Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 16 (2014) 163 – 171

Procedia Technology

CENTERIS 2014 - Conference on ENTERprise Information Systems / ProjMAN 2014 - International Conference on Project MANagement / HCIST 2014 - International Conference on Health and Social Care Information Systems and Technologies

A multi-agent supply chain simulation analysis through a statistical mixed model

Alexandra Seco[a]*, Carlos Vieira[b]

[a] School of Technology and Management, *Polytechnic Institute of Leiria, 2411-901 Leiria, Portugal*
[b] *School of Technology and Management, Polytechnic Institute of Leiria, 2411-901 Leiria, Portugal*

Abstract

Supply chain coordination is a problem that arises in a supply chain and it is necessary to understand and analyse it as a set of dependencies both in physical flows, as well as in informational flows. The need to manage these dependencies is important for a company's success. In this paper, data from a supply chain simulation model, which is based on the multi-agent system simulation approach, was considered. A pharmaceutical case study was used to help and explore various coordination mechanisms, under different operational conditions. These scenarios were previously proposed and presented in the work of Vieira [9], and it becomes necessary to analyse and choose the best procedure to coordinate the supply chain, which will be done using a statistical model. In each scenario, data reveal within-group correlation and statistical mixed models are a flexible way to model this dependence, and, therefore, are used in this case study and presented in this paper.

Peer-review under responsibility of the Organizing Committee of CENTERIS 2014.

Keywords: supply chain coordination; multi-agent system; scenario simulation; within-group correlations; linear mixed models; nonlinear mixed models.

* Corresponding author. Tel.: +351 96 51 44 906; fax: +351-244-820-310.
E-mail address: alexandra.seco@ipleiria.pt

Peer-review under responsibility of the Organizing Committee of CENTERIS 2014.
doi:10.1016/j.protcy.2014.10.079

1. Introduction

A supply chain (SC) is a network of trading partners involved, through upstream and downstream linkages, in different processes and activities to produce value in the form of products and services delivered to the ultimate consumers so as to provide profit for each SC member (Christopher [1]). These SC partners operate subject to different sets of constraints and objectives, which are highly interdependent when it comes to improving performance of the SC in terms of objectives, such as on-time delivery, quality assurance, and cost minimization. Toward cooperative efforts to synchronize and converge intra-firm and inter-firm operational and strategic capabilities into a unified whole, a new business management philosophy, called Supply Chain Management (SCM) has appeared. SCM is defined by Mentzer *et al*. [6] as the systemic, strategic coordination of the traditional business functions and the tactics across these business functions within a particular company and across businesses within the SC, for the purpose of improving the long-term performance of the individual companies and the SC as a whole. Coordination, as an essence of SCM, is the management of dependencies between activities. These dependencies stem from the lack of ability to control all the necessary conditions to achieve a desired outcome. Coordination may take place within operations, across functions (cross-functional) or between organizations (inter-organizational). Its purpose is to collectively achieve goals that individual actors cannot meet.

Therefore, supply chain coordination (SCC) offers means to understand and analyse a SC as a set of dependencies both in physical flows, as well as in informational flows. The need to manage these dependencies is important for a company's success. SCC is a vehicle for redesigning decision rights, workflows, and resources among SC members so as to leverage improved performance (Lee and Whang [5]).

In this paper, a SC simulation model, which is based on the multi-agent system (MAS) simulation approach, where the performance of multi-product, multi-echelon SC, subject to different sets of operational constraints, was considered to explore various coordination mechanisms. This SC simulation model, over a pharmaceutical SC case study, provides an approach which may help to explore various coordination mechanisms, under different operational conditions, to coordinate the SC.

Besides that, analysing final results and concluding remarks on each scenario are needed in the simulation. For this purpose, a statistical mixed model was considered to achieve proper conclusions and is presented in this paper.

2. The simulation model

A Multi-agent System (MAS), a branch of distributed artificial intelligence, is a system of interacting agents, where an intelligent agent is defined as "a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives" (Wooldridge *et al.* [10]). When applied to SC, agents are largely seen as intelligent components for automating actions of operation, planning, optimizing, controlling in an ambience of frequent exchange of data with others. Agents can be modelled to represent organizations, functions, resources, and even human beings.

