Forecasting Book Club
University in Novi Sad,
Faculty of Technical Sciences

Chapter 7: Inventory Control, Aggregation, and Hierarchies

Dr. Dejan Mircetic

University of Novi Sad, Republic of Serbia;

AGENDA

- **7.1 INTRODUCTION**
- **7.2 IDENTIFYING REORDER LEVELS AND SAFETY STOCKS**
- **7.3 ESTIMATING THE PROBABILITY DISTRIBUTION OF DEMAND**
  - 7.3.1 Using Prediction Intervals to Determine Safety Stocks
- **7.4 WHAT IF THE PROBABILITY DISTRIBUTION OF DEMAND IS NOT NORMAL?**
  - 7.4.1 The Log-Normal Distribution
  - 7.4.2 Using the Poisson and Negative Binomial Distributions
- **7.5 TEMPORAL AGGREGATION**
- **7.6 DEALING WITH PRODUCT HIERARCHIES AND RECONCILING FORECASTS**
  - 7.6.1 Bottom-Up Forecasting
  - 7.6.2 Top-Down Forecasting
  - 7.6.5 Issues and Future Developments
  - Additional slides
- **7.7 WRAP-UP**

7.1 INTRODUCTION

- In many organizations, forecasts guide decisions on how much safety, or buffer, stock to hold to reduce the probability of stock-outs to an acceptable level.
- There is a direct connection between the accuracy of forecasting process and the accuracy of the decisions based on them, Fig. 1.
- The point forecasts we have focused on in the preceding chapters are not sufficient to provide this guidance – we also need to know how much uncertainty surrounds the point forecasts.
- There is a direct connection between the accuracy of forecasting process and the accuracy of the decisions based on them, Fig. 1.

Fig. 1. Processes in supply chain (SC) which are driven by demand forecasts

Demand forecasts & inventory decisions

- It is important to distinguish between a demand forecast and an inventory decision!
- A demand forecast is an estimate of what is likely to happen to sales in a future period. An inventory decision is a manager's choice of how much stock should be held in light of the forecast.

7.2 Identifying reorder levels and safety stocks

- When deciding on inventory levels, managers have to balance two costs – holding and stock-out costs. The cost of holding inventory includes storage and insurance costs and the financial returns that could have been earned had money been invested elsewhere rather than in inventory. There is also a risk that stock will deteriorate and, hence, lose its value during storage, or that it will become obsolete before it is sold.

7.2 Identifying reorder levels and safety stocks

- A stock-out occurs when there is demand for a product that we cannot meet.
- Stock-out costs can be more difficult to measure, but they include loss of profit resulting from demand that cannot be fulfilled, the cost of emergency production runs, and the losses arising from damaged customer goodwill.
- It usually makes no sense to have a policy of never having a stock-out – unless, for example, the product is a life-saving drug – because the holding costs would be unduly expensive.
- As a result, organizations often aim to meet demand on, say, 95% of occasions so they would only disappoint customers in one out of every twenty periods on average.

Inventory policies:

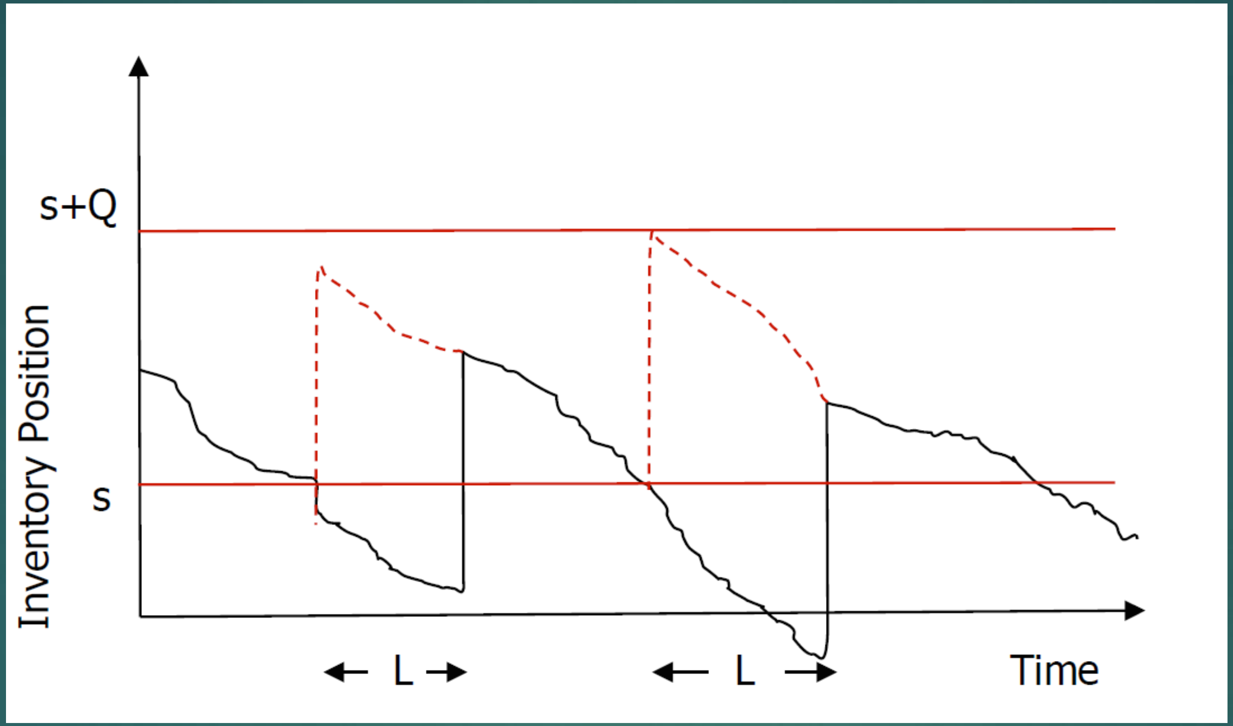
Continuous review inventory system:

- We constantly monitor stock levels. We allow these to be depleted until they reach a level where it is necessary to place an order for replenishments. This is called the reorder level (or reorder point).

