Business forecasting using hybrid approach: Forecasting using deep learning and time series methods

Larry Orimoloye

Pasi Helenius

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

THE POWER TO KNOW.

Business Forecasting

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Business Forecasting

Cost Reduction

- Reduced Inventory
- Reduced Waste and Scrap
- Less Downtime
- Automation and Operational Efficiency

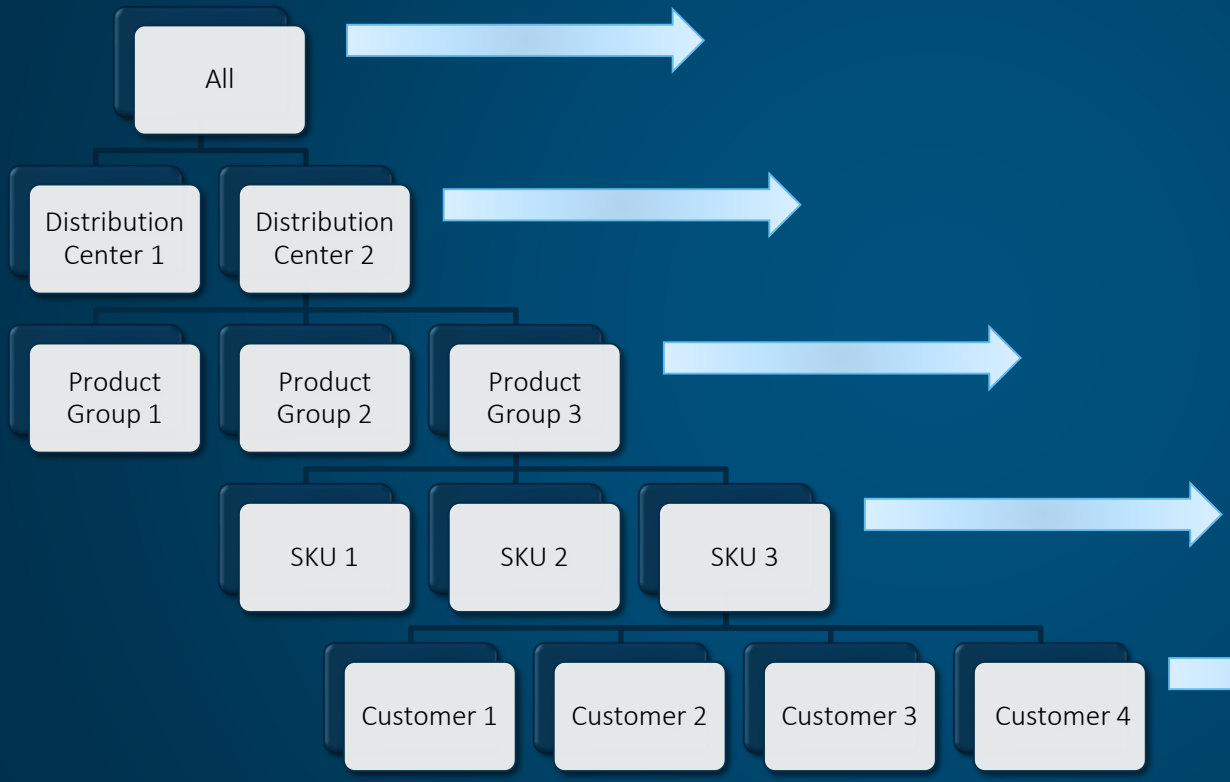
Increased Sales

- Less Out of Stock
- Improved Customer Satisfaction

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Hierarchical Forecasting

Example Hierarchy

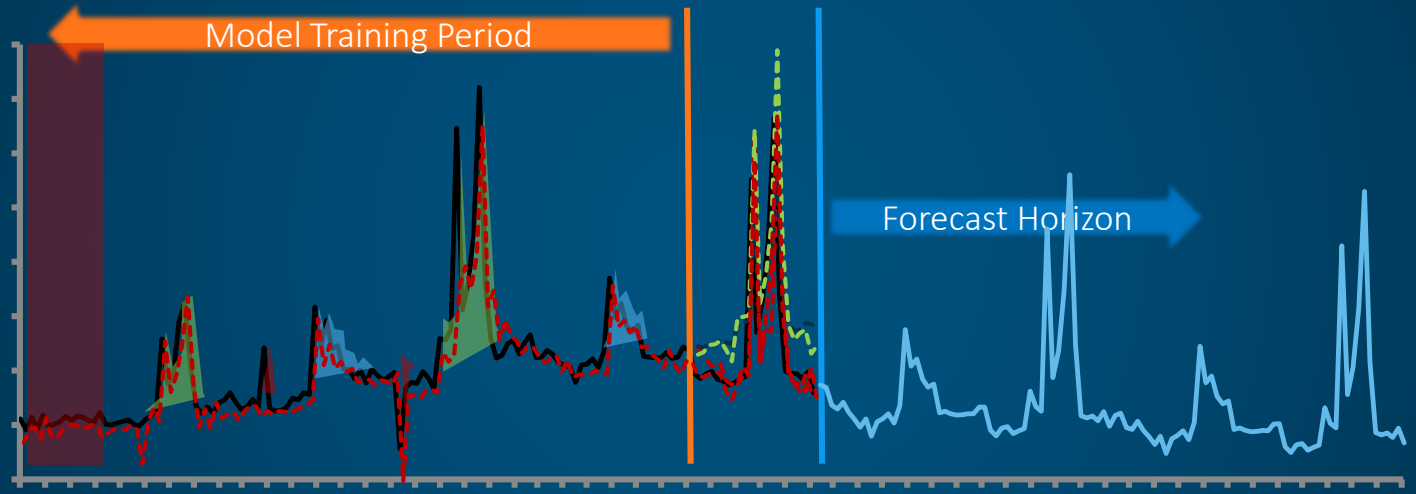


- Different Forecast models created at ***each*** level of the hierarchy for ***each*** accumulated timeseries
- Forecasts ***reconciled*** to one another

Copyright © SAS Institute Inc. All rights reserved.

