DEMAND PREDICTION MODEL OF AN ORGANIZATION

Oroselfia Sánchez[(a)], Idalia Flores[(b)]

[(a),(b)]Facultad de Ingeniería. División de Posgrado. Universidad Nacional Autónoma de México.

[(a)]oroselfia.sanchez@yahoo.com.mx, [(b)]idalia@unam.mx

ABSTRACT
Nowadays, the organization resources management should be more accurate to avoid delay risks and penalties. One start point for planning are historical data, which form the base for forecasting the future demand. This demand forecast is used to plan resources acquisition and resources management. Frequently, the management systems of the organization do not allow detailed historical record of activities, which causes insufficient data to forecast demand and other important aspects.

In this paper, data of a petroleum private enterprise are treated. The purpose of this work is to know all the feasible demand scenarios that may occur through a mathematical model based in a Poisson distribution from a scorecard. The objective of the model is to forecast demand of two types of services offered by the enterprise. Finally, a simulation model is used to validate this mathematical model.

Keywords: Poisson Distribution, Demand, Forecasting Scorecard, Risks, Simulation.

1. INTRODUCTION

In organizations, budgets for acquisition of material and human resources depend on demand of future services and products. A good prediction of such demand allows a more detailed planning of the required resources. When organizations do not have the necessary capacity to produce products or provide services failures appear, causing penalty costs. This is important, because costs of prevention are lower than those originated from corrections, since the possible impact may be difficult to assess.

Some common methods employed to forecast demands can be qualitative, such as expert opinions, surveys to costumers, Delphi methods (Middendorf,1973), and quantitative of classical and Bayesian kind (Niu, Zhao and Liu, 2009), such as random variables and time series models (Qiu, Suganthan and Amaratunga, 2016), causal and stochastic models (Ma, Wu, Khanwala, Li and Dang, 2015), Markov chains (Wan, Zhang and Dai, 2014), networks (Zou, Huang, Chen and Qu, 2011), simulations (Chen, Lu and Shao, 2010), among others. Moreover, some methods should be adapted in order to obtain data that helps to solve the particular issues of the organization, including constrictions addition and system properties.

In this work, we propose a mathematical model adjusted to the necessities of an organization with the purpose of obtaining expected demand events. The model uses as input data those historically recorded in a scorecard. It contains the number of two different kinds of services based on historical data and expert group predictions in a monthly frequency.

2. THE ORGANIZATION AND ITS DATA

In this work, a demand forecast model is established for a petroleum private organization, whose activity is oil well cementing. This company offers two types of services, oil well cementing (i) and pumping jobs (j). The first one consists in building a cement wall to stabilize and isolate the oil well, while the second one is the pumping of fluids from the surface to the bottom of the oil well. The resources of the organization are shared by both kind of jobs.

Jobs i and j performed each month are recorded in a scorecard format. It does not specify situations that occur by day, and then details of demand in worst scenarios are unknown. Scorecard data recorded for a year is shown in Table 1, scorecard data.

Managers would like to know, the maximum and minimum limits of jobs that probably occur by day and by job type using the scorecard data. With this information, the quantity of resources could be planned. Quantitative results obtained could help to analyze and improve predictions that otherwise are only based on beliefs.

3. MATHEMATICAL MODEL

3.1. Obtaining the model.

The objective of the mathematical model is to estimate the number of days that i cementing and j pumping jobs occur during a given month. According to expert group opinions and historical data recorded in scorecards, feasible scenarios are given by $i = \{0,1,...,6\}$ and $j = \{0,1,...,6\}$. Possible events are represented by a pair (i,j) that specify the number of cementing and pumping jobs in a given day. Calculation of occurrence probability of every event by month could be performed to obtain the number of days of each situation by month. We use a Poisson distribution to appraise probability of each kind of job in the period of one month. However, Poisson distribution do not consider constrictions of the system. In consequence, we add such restrictions to the model, to adapt it to data results required. In particular, Poisson distribution gives results of real type; however we need to count the number of days using natural numbers. Poisson distribution is then employed to establish inferior (d_{ij}^-) and superior (d_{ij}^+) bounds for the number of days that occur every event. These bounds allow the adjustment of data to integer numbers with the purpose to determine the Total Number of Days per month that each event can occur. The superior limit d_{ij}^+ is given by $d_{ij}^+ = round\left(\bar{d}_{ij} + \varepsilon\right)$ and the inferior is defined by $d_{ij}^- = round\left(\bar{d}_{ij} - \varepsilon\right)$, where ε is a value chosen, varying arbitrary between 0.2 and 0.3 and $\bar{d}_{ij}$ is the average number of days that event is presented per month, as obtained from the Poisson distribution.