Periodic review system:

- Stock levels are only reviewed periodically at particular times (e.g., once a month).

Continuous review inventory system:



- Usually the order will take some time to be delivered. The time between placing the order and receiving the new supplies is called the lead time. We need to set the reorder level to ensure that we still have sufficient stocks to keep us going while we await the delivery – albeit allowing for a predetermined risk of a stock-out. To achieve this, ideally we need to know the probabilities of different levels of demand in the lead time.

Continuous review inventory system:

$$s = \mu_{DL} + k\sigma_{DL}$$

Reorder Point

Forecasted mean demand over the lead time

Safety Stock

k = SS factor

σ_{DL} = RMSE of forecast

- The normality assumption of forecasting distribution during the lead time.
- Expected (or mean) level of demand in the lead time is 2,200 units, but it can vary to below 1,400 and above 3,000 units.
- The diagram suggests that a demand of about 2,800 units is exceeded only 5% of the time, so if we want to achieve that customer service level, we should have 2,800 units in stock at the start of the lead time.

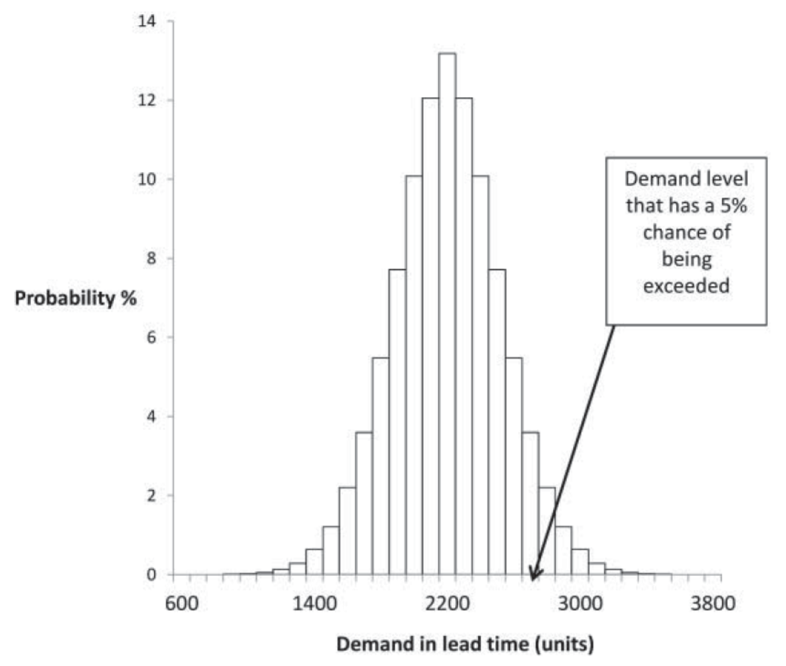


Figure 7.1 Normal probability distribution of demand

Continuous review inventory system:

- Figure 7.2 shows probability distributions of demand in the period-to-be-forecast for different standard deviations.

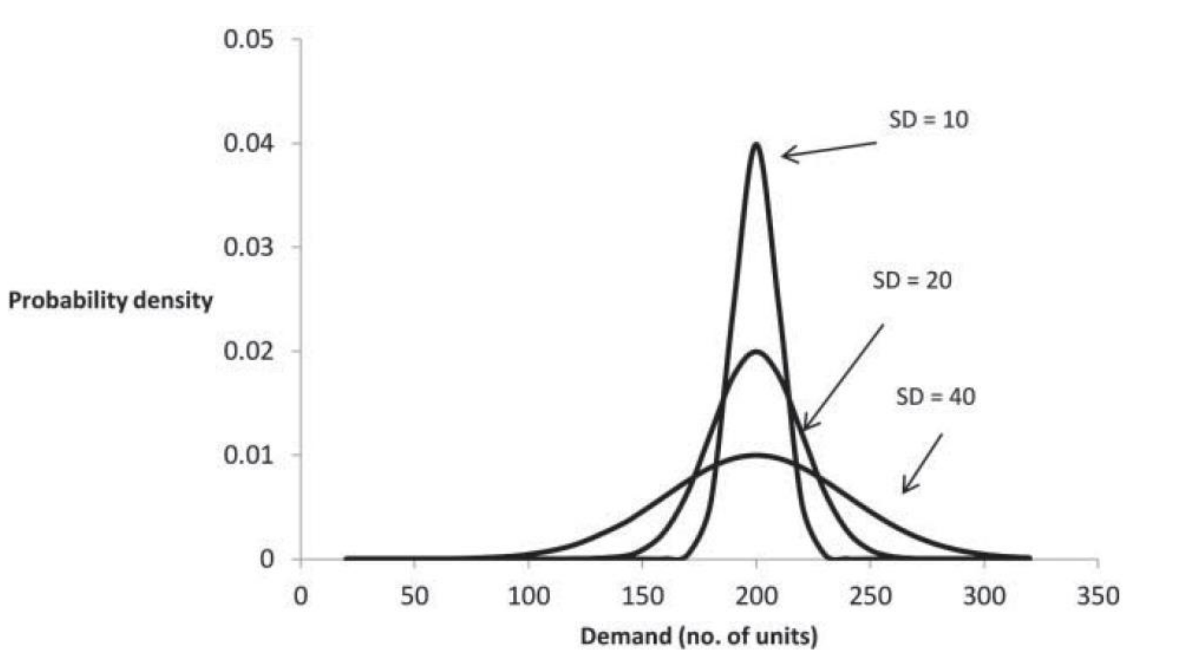


Figure 7.2 Normal distributions with different standard deviations

- The diagram shows that the distribution has a greater spread when the standard deviation is larger. This drastically increases the safety stock!

Continuous review inventory system:

- Once we know the expected demand and the standard deviation for the lead time, we can determine the amount of stock we need to meet a given customer service level by using the following formula:

Reorder level = Expected demand in lead time + Z × Standard deviation.

- Here Z depends on the service level required and can obtained from published tables. Table 7.1 shows some typical values.