Automated Large Scale Time Series Forecasting

6



Error	
13.64%	Model1: *ARIMA*, f(History, Outliers, Price, Promotions, Inventory, Christmas, Black Friday, Catalog)
11.05%	Model2: *Exponential Smoothing*, f(History, Seasonality)
8.58%	Model3: *UCM*, f(History, Outliers, Price, Christmas, Catalog)

Trend

Seasonality

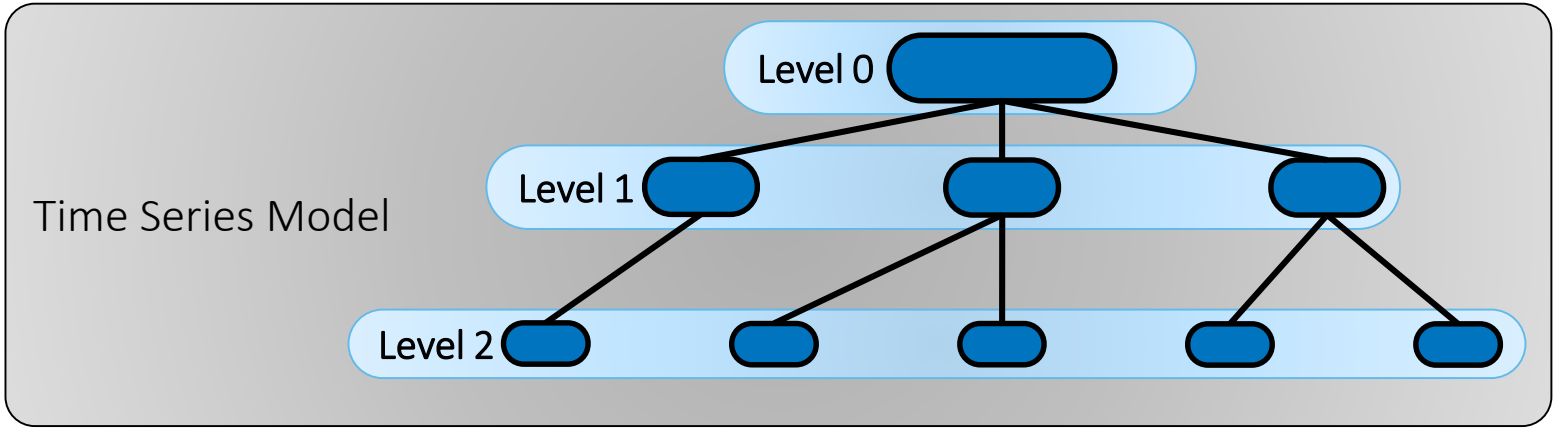
Outliers

Copyright © SAS Institute Inc. All rights reserved.

Time Series Approach

Hierarchical Forecasting Strategy

- Model all levels with time series models
  - ARIMA
  - ESM
  - UCM



Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Time Series Approaches

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

ARIMA (Box-Jenkins)

AutoRegressive

Integrated

Moving Average

AR I MA

Trend

$$y_t = \mu_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + e_t$$

P's

Effect of error terms

$$y_t = \mu_0 + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}$$

Q's

$$y_t - y_{t-1} = z_t$$

D's

Differencing

$$z_t = \mu_0 + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

Extending ARIMA

- Seasonality (S)

- Input Variables (X)
  - Promotion, discount, GDP, weather

$$z_t = \mu_0 + \beta x_t + \phi_1 z_{t-1} + \phi_2 z_{t-2} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-22}$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Coefficients ⇨ Transfer Functions

- Ordinary Regression

$$X_{1,t} \Rightarrow \beta_1 \Rightarrow Y_t$$

Time Series Regression

$$X_{1,t} \Rightarrow \frac{\omega(B)}{\delta(B)} \Rightarrow Y_t$$

Transfer Functions

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

$$(1-B)^d(1-B^S)^D g(Y_t)$$

$$= \mu + \sum_{i=1}^{k} B^{m_i}(1-B)^{d_i}(1-B^S)^{D_i} \frac{\omega_i(B)\omega_{S,i}(B^S)}{\delta_i(B)\delta_{S,i}(B^S)} f_i(X_{i,t}) + \frac{\theta(B)\Theta(B^S)}{\phi(B)\Phi(B^S)}\varepsilon_t$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

Trend

$$\boxed{(1-B)^d}(1-B^S)^D g(Y_t)$$

$$= \mu + \sum_{i=1}^{k} B^{m_i} \boxed{(1-B)^{d_i}} (1-B^S)^{D_i} \frac{\omega_i(B)\omega_{S,i}(B^S)}{\delta_i(B)\delta_{S,i}(B^S)} f_i(X_{i,t}) + \frac{\theta(B)\Theta(B^S)}{\phi(B)\Phi(B^S)} \varepsilon_t$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

Seasonal

$$(1-B)^d \boxed{(1-B^S)^D} g(Y_t)$$

$$= \mu + \sum_{i=1}^{k} B^{m_i} (1-B)^{d_i} \boxed{(1-B^S)^{D_i}} \frac{\omega_i(B)\boxed{\omega_{S,i}(B^S)}}{\delta_i(B)\boxed{\delta_{S,i}(B^S)}} f_i(X_{i,t}) + \frac{\theta(B)\boxed{\Theta(B^S)}}{\phi(B)\boxed{\Phi(B^S)}} \varepsilon_t$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