Table 1: Total Number of each type of Jobs by month (Data from Scorecard Organization and Simulation results)

Month	**scorecard data**		**simulation results (average)**	
	(i)	**(j)**	**(i)**	**(j)**
January	33	51	32.9065	50.3874
February	29	46	28.9184	45.4272
March	42	56	41.9151	55.7194
April	48	53	47.517	53.301
May	47	47	46.6829	46.717
June	59	64	58.872	64.5
July	51	77	51.0167	77.4845
August	46	32	45.8025	31.5642
September	56	10	55.542	9.87
October	53	46	52.9759	45.7405
November	47	35	46.41	34.881
December	64	28	63.2989	27.962

In other words, the number of days that an event is expected to occur during a month is obtained through the determination of an integer-distribution that is closer to the real-Poisson-distribution. The mathematical model seeks the achievement of the following restrictions:

- The sum of days of occurrence of each event is equal to the number of days in a month.
- The number of days of occurrence of each event is less than the days calculated in the upper limit and greater than the days calculated in the lower limit.
- The addition of the number of days that each event occurs per month, multiplied by the number of cement works, is equal to the total number of cementations registered in the scorecard.
- The addition of the number of days that each event occurs per month, multiplied by the number of pumping jobs is equal to the number of total pumping jobs registered in the scorecard.

Considering the pairs (i,j), the model to obtain the demand by day and by month is the following,

$$\max z = \sum_{i=0}^{6}\sum_{j=0}^{6} p_{ij} d_{ij}^{(m)} , \tag{1}$$

subject to

$$\sum_{i=0}^{6}\sum_{j=0}^{6} d_{ij} = D_m \tag{2}$$

$$d_{ij} \geq d_{ij}^+ \tag{3}$$

$$d_{ij} \leq d_{ij}^- \tag{4}$$

$$\sum_{j=0}^{6}\sum_{i=0}^{6} i d_{ij} = T_{cmt} \tag{5}$$

$$\sum_{i=0}^{6}\sum_{j=0}^{6} j d_{ij} = T_{pump} \tag{6}$$

$$p_{ij}, d_{ij} \geq 0 \tag{7}$$

where,

i = Number of Cementing jobs

j = Number of Pumping jobs

p_{ij} = Probability of each event

d_{ij} = Number of days

D_m = Total number of days that occur the pair (i,j)

T_{cmc} = Total number of Cementing jobs in a month (scorecard)

T_{pmp} = Total number of Pumping jobs (scorecard)

The mathematical model can be solved through MS-Excel Solver to obtain data of the combination of cementing and pumping jobs by day and by month.

3.2. Validation of the Model.

In order to validate the mathematical model, a simulation in SIMIO platform was made.

3.2.1. Simulation

The simulation models a generation of cementing jobs *(i)* and pumping jobs *(j)* month by month.

Figure 1 shows how the simulation works. In each run the number of jobs of each kind (i and j) per day are stored. At the end of the month it is determined the number of days that happened i, cementing and j pumping jobs. Results are then compared with those predicted by the mathematical model. The processes were set into SIMIO following the logic shown in Figure 2.

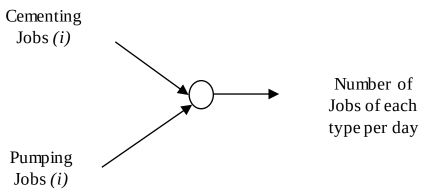

Figure 1: Conceptual Simulation Model

4. ANALYSIS OF THE RESULTS

4.1. Mathematical Model results

The search of information about a system can be done with different approaches by the managers. Some are based on the statistical mode per month and observe possibilities around it. Others, analyze all the different scenarios which can happen in one day. It depends on the approach of the managers what data are chosen to make decisions.

