Roman SZOSTEK[1]

MODIFICATION OF HOLT'S MODEL EXEMPLIFIED BY THE TRANSPORT OF GOODS BY INLAND WA-TERWAYS TRANSPORT

In the paper there were presented the modifications of Holt's method. Firstly, it was assumed that the values of the parameters appearing in Holt's model do not need to be limited, as is commonly assumed, to the range of [0, 1]. Secondly, it was proposed the more precise way of forecting of range series for more distant time. The aim of the paper is to propose some modifications to Holt's model which will allow for better forecasting. The paper presents an idea and the effects of the proposed modifications exemplified by the data on the transport of goods by inland waterways transport in Poland.
Keywords: Holt's method, optimization, forecasting.

1. INTRODUCTION

Holt's method is one of exponential smoothing methods. It is based on the smoothing of the analyzed time series by means of a moving average and it is used to smooth the time series in which there is the tendency of random fluctuations. Thanks to the series smoothing it is obtained a piece of information about its properties used to determine the forecast.

Exponential smoothing may be based on different models, respectively matched to the course of the series. In addition to Holt's model it is used a simple model of exponential smoothing as well as Winter's model.

2. HOLT'S MODEL

Holt's model allows for smoothing time series, in which there is the trend and random fluctuations. The values of the forecasted series are indicated by symbols $x_0, x_1, ..., x_{n-1}$. This model has two parameters α and β and the following form:

$$F_1 = x_1, \tag{1}$$

$$S_1 = x_1 - x_0, \tag{2}$$

$$F_t = \alpha x_t + (1-\alpha)(F_{t-1} + S_{t-1}), \tag{3}$$

$$S_t = \beta(F_t - F_{t-1}) + (1-\beta)S_{t-1}, \tag{4}$$

where: $t = 2, ..., n-1$,
    F_t –   the smoothed value of time series
    S_t –   the smoothed value of the trend increment on the moment t,

---

[1] PhD, Eng., Roman Szostek, Department of Quantitative Methods, The Faculty of Management, Rzeszow University of Technology.

α, β –   model parameters.

The values of F_t and S_t are calculated in a recursive manner. Forecast of future values of a series are calculated as follows:

$$x^*_{n+k-1} = F_{n-1} + k \cdot S_{n-1}, \qquad (5)$$

where: $k = 1, 2, 3, \ldots$

Holt's model parameters α and β are chosen as to minimize the errors of the expired forecast. For this purpose the values of any specific parameter are assumed and determined in accordance with (5) for $k = 1$ of the expired forecast

$$x^*_t = F_{t-1} + S_{t-1} \qquad (6)$$

for the moment of time t ($t = 2, 3, \ldots, n-1$) on the basis of the value of the time series from the earlier period ($x_0, x_1, \ldots, x_{t-1}$). This forecast can be compared with the actual values of a x_t series. The resulting differences are the errors of the expired forecast which gives the model for the adopted parameters α and β. As a measure of the quality of the method it should be assumed the average from the errors of the expired forecas. It can be an average mean

$$J_1 = \frac{1}{n-6} \sum_{t=6}^{n-1} \left| F_{t-1} + S_{t-1} - x_t \right| \qquad (7)$$

or qudratic mean

$$J_2 = \sqrt{\frac{1}{n-6} \sum_{t=6}^{n-1} (F_{t-1} + S_{t-1} - x_t)^2} \; . \qquad (8)$$

Finally, among all the possible values of the parameters α and β there should be found those that give the smallest error value J_1 or J_2. In this way, there are determined the optimum values of model parameters, i.e. its optimization is carried out. The value of J_* is a measure of forecast error determined by the model.

It is commonly assumed that $\alpha \in [0, 1]$ and $\beta \in [0, 1]$. It seems, however, that this limitation is unnecessary, therefore in the carried out calculations it was omitted. In this way it was assumed that the best model parameters are such for which the model determines the expired forecast in the best way irrespectively whether these parameters are higher or lower than the unity.

The evaluation of the optimum values of parameters α and β is conducted by means of numarical methods.

3. DATA ANALYSIS

The analysed time series was presented in table 1. It contains the information on the data on the transport of goods by inland waterways transport in Poland in annual periods.