Stationary

$$(1-B)^d(1-B^S)^D g(Y_t)$$

$$= \mu + \sum_{i=1}^{k} B^{m_i}(1-B)^{d_i}(1-B^S)^{D_i} \frac{\omega_i(B)\omega_{S,i}(B^S)}{\delta_i(B)\delta_{S,i}(B^S)} f_i(X_{i,t}) + \frac{\theta(B)\Theta(B^S)}{\phi(B)\Phi(B^S)}\varepsilon_t$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

(S)ARIMA(X)

Input Variables

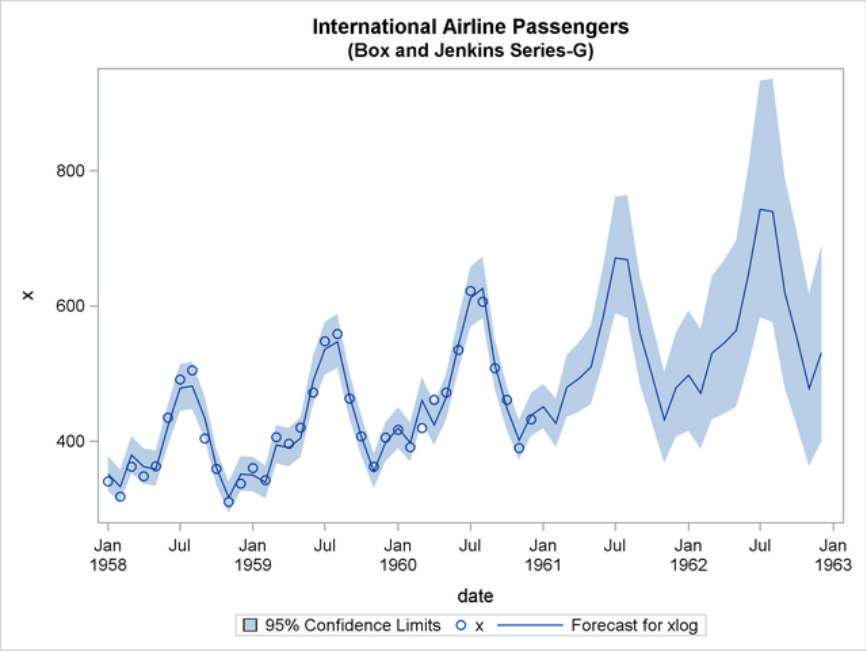
$$(1-B)^d(1-B^S)^D g(Y_t)$$

$$= \mu + \boxed{\sum_{i=1}^{k} B^{m_i}(1-B)^{d_i}(1-B^S)^{D_i}\frac{\omega_i(B)\omega_{S,i}(B^S)}{\delta_i(B)\delta_{S,i}(B^S)} f_i(X_{i,t})} + \frac{\theta(B)\Theta(B^S)}{\phi(B)\Phi(B^S)}\varepsilon_t$$

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

ARIMA(X)

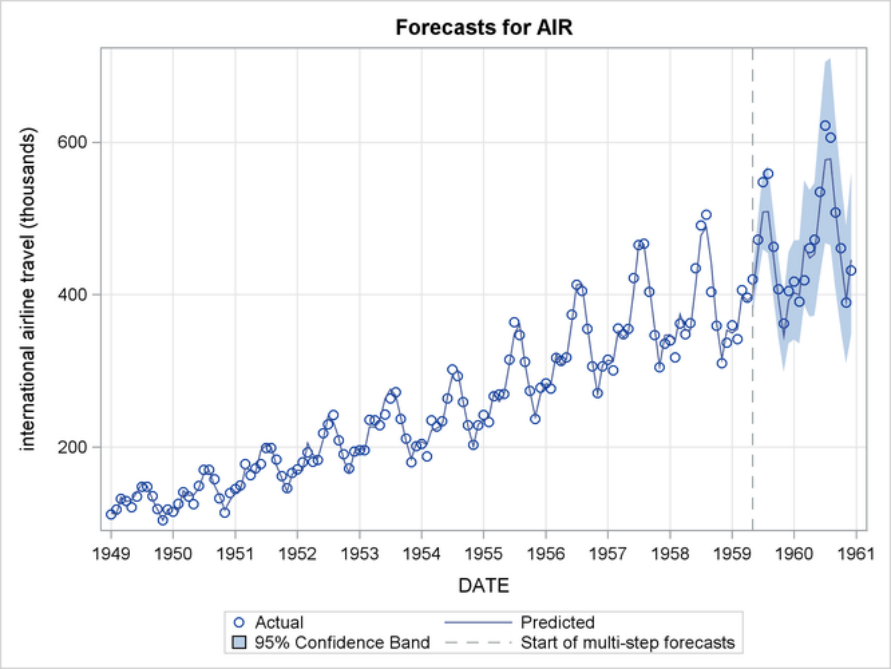
- ARIMA can be automed
  - Good packages available
- ARIMA can be very powerful
- Can be generalized to complex input – output relationships
- Challenges with many candidate inputs



Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Exponential Smoothing

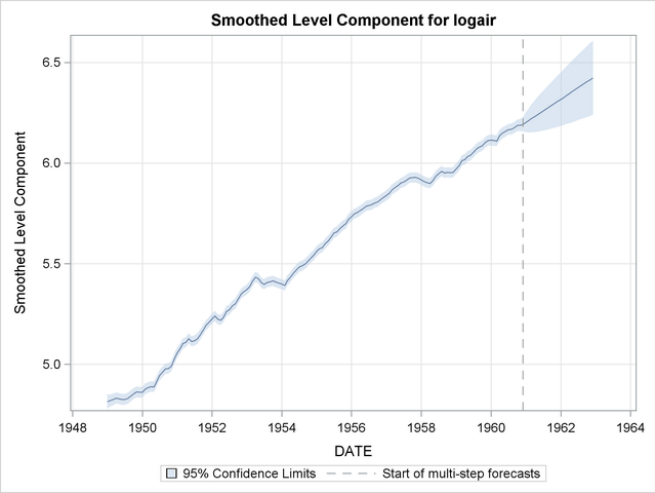
- Simple method
- Works well in some cases like short term and simple structure
- No input variables

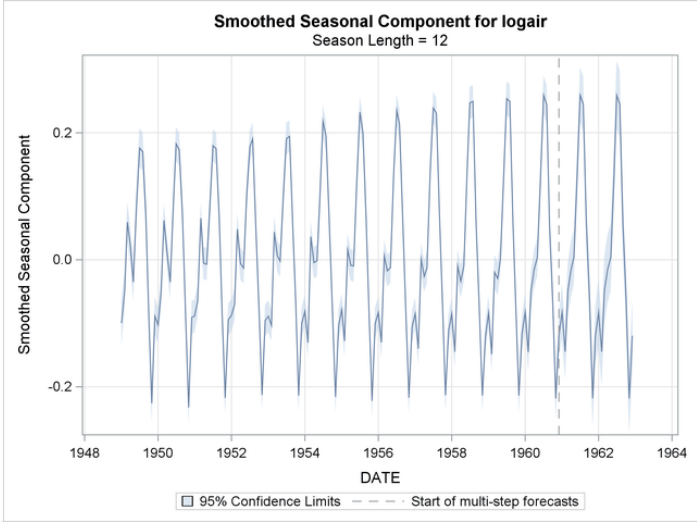


Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Unobserved Component Model

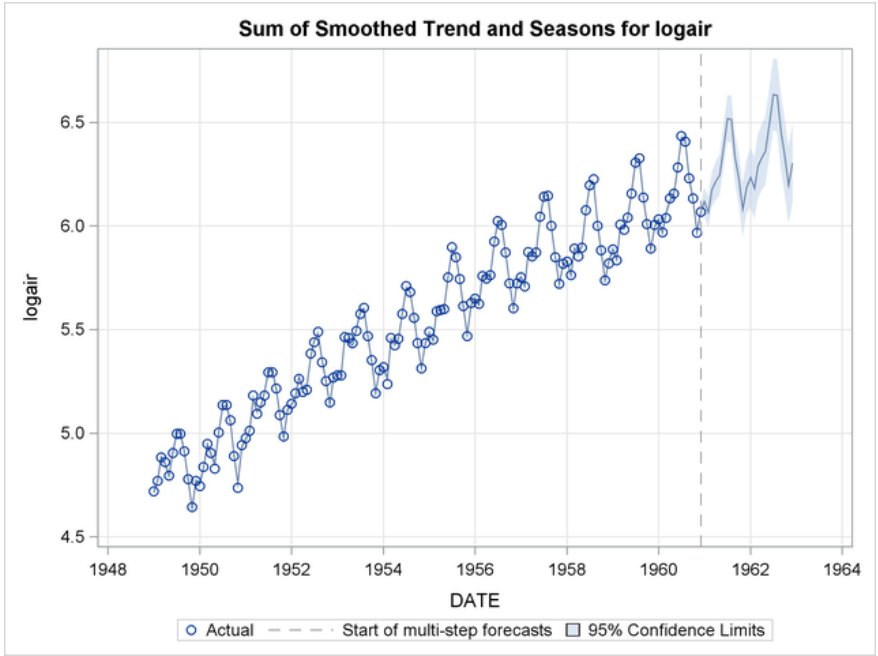
- Works with input variables
- Trend, seasonal and cyclical components are the *unobserved* components
- Decomposes time series and models components separately





Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

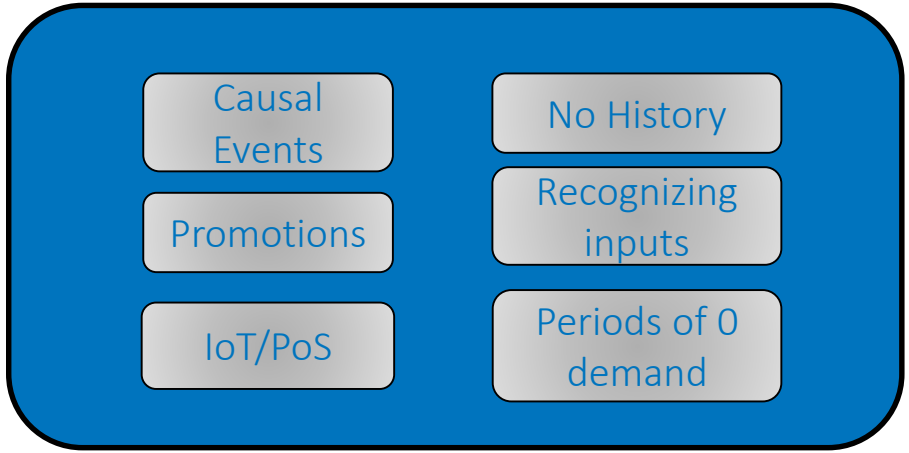
Unobserved Component Model



Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

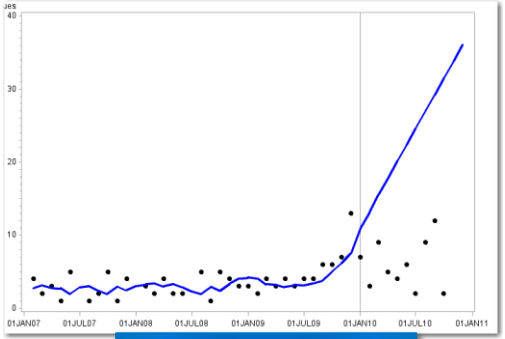
Challenges of Traditional methods

- Sparse and noisy data
- Variety and amount of input data
- Rise of Machine Learning methods