A proposed MAS model, called MASCS, was previously developed in order to analyse coordination issues and global performance of SC systems under different inventory, production, procurement and shipments policies scenarios (Vieira [9]).

The MASCS model was implemented using the Java Agent Development Framework (JADE™). JADE™ is a software development framework aimed at developing MAS and applications conforming to the Foundation for Intelligent Physical Agents (FIPA) standards, which are intended to promote the interoperation of heterogeneous agents and the services that they can represent.

The dynamics of a SC of a vertically integrated multinational pharmaceutical company was studied and simulated. The performance measures of their prime products are estimated based on the applied coordination mechanism, inventory management and production policies. A diagram representing the pharmaceutical SC is shown in Figure 1.

The SC considers four final products being sold on two markets: *Pack A* and *Pack B* on the Japanese market; *Pack C* and *Pack D* on the USA market. These products are packs of tablets made with an Active Ingredient (*AI*), which is produced in the primary plant in Europe via five synthesis stages (Table 1). The secondary plant in Asia produces *Pack A* and *Pack B* via two steps: formulation (i.e. the tablet production from *AI*) and then packing. It also produces bulk products Tablet *EurExp* and Tablet *GenExp* to fulfil fixed supply agreements with pharmaceutical industry. *Pack C* and *Pack D* are produced in the secondary plant in America also via the two steps of formulation and packing. The four final products, *Packs A, B, C* and *D*, are shipped directly to the distribution centers in Japan and the USA. The transportation lead times between SC nodes are also presented in Figure 1.

