УДК 330.43
JEL classification: C32

Zadko K. V.
ORCID ID: 0000-0003-2097-3582

Zhukovska O. A.
Cand. physics and mathematics, Associate Professor
ORCID ID: 0000-0003-1110-9696

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

FORECASTING MODELING OF THE INPUT LOAD ON THE GOVERNMENT CONTACT CENTRE

ПРОГНОЗНЕ МОДЕЛЮВАННЯ ВХІДНОГО НАВАНТАЖЕННЯ НА КОНТАКТНИЙ ЦЕНТР ДЕРЖАВНОЇ УСТАНОВИ

The article conducted a model study of the load on the call centre of the state budget institution for its more efficient work. The relevance of this study is of particular importance in the context of the growing need for an uninterrupted process of communication between companies and customers with an increase in the cost of labour resources, the cost of their software and hardware. This problem can be solved by ensuring a balance of the necessary and sufficient amount of labour resources. With an increase in the efficiency of use of labour resources, their costs will decrease and the index of customer loyalty and satisfaction will grow. For this purpose, based on statistical data, forecast models are constructed. As a result of the predictive analysis of the input load data, the entire sample was reorganized in this way: the initial data were divided only by the days of the week, and the same data was divided by time for all days of the week. The reason for this separation was the impossibility of constructing an adequate forecast model for the entire sample due to the overlap of fluctuations in both time and days of the week, as well as the presence of non-stationarity caused, as the study showed, by a change in data depending on time and days of the week. However, such non-stationarity in the data did not allow building either an autoregressive model or a Winter's model. Autoregressive – due to the non-removable non-stationarity by the difference method, the Winter's model – due to the small trend. The imposition of oscillations did not allow them to be accurately simulate. A sample autocorrelation and a private sample autocorrelation function were constructed and the structure of time series was revealed for each of the obtained samples. Seasonal ARIMA models based on the Box-Jenkins approach were built, the adequacy of the developed models was studied and forecasts of the possible input load were made. The results of the forecasting carried out in this work were used to develop recommendations for improving the efficiency of use of labour resources at the enterprise under study.

Keywords: call centre, planning, forecasting, AR-models, seasonality, manpower.

У статті проведено модельне дослідження навантаження на колл-центр державної бюджетної установи для його більш ефективної роботи. Актуальність даного дослідження набуває особливого значення в умовах зростання необхідності забезпечення безперебійного процесу комунікації компаній з клієнтами при зростанні вартості трудових ресурсів, вартості їх програмного і апаратного забезпечення. Дане завдання може бути вирішена за рахунок забезпечення балансу необхідного і достатнього кількості трудових ресурсів. За умови збільшення ефективності використання трудових ресурсів відбувається зменшення витрат і, відповідно, спостерігається зростання індексу клієнтської лояльності і задоволеності. З цією метою на основі статистичних даних побудовані прогнозні моделі. В результаті предиктивного аналізу даних вхідного *навантаження вся вибірка була реорганізована таким чином: вихідні дані були розділені тільки за днями тижня, і ці ж дані були розділені за часом за всі дні тижня. Причиною такого поділу послужила неможливість побудови адекватної прогнозної моделі для всієї вибірки через накладення коливань і за часом і за днями тижня, а також наявність нестаціонарності, викликаної, як показало дослідження зміною даних в залежності від часу і днів тижня. Однак така нестаціонарность в даних не дозволяла побудувати ні авторегресійну модель, ні модель Вінтарса. Авторегресійну – із-за неможливості позбавитись нестаціонарності методом різниць, модель Вінтарса – через слабкий тренд. Також накладення коливань не дозволяло точно їх змоделювати. За кожною з отриманих вибірок була побудована вибіркова автокореляційна і часткова вибіркова автокореляційна функції, виявлена структура часових рядів. На основі підходу Бокса-Дженкінса були побудовані сезонні ARIMA моделі, проведено дослідження адекватності розроблених моделей, побудовані прогнози можливого вхідного навантаження. Результати прогнозування, проведеного в даній роботі, було використано для розробки рекомендацій щодо підвищення ефективності використання трудових ресурсів на досліджуваному підприємстві.*

Ключові слова: колл-центр, планування, прогнозування, AR-моделі, сезонність, трудові ресурси**.**

Introduction. For the effective organization of the call center, the manager needs to correctly plan the possible load in order to ensure customer satisfaction with minimal maintenance costs.

