ESTIMATING AUTOREGRESSIVE MODEL USING GOAL PROGRAMMING

Efuwape B.T.[1], Abdullah K-K. A[1], Hammed F. A.[1], Obadina O.G.[1,2] and Awosanya A.[1]

[1] Olabisi Onabanjo University, Department of Mathematical Sciences, Faculty of Science, Ago-Iwoye, Nigeria
[2] University of Ibadan, Department of Mathematical Sciences, Faculty of Science, Ibadan, Nigeria

Corresponding author: Obadina Olugbenga, olugbenga.obadina@yahoo.com

ABSTRACT: *An autoregressive (AR) model predicts future behavior based on past behavior. It is used for forecasting when there is some correlation between values in a time series the values that precede and succeed them. Goal programming (GP) model is a simple extension and modification of the linear programming technique that provides a simultaneous solution of a system of complex objectives rather than a single one. The study proposes goal programming as a method to estimate the parameters of the autoregressive model. The findings indicates that in both data sets, the Mean Absolute Percentage Error (MAPE) values using the GP prediction were lower than those obtained using the AR(1) model. Based on these results, it can be deduced that prediction equations obtained using the GP approach were more accurate than those obtained using the AR(1) model. This is because, using the GP approach, the problem can be restated as to minimize the sum of absolute errors rather the sum of squares of the error in the case of the AR(1) model.*

KEYWORDS: *Method of least squares, outliers, goal programming*

1. INTRODUCTION

In the recent years, inverse optimization problems attracted many operations research specialists and different kind of inverse problems have been developed by researchers (Ahuja & Orlin, 2001). On the other hand, time-series modeling and forecasting continues to be an important area in both academic research and practical application. For example, Inniss (2006) developed a seasonal clustering technique for determining clusters of time series data. His model was also applied to weather and aviation data to determine probabilistic distributions of arrival capacity scenarios, which can be used for seasonal forecasting and planning (Inniss, 2006). As the practical application of time-series modeling Bermu´dez, Segura, and Vercher (2006) suggested a non- linear multi-objective problem for forecasting of time series based on soft computing.

In time series analysis, generally, historical observations on the item to be predicted are collected and analyzed to specify a model to capture the underlying data generating process and to use the model for predicting the future. Depending on the theory or assumption about the relationship in the data there are two different approaches widely used in time series forecasting (Box, Jenkins, & Reinsel, 1994). The traditional approaches such as the time-series regression, exponential smoothing, and autoregressive integrated moving average (ARIMA) are based on linear models. That is, they assume the future value of a time series is linearly related to the past observations (Box et al., 1994). The ARIMA model is representative of linear models and has achieved great popularity since the publication of Box, Jenkins, and Reinsel's classic book (Box *et al*., 1994). Mohammadi, Eslami, and Kahawita (2006) introduced an autoregressive moving average (ARMA) model for river flow forecasting using a goal programming methodology for parameter estimation. Two main disadvantages in goal programming are the mathematical expression of goals and constraints and simultaneously optimizing all goals (Arikan & Gungor, 2001). Traditionally, linear statistical forecasting methods have been widely used in many real-world situations including forecasting (Mohammadi et al., 2006) since linear models are easy to develop and implement. They are also simple to understand and interprete, (Bazaraa, Jarvis, & Sherali, 2005). Taylor (2007) forecasted the daily supermarket sales using an exponential weighted regression time-series model.

On the other hand, in the field of operational research many applications have been reported using linear goal programming problem that first was introduced by Zhang and Liu (2016) and further improved by Huang and Liu (2019). The first application of linear goal programming was the shortest paths problems as developed by Burton and Toint (2012a, 2012b). Other applications include the shortest arborescence problem (Hu & Liu, 1998), maximum capacity problems (Yang & Zhang, 2018) and the maximum flow and minimum cut problems (Burkard, Klinz, & Zhang, 2001; Yang, Zhang, & Ma, 2017). Many more applications of linear goal programming have been reported by Huang and Liu (2019).

Despite the large uses of linear programming in forecasting models there are little or no application reported in the literature using goal programming to estimate parameters of short term models.

The aim of the study is to propose a parameter estimation of autoregressive model using goal programming. The specific objectives of this study are:

1. To estimate autoregressive model of order one.
2. To make comparison on the performance of parameters AR(1) and parameter of AR(1) using goal programming using MAPE.

