Forecasting Methods in Logistics Problems

Anna V. Zykina
Omsk State Technical University
Omsk, Russia, Mira Ave. 11, 644050
avzykina@mail.ru

Angela A. Kanunnikova
Omsk State Technical University
Omsk, Russia, Mira Ave. 11, 644050
marymind99@gmail.com

Abstract

The article considers the use of forecasting methods in solving problems of delivering bottled water to customers. It also examines the method of regression analysis of data. Moreover, a mathematical model is developed to predict the amount of water in the following months. To determine the optimal allocation of financial resources, a multi-index transport problem is constructed. The constructed models are implemented in Python.

1 Introduction

Logistics is the science of managing material flows, as well as related information, financial, service and other flows, to optimize the functioning of economic systems through the efficient utilization of all types of resources [Sak2018].

The main strategic goals of logistics for this study are:

1) to ensure a continuous flow of external resources needed for the operation of the enterprise;

2) to optimize the overall cost of resources of the enterprise, as well as to manage the stocks of purchased products;

3) to manage suppliers, namely, their selection and evaluation, and development of relationships with them;

4) to coordinate the procurement process with other functional areas of logistics,e. g. the logistics of production, distribution, and transportation.

This study uses methods of mathematical statistics and optimization methods to achieve its goals.

In market conditions, several factors determine the relevance of companies' use of logistics.

First of all, there is an economic factor. Every company aims to maximize profits, so the focus is on finding ways to reduce the time and money costs that come with products transportation.

Another factor is the organizational-economic one. The market is developing, new organizational forms are emerging, and the development of interaction between manufacturers, consumers, and transport companies is becoming increasingly important.

The rapid development of technology, as well as the growth of the information sector of the economy, contributes to the search for different methods to increase company profits and reduce its costs [Pic2020, Mas2006]. Opportunities and tools for the management of logistics processes are emerging. New theories and methods of operation research are being developed that can help to optimize the processes.

Constant development of market relations, technical opportunities, as well as a wide range of goals and objectives of logistics, makes this field relevant to the study and highlights new challenges to be addressed.

In: Sergei S. Goncharov, Yuri G. Evtushenko (eds.): Proceedings of the Workshop on Applied Mathematics and Fundamental Computer Science 2021 , Omsk, Russia, 24-04-2021, published at http://ceur-ws.org

2 Problem statement

Currently, many companies propose products' delivery services. One group of such companies are those delivering bottled water.

In modern society, the availability of clean drinking water in the workplace is one of the requirements of sanitary and epidemiological regulations. The number of business and industrial enterprisers, where these regulations must be followed, is quite large. Therefore, water delivery companies are in high demand.

Since water delivery companies, like any other company, strive to minimize costs and improve the quality of their services, their work is carried out with the use of logistics processes.

There are several main work areas for logisticians:

– resource procurement;

– production;

– end product sales.

When the subject area was studied, two tasks were identified:

– forecast of water demand;

– optimization of the bottled water delivery to customers.

A mathematical model was developed for each problem, parameters were selected, and the results were interpreted.

The essence of modeling using economic-statistical methods is that the forecast indicator is determined through specific models that show its functional dependence on certain factors [Chi2020, Ash2019].

Problem 1. Forecast of water demand.

The water delivery company should anticipate what demand is expected in the future. The option to prepare excess products is not suitable in the case of bottled water. Firstly, the warehouse has a limited capacity. Secondly, bottled water has a limited expiry date even when closed, so it is necessary to make a certain amount of water so as not to suffer losses.

Consumer companies do not always order the same amount of water each time. This can be influenced by various factors. Therefore, the following forecast task is being investigated: how much water the company will order, or how much water the company should order to provide sufficient water amount for employees.

To be able to make a forecast, it is necessary to build a model which will be used to predict the amount of water. Three types of parameters will be used for constructing the model: the number of employees, the number of working days and the time of the year. Let us build a regression model using these parameters. Linear regression is a standard representative of machine learning algorithms. Regression belongs to the class of supervised learning tasks, where a particular target variable must be predicted for a given set of characteristics of the object under observation [Lin2020].

It is necessary to predict the amount of water, so the data on the number of bottles ordered is used as the dependent variable y. The independent variables are the number of employees in the company x_1, the number of working days in a month x_2, and the time of the year $z_1,\ z_2,\ z_3$.