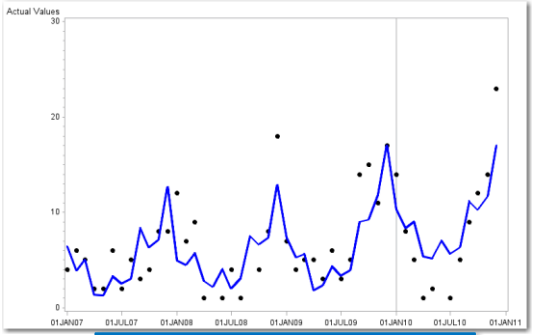


Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

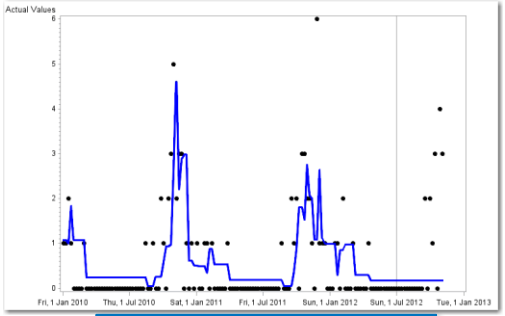
Examples



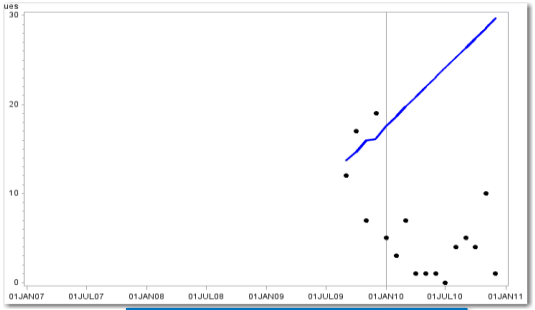
Year-round Basic



Year-round Seasonal



Highly Seasonal



Short

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Machine Learning vs Statistical approach

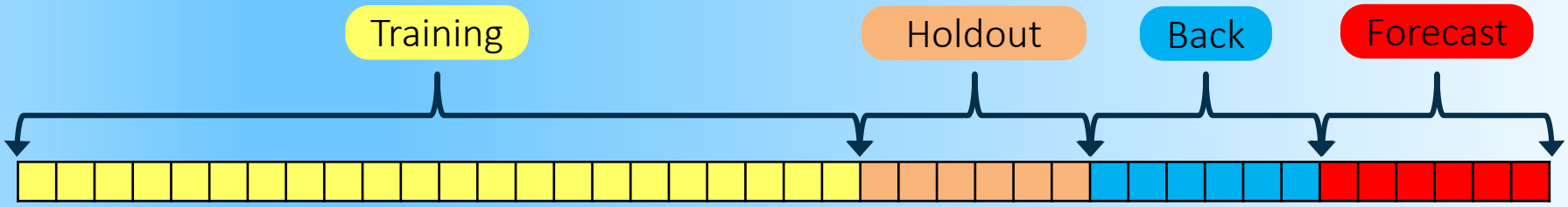
Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Forecasting with Neural Networks

Partitioning

Machine Learning: Random Sampling

Forecasting: Ordered Sampling

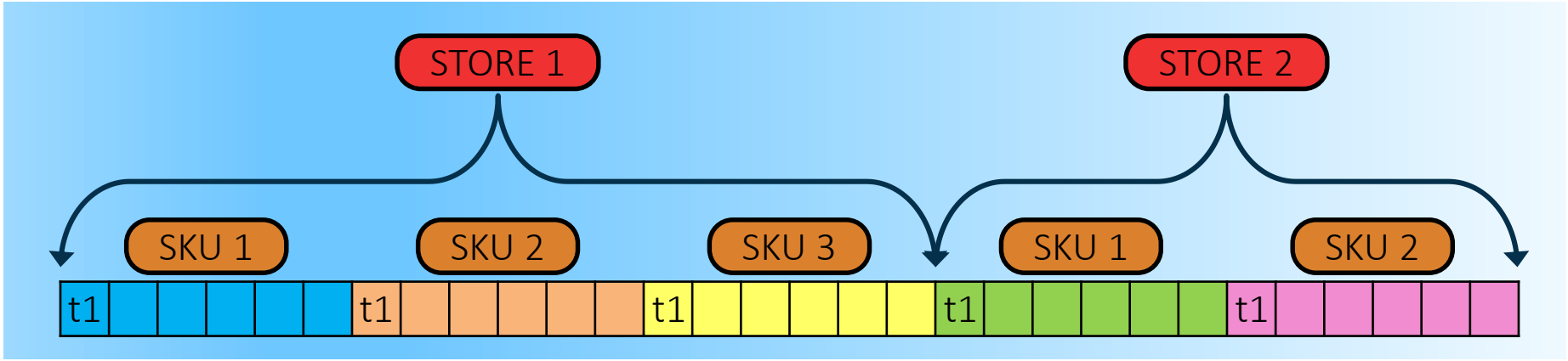


Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Forecasting with Neural Networks