Table 7.1 *Z* values for Different Customer Service Levels

Probability of Stock-Out	*Z*
0.5%	2.58
1.0%	2.33
2.0%	2.05
2.5%	1.96
4.0%	1.75
5.0%	1.65
10.0%	1.28
20.0%	0.84

Continuous review inventory system (example):

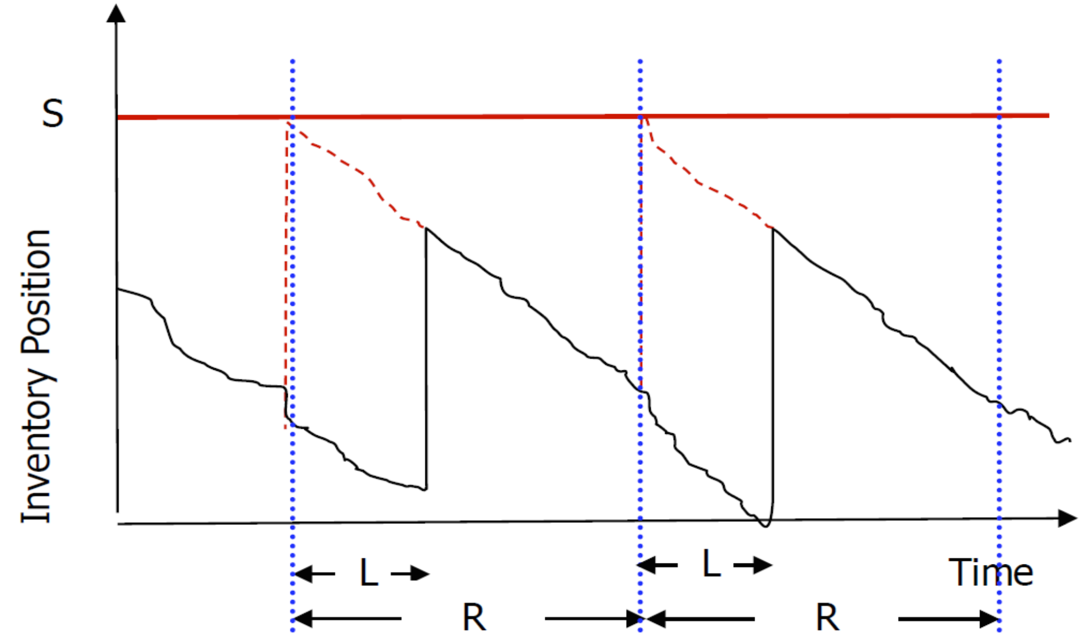
- For example, if the expected demand for a lead time is 450 units, the standard deviation is 50 units, and we want to allow a 2% chance of a stock-out, then:

$$\textit{Reorder level} = 450 + 2.05 \times 50 = 553 \textit{ units}$$

- Which indicates that we should be carrying 553 − 450 = 103 units of safety stock.
- It can be seen that the standard deviation reflects the amount of uncertainty we have about demand during the lead time, and the larger it is, the greater the amount of safety stock we will need to achieve a given customer service level.

Periodic review system:

- Stock levels are only reviewed periodically at particular times (e.g., once a month). If the review suggests that replenishments are needed, an order is placed. In this system, we need to ensure that the size of our order is sufficient to brings stocks up to a level that will meet demand, not just for the review period, but also for the subsequent lead time.
- Hence, we need a probability distribution of demand for the period covered by the review interval plus the lead time.



7.3 Estimating the probability distribution of demand

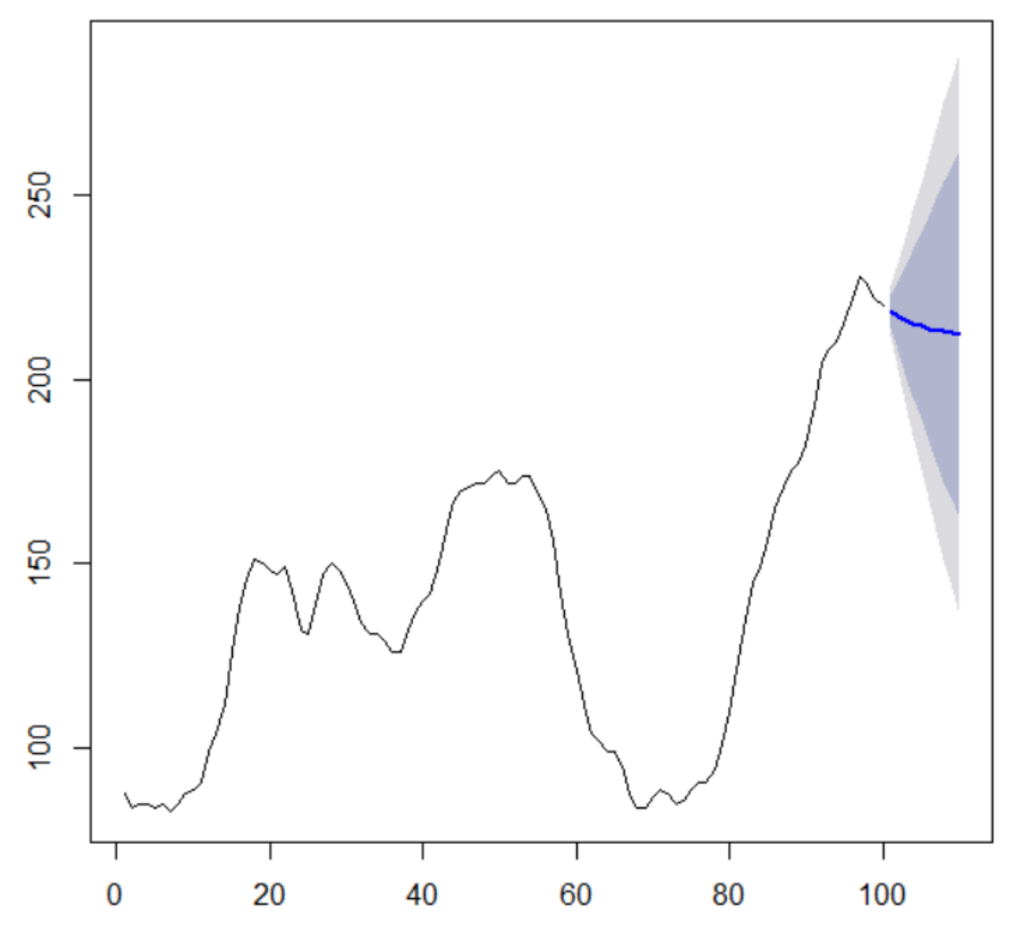
- How can forecasting software be used to estimate the probability distribution of future demand?
- The expected demand in the lead time is simply equal to the point forecast. However, estimating the standard deviation is more challenging!
- The simplest approach is to calculate the root mean square error (RMSE) of past forecasts (see Chapter 3) and use the square root of this as the estimate of the standard deviation.