The identification of patterns in the change in the load coming to the service, as well as the selection and study of the mathematical model that most accurately describes the customer service process are tasks whose solution will improve the quality of the services provided. This will allow you to effectively organize the work of the center: reduce the cost of maintaining a call-center, increase customer loyalty, and optimize the staff. It should be noted that since the organization of Call centers in each company has its own specifics, it is impossible to build a universal mathematical model for optimizing their work. But, based on the analysis of the statistical data of an individual company, it is possible to predict the load on the call center and, accordingly, organize work more efficiently.

In 1909, Danish scientist A.K. Erlang published the work "Probability Theory and Telephone Conversations" on the study of the throughput of a fully accessible bundle of lines serving the simplest call flow with losses and with anticipation. This work served as an impetus for scientists T. Engset, T. Fry, A.N. Kolmogorov, A. Ya. Khinchin, K. Palm, and others, whose work was associated with the confirmation, development, or refutation of its results [1]. Nowadays, the problem of organizing call-center centers in the pool is highlighted in the robots of Samolyubovo B.A. [2], Goldstein B.S. that of Freikman V.A. [3]. The robots of forecasting the subscriber's call to the call center were examined by the robots Soluyanova A.V., Yurkina Yu.V [4], as well as by the robots Sivan Aldor-Noiman1, Paul D. Feigin and Avishai Mandelbaum [5] and Yu.A. Kryukov, D.V. Chernyagin [6]

Typically, the calculation of the required number of operators is carried out according to the Erlang formula, for the application of which it is necessary to indicate the following values: average time of the operator's conversation with the client; average call post-processing time spent by the operator on processing after the call is completed; The maximum target delay for answering a call is the time that the client will wait for a response on the line, as well as the number of calls per hour.

To determine the number of calls, it is necessary to analyse the degree of load at different times and on different days and build a predictive model. Note that any research of call centres must necessarily begin with a thorough analysis of the data. It follows that parameter prediction is a prerequisite for an agreed level of service and efficient operation. To build such a predictive model, it is necessary to carefully analyse the data, determine its structure and verify the adequacy of the model.

Setting objectives. The purpose of this work is to analyse and build a forecast load on the call centre.

Methodology: integrated model of autoregression, moving average model, model and methodology of time series analysis.

Results of the research. Based on the statistics of the load on the call centre of the state institution N for three months of work (Fig. 1), we will build a forecast model. Therefore, analyse the data and determine the structure of the time series.

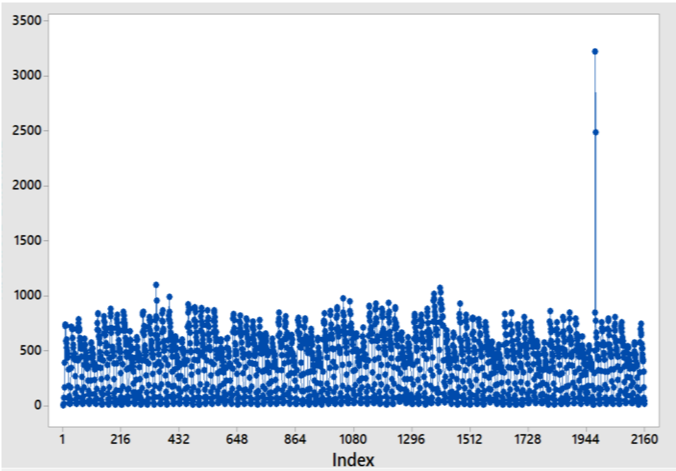

Figure 1 – Call centre load schedule

Visual analysis of the input load graph (Fig. 1) does not allow to determine the presence or absence of trend and expressed fluctuations. However, you can see significant ejection and a few minor ones. After finding out the cause of the peak emissions, they were removed from the investigated time series.

