44th International Symposium On Forecasting | Dijon, France

Does Forecast Accuracy Even Matter – Findings from a Retail Dataset

Dr. Jim Hoover
Jack Faricy Professor
Director, Business Analytics & Artificial Intelligence Center
Warrington College of Business
University of Florida
jim.hoover@warrington.ufl.edu

UF BUSINESS ANALYTICS & ARTIFICIAL INTELLIGENCE CENTER

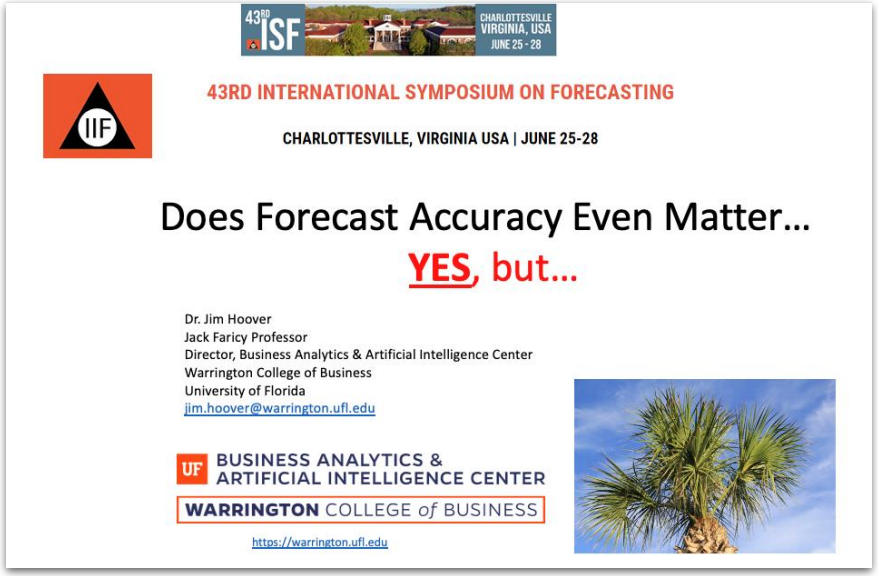
WARRINGTON COLLEGE of BUSINESS

https://warrington.ufl.edu

Does Forecast Accuracy Even Matter?

- Commercial discussion
  - Steutermann (2017) - Gartner
    - 1% improvement in forecast accuracy =
      - 2.7% inventory reduction
      - 3.2% transportation cost reduction
      - 3.9% inventory obsolescence reduction
- ISF 2023 – Charlottesville, VA
    - Methodology reviewed
    - Single SKU example presented
    - Next steps to include expansion to entire 30,490 SKUs from the M5 competition
    - “ABC” analysis

Foresight Issue 68, Spring 2023
Special Feature:
Does Forecast Accuracy Even Matter?

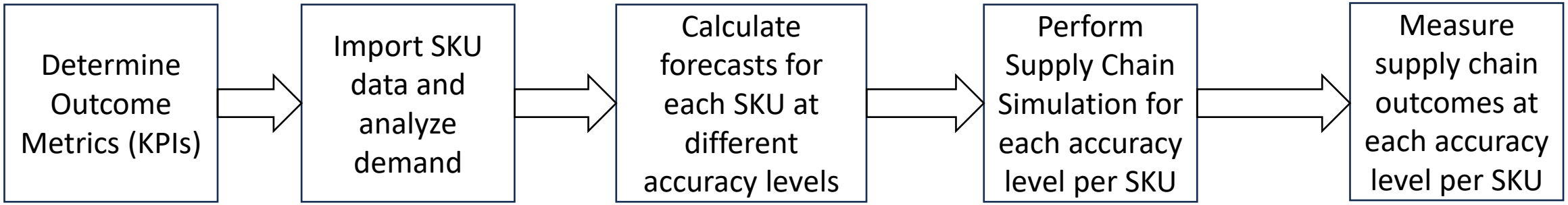


Does Forecast Accuracy Even Matter?

- Forecasting and S&OP processes
  - Expectations of management about forecast accuracy?
  - Limited human resources to perform forecast improvements
  - High number of SKU forecasts when measured at a store / SKU level
    - Walmart has over 4,600 stores and 600 Sam's Club locations in the U.S.[1]
    - Each store has an average of 142,000 SKUs[2]
    - Walmart has 210 distribution centers[3]
    - Forecasting SKUs monthly at the store level represents 738,000,000 forecasts
    - Forecasting SKUs across Walmart > 1 billion SKUs per month

1) https://www.cnbc.com/2024/01/31/walmart-plans-to-add-150-more-stores-across-us.html
2) https://inventory-management.com/inventory-management/inventory-management-at-walmart/
3) https://corporate.walmart.com/about

Simulation Methods to Determine Outcomes



- Metrics evaluated:
  - Average inventory levels
  - Backorder periods
  - Safety stock levels

- M5 competition data from Walmart
- 30,490 SKUs
- Determine A,B,C assignments

- Forecasts with errors were perturbated around actuals at 50%, 40%, 30%, 20%, 10%, 5%, 2%, 1%, and 0% error for each period

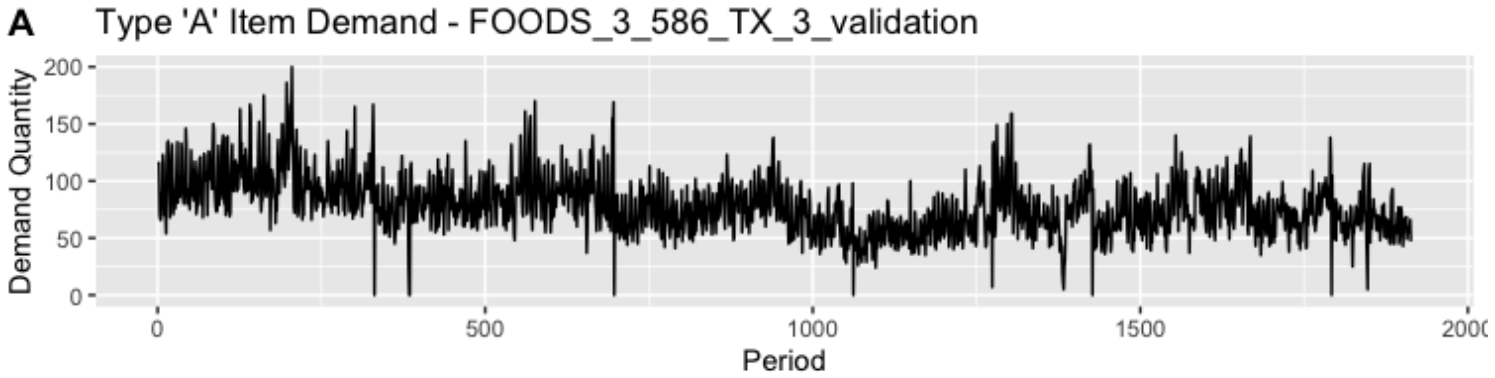
- Accuracy was calculated using Mean Absolute Deviation (MAD) to conform to supply chain calculations
- Outcomes simulate the SAP supply chain calculations

- Average inventory levels and backorder periods were compared to the forecast accuracy levels at each A,B,C category