7.3 Estimating the probability distribution of demand

- However, this approach has a number of limitations, which cause it to underestimate the future variation in demand and hence lead to safety stocks that are too low, resulting in poorer service than we intended. There are three reasons for this:
  - if we use the RMSE from the in-sample periods (i.e., the fitting periods), this will tend to underestimate typical forecast errors.
  - Even if we use the RMSE for the hold-out sample, this will still tend to underestimate the standard deviation. We may have too few forecasts for the hold-out periods to get a reliable estimate of forecast accuracy.
  - The RMSE may only be measured for one-period-ahead forecasts, but the lead time might be a multiple or fraction of the typical period. Rescaling the RMSE to its equivalent for the lead time period is not easy because sales figures and forecast errors are likely to be autocorrelated.

7.3.1 Using Prediction Intervals to Determine Safety Stocks

- Some software products do not suggest safety stock levels, but they do produce forecasts in the form of prediction interva
- A prediction interval is a range of dema a given probability of capturing the actual c on interval of 90 to 110 units should have ual demand in its range.
- Two scenarios:
  - Lead time forecasting probability distribut
  - Lead time forecasting probability distribut



7.4 What if the probability distribution of demand is not normal?

- Demand will never exactly conform to the normal distribution, but it can provide a good approximation in many practical contexts. This is likely to be true when products sell in large numbers and the probabilities of demand exceeding, or falling below, the point forecast are roug

7.4.1 The Log-Normal Distrib

- Here the distribution is highly skewed a are most frequent, there is a possibility have intermittent demand, the size of c sometimes be modeled by a log-normal

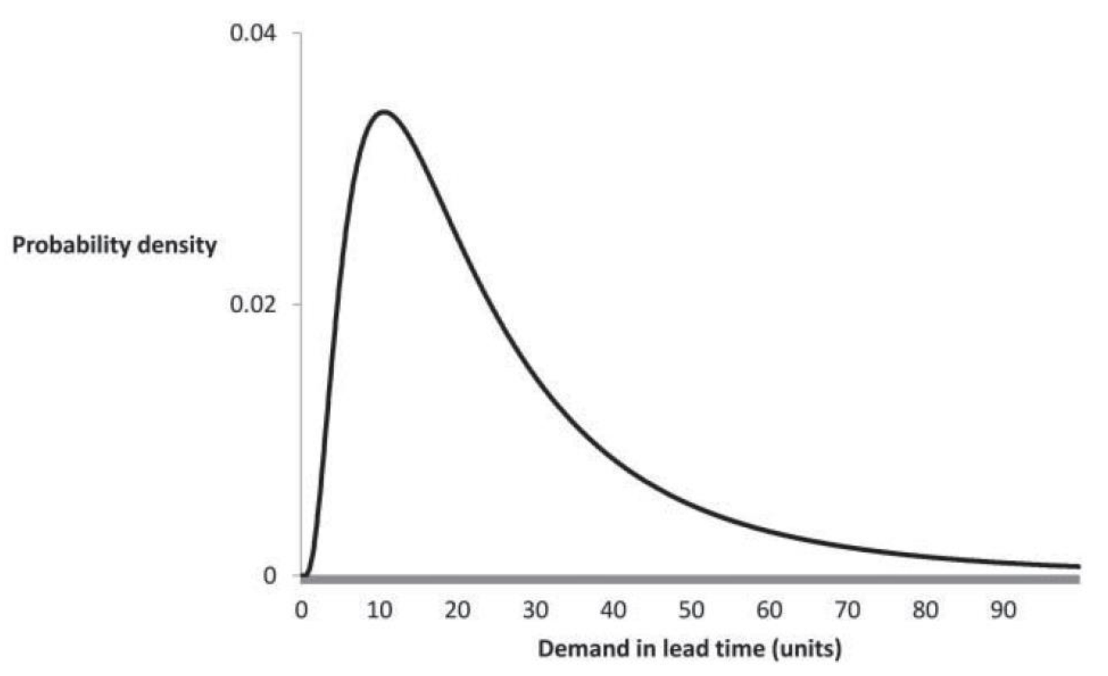


Figure 7.3 A log-normal distribution of demand

7.4.2 Using the Poisson and Negative Binomial Distributions

- The approximation provided by the normal distribution is likely to be poor when demand for a product is relatively small and it is only sold in whole number quantities. In this case, we need to use a discrete distribution – a distribution that assumes that fractional units cannot be sold. Two such distributions are available in some software products. These are the Poisson and the negative binomial distributions.

7.4.1 The Poisson distribution

- The Poisson distribution gives the probabilities that demand in a given period will equal different whole-number quantities. It assumes that the demand occurs randomly over the period in question. To compute the probabilities, we only need to know the mean demand for the period. The standard deviation is equal to the square root of the mean.

7.4.1 The Poisson distribution

- Notice that, unlike the normal distribution, which is symmetrical, the Poisson distribution is skewed with a long tail to the right. This means that any prediction intervals will not be symmetrical around the point forecast.

