MONASH University

MONASH
BUSINESS
SCHOOL

Feature-based Time Series Forecasting

Thiyanga S. Talagala

13 March 2019

Joint work with

Rob J Hyndman

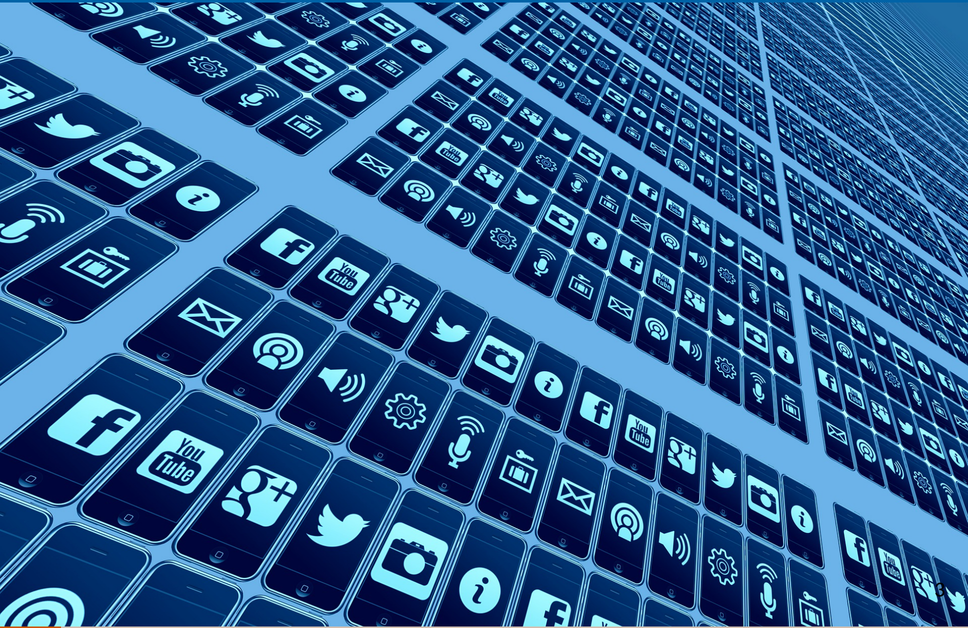
George Athanasopoulos

Pablo Montero-Manso

Feng Li

Yanfei Kang

Introduction

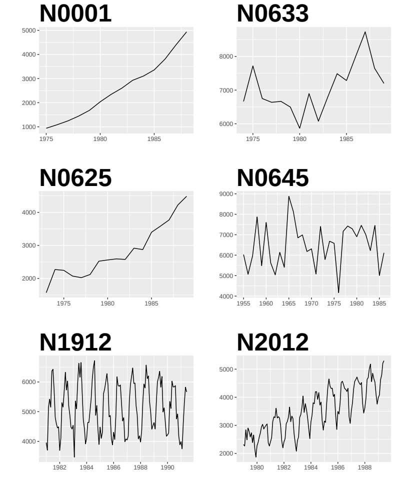


Time series features

- Transform a given time series $y = \{y_1, y_2, \cdots, y_n\}$ to a feature vector $F = (f_1(y), f_2(y), \cdots, f_p(y))'\}$.

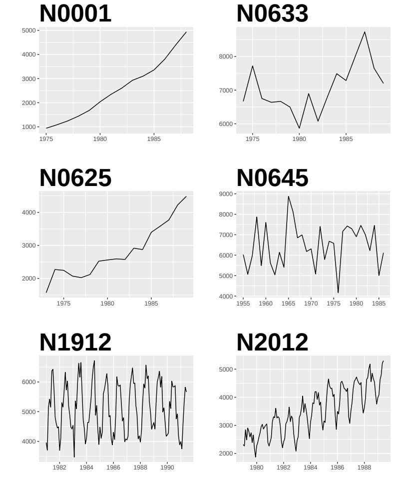
Time series features

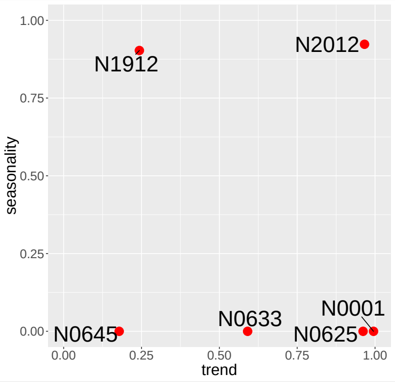
- Transform a given time series $y = \{y_1, y_2, \cdots, y_n\}$ to a feature vector $F = (f_1(y), f_2(y), \cdots, f_p(y))'\}$.



Time series features

- Transform a given time series $y = \{y_1, y_2, \cdots, y_n\}$ to a feature vector $F = (f_1(y), f_2(y), \cdots, f_p(y))'\}$.

