Anna Bachurina, Moscow, Russian Federation, 2013

Model validation: matching data and structure to behavior through partial model calibration in Group Model Building setting

Abstract

This paper presents a framework for model validation and calibration, while employing ERP systems' data and expert knowledge of Group Model Building sessions' participants. This framework was applied in the project for a staffing company, and I will use mini case studies from it to illustrate my approach. In order to build this model various sources of knowledge were used: statistical data, market reports, knowledge of company representatives obtained through interviews and group model building sessions, and system data stored in various IT systems. Combining those sources of information it is possible not only to get more knowledge about the system, but also validate and calibrate it. Based on this project experience and related literature some practical recommendations were developed and their applications demonstrated.

Introduction

Nowadays enterprises are overloaded with data, which is collected and stored in databases. More and more information is available for decision support. However, data analysis is not only time-consuming and demanding to computational resources, but also sometimes it is difficult to define which data is worth to collect and analyze, and which methods are applicable for particular analytical purposes. Since the arrays of data stored in Enterprise Resource Planning (ERP) systems are huge, people get lost in it and have problem to recall not only the figures, but also trends: is this value increasing or decreasing. While synthesizing data with expert knowledge, it can make more sense and consequently can lead to a better understanding of the system.

Group model building has emerged as a methodology for not only gathering data from people, but also capturing their interpretations of the causality present in the system (Vennix, 1995). Moreover, group model building allows for overcoming the human's tendency of seeing pieces of the system instead of whole. There are several reasons for that. One of them is that "many people are not trained in systems thinking. Another is the limited information processing capacity of the human mind" (Vennix, 1995, p. 41).

Information necessary to build System dynamics models can be of a qualitative and quantitative nature, and can be obtained from three different sources: mental, written and numerical databases (Forrester, 1991).

Although qualitative data is recognized as the main source of knowledge about system structure and governing policies (Ford and Sterman, 1998), "ignoring numerical data or failing to use statistical tools when appropriate is sloppy and lazy" (Sterman, 2002). Moreover, avoiding usage of numerical values the chance that "the insights you derive from your model will be wrong or harmful to the client" increases (Sterman, 2002).

According to the definition given by Benson & Davis (2008, p. 8) data are "raw, unprocessed streams of facts" and data turn into information after it "processed and shaped in a meaningful form useful for a person or computer". However, he admits that "raw data is a relative term as data processing may have a number of stages, so the output of one processing stage can be considered raw data for the next" (Benson & Davis, 2008, p. 8).

The sequence of the stages through which information and data flows is presented on figure 1. The arrays of data stored on ERP system are organized in OLAP (on-line analytical processing) cubes and need to be extracted to feed the system dynamics model. However this data needs processing – data mining. The output of this process (values of parameters in the model) is used as an input for the simulation.

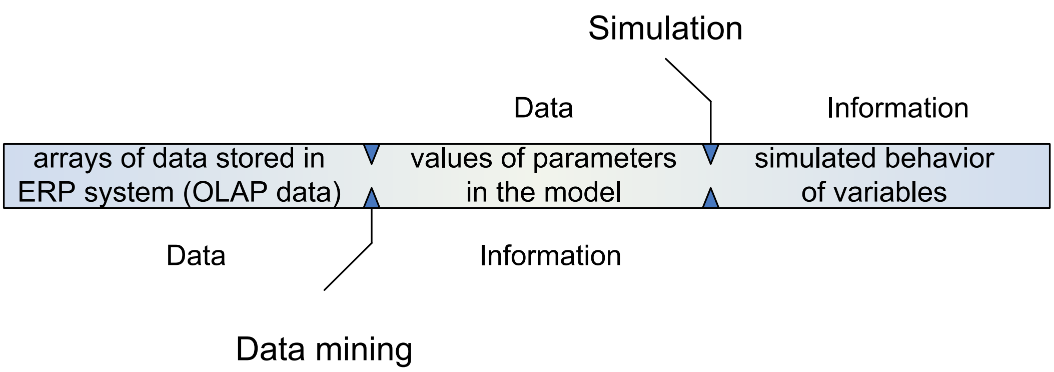

Figure 1. Data and information in modeling process using ERP data.

Exported spreadsheets contain a lot of data, which might need to be reorganized for its further usage. "Preparation of the data set involves selection of the data sources, integration of heterogeneous data, cleaning the data from errors, assessing noise, dealing with missing values etc." (Mannila, p. 2). When data entries are processed "some utility value or meaning is added to raw data to transform it into business information" (Benson & Davis, 2008, p. 8).

