

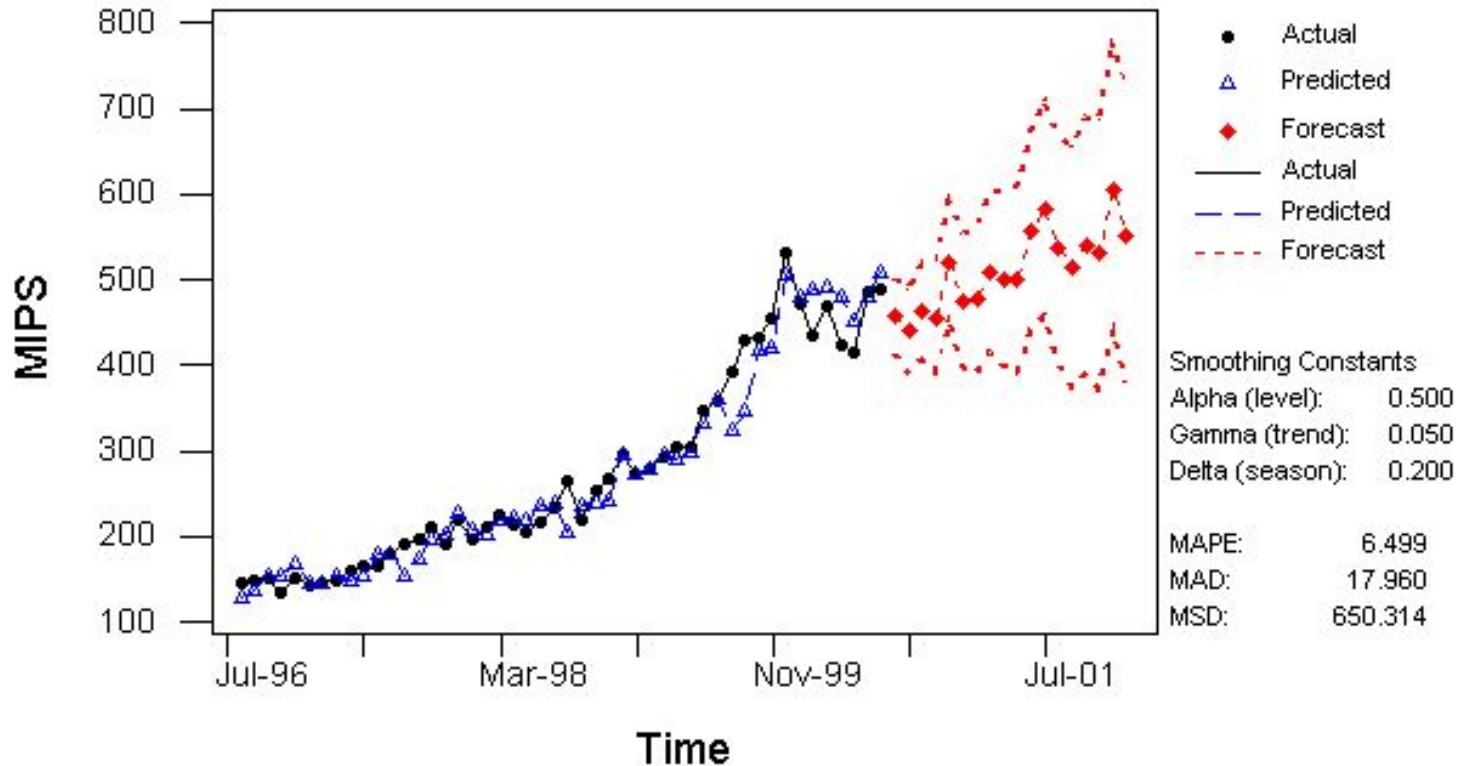
Time Series Forecast

...

