IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

# FORECASTING RAILWAY PASSENGERS DEMAND USING HOLT-WINTER METHOD WITH R STATISTICAL TOOL

**Dr. M.Rani Reddy**

Academic Consultant, Dept. of Computer Science, V.S.U.P.G. Centre, Kavali, A.P., India

*Email:* ranimarri@gmail.com

**Keywords:** *Holt-Winter Method, R Statistical Tool, Railway passengers*

----------------------------------------

**A B S T R A C T**

Holt-Winter Double exponential smoothing is a statistical method that can be used to forecast the demand of any object with time and seasonality statistics. This paper explores how we can use that method adopted to railway passengers' demand forecasting with the help of a software tool R which can support both Arithmetic and Statistical methods. This tool simply accepts the Input in proper format and forecasts the demand without the assistance of any other layouts which can be used in normal statistical methods to produce the result. R is a built in tool with graphical layouts. This paper explains how we can forecast demand with Exponential smoothing method using R tool.

***Citation:*** *M.Rani Reddy (2019). Forecasting Railway Passengers Demand Using Holt-Winter Method With R Statistical Tool: International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281, 2 (8), August, 2019, # Art.1811, pp 1-8*

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

## Introduction

The method Holt-Winters is a time series pattern variation method. Prediction or forecast everywhere needs a design, and this method is a type of triple orientations of the time ranges: an average, a slope or trend along with time, and a cyclic iterative behavior or seasonality. This model uses "double exponential smoothing "to encrypt a number of values from the previous and using them to forecast "typical" results in the future and present. The three objectives of the time range behavior: trend, seasonality and value are represented as three kinds of exponential smoothing, so Holt-Winters are also named as triple/double exponential smoothing. The method forecasts a presence or value of future by calculating the mixed results of these two or three affects. The method considers various attributes: one for every smoothing (α, β, γ), the period of a season, and the total number of intervals in a season. Seasonality can be complex. A season is a constant value of the time period that holds the complete iteration within the season itself, there are time intervals, which is the granularity of forecast. If you want to design a value for each hour of each day in a week, your season is multiples, i.e. 168 hours long and your time period is one hour.

### 2. Holt's Method:

2.1 Double Exponential Smoothing

The idea invoked to smooth the series with trend:

time-series separation base from trend effects is called ``De-trend"

Smoothing base with α

Smoothing trend predictions with β

Smoothing the base prediction Bt

Smoothing the trend prediction Tt

Predict k time intervals into future Ft+k using trend and base

when k=1

### 3. Winter's Method

Exponential Smoothing w/ Trend and Seasonality

Ideas behind smoothing with trend and seasonality:

"De-trend and de-seasonalize" time ranges by deviating trend from base and seasonal parameter influences

Smoothing base with α

Smoothing trend predicts with β

Smooth seasonality predicts with γ

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

IJAMSR

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Assume m seasons in a cycle

months in a year 12

quarters in a month 4

months in a quarter 3

### 3.1 Winter's Method:

Exponential Smoothing w/ Trend and Seasonality

Smooth the base forecast Bt

Smooth the trend forecast Tt

Smooth the seasonality forecast St

Forecast Ft with trend and seasonality

Smooth the trend forecast Tt

Smooth the seasonality forecast St

## 4. Methods Explanation:

### 4.1 Holt's Method ( Double Exponential Smoothing)

This method is achieved by including a double exponential smoothing model to capture the trend (either downwards or upwards). The method considers the following parameters for all I > 1

B1=D1, T1=0

where $0 < \alpha \leq 1$ and $0 \leq \beta \leq 1$.

An alternative form of these equations is

B1=D1 , T1=0

αei

+α βei

Where ei = Di –(Bi-1+Ti-1)=Di–Fi-1

Note that if β = 0, then the Holt model is equivalent to the Single Exponential Smoothing model.