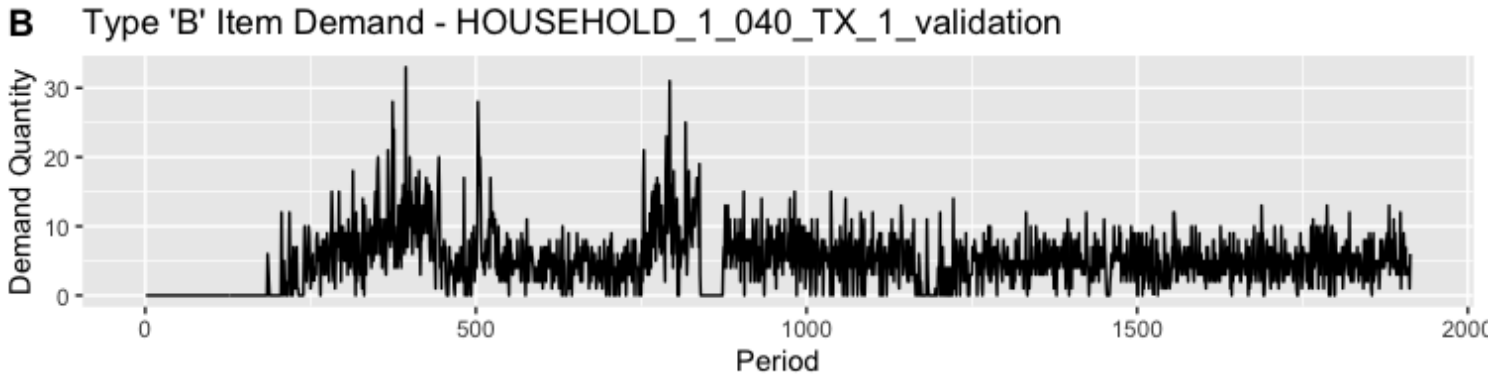
Data Profile and ABC Analysis

Example SKUs from each A,B, C level

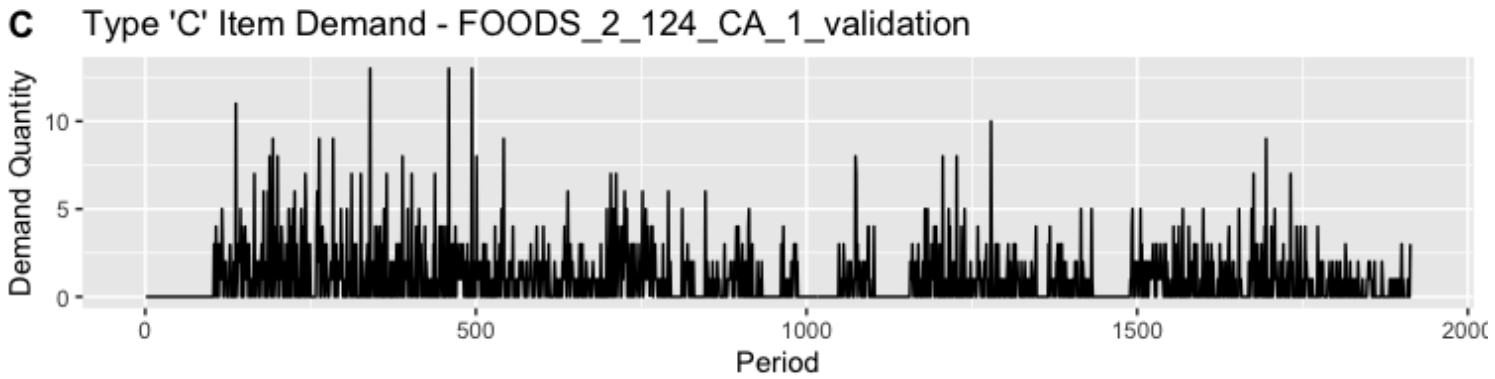
>=10



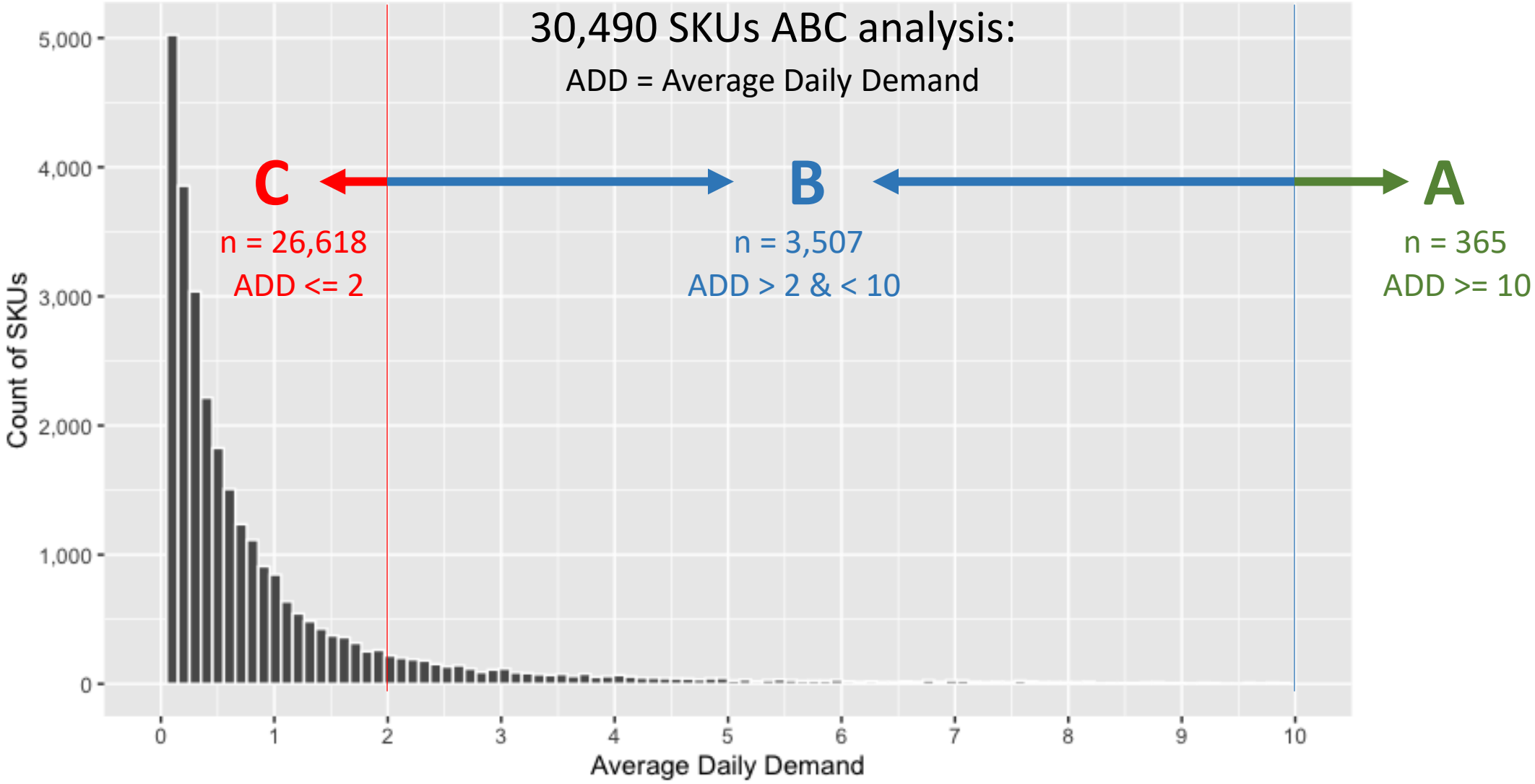
>2 & <10



<=2

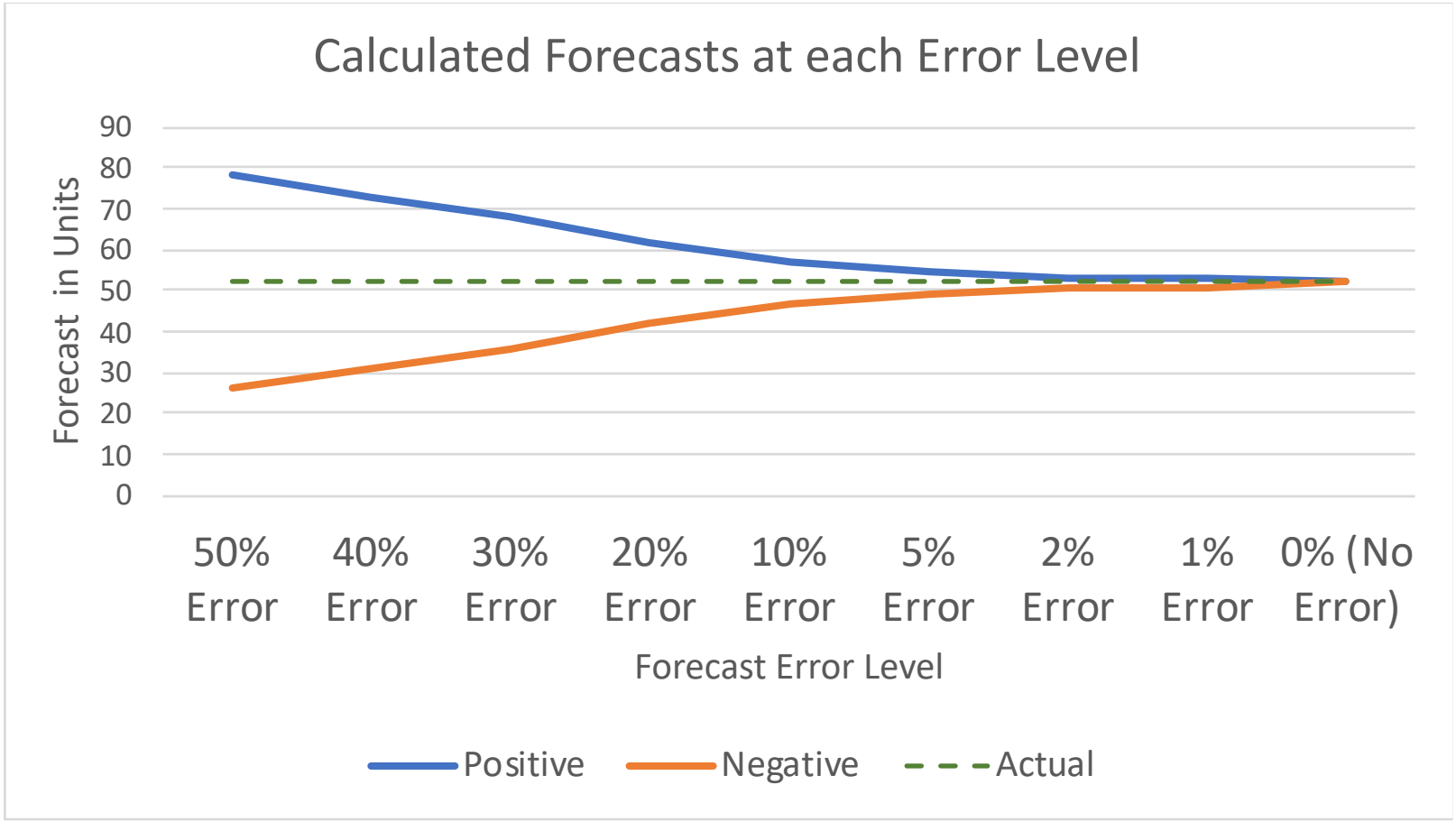


Data Profile and ABC Analysis



https://www.kaggle.com/competitions/m5-forecasting-accuracy/data

Simulating Forecast Levels of Accuracy



- By preserving forecasts at each error level for each period for each SKU, comparisons of different supply chain outcomes can be calculated

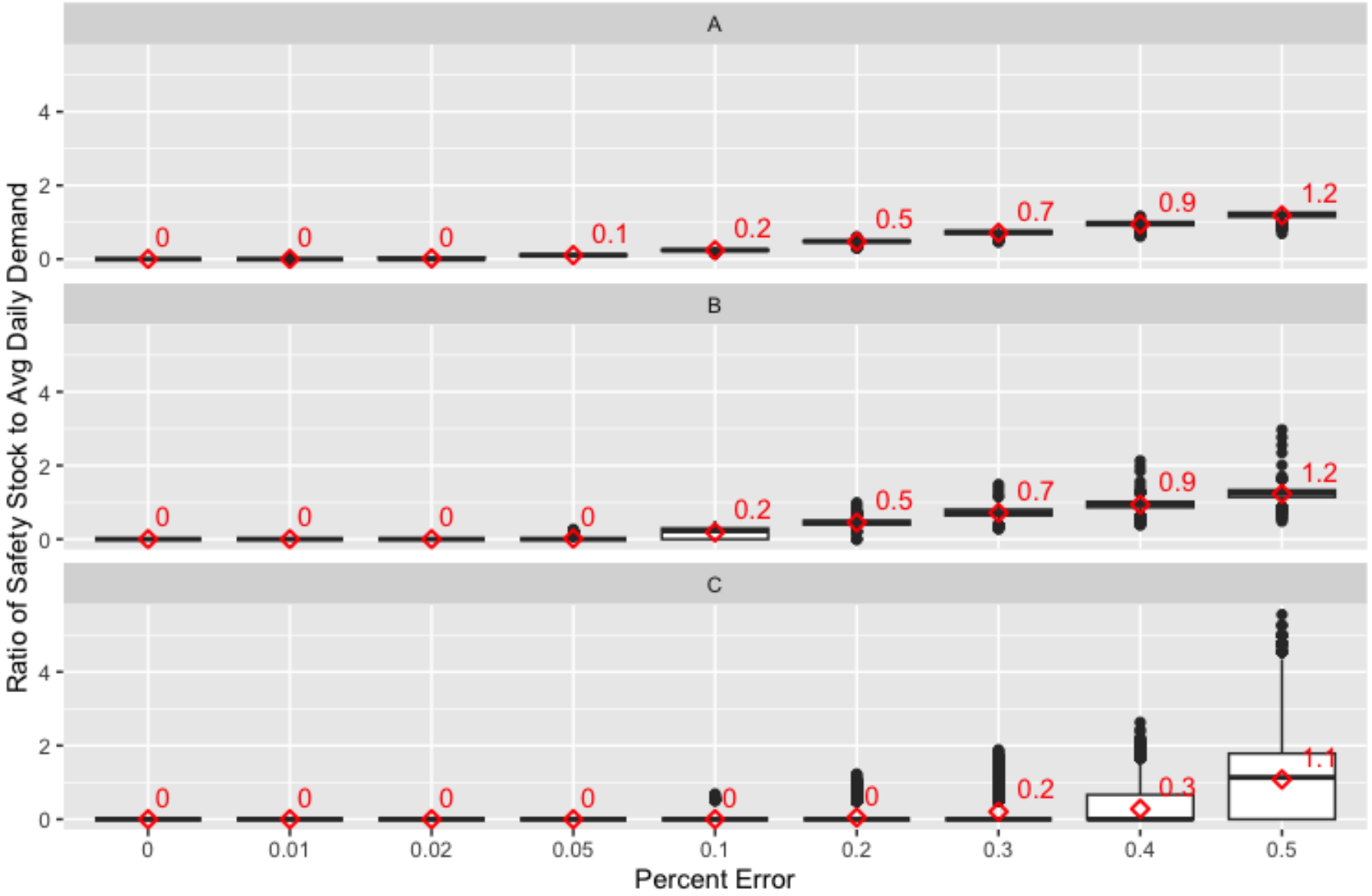
Supply Chain Simulation Assumptions

- Monthly levels calculations and use
  - Constant values for decisions updated periodically
  - Many firms update monthly
  - Why?
- SAP Supply Chain Calculations
  - ROQ – Optimal Order Quantity
    - Limits the number of orders trading off costs of transactions
  - Lot size
  - Reorder Review Period, Reorder Point, Inventory Position
  - Safety Stock
  - Service Level
  - Backorders permitted?