The values of the optimum parameters of Holt's model are different depending on the applied quality measure. The parameters grid was checked with the accuracy of 0.0001. The forecasts were determined for the next future periods. Fo the linear meaure (7) the optimum are parameters

$$\alpha = 1.2428 \text{ and } \beta = -0.0336. \qquad (9)$$

For the squared measure (8) the optimum are parameters
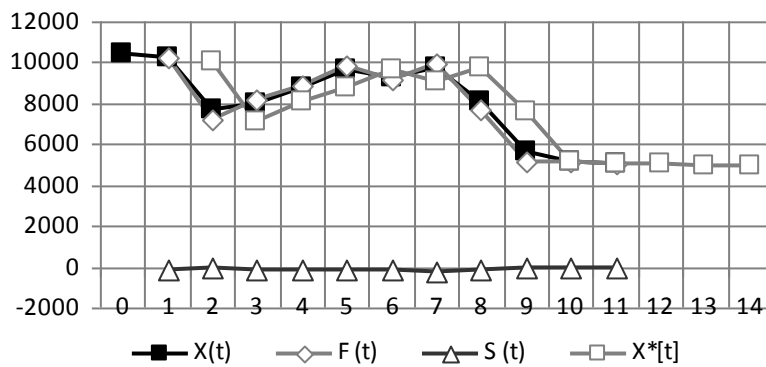
$$\alpha = 1.3884 \text{ and } \beta = -0.0145. \tag{10}$$

Table 1. Passengers air transport in Poland

Year	Transport (thousand tonnes)	Year	Transport (thousand tonnes)
2000	10433	2006	9271
2001	10255	2007	9792
2002	7729	2008	8109
2003	7968	2009	5655
2004	8747	2010	5141
2005	9607	2011	5093

Source: Central Statistical Office (www.stat.gov.pl) [1][2].

Time series together with the forecast for the measure (7) was presented in fig.1.
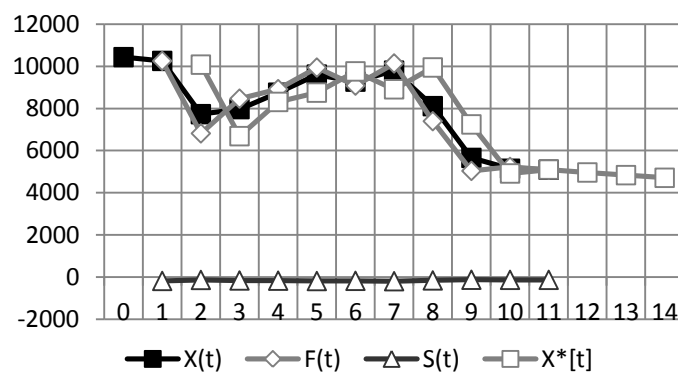
Fig. 1. Forecast, when in model $\alpha = 1.2428$ and $\beta = -0.0336$.

Source: Own study.

Time series together with the forecast for the measure (8) was presented in fig.2.

Fig. 2. Forecast, when in model $\alpha = 1.3884$ and $\beta = -0.0145$.

Source: Own study.

For comparison, when the parameters value of the model α and β are limited to the range [0, 1], then the optimum values of these parameters for the linear measure (7) are
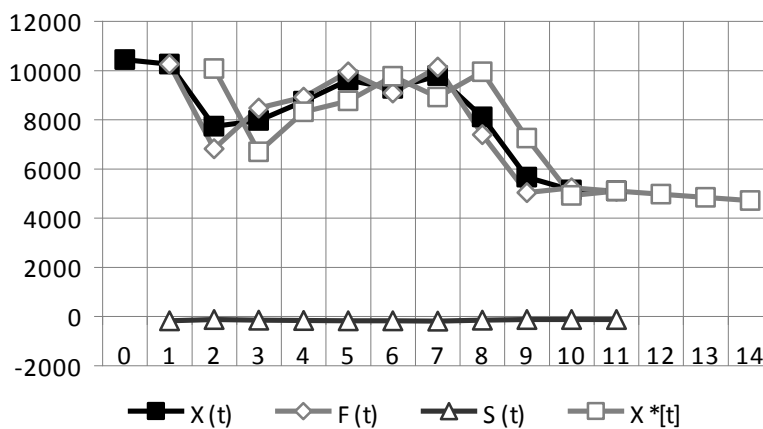
$$\alpha=1.0000 \text{ and } \beta=0.0000, \tag{11}$$

whereas for the squared measure (8) they are

$$\alpha=1.0000 \text{ and } \beta=0.0372. \tag{12}$$

Time series together with the forecast for the case when $\alpha = 1.0$ and $\beta = 0.0$ was shown in fig. 3.