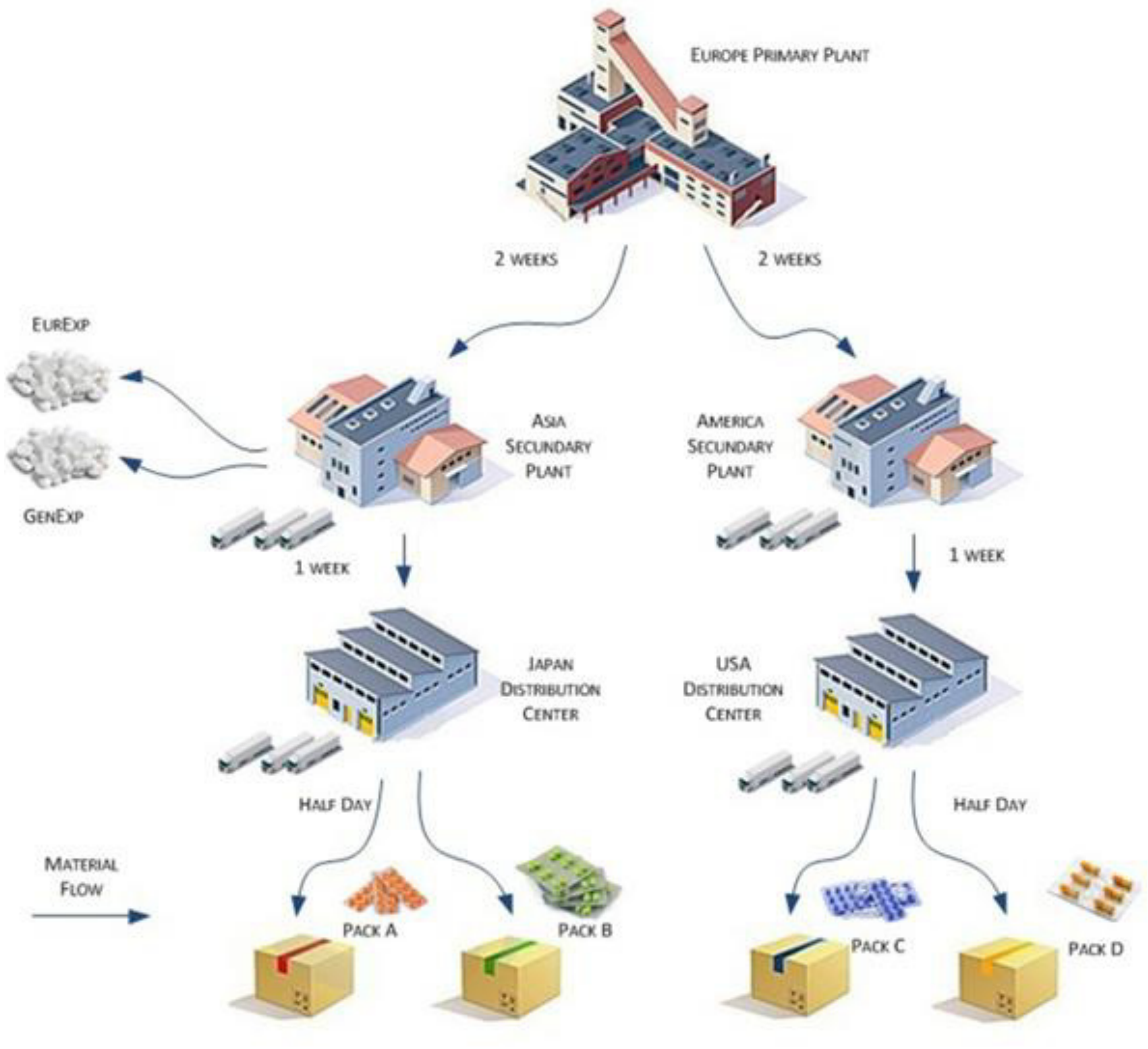

Fig. 1. Pharmaceutical case study SC overview.

The bill of materials, batch sizes and process durations at the various the SC secondary plants are illustrated in Table 1. Further complementary descriptive information about the SC case study can be found in Vieira [9].

Table 1. Bill of Materials, Batch Sizes and Process Durations at Asia and America Secondary Plant.

				Process Duration	
Product	Formulation	Packing	Batch Size	Formulation	Packing
Pack A	20 mg AI in 1 Tablet A	12 Tablets A in 1 Pack A	9000 packs	0,030	0,625
Pack B	30 mg AI in 1 Tablet B	30 Tablets B in 1 Pack B	6000 packs	0,050	0,415
Pack C	30 mg AI in 1 Tablet C	30 Tablets C in 1 Pack C	30000 packs	0,575	0,520
Pack D	30 mg AI in 1 Tablet D	30 Tablets D in 1 Pack D	42000 packs	0,788	0,730

In our statistical model we only consider *Pack A* product; for *Pack B*, *C* and *D* the model would be similar.

Six scenarios were simulated (scenarios S1 to S6) to explore the application of different collaborative initiatives. These scenarios can be further classified in two sets according to the replenishment control policies types applied along the SC nodes. The first set (scenarios S1 to S4) use instantaneous replenishment control policies; the second set of scenarios (S5 and S6) use time-phased replenishment control policies at SC nodes.

Several performance measures were implemented to evaluate the pharmaceutical SC performance under each scenario. One recommended performance measure, *CSL,* by Hung *et al.* [3], was adopted in order to compare the achieved SC performance between the current scenarios and the ones from the original case study performed by this author. This performance measure is taken to be the ratio between the total quantity sold directly from shelf over the total quantity ordered, for a given product at any given SC node for the time horizon length (in the present case: 104 weeks). It is our goal to perform a statistical model which will allows us to model this measure under different scenarios and therefore achieve conclusions about significant differences between them.

In figure 2 we can see the data obtained under simulation with the Java Agent Development Framework (JADE™), an open source toolkit, developed by Telecom Italia Laboratories, where messages exchanged by MASCS agents were specified using the Agent Communication Language (ACL) format (FIPA 2001).

