weekly, daily, hourly.
- Point forecast and prediction intervals assessed.
- Code must be public
- 248 registrations, 50 submissions.

M4 competition: 2018

- January – May 2018
- 100,000 time series: yearly, quarterly, monthly, weekly, daily, hourly.
- Point forecast and prediction intervals assessed.
- Code must be public
- 248 registrations, 50 submissions.

Winning methods

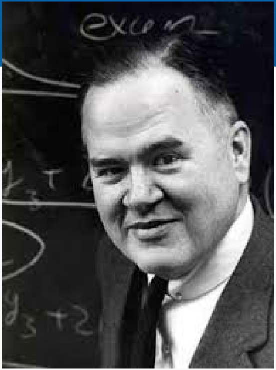
1. Hybrid of Recurrent Neural Network and Exponential Smoothing models
2. FFORMA: Feature-based forecast combinations using xgboost to find weights

Outline

Key idea

Cognostics

Computer-produced diagnostics
(Tukey and Tukey, 1985).

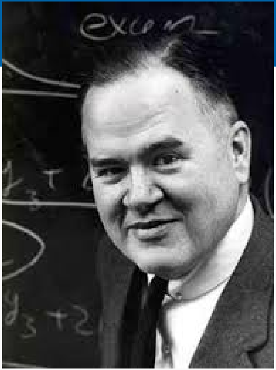


John W Tukey

Key idea

Cognostics

Computer-produced diagnostics (Tukey and Tukey, 1985).



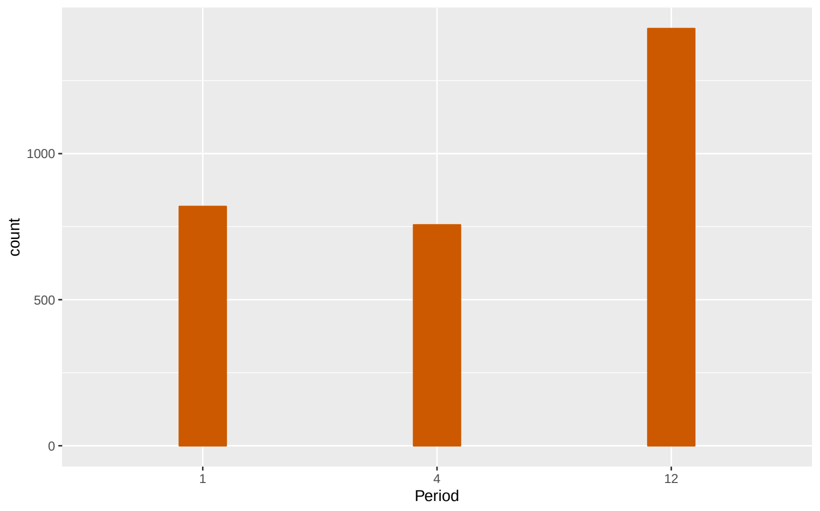
John W Tukey

Examples for time series

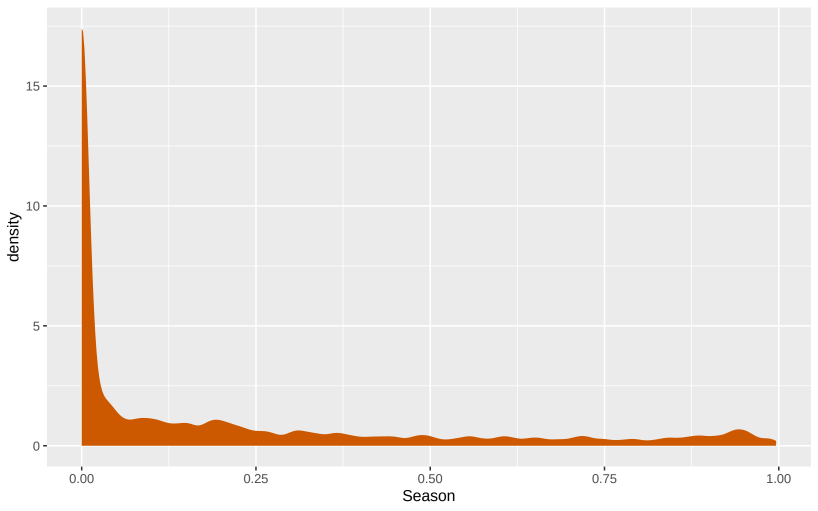
- lag correlation
- size and direction of trend
- strength of seasonality
- timing of peak seasonality
- spectral entropy

Called “features” in the machine learning literature.