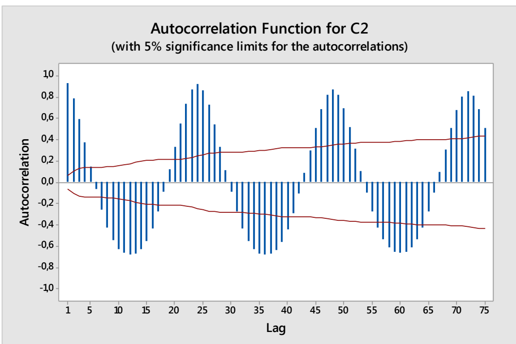

Figure 2 – Autocorrelation function of time series

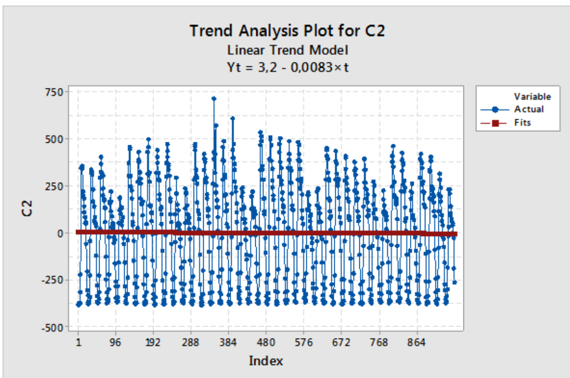

Figure 3 –Trend analysis of load distribution output data

Visual analysis of the graph of the autocorrelation function (Fig. 2) of the time series of load distribution allows to conclude the non-stationarity of the process, as well as the presence of a seasonal component with a period of 24, which corresponds to the number of hours per day (Fig. 2). Note that the stationary time series correlogram declines rapidly with increasing order after the first few values. If the schedule falls short enough, then the time series can be considered non-stationary. The time series under study are uniquely non-stationary with a declining correlogram. This is the situation observed in Fig. 2. At the same time, the trend analysis showed a rather weak trend (Fig. 3), while a clear trend is needed to build a forecast model using the Vintars method, and a stationary time series is needed to construct autoregressive models. Research on stationary time series first and second differences did not provide the desired result. Therefore, we reorganize the original time series as follows: we separate the data of each day of the week and the data for all days of the week, but for a specified time.

In Fig. 4, we can see a time series of inbound call centre values for only one day - Monday for 13 weeks (sample length 312).

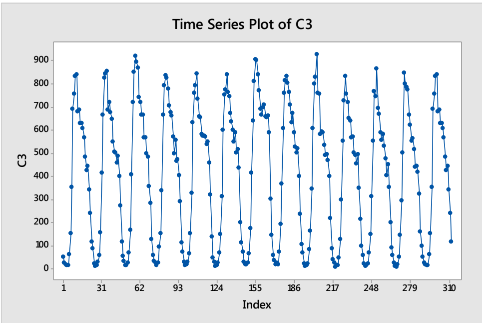

Figure 4 – Incoming call centre data for 13 Mondays

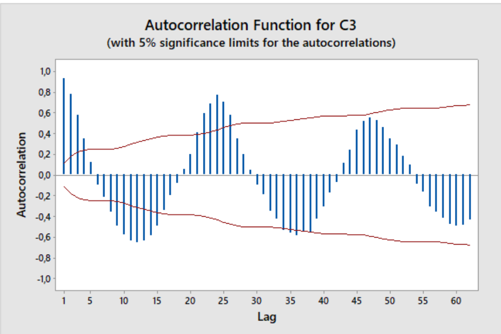

Figure 5– Autocorrelation function for call centre incoming load data for 13 Mondays

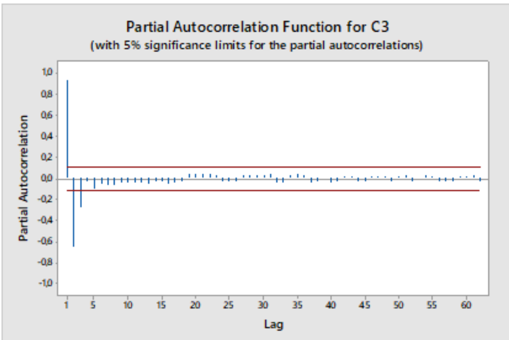

Figure 6 – Partial autocorrelation function for call centre incoming load data for 13 Mondays

The autocorrelation function (Fig. 5) is slowly decreasing harmonically, indicating that the time series is unstable and the seasonal component is present. The coefficients of the partial autocorrelation function (Fig. 6) become much smaller after the second lag and statistically insignificant after the third lag. Using the elements of the Box-Jenkins method methodology, we assume that this process will be successfully described by the ARIMA (2,1,0) $(0,2,2)_{24}$ model.