To build the model, monthly data for one company have been taken for two years. The method of least squares is used to find the model parameters. Instead of the season variable, which is nominative, let us introduce three dummy variables: winter month or not; spring month or not; summer month or not.

As a result of the calculation, the following regression equation was obtained:

$$y = -6,0534 + 1,5135x_1 + 1,7924x_2 - 1,8131z_1 + 1,1467z_2 - 0,9204z_3.$$

To evaluate the quality measure of the regression equation, the determination coefficient R^2 was calculated. In this case, $R^2 = 0.86$, which means that the resulting regression equation explains the change in the resulting variable y by 86. The remaining 14 are accounted for by unaccounted factors.

To check the regression equation for significance, an F-test is used. The "zero" hypothesis H_0 of statistical insignificance of the regression equation is put forward. In the model, criterion $F_{obsrv} = 28.284$. Using F-criterion table we find the critical value $F_{crit}(0.01; 5; 23) = 2.64$. Compare the two values: $F_{obsrv} > F_{crit}$. This means that the zero hypothesis is rejected, and the equation is significant.

To check the correctness of the construction and further automate the calculation of the forecast, the model is implemented using the Python language. The language allows one to work with big data, build various models and present a visualization of the resulting calculations. To build the model, the *Linear Regression*() function from the *sklearn* library is used.

The data for the construction are loaded from the excel-file (Figure 1).

```
from sklearn import linear_model
excel_data_df = pd.read_excel('/content/sample_data/test_regr1.xls')
excel_data_df.columns = ['Y', 'X1', 'X2', 'X3', 'X4']
excel_data_df.head()
```

	Y	X1	X2	X3	X4
0	85	45	17	1	2019-01-01
1	106	48	20	1	2019-02-01
2	108	48	20	2	2019-03-01
3	102	50	22	2	2019-04-01
4	91	41	18	2	2019-05-01

Figure 1: Data view in the file

As a result of using the function, the following data were obtained for the model:
– intercept is free coefficient;
– slope is the slope of the linear regression, are coefficients at independent variables of the model;
– and coefficient of determination R^2 (Figure 2).
As seen from $Figure2$, the data from manual calculations and those made with Python are the same.

```
intercept: -6.053404055739961
slope: [ 1.51349773  1.79244135 -1.81307922  1.14667067 -0.92041444]
coefficient of determination: 0.8631556745017089
```

Figure 2: Formation of the mathematical model in Python

Using the model, let us predict the amount of water for the next few months. To do this, data should be entered for each of the variables. The result will be a forecast of how many bottles of water the company will need in a given month (Figure 3).

```
Input the number of employees: 36
Input the number of working days: 23
Input the prediction month: December
Input the year: 2020
Prediction of the water bottles number: [87]
```

Figure 3: Data entry and forecasting

To view the dynamics of the water orders, let us build a diagram with the data used to find the model, and mark the point of the predicted quantity on the diagram (Figure 4).

The implementation results in a total of 87 bottles of water for 36 employees for December 2020, which has 23 workdays.

Problem 2. Bottled water delivery to customers.

As a rule, the company's costs associated with the transportation of products are not predicted. The calculation of operating efficiency is based on the actual costs of the transportation division. This makes it impossible to estimate costs in advance when planning transportation. Moreover, the task of determining the optimal route for the transportation organization, together with the task of loading transport and product allocation, are the most important in transport logistics. The task of logistics management of product flows to minimize costs and maximize profits in grocerycompanies is a complex methodological problem, since the technological chain of food production includes a large number of interacting participants. It is necessary to include all participants of production, their interaction, peculiarities of organization and operating technology [Mok2020].