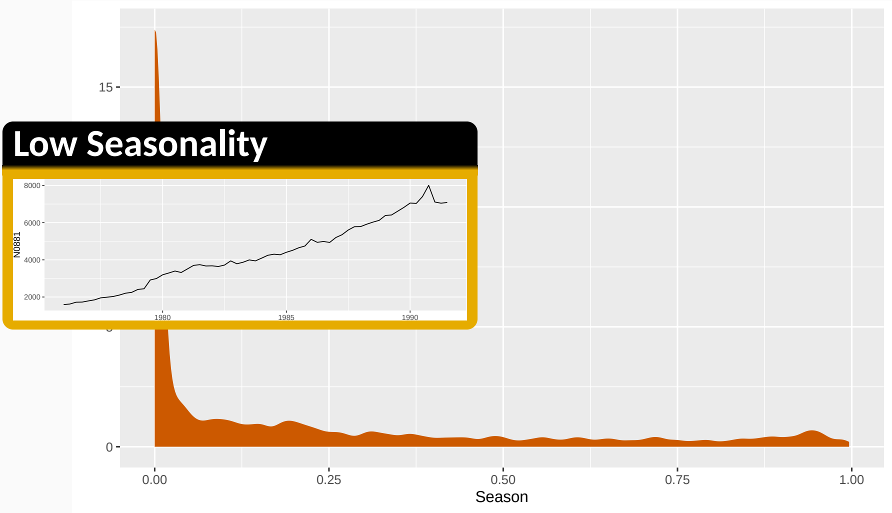
Distribution of Period for M3



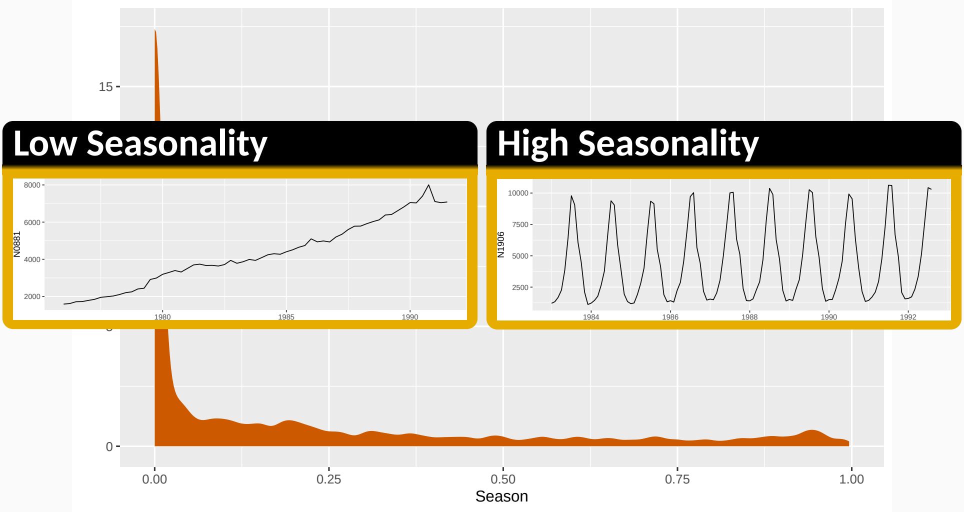
Distribution of Seasonality for M3



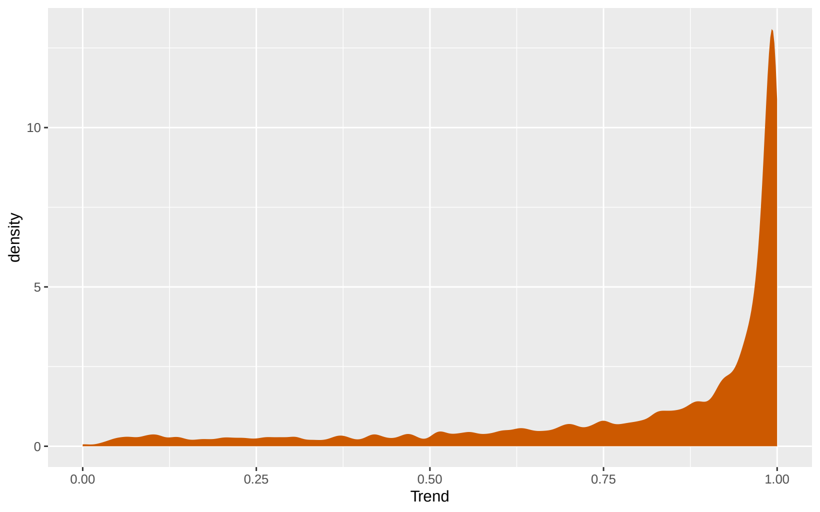
Distribution of Seasonality for M3



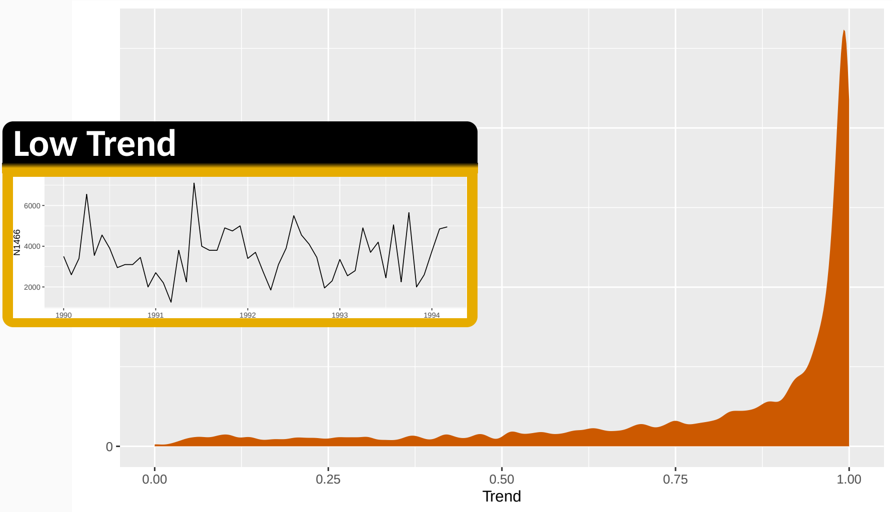
Distribution of Seasonality for M3



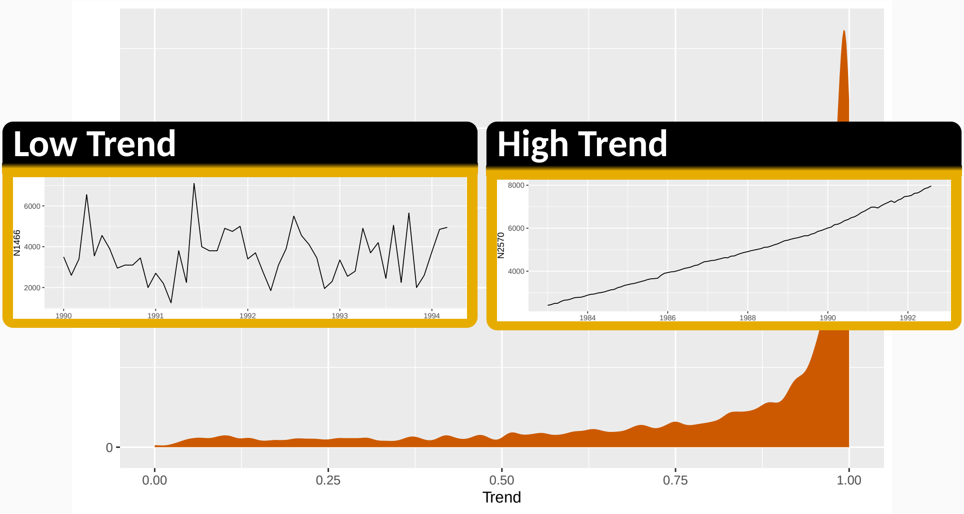
Distribution of Trend for M3



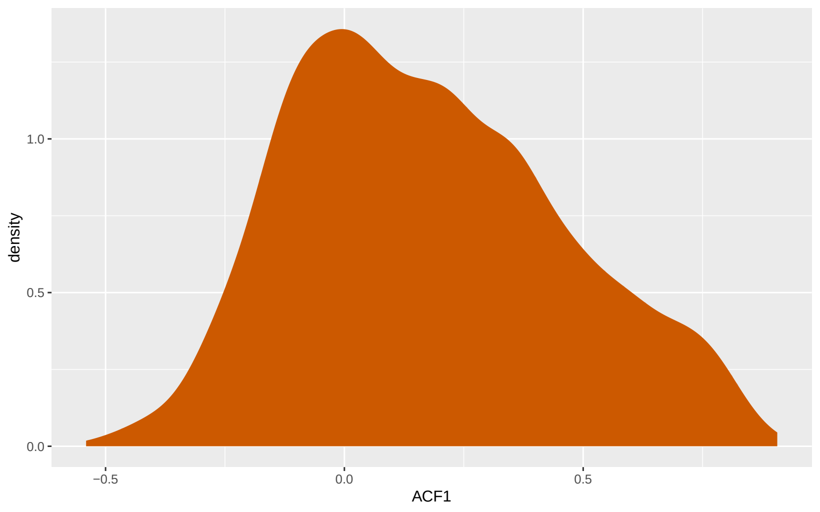
Distribution of Trend for M3



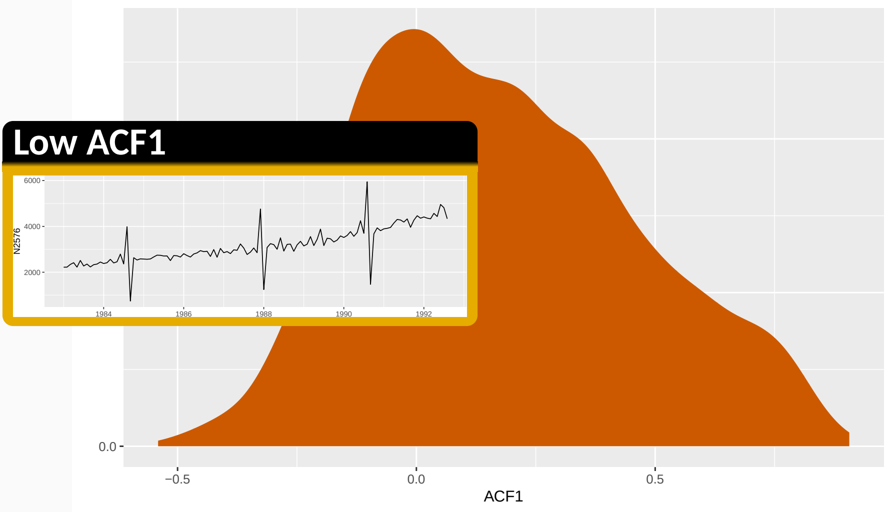
Distribution of Trend for M3