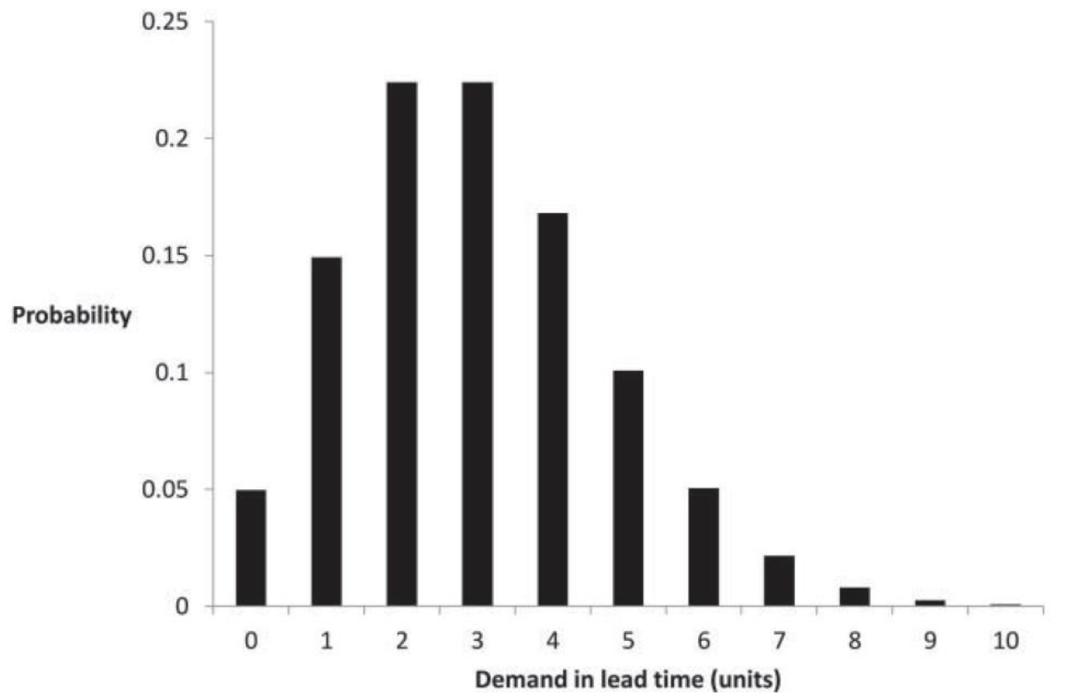


Figure 7.4 Poisson distribution with a mean of 3

7.4.1 The Poisson distribution (example)

Table 7.3 Intermittent Demand with a Poisson Distribution

Month	1	2	3	4	5	6	7	8	9	10
Demand (units)	1	0	1	2	0	1	2	3	0	1

- The software made a forecast that the expected (mean) demand in month 11 would be 1 unit (this is the point forecast), and it used the Poisson distribution to produce a 95% prediction interval of **0 to 3** units.

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

where
x = 0, 1, 2, 3, …
λ = mean number of occurrences in the interval
e = Euler's constant $\approx$ 2.71828

Table 7.4 Poisson Probabilities for Lead Time Demand

Demand in Lead Time (Units)	Probability	Probability of Exceeding This Demand
0	0.368	0.632
1	0.368	0.264
2	0.184	0.080
3	0.061	0.019
4	0.015	0.004
5	0.003	0.001
6	0.001	0.000

The Negative binomial distribution

- Sometimes the variation in demand for low-demand items sold in whole-number quantities is too great for the Poisson distribution to cope with. As we saw, the Poisson distribution assumes that the standard deviation of demand is equal to the square root of the mean demand. When the standard deviation is much greater than this, the negative binomial distribution can be used as an alternative.
- It can be seen that the distribution is highly skewed and any prediction intervals estimated from it are therefore highly asymmetric around the expected sales or point forecast.

