

## **IMPROVING HISTORICAL DEMAND DATA THROUGH SIMULATION**

Oroselvia Sánchez

Universidad Nacional Autónoma de México  
Av. Universidad 3000, Cd. Universitaria  
Coyoacán, 04510, Ciudad de México

Idalia Flores

Universidad Nacional Autónoma de México  
Av. Universidad 3000, Cd. Universitaria  
Coyoacán, 04510, Ciudad de México

### **ABSTRACT**

Organizations collect data during different periods so that they can use them for management and business purposes. However, the data do not always come in the most suitable form for analysis, and often needs to be prepared, for which there are a variety of methods, including simulation. This paper presents a case where simulation is used as a tool to get insights into demand, based on historical data. Through simulation, we extract the most frequent demand events for two types of jobs together with the worst events. The simulation model is based on the historical data from a private oil company that operates in Mexico. In addition, we show how simulation results improve the information about Scorecard data recorded during a year of work.

### **1 INTRODUCTION**

Organizations generally record demand data to use them for forecasting possible future requirements in order to be in a position to avoid or prevent risks such as penalties, delays, insufficient capacity, among other things. However these data normally need to be pretreated to fit the requirements of the analysis.

To facilitate this, many organizations have robust record systems where employees record data in a timely fashion; however every step of an operation is not always recorded. This is a persistent problem in several organizations, meaning that the historical data has to be prepared before its analysis. But what happens when the recorded information needs to be used but is not in the right form for the analysis? Nowadays, a lot of companies prefer not to use this information while others use a variety of methods to get as many insights as possible for planning their future activities; methods such as networks (Zou et al. 2011); exponential smoothing (Mohammed et al. 2017); time series models (Qiu et al. 2016); causal and stochastic models (Ma et al. 2015) and simulations (Chen et al. 2010).

The veracity of the information can open up the possibility of more accurate planning of resources, budget and possible new locations, new job positions or scheduling of activities.

In this paper, historical data from a scorecard is analyzed and prepared using the simulation method, as it enables us to generate a lot of different possible scenarios for the model entities. The purpose of this analysis is to identify risks by obtaining demand events that can be presented on a day-by-day basis for every month of the year.

### **2 HISTORICAL DATA AVAILABLE FOR DEMAND.**

The demand data we analyzed is the historical data from an organization that cements oil wells in different states of Mexico. Basically, it offers two types of services: the first being the Cementing Job (*i*) that includes the design of cement slurry and building a circular wall inside the oil well, while the second one refers to the Pumping Job (*j*), which consists of the leasing of resources. In order to meet the demand for both services, the organization uses the same resources for both types of services, *i* and *j*. The total num-

ber of each kind of job per month during a year is recorded on a scorecard, whose data are given in Table 1.

In practice, the scorecard is used by the organization to forecast the future behavior of its work. It is not uncommon to made mistakes that mean that there are insufficient or an excess of resources in certain locations. This happens because the total number of jobs recorded does not show the possible events that could happen day by day or the frequency of each event. For example, we cannot see in Table 1 how many jobs were requested of the company on July 3rd, or during how many days in any month the resources were used at maximum capacity. These details are omitted in these kinds of records.

TABLE 1. SCORECARD DATA

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<i>i</i>	33	29	42	48	47	59	51	46	56	53	47	64
<i>j</i>	51	46	56	53	47	64	77	32	10	46	35	28

In this paper, starting from Scorecard data, a model is established to obtain events information which will give the organization more insight into how its demand behaves. Especially the more frequent events and those that only happen sporadically but could cause operational risks for the organization, such as delays, penalties and non-compliance.

### 3 SIMULATION MODEL FOR DEMAND.

The purpose of the simulation model is to generate the most frequent events and find the ones that could be critical for the company. The demand simulation model is executed for each month of the year where the demand has been simulated according to Table 1 data.

#### 3.1 Relevant aspects of the Simulation Model

SIMIO is the simulation platform we chose to execute the model because its characteristics are suitable for developing the model and using it with other models later on.

The simulation model has servers to generate entities, representing jobs *i* and *j*. Each server responds to distributions that complies Scorecard data (Table 1). Both servers are connected to a Sink counting the number of jobs of each kind per day. As result of the model, the number of days that occur *i* cementing and *j* pumping jobs are obtained monthly.

The model records the results starting from day 1200 for the purpose of report results. The simulation results have been determined after 1000 days of simulation time.

TABLE 2. AVERAGES OF WORKS GENERATED BY SIMULATION MODEL

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<i>i</i>	32.9065	28.9184	41.9151	47.517	46.6829	58.872	51.0167	45.8025	55.542	52.59759	46.41	63.2989
<i>j</i>	50.387	45.427	55.719	53.301	46.717	64.5	77.485	31.564	9.87	45.741	34.881	27.962

The results of the validation of the model are illustrated by comparing Table 2 with Table 1, which gives the averages for the runs of the model per job for every month.

Besides the simulation model, a mathematical model has been constructed, enabling us to compare the resulting events.

### 3.2 Simulation Model Results

The simulation model generates a combination of Jobs  $i$  and  $j$  with the corresponding frequency of each event for every month of the year. This combination of a number of jobs allows us to group the most frequent events and enables us to understand the demand per month. At the same time, the information about events provides us with information that is useful for resources management.

With detailed information available, it is possible to know those events or scenarios which are not frequent but can happen in some day of the year. Figure 1 also shows nearby events to probability region of events but with one time of difference in the frequency obtained. Those events can be analyzed for the purpose of knowing how the organization will react if they happen, and consequently an action plan can be formulated.

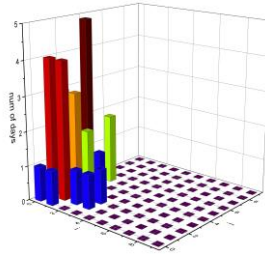


Figure 1. Number of jobs  $i$  and  $j$  generated in January simulation.

The importance of preparing the data using the methodology presented in this paper is that the knowledge of information details can avoid operational risks by reducing the surprise factor of the system.

## 4 CONCLUSIONS AND DISCUSSION

Due to the changing environment and the resilience that systems need to have in respect of their own reality, more robust approaches and tools can offer better results. In this sense, Risk Identification, using tools that provide more detailed information about a system, offers great insights into the various risk situations the organization may face. Simulation is a tool that allows us to analyze a system and its information better and faster, allowing the emulation of unexpected changes with possible behavior of a system.

## 5 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*.
- 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.
- Mohammed, J., Bahadoorsingh, S., Ramsamooj, N, and Sharma, C. 2017. “Performance of Exponential Smoothing, a Neural Network and a Hybrid Algorithm to the Short Term Load Forecasting of Batch and Continuous Loads”. *PowerTech, 2017 IEEE Manchester*.
- 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