The data for the study is obtained from CDK Integrated Industries, producers of porcelain tiles of different sizes and luxury sanitary wares. The data collected are total production output and sales, between January, 2017 to December, 2018.

2. LITERATURE REVIEW

Gupta and Kumar (2015) carried out a study to forecast the packaged food product sales using mathematical programming. The scope of the subject is wide and the techniques chosen reflect particular interests and concerns. In this study Linear Programming is used to estimate the parameters of time series forecast by minimising one error index (MAD, MAPE, MPE) and they are compared with the time series forecasting method. Linear Programming is used to estimate the parameters of times series forecast with optimisation objectives to minimise forecasting error and it is compared them with the traditional time series forecasting models. The linear programming approach improves the accuracy of forecast and outperforms all the other techniques. The use of a mathematical programming provides a formal, logical way of thinking about this decision process. This should increase the understanding of this problem area and increase the quality of decisions.

Vyskocil (2018) combine methods of time-series prediction together with stochastic optimization concepts. They create a flexible inventory control pipeline, which is capable of generating goods ordering decisions that consider demand uncertainty, goods durability, shortage costs, warehousing costs, and both fixed and per-unit ordering costs. The pipeline is enriched with the progressive hedging decomposition algorithm, which helps to reduce computation times and improves the capability of our models to reduce risks of unexpected demand outcomes.

Mahdiraji et al. (2019) discusses the primary criterion that influences electricity consumption and uses the singular spectrum analysis based on these factors to predict use. Besides, a fuzzy regression model is represented to optimise function. Results of optimisation show a considerable reduction in comparison with SSA forecasting method, indicating the efficiency of the offered method. Eventually results considerably assume that attention to a way of construction and improvements of the culture of use is a priority of the persons making decisions to reduce radical electric consumption in Iran and to become more optimistic concerning management of an electrical network.

Mohammadi *et al*. (2006) used goal to minimize the error for a specific season of the year as well as for the complete series. Goal programming (GP) was used to estimate the ARMA model parameters. Shaloo Bridge station on the Karun River with 68 years of observed stream flow data was selected to evaluate the performance of the proposed method. The results when compared with the usual method of maximum likelihood estimation were favourable with respect to the new proposed algorithm.

Mohammadi, et al. (2006) proposed a goal programming method of estimating the parameters of ARMA model to forecast river flow River flow. The results when compared with the usual method of maximum likelihood estimation were favourable with respect to the new proposed algorithm.

Maizah et al. (2005) proposed a goal programming as a method to estimate regression model parameters when outlier must be included in the analysis

3. METHODOLOGY

Estimating AR Model Parameters Using Goal Programming

Recall that for the OLS estimation,

$$y_t = X_t'\beta + \varepsilon_t, t = 1, \dots, T \qquad (1)$$

With the assumptions

1. $E(\varepsilon_t|X_t) = 0$
2. $E(X_tX_t')$ is non singular
3. $(y_t, X_t')'$ is stationary and weakly dependent

$\hat{\beta} \xrightarrow{p} \beta$ as $T \to \infty$

For the AR(1) model, that is the first order process, which means that the current values is based on the immediate preceding value.

$$y_t = \phi y_{t-1} + \varepsilon_t,\ t = 1, \dots, T \qquad (2)$$

$\varepsilon_t \sim i.i.d\ N(0, \sigma^2)$ y_0 fixed / known

Equivalently, between (1) and (2), we have;

$$X_t = y_{t-1}, \beta = \phi$$

Hence,

$$\hat{\phi}_{OLS} = \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}y_t}{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2}$$

We have to show that our $AR(1)$ model satisfy consistency of the OLS estimation.

Assumption: $|\phi| < 1 \Rightarrow y_t$ is stationary and weakly dependent.

1. $E(y_t|y_{t-1}) = 0$

$$y_{t-1} = \phi y_{t-2} + \varepsilon_t$$
$$\vdots$$
$$= \phi^{t-1} y_0 + \sum_{t=1}^{T} \phi^{t-1-i} \varepsilon_t = f(\varepsilon_t, \dots, \varepsilon_{t-1}, y_0)$$

This variable is by construction independent of ε_t as we have assumed that ε_t are independent over time. Hence, $E(\varepsilon_t|y_{t-1}) = E(\varepsilon_t) = 0$

2. $E(y_{t-1}^2)$ is non singular.

$$0 \leq E(y_{t-1}^2) < \infty$$

Recall that,

$$E(y_{t-1}^2) = \frac{\sigma^2}{1-\phi^2} \text{ if } |\phi| < 1.$$

3. Since $|\phi| < 1$, (y_t) is stationary and weakly dependent.

If $|\phi| < 1$, then $\hat{\phi}_{OLS} \xrightarrow{p} \phi$

In estimating $AR(1)$ parameter using goal programming.

