

Grzegorz Chodak*, Witold Kwaśnicki*

GENETIC ALGORITHMS IN SEASONAL DEMAND FORECASTING

The method of forecasting seasonal demand applying genetic algorithm is presented. Specific form of used demand function is shown in the first section of the article. Next the method of identification of the function parameters using genetic algorithms is discussed. In the final section an example of applying proposed method to forecast real demand process is shown.

1. INTRODUCTION

Demand forecasting is one of the most important phase of a firm's decision making process. Management of material flow, which is based on just-in-time strategy, would not be possible to achieve without precise estimation of the demand. We can distinguish two general demand forecasting methods, namely quantitative and qualitative methods [2, 3, 6]. The advanced computer software that enables data computing and graphic presentation of data can support both of them. Artificial intelligence (AI) techniques (such as: genetic algorithms, fuzzy logic and neural networks) will be extensively applied in the next stage of forecasting methods development. Applying AI techniques should improve correctness of seasonal demand forecasting. A presentation of seasonal demand forecasting method and its applicability is the main objective of this article. Identification of demand function's parameters is based on genetic algorithms (GA) approach.

2. DEMAND FUNCTION

Shape of a demand function should be inferred from internal and external environments of each firm. It is impossible to propose general formula of a demand function applicable to all economic conditions of different firms. In some firms demand could depend mainly on one factor but in other firms tens of important factors influencing demand can be identified. Therefore the main problem is to point out that factors correctly.

*Industrial Engineering and Management Institute, Wrocław University of Technology

Frequently, as the first approximation, it is assumed that the demand function is a function of time (e.g. linear, logarithmic, exponential etc) representing a trend of demand. It is well known that the price is one of the most important factors influencing demand of almost all products and services. What we propose in this article is a combination of classical trend approach (i.e., demand as a function of time) and the economic textbook demand function (i.e., demand as a function of price). At the first stage of development of this approach we propose rather simple form the demand function, namely:

$$D = \frac{M(t)}{p^e}$$

D – demand (of products or services);

p – price;

t – time;

e – price elasticity.

$M(t)$ is function of time and it can represent a trend but when the demand is seasonal the function $M(t)$ is periodical. To include periodicity in our demand function we use the following formula:

$$D = \frac{C + B \cdot t + A \cdot \sin(\omega \cdot t + \varphi)}{p^e} \quad (1)$$

D - demand;

A – amplitude of fluctuation;

t – time;

ω - frequency (representing periodicity);

φ - horizontal offset

C – vertical offset

B – trend factor

e – price elasticity.

The above function has six parameters (A , ω , φ , C , B , e) which values ought to be identified using real data. Applying analytical method for evaluation (identification) values of these parameters is rather impossible. Therefore we propose to apply well-known and efficient approach of GA to identify that parameters.

2. IDENTIFICATION VALUES OF DEMAND FUNCTION PARAMETERS

Mathematical form of our demand function can be considered as a complex one. Therefore we need to look for a more effective method of identification. Using techniques of artificial intelligence seems to be very promising to solve this problem and within AI

techniques genetic algorithms, which are searching procedures based on natural evolution, seems to be the best one. The GA is defined by its genetic operators acting on bit strings (examples of the operators are: crossover, inversion, and mutation) and its method of credit allocation (fitness evaluation and selection). Searching for optimal solution is possible thanks to the genetic operators, which allow for small and large steps on the route search forward the optimum in a parameter space. It is done by exchanging sections of each individual's descriptive bit string (the chromosome), inverting selections of a given bit string, or mutating a bit string by altering one or more bits. The credit-allocation scheme is based on a rule that individual has different probabilities of entering into the next generation accordingly to its relative fitness.

Advantage of applying GA for identification of demand function's parameters is that there is no constraints on a shape of demand function and also on a number of parameters. The main disadvantage of GA is a lack of guarantee that global optimum is found within limited period of time. There is a possibility to adjust GA to specific requirements of real problems. We can say that GA technique is very elastic one. Here we will not discuss details of GA, it can be found in many publications, e.g., [1, 4, 5, 7].

3. APPLYING GA TO IDENTIFICATION OF DEMAND FUNCTION'S PARAMETERS

First step in applying GA to our problem of parameters' identification is to define a fitness function (F). It ought to be a measure of "distance" of model generated data and collected real data. In our experiments we use the following fitness function:

$$F = \sum_t \left| (D_t - \left(\frac{C + B \cdot t + A \cdot \sin(\omega \cdot t + \varphi)}{p_t^e} \right)) \right| \quad (3)$$

t – time;

D_t – real sell in t time;

p – price;

$A, B, C, \omega, \varphi, e$ – the same as in 2 equation – these values of parameters should be identified.