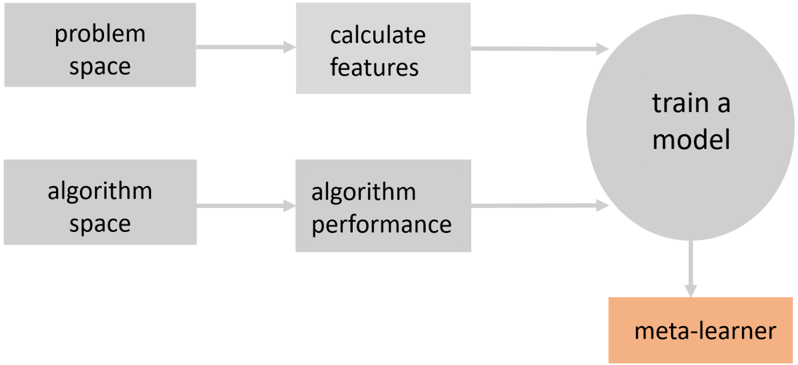




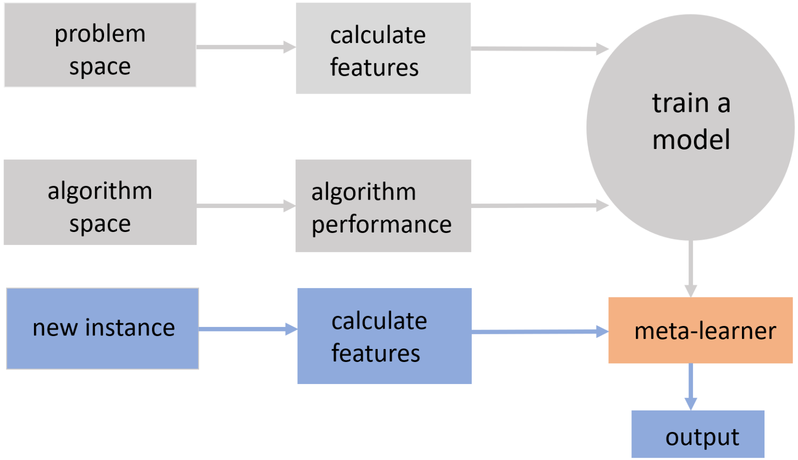
Features used to select a forecasting model

- 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.

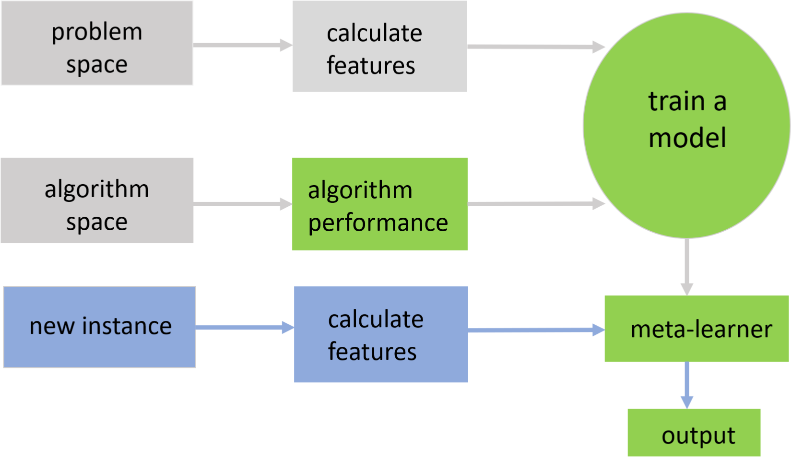
Meta-learning



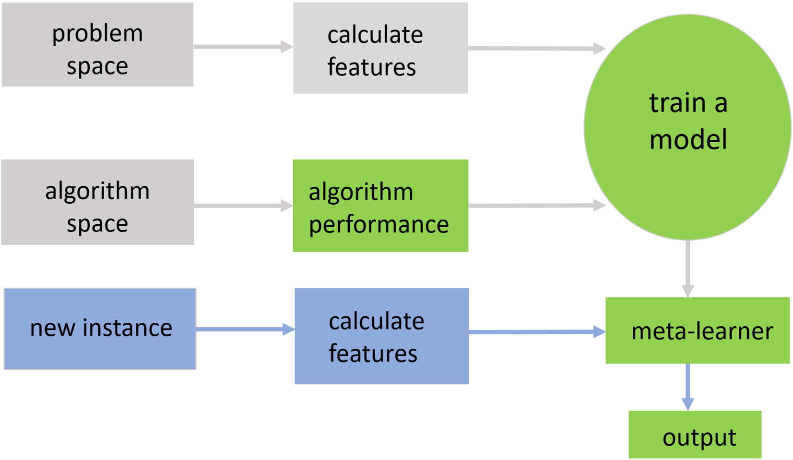
Meta-learning



Feature-based forecasting algorithms

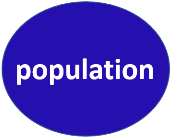


Feature-based forecasting algorithms

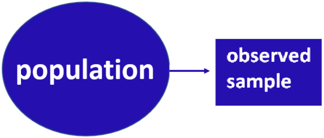


- three algorithms: FFORMS, FFORMA, FFORMPP

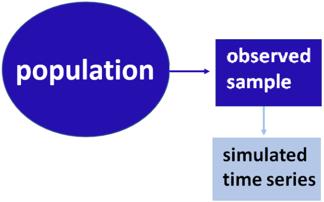
FFORMS: Feature-based FORecast Model Selection