Feb 03, 2016

minhtran & khangpham

Time Series Forecast at a Glance



Agenda

- Problem Statement
- Concepts
- Scope
- Methodology
- Meta Algorithm
- Forecast APIs & Demo
- Experimental Results
- Conclusions and Future Work

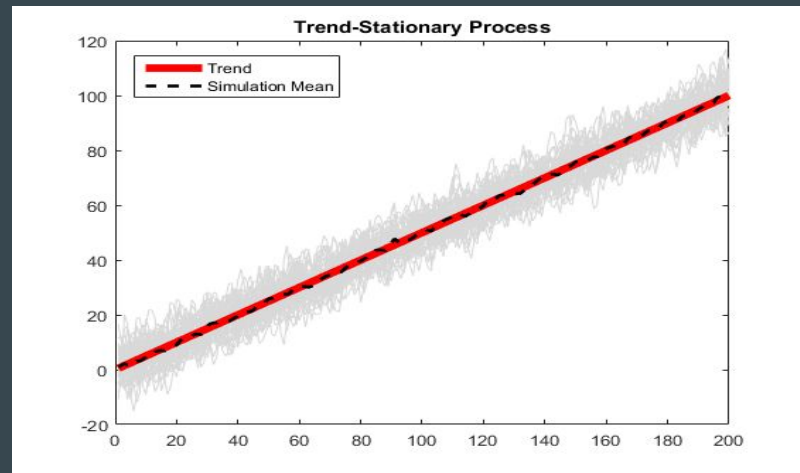
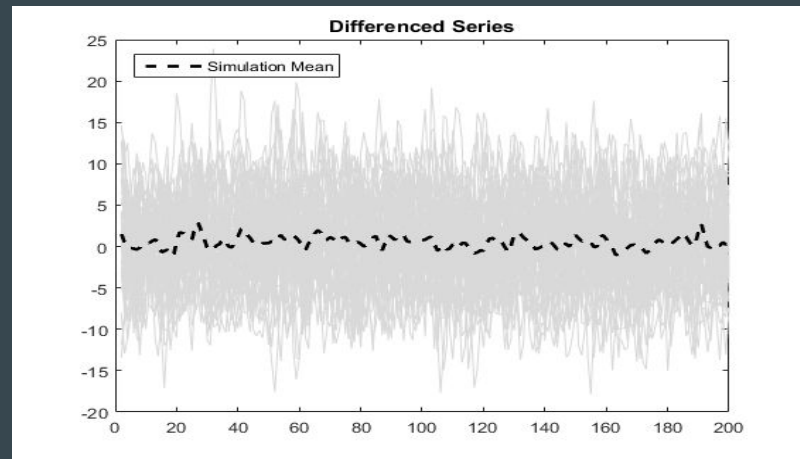
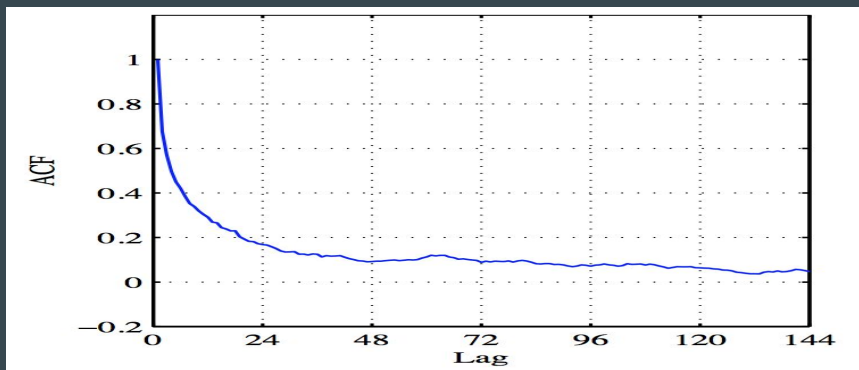
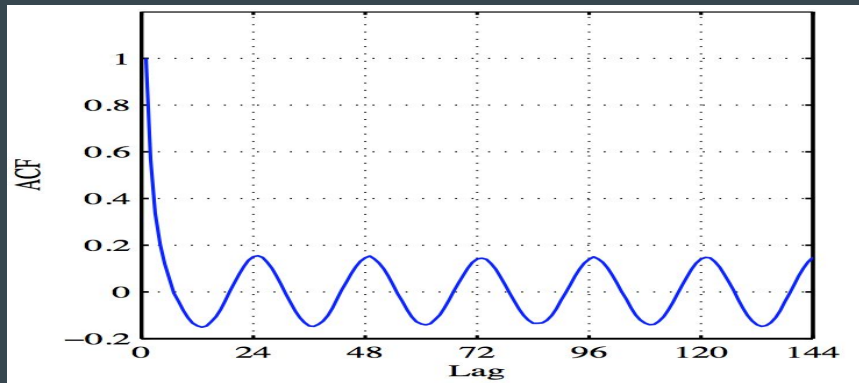
Problem Statement