The construction of the ARIMA(2,1,0)(0,2,2)$_{24}$ model was implemented in the Minitab software environment. In Fig. 7 we can see a listing of the program for the construction and analysis of the model ARIMA(2,1,0)(0,2,2)$_{24}$.

```
Session

Final Estimates of Parameters

Type         Coef  SE Coef      T      P
AR   1    -0,5388   0,0644  -8,37  0,000
AR   2    -0,1260   0,0630  -2,00  0,007
SMA  24    1,5835   0,0559  28,35  0,000
SMA  48   -0,6683   0,0673  -9,93  0,000


Differencing: 1 regular, 2 seasonal of order 24
Number of observations:  Original series 312, after differencing 263
Residuals:    SS =  373941 (backforecasts excluded)
              MS =  1444  DF = 259


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12     24     36     48
Chi-Square    8,8   23,3   34,2   46,5
DF              8     20     32     44
P-Value     0,402  0,321  0,303  0,230


Forecasts from period 312

                     95% Limits
Period  Forecast    Lower    Upper  Actual
   313    89,956   89,956   97,074
   314    59,751   54,638   64,864
   315    47,418   43,097   51,739
   316    48,712   44,087   53,337
   317    51,466   45,854   57,078
   318    94,310   84,509  104,111
   319   182,390  170,085  194,695
   320   346,868  328,663  365,073
   321   601,351  577,192  625,510
   322   812,450  779,846  845,054
```

Figure 7 – Listing the result of building and analysing the ARIMA (2,1,0)(0,2,2)$_{24}$

The ARIMA(2,1,0)(0,2,2)$_{24}$ model 24 describes the data structure well. Ljung-Box statistics for groups of intervals 12, 24, 36 and 48 are not significant, which shows great importance. The autocorrelation of residues does not have any visible structure (Fig. 8). The correspondence of the constructed model is confirmed also by the normality of the distribution of residues (Fig. 9) and the finding of points in a certain corridor on the plot of the dependence of the residuals on the predicted values (Fig. 10).

Thus based on the data considered, we have not grounds for claiming that the model is incorrect. Moreover, the model can be recommended to build a short-term forecast. Such a forecast is provided in the program listing (Fig. 7) and on the graph (Fig. 11). Data analysis for other days of the week showed the same pattern of regression models ARIMA(2,1,0)(0,2,2)$_{24}$.