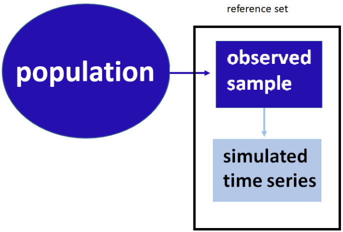
FFORMS: observed sample



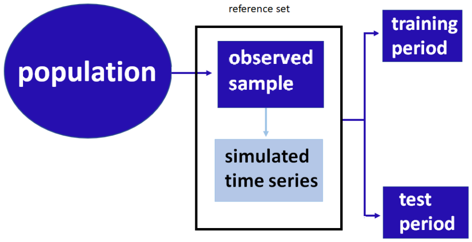
FFORMS: simulated time series



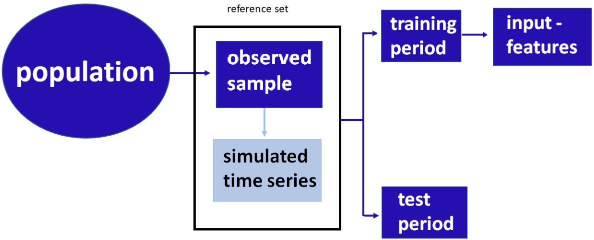
FFORMS: reference set



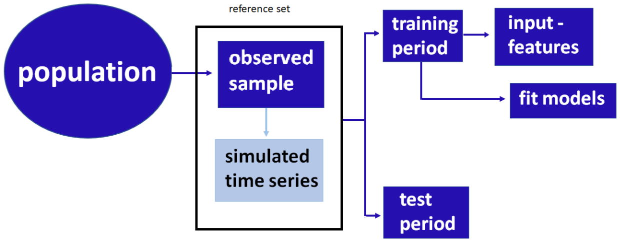
FFORMS: Meta-data



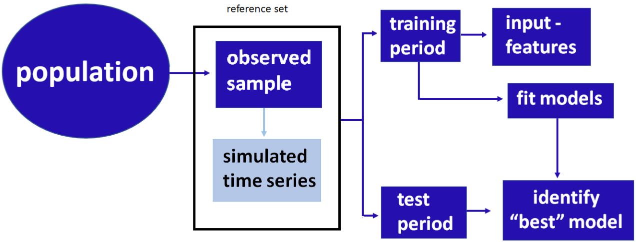
FFORMS: Meta-data



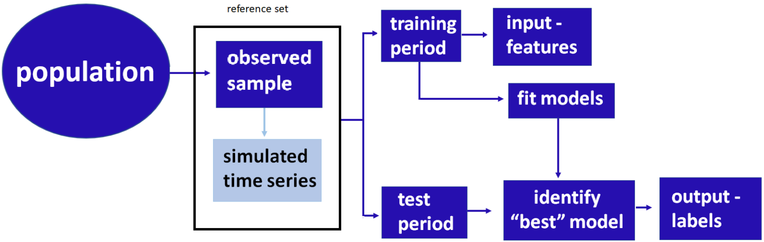
FFORMS: Meta-data



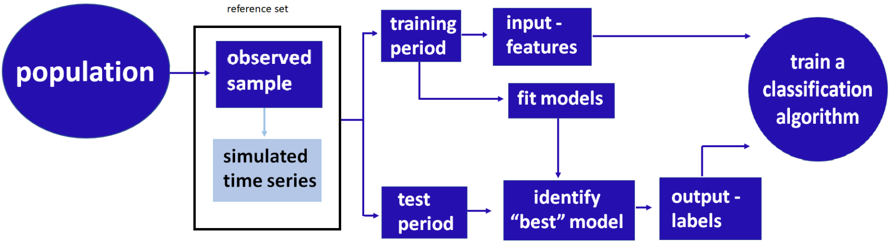
FFORMS: Meta-data



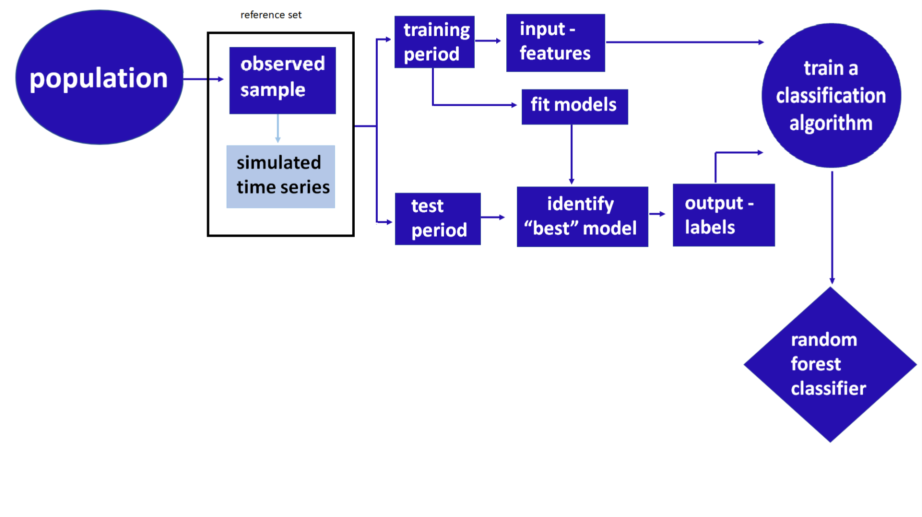
FFORMS: Meta-data



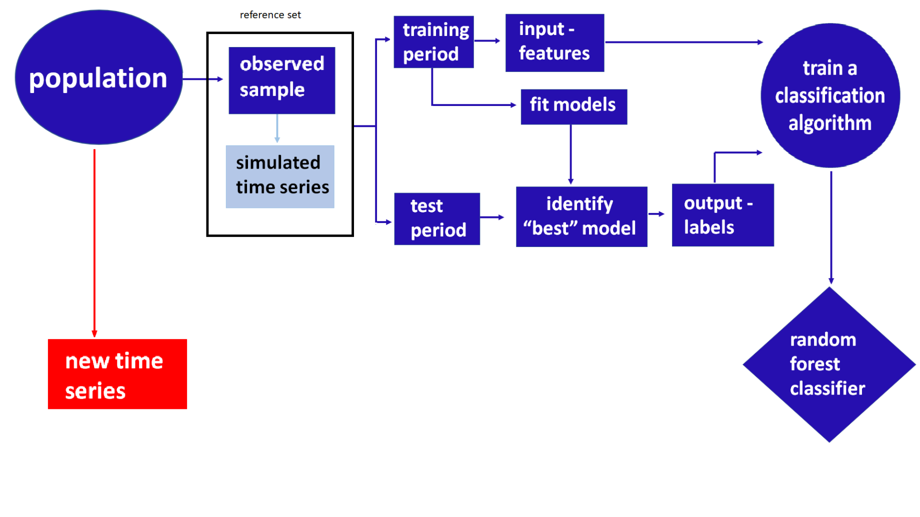
FFORMS: Meta-data



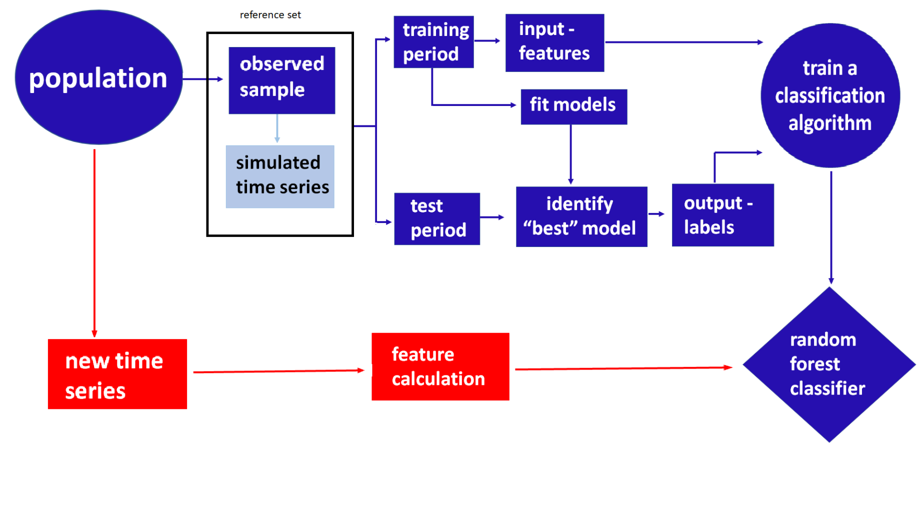
FFORMS: Random-forest classifier



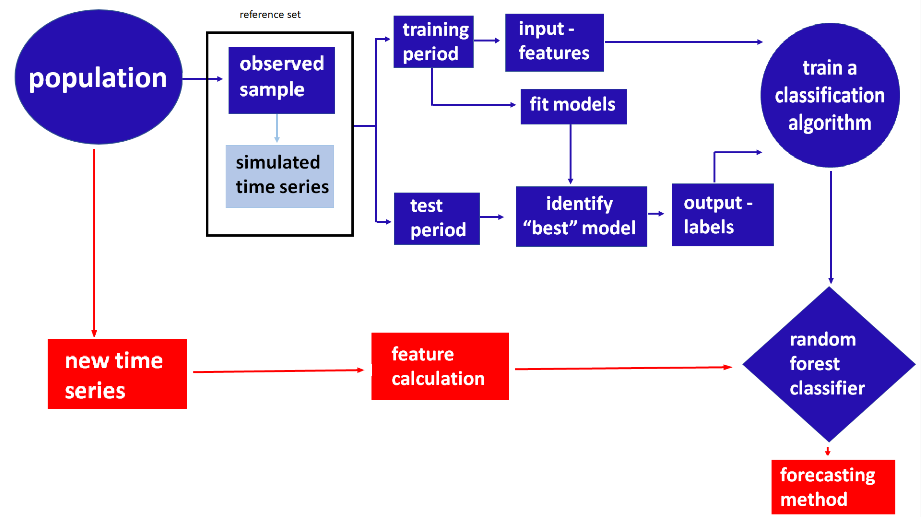
FFORMS: “online” part of the algorithm