BY Groups

Five time series delineated by the BY variables SKU and STORE



Time Series Approach: Generate a unique model for each series

Neural Net Approach: Generate one model and include BY variables as categorical input data

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

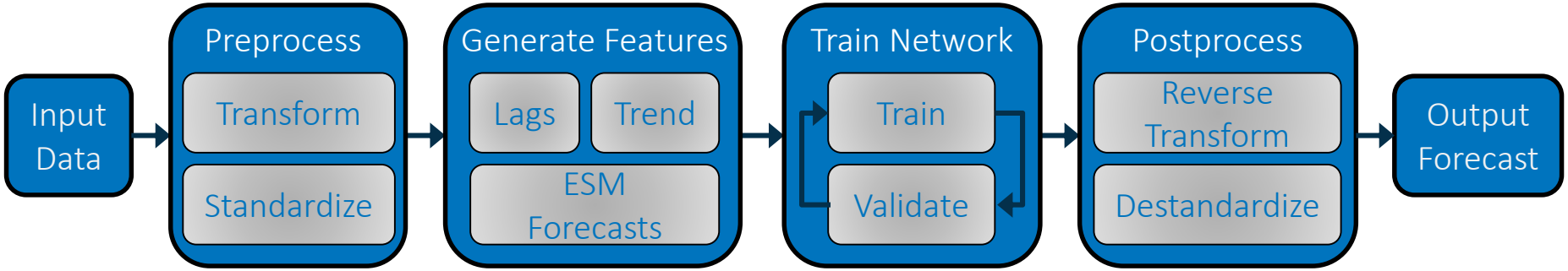
Panel Series Neural Network

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Panel Series Neural Network (PSNN)

PSNN Modeling Strategy

- Standardizes and/or transforms data
- Generates input features from time series data: lags, trend & Exponential Smoothing Model (ESM) forecasts
- Trains a fully connected multilayer perceptron neural network to generate forecast



Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

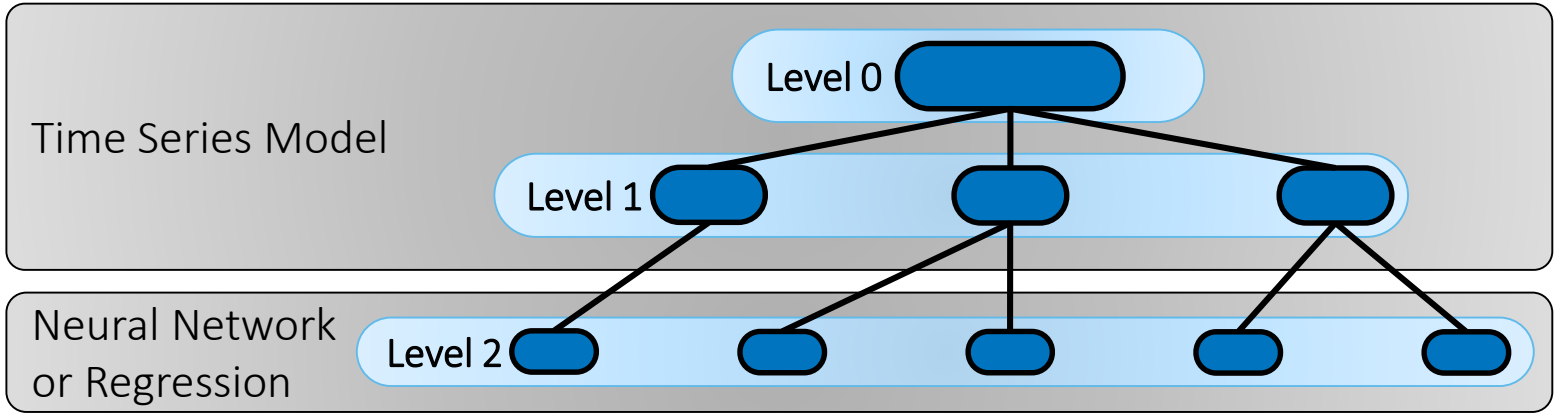
Multistage Forecasting

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Multistage Model

Multistage Forecasting Strategy

- Model lower levels of hierarchy where series may be more nonlinear with a neural network or regression
- Model higher levels with time series models
- Reconcile forecasts to produce final forecast



Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Multistage Model

Aggregation

- Reduce noise
- Solve data sparsity issue

Pooling data

- Borrowing information for estimation

Machine Learning

- Machine Learning at low level capture features from lower level time series

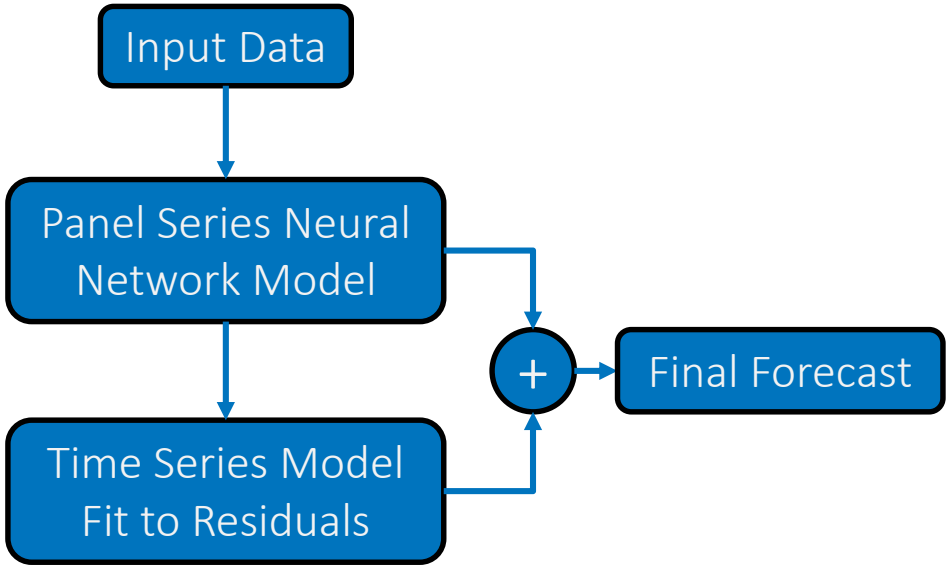
Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Stacked Forecasting