Forecast Error & Safety Stock Impact

Outcome Measure: Ratio of Safety Stock to Average Daily Demand

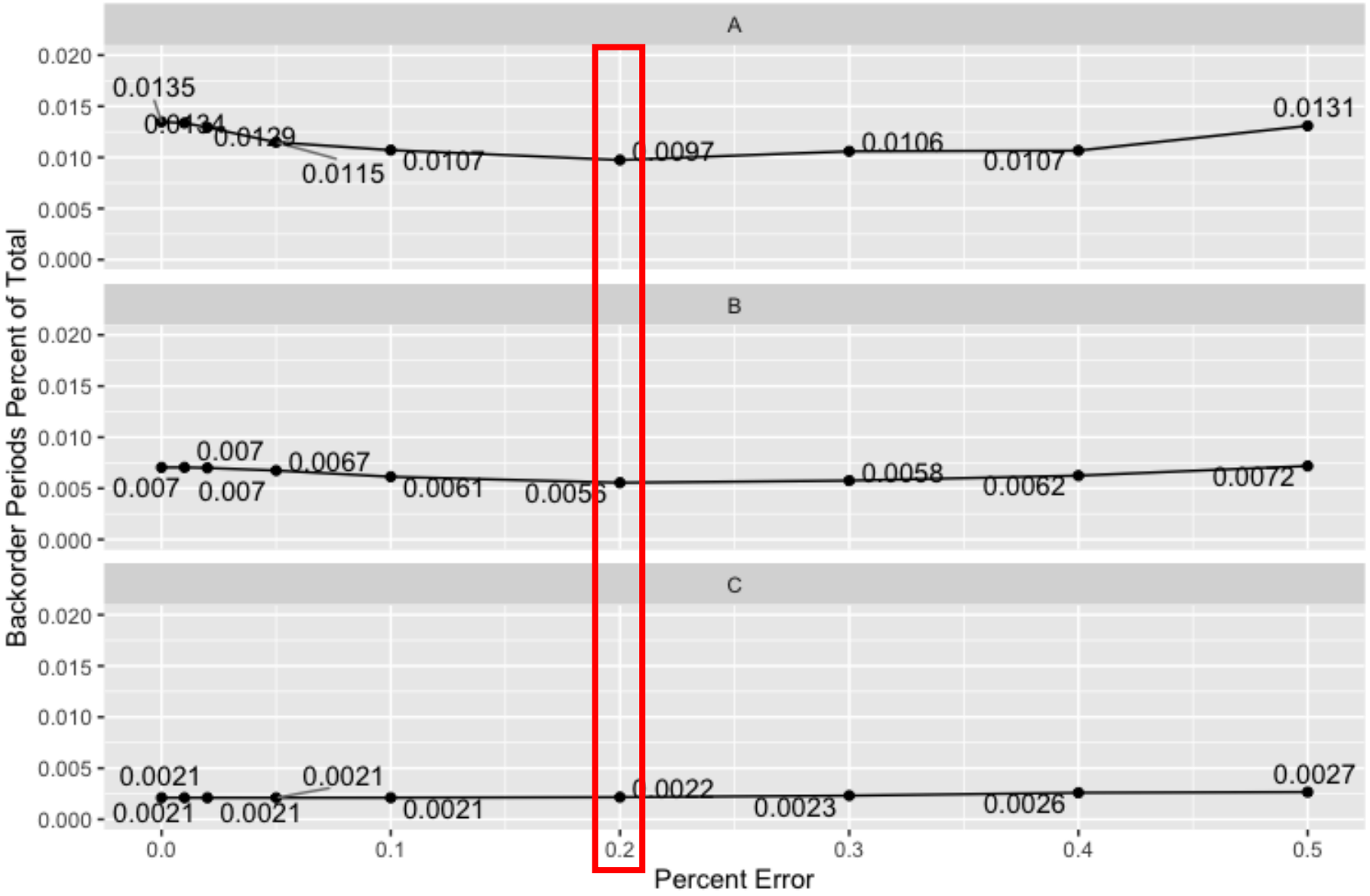
Conclusion: Similar pattern between A, B, C except that C items have little safety stock until higher levels of forecast error



Forecast Error & Backorder Periods

Outcome Measure: Average Number of Periods with Backorders

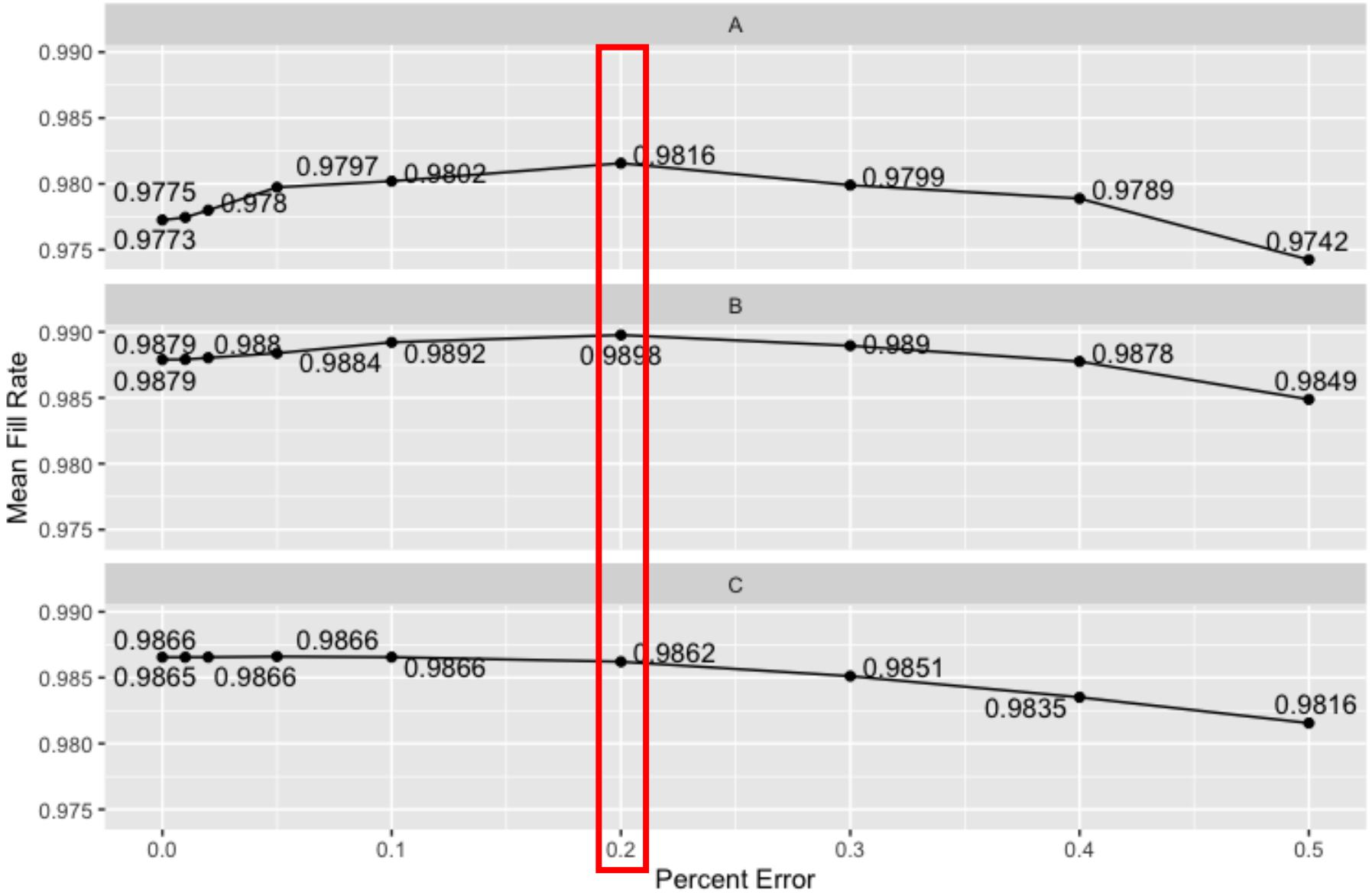
Conclusion: Backorder periods decline until forecast accuracy exceeds 80%... then backorders begin to go up again… “safety stock effect”



Forecast Error & Fill Rates

Outcome Measure:
Mean Fill Rates

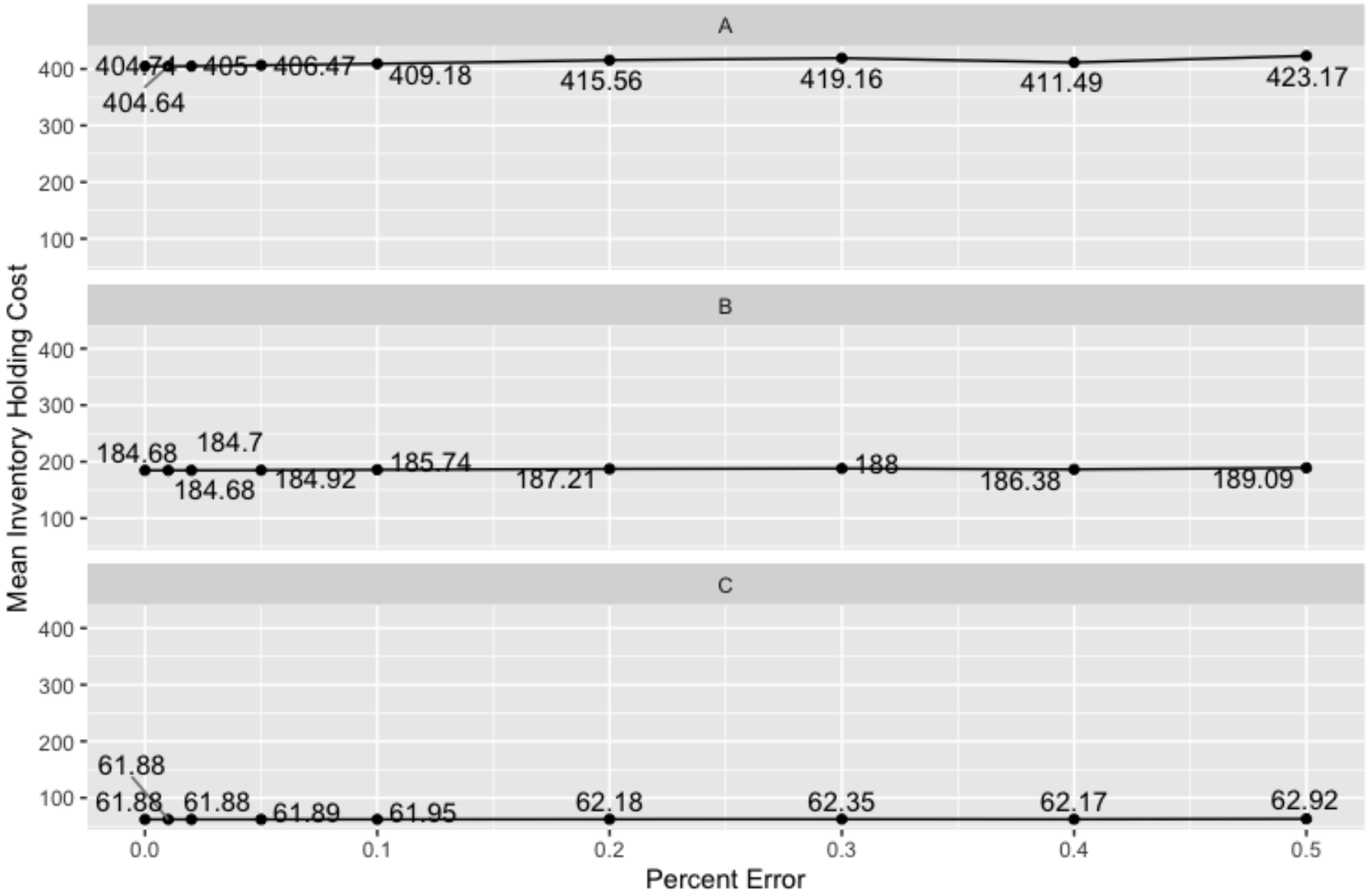
Conclusion:
Fill rates increase until 80% forecast accuracy achieved then decline…
biggest impact on categories A & B…
category C has no impact



Forecast Error & Inventory Holding Costs

Outcome Measure: Inventory Holding Costs

Conclusion: Inventory Holding Costs decline as forecast accuracy increases across all A,B, C categories



Conclusions Summary

- As accuracy goes up:
  - Safety stock and inventory goes down
    - Savings from inventory holding cost perspective
  - Backorder periods can go up when accuracy is very high
    - Due to Safety Stock → 0 and monthly calculated reorder quantity (Q*)
    - The ratio of Safety Stock to Order Quantity and demand decreases
    - Mostly affects higher demand items and not low intermittent demand items

Implications

- Implications for forecast accuracy improvement:
  - Forecast accuracy does matter, **but…**
  - Supply chain planning configurations matter a great deal… maybe more
    - Set A, B, C classifications on SKUs
    - Consider setting floors on safety stock levels for key items
  - These results will change as the supply chain planning configurations change
  - Spend limited demand planning and forecasting improvement efforts on forecast improvements on A & B category items
    - Don’t try to improve forecasts where accuracy is at or above 80% accuracy

Code for Completing Supply Chain Simulations

- Code hosted on GitHub – Hyperlink
- Each SKU approximately 1 minute to complete analysis
- 30,490 SKU's – 2 weeks to analyze
- Analysis only needs to be completed once

QR Code for GitHub Repository

Appendix

Simulation Methods to Determine Outcomes

- Assumptions must be made about the supply chain replenishment process – Theodoru & Spiliotis (2022)
- SAP provides one of the most utilized supply chain planning systems – enlyft (2023)
    - Utilized the supply chain formulas from SAP's Sourcing and Procurement (Materials Management) – SAP S/4 HANA | 2020 FP502 (May 2021)
    - Key formulas contained in Appendix:
        - Mean absolute deviation (MAD) utilizes an exponentially smoothed MAD
        - Standard deviation = MAD x 1.25
        - Service Level
        - Safety Stock
        - Economic Order Quantity (EOQ)
        - Reorder Level (assumes SAP is configured to use the R, Q replenishment policy)