Fig.3. Forcast, when in model $\alpha =1.0$ and $\beta =0.0$.

Source: Own study.

Table 2. Calculation results. Comparison of three solutions

t	X(t)	α=1.2428 β=-0.0336			α=1.3884 β=-0.0145			α=1,000 β=0,000		
		F(t)	S(t)	X*[t]	F(t)	S(t)	X*[t]	F(t)	S(t)	X*[t]
0	10433	-	-	-	-	-	-	-	-	-
1	10255	10255	-178	-	10255	-178	-	10255	-178	-
2	7729	7158.9	-80.0	10077	6817.0	-130.7	10077	7729	-178	10077
3	7968	8183.9	-117.1	7079.0	8465.8	-156.5	6686.3	7968	-178	7551
4	8747	8912.2	-145.5	8066.8	8917.0	-165.4	8309.3	8747	-178	7790
5	9607	9811.0	-180.6	8766.7	9939.2	-182.6	8751.7	9607	-178	8569
6	9271	9183.7	-165.6	9630.5	9082.4	-172.8	9756.7	9271	-178	9429
7	9792	9979.9	-197.9	9018.2	10134.7	-190.6	8909.6	9792	-178	9093
8	8109	7702.8	-128.0	9782.0	7396.2	-153.6	9944.2	8109	-178	9614
9	5655	5188.9	-47.9	7574.8	5038.4	-121.7	7242.6	5655	-178	7931
10	5141	5141.0	-47.9	5141.0	5228.1	-126.2	4916.7	5141	-178	5477
11	5093	5093.0	-47.9	5093.1	5089.5	-126.0	5102.0	5093	-178	4963
16				5045.1			4963.5			4915.0
17		forecast:		4997.3	forecast:		4837.6	forecast:		4737.0
18				4949.5			4711.5			4559.0
lin. error		**787.7**			837.4			850.7		
squ. error		1096.4			**1076.5**			1161.1		

Source: Own study.

In table 2 there were shown the results of calculations for the above three cases. The optimum values of quality measures were shown in bold in the table.

Limitation of the values of the parameters α and β to the range [0, 1] makes that to the large family of a time series it is assigned the same Holt's model with parameter $\alpha = 1.0$ (or $\beta = 1.0$). That is so for the considered range. In this context, a wide variety of series is described by the same model. At the same time the parameters of the Holt's model can be selected in order to determine the expired forecast in a more accurate way. These parameters should be used for forecasting.

4. FORECAST FOR FURTHER INSTANTS OF TIME

In the traditional approach presented in the form of equations (1)-(5) the optimum parameters of the model are based on the expired forecast calculated for one step forewards. If, however, by using the model the forecast for k steps forwards needs to be determined, it should be performed the optimization model taking into account the expired forecast for *k* steps forwards. Thus, it was suggested that for every instant of time which the forecast is to be determined, it should be determined another model, each time on the basis of different quality measure.

If the prediction is to be performed on the na *k*-th step forwards, then the model parameters should be determined by minimizing a measure of the quality of

$$J_1(k) = \frac{1}{n-5-k} \sum_{t=5+k}^{n-1} \left| F_{t-k} + k \cdot S_{t-k} - x_t \right| \tag{13}$$

or

$$J_2(k) = \sqrt{\frac{1}{n-5-k} \sum_{t=5+k}^{n-1} (F_{t-k} + k \cdot S_{t-k} - x_t)^2} \ , \tag{14}$$