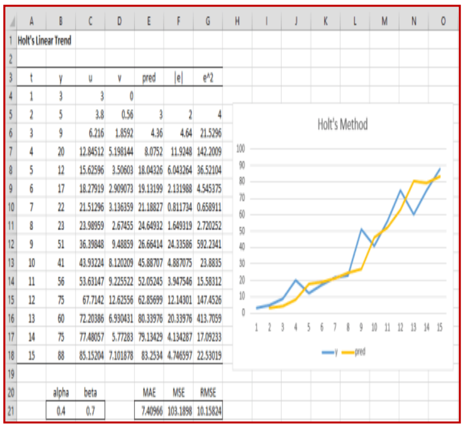
For Example using Holt's Method where α = .4 and β = .7.

The result is shown in the following Figure Using excel. Here the cell C4 contains the formula =B4, cell D4 contains the value 0, cell C5 contains the formula =B$21*B5+ (1-B$21) * (C4+D4), cell D5 contains the formula =C$21* (C5-C4) + (1-C$21) *D4 and cell E5 contains the formula =C4+D4. Here we use the details of railway passengers' monthly income without seasonal variations.

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281 Volume 2, Issue 7, August, 2019

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281



Holt's Linear Trend

| t | y | u | v | pred | \|e\| | e^2 |
|---|---|---|---|---|---|---|
| 1 | 3 | 3 | 0 | | | |
| 2 | 5 | 3.8 | 0.56 | 3 | 2 | 4 |
| 3 | 9 | 6.216 | 1.8592 | 4.36 | 4.64 | 21.5296 |
| 4 | 20 | 12.84512 | 5.198144 | 8.0752 | 11.9248 | 142.2009 |
| 5 | 12 | 15.62596 | 3.50603 | 18.04326 | 6.043264 | 36.52104 |
| 6 | 17 | 18.27919 | 2.909073 | 19.13199 | 2.131988 | 4.545375 |
| 7 | 22 | 21.51296 | 3.136359 | 21.18827 | 0.811734 | 0.658911 |
| 8 | 23 | 23.98959 | 2.67455 | 24.64932 | 1.649319 | 2.720252 |
| 9 | 51 | 36.39848 | 9.48859 | 26.66414 | 24.33586 | 592.2341 |
| 10 | 41 | 43.93224 | 8.120209 | 45.88707 | 4.887075 | 23.8835 |
| 11 | 56 | 53.63147 | 9.225522 | 52.05245 | 3.947546 | 15.58312 |
| 12 | 75 | 67.7142 | 12.62556 | 62.85699 | 12.14301 | 147.4526 |
| 13 | 60 | 72.20386 | 6.930431 | 80.33976 | 20.33976 | 413.7059 |
| 14 | 75 | 77.48057 | 5.77283 | 79.13429 | 4.134287 | 17.09233 |
| 15 | 88 | 85.15204 | 7.101878 | 83.2534 | 4.746597 | 22.53019 |

| alpha | beta | MAE | MSE | RMSE |
|---|---|---|---|---|
| 0.4 | 0.7 | 7.40966 | 103.1898 | 10.15824 |



**Figure 4.1 – Holt's Trend Analysis without seasonal variations**

### 4.2 Winter's method:

In the **Holt Winters Method (Triple Exponential Smoothing**), we add a seasonal component to the Holt's Trend Model. We explore two such models: the multiplicative seasonality and additive seasonality models. We consider the first of these models.

Let *c* be the length of a seasonal cycle. Thus m = 12 for months in a year, m = 7 for days in a week and m = 4 for quarters in a year. The method considers the below stated formula for all $i > c$

$$B_i = \alpha \frac{D_i}{S_{i-m}} + (1-\alpha)(B_{i-1} + T_{i-1})$$

$$T_i = \beta(B_i - B_{i-1}) + (1-\beta)T_{i-1}$$

$$S_i = \gamma \frac{D_i}{B_i} + (1-\gamma)S_{i-m}$$

where $0 < \alpha \leq 1$, $0 \leq \beta \leq 1$ and $0 \leq \gamma \leq 1-\alpha$.

The $B_i$ values represent the baseline, the $T_i$ values represent the trend (i.e. slope) and the $S_i$ values represent the seasonality component. In the multiplicative model, for any consecutive *c* periods of time, the sum of the $S_i$ values is approximately equal to 1.