Thus using expert knowledge we can get access to right data sources, understand what stands behind each number and transform it in a way that this data can be used as an input for System Dynamics model.

Fayyad et al. (1996) distinguish two knowledge discovery goals: verification and discovery. “With verification, the system is limited to verifying the user’s hypothesis; with discovery the system autonomously finds new patterns”.

Both discovery goals can be achieved when matching data and structure to behavior through partial model calibration process in Group Model Building setting.

Framework description

The first step of the process is building the preliminary model during group model building session(s). The involvement of experts from the areas related to the problem ensures that, even though it is difficult to capture detail complexity during the first session, the main causalities would be captured and mapped as a holistic system. In addition, educated guesses on parameters of the system and behaviors over time need are to be made.

Second step is specification of data requirements, which is based on the variables captured in Group Model Building session. It is important to notice that not all the needed data can be available, due to the fact that it was not collected in the past.

Third step is division of the model into blocks – sub models. Those blocks are interconnected and the criterion for defining each block is availability of input and output historical data for each of those. In the case study the model is divided into seven interconnected sub models. Those sub models and the connections between them are shown on figure 2. All blocks are connected through certain variables, which are written above the arrows. Reference modes are available for each of those values. Availability of those reference modes allowed to calibrate and validate every sub model in order to ensure that it behavior corresponds to the reference mode when being fed by reference data produced by another sub model. The process of calibration will be described in the next chapter.

It is worth noticing that the more historical data are available, the more subsections can be isolated and calibrated individually, the more solid those subsectors are. Walker and Wakeland (2011, p.7) point out that “the technique of building solid submodels is probably one of the most important keys” to getting a model working properly. “Given the size of the model, discrepancies in the outputs could frequently have been caused by a number of different parts of the model" (Walker and Wakeland, 2011, p. 7). The more subsectors are isolated and calibrated, the more "stability points " in the system.

The sub models "Part-time labor market" and "Flexworkers working" will be used to illustrate my approach.

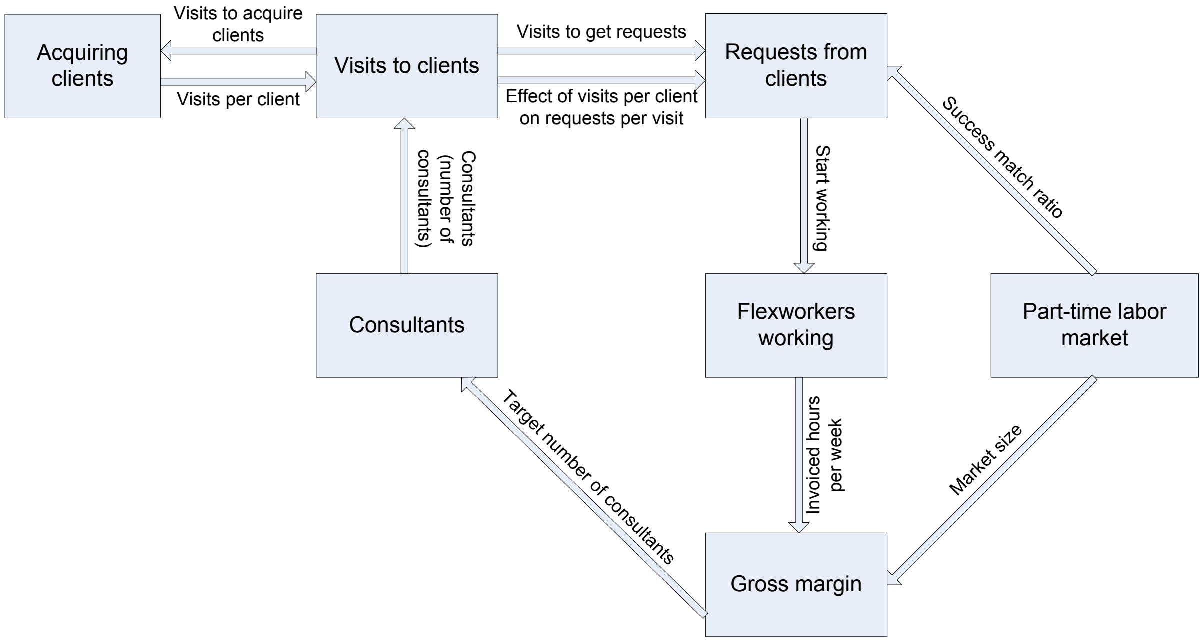

Figure 2. Sub sectors of the model

Fourth step is verification of existing and estimation of missing parameters in each sub model. It is made while employing data mining or calibration process, which is described in detail in the next chapter.