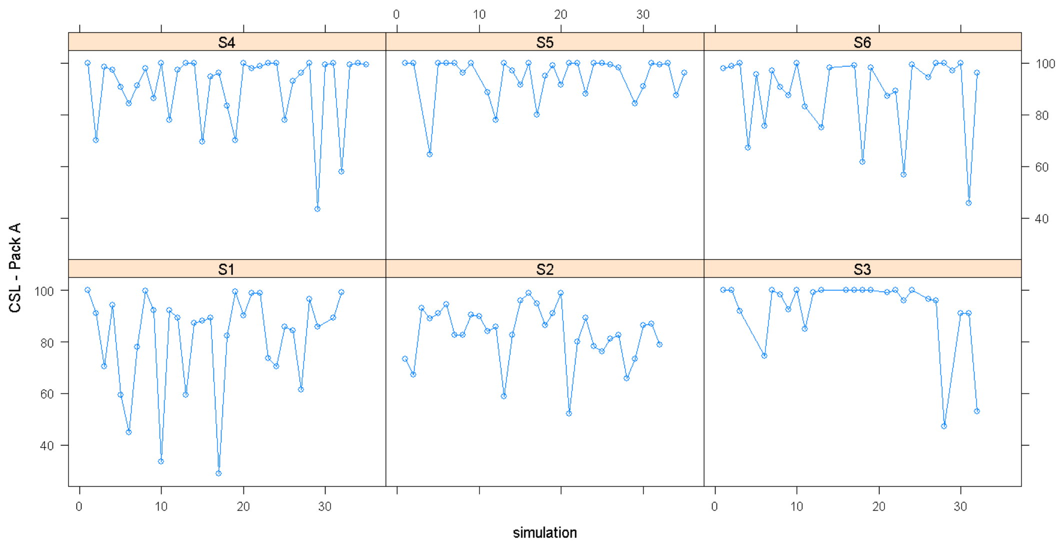

Fig. 2. Pack A CSL measure by scenario.

Our aim is to develop a statistical model to analyse this data in order to find the best scenario that coordinates the SC. As we can see, within each scenario we have several observations generated by the simulation. Therefore, these observations should be correlated, and the usual statistical models are not adequate to model this data, since they state that the observations are independent. Linear mixed models are a flexible way to model grouped data with within-group correlations and will be presented and used here.

3. The statistical mixed model

In many applications grouped data reveal within-group correlation. Examples are longitudinal data and repeated measures data. We can find in literature several examples where ignoring the group structure can lead to imprecise estimates, confidence intervals and significant tests. Grouped data should be modelled respecting its particular structure.

For continuous data there are several available models which are used. These include mixed models (Laird *et al.* [4], Seco [8]) which embody fixed and random effects. These models are based on the Multivariate Normal distribution, which has friendly properties, as the marginal and conditional distributions are still Normal.

Let $\boldsymbol{y_i}$ be a vector of dimension n_i, i = 1, ... , *M*, a vector observations grouped in *M* groups. For a single level of grouping, the linear mixed model proposed by Laird *at al.* [4] is of the form,

$$\boldsymbol{y}_i = \boldsymbol{X_i\beta} + \boldsymbol{Z_i b_i} + \boldsymbol{\epsilon_i},\ i=1,\ldots,M, \\ b_i \sim N(\boldsymbol{0},\boldsymbol{\Sigma}),\ \boldsymbol{\epsilon_i} \sim N(\boldsymbol{0},\sigma^2\boldsymbol{I}), \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of fixed parameters (fixed effects) of dimension p , $\boldsymbol{b_i}$ is a vector of random parameters (random effects) of dimension q , $\boldsymbol{X_i}$ and $\boldsymbol{Z_i}$ are model matrices of order $n_i \times p$ e $n_i \times q$, respectively. The columns of $\boldsymbol{Z_i}$ are generally a subset of the columns $\boldsymbol{X_i}$.

$\boldsymbol{\epsilon_i}$ is a vector of dimension n_i, called residual "error " within-groups or simply by error within-groups.

It is assumed that $\boldsymbol{b_i}$ and $\boldsymbol{\epsilon_i}$ have multivariate Normal distribution $N(\boldsymbol{0},\boldsymbol{\Sigma})$ and $N(\boldsymbol{0},\sigma^2\boldsymbol{I})$, respectively, independent for different *i* (groups) and independent of each other, that is, their covariance is zero.

The assumption that $var(\boldsymbol{\epsilon_i}) = \sigma^2\boldsymbol{I}$, i = 1 , ... *M* , can be generalized to other variance structure.