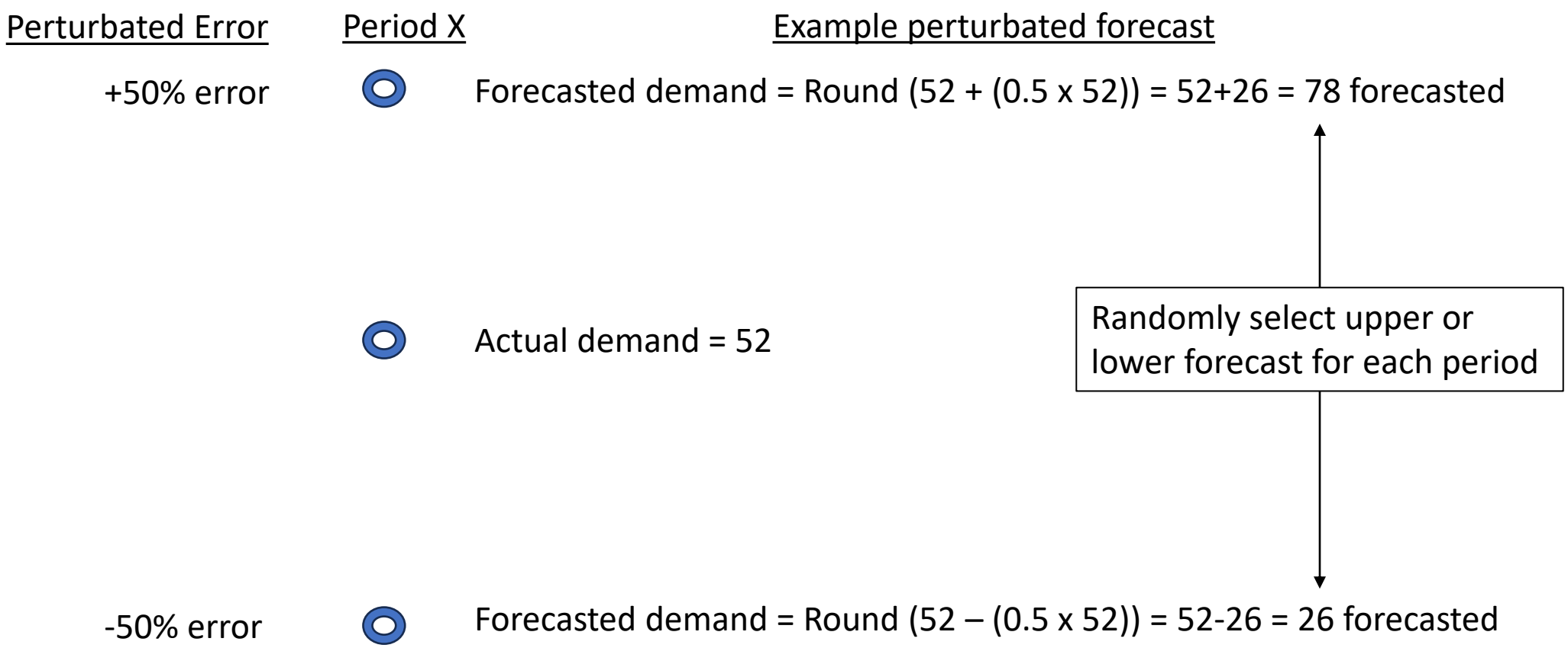
Simulation Methods to Determine Outcomes*

	Step 1	Step 2	Step 3	Step 4	Step 5
Kolassa (2022)	• Decide on forecast supported business process(es) outcomes	• Collect time series (and other data) for simulation	• Simulate ***multiple*** forecasts for each time series	• For each time series and for each forecast for the series: ▪ Calculate forecast accuracy using your various KPIs ▪ Simulate the subsequent business process that relies on the forecast and derive the business outcome KPIs	• Plot or otherwise analyze the relationship between forecast accuracy and business outcomes
Hoover (2023)	• In our case, we evaluated: ▪ Average inventory levels ▪ Backorder periods ▪ Safety stock levels	• We utilized the M5 competition data from Walmart	• Forecasts with errors were perturbated around actuals at 50%, 40%, 30%, 20%, 10%, 5%, 2%, 1%, and 0% error for each period	▪ Accuracy was calculated using Mean Absolute Deviation (MAD) to conform to supply chain calculations ▪ Outcomes simulate the SAP supply chain calculations	• Average inventory levels and backorder periods were compared to the forecast accuracy levels

*Simulation Framework Adapted from Kolassa (2022)

Simulating Forecast at Levels of Accuracy

- Kolassa (2022) – calculate multiple forecasts for each period
- Hoover (2023) – perturbate around known forecast error levels



Simulating Forecast at Levels of Accuracy

Forecast Error Level	Forecast	Error
Actual value	52	
+50%	78	26
+40%	73	21
+30%	68	16
+20%	62	10
+10%	57	5
+5%	55	3
+2%	53	1
+1%	53	1
0% (Perfect)	52	0

- Randomly select either the positive or the negative forecast value at each forecast error level
- Store forecasts for each period at each error level
- Utilize the forecasts errors within supply chain calculations

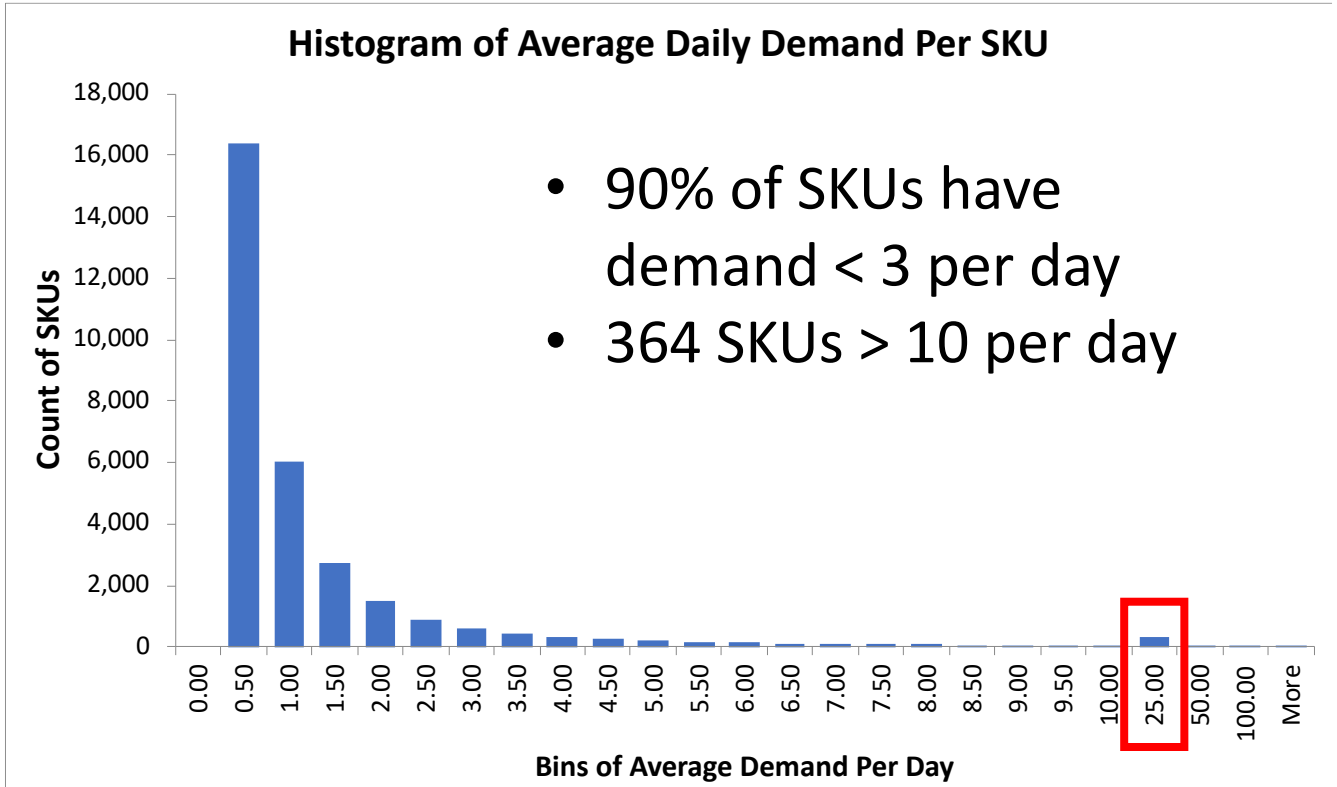
Forecast Error Level	Forecast*	Error
Actual value	52	
-50%	26	-26
-40%	31	-21
-30%	36	-16
-20%	42	-10
-10%	47	-5
-5%	49	-3
-2%	51	-1
-1%	51	-1
0% (Perfect)	52	0

*Lower Forecast Bounded by 0 (zero)

Data Used – M5 Competition SKUs

- 30,490 different SKUs (provided by Walmart for the competition)
- 1,913 daily demand observations per SKU
- Daily demand periods (1/29/2011 – 4/24/2016)
  - Demand is at the store / SKU level
- Average daily demand is low
  - Lots of periods with 0's

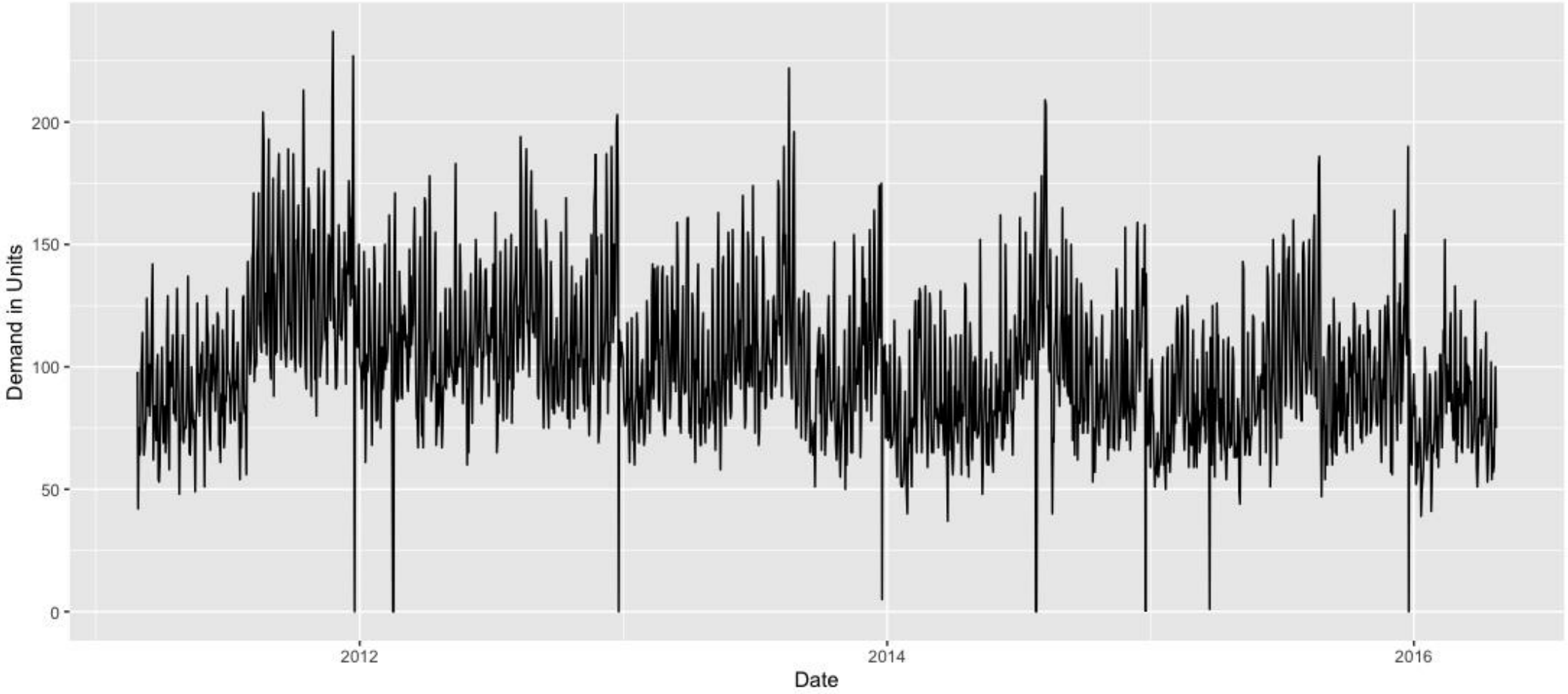
https://www.kaggle.com/competitions/m5-forecasting-accuracy/data