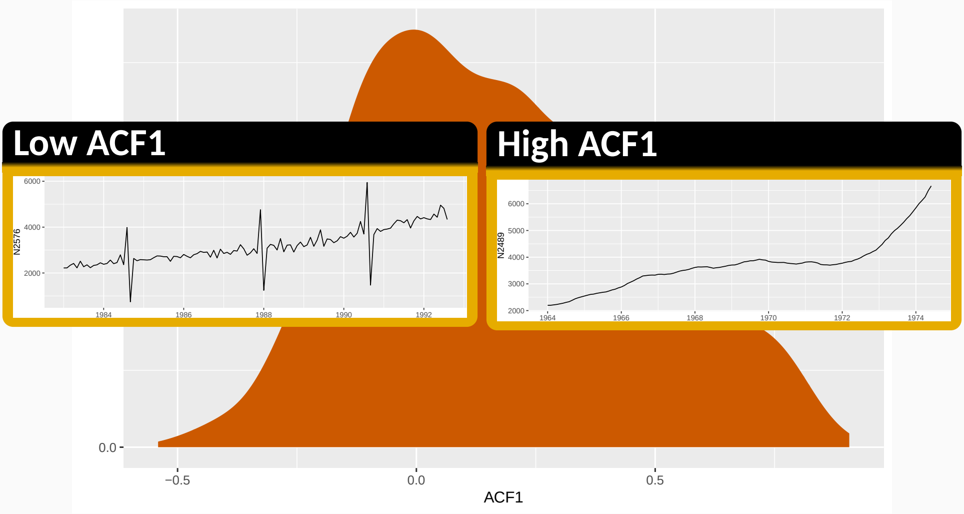
Distribution of Residual ACF1 for M3



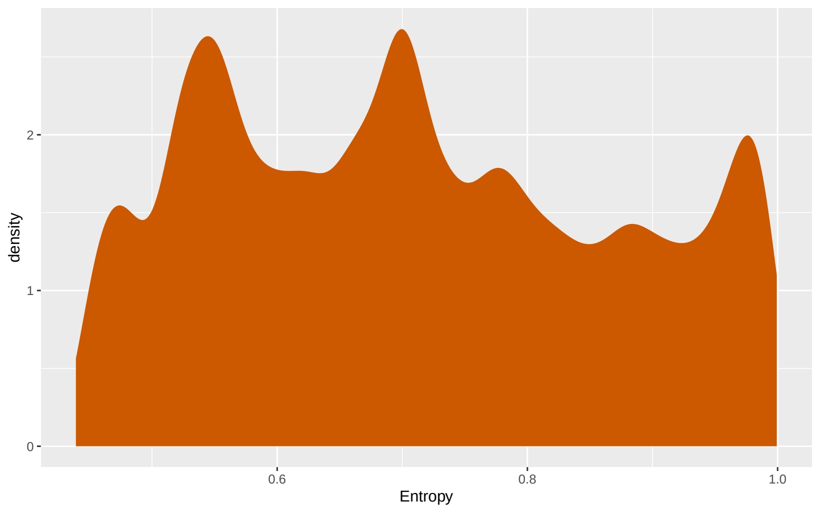
Distribution of Residual ACF1 for M3



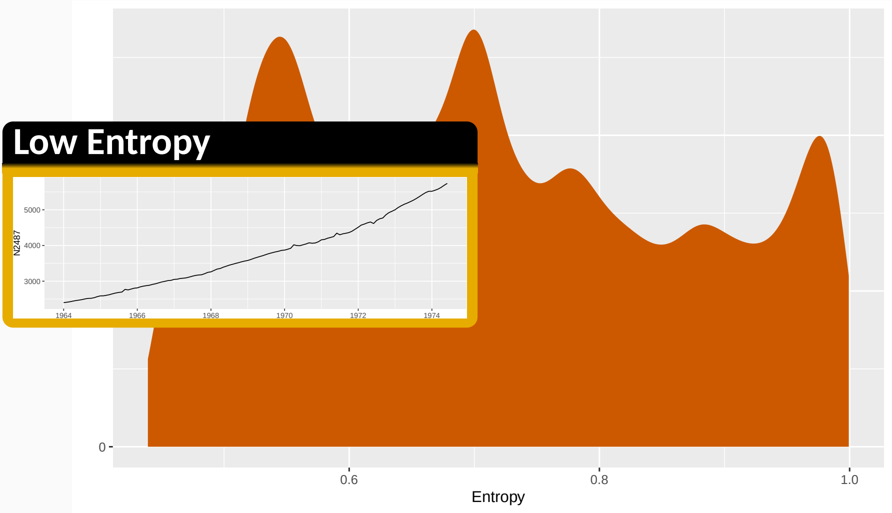
Distribution of Residual ACF1 for M3



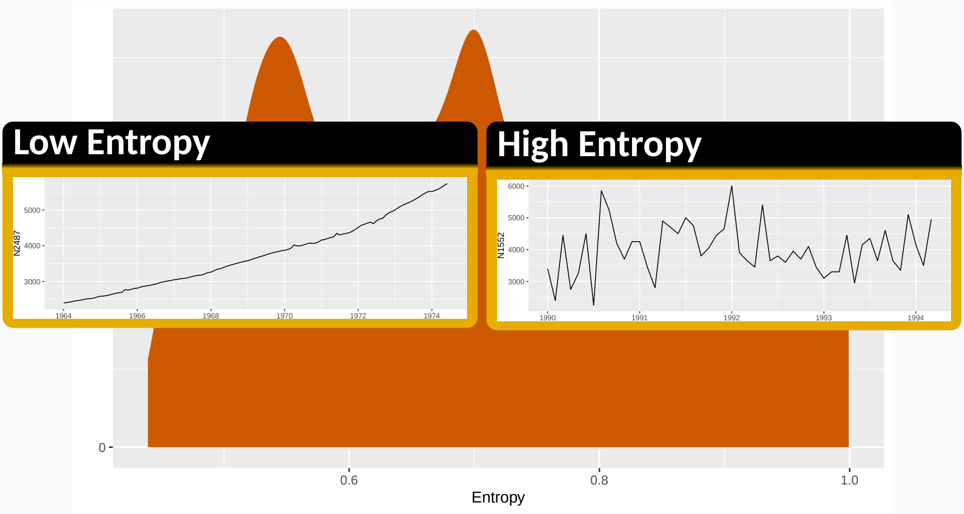
Distribution of Spectral Entropy for M3



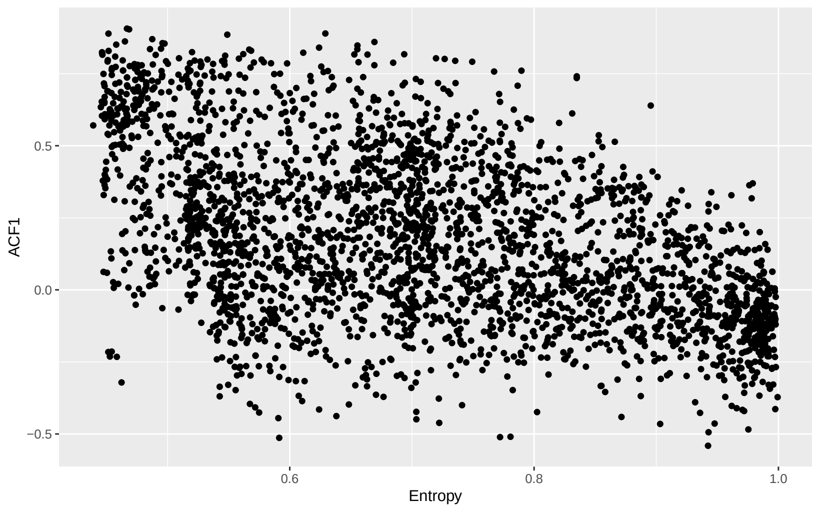
Distribution of Spectral Entropy for M3



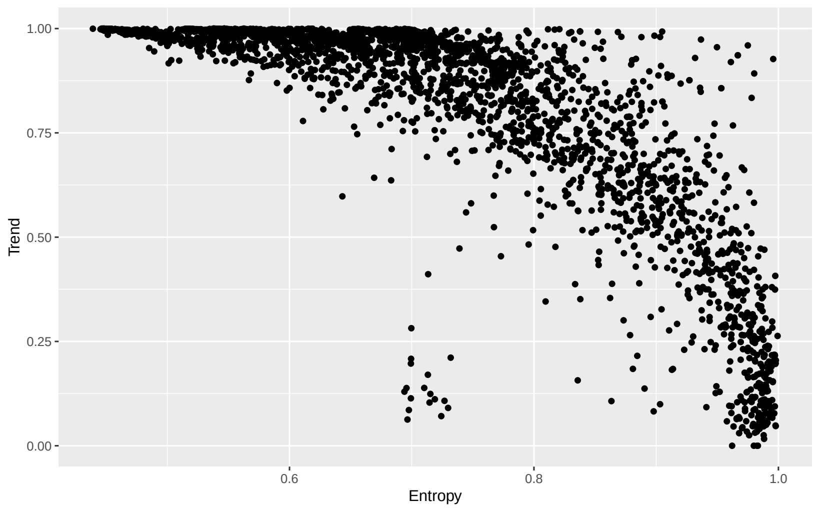
Distribution of Spectral Entropy for M3



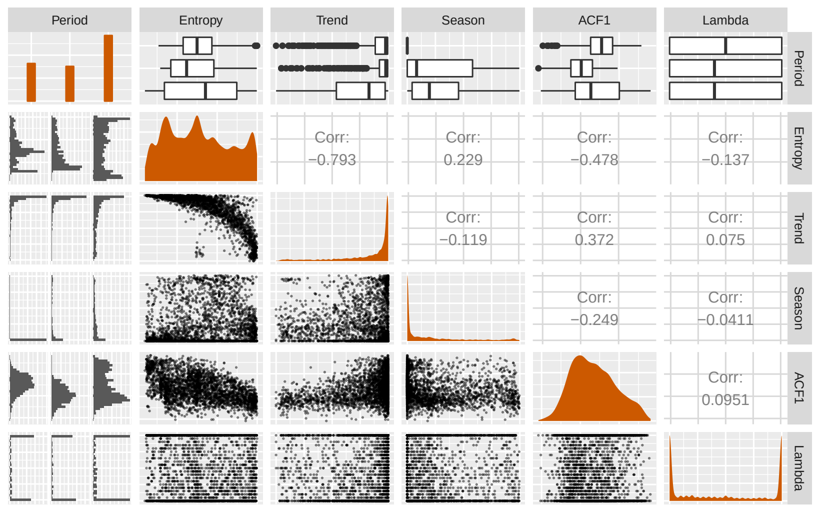
Feature distributions