where: $k = 1, 2, 3, \ldots$

The calculations were made with an application of C++ of the following code:

```
int main(int argc, char *argv[])
{double X[12]={10433, 10255, 7729, 7968, 8747, 9607, 9271, 9792, 8109, 5655, 5141,
5093};
double F[12], S[12], F_opt, S_opt;    // the table needs to be of the n size
double J2, J1, a_opt, b_opt, J2_min=999999999, J1_min=999999999;
double a, b, krok=0.0001;
int t, n=12, k=1;   // k means on how many steps forwards is the forecast, k=1, 2, 3, ...

F[1]=X[1];  S[1]=X[1]-X[0];
for (a=0.0; a<=1.0; a=a+krok)
{    for (b=0.0; b<=1.0; b=b+krok)
    {    J2=0; J1=0;
        for (t=2; t<n; t++)
            {    F[t]=a*X[t]+(1-a)*(F[t-1]+S[t-1]);
                S[t]=b*(F[t]-F[t-1])+(1-b)*S[t-1];
                if (t>=(5+k))
```

```
            {    J2=J2+pow((F[t-k]+k*S[t-k]-X[t]),2);   //squared error
                 J1=J1+fabs(F[t-k]+k*S[t-k]-X[t]);       //linear error
            }
        }
        if (J1 < J1_min)                         //minimization of linear error
        {    J2_min=J2; J1_min=J1;
             a_opt=a; b_opt=b;
             F_opt=F[n-1]; S_opt=S[n-1];
        }
    }
}
J1_min=J1_min/(n-5-k);                           //linear error
J2_min=sqrt(J2_min/(n-5-k));                     //squared error
cout << "optimum a = " << a_opt << endl;
cout << "optimum b = " << b_opt << endl;
cout << "forecast for the moment " << n+k-1 << " = " << F_opt+k*S_opt << endl;
cout << "linear error J1 = " << J1_min << endl;
cout << "squared error J2 = " << J2_min << endl;
system("PAUSE");
}
```
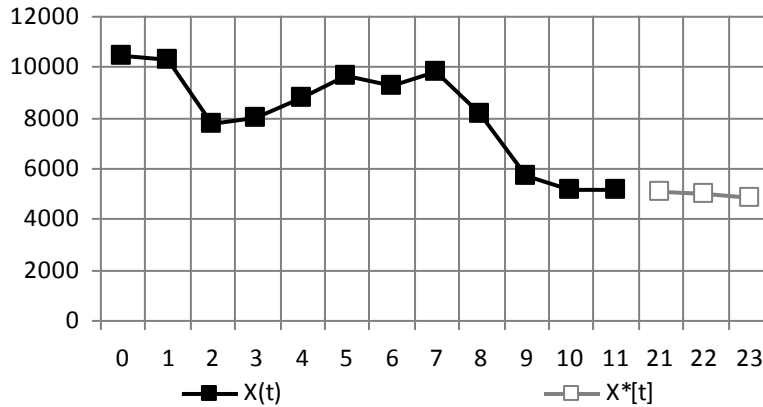
Forecast value for k steps forwards is calculated according to equation (5). The results of calculations for the quality measure (13) are presented in table. 3 and in figure 4. In the case of the modified method the forecasts do not have to be a straight line, as it is always in case of the conventional method.

Fig. 4. Forecast for the modified quality measure $J_1(k)$

Source: Own study.

Thanks to the presented modification there were obtained lower average errors of expired forecast. They are summarized in table. 4 (for linear quality measure). These errors for the traditional method were determined for a model in which α and β have the values (9), but according to the equation (13). Yet, for the modified method the errors of expired forecast come from table 3.