The $\boldsymbol{\Sigma}$ matrix, variance-covariance matrix, must be symmetric and positive semi-definite, but here will be assumed that this is always positive definite. This restriction implies that, under other conditions, the model can be redefined so that it includes a positive definite matrix of lower dimension. The random effects $\boldsymbol{b_i}$ have mean value zero, so any random effect of nonzero average value should be included in the fixed component of the model.

The linear mixed model can be seen as an extension of the classical linear model, in which is considered an additional "error" reflecting the correlations between observations belonging to the same group. Because observations made on the same group share the random effect $\boldsymbol{b_i}$, they are correlated.

The estimators of fixed $\boldsymbol{\beta}$ parameters are obtained through the likelihood function associated with the mixed model, but the estimation of variance components (σ^2 and $\boldsymbol{b_i}$) by the this method raises some problems. One of the properties of the maximum likelihood method is that, when estimating the variance components, it does not take into account the degrees of freedom involved in estimating $\boldsymbol{\beta}$. Thus, maximum likelihood estimators, in general, have no minimum variance. A workaround for this problem is to use the restricted likelihood function, instead of the usual function, and obtain the restricted maximum likelihood estimators designated by REML estimators. The idea is to maximize only the likelihood function of the parameters comprising the variance.

The criteria most commonly used in inference analysis in this type of models, which presents the greatest advantages and lead to choose the best model, are the information criteria *AIC* (*Akaike Information Criterion*) and *BIC* (*Bayesian Information Criterion*), first proposed by Akaike in 1974, based in log-likelihood or in restricted likelihood function. Applying these criteria to compare models, one should choose the model that presents lower *AIC* or *BIC*. The two criteria are very similar, with a slight difference. The *BIC* criterion is more sensitive to the number of parameters included in the model, penalizing those which, with equal likelihood, use more parameters. However, when comparing models with different fixed components, the *AIC* and *BIC* criteria should be based on maximum likelihood estimation.

Linear mixed models can also be generalized to nonlinear models, considering a predict function of form (Davidian and Giltinan [2]):

$$\boldsymbol{y}_i = g(\boldsymbol{X_i\beta} + \boldsymbol{Z_i b_i}) + \boldsymbol{\epsilon_i},\ i=1,\ldots,M. \tag{2}$$

The parameter estimation is made using the likelihood function numerically optimized, since its complexity increases substantially.

For application of the mixed model in the SC simulation, we used the *R* (CRAN) statistical software, with the library *nlme* developed by Pinheiro and Bates [7], which provides the methods mentioned.

Fitting the usual linear regression model, it provides to be unsuitable for modeling data, showing large residuals and normal deviations. Data suggest a variation between scenarios that is not considered in usual regression model. On the other hand, using a linear regression model for each scenario is also not a good solution, as it is the opposite of the simple solution provided by the linear regression model (as it considers at least six parameters). Furthermore both solutions ignore dependencies within each group.

Mixed models are a good compromise between the two mentioned previous models, as they have a fixed part and a random part to model within-group correlations, without ignoring them and also respecting its shared characteristics.

Starting to fit a linear mixed model with the *intercept* and *simulation* variables, it was found that the *intercept* was the only significant covariate as well as its corresponding random effect, that is, we considered the model:

$$y_{ij} = (\beta_0 + b_{i0}) + \epsilon_{ij},\ i=1,\dots,6, j=1,\dots,n_i \quad (3)$$
$$b_{i0} \sim N(0, \sigma_0^2),\ \epsilon_{ij} \sim N(0, \sigma^2).$$

Getting:

$$\hat{\beta}_0 = 88.39\ (t\text{-value} = 40.76) \quad \hat{\sigma}_0 = 4.66 \ ; \hat{\sigma} = 13.97 \ ;$$
$$AIC = 1486.81;\ BIC = 1496.41.$$

Finding that the *intercept* is the most important variable, and random, could mean that the simulation model was successful, since it evidences a simulated situation on a regular basis and with common characteristics that only differ from scenario to scenario, as suggested by the random *intercept*.