From $y_t = \phi y_{t-1} + \varepsilon_t$

$$\varepsilon_t = y_t - \phi y_{t-1}$$

The general goal programming model can be expressed as follows.

Minimize

$$Z = \sum_{i=1}^{m} (d_i^- + d_i^+)$$

Subject to the linear constraints.

Goal constraints:

$$\sum_{j=1}^{n} a_{ij} x_j + d_i^- - d_i^+ , i = 1,2, \dots, m$$

System constraints:

$$\sum_{j=1}^{n} a_{ij} x_j \begin{bmatrix} \leq \\ = \\ \geq \end{bmatrix} b_i, i = m+1, \dots, m+p$$

With $x_j, d_i^-, d_i^+ \geq 0$ for $i = 1,2, \dots, m$ and $j = 1,2, \dots, n$

Equivalently, $d_i^- - d_i^+ = \varepsilon_t$ and $y_t = b$

Rewriting the $AR(1)$ in terms of goal programming, we have;

Minimize

$$Z = \sum_{i=1}^{m} (d_i^- + d_i^+)$$

Subject to the linear constraints.

Goal constraints:

$$\phi y_{t-1} + \varepsilon_t$$

System constraints:

$$\phi y_{t-1} \begin{bmatrix} \leq \\ = \\ \geq \end{bmatrix} y_t,$$

Note: There is only one goal

4. RESULTS

The data adopted for this study is presented in Table 1. The variables used in this study are discussed in this subsection to establish their meaning.
Production of Porcelain tiles: This is a US term, and defined in American Society for Testing and Materials standard C242 as a ceramic mosaic tile or paver that is generally made by dust-pressing and of a composition yielding a tile that is dense, fine-grained, and smooth, with sharply-formed face, usually impervious. The colours of such tiles are generally clear and bright. In this study, the total production output and sales are used.

Table 1: Total Production Output and total sales of Porcelain tiles

MONTHS	TOTAL PRODUCTION OUTPUT(SQM)	TOTAL SALE (A)
2017:01:00	164303	36450
2017:02:00	148302	113236
2017:03:00	155196	130436
2017:04:00	172691	77225
2017:05:00	142988	137276.46
2017:06:00	78433.01	129460.55
2017:07:00	45488	74711.95
2017:08:00	206030.87	109905
2017:09:00	209030.87	178450.88
2017:10:00	203930.01	151552.88
2017:11:00	209030.87	201067
2017:12:00	180509.22	177063.5
2018:01:00	290533.61	98406.41
2018:02:00	299527.07	240094.9
2018:03:00	302368.47	292836.8
2018:04:00	307037.08	271085.46
2018:05:00	209648.55	202475.51
2018:06:00	306587.61	180073.95
2018:07:00	287108.63	302841.23
2018:08:00	298781.99	200311.3
2018:09:00	306909.15	250393.89
2018:10:00	309919.15	289063.89

Source: CDK Ceramic Industries, 2019

4.1 Descriptive Analysis

Descriptive Statistics

The descriptive statistics of the monthly time series variables used in the study are presented in Table 4.2 below. Descriptive statistics describes the basic features of the data and form the basis of virtually every quantitative analysis of data in a study. This is of essence as it profiles the characteristics of the variables engaged in the country and suggests possible pliable models for the study. It is observed that all the mean and median values are positive and similar, which suggests that the distribution of the individual variables are normal (i.e., bell-shaped). Skewness is a measure of the symmetry and asymmetry nature of random individual variable about its mean, which can be positive or negative, or even undefined depending on the biasness of the tails (Gujarati, 1999). The skewness of production output and sales, which are -0.6 and 0.04 respectively. In addition, the Kurtosis is a parameter that describes how tall and sharp the central peak is, relative to a standard bell curve. The kurtosis of production output and sales are below 3.0, which means that they have normal frequency (platykurtic). This means that their central peak is lower and broader.

The descriptive statistics of the annual time series variables used in the study are presented Table 2.