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Stacked Forecasting

Neural Network + Time Series



Stacked Modeling Strategy

1. Fit PSNN model to input data
2. Fit time series model to residuals from step 1
3. Add forecasts from each model together to obtain final forecast

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Demo

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

date	Sku	Store	Vendor	sales	Var1	Var2	Var3	Var4	V
24APR2011	Product2_auto	F26	Reseller_ID9	701	69	27081	427	143	6
01MAY2011	Product2_auto	F26	Reseller_ID9	728	66	27157	427	147	7
08MAY2011	Product2_auto	F26	Reseller_ID9	722	68	27200	427	145	7
15MAY2011	Product2_auto	F26	Reseller_ID9	702	69	27229	427	144	6
22MAY2011	Product2_auto	F26	Reseller_ID9	706	69	27282	427	144	6
29MAY2011	Product2_auto	F26	Reseller_ID9	729	65	27354	427	145	7
05JUN2011	Product2_auto	F26	Reseller_ID9	689	72	27363	427	142	6
12JUN2011	Product2_auto	F26	Reseller_ID9	735	66	27458	427	146	7
19JUN2011	Product2_auto	F26	Reseller_ID9	715	67	27487	427	145	7
26JUN2011	Product2_auto	F26	Reseller_ID9	661	73	27482	427	141	6

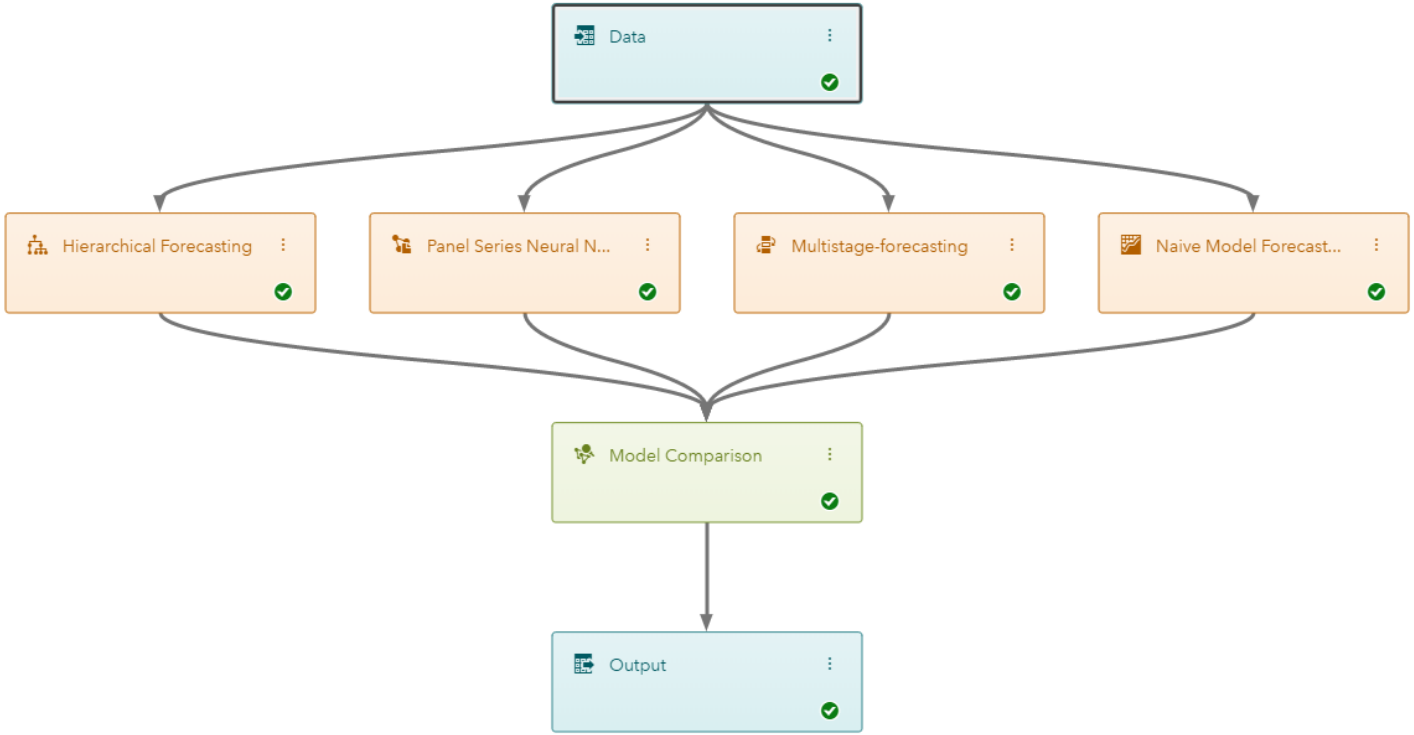
Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Model Studio - Build Models

Wine

Data | Pipelines | Pipeline Comparison | Overrides

Pipeline 1 | Hierarchical



Copyright © SAS Institute Inc. All rights reserved.