However, the standardized residuals were in the interval *[-3.85, 1.24]* suggesting that the model should be optimized (figure 3).

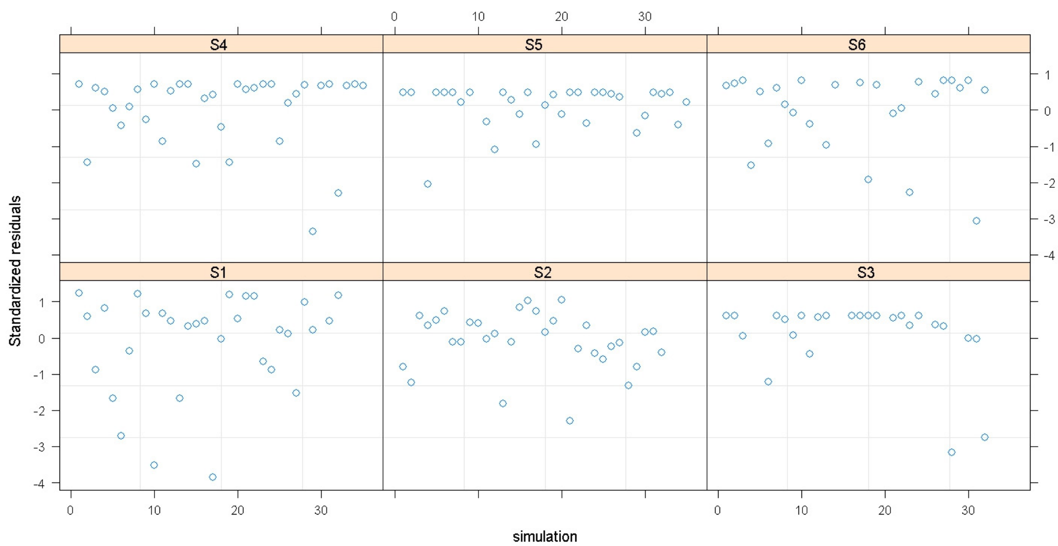

Fig. 3. Standardized residuals for the fitted linear mixed model.

Since the residuals are over *-2*, we shall introduce a variable modelling them in the model predictor, creating a factor variable with two levels: *1* if it is a outlier (outside the residual range *[-2,2])* and *0* if not. This *outlier* variable was then estimated in the model:

$$y_{ij} = (\beta_0 + b_{i0}) + \beta_1 outlier_{ij} + \epsilon_{ij},\ i=1,\ldots,6, \qquad (4)$$
$$b_{i0} \sim N(0, \sigma_0^2),\ \epsilon_{ij} \sim N(0, \sigma^2).$$

This model has a *AIC= 1346.64*, *BIC=1359.41*, and the *ANOVA* comparison between this model and the previous one, based on the maximum likelihood (since they have different fixed parts), shows that this latter is a better model.

Being *outlier* a significant variable also indicates that, once again, the simulation takes into account the real situation in a SC, in which there are frequent abrupt situations of high demand, or the opposite.

However, since the model only has the *intercept* and the *outlier* variables in the fixed part, it predicts values are poor since they don't show much variation between simulations (figure 4).