- Demands from customers
- First version of a time series forecaster is implemented in BigApp
- Lots of disadvantages
 - Limited simple algorithm
 - Low accuracy
 - No choice for customers
 - Lack of data analysis
 - Lack of explanation to customers
 - Lack of auto-support to customers
- A new time series forecaster needs to be researched, developed and integrated into BigApp

New Time Series Forecaster

- Designed based on characteristics of a time series
- Designed to overcome limitations of the current forecaster
 - Provide several different core algorithms, each corresponds to a different characteristic
 - High accuracy
 - Provide several choices for customers
 - Auto-support to customers based on data analysis
 - Provide explanation to customers

Concepts

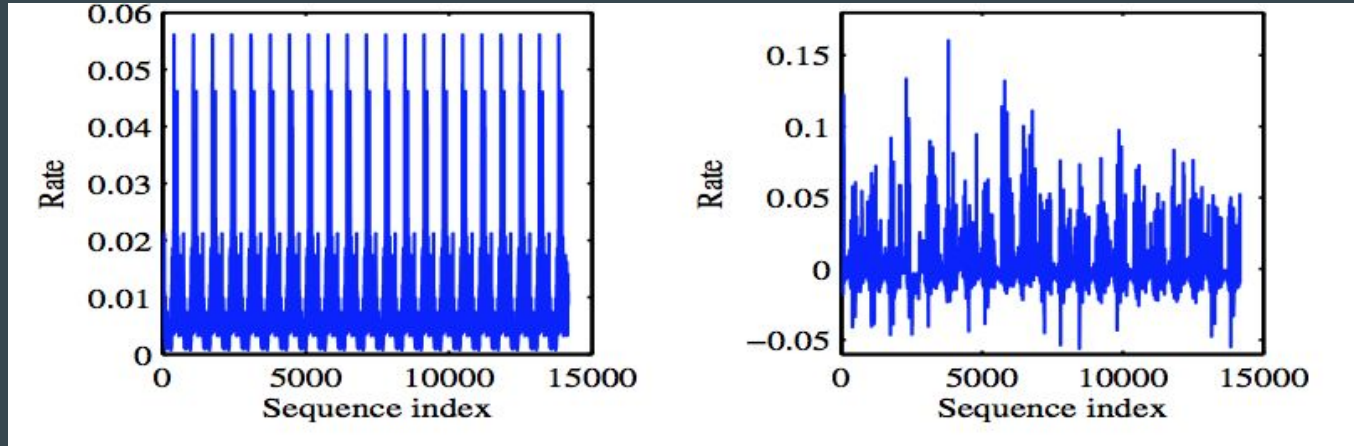


Scope

- We focus on the following 3 characteristics:
 - periodicity
 - trend-stationary
 - level-stationary
- Assumption: data set consists of only a time series (single target variable without other predictors)