FFORMS: “online” part of the algorithm



FFORMS: “online” part of the algorithm



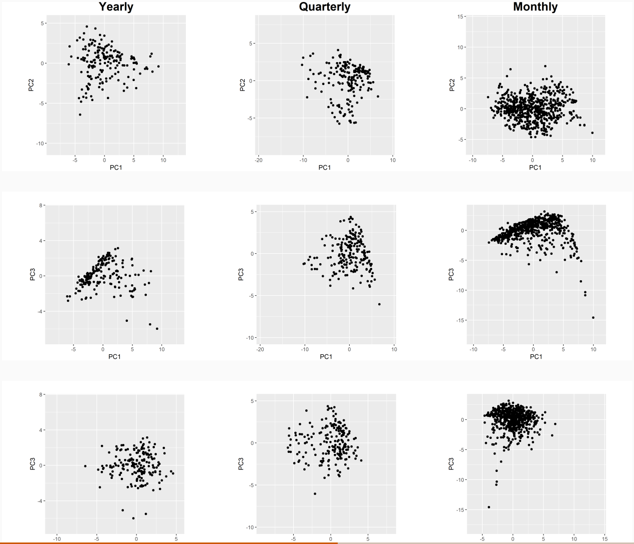
Application to M competition data

- Proposed algorithm is applied to yearly, quarterly and monthly series separately.
- We run two experiments for each case.

	Experiment 1				Experiment 2			
	Source	Y	Q	M	Source	Y	Q	M
Observed series	M1	181	203	617	M3	645	756	1428
Simulated series		362000	406000	123400		1290000	1512000	285600
New series	M3	645	756	1428	M1	181	203	617

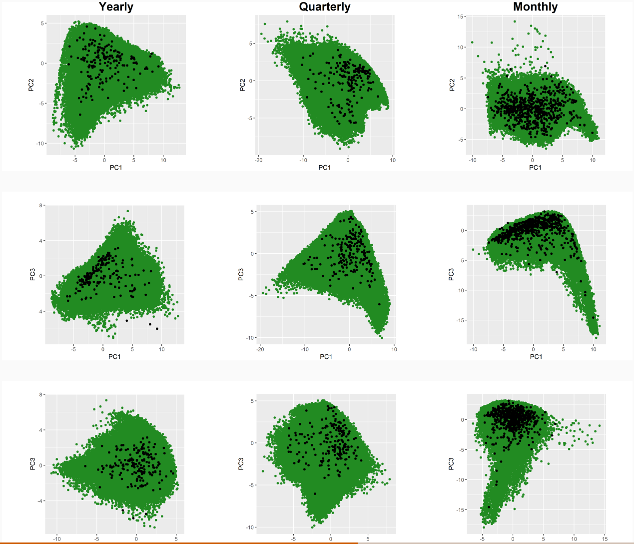
Experiment 1: Distribution of time series in the PCA space

observed - M1



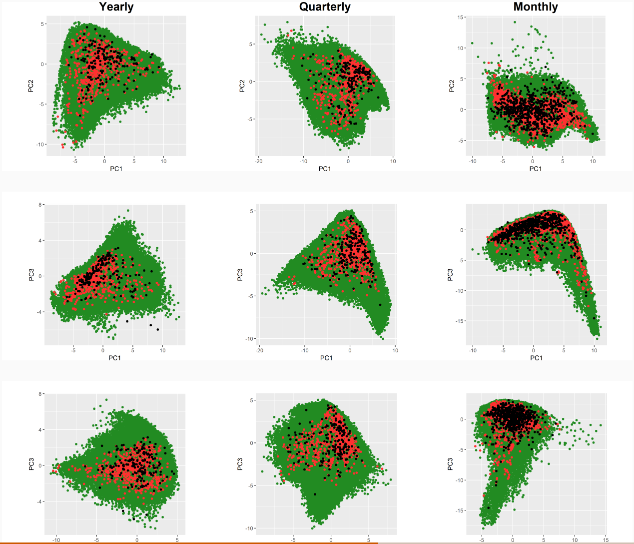
Experiment 1: Distribution of time series in the PCA space