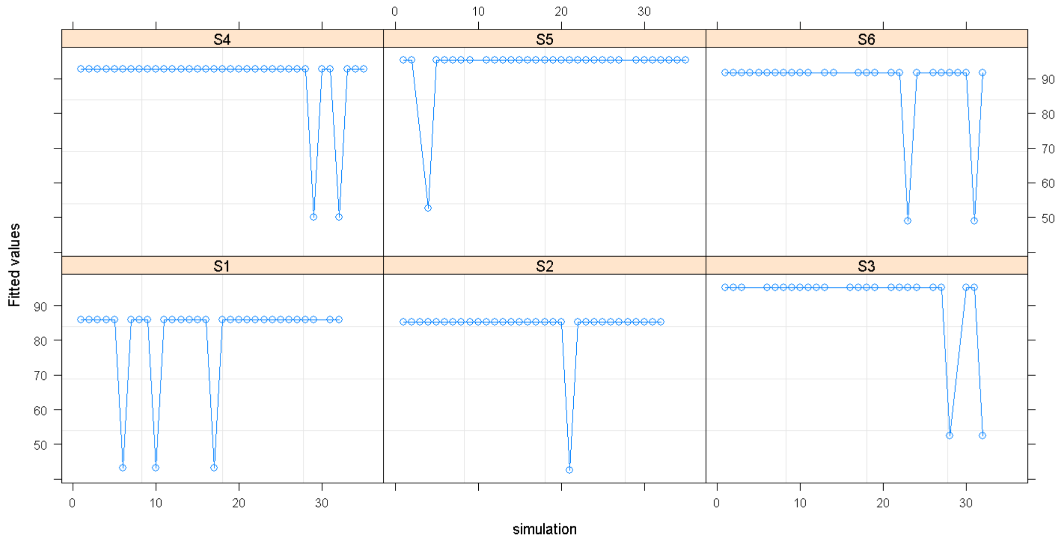

Fig. 4. Predicted values from the previous model.

Therefore it will be considered a nonlinear function $sin(w \times simulation) + cos(w \times simulation)$ in the predictor, where w is a parameter to be estimated, for better modeling the simulated observations. The model in now a nonlinear mixed model:

$$y_{ij} = (\beta_0 + b_{i0}) + \beta_1 outlier_{ij} + \sin(w_i \times sim_{ij}) + \cos(w_i \times sim_{ij}) + \epsilon_{ij},\ i=1,\ldots,6, \qquad (5)$$
$$b_{i0} \sim N(0, \sigma_0^2),\ \epsilon_{ij} \sim N(0, \sigma^2),$$

where sim_{ij} represents the j simulation from the i scenario. Observing the standard residuals of model (4), new observations were added as outliers in this adjustment (those whose standard residuals were again outside the range [-2,2]). The estimates obtained are:

$$\hat{\beta}_0 = 92.82\ (t - value = 60.07);\quad \hat{\beta}_1 = -34.99\ (t - value = -20.39);$$
$$\hat{w} = 0.95\ (t - value = 33.40);\ \hat{\sigma}_0 = 3.45\ ;\ \hat{\sigma} = 7.55;$$
$$AIC = 1274.23;\ BIC = 1290.25.$$

And the selected model predicted values are shown in the next figure.

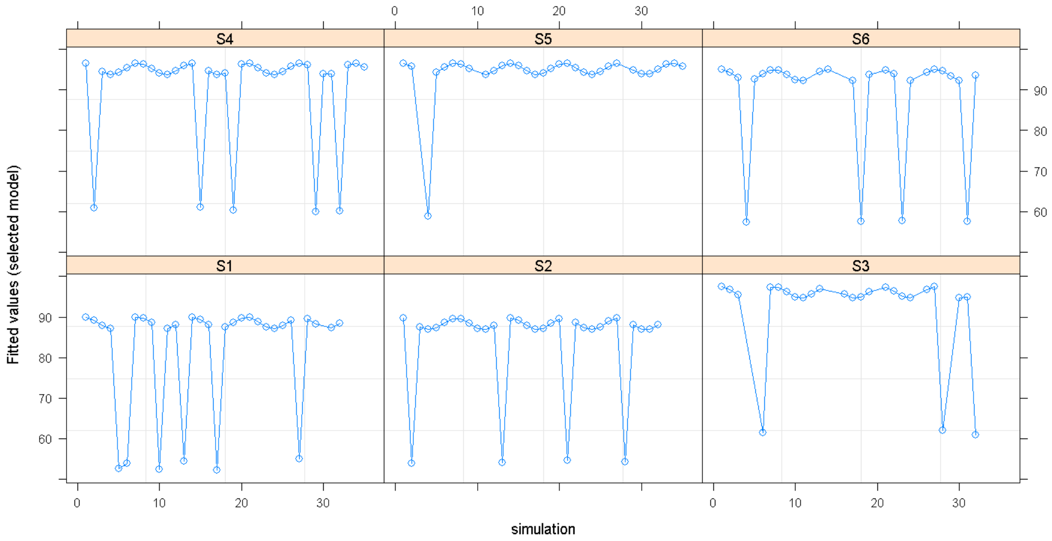

Fig. 5. Predicted values with the nonlinear mixed model.