Feature distributions



Feature distributions



Outline

Features used to select a forecasting model

- length
- strength of seasonality
- strength of trend
- linearity
- curvature
- spikiness
- stability
- lumpiness
- parameter estimates of Holt's linear trend method
- spectral entropy
- Hurst exponent
- nonlinearity
- parameter estimates of Holt-Winters' additive method
- unit root test statistics
- crossing points, flat spots
- peaks, troughs
- ACF and PACF based features - calculated on raw, differenced, and remainder series.
- ARCH/GARCH statistics and ACF of squared series and residuals.

Features used to select a forecasting model

Why these features?

- Hyndman, Wang and Laptev. “Large scale unusual time series detection” (ICDM 2015).
- Kang, Hyndman & Smith-Miles. “Visualising forecasting algorithm performance using time series instance spaces” (IJF 2017).
- Talagala, Hyndman and Athanasopoulos. “Meta-learning how to forecast time series” (2018).
- Implemented in the tsfeatures R package

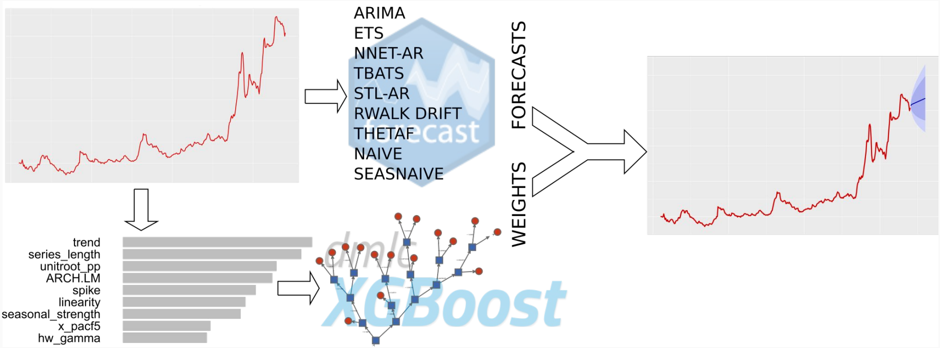
FFORMS: Feature-based FORecast Model Selection