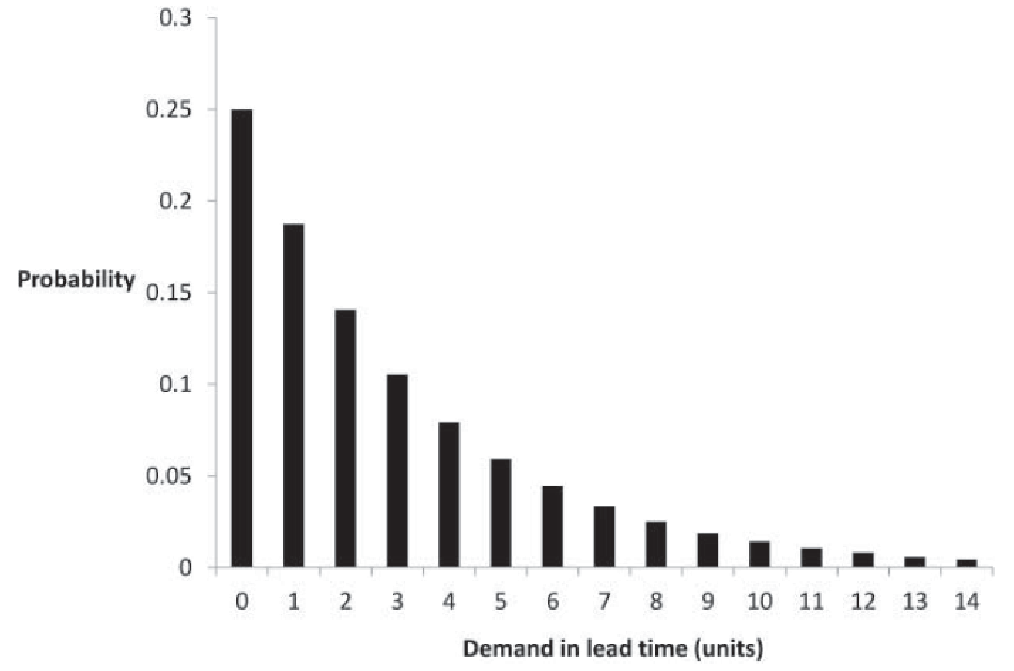


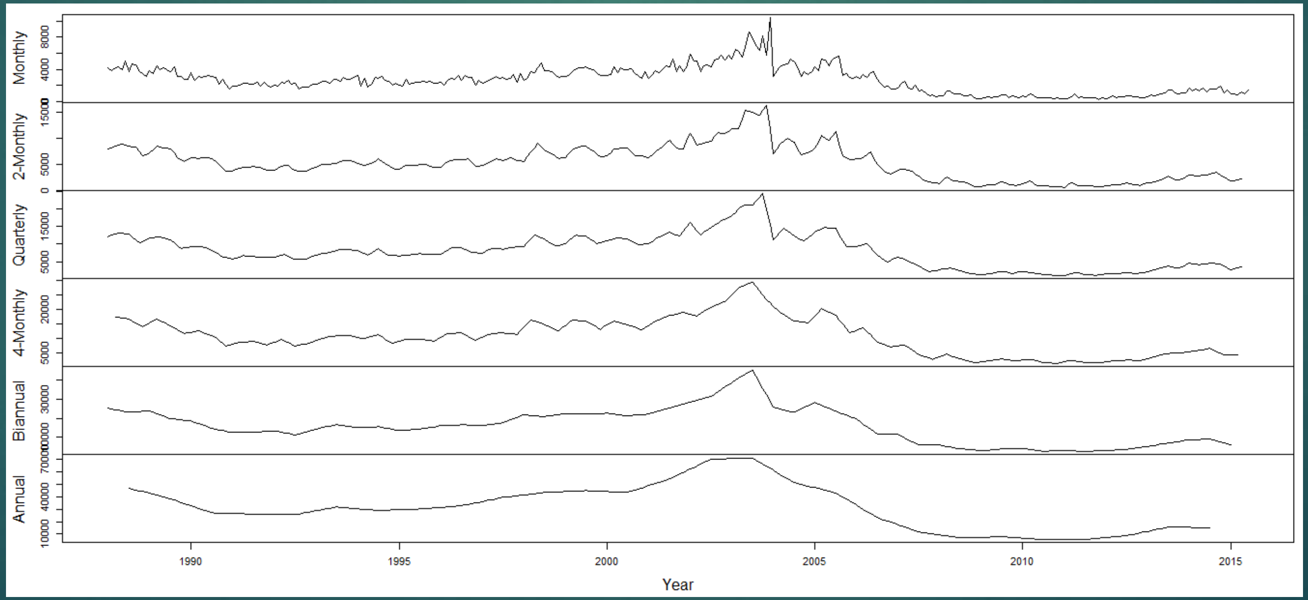
Figure 7.5 Negative binomial distribution with a mean of 3 and standard deviation of 3.5

7.5 Temporal Aggregation:

- Sales histories recorded at the daily or weekly level are often highly volatile. They may even show the characteristics of intermittent demand with some days or weeks recording zero sales.
- By aggregating sales from high-frequency time buckets (e.g., days) to lower-frequency buckets (e.g., weeks of months), we can end up with smoother patterns of sales and less uncertainty about future demand, making forecasting easier and more accurate.
- Temporal aggregation can be:
  - Non-overlapping and
  - Overlapping.

Non-overlapping temporal aggregation

- Non-overlapping temporal aggregation is an approach that transforms a time series from **higher** into **lower** frequencies strengthening or attenuating different elements.
- This is achieved by dividing the time series into consecutive non-overlapping buckets of time with length *m*, where *m* is the aggregation level. A new series is then **generated through the summation** (bucketing) of **every *m* periods** of the high-frequency time series.



Temporal aggregation of series from monthly to annual level.

Temporal aggregation

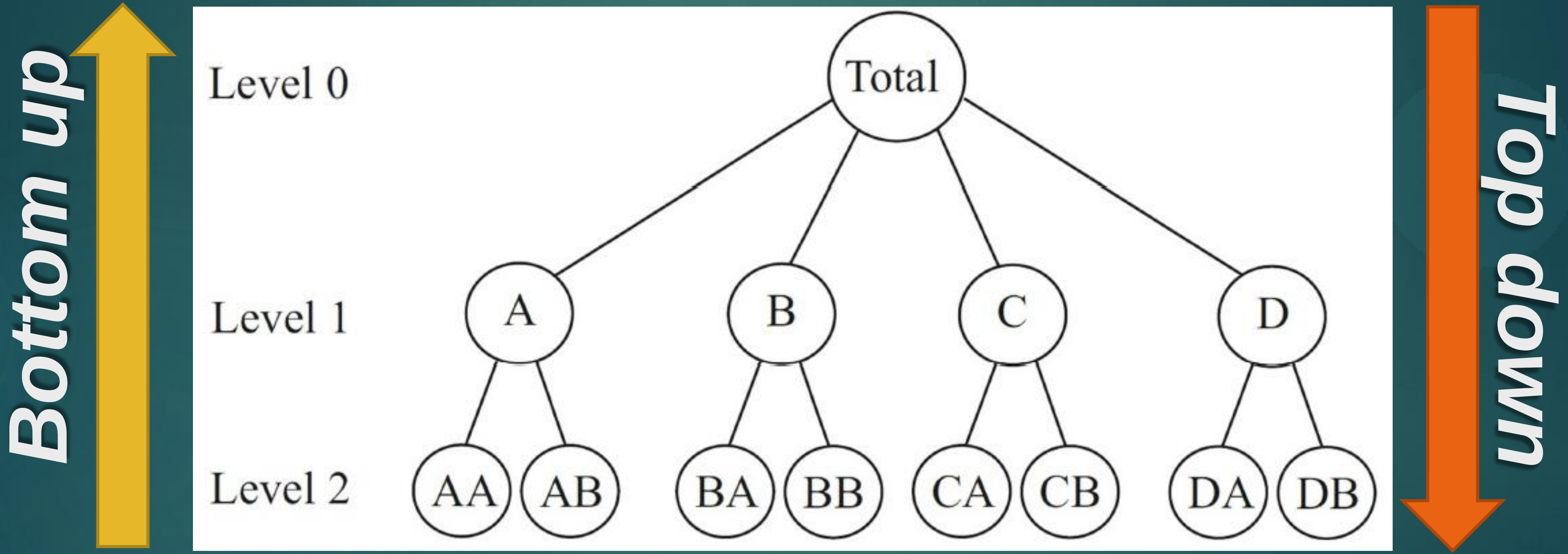
- Literature suggests that temporal aggregation can be beneficial in improving forecast accuracy. The reduction in the number of observations available to the forecasting method is outweighed by the reduced volatility and uncertainty associated with the forecasts.
- Is that always good? Or should we sometimes stay on the original level and forecasted to the period of interest. Should we stay (do not temporally aggregate), or should we go (temporally aggregate given series)?
- Further research on this topic is needed since the initial results in ***The impact of time series characteristics on temporal aggregation performance***, by **Bahman Rostami-Tabar and Dejan Mircetic** demonstrate different results….