The diagnosis of the model regarding the assumption of independence of within-groups errors ($\boldsymbol{\epsilon_i}$) and Normal distribution shows no major problems. The independence of $\boldsymbol{\epsilon_i}$ was confirmed with the empirical autocorrelation function (Pinheiro and Bates [7]) from the residuals of the model. And the estimates effects, by scenario, are:

Table 2: Estimates of the select mixed model.

Estimates	*Intercept* (fixed and random)	*Outlier* (fixed)	*w* (fixed)
Scenario 1	88.54	-34.99	0.95
Scenario 2	88.28	-34.99	0.95
Scenario 3	96.08	-34.99	0.95
Scenario 4	95.18	-34.99	0.95
Scenario 5	95.22	-34.99	0.95
Scenario 6	93.65	-34.99	0.95

As we can see, scenario 1 and 2 are the worst scenarios for SC coordination. Scenario 3 is the best, followed by scenario 4 and scenario 5. From this analysis, it can be concluded that the best option to coordinate the SC is to adopt the proposed solution for scenario 3.

4. Conclusions

In this paper, data from six scenarios were considered, previously obtained by a pharmaceutical simulation case study. These coordination scenarios represent different solutions proposed by Vieira [9] and the aim was to find the solution that statistically indicates which is the best coordination scenario to use in the SC case study. For this purpose, statistical mixed models were applied, because they are a flexible way to model within-groups dependence.

The model that best fits the situation includes a random component with an *intercept*, and a fixed component being constituted by an *intercept*, a factor variable *outlier* and a nonlinear trigonometric function to better reflect the variability of observations. Thus, a flexible and parsimonious model was found with few parameters, taking into account the common scenarios features while also taking into account the differences between them.

The most important covariates are the random *intercept* and the *outlier* factor. The fact that o*utlier* is a significant variable indicates that the simulation takes into account the real situation in a SC, where outliers often occur. The *intercept* fixed and random means that the simulation model was successful, since it evidences a simulated situation on a regular basis and with common characteristics that only differ from scenario to scenario, as suggested by the random *intercept*. Its estimates allow us to conclude that scenario 3 is the most suitable for the coordination of the SC, followed by scenarios 5 and 4.

References

[1] Christopher M. Logistics and supply chain management. Financial Times Prentic Hall; 2005.

[2] Davidian M. and Giltinan DM. Nonlinear Models for Repeated Measurement Data: An Overview and Update. JABES, 8, p.387-419; 2003.

[3] Hung WY, Samsatli NJ, and Shah N. Object-oriented dynamic supply-chain modelling incorporated with production scheduling. European Journal of Operational Research, 169 (3); 2006. p.1064–1076.

[4] Laird NM, Ware JH and Fitzmaurice GM. Applied Longitudinal Analysis. 2rd ed. Wiley; 2011.

[5] Lee HL and Whang S. Information Sharing in a Supply Chain. International Journal of Manufacturing Technology and Management 2000, 1, p. 79–93.

[6] Mentzer JT, DeWitt W, Keebler JS, Min S, Nix NW, Smith CD, and Zacharia ZG. Defining Supply Chain Management. Journal of Business Logistics, 22 (2); 2001. p. 1–25.

[7] Pinheiro JC and Bates DM. Mixed-Effects Models in S and S-PLUS. 2rd ed. Springer-Verlag; 2009.

[8] Seco A. Modelação de Dados Agrupados nas Ciências do Ambiente. PhD Thesis in Mathematics and Statistics ISA, UTL, Lisbon; 2003.

[9] Vieira CM, Barbosa-Povoa AP, and Martinho C. Supply Chain Coordination Analysis through a Multi-Agent System. PhD Thesis in Engineering and Management. IST, Lisbon; 2013.

[10] Wooldridge M, Jennings NR, and Kinny D. The Gaia Methodology for Agent-oriented Analysis and Design. Autonomous Agents and Multi-Agent Systems, 3 (3); 2000. p. 285–312.