The predictions for the data elements $F_i$ is given by

$$F_i = (B_{i-1}+T_{i-1})S_{i-m}$$

For forecasts at future times, we use the form

$$F_{i+h}=(B_i+hT_i)S_{i+h-mh'}$$

where $h'$ =INT(($h$–1)/m)+1 and m=c.

The initial values, i.e. where $1 \leq i \leq c$, are given by $B_1=u_c$ , $T_1=v_c$, $y_i=D_i$

$$u_c = \frac{1}{c}\sum_{j=1}^{c} y_j \qquad v_c = 0 \qquad s_i = \frac{y_i}{u_c}$$

Alternatively, we can set the initial trend value by using the average slope for the first two years, namely:

$$v_c = \frac{1}{c}\sum_{j=1}^{c} \frac{y_{c+j} - y_j}{c} = \frac{1}{c^2}\left[\sum_{j=1}^{c} y_{c+j} - \sum_{j=1}^{c} y_j\right]$$

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281 Volume 2, Issue 7, August, 2019

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

IJAMSR

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

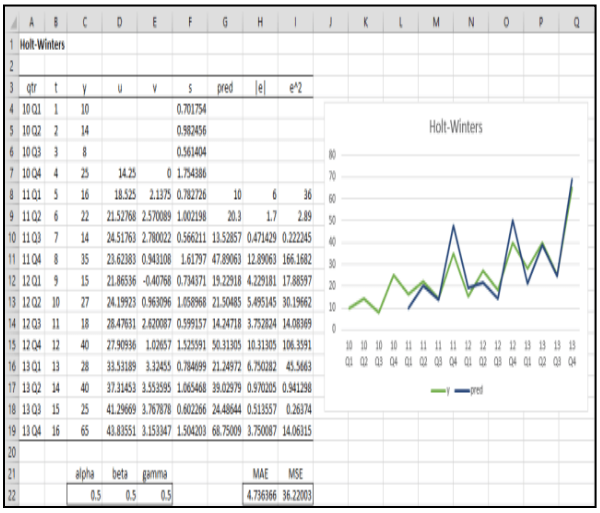
Note that if $\gamma = 0$, then the Holt-Winters model is equivalent to the Holt model and if $\beta = 0$ and $\gamma = 0$, then the Holt-Winters model is equivalent to the Single Exponential Smoothing model.

**For Example** Calculate the forecasted values of the time series shown in range C4:C19 of below figure 4.2 using the Holt-Winter method with $\alpha = .5$, $\beta = .5$ and $\gamma = .5$.

The result is shown in Figure 5.3 First we calculate $s_1$, $s_2$, $s_3$, $s_4$, where $c = 4$, as shown in range F4:F7. We do this by inserting the formula =C4/AVERAGE(C$4:C$7) in cell F4, highlighting the range F4:F7.

Next, we calculate $u_c$ and $v_c$ by placing the formula =C7/F7 in cell D7 and the value 0 in cell E7.

We now insert the formula =C$22*C8/F4+(1-C$22)*(D7+E7) in cell D8, the formula =D$22*(D8-D7)+(1-D$22)*E7 in cell E8, =E$22*(C8/D8)+(1-E$22)*F4 in cell F8 and the formula =(D7+E7)*F4 in cell G8, and then highlight the range D8:F19.



| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Holt-Winters | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | qtr | t | y | u | v | s | pred | \|e\| | e^2 |
| 4 | 10 Q1 | 1 | 10 | | | 0.701754 | | | |
| 5 | 10 Q2 | 2 | 14 | | | 0.982456 | | | |
| 6 | 10 Q3 | 3 | 8 | | | 0.561404 | | | |
| 7 | 10 Q4 | 4 | 25 | 14.25 | 0 | 1.754386 | | | |
| 8 | 11 Q1 | 5 | 16 | 18.525 | 2.1375 | 0.782726 | 10 | 6 | 36 |
| 9 | 11 Q2 | 6 | 22 | 21.52768 | 2.570089 | 1.002198 | 20.3 | 1.7 | 2.89 |
| 10 | 11 Q3 | 7 | 14 | 24.51763 | 2.780022 | 0.566211 | 13.52857 | 0.471429 | 0.222245 |
| 11 | 11 Q4 | 8 | 35 | 23.62383 | 0.943108 | 1.61797 | 47.89063 | 12.89063 | 166.1682 |
| 12 | 12 Q1 | 9 | 15 | 21.86536 | -0.40768 | 0.734371 | 19.22918 | 4.229181 | 17.88597 |
| 13 | 12 Q2 | 10 | 27 | 24.19923 | 0.963096 | 1.058968 | 21.50485 | 5.495145 | 30.19662 |
| 14 | 12 Q3 | 11 | 18 | 28.47631 | 2.620087 | 0.599157 | 14.24718 | 3.752824 | 14.08369 |
| 15 | 12 Q4 | 12 | 40 | 27.90936 | 1.02657 | 1.525591 | 50.31305 | 10.31305 | 106.3591 |
| 16 | 13 Q1 | 13 | 28 | 33.53189 | 3.32455 | 0.784699 | 21.24972 | 6.750282 | 45.5663 |
| 17 | 13 Q2 | 14 | 40 | 37.31453 | 3.553595 | 1.065468 | 39.02979 | 0.970205 | 0.941298 |
| 18 | 13 Q3 | 15 | 25 | 41.29669 | 3.767878 | 0.602266 | 24.48644 | 0.513557 | 0.26374 |
| 19 | 13 Q4 | 16 | 65 | 43.83551 | 3.153347 | 1.504203 | 68.75009 | 3.750087 | 14.06315 |
| 20 | | | | | | | | | |
| 21 | | | alpha | beta | gamma | | | MAE | MSE |
| 22 | | | 0.5 | 0.5 | 0.5 | | | 4.736366 | 36.22003 |



**Figure 4.2 – Holt-Winters Multiplicative Method**

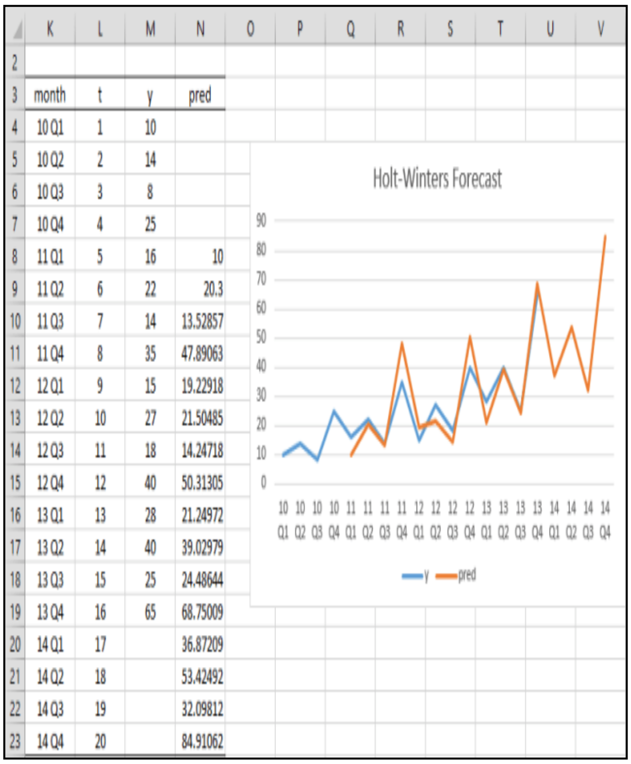
Forecast the y values for 2016 from previous data of railway passengers' income in 150 stations of south central railways Vijayawada division (i.e. the next 4 quarters).

The result is shown in Figure 5.4 the values through 2015 are copied from above. The forecasted value for Q1 of 2016 is 36.87209 (cell N20), as calculated by the following formula with reference to cells in Figure 5.4.
=(D$19+(L20-L$19)*E$19)*F16

The other three forecasted values are calculated by highlighting the range N20:N23 and pressing.