observed - M1 simulated



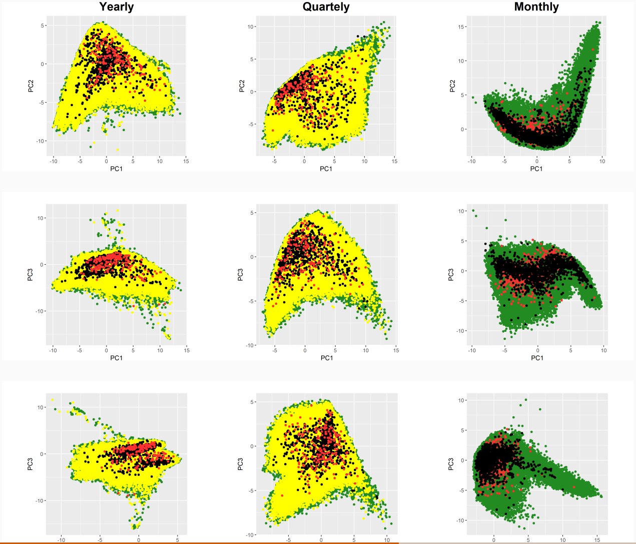
Experiment 1: Distribution of time series in the PCA space

observed - M1 simulated new - M3

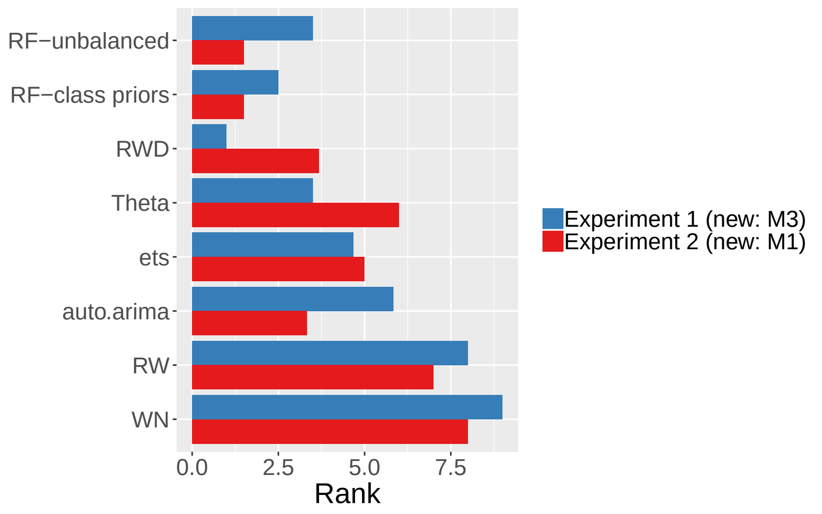


Experiment 2: Distribution of time series in the PCA space

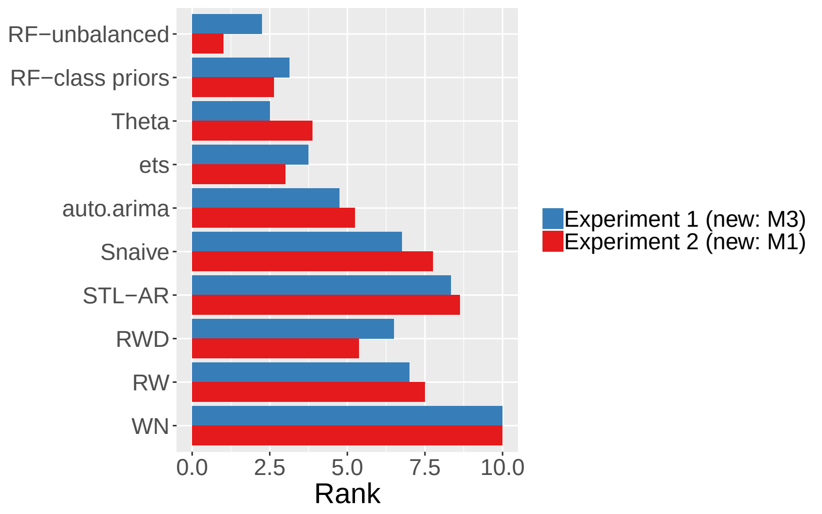
observed - M3 | simulated | subset | new - M1