Methodology

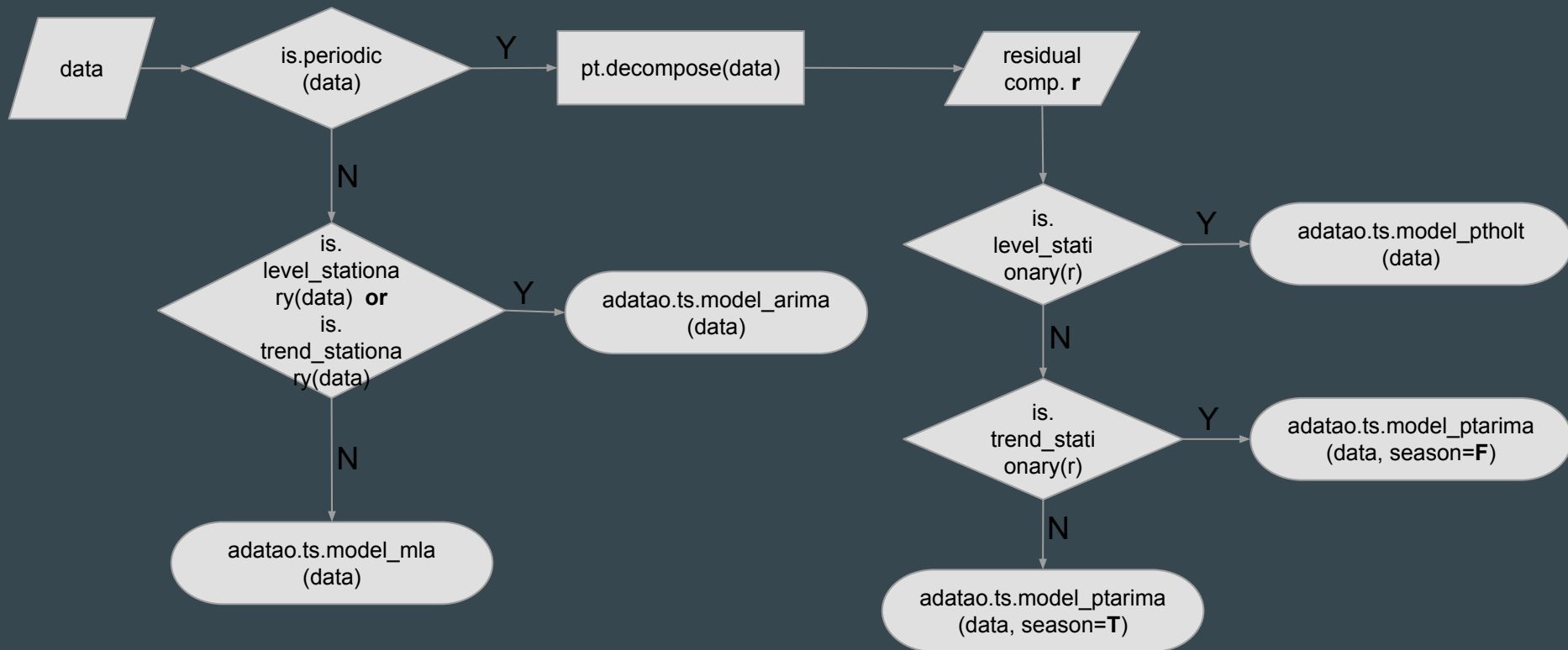
- We based our forecaster on
 - the arima algorithm
 - the holt algorithm
 - machine learning approach (svr - ongoing)
 - the theory of periodicity transform, a third kind of transform in addition to fourier transform and wavelet transform



Methodology

- Many choices for customers by core algorithms
 - periodicity transform + arima (season = true) \Rightarrow ptarima(season = true)
 - periodicity transform + arima (season = false) \Rightarrow ptarima(season = false)
 - periodicity transform + holt \Rightarrow ptholt
 - arima (auto.arima, the advanced version of the arima family)
- Auto support by a meta algorithm
 - analyse time series characteristics
 - explain to customers why a core algorithm should be selected for forecasting