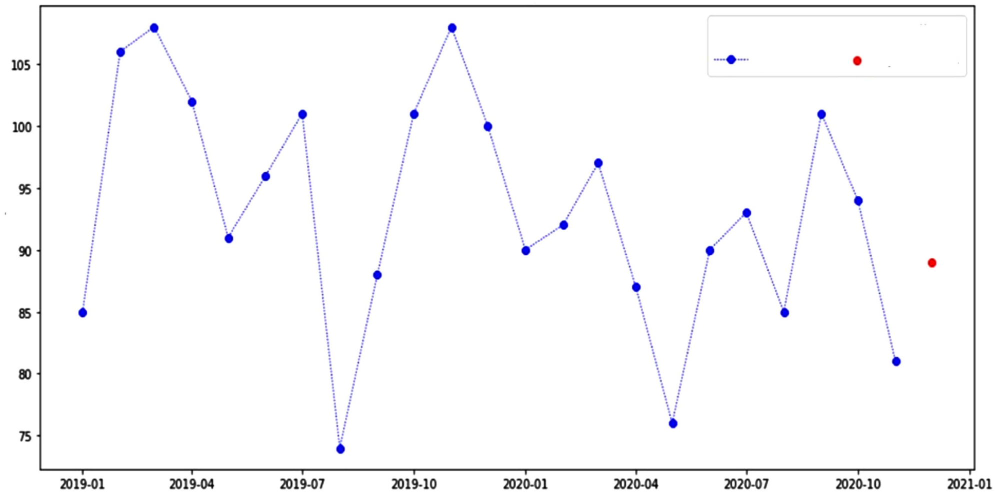

Figure 4: Schedule of changes in water delivery

Based on the forecast of water demand, the problem of delivering bottled water to customers can be solved as a linear programming problem of transport type. Let us consider the formulation of the mathematical model of the problem.

There are n kinds of water bottles, m shipment points, l consumer delivery points, g vehicles of the first type, and f vehicles of the second type, which will be used to deliver products.

Set constant parameters:

$-a_{ij}$ is the stock of the i-th type of product in the j-th point;

$-b_{ik}$ is the demand of the k-th consumer of the i-th type of product;

$-d_{ip}$ is thethe amount of the i-th type of product, that can be loaded in the transport of the first type;

$-e_{it}$ is the amount of the i-th type of product that can be loaded in the transport of the second type;

$-c_{ijkp}$ is transport costs per unit of the product of the i-th type of the j-th supplier delivered to the k-th consumer by transport of the first type;

$-q_{ijkt}$ is transport costs per unit of the product of the i-th type of the j-th supplier delivered to the k-th consumer by the second type of transport.

The following parameters are selected as variables:

$-x_{ijkp}$ is the amount of the i-th type of product from the j-th shipment point to be delivered to the k-th consumer by the first type of transport;

$-y_{ijkt}$ is the amount of the i-th type of product from the j-th shipment point to be delivered to the k-th consumer by the second type of transport.

The mathematical model of the problem has the form:

$$x_{ijkp} \geq 0,\ y_{ijkt} \geq 0,\ i = \overline{1,n}\ , j = \overline{1,m}\ , k = \overline{1,l},\ p = \overline{1,g},\ t = \overline{1,f}, \tag{1}$$

$$\sum_{k=1}^{l} \sum_{p=1}^{g} x_{ijkp} + \sum_{k=1}^{l} \sum_{t=1}^{f} y_{ijkt} \leq a_{ij},\ i = \overline{1,n}\ , j = \overline{1,m}, \tag{2}$$

$$\sum_{j=1}^{m} \sum_{p=1}^{g} x_{ijkp} + \sum_{j=1}^{m} \sum_{t=1}^{f} y_{ijkt} = b_{ik},\ i = \overline{1,n},\ k = \overline{1,l}, \tag{3}$$

$$\sum_{i=1}^{n} x_{ijkp} \leq d_{ip},\ j = \overline{1,m},\ k = \overline{1,l},\ p = \overline{1,g}, \tag{4}$$

$$\sum_{i=1}^{n} y_{ijkt} \leq e_{it},\ j = \overline{1,m},\ k = \overline{1,l},\ t = \overline{1,f}, \tag{5}$$

$$L = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}\sum_{p=1}^{g} c_{ijkp}\ x_{ijkp}\ +\ \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}\sum_{t=1}^{f} q_{ijkt}\ y_{ijkt}\ \rightarrow min \tag{6}$$

All elements x_{ijkp} and y_{ijkt} must be either positive or equal to zero (1). This means that from the shipment point a certain amount of a certain type of product will be transported to the consumer, or the product will not be transported to them. The solution matrix cannot contain negative numbers.

Condition (2) describes the constraint that the amount shipped from the shipment point on all types of transport should not exceed the total stocks of this shipment point, and condition (3) means that the number of shipped products must meet the demand of the consumer. Conditions (4, 5) refer to the restrictions on the capacity of transport that will be used for delivery: the amount of transported products must not exceed the volume that the vehicle can accommodate.

Target function (6) of the problem sets the minimization of transport costs in the delivery of products.