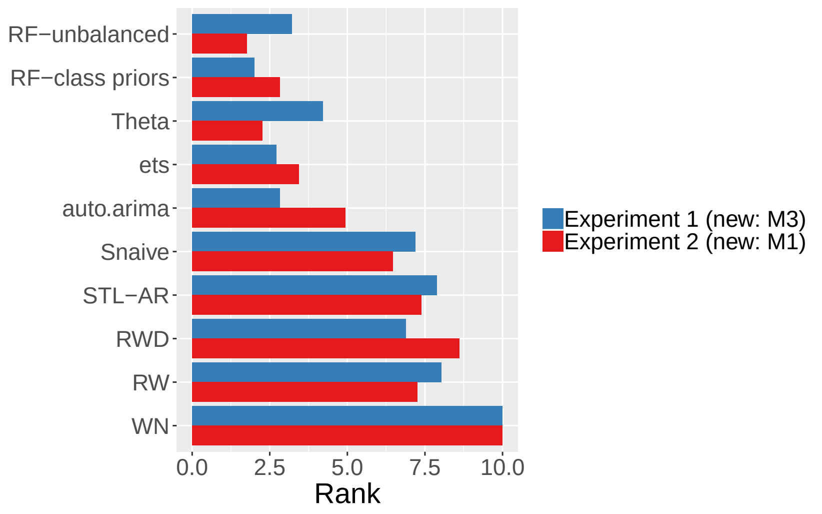
Results: Yearly



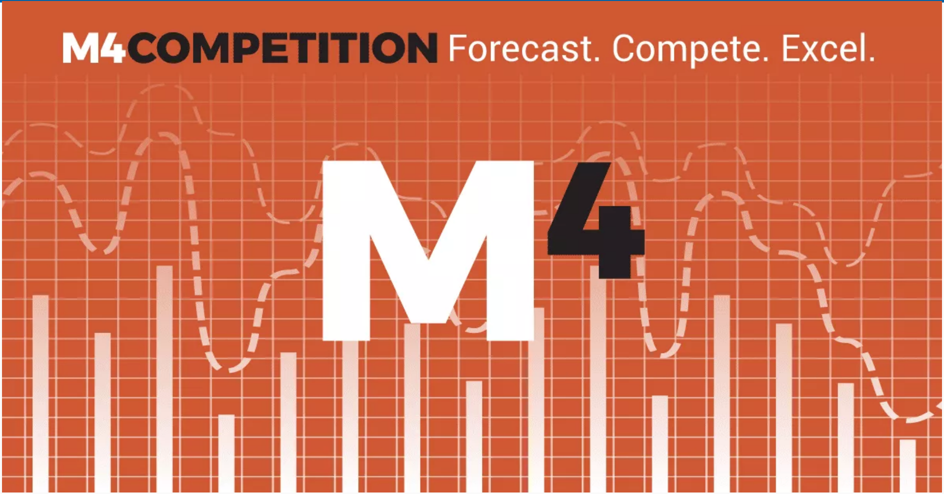
Results: Quarterly



Results: Monthly

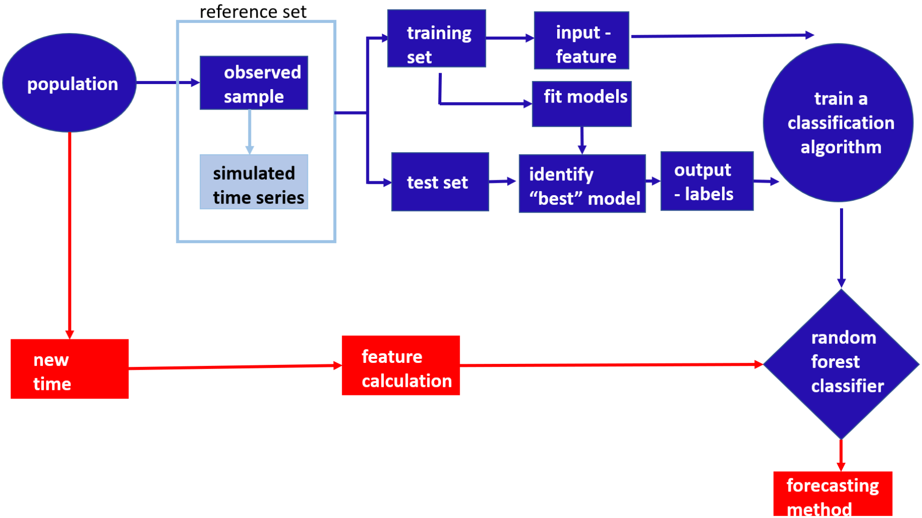


M4 Competition: 2018

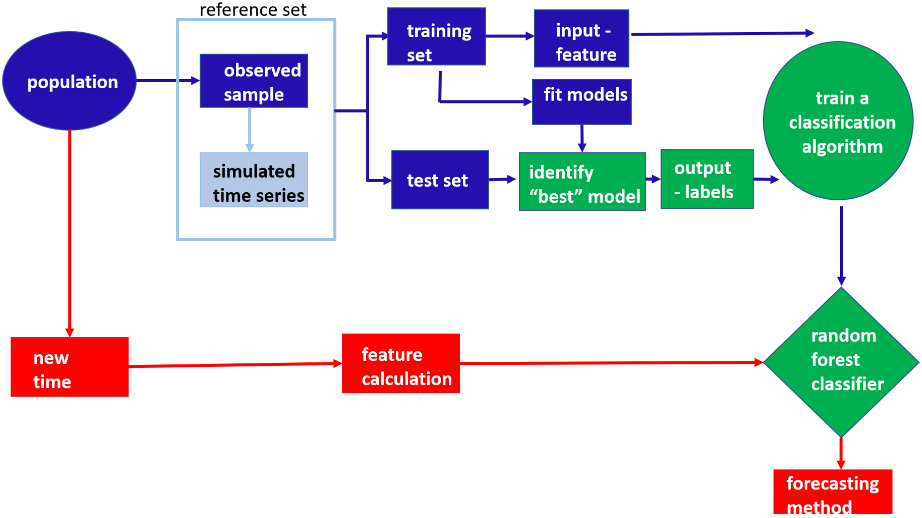


- 100,000 time series: yearly, quarterly, monthly, weekly, daily, hourly

FFORMS: Feature-based FORecast Model Selection

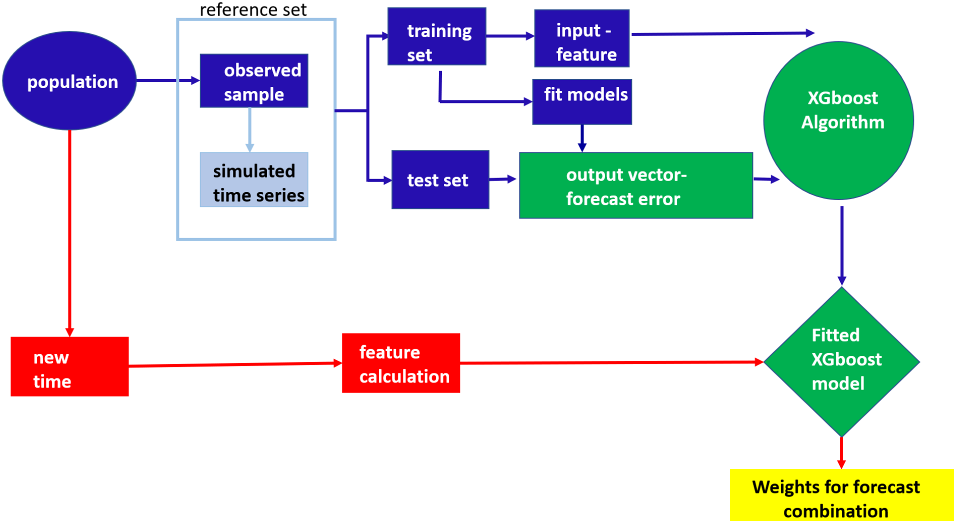


FFORMS: Feature-based FORecast Model Selection



- optimization criterion: classification accuracy

FFORMA: Feature-based FORecast Model Averaging



- optimization criterion: forecast accuracy

FFORMA: Models included

- naive
- random walk with drift
- seasonal naive
- theta method
- automated ARIMA algorithm
- automated exponential smoothing algorithm
- TBATS model
- STLM-AR Seasonal and Trend decomposition using Loess with AR modeling of the seasonally adjusted series
- neural network time series forecasts

FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.

FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.
- Optimization criterion: forecast accuracy

FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.
- Optimization criterion: forecast accuracy
- The probability of each model being best is used to construct a model weight.

FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.
- Optimization criterion: forecast accuracy
- The probability of each model being best is used to construct a model weight.
- A combination forecast is produced using these weights.

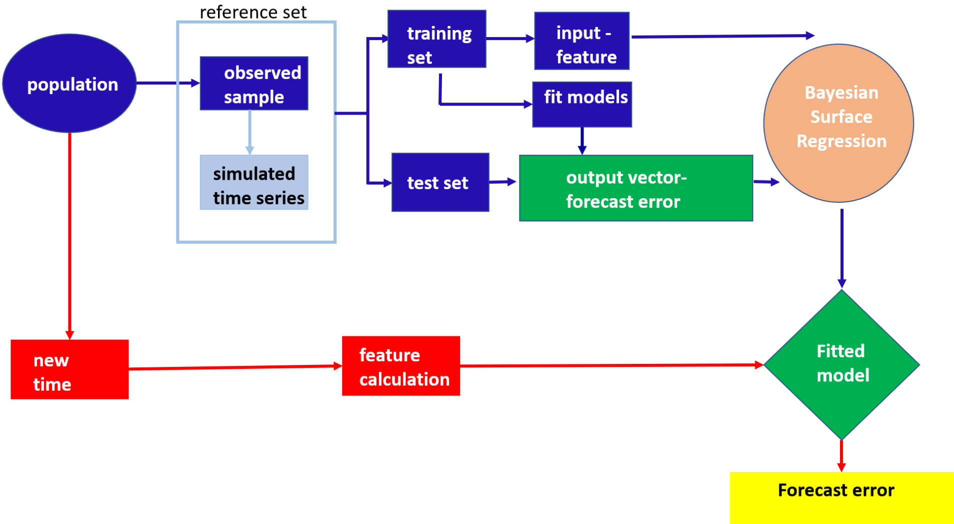
FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.
- Optimization criterion: forecast accuracy
- The probability of each model being best is used to construct a model weight.
- A combination forecast is produced using these weights.
- 248 registrations, 50 submissions

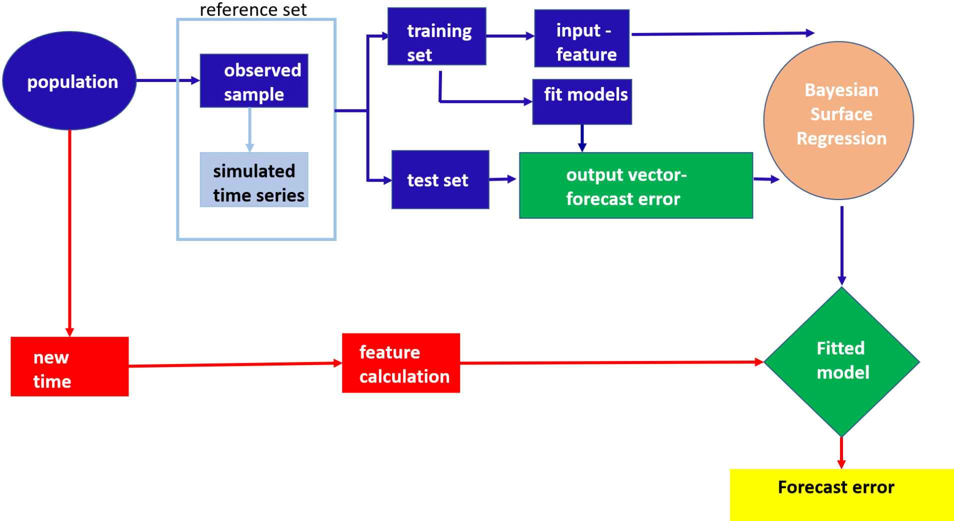
FFORMA: Feature-based FORecast Model Averaging

- Like FFORMS but we use xgboost rather than a random forest.
- Optimization criterion: forecast accuracy
- The probability of each model being best is used to construct a model weight.
- A combination forecast is produced using these weights.
- 248 registrations, 50 submissions
- Came second in the M4 competition

FFORMPP: Feature-based FORecast Model Performance Prediction



FFORMPP: Feature-based FORecast Model Performance Prediction



- use the minimum predicted MASE to select forecast method(s)

FFORMPP: Feature-based FORecast Model Performance Prediction

- We use Efficient Bayesian Multivariate Surface Regression

FFORMPP: Feature-based FORecast Model Performance Prediction

- We use Efficient Bayesian Multivariate Surface Regression
- **Y:** forecast error of each method, we take the correlation structure between the errors into account.

FFORMPP: Feature-based FORecast Model Performance Prediction

- We use Efficient Bayesian Multivariate Surface Regression
- **Y:** forecast error of each method, we take the correlation structure between the errors into account.
- Why Efficient Bayesian Multivariate Surface Regression?
  - handles interactions and nonlinear relationships between features
  - allows the knot locations to move freely in the feature space, thus a less number of knots are usually used

Application to M4 Competition

- Composition of the time series in the reference set and collection of new time series

Frequency	Reference set			New series
	M1	M3	Simulated	M4
Yearly	181	645	10000	23000
Quarterly	203	756	10000	24000
Monthly	617	1428	10000	48000
Weekly	-	-	10000	359
Daily	-	-	10000	4227
Hourly	-	-	10000	414