- 90% of SKUs have demand < 3 per day
- 364 SKUs > 10 per day

Simulation Methods to Determine Outcomes

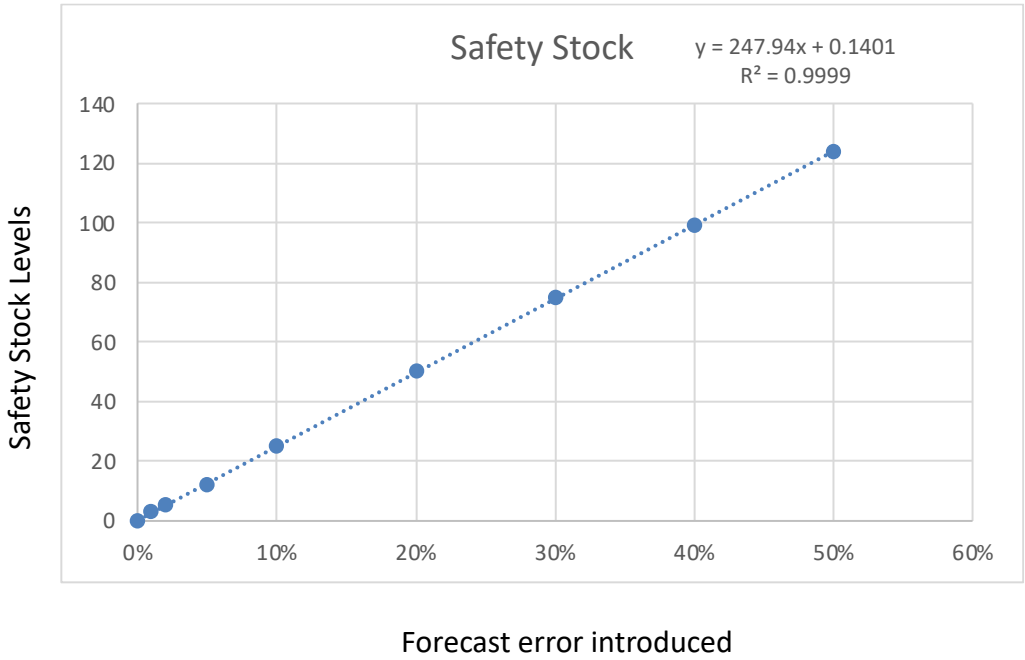
- SKU "FOODS_3_586_TX_2_validation" demand observations over the simulation period



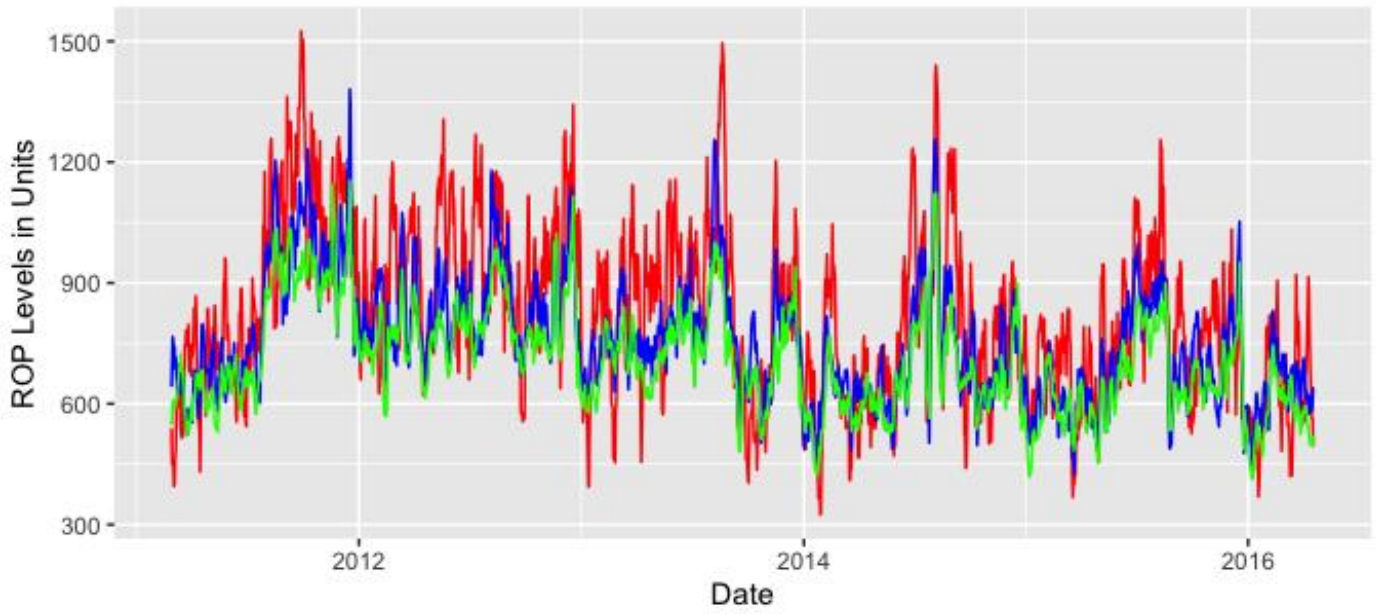
Supply Chain Results – Safety Stock Levels

- Safety Stocks (SS) are calculated with a monthly periodicity
- Monthly "levels" calculations are common in commercial supply chain software (like SAP)

% Error Level	Average SS
50% error	124 units
20% error	50 units
0% error	0 units



Supply Chain Results – Reorder Points



Red = ROP at 50% error

Blue = ROP at 20% error

Green = ROP at 0% error

- Reorder points are calculated using the forecasts over lead time at each date
- Lead time was an assumption of the model = 7 days

% Error Level	Average ROP
50% error	822 units
20% error	759 units
0% error	709 units

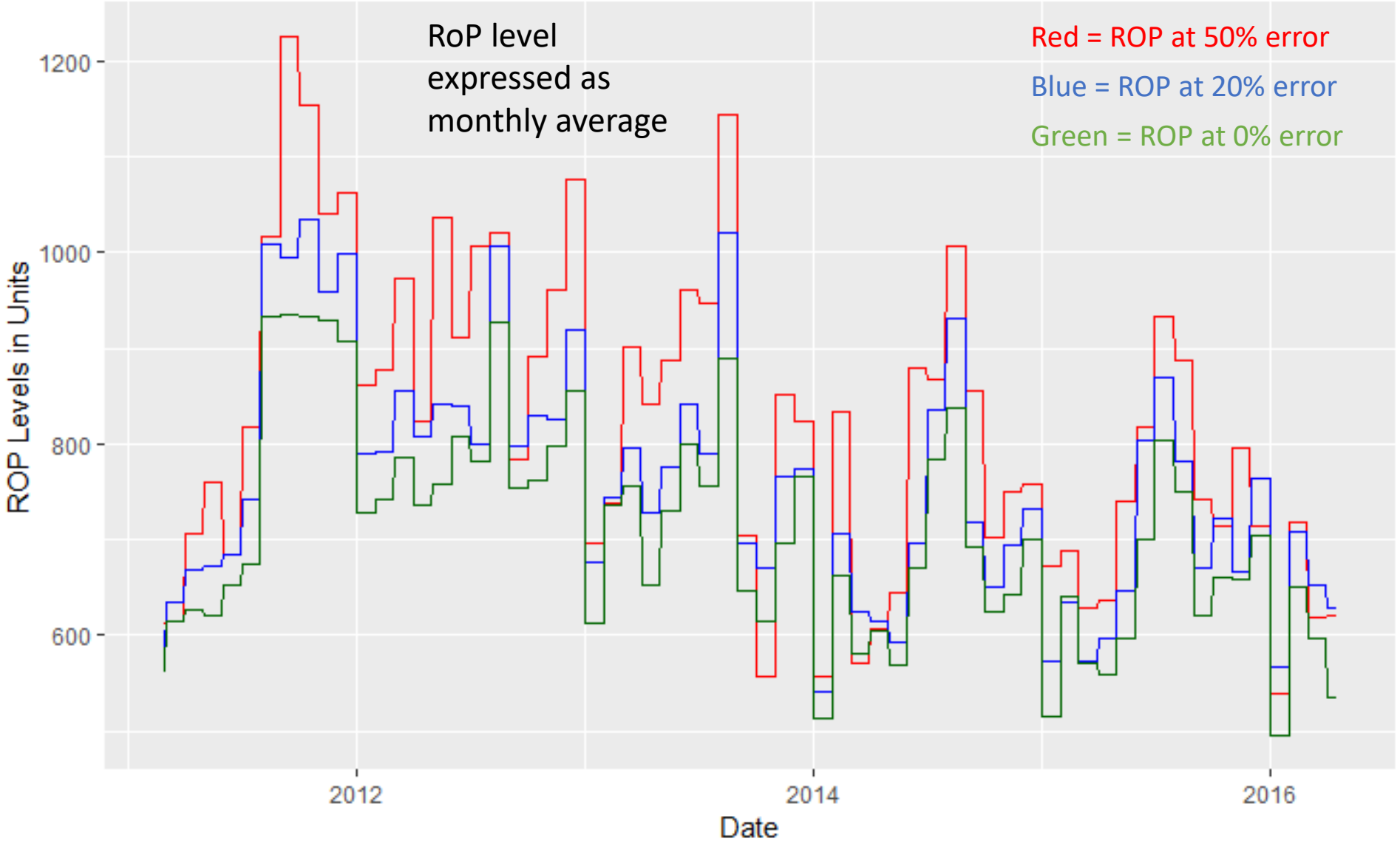
Automatic Reorder Point Calculation and Formula:

- Automatic reorder point = safety stock + daily requirement * replenishment lead time

https://community.sap.com/t5/enterprise-resource-planning-blogs-by-members/automatic-reorder-point-mrp-type-v2-or-vm-planning-part-1/ba-p/13569776

Supply Chain Results – Reorder Points

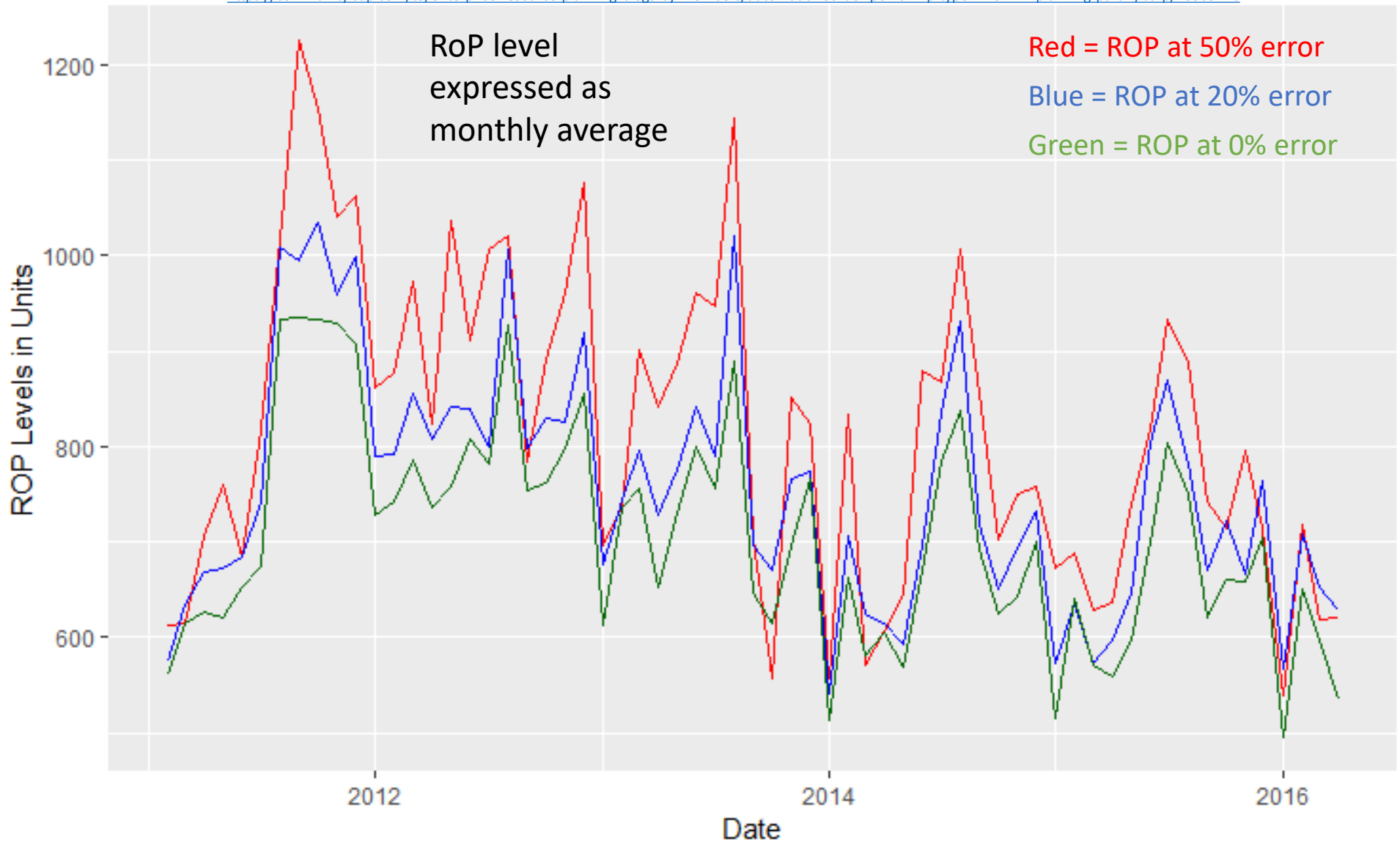
https://community.sap.com/t5/enterprise-resource-planning-blogs-by-members/automatic-reorder-point-mrp-type-v2-or-vm-planning-part-1/ba-p/13569776