The structure of individuals in our GA is following: every individual consists of 1 chromosome divided into six segments. Each segment represents one identified parameter. Chromosome is composed of 60 bits, divided into 10 bits segments, the segments code values of amplitude, vertical and horizontal offset, frequentness, price elasticity factor and trend factor. We use binary code and every gene is coded as one bit. An example of individual is presented in Figure 1.



Figure 1 GA individual

The size of chromosome is compromise between exactness of parameters representation and quickness of computing. The greater number of bits for each parameter representation

allows for more exact evaluation of its value but length of calculation (simulation) to reach the more exact value grows enormously.

We use two genetic operators, namely:

- mutation – invert one bit in a chromosome,
- crossover - exchange of bits string between two chromosomes.

The next crucial phase of GA application is proper definition of the credit-allocation scheme. It ought to be based on relative fitness over the entire population of solutions to the problem so that a given individual has a probability of entering into the next generation according to its relative fitness. In our experiments we use modified roulette method. Each individual (chromosome) has assigned piece of a roulette wheel, which size is proportional to its fitness. The whole wheel represents a sum of fitness of all individuals of the population [5]. Additionally the individual with best fitness is chosen to the next generation (the rule “don’t lose the best”).

To make GA search more effective it is important to estimate the scope of the parameters’ domain. We can use our heuristic knowledge about the explored demand. Namely, to evaluate the range of amplitude (A) one can estimate maximum and minimum amount of sell. Maximum vertical offset (C) ought to be located between maximum of demand function and zero. The range of horizontal offset (ϕ) should be placed in the range $(0-2\pi)$. The scope of variability of price elasticity (e) is related to individual good and can be set up by a sell manager (plausible values are 0 to 2). Range of frequency (ω) can also be set by manager who can predict the seasonal demand. Accepted values of the trend parameter are related to dynamics of increase (decrease) of sell and for static environment it can set up between 1 and -1 .

If one of parameters values got from optimisation experiment is getting closer to the limit of range of domain, the range should be shifted in such way that value of parameter will be in the centre of new range.

If the ranges of parameters are set up accurately, the values of identified parameters will also estimated quicker and more precisely. It is important when we make calculations (simulation) on not very powerful computer (like PC class machine).

Definition of the stopping condition is the important for proper GA application. There are two possibilities of stopping the algorithm: when the settled number of generations is exceeded or when fitness is less then settled error. Both criteria are used in our algorithm.

4. SIMULATION EXPERIMENTS

We make our simulation experiments using real data from middle-size bakery “*Interpiek*”. The analysed good is “*Wroclaw roll*”.

After many test experiments we set up the following parameters of GA:

generation_size = 60

crossing_over_probability = 0,2

mutation_probability = 0,05

every_feature_size = 10

generations_amount = 500

Table 1 shows monthly sells of *Wroclaw roll* (from March 98 to April 99). Price is a weighted average calculated from daily sell prices in each month. Collected data suggest that bakery market is rather stable one but also seasonal demand is clearly visible.

Table 1 Sell of „Wroclaw roll” and it’s wholesale price

Month	Sell [th]	Price [zł]
0	44.32	0.26677
1	46.59	0.26630
2	50.11	0.26318
3	49.47	0.26288
4	42.73	0.26280
5	45.32	0.26235
6	47.85	0.26225
7	48.04	0.26246
8	41.76	0.26253
9	42.35	0.26259
10	44.71	0.26267
11	46.57	0.26291

For data from Table 1 we run an experiment to identify parameters of demand function given by equation (2). The result of identification experiment (10 best individuals) are presented in Table 2.

Table 2 Results of simulation (identification) experiment

Amplitude A	Frequency ω	Elasticity e	Horizontal offset ϕ	Vertical offset C	Trend factor B	Fitness F
1.22361	1.49169	0.76637	4.65005	47.46334	-0.28935	5.86275
3.40013	1.48192	0.02444	4.62659	47.46334	-0.25904	5.88233
2.52664	1.49853	0.21603	4.71261	47.39003	-0.33138	5.92691
2.74682	1.47898	0.16911	4.71652	47.27273	-0.26784	6.01847
2.38948	1.50147	0.25806	4.69306	47.40469	-0.31769	6.02310
2.06462	1.46725	0.35484	4.68133	47.44868	-0.29521	6.04227
3.46149	1.49951	0.03910	4.62268	47.36070	-0.31085	6.04354
1.98882	1.49658	0.43891	4.61486	47.37537	-0.32551	6.04656
1.98882	1.49658	0.44282	4.61486	47.37537	-0.32160	6.05567
1.95995	1.49658	0.37634	4.70870	47.37537	-0.33333	6.06452

The best individual (i.e. the bolded one in Table 2) is used to calculate the overall error (defined by equation (4)) and the partial errors are presented in Table 3.

$$E = \sum_t |Dr_t - Dp_t| \quad (4)$$

Dr_t – real demand;

Dp_t – demand counted using identified values of parameters (eq. (2));

t – time.