Table 2: Descriptive Statistics of Variables

Variable	Total Production Output (SQM)	Total Sale (A)
N	24	24
Mean	226884	185522
SE Mean	16290	16699
StDev	79805	81809
Variance	6.4E+09	6.7E+09
Minimum	45488	36450
Maximum	310407	305262
Range	264919	268812
Skewness	-0.6	0.04
Kurtosis	-0.54	-1.12

Source: Author's Computations

Time Plot

The descriptive analysis was used to summarized the characteristics of the variables consider in this research work with a view of showing the important features of each of the variables through the use of time plot.

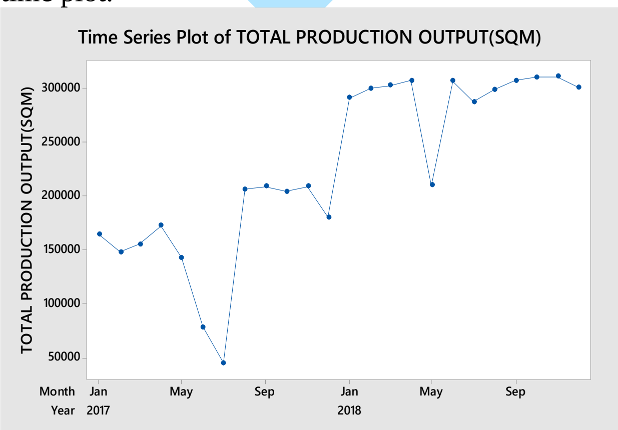


Fig. 1 Time plot of production output, 2017-2018

The time plot for the production output shows a short-term movement of the value in the series in different direction over the period considered. This movement is characterized by a sinusoidal increase in the values of the production output over the period of time. This movement is referred to as secular variation or secular movement. By fitting a straight line freely by hand on the plotted points on the time plot for production output stretching over the period, this plotted point forms a line and this line is the trend of the time plot for production output.

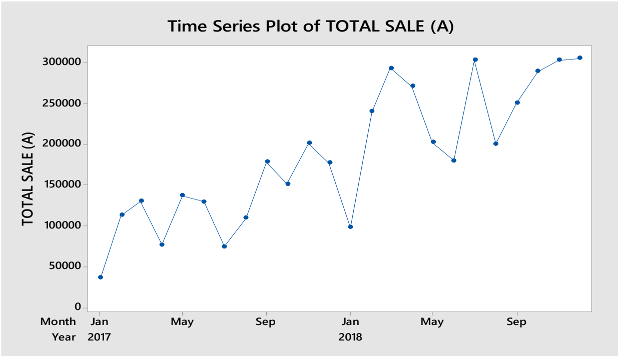
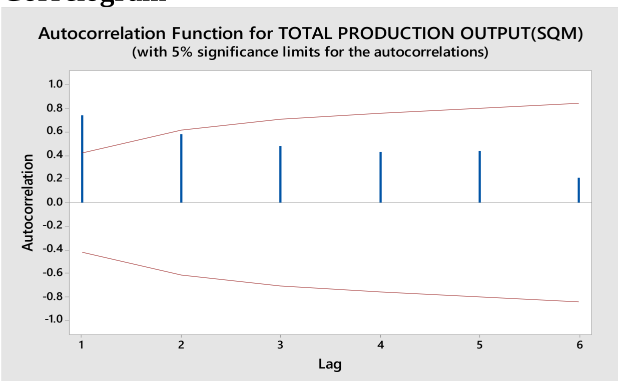
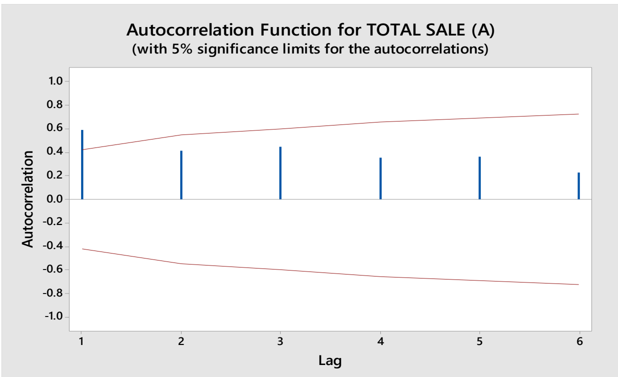


Fig. 2: Time plot of sales, 2017-2018

The time plot for sales shows a short-term movement in the series in different direction over the period considered. By fitting a straight line freely by hand on the plotted points on the time plot for sales stretching over the period, this plotted point forms a line and this line is the trend of the time plot for sales.