% Error Level	Average ROP
50% error	822 units
20% error	759 units
0% error	709 units

Supply Chain Results – Reorder Points

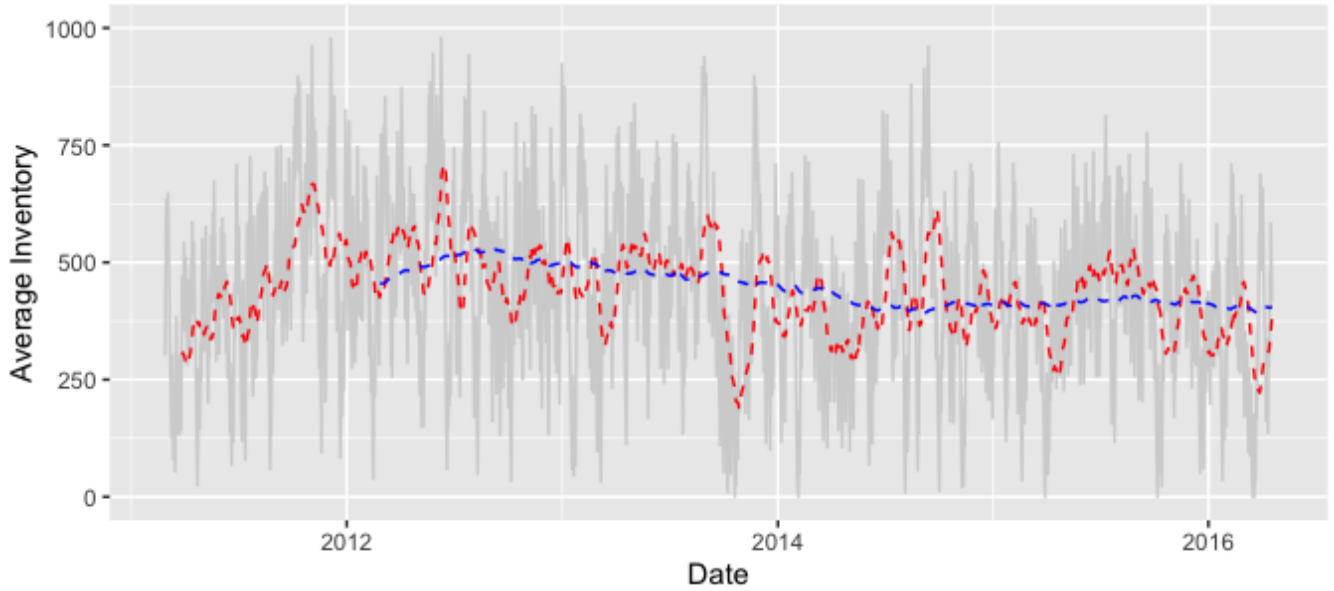
https://community.sap.com/t5/enterprise-resource-planning-blogs-by-members/automatic-reorder-point-mrp-type-v2-or-vm-planning-part-1/ba-p/13569776



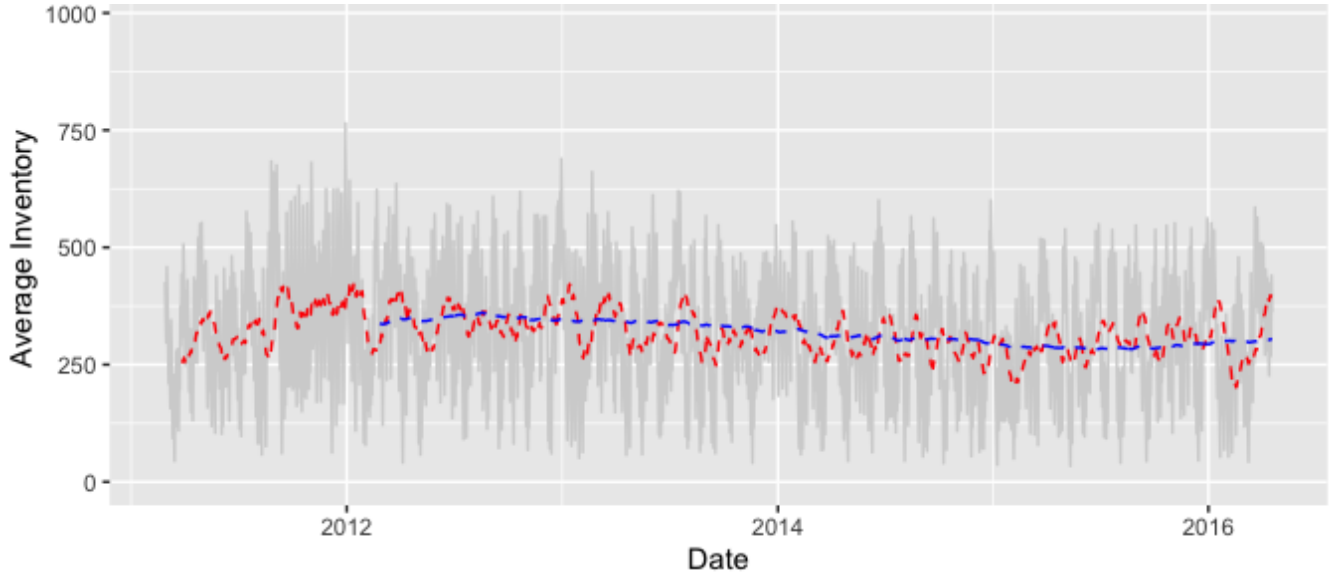
% Error Level	Average ROP
50% error	822 units
20% error	759 units
0% error	709 units

Supply Chain Results – Average Inventory

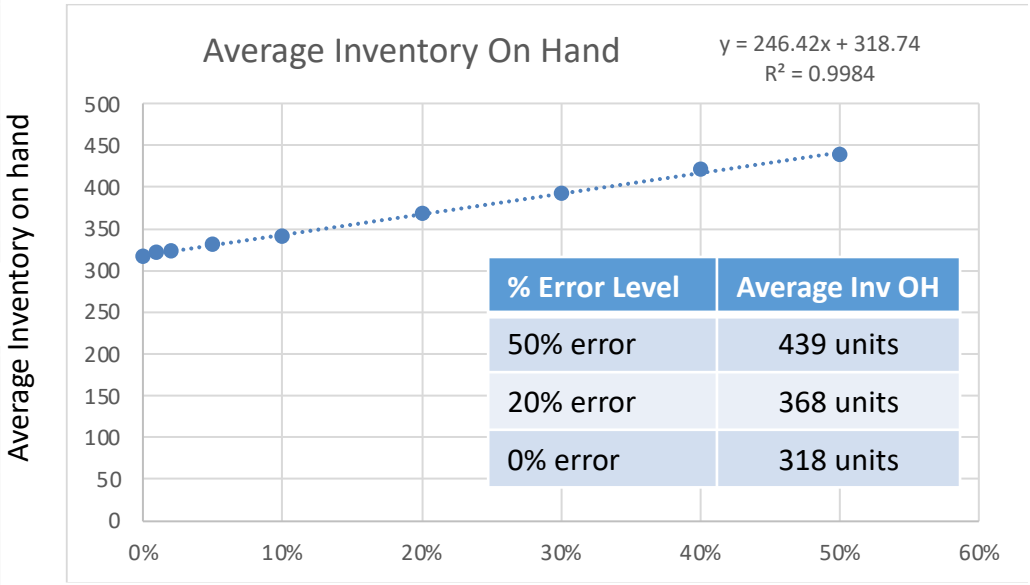
50% error



0% error



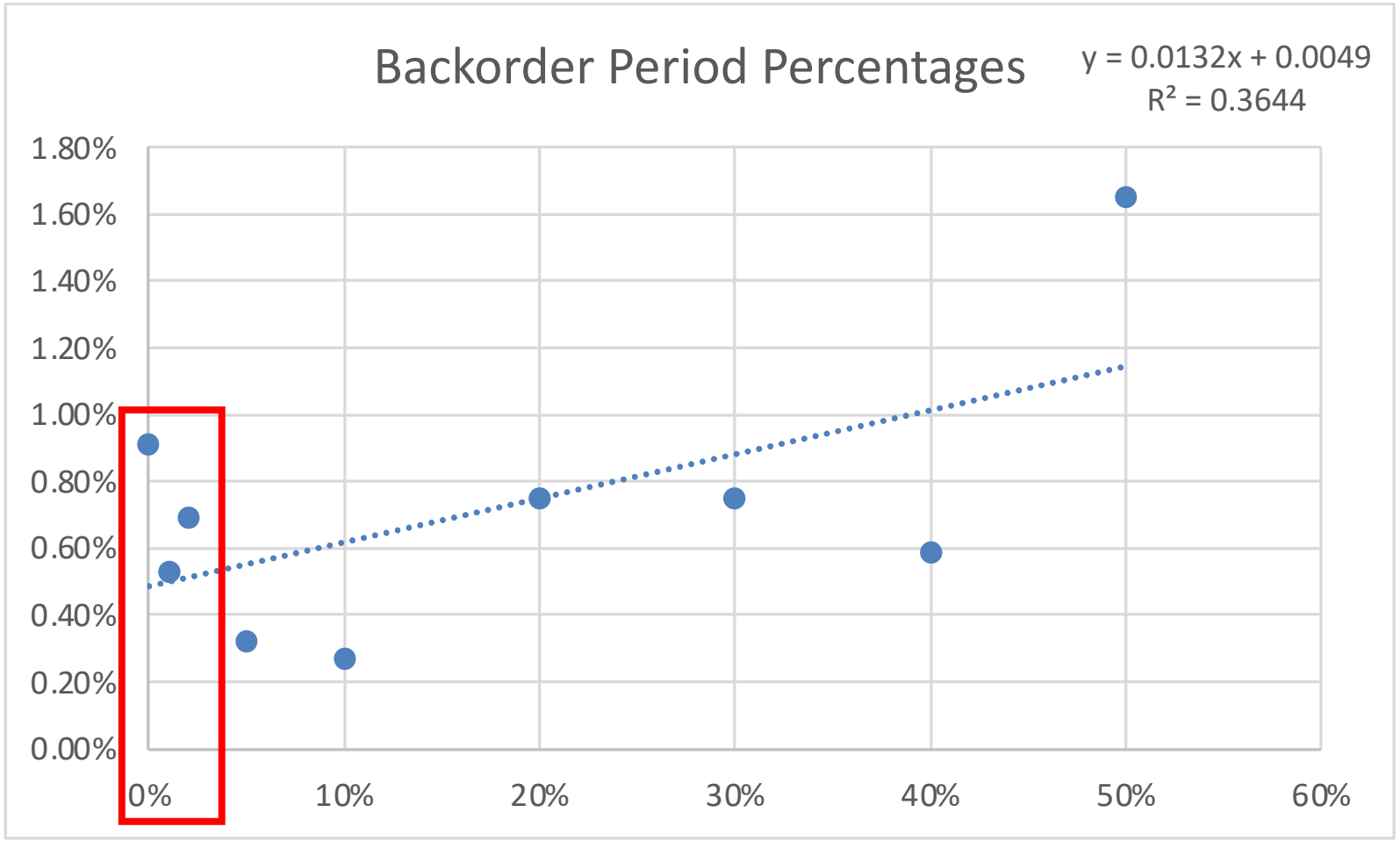
- Average Inventory on hand goes down as accuracy goes up
- 1% reduction in forecast error ~ 0.8% reduction in Average Inventory on hand