Meta Algorithm



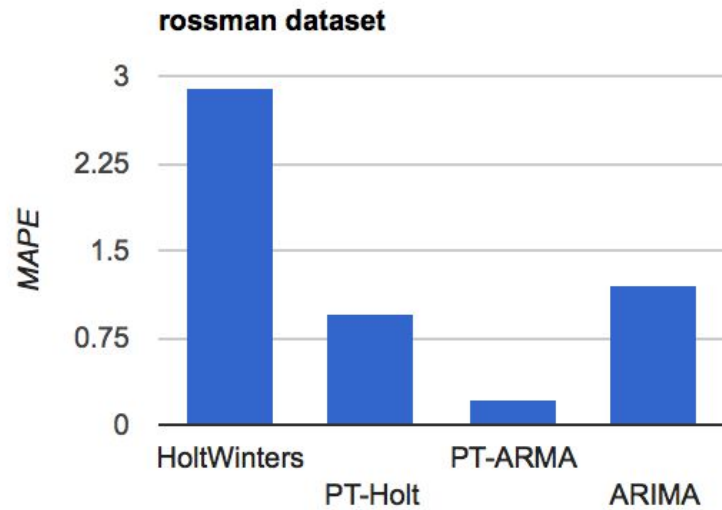
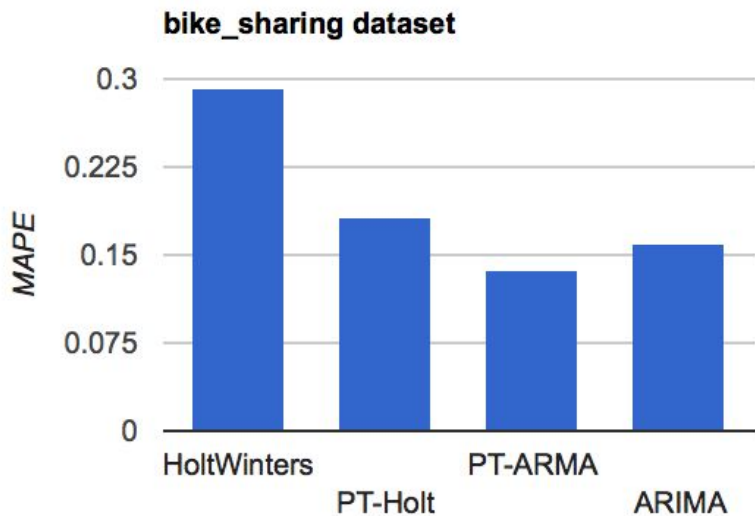
Forecast APIs and Demo

- A generic forecaster with the meta algorithm
 - `adatao.ts_train`
- Select a specific forecaster
 - `adatao.ts.model_arima`
 - `adatao.ts.model_ptarima`
 - `adatao.ts.model_ptholt`
 - `adatao.ts.model_mla`
- Forecast with a model
 - `adatao.ts_forecast`
 - `adatao.ts.predict_mla`
- Demo
 - `rossman`

Experimental Results

DataSet	Number of data points	Train/Test 7 days	Train/Test 10 days	Train/Test 30 days
bike_sharing	456	449/7	446/10	426/30
rossman	942	935/7	932/10	912/30
mobivi	205	198/7	195/10	175/30
flightinfo	7763	7756/7	7753/10	7733/30
nyctaxi	730	723/7	720/10	700/30
amazon_bills	365	358/7	355/10	335/30

Test Days = 30



Full Results

Shared in google doc: [https://docs.google.](https://docs.google.com/spreadsheets/d/1_9R8rduLeNldlX_UEmup5QyZLmZ5tdoUt589fEKvNxA/edit#gid=0)

[com/spreadsheets/d/1_9R8rduLeNldlX_UEmup5QyZLmZ5tdoUt589fEKvNxA/edit#gid=0](https://docs.google.com/spreadsheets/d/1_9R8rduLeNldlX_UEmup5QyZLmZ5tdoUt589fEKvNxA/edit#gid=0)

Integrate to BigApp



Auto-Support

DataSet	Recommended by adatao.ts_train	Best algorithm
amazon_bills	adatao.ts.model_arma	adatao.ts.model_arma
bike_sharing	adatao.ts.model_ptarima	adatao.ts.model_ptarima
rossman	adatao.ts.model_ptarima	adatao.ts.model_ptarima
mobivi	adatao.ts.model_ptarima / adatao.ts.model_arma	adatao.ts.model_ptarima / adatao.ts.model_arma
flightinfo	adatao.ts.model_ptarima	adatao.ts.model_arma
nyctaxi	adatao.ts.model_ptarima	adatao.ts.model_ptholt

Conclusions and Future Work

- A new time series forecaster was developed
 - Support multiple algorithms
 - Auto-support in recommending algorithms to customers
 - High accuracy
- Work to improve
 - Focus more on the long-range dependence of the time series
 - Take into account machine learning approaches
 - Mine more into the utilization of the periodicity transform theory

Thank you !!