Correlogram





Autoregressive Models

AR(1) Model for Total Production Output

Type	**Coef**	**SE Coef**	**T-Value**	**P-Value**
AR 1	0.668	0.163	4.11	0.000
Constant	1.7685	0.0327	54.13	0.000
Mean	5.3224	0.0983		
MAPE	0.0255007			

The AR(1) model obtained in given as:
$y_t = 0.668y_{t-1}$

AR(1) for Sales

Type	**Coef**	**SE Coef**	**T-Value**	**P-Value**
AR 1	0.769	0.151	5.08	0.000
Constant	1.1954	0.0369	32.39	0.000
Mean	5.175	0.160		
MAPE	0.0326417			

The AR(1) model obtained in given as:
$y_t = 0.769y_{t-1}$

Estimating AR(1) with Goal Programming

There is only one goal. The goal is to predict the total product output.

Total Production Output

Decision variable analysis	Value	MAPE
y_{t-1}	.58	0.0234109

The GP result is $y_{t-1} = 0.58$. From the result, we can write the predicted equation as

$$y_t = 0.58y_{t-1}$$

Sales

Decision variable analysis	Value	MAPE
y_{t-1}	.76	0.03123587

The GP result is $y_{t-1} = 0.76$. From the result, we can write the predicted equation as

$$y_t = 0.76y_{t-1}$$

When choosing between competing models or when evaluating an existing model, measures that summarize the overall accuracy provided by the model (s) should be used. Generally, the closer the predicted is to actual, the more accurate the model is. Thus, the quality of a model can be evaluated by examining the series of predicted error. Since MAPE is measured as a percentage, it is particularly useful for comparing the performance of a model in different units. A large value of MAPE means that the value of the error is large.

CONCLUSION

The study developed as parameter estimation method using goal programming. The data used for the study were obtained CDK Ceramic Industries, makers of high quality ceramic tiles and sanitary wares. Obtained was for total production output and sales for periods, January, 2017 to December, 2018.

The data obtained were modelled using Autoregressive model of order one and the parameters obtained. Also, the goal programming was used to obtain the parameters of the AR(1) model. Both methods were compared, using the MAPE (Mean Absolute Percent Error).

In both data sets, the MAPE values using the GP prediction were lower than those obtained using the AR(1) model. Based on these results, it can be deduced that prediction equations obtained using the GP approach were more accurate than those obtained using the AR(1) model. This is because, using the GP approach, the problem can be restated as to minimize the sum of absolute errors rather the sum of squares of the error in the case of the AR(1) model.

REFERENCES

[1] Atmaca, E.; Pehlivan, C.; Aydog˘ du, B.; Yakici, M. (2012). Nurse Scheduling Problem and Application. Erciyes Univ. J. Inst. Sci. Technol., 28, 351–358.

[2] Azaiez, M.N.; Al Sharif, S.S. (2005). A 0-1 Goal Programming Model for Nurse Scheduling. Comput. Oper. Res., 32, 491–507.

[3] Bag˘, N.; Özdemir, N.M.; Eren, T. (2012). 0-1 Goal Programming and ANP Method and Nurse Scheduling Problem Solution. Int. J. Eng. Res. Dev., 4, 2–6.

[4] Bard, J.F.; Binici, C. (2003). Staff Scheduling at the United States Postal Service. Comput. Oper. Res., 30, 745–771.

[5] Bedir, N. (2018). Shift Scheduling Problems Solution with Combined Ahppromethee and Goal Programming Methods: An Aplication in Hydroelectric Power Plant. Master 's Thesis, Kirikkale University, Kirikkale, Turkey.

[6] Bektur, G.; Hasgül, S. (2013). Workforce Scheduling Problem According to Seniority Levels: An Application for A Service System. Afyon Kocatepe Univ. J. Fac. Econ. Adm. Sci., 15, 385–402.

[7] Castillo, I.; Joro, T.; Li, Y.Y. (2009). Workforce Scheduling with Multiple Objectives. Eur. J. Oper. Res., 196, 162–170.

[8] Goal Programming estimates of Linear Regression Parameters, MATEMATIKA, 21(2), 101-112.

[9] Güngör, I˙. (2002). A Model Proposal for Nursing Assignment and Scheduling. Süleyman Demirel Univ. J. Fac. Econ. Adm. Sci., 7, 77–94.