% Error Level	Average Inv OH
50% error	439 units
20% error	368 units
0% error	318 units

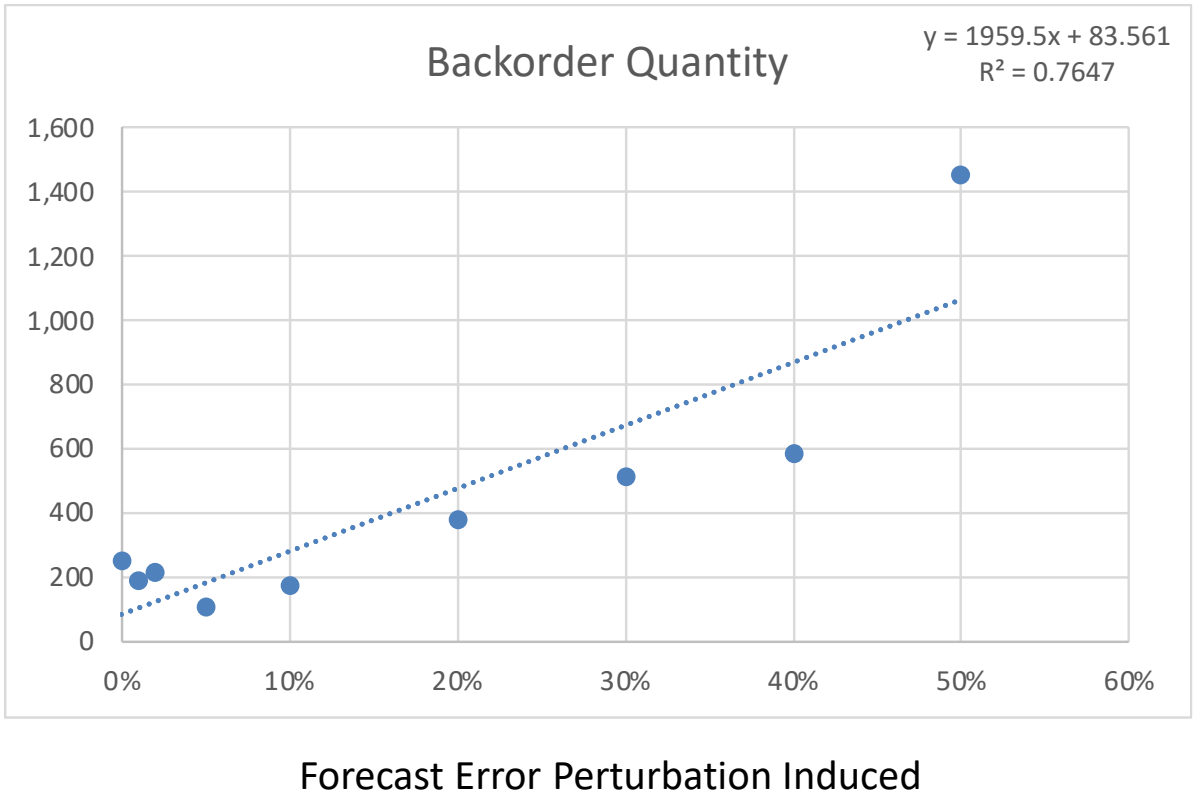
- Average on hand primarily driven by reorder quantity Q*

Outcome Measures – Backorder Periods



- Backorder periods are not well correlated to forecast error

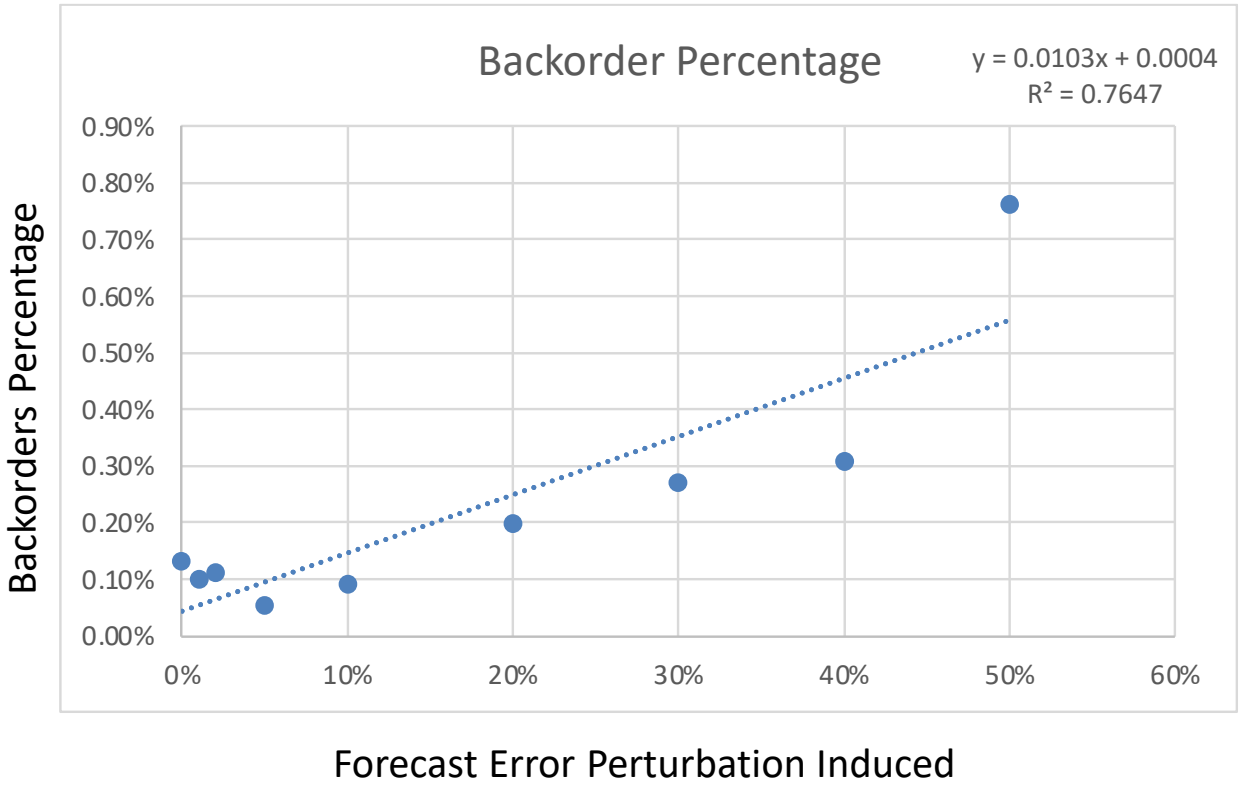
Outcome Measures – Backorder Quantities



% Error Level	Backorder Quantity
50% error	1,450 units
20% error	376 units
5% error	105 units
0% error	248 units

- Backorder quantities seems to increase with forecast error below 5%

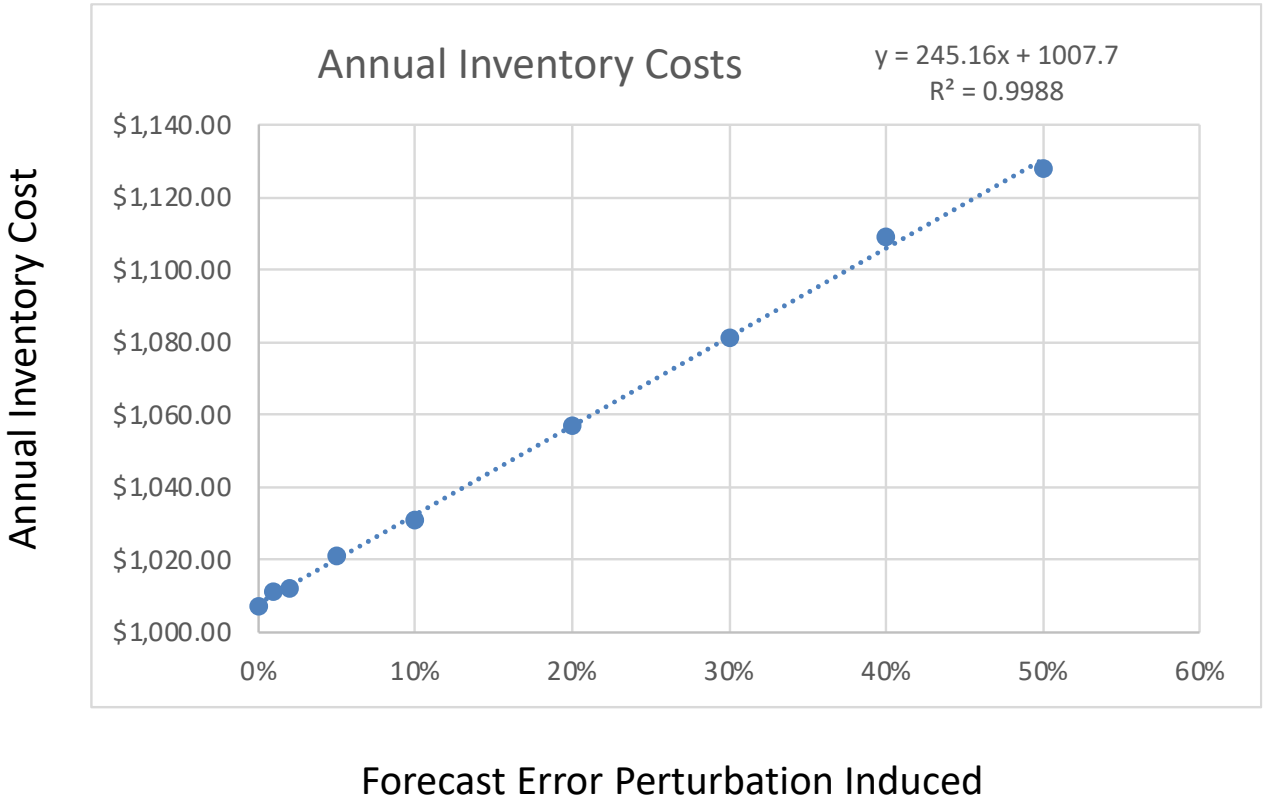
Outcome Measures – Backorder Quantity Percentage



% Error Level	Backorder Percentage
50% error	0.76%
20% error	0.20%
5% error	0.06%
0% error	0.13%

- Backorder quantity percentage seems to increase with forecast error below 5%

Outcome Measures – Inventory Holding Costs



% Error Level	Inventory Costs
50% error	$1,128
20% error	$1,057
5% error	$1,021
0% error	$1,007

- Linear relationship between annual inventory costs and forecast error introduced
- $2.45 per every 1% reduction in forecast error

Code for Completing Supply Chain Simulations

- Code hosted on GitHub – Hyperlink
- Each SKU approximately 1 minute to complete analysis
- 30,490 SKU's – 2 weeks to analyze
- Analysis only needs to be completed once

QR Code for GitHub Repository

Formulas - MAD

- Formulas 13 & 14 from SAP documentation

Mean Absolute Deviation for Initialization

$$(13) \quad MAD = \frac{1}{n} \sum_{t-1}^{n} \left| P(t) - V(t) \right|$$

n = Number of periods for initialization

n >= 4

Mean Absolute Deviation for the Ex-Post Forecast

$$(14) \quad MAD(t) = (1 - \delta) * MAD(t-1) + \delta * \left| V(t) - P(t) \right|$$