Fifth step is presentation of the refined model, which contains data and parameter estimations from experts and data mining/calibration process. The value of each parameter should fit to the data stored in the system (if available or can be inferred from system data), outcome of calibration process (assuming that the structure of calibrated system is right) and mental models of the experts.

If there is no correspondence between those values, several issues can arise. Depending on the issue, different troubleshooting approaches are applied. The matrix for those classes of problems is presented on figure 3. N/A means this value is not available, A, B and C are symbolic codes for different values.

Parameter estimation based on data	A	N/A	A	A	A	A

Parameter estimation based on calibration process outcome	B	A	N/A	B	B	A
Parameter estimation based on experts' educated guesses	C	B	B	A	B	B
Troubleshooting method	1	2	3	4	5	6

Figure 3. Troubleshooting matrix

1. All three values are very different.

The sub model needs to be revised totally: data needed for input and output, data needed for parameter estimation, causal links in the model and equations.

2. The parameter estimated by experts doesn't correspond to the value received in the calibration process. No data is not available to estimate this parameter

The experts need to re-evaluate the new value of the parameter, taking into account the value from calibration process. If experts accept new value, this is a contribution to the learning about the system. If the value is not accepted, the causal structure of the sub system is revised and data to feed this substructure is presented and verified.

3. The parameter estimated by experts doesn't correspond to the estimation based on data. No calibration is made for this parameter.

The experts are presented with data used for estimation of the parameter and explanation of how this parameter was estimated. If expert confirm it, the new value it is accepted. If data or estimation methods are not acceptable, the parameter is revised.

4. The value estimated by experts corresponds to data, but the result of calibration is different.

The causal structure of the sub model need to be re-defined, the input and output data verified.

5. The outcome of calibration corresponds to parameter estimations by experts. However, data doesn't confirm it.

It means that the data for parameter estimation is not appropriate. For the model building purpose it can be ignored. However, this issue can be taken into account by a company to investigate why the data is wrong.

6. Data supports the outcome of calibration, but experts name different value.

This is the case when experts usually accept a new parameter value, because of supporting data.

The described approach to model validation corresponds to earlier mentioned knowledge discovery goals: discovery and verification. All three values were first “discovered” and afterwards verified using one another, what brings more confidence into the model.

The proposed framework involves an iterative approach until each block of the model is calibrated and validated. Next steps would involve overall testing of the model and using it to solve problems it is designed for.

Partial model calibration illustrated

The partial model calibration approach will be explained using a case study from the project conducted for a staffing company. Since its performance defined both by internal policies and economic situation, the impact of economic growth on part time labor market needed to be investigated. The development of number of hours by agency workers and GDP growth is shown on figure 4.

It is worth noticing that the idea that the growth in the number of hours (not the number of workers or the hourly wage) is directly affected by GDP growth was taken from a group model building session and supported by reports on the industry.