The model proposed in this paper, seeks the most feasible scenarios that may be presented based on scorecard data. Mathematical model was solved in MS-Excel Solver allowing to obtain the expected events per month. Maximization is performed for each month separately based on their corresponding scorecard data. Figure 3 shows the results, in color scale, of the mathematical model for each situation per month. Color scale goes from cero to seven, which can be interpreted as the number of times the event was obtained during month.

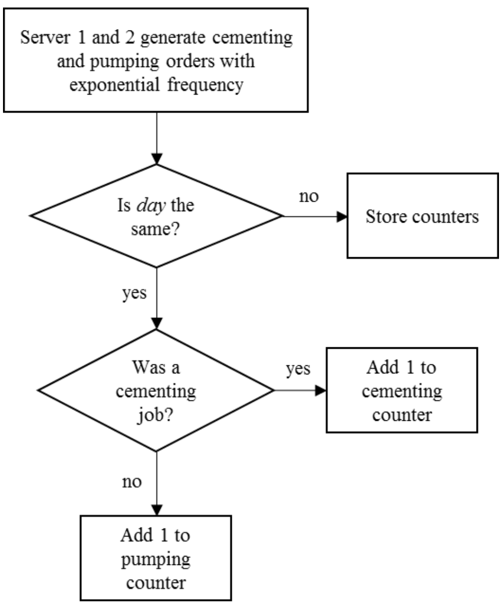

Figure 2: Logic structure of the simulation model

All these data results demonstrate a plethora of different possible scenarios or events that the organization could present during a year from the scorecard historical data. We can see in Figure 3 the results month by month, presenting different demand behavior between them. The most representative results are explained below.

In January, the expected events that can happen oscillates in the combination of $i = (0,1,...,3)$ and $j = (0,...,4)$. According to the numerical results, the events that are most likely to occur during the month are around 1 pumping job and 1 cementing job. However, an event which could demand a great quantity of resources is the pair $i, j = (3,3)$

The enterprise policy will influence resources management strategy. Will the organization be prepared for the hardest days of the month? Or for the most frequent days? In the first case, the maximum capacity of their resources will not be occupied in several days, which generates costs per day of non-occupation. In the second case, the company may require external resources causing variable costs.

Events obtained for June and July are different from January. June events are widespread combination of *i* and *j* indicating a high number of days in which there are several events that require more resources, this means that organization may have more cementing and pumping jobs in a day during several days. Both in June and in July, the variation within the same month is high which generates more management and scheduling uncertainty.

In September opposite happens because is the month with lowest reported jobs (see Table 1 scorecard data) in comparison to other months due to a few number of required pumping jobs. However, the obtained high number of cementing jobs causes that possibilities with 0 to 5 cementing jobs become feasible.

Events resulting from the model may show the months with fewer down time. How does it help us to know this data? If there are months with high demand of services, activities as maintenance, capacitation or preparation of resources must be anticipated by the organization. If the information provided to the stakeholders includes those months with more jobs, then contingency plans may be determined in advance. Thus, considering acquisition of outsourcing maintenance, personal and resources. The previous prevent the company to possible risks ensuring to avoid the surprise factor.

4.2. Simulation Model Results

The results of simulation indicate that expected events may be only slightly deviated from those predicated by the mathematical model. Such deviation is represented by the area enclosed by the dashed line in Figure 3 for January. In fact, some events out of the area also occur but with very low frequency and are related to not including restriction in the number of jobs suggested by expert opinions. For example, feasible sets for *i* and *j* were not included. In general, situation results concords with the mathematical model in determining the most probable events.

For example, in Figure 3 the results of the simulation do not exceed the white dotted line that can be appreciated for the month of January. This validates that mathematical model represents in a correct way the most frequent scenarios and the combinations of these.

ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

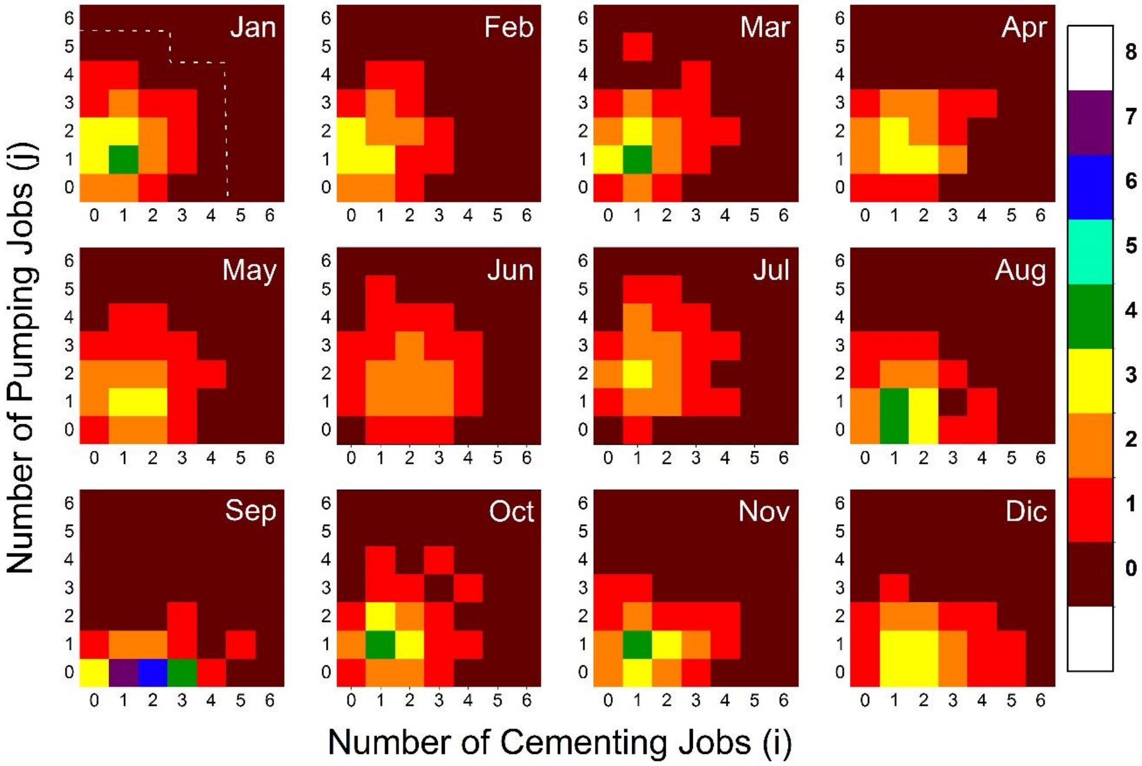

Figure 3: Mathematical model results per month

An important highlight is that both, mathematical and simulation modelling are a powerful tool to analyze the demand. But in this case, through simulation as a tool, a lot of runs must be done to determine the same feasible region obtained by the mathematical model. However, one of the advantage of simulation is that it invites to us to analyze more the system. For example, simulation gives explicit idea of worst scenarios, that may be considered as anomalous by the mathematical model. They occur rarely, but it does not mean that the company will never face those events.

In addition to the previous results, Table 1 shows a comparison between the average of number of jobs resulting from simulation, versus historical scorecard data. The averages were obtained in a period of 1000000 days (simulation time) for every month.

4.3. The worst scenarios

When the resources are planning, the scenarios or situations which are commonly considered are the most frequent. Determination of worst scenarios proposes to have a broader paradigm of demand behavior to make better decisions. It is always good to know which are the worst scenarios that organization could have before they happen. The resources must be managed to comply most jobs, but also should be useful to have a plan that allows a rapid response to a contingency. In this sense, the analysis of the worst scenarios is a powerful tool. A lot of risk appear when bad decisions are taken, and with high costs consequences.

Figure 3 shows the results obtained through the mathematical model. How can you accomplish a worst scenarios analysis with this type of data? In this framework, we could do through two options. The first one, is determining those events closers to the delimited area in Figure 3 (for each month). For example, for January the worst scenarios may be follow the pairs of *i,j*: (0,5), (1,5), (2,4), (3,4), (4,4), (4,3), (4,2), (4,1), (4,0) and (0,3). Those events, are very closer to the occurrence region and even the maximization results indicate that they do not occur, it does not that this will not happen in the real world.

If it is desired to explore even more the system, as a second option simulation can be used to find worst scenarios outside of the white dotted line in Figure 3. This results could present some ideas of the events which are considered as not feasible for the expert group but in real world have a tiny ocurrence probability.