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281



| month | t | y | pred |
|---|---|---|---|
| 10 Q1 | 1 | 10 | |
| 10 Q2 | 2 | 14 | |
| 10 Q3 | 3 | 8 | |
| 10 Q4 | 4 | 25 | |
| 11 Q1 | 5 | 16 | 10 |
| 11 Q2 | 6 | 22 | 20.3 |
| 11 Q3 | 7 | 14 | 13.52857 |
| 11 Q4 | 8 | 35 | 47.89063 |
| 12 Q1 | 9 | 15 | 19.22918 |
| 12 Q2 | 10 | 27 | 21.50485 |
| 12 Q3 | 11 | 18 | 14.24718 |
| 12 Q4 | 12 | 40 | 50.31305 |
| 13 Q1 | 13 | 28 | 21.24972 |
| 13 Q2 | 14 | 40 | 39.02979 |
| 13 Q3 | 15 | 25 | 24.48644 |
| 13 Q4 | 16 | 65 | 68.75009 |
| 14 Q1 | 17 | | 36.87209 |
| 14 Q2 | 18 | | 53.42492 |
| 14 Q3 | 19 | | 32.09812 |
| 14 Q4 | 20 | | 84.91062 |



**Figure 4.3 – Holt-Winters' Multiplicative Forecast**

## 5. Holt – Winter's Method Using R

The statistical tool R will greatly assist for Holt-Winter smoothing model and predict. We make use of this function, "*HoltWinter & predict. HoltWinter*", to predict passengers demand based on past time periods data. HoltWinter function usage in R is easily understandable.

**Syntax**

HoltWinters(x, alpha = "NULL", beta = "NULL", gamma ="NULL", seasonal = c("additive", "multiplicative"), start.periods = 2, l.start = "NULL", b.start = "NULL", s.start = "NULL", optim.start = c(alpha = 0.3, beta = 0.1, gamma = 0.1), optim.control = list())

**Parameters**

X: class object ts

Alpha: Holt-Winters alpha value.

Beta: Holt-Winters beta value. If set to FALSE, the function will be simple exponential smoothing.

Gamma: seasonal component gamma value. If set to FALSE, it was a non-seasonal model.

Seasonal: Character parameter to select a model "additive" (the default) or "multiplicative". The first few characters are sufficient.

start.periods: Beginning intervals in the start values auto detection which is >=2.

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

l.start: level Start value (a[0]).

b.start: trend Start value (b[0]).

s.start: start values Vector for the component of seasonal variation (s1[0]…sp[0])

optim.start: The named components alpha, beta, and gamma vector holding the beginning values for the optimization. Only the needed values should be stated. Ignore in case of single-parameter.

optim.control: Minimal list with extended control variables passed to optimize if this is used. Ignored in case of single-parameter.

**5.R-Code for passengers seasonal trend analysis using Holt-Winter's Method**

```
> library(RODBC)

> channels <- odbcConnect("railwaydb", uid="rani", case="tolower")

## data frame is loaded within the warehouse

> data(stations)

> sqlSave(channels, stations, rownames = "stationname", addPK = TRUE)

> rm(stations)

## listing the relations in the warehouse

> sqlTables(channels)

2 stations TABLES

## list it

> sqlFetch(channels, "stations", rownames = "statationname")

Kakinadaport,kakinadatown,gudur,Nellore,----

> m->sqlQuery(channels, "select stationname, from stations where income > 1,00,00,000 and <3,00,00,000")

>a->ts(m, start = c(2000, 1), frequency = 12)

>hw->HoltWinters(a)

>forecast->predict(hw, n.ahead = 12, prediction.interval = T, level = 0.95)

>plot(hw, orecast)

>sqlDropTable(stations)

>sqlClose(channel)

>Close(forecast)
```

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281 Volume 2, Issue 7, August, 2019

IJAMSR 2 (8) www.ijamsr.com CrossRef: https://doi.org/10.31426/ijamsr.2019.2.8.1811