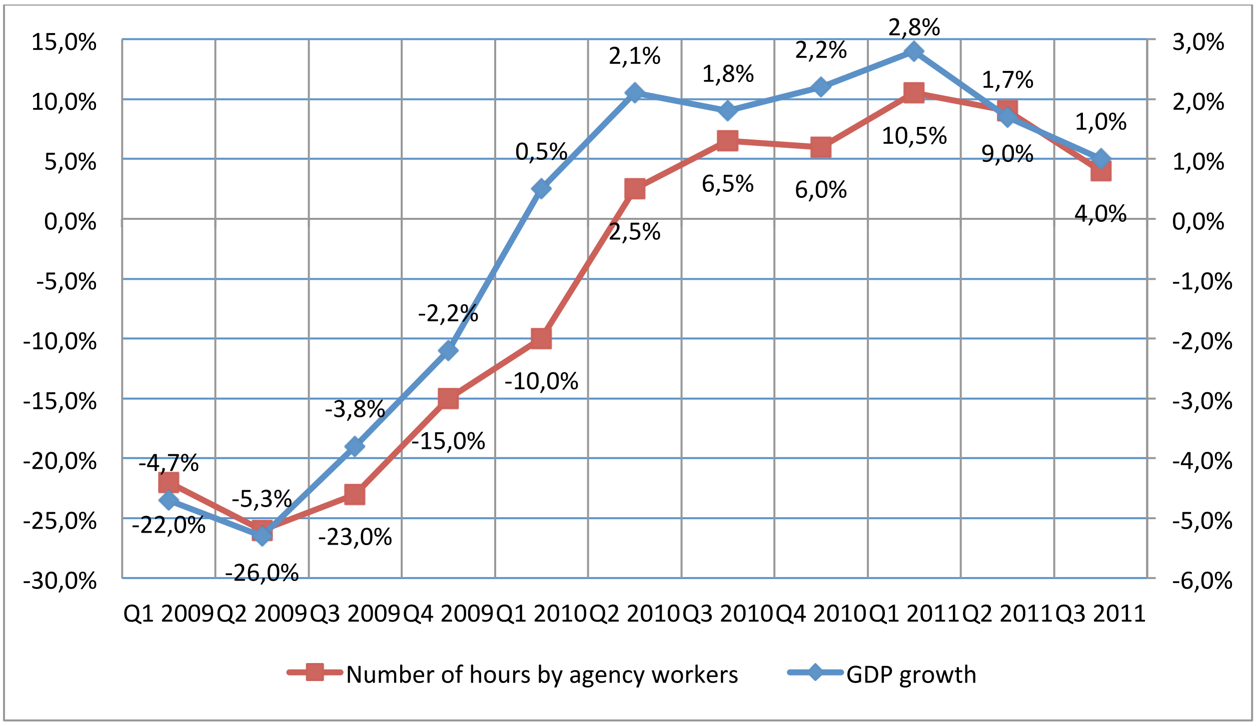

Figure 4. GDP growth and Number of hours by agency workers

We can see a delay between those variables, and a strong interrelationship. However, in order to put this into equation we need to estimate those parameters.

The structure of the sub model is shown on figure 5. Below each of those sectors would be explained in more detail. The variables which are not linked to the model and contain _RM in their name is reference data.

The proposed framework includes three elements: calibration sub-model, core model and parameters to estimate.

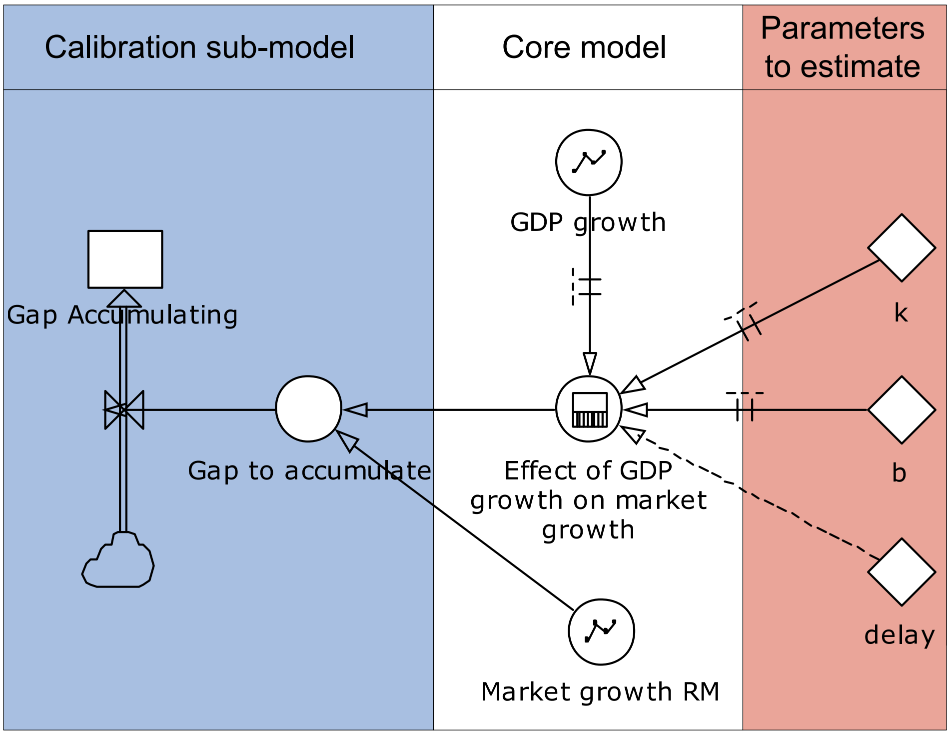

Figure 5. Sub model structure. GDP growth and Number of hours by agency workers

The core model contains elements, which are included into the modeled system and knowledge about those elements. This knowledge is derived from participants during GMB sessions and interviews, and from papers published by other experts in the field. In this case the knowledge about the dependency of part time labor market growth on GDP growth. Another source about knowledge for Core model is the reference data which can be obtained from statistical databases, market reports and system data about the historical performance of the company stored in the various repositories for information (ERP - Enterprise Resource Planning systems, BI – Business Intelligence solutions, etc). In this case data about GDP is exported from cbs.nl (statistics Netherlands agency) online database, and market growth data is obtained from the reports by the Federation of Private Employment Agencies (ABU).