Hierarchical Forecasting

Description:

Generates forecasts for each level of the specified hierarchy.

- Task
  - Forecast task: Diagnose
- Diagnostics: Changes require re-dia...
- Model Generation: Changes require ...
  - ARIMAX
    - Identification order for inputs and events: Create two models, each of...
    - Refinement
      - Factor option: INPUT
      - Significance level: 0.4
  - ESM
    - ESM model: BEST
  - UCM
  - External Models

Panel Series Neural Network

Description:

Generate forecast using fully connected neural network model.

- Task
  - Forecast Task: Diagnose
- Feature Generation: Changes require r...
  - Dependent variable lags: 4
  - Independent variable lags: 4
  - [x] Seasonal dummy variables
  - [x] ESM Forecast of dependent variable
  - Dependent variable trend: Damped trend
- Model Generation: Changes require r...
  - Model Initialization: Changes requi...
  - Model Training: Changes require r...
  - Enable Autotune: Changes ...
- Model Validation: Changes require re-...
- Forecast Properties: Changes will upd...

Multistage-forecasting

Description:

Generate forecasts using multistage forecasting model.

Code editor: Open

Stage 1 high level: 1

Stage 1 low level:

Calculate statistics of fit over an out of sample range: 6

Extend the lead beyond the out of sample range: Yes

Size of data to be used for holdout: 0

Percentage of data to be used for holdout: 0

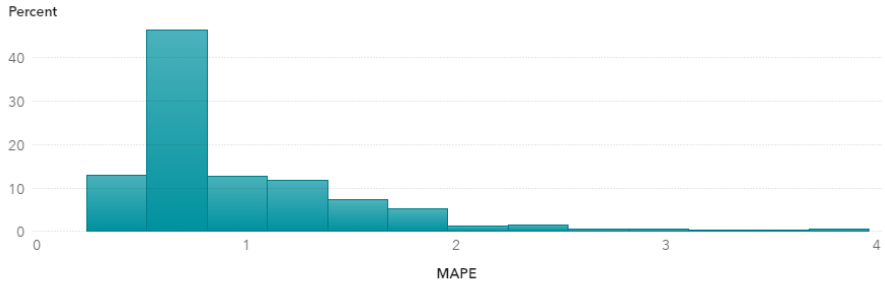
- Time series forecast model
  - Transformation

Specify whether to include ARIMAX model for diagnosis: Yes

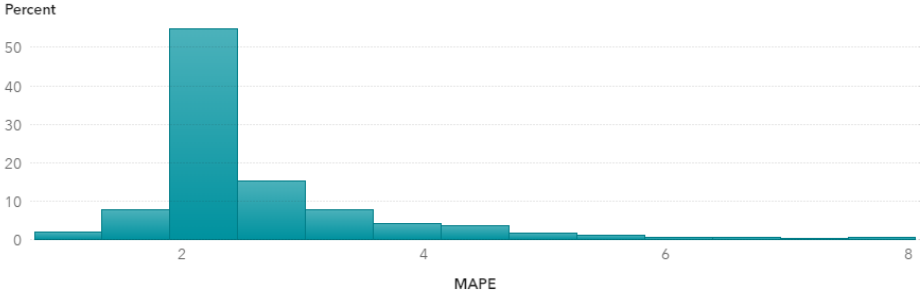
Specify whether to include ESM model

Champion	Model Name	Status	WMAE	WMAPE
	Multistage-forecasting	Successful	4.2062	0.5742
	Hierarchical Forecasting	Successful	16.2033	2.2743
	Panel Series Neural Network	Successful	8.6146	1.1818
	Naive Model Forecasting	Successful	21.3517	2.9845

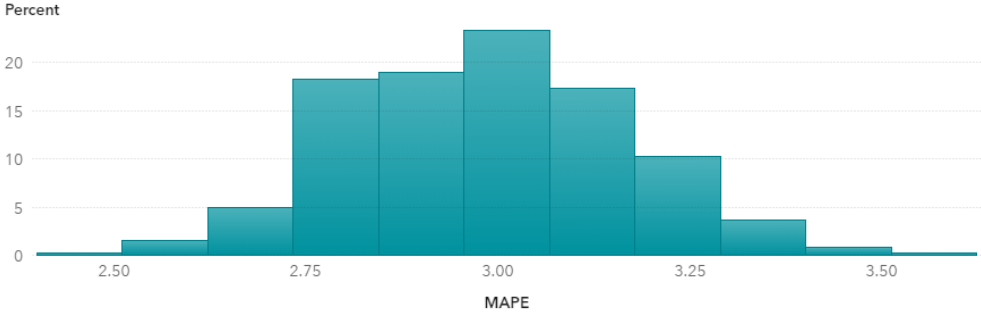
MAPE Distribution (Multistage-forecasting)



MAPE Distribution (Hierarchical Forecasting)



MAPE Distribution (Naive Model Forecasting)



A Hybrid Approach for Forecasting Modelling Strategies

Hierarchical Forecasting (ARIMA, ESM, UCM)

Panel Series Neural Network

Multistage Forecasting

Stacked Forecasting

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Further Reading

- A Multistage Modeling Strategy for Demand Forecasting:
  - http://support.sas.com/resources/papers/proceedings16/SAS5260-2016.pdf

- SAS Visual Forecasting – Product Page and Fact Sheet:

  - https://www.sas.com/en_us/software/visual-forecasting.html
  - https://www.sas.com/content/dam/SAS/en_us/doc/factsheet/sas-visual-forecasting-109226.pdf

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

Thank You!

sas.com

Company Confidential – For Internal Use Only
Copyright © SAS Institute Inc. All rights reserved.

THE POWER TO KNOW.