Application to M4 Competition

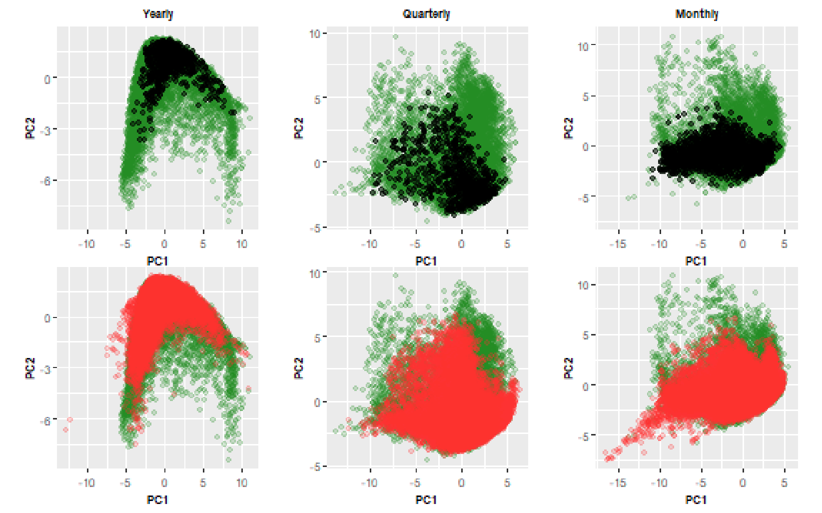
- Composition of the time series in the reference set and collection of new time series

Frequency	Reference set			New series
	M1	M3	Simulated	M4
Yearly	181	645	10000	23000
Quarterly	203	756	10000	24000
Monthly	617	1428	10000	48000
Weekly	-	-	10000	359
Daily	-	-	10000	4227
Hourly	-	-	10000	414

Simulated data

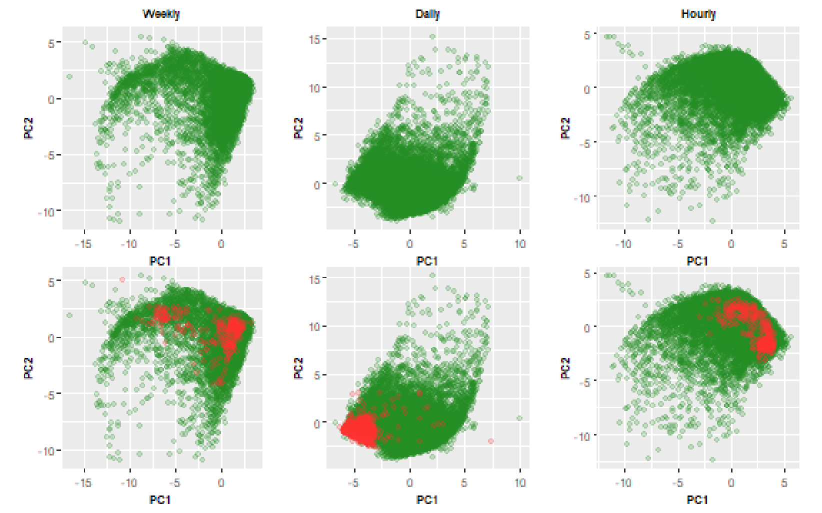
- Augmenting the observed sample with simulated time series from mixture autoregressive (MAR) models
- Proposed by Kang, Hyndman and Li (2018), Efficient generation of time series with diverse and controllable characteristics
- R package: https://github.com/ykang/tsgeneration

Distribution of time series in the PCA space



- black: observed, green-simulated, orange-M4

Distribution of time series in the PCA space



- green-simulated, orange-M4

Results: forecast accuracy based on MASE

	Yearly	Quarterly	Monthly	Weekly	Daily	Hourly
FFORMPP-combination*	**3.07**	**1.13**	**0.89**	**2.46**	3.62	**0.96**
FFORMPP-individual	3.37	1.17	1.05	2.53	4.26	1.06
auto.arima	3.40	1.17	0.93	2.55	-	-
ets	3.44	1.16	0.95	-	-	-
theta	3.37	1.24	0.97	2.64	3.33	1.59
rwd	**3.07**	1.33	1.18	2.68	3.25	11.45
rw	3.97	1.48	1.21	2.78	**3.27**	11.60
nn	4.06	1.55	1.14	4.04	3.90	1.09
stlar	-	2.02	1.33	3.15	4.49	1.49
snaive	-	1.66	1.26	2.78	24.46	2.86
tbats	-	1.19	1.05	2.49	**3.27**	1.30
wn	13.42	6.50	4.11	49.91	38.07	11.68
mstlarima	-	-	-	-	3.84	1.12
mstlets	-	-	-	-	3.73	1.23
combination (median)	3.29	1.22	0.95	2.57	3.52	1.33
combination (mean)	4.09	1.58	1.16	6.96	7.94	3.93

FFORMPP-combination*: based on median forecasts of the four algorithms with minimum predicted MASE

Results: FFORMS, FFORMA, FFORMPP

- forecast accuracy based on MASE

	Yearly	Quarterly	Monthly	Weekly	Daily	Hourly
FFORMPP-combination	3.07	1.13	**0.89**	2.46	3.62	0.96
FFORMA-combination	**3.06**	**1.11**	**0.89**	**2.10**	**3.34**	**0.81**
FFORMPP-individual	3.37	1.17	1.05	2.53	4.26	1.06
FFORMS-individual	3.16	1.20	0.98	2.31	3.56	9.33

Algorithm complexity and time to calculate forecasts

FFORMS < FFORMPP < FFORMA

Discussion and Conclusions

- Three feature-based algorithms for large scale forecasting.

Discussion and Conclusions

- Three feature-based algorithms for large scale forecasting.
- Each of these algorithms uses a meta-learning approach to guide the way the forecasts are computed.

Discussion and Conclusions

- Three feature-based algorithms for large scale forecasting.
- Each of these algorithms uses a meta-learning approach to guide the way the forecasts are computed.
- Future directions:
  - Combination of FFORMS, FFORMA, FFORMPP
  - Density forecast

R packages and papers

R packages

- **seer**: FFORMS
  github.com/thiyangt/seer
- **M4metalearning**: FFORMA.
  github.com/robjhyndman/M4metalearning
- **fformpp**: FFORMPP
  github.com/thiyangt/fformpp

Papers

Available from **robjhyndman.com**

email: thiyanga.talagala@monash.edu