Parameters to estimate are variables for which experts can define a certain range, and which need to be established as an output of calibration process.

The calibration sub model is designed in order to obtain historical data fit with simulated behavior through search for optimal set of parameters ensuring that fit.

Calibration is done thorough minimizing accumulated squares of the gap between simulated behavior and reference data. The tool, which was used in order to execute calibration process, is Powersim Solver - built in optimization tool using evolutionary search algorithm. In order to launch it “decisions” and “objectives” are to be set.

“Decision” is a set of parameters, which should be found as an output of calibration. The range within which those values can be set according to the knowledge obtained while interviews and GMB sessions.

“Objective” is defined as an accumulated absolute value of the gap between historical data and simulated behavior.

In case when there is high level of confidence in some parameters, they are not included into optimization task. “The most serious difficulty with a large number of calibration parameters, however, is the increased difficulty in detecting formulation errors" (Oliva, 2001).

Basically this technique is very similar to building regression models, but also captures delay (and not only captures, but also calculates the value of this delay) and can capture feedbacks in some complex sub models (nest case study).

The results of the calibration process are shown below (figure 6).

The delay between growth of GDP and growth of part-time labor market is 2.07 months. This value was presented and confirmed by participants during a group model building session. The period of two month includes the period within which companies realize the increased need for flexible labor, place requests for flexworkers and this request is fulfilled after a period of searching for candidates suitable for this temporary job.

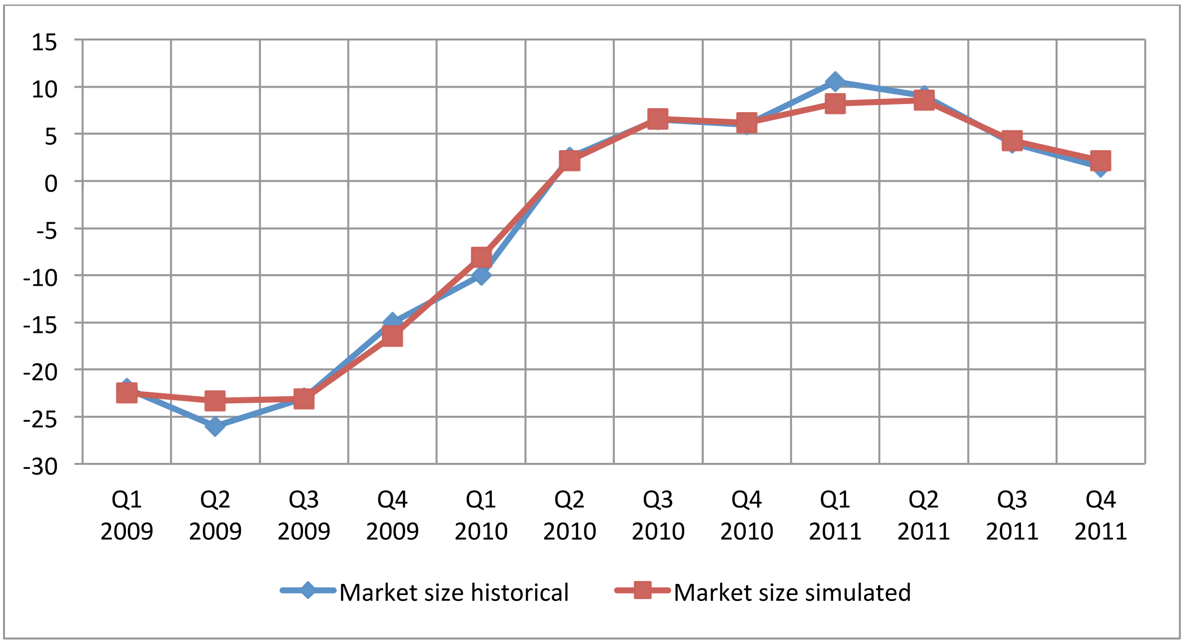

Figure 6. Market size: historical and simulated data

The results of the simulation show very good fit and the hypothesis about dependence of part time marker size of GDP growth is confirmed.

Case study II

One more example of using partial model calibration method is provided in this section. The sub model “Flexworkers working” is shown on figure 7. As we can see the parameter is estimate is “length of the contract”.