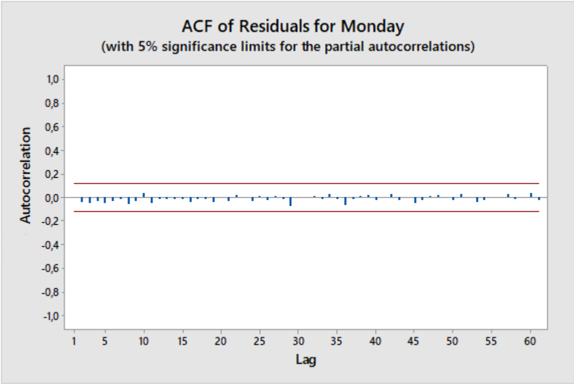

Figure 8 – Autocorrelation function of residuals of ARIMA$(2,1,0)(0,2,2)_{24}$

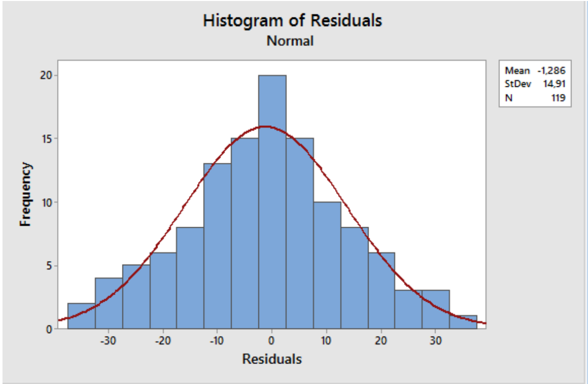

Figure 9 – Histogram of ARIMA $(2,1,0)(0,2,2)_{24}$ model residuals

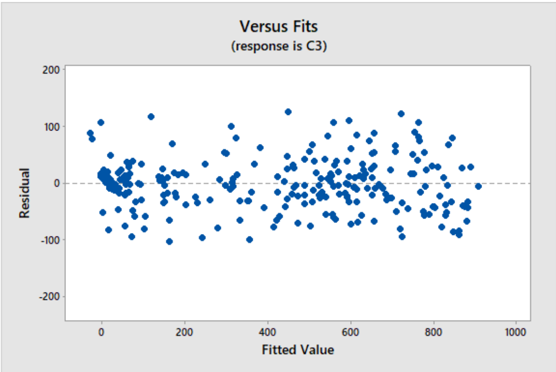

Figure 10 – The plot of the residuals of the model ARIMA $(2,1,0)(0,2,2)_{24}$ on the predicted values

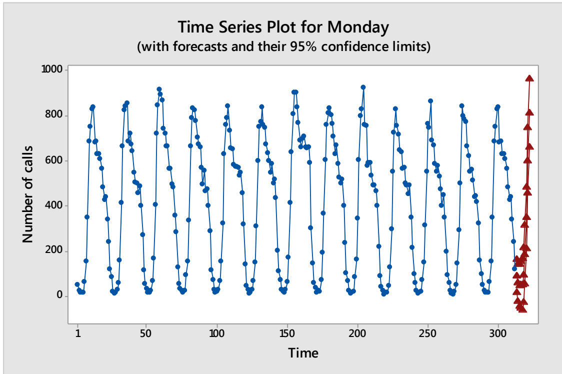

Figure 11 – Graph of the results of the simulation forecast by ARIMA$(2,1,0)(0,2,2)_{24}$

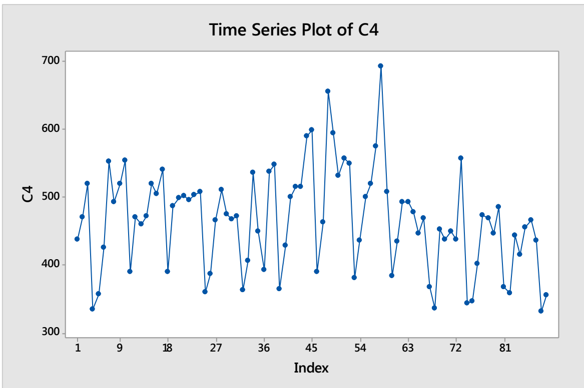

Figure 12 – Call centre data load at 4pm

The next step is to look at the data for all days of the week for 89 days at the same time – 4 pm (Fig. 12). From the analysis of graphs of autocorrelation and partial autocorrelation function for data of incoming load on the call centre at 16 o'clock (Fig. 13) we conclude about the frequency of the time series with period 7, which corresponds to the length of the week. The data are characterized by gradually harmonic attenuation of the autocorrelation coefficients and show a slight trend.