Aggregation level	RMSE	MAE	MPE	***MAPE***
Annual level (%)	243 (0.5%)	189 (0.39%)	2884 (6%)	***39651 (82.60%)***
Monthly level (%)	47757 (99.49%)	47811 (99.60%)	45116 (93.99%)	***8349 (17.39%)***
Total (%)	48000 (100%)	48000 (100%)	48000 (100%)	48000 (100%)

7.6 Dealing with product hierarchies and reconciling forecasts

- Often we need to forecast sales at different levels of aggregation. For example, a vehicle manufacturer may need to forecast the monetary value of its total sales, the value of sales in each product class (cars, van, and buses), and the value of sales within each class (e.g., by models of car).
- The problem is that if we make separate forecasts for each level of the hierarchy, it is unlikely that the forecasts will be consistent. Add up the individual forecasts of sales of cars, vans, and buses, and they will almost certainly be different from a separate forecast of total sales.
- Reconciliation methods attempt to address this problem. Three methods available in many software products are bottom-up, top-down, and middle-out forecasting.

HIERARCHICAL FORECASTING APPROACHES



7.6.1 Bottom-up

The advantages of bottom-up forecasting are as follows:

- Forecasting for the individual items at the lowest level of the hierarchy allows the forecasts to take into account conditions that are specific to each item.
- We can incorporate local knowledge into a forecast. For example, the managers of an individual supermarket can build their expertise on their local market into the forecast. Staff at the head office are unlikely to have this knowledge.

However, bottom-up forecasting also has three disadvantages:

- Sales for individual products may be volatile or intermittent, or sales histories may be short.
- It might be difficult to detect relationships between potential sales drivers and the sales of an individual product.
- We regularly need to make separate forecasts for each item at the lowest level of the hierarchy. This can be demanding if the number of these items is large.

7.6.1 Top-Down forecasting

The advantages of top-down forecasting are as follows:

- Generates only one forecast on the top level.

The disadvantages of top-down forecasting:

- We don’t use knowledge about the sales patterns of individual products.
- Researchers have shown that, even if the top-level forecasts are unbiased, the forecasts for the lower level items are bound to be biased.

7.6.5 Issues and Future Developments

- It can be seen that none of the aforementioned methods of dealing with product hierarchies is perfect. Each has disadvantages, and there is no clear evidence on which method is best. The most accurate approach in any specific situation can only be found by trial and error.
- In addition, the methods are designed only to produce point forecasts – the issue of how to generate reconciled forecasts in the form of prediction intervals is a topic currently being explored by researchers.
- One development that may soon be implemented in commercial software uses a new approach called optimal reconciliation.

Additional slides

- Which hierarchical model to use?
- Bottom up? Top down? Optimal combination?
- Large scale numerical studies can possible provide answers.

NUMERICAL SIMULATION

- The simulation is repeated for 500 times, generating 500 different scenarios of bottom level series. Each series in the bottom level has 100 observations. The data for each series at the bottom level are generated from the *Autoregressive Integrated Moving Average* (ARIMA) process.

Results of the numerical simulation

- Accuracy of different HF models measured through the RMSE error.