International Journal of
Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

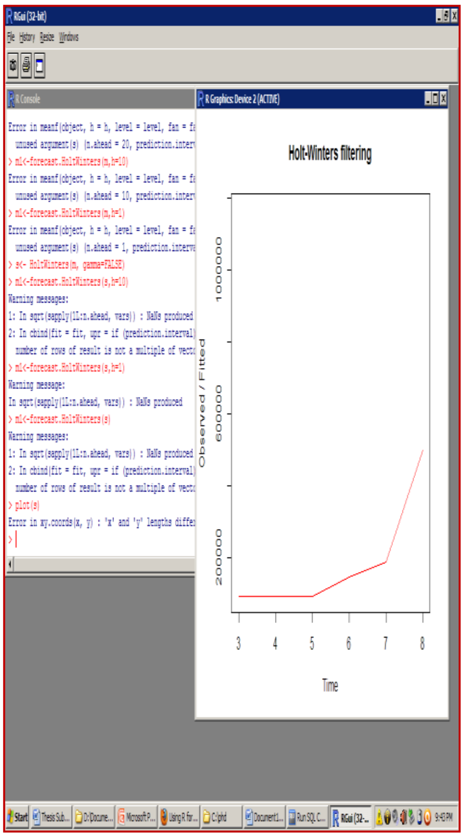



**Figure 5.1 Holt-Winter's seasonal Trend using R**

This is a simple code which explains how can we connect the data which is available in ORACLE through RODBC package available in R.after getting and filtering the data through the sqlqueries it can be saved in a temporary data store and then submitted to the functions forecast, Holt-Winters and predict as a parameter to forecast it can calculate using the syntax of HoltWinter's and produces the result then it will be plotted exploring the results of the function through a graph.

The graph shows total there are five categories of stations and only one station which is up on the graph can become category A1 grade in near future, but the others can take how much time depending on their earnings and amenities available on them can be shown in the graph. This is how the Holt-Winter's method can be used through R can forecast the demand.

Here the other categories A, B, C, D, E are of lower than A1. The government has estimated to put 375 stations as ADARSH STATIONS throughout the country. A station should categorize as ADARSH it should possess some features for that purpose these estimations are made and need to be forecasted within this division how many stations may reach that goal in how much time. For that purpose this work was done and estimations made clear at the end of this thesis.

**References:**

1) Yvonnick Noel. *Psychologie statistique avec R*. Pratique R. Springer, Paris, 2013.
2) Stano Pekar and Marek Brabec. *Moderni analyza biologickych dat. 2. Linearni modely s korelacemi v prostredi R [Modern Analysis of Biological Data. 2. Linear Models with Correlations in R]*. Masaryk University Press, Brno, 2012. ISBN 978-80-21058-12-5. In Czech.
3) Sarah Stowell. *Instant R: An Introduction to R for Statistical Analysis*. Jotunheim Publishing, 2012. ISBN 978-0-957-46490-2.
4) Bernhard Pfaff. *Financial Risk Modelling and Portfolio Optimisation with R*. Wiley, Chichester, UK, 2012. ISBN 978-0-470-97870-2. ]
5) Dimitris Rizopoulos. *Joint Models for Longitudinal and Time-to-Event Data, with Applications in R*. Chapman & Hall/CRC, Boca Raton, 2012. ISBN 978-1-4398-7286-4.
6) Hyndman, R.J., Koehler, A.B., Ord, J.K., Snyder, R.D. (2008) *Forecasting with exponential smoothing: the state space approach*, Springer-Verlag: New York. http://www.exponentialsmoothing.net.
7) Hyndman, R.J., Athanasopoulos (2012) *Forecasting: principles and practice*, OTexts: Melbourne, Australia.
8) Shravan Vasishth and Michael Broe. *The Foundations of Statistics: A Simulation-based Approach*. Springer, 2010. ISBN 978-3-642-16312-8.
**9)** Robert A. Muenchen and Joseph M. Hilbe. *R for Stata Users*. Statistics and Computing. Springer, 2010. ISBN 978-1-4419-1317-3., Rob Kabacoff. *R in Action*. Manning, 2010.