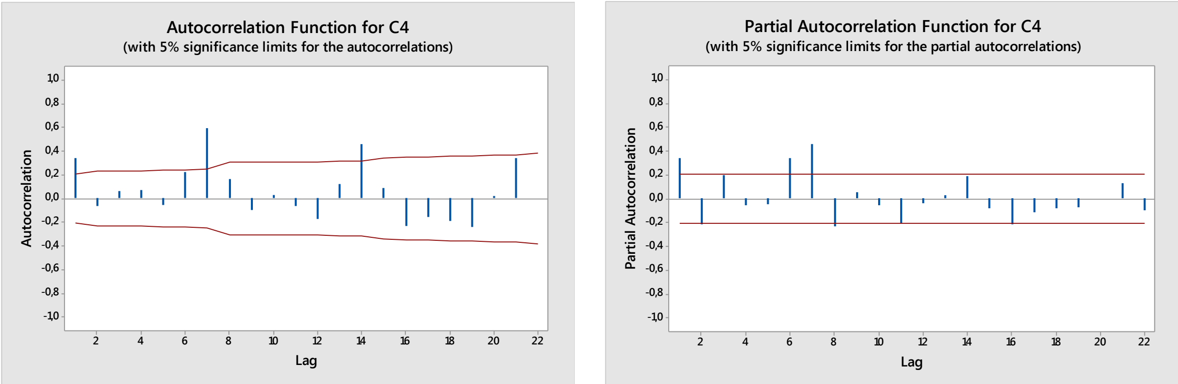

Figure 13 –Autocorrelation and partial autocorrelation functions for call centre incoming load data at 4 pm

```
Session

Final Estimates of Parameters

Type          Coef   SE Coef       T       P
AR    1    -0,2611    0,1316   -1,98   0,002
AR    2    -0,3380    0,1413   -2,39   0,002
SMA   7     1,0979    0,1994    5,51   0,000
SMA   14   -0,4027    0,2512   -1,60   0,001


Differencing: 1 regular, 2 seasonal of order 14
Number of observations:  Original series 89, after differencing 60
Residuals:    SS =  341311  (backforecasts excluded)
              MS =   6095  DF = 56


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag              7      14      21      28
Chi-Square     9,6    20,6    32,2    44,6
DF               8      20      32      44
P-Value      0,395   0,421   0,507   0,480


Forecasts from period 89

                          95% Limits
Period  Forecast     Lower     Upper  Actual
    90    342,24    316,24    368,24
    91    407,75    377,75    437,75
    92    430,26    398,26    462,26
```

Figure 14 – Listing ARIMA $(2,1,0)(0,2,2)_{24}$ model building and analysis for call centre business at 4pm

Thus, the series is non-stationary with a seasonal component and, based on the graphs of autocorrelation and partial autocorrelation functions, we apply to it the ARIMA $(2,1,0)(0,2,2)_{24}$ model.

The results of the calculation are shown in Fig.14. The histogram of the residuals (Fig. 15) shows an approximately normal distribution, and the graph (Fig. 16) of the correlation coefficients of the residuals from the estimated values indicates that there is no influence of time and the residuals have a horizontal structure. Autocorrelations of model residuals (Fig. 17) and Ljung-Box statistics (Fig. 14) show random behaviour of errors. Therefore, we can conclude that the predictive value is valid. We also note that the structure of the forecast (Fig. 18) is similar to the temporal structure of real observations.

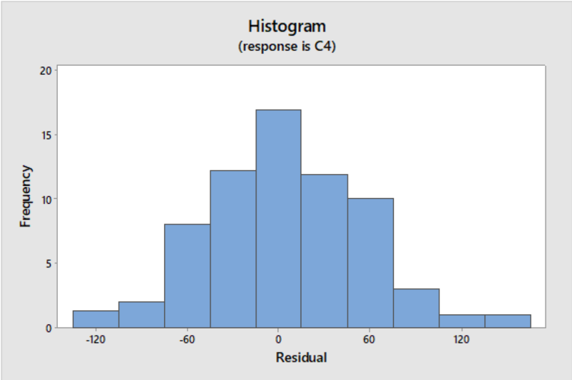

Figure 15 – Histogram of ARIMA (2,1,0)(0,2,2)$_{24}$ residuals for call centre data at 16.00

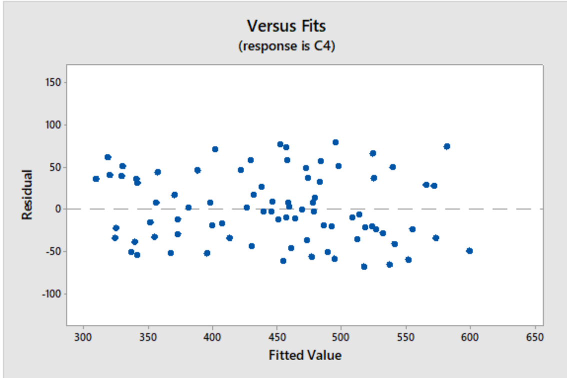

Figure 16 – ARIMA(2,1,0)(0,2,2)$_{24}$ residuals plot from predicted values for call centre work at 16.00