With the results of this analysis, the managers plan what to do if the events with the color representing zero in Figure 3. Would the company refuse to provide the service? Or would they risk being penalized for lack of resources? Would request external resources with variable costs? How much they would this cost to the organization?

The point as a manager is to keep a curious thought regarding the situations that may occur. Different techniques used to analyse the demand allow to have more insights of the system and prevent organizations to incur in operational or other kind of risks.

5. CONCLUSIONS

Commonly, organizations deal with missing data for the planning of their systems and the lack of information generates high uncertainty. This method presented in this paper, can help to obtain more information about the probable demand situations that can occur beyond based on historical data.

In this work, a mathematical model was proposed and developed to analyse the demand from an historical scorecard (Table 1, scorecard data). The number of jobs per month generated by mathematical model, are the same as those in Table 1 scorecard data because it is a constriction of the model. The results were shown in feasible scenarios form with the corresponding frequency obtained (Figure 3) and validated through simulation model. Besides, the resulting events give us insights about the demand in systems.

Through mathematical model it was possible to obtain a broad picture of the demand behaviour. Also, the addition to the model own features permit to show data in a form useful to make decisions in management. The combination of techniques allows to discover the details of the resulting events.

A more careful analysis of the data available has the purpose of give more information of the system preventing delays, penalties and operational risks. This model helps to reduce lack of knowledge, which could decrease uncertainty and risks by helping prevention decisions.

ACKNOWLEDGMENTS

This research was supported by UNAM PAPIIT grant IT102117. The first author specially acknowledges the support from CONACYT scholarship program.

REFERENCES

Chen M., Lu J. and Shao Y., 2010. An application of Monte Carlo Simulation: Activity Based Urban Travel Demand Forecast Modeling. International Conference on Computer Design and Applications, 4, V4-251–V4-254.

Ma R., Wu R., Khanwala M., Li D. and Dang S., 2015. Power Demand Forecasting using Stochastic Model: Parameter Estimation. Modern Electric Power Systems.

Middendorf W.H., 1973. A modified Delphi method of solving business problems. IEEE Transactions on Engineering management, EM-20 (4), 130–133.

Niu L., ZhAO J. and Liu M., 2009. Application of Relevance Vector Regression Model based on Sparse Bayesian Learning to Long-term Electricity Demand Forecasting. Proceedings of the 2009 International Conference on Mechatronics and Automation. 2363–2367.

Qiu, X. and Suganthan P.N., 2016. Electricity Load Demand Time Series forecasting with empirical Mode Decomposition based Random Vector Functional Link Network. International Conference on systems, Man and Cybernetics, 001394 – 001399.

Wan X., Zhang Q. and Dai G., 2014. Research on Foresting method of natural gas demand based on GM, Model and Markov Chain. International Conference on Cognitive Informatics & Cognitive Computing, 436 – 441.

Zou A.Q., Huang G., Chen J.L. and Qu Y.Q., 2011. The Inventory Demand Forecasting Model of the Regional Logistics Network in Supply Chain. Industrial Engineering and Engineering Management, 1565 – 1569

AUTHORS BIOGRAPHY

Oroselfia Sánchez Sánchez studied Chemical Engineering at Faculty of Chemistry of Universidad Veracruzana (UV), then a Specialization in Quality Systems at Universidad Nacional Autónoma de México (UNAM) obtaining the Alfonso Caso Medal. She has worked on Petroleum Industry in Quality, Operation and Management areas. She received a Master in Operational Research with honors at UNAM. She was also a Professor of the Specialization in Quality Systems at UNAM. She is currently studying a PhD in Operations Research at UNAM.

Idalia Flores de la Mota received a Master with honors, being awarded the Gabino Barreda Medal for the best average of her generation, in the Faculty of Engineering of the UNAM, where she also obtained her Ph.D. in Operations Research. Dr. Flores is a referee and a member of various Academic Committees at CONACYT as well as being a referee for journals such as Journal of Applied Research and Technology, the Center of Applied Sciences and Technological Development, UNAM and the Transactions of the Society for Modeling and Simulation International. She is a full time professor at the Postgraduate Program at UNAM and her research interests lie in simulation and optimization of production and service systems, based on this research she has been editor of two books edited by Springer.