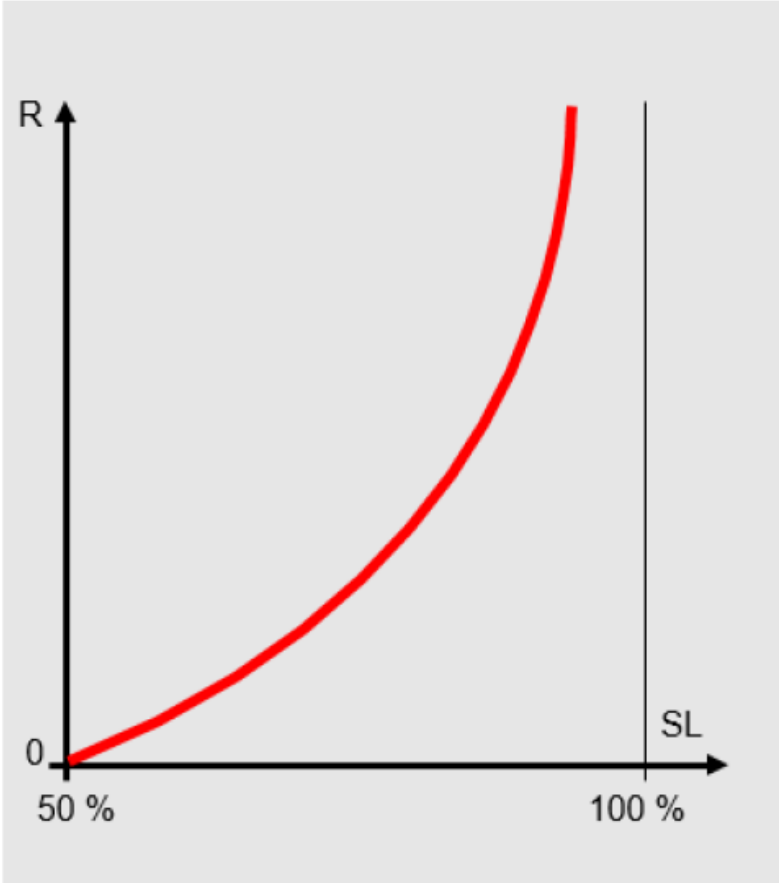
alpha = 0.2

Formulas – Service Level (Look-up Table)

- Supports Safety Stock Calculation

5/25/2023

The following figure shows that, without safety stock, customer demand can be satisfied by 50%. It also shows that it is almost impossible to satisfy customer demand 100% of the time. Factor R describes the relationship between forecast accuracy and service level (SL).



SL (%)	Factor R
50	0
55	0.16
60	0.31
65	0.49
70	0.65
75	0.84
80	1.05
85	1.30
90	1.60
95	2.06
98	2.56
99	2.91
99.5	3.20
99.8	4.00

Formulas – Safety Stock

- Safety Stock Calculation – Formula 17

If replenishment lead time is greater than the forecast period by factor W, then the mean absolute deviation (MAD) is recalculated for this period (formula 17).

Formula 17

$$SS = R \times \sqrt{W} \times MAD$$

$$W = \frac{\text{Delivery Time (in Days)}}{\text{Forecast Period (in Days)}}$$

Formulas – Reorder Level (ROP)

Reorder Level Calculation

The reorder level is defined as the sum of the safety stock plus the requirement forecast within the replenishment lead time (see formula (17)).

Example

A forecast was carried out on a monthly basis. A month has 30 days in the case of external procurement.

Safety stock		100
Forecast	1st subsequent period	200
	2nd subsequent period	300
	3rd subsequent period	400
Replenishment lead time	40 days	0/30
		30/30 + 10/30

(an entire monthly requirement + a part of the following month)

Reorder level = 100 + 30/30 * 200 + 10/30 * 300 = 400

Formulas – Classic Wilson's EOQ

Blog Post by SAP on calculation of the EOQ - Hyperlink

Calculated monthly using levels update

APICS defines EOQ as "a type of fixed order quantity model that determines the amount of an item to be purchased or manufactured at one time." The basic premise is to balance costs on a a production or ordering schedule where small frequent batches incur a fixed order or changeover cost and large batches have excess inventory carry costs. Borrowing from APICS again we see the basic EOQ formula to be:

Economic Order Quantity

$$Q^* = \sqrt{\frac{2DK}{h}}$$

where:

Q* = Economic Order Quantity

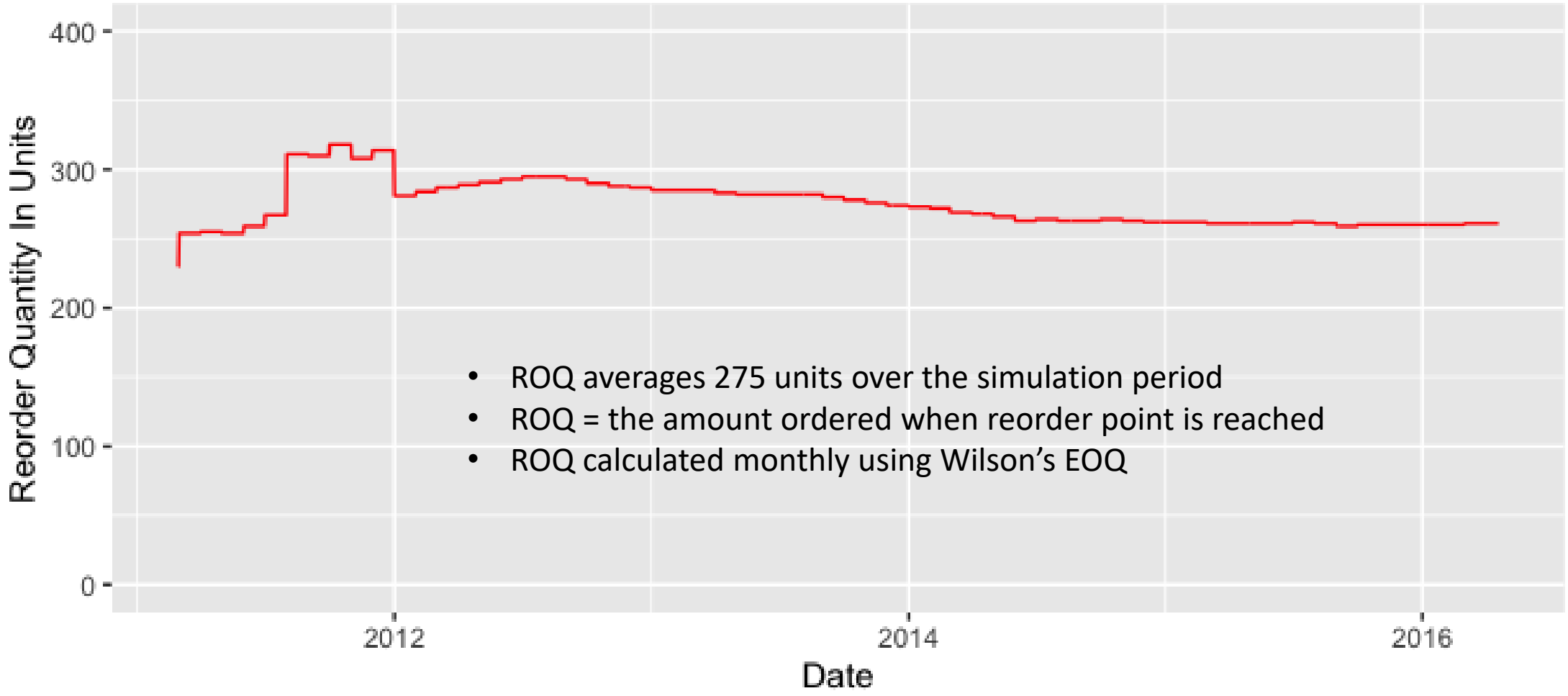
D = Annual Demand

K = Ordering Cost (Or changeover / setup costs)

h = inventory holding cost per unit

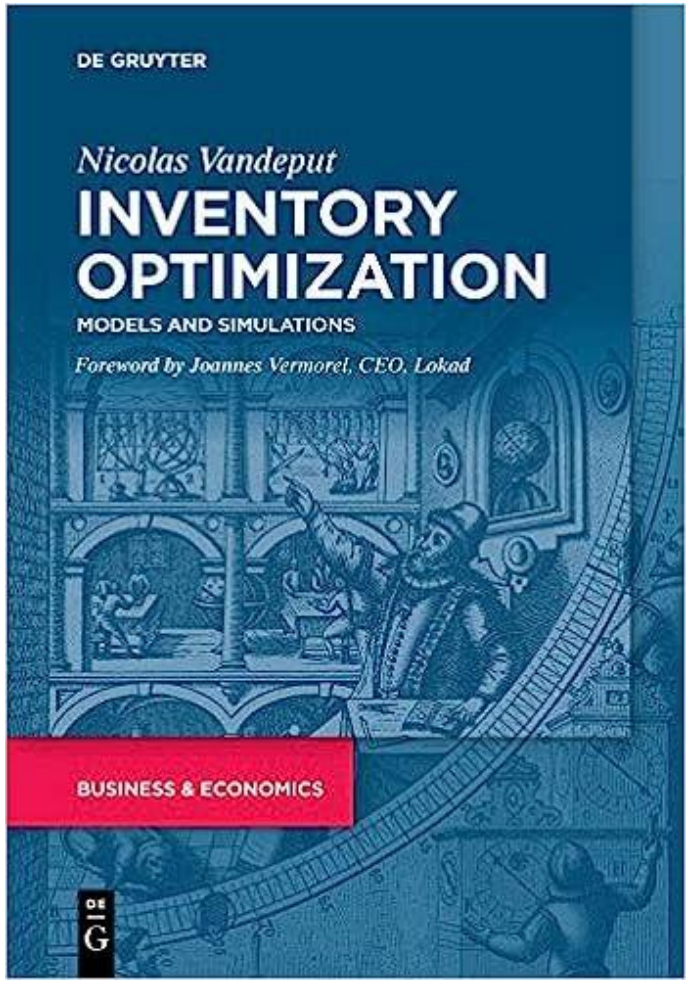
Note: Other methods for calculating the order quantity in SAP exist (e.g., SAP APO calculation of EOQ and safety stock together using a normal curve). This method requires a substantial amount of knowledge of the system configuration. The Wilson EOQ method is discussed and suggested as a potential setting.

Formulas – EOQ results

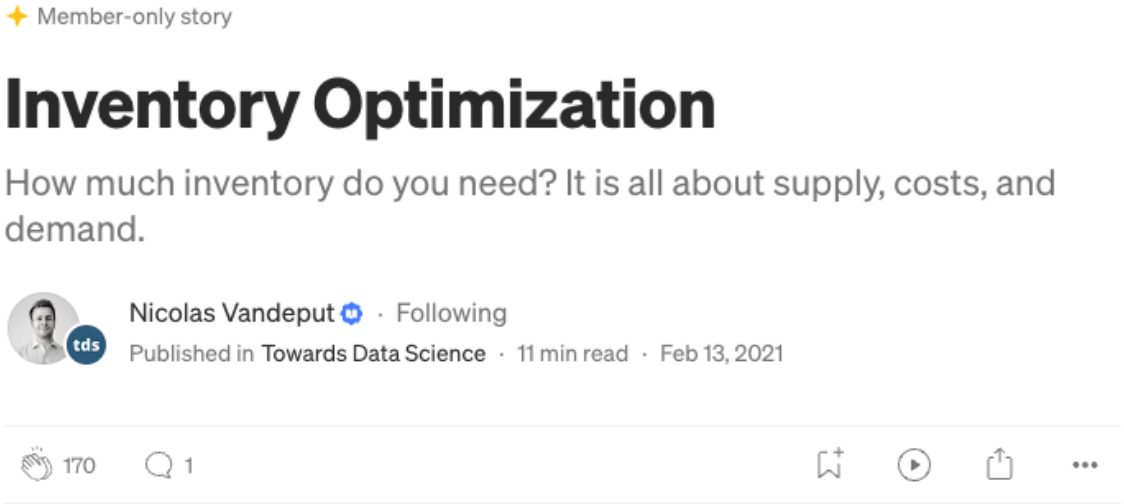


- ROQ averages 275 units over the simulation period
- ROQ = the amount ordered when reorder point is reached
- ROQ calculated monthly using Wilson's EOQ

Other References



Great book by Nicolas Vandeput on setting up simulations - Hyperlink



Thorough article on Medium.com by Nicolas Vandeput on concepts - Hyperlink