[10] Gür, S¸.; Eren, T. (2018). Scheduling and Planning in Service Systems with Goal Programming: Literature Review. Mathematics, 6, 265.

[11] Heimerl, C.; Kolisch, R. (2010). Scheduling and Staffing Multiple Projects with a Multi-Skilled Workforce. OR Spectr., 32, 343–368.

[12] Ibrahim, H. Y. & Omotesho, O. A. (2011). Optimal Farm Plan for Vegetable Production Under FADAMA in North Central Nigeria, Trakia Journal of Sciences, 9(4); 43-49.

[13] Karaatli, M. (2010). Multi-Purpose Manpower Scheduling in Fuzzy Environment: An Application for Nurses. Ph.D. Thesis, Süleyman Demirel University, Isparta, Turkey.

[14] Koç, E. (2018). An Exact Approach for a Dynamic Workforce Scheduling Problem. Master 's Thesis, Özyeg˘ in University, I˙stanbul, Turkey.

[15] Koçtepe, S.; Bedir, N.; Eren, T.; Gür, S¸. (2018). Solution of Personnel Scheduling Problem for Organization OfficersWith 0-1 Integer Programming. J. Econ. Bus. Manag., 2, 25–46.

[16] Koruca, H.I˙. (2010). Development of A Simulation-Based Shift Planning Module. J. Fac. Eng. Archit. Gazi Univ., 25, 469–482.

[17] Olive, J.T. (2009). A Proposed Mathematical Model for the Personnel Scheduling Problem in a Manufacturing Company. Master 's Thesis, Istanbul Technical University, I˙stanbul, Turkey.

[18] Özcan, E.C.; Varli, E.; Eren, T. (2017). Goal Programming Approach for Shift Scheduling Problems in Hydroelectric Power Plants. J. Inf. Technol., 10, 363–370.

[19] Özder, E.H.; Özcan, E.; Eren, T. (2019). Staff Task-Based Shift Scheduling Solution with an ANP and Goal Programming Method in a Natural Gas Combined Cycle Power Plant. Mathematics, 7, 192.

[20] Özder, E.H.; Varli, E.; Eren, T. (2017). A Model Suggestion for Cleaning Staff Scheduling Problem with Goal Programming Approach. Karadeniz J. Sci., 7, 114–127.

[21] Seçkiner, S.U.; Gökçen, H.; Kurt, M. (2007). An Integer Programming Model for Hierarchical Workforce Scheduling Problem. Eur. J. Oper. Res., 183, 694–699.

[22] Sungur, B. (2008). Integer Programming Model for Fuzzy Shift Scheduling Problems. Erciyes Univ. J. Fac. Econ. Adm. Sci., 30, 211–227.

[23] Tapkan, P.Z.; Özbakir, L.; Kulluk, S.; Telciog˘lu, B. (2018). Modelling and solving railway crew rostering problem. J. Fac. Eng. Archit. Gazi Univ., 33, 953–965.

[24] Topalog˘lu, S. (2009). A Shift Scheduling Model for Employees with Different Seniority Levels and an Application in Healthcare. Eur. J. Oper. Res., 198, 943–957.

[25] Ünal, F.M.; Eren, T. (2016). The Solution of Shift Scheduling Problem by Using Goal Programming. Acad. Platf. J. Eng. Sci., 4, 28–37.

[26] Varli, E. (2017). Solution of the Shift Scheduling Problem for the Foremans in the Manufacturing Sector with Ahp-Goal Programming. Master's Thesis, Kirikkale University, Kirikkale, Turkey.

[27] Varli, E.; Alag˘as¸, H.M.; Eren, T.; Özder, E.H. (2017). Goal Programming Solution of the Examiner Assignment Problem. Bilge Int. J. Sci. Technol. Res., 1, 105–118.

[28] Varli, E.; Ergis¸i, B.; Eren, T. (2017). Nurse Scheduling Problem with Special Constraints: Goal Programming Approach. Erciyes Univ. J. Fac. Econ. Adm. Sci., 49, 189–206.

[29] Wong, G.Y.C.; Chun, A.H.W. (2004). Constraint-Based Rostering Using Meta-Level Reasoning and Probability-Based Ordering. Eng. Appl. Artif. Intell., 17, 599–610.

[30] Zinouri, Nazanin, "Improving Healthcare Resource Management through Demand Prediction and Staff Scheduling" (2016). All Dissertations. 1682