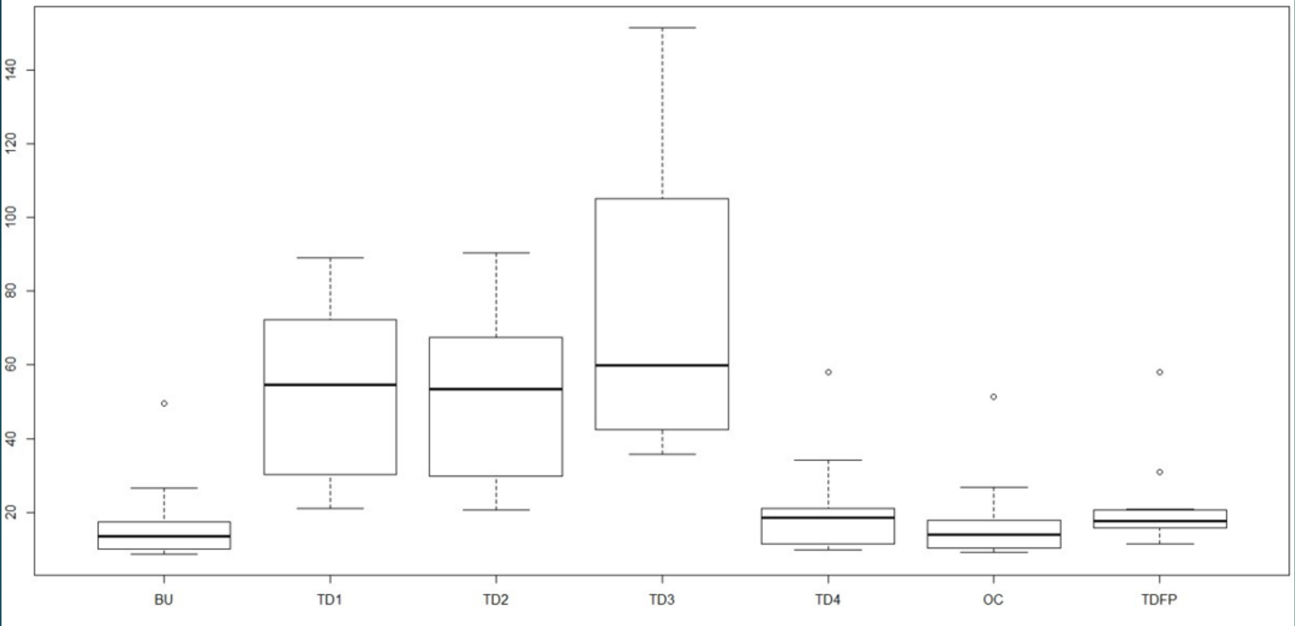
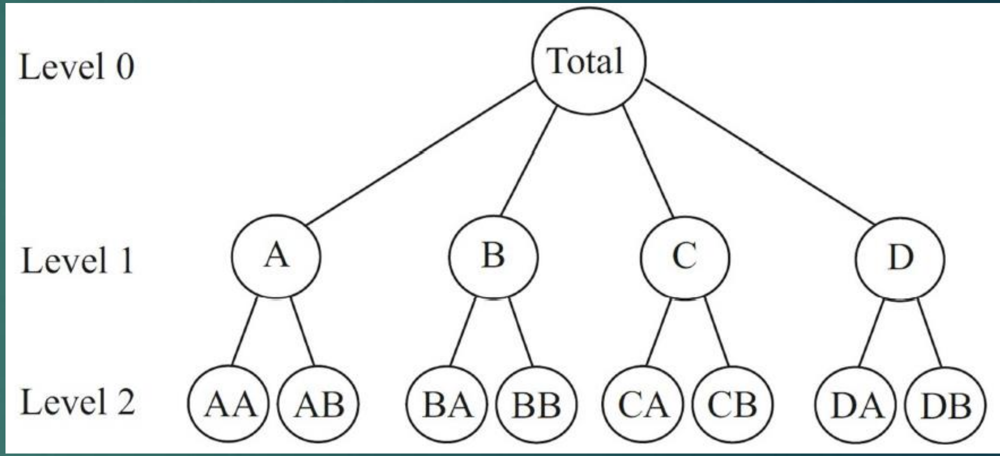
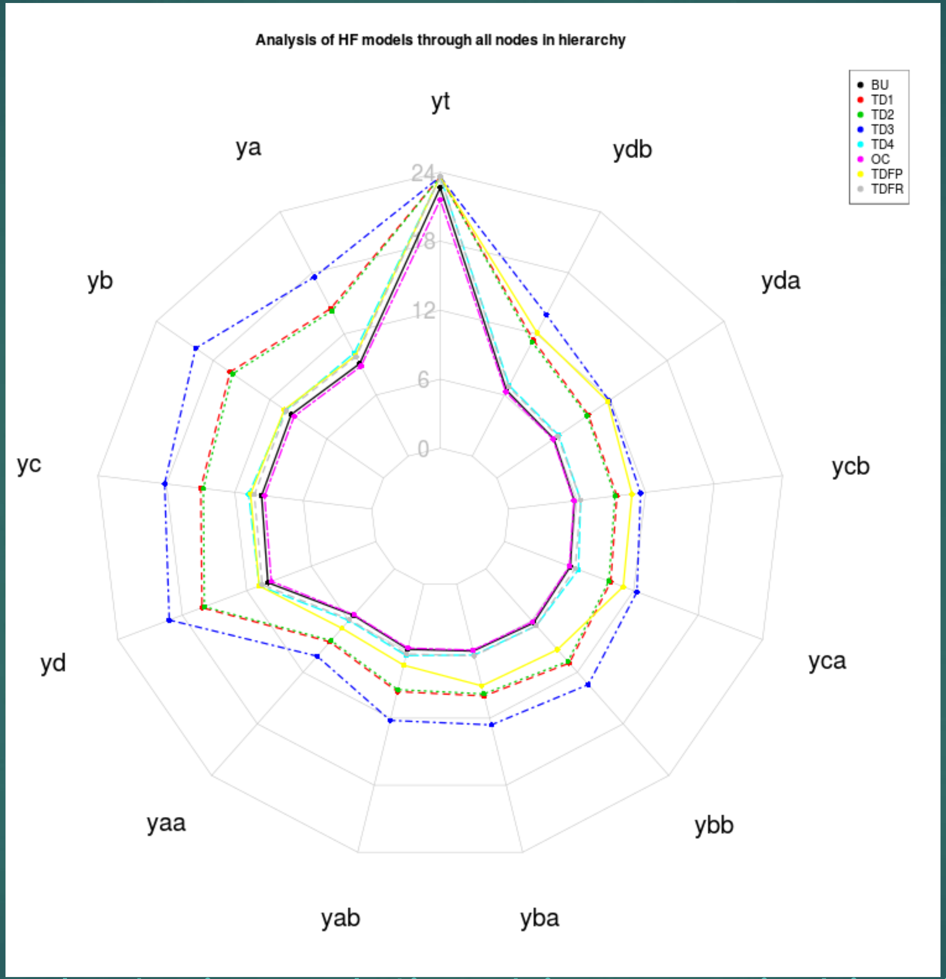


Fig. 2. RMSEs of competing models tested on simulated series.



Note: Bottom Up, Optimal Combination and Top down forecasted proportions produced the most accurate forecasts.

Which model to use?



https://supplychainanalytics.shinyapps.io/simulation_study/

More details: https://www.bahmanrt.com/publication/18w1/

7.7 WRAP-UP

- Forecasts can provide guidance on when to order new supplies to replenish stocks of items. However, a point forecast alone cannot provide this guidance. We also need to know the amount of uncertainty associated with the forecast. This is represented by a probability distribution.
- Forecasting packages will generally either use one of these distributions to calculate directly the stock needed to achieve a given customer service level or use the distribution to supply a forecast in the form of a prediction interval. We can use the prediction interval to determine how much safety stock is needed as long as we know the probability distribution that was used to obtain it.
- Some forecasting packages have a tendency to underestimate the level of uncertainty associated with future demand.
- When products form hierarchies, the forecasts at different levels are rarely consistent. A number of reconciliation processes are available, such as top-down and bottom-up forecasting.

Thank you for your attention!

Questions and comments?

Author contact information:
Dejan.Mircetic@uns.ac.rs

https://www.researchgate.net/profile/Dejan_Mircetic
http://www.ftn.uns.ac.rs/1417837974/dejan-mircetic