Even though participants provided me with an estimate of average duration of the contract, when putting this number into the model, the behavior did not make sense. The ERP data did not explicitly capture delays between the moment a flexworker starts and stops working.

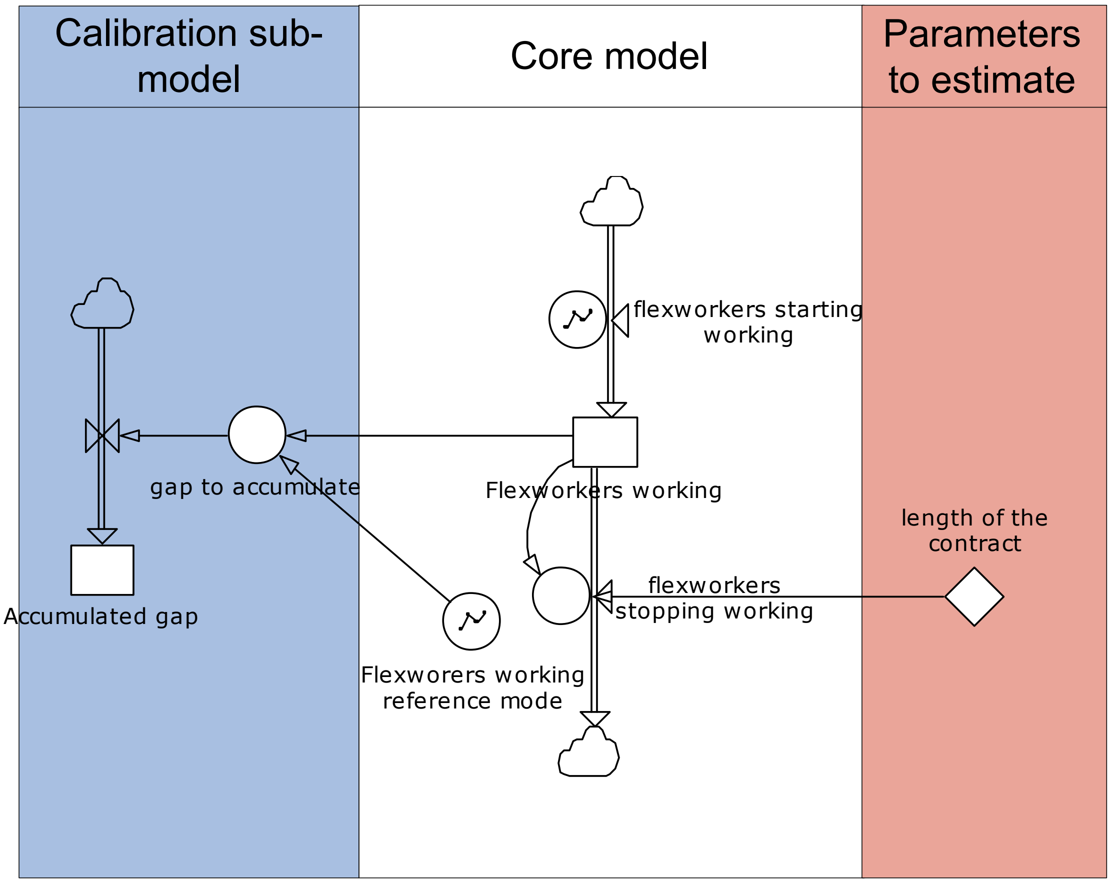

Figure 7. Flexworkers working

The layout of the data exported to a spreadsheet is presented on figure 8. The number of flexworkers working for a company is registered every quarter and the dynamics (how many people were hired and quit) can't be derived from it. Consequently, if the length of the contract is high, there are less hires and quitting, if duration is low, there is higher turnover.

Time period	Unit name	N of consultants	N of flexworkers working
Q1 2010	Unit A	2	20
Q2 2010	Unit A	2	50
Q3 2010	Unit A	1	60
Q4 2010	Unit A	0	10

Figure 8. An ambivalence of data in the excel sheet

The results of calibration in order to estimate the length of the contract are shown on figures 9 (if the duration of contract is high) and 10 (if the duration of contract is low).