Table 3. Calculation results for the modified quality measure $J_1(k)$

t	X(t)	k=1 $\alpha=1.2428$ $\beta=-0.0336$			k=2 $\alpha=1.7249$ $\beta=-0.0168$			k=3 $\alpha=0.0016$ $\beta=47.1423$		
		F(t)	S(t)	X*[t]	F(t)	S(t)	X*[t]	F(t)	S(t)	X*[t]
0	10433	-	-	-	-	-	-	-	-	-
1	10255	10255.0	-178.0	-	10255.0	-178	-	10255.0	-178.0	-
2	7729	7158.9	-80.0	10077.0	6026.9	-110.0	-	10073.2	-355.1	-
3	7968	8183.9	-117.1	7079.0	9454.8	-169.4	9899.0	9715.3	-487.1	-
4	8747	8912.2	-145.5	8066.8	8356.7	-153.8	5807.0	9227.5	-523.4	9721.0
5	9607	9811.0	-180.6	8766.7	10624.8	-194.5	9116.0	8705.5	-455.3	9007.9
6	9271	9183.7	-165.7	9630.5	8430.6	-160.9	8049.1	8251.8	-378.3	8254.0
7	9792	9979.9	-197.9	9018.2	10895.5	-205.0	10235.9	7876.6	-233.6	7657.2
8	8109	7702.8	-128.0	9782.0	6237.7	-130.2	8108.8	7643.7	-198.5	7339.6
9	5655	5188.9	-47.9	7574.8	5327.0	-117.1	10485.5	7442.4	-333.5	7116.9
10	5141	5141.0	-47.9	5141.0	5091.0	-115.1	5977.3	7105.8	-481.9	7175.8
11	5093	5093.0	-47.8	5093.2	5177.8	-118.5	5092.8	6621.4	-597.4	7048.4
16		prognoza:		5045.1						
17					prognoza:		4940.9			
18								prognoza:		4829.2
Forecast error.		$J_1(1) =$		787.7	$J_1(2) =$		1222.2	$J_1(3) =$		1555.4

Source: Own study.

Table 4. Comparison of average errors of expired forecast for linear quality measurement

	Forecast for k steps forwards		
	k=1	k=2	k=3
Traditional method	787.7	1464.2	2665.8
Modified method	787.7	1222.2	1555.4

Source: Own study.

Based on the criterion which is the average error of expired forecast, one can conclude that the modified method allows for the determination of more reliable forecasts.

5. CONCLUSIONS

In the classic applications of Holt's model the values of its parameter are restricted to the range [0, 1]. This approach has been used for example in a Statistica package. In the paper it was proposed not to restrict the parameters of the model to the range [0, 1]. Such models can better determine the expired forecast. So they are a better way to determine future forecasts.

The limitation of the parameters values α and β to the range [0, 1] probably resulted from the idea that the values of series F_t and S_t are in a certain percentage of the previous values of these series, and the rest of the values of another factor (in accordance with (3) and (4)). Resignation of reducing the value of these parameters is a natural generalization of the method.

In the cases where in the model the parameter $\alpha>1$ was obtained, the series F_t was not always the smoothing of x_t series. For some of the analyzed examples the values of F_t series oscillated at around the value of x_t series. In that case F_t series did not smooth but on the contrary sharpened the fluctuation of the original series. However, always modified

models determined the expired forecast in a better way, so they are a better tool to determine future forecasts.

In the paper it was also proposed a modified way of Holt's model optimization. It consists in independent optimization of some models, one for each forecast for the *k*-th period forwards. In this way there are determined so many Holt's models on how many steps forwards the forecast is made. Such a modified model determines the expired forecast with less average error. It is, therefore, a better way to determine future forecast. Thus obtained forecast doe not need to be a straight line (Fig. 4), as is always in the classic approach.

6. REFERENCES

[1]   Central Statistical Office – Transport – activity results,
       http://www.stat.gov.pl/cps/rde/xbcr/szczec/ASSETS_sygnalna_transport_wodny_2011_pop.pdf
[2]   Central Statistical Office – Transport – activity results,
       http://www.stat.gov.pl/cps/rde/xbcr/gus/PUBL_til_transport_wyniki_dzialalnosci_2010.pdf

MODYFIKACJA MODELU HOLTA NA PRZYKŁADZIE PROGNOZY TRANSPORTU ŁADUNKÓW W ŻEGLUDZE ŚRÓDLĄDOWEJ W POLSCE

W artykule przedstawiony został zaproponowany przez autora zmodyfikowany model Holta. Pierwsza modyfikacja polega na przyjęciu, że parametry występujące w modelu Holta nie muszą być ograniczone, jak się powszechnie przyjmuje, do przedziału [0, 1]. Druga modyfikacja wiąże się z precyzyjniejszym sposobem prognozowania wartości szeregu dla bardziej odległych chwil czasu. Celem opracowania jest przedstawienie modyfikacji modelu Holta, które pozwalają na uzyskanie lepszych prognoz. W artykule zaprezentowano ideę oraz przedstawiony został przykład zastosowania zmodyfikowanej metody do prognozowania wielkości transportu ładunków w żegludze śródlądowej w Polsce.
Słowa kluczowe: metoda Holta, optymalizacja, prognozowanie.