The sum of error equals to 5863 *Wrocław rolls* and the average error equals to 586 – it means that the error is relatively small, namely 1.28%.

Table 3 Compare of real sell and sell calculated with identified parameters

Month	Identification	Sell	Error
0	44101	44320	219
1	46698	46590	108
2	50206	50110	096
3	47601	49470	1869
4	43140	42725	415
5	44508	45318	810
6	48660	47845	815
7	47409	48041	632
8	42530	41755	775
9	42475	42351	124
		Sum:	5863

Graphic representation (see Figure 2) also suggests that the fitting of the model and its prediction is adequate and acceptable but more systematic study of verification of proposed method need to be done. The proposed approach ought to be tested on different types of goods and for wide range of market structures.

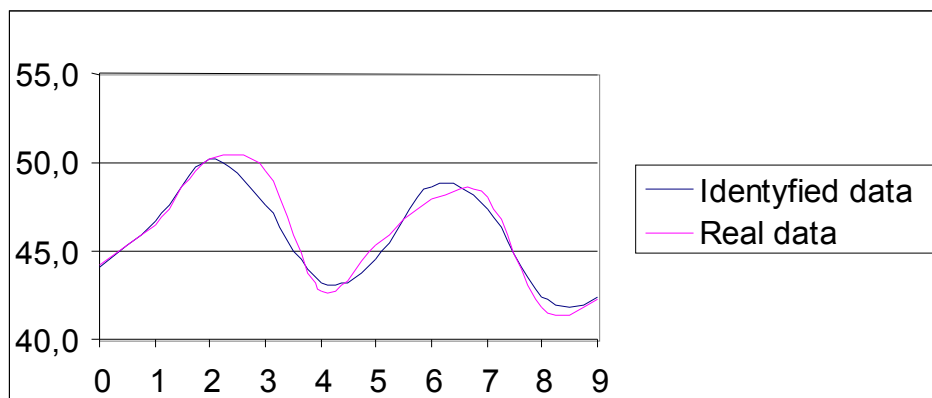


Figure 2 Real and predicted sell of *Wrocław rolls*

A prediction for next two months has been made using identified values of parameters. The results of this prediction are presented in Table 4. Comparing the prediction with real sell of *Wrocław rolls* gives the sum of error in next 2 months equal to 2548, i.e., 2.8% of real sell.

Table 4 Demand forecast on next months with no change of price

Month	Real sell	Demand forecast	Error	Price [zł]
10	44710	46811	2101	0.26259
11	46572	47019	447	0.26259

5. SUMMARY

Our results suggest that in some cases identification of demand function parameters using genetic algorithms can be more effective than in standard seasonal demand forecasting (e.g. Winter model [6]). In the presented article we have explored a situation when demand function is seasonal and has linear trend but our proposition is more general – there is no limitation to adjust the shape of the demand function to specific requirements, e.g. when double seasonalness is observed, or the trend is non-linear. Results of our experiments based on different sets of real data suggests good applicability of our proposition but it has to be said that this method does not work well for all data. There can be two general reasons of unsatisfied results of forecasting using our approach: not correct shape of demand function or not appropriate method of parameters identification. The results of presented method were discussed with “*Interpiek*” management officers. The results of the discussions are very promising, in opinion of the officers the proposed approach can be very useful in supporting decisions making on amount of *Wrocław rolls* production and can make the production more effective.

REFERENCES

- [1] GOLDBERG D., *Algorytmy genetyczne i ich zastosowania*, Wydawnictwo Naukowo-Techniczne, Warszawa 1998
- [2] E. NOWAK (ed.), *Prognozowanie gospodarcze – metody, modele, zastosowania, przykłady*, Warszawa 1998
- [3] DITTMANN P., *Metody prognozowania sprzedaży w przedsiębiorstwie*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1998
- [4] HOLLAND H. J. *Algorytmy genetyczne*, Świat Nauki, 9/92
- [5] RUTKOWSKA D., PILIŃSKI M., RUTKOWSKI L., *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*, Wydawnictwo Naukowe PWN, Warszawa 1997
- [6] SARIUSZ-WOLSKI Z., *Ilościowe metody zarządzania logistycznego w przedsiębiorstwie*, Toruńska Szkoła Zarządzania, Warszawa 1997
- [7] DRESS W.B., *On modeling and Controlling Intelligent Systems*, Instrumentation and Controls Division, Oak Ridge National Laboratory