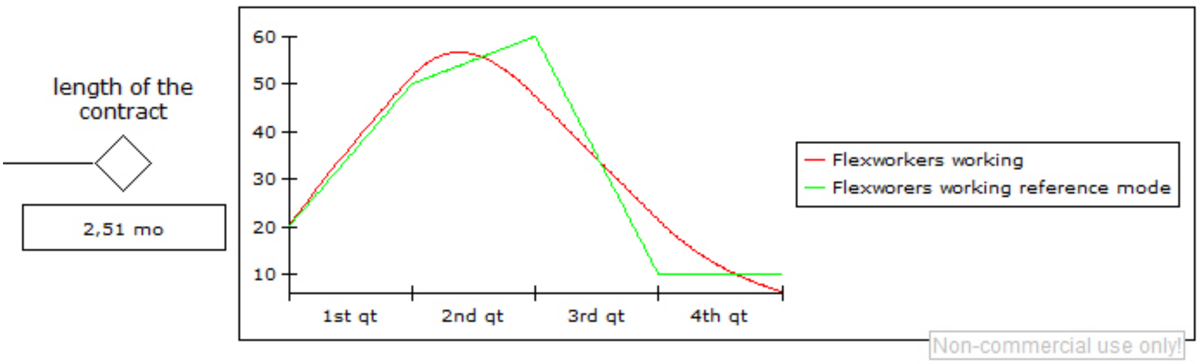

Figure 9. High length of the contract. Flexworkers working

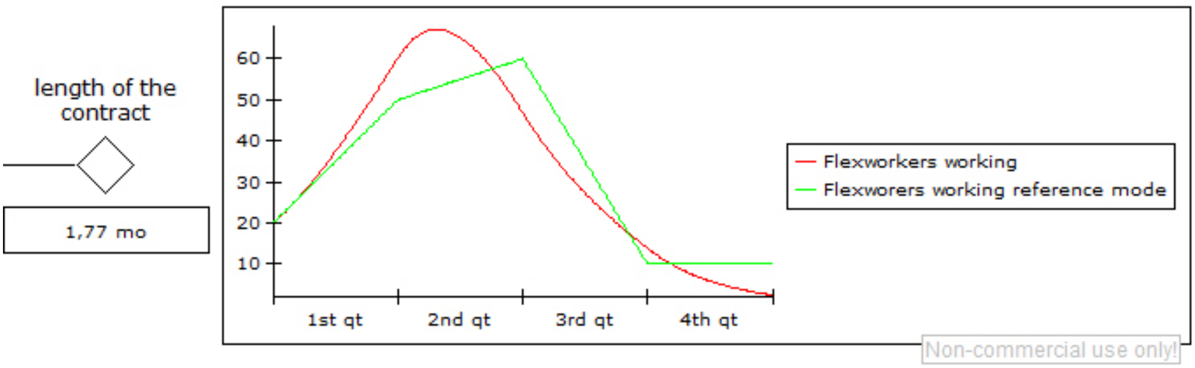

Figure 10. Low duration of the contract. Flexworkers working.

The results derived from this sub model were reported during a group model session and the value lying within “shortest” and “longest” contract durations was confirmed to be included into the model.

This case study has shown the possibility not only to put existing data into the system dynamics model, but also derive missing parameters based on existing information, when application of classical tools is problematic because of existing delays and feedback loops.

Conclusions

This paper described validation process while combining historical data, expert knowledge on causalities about the system and experts’ educated guesses on the parameter values. The framework involves an iterative approach and is applied until all parameters are established in order to ensure that structure matches the data.

For that purpose the calibration process based on the optimization of the error between reference mode and simulated behavior was applied. Oliva (2001) point out that "This process, however, assumes that the model structure (equations) is known, and that all uncertainty resides in the parameter values." (p.6). Thus if the structure designed based on mental models of the participants is right, and the data is valid, the missing parameter derived out of that sub mode cannot be wrong. If it is wrong, the data or causal structure needs to be revised.

Fayyad et al. (1996) state that "if one searches long enough in any data set (even randomly generated data) one can find patterns that appear to be statistically significant, but in fact, they are not" (Fayyad et al. 1996, p. 4). In addition, patterns can be statistically significant, but practically to finding a practical use for those patterns can be problematic. Sterman (2000) gives an example of correlation between ice cream sales and murders which both peak on summer. However, there is no practical application of this pattern. Thus, the confirmation of patterns needs to be confirmed by experts in the field.

The synthesizing data with participant knowledge proved to give added value on different stages of the process. The formulation of problem should be supported by reference mode. The hypothesis testing combines compliance to the mental models of participants with the data verification. The analysis stage involves analyzing parameters derived from GMB sessions and data. Thus those sources of knowledge not only complimentary, some knowledge can be generated only when both sources are available. Combination of those allows to reach a higher level of confidence in the model.