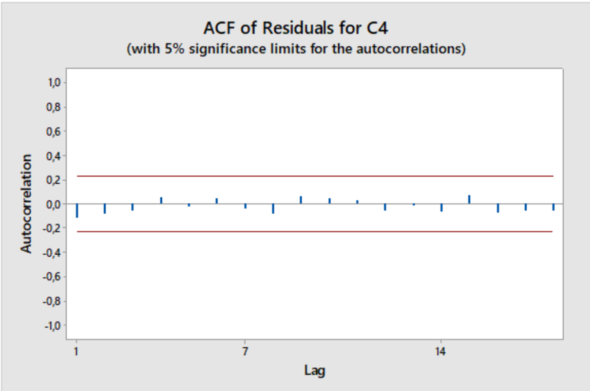

Figure 17 – Autocorrelation function of ARIMA (2,1,0)(0,2,2)$_{24}$ model residuals for call centre data at 16.00

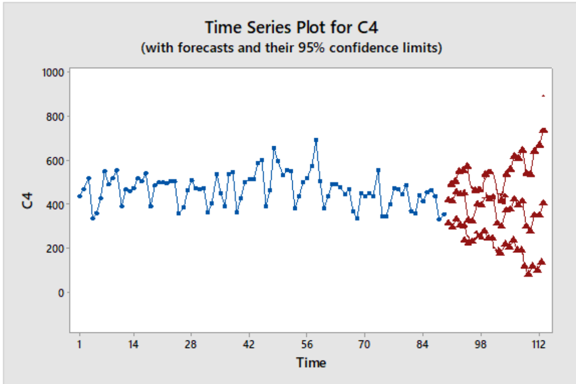

Figure 18 – The results of the simulation forecast by ARIMA (2,1,0)(0,2,2)$_{24}$ for call centre work at 16.00

Based on the data obtained as a result of forecasting the input load for the developed data models, it has been suggested that a number of changes be made to the cost of resources by Company N. Currently, the company uses 30 units of labour and the corresponding amount of equipment per hour per day. At the enterprise there are 3 changes lasting 8 hours. After conducting the study, it was recommended to reduce the number of employees by 8-hour shifts to 20 and to introduce a new 5-hour workforce with 15 employees during peak load times, to train part-time staff to provide rapid response to unplanned bursts.

Due to the test implementation of the recommended changes, it was confirmed that labour costs and workplace costs were reduced, customer satisfaction with the call centre was increased by reducing waiting times in the queue and reducing the number of rejected calls.

Conclusions. In the course of this work, the statistics of the call centre work were analysed. The data were distributed according to certain criteria, which is conditioned by the analysis of graphs of sample and private sample autocorrelation. During the study, ARIMA forecast models were built for each day of the week, the model presented was analysed, the residuals were evaluated and its adequacy confirmed for the initial input.

The results of the forecasting work used in this work were used to develop recommendations for improving the efficiency of labour use in the enterprise under study.

References:

1. Teoryia teletrafyka [Elektronnyi resurs]. – Update date: 11.07.2019.– URL: https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B8%D1%8F_%D1%82%D0%B5%D0%BB%D0%B5%D1%82%D1%80%D0%B0%D1%84%D0%B8%D0%BA%D0%B0.
2. Samoliubova A. B. Call-Center na 100%. Praktycheskoe rukovodstvo po orhanyzatsyy tsentra obsluzhyvanyia vyzovov / A. B. Samoliubova. – Moskva: Alisha Biznes Buks, 2004. – 309 s.
3. Holdshtein B. S. Call-tsentry y kompiuternaia telefonyia / B. S. Holdshtein, V. A. Freikman., 2002.
4. Soluianov A. V. Prohnozyrovanye nahruzky, postupaiushchei v Call-tsentr. Vestnyk sviazy. / A. V. Soluianov, Yu. V. Yurkin. // Vestnyk sviazy. – 2012. – №5.
5. S. Aldor-Noiman. WORKLOAD FORECASTING FOR A CALL CENTER: METHODOLOGY AND A CASE STUDY / S. Aldor-Noiman, Paul D. Feigin, A. Mandelbaum. // The Annals of Applied Statistics. – 2009. – №4. – C. 1403–1447.
6. Yu.A. Kriukov. Model prohnozyrovanyia znachnyi trafyka / Yu.A. Kriukov, D.V. Cherniahyn., 2011. –S. 41-49.