This problem is a multi-index transport problem of linear programming. To form the cost table, a three-index matrix is constructed.

3 Numerical experiment

When compiling a data table for further formation of the solution in this example, a different type of transport for delivery is given for each type of product.

Figure 5 shows a table of input data.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2			L1_1G	L1_2F	L2_1G	L2_2F	L3_1G	L3_2F	a		x1=	3	x4=	7
3	M1	M1_1	x1	-	x2	-	x3	-	200		x2=	4	x5=	4
4		M1_2	-	y1	-	y2	-	y3	160		x3=	5	x6=	6
5	M2	M2_1	x4	-	x5	-	x6	-	240					
6		M2_2	-	y4	-	y5	-	y6	190		y1=	7	y4=	5
7		b	87	40	108	0	0	80			y2=	2	y5=	3
8			92	50	92	50	92	50			y3=	5	y6=	2

Figure 5: The table of input data

Here:

– L1, L2, L3 are consumers;

– M1, M2 are shipment points;

– 1G, 2F are bottles of the 1st and 2nd types and the corresponding type of transport for delivery;

– $M1_1$, $M1_2$ are shipments of bottles of the 1st and 2nd types from the first point;

– $M2_1$, $M2_2$ are shipments of bottles of the 1st and 2nd types from the second point.

The transport variables are recorded in the cells at the intersection. In places, where the delivery of certain products will not be made by this type of transport, lines are drawn.

Columns L and N contain the transport costs per unit of transported products. The problem has the following target function

$$F = 3x_1 + 4x_2 + 5x_3 + 7x_4 + 4x_5 + 6x_6 + 7y_1 + 2y_2 + 5y_3 + 5y_4 + 3y_5 + 2y_6 \rightarrow min$$

As a result of solving the problem, the following volumes of product delivery to the consumer and the total costs are obtained (Figure 6).

To solve the problem, the *Scipy* library for the Python programming language is used. To solve linear programming problems, *Scipy* offers the linprog function (included in the optimize component). The function looks as follows: $scipy.optimize.linprog(c, A\_ub, b\_ub, A_eq, b\_eq)$ and solves the problem by the simplex method.

```
x [1] = 87
x [2] = 92
x [3] = 0
x [4] = 0
x [5] = 16
x [6] = 0
y [1] = 0
y [2] = 0
y [3] = 30
y [4] = 40
y [5] = 0
y [6] = 50
        1143.0
```

Figure 6: Results

4 Conclusion

Using a regression model, a forecast of the demand for the number of bottles has been obtained, these forecast results have been used to formulate and solve the linear programming problem of transport type. In this case not only the need for the number of water bottles has been forecast, but also the company's costs for water delivery.

References

[Sak2018] R. L. Sakhapov, R. V. Nikolaeva, M. H. Gatiyatullin, M. M. Makhmutov. Modeling of regional transport and logistics systems. *Journal of Physics: Conference Series*, 1141:012133, 2018.

[Pic2020] A. Pichugina. Mathematical Modeling Of Transport Logistics Problems. *2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC)*, 1-5, 2020.

[Mas2006] R. Mason, C. Lalwani. Transport integration tools for supply chain management. *International Journal of Logistics Research and Applications*, 9(1): 57-74, 2006.

[Chi2020] T. I. Chinaeva, E. I. Larionova, V. V. Narbut. Modelling and forecasting the Net Income Dynamics. *Proceedings of the 10th International Scientific and Practical Conference named after A. I. Kitov ¡¡Information Technologies and Mathematical Methods in Economics and Management (IT&MM-2020)¿¿*, 77–94, October, 2020.

[Ash2019] N. Ashford. Problems With Long Term Air Transport Forecasting. *Journal of Advanced Panspori-aiion*, 19(2): 101–114, 2019.

[Lin2020] G. Linets, R. Voronkin, S. Govorova, I. Palkanov, C. Grilo. The regression analysis of the data to determine the buffer size when serving a self-similar packets flow. *Proceedings of the International Workshop on Data Mining and Knowledge Engineering*, 60–69, October 2020.

[Mok2020] A. A. Mokhnenko, V. Babenko, O. Naumov, I. Perevozova, O. Fedorchuk. Mathematical-Logistic Model of Integrated Production Structure of Food Production. *Proceedings of the 16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer*, 2: 446–454, October, 2020.