- Using large collection of time series, each series split into training and test sets.
- Features computed on training data.
- All forecasting methods fitted to training data, and forecasts obtained for test data period.
- Forecast accuracy for each method/series computed from test data.
- Train a random forest to identify the most accurate forecasting method for a given time series using only a vector of features.

FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but using gradient boosted trees (xgboost) rather than random forest.
- Trained on temporal holdout version of M4 dataset, where size of test sets equal to required forecast horizons
- Pptimization criterion: forecast accuracy not classification accuracy.
- Probability of each model being best is used to construct model weights for combination forecast.
- 5 days computing time.

FFORMA: Feature-based FORecast Model Averaging



M4 competition results (based on average OWA)

1st	0.821	
2nd	0.838	(FFORMA)
3rd	0.841	

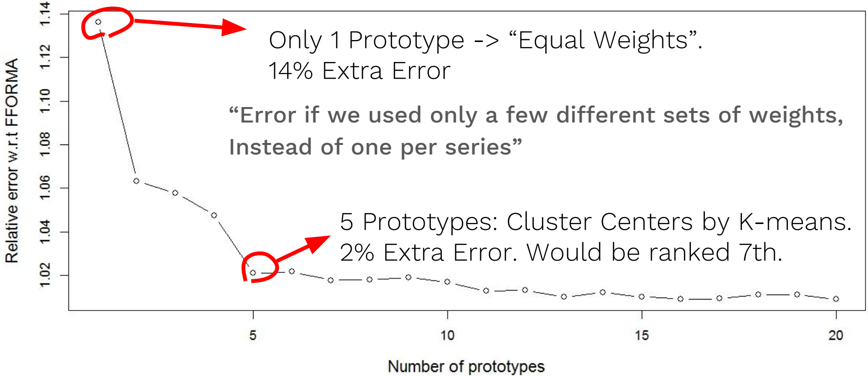
FFORMA: Feature-based FORecast Model Averaging