References

Ariza C., Graham A., 2002. Quick and rigorous, strategic and participative: 12 ways to improve on the expected tradeoffs. Proceedings of the System Dynamics conference in 2002.

Barlas Y. 1996. Formal aspects of model validity and validation in system dynamics/ System dynamics review. Vol. 12-3. P. 183-210

Benson V., Davis K., 2008. Business Information management. Ventus Publishing Aps

Berkhout E., Berg E., 2010. Bridging the gap: international database on employment and adaptable labor. Seo economic research

Better M., Glover F., Laguna M. 2007. Advances in analytics: integrating dynamics data mining with simulation optimization. IBM Journal of Research and development 5 feb 2007

Burmeister L., Goeken M. 2008. Combining system dynamics and multidimensional modeling – a metamodel based approach. Proceedings of the Fourteenth Americas Conference on Information Systems, Toronto, ON, Canada August 14th-17th 2008

CIETT report, 2002. Rationale of agency work: European labor suppliers and demanders' motives to engage in agency work .

CIETT, BCG report, 2011. Adapting to change

Coe N.M., Jones K., Ward K. 2010. The business of temporary staffing: a developing research agenda. Geography Compass 4/8

Fayyad U., Piatetsky-Shapiro G., Smyth P., 1996. From data mining to knowledge discovery in databases. AI magazine. Fall 1996.

Ford D.N., Sterman J.D., 1998. Expert knowledge elicitation to improve formal and mental models reference1998. System Dynamics Review 1998

Forrester J.W., 1991. System dynamics and the lessons of 35 years. A chapter for the Systemic Basis of Policy Making in the 1990s

Gonzalez F. 2001. Integrating Dynamic Simulations with an OLAP data. Proceedings of the 19th System Dynamics Conference. Atlanta: System Dynamics Society

Hauke U., 2001. Powersim's business modeling and simulation tools are built in to SAP SEM. Sap Insider October December 2001

IBM Cognos Business Insight Advanced: User Guide

Lee T., Zagonel A., Andersen D. F., Rohrbaugh J.W., Richardson G.P. 1998. A judgement approach of estimating parameter in group model building: a case study of social welfare reform at dutchless county. Proceedings of the System dynamics conference 1998.

Lyneis J.M. (1980), Corporate Planning and Policy Design: A System Dynamics Approach, MIT Press: Cambridge, MA.

Mannila H., 1996. Data mining: machine learning, statistics and databases. Scientific and Statistical Database Systems, 1996. Proceedings., Eighth International Conference on

Medina-Borja A., Pasupathy K.S., 2007. Uncovering complex relationships in system dynamics modeling: exploring use of cart, chaid and sem.

OECD Employment outlook 2004

Oliva R., Model calibration as a testing strategy for system dynamics models, European Journal on Operational Research Vol 151, p.552 - 568

Pedersen T.B., Jensen C.S., 2001. Multidimensional database technology. Distributed Systems Online (IEEE): 40–46.

Philips K., Eamets R. 2007. Approaches to flexicurity: EU models. European Foundation for the Improvement of Living and Working Conditions

Ployhart R.E., 2006. Staffing in 21st century: new challenges and strategic opportunities. Journal of management Vol. 32, p. 868-897

Rashid M.A., Hossain L., Patrick J.D., 2002. The evolution of ERP systems: a historical perspective. Idea Group Publishing

Rouwette, E.A.J.A. & Vennix, J.A.M. (2010). Improving operations management by synthesizing participant knowledge and system data. Proceedings System Dynamics Conference Seoul. Korea: System Dynamics.2010

Schmid G., Modracj S., 2008. Employment Dynamics in Germany: lessons to be learned from the Hartz reforms. Proceedings of the conference of the GTZ Sector Network for Sustainable Economic Development in MENA: “Employment: Challenge for Economic Development in the MENA Region” in Istanbul, 7 November 2007.

Shah S., Horne. A., Capella J., 2012. Why good data won’t guarantee good decisions. Harvard Business Review April 2012 Torres M.D.S., Lechon R.F. A dynamic model of labor market. 1995. Proceedings of the System Dynamics Conference.

Sterman J.D., 2002. All models are wrong: reflections on becoming a systems scientist. Jay Wright Forrester Prize Lecture, 2002

Vennix, J. (1996). Group model building. Facilitating team learning using system dynamics. Chichester: Wiley.

Walker R., Wakeland W., 2011. Calibration of Complex System Dynamics models: a practitioner’s report. Proceedings of System Dynamics conference 2011