Models included

1. Naive
2. Seasonal naive
3. Random walk with drift
4. Theta method
5. ARIMA
6. ETS
7. TBATS
8. STL decomposition with AR for seasonally adjusted series
9. Neural network autoregression

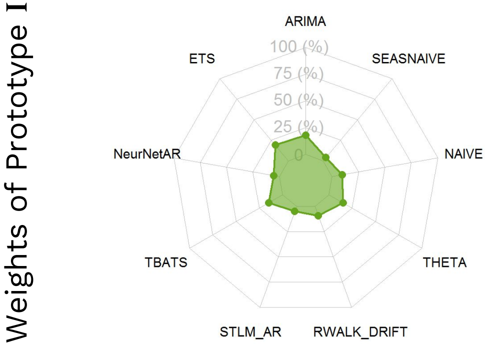
FFORMA: Feature-based FORecast Model Averaging

Looking for Prototypes in the weights



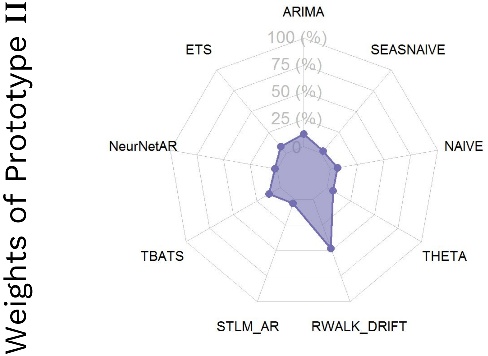
FFORMA: Feature-based FORecast Model Averaging

“Roughly Equal Weights”. 40000 Series in M4



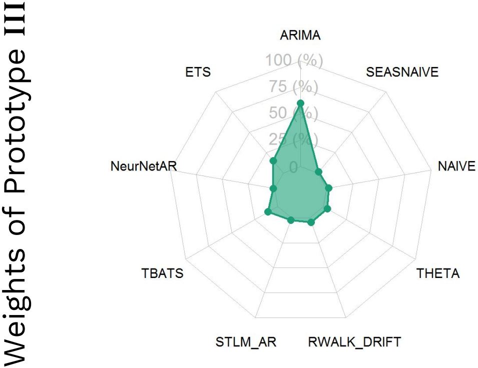
FFORMA: Feature-based FORecast Model Averaging

“Mostly RandomWalk Drift”. 20000 Series in M4



FFORMA: Feature-based FORecast Model Averaging

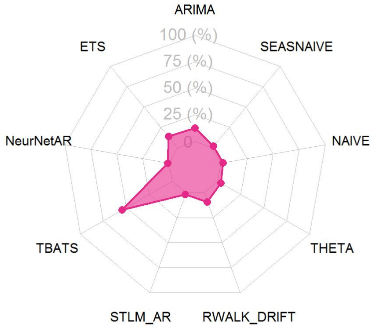
“Mostly ARIMA”. 16000 Series in M4



FFORMA: Feature-based FORecast Model Averaging

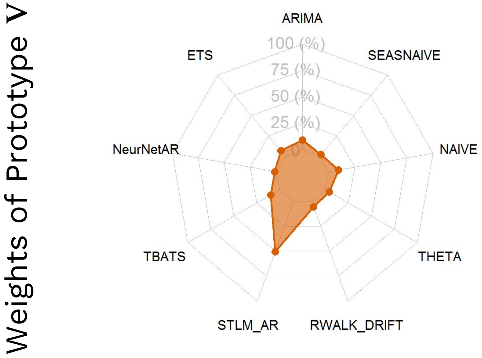
“Mostly TBATS”. 13000 Series in M4

Weights of Prototype IV



FFORMA: Feature-based FORecast Model Averaging

"Mostly STLM-AR". 8000 Series in M4



Papers and packages

R packages

- **tsfeatures**: Calculating time series features.
  `github.com/robjhyndman/tsfeatures`
- **seer**: FFORMS — selecting forecasting model using features.
  `github.com/thiyangt/seer`
- **M4metalearning**: FFORMA – forecast combinations using features to choose weights.
  `github.com/robjhyndman/M4metalearning`

Papers

Available from `robjhyndman.com`

Acknowledgements

Kate Smith-Miles

Yanfei Kang

Earo Wang

Thiyanga Talagala

George Athanasopoulos

Pablo Montero-Manso