Value of Smart Data for Supply Chain Decisions in a Data Rich, Uncertain World

Ganesh Sankaran[1,2], Martin Knahl[1], Guido Siestrup[1], Ismini Vasileiou[2]

[1] Hochschule Furtwangen University, Robert-Gerwig-Platz 1, 78120 Furtwangen, Germany
[2] University of Plymouth, Drake Circus, Plymouth, Devon PL4 8AA, United Kingdom

Abstract. Data-driven decisions are becoming increasingly relevant for supply chains as traditional paradigms are being replaced with concepts and approaches more suited for the advent of big data. However, the prevailing consensus is that companies are struggling to cope with an overabundance of data, which presents the following pertinent question: how to efficiently analyze data applying filters of relevance and insightfulness to make effective decisions? There is currently a lack of research focus on providing quantitative tools to do such analyses. This paper, besides offering thoughts on decision-making uncertainty in a digital supply chain context, describes an approach to address the research gap. The approach (which involves developing a quantitative model) is further elucidated by utilizing an example in the agricultural supply chain that illustrates how value of data can be quantified by measuring the performance impact of insights delivered using uncertainty reduction as the leverage.

Keywords: Supply Chain Management, Digital Technologies, Big Data Analytics, Uncertainty, Data Driven Decision Making, Value of Data.

1 Decision Making Under Uncertainty in Digital Supply Chains

In designing and analyzing supply chain processes, the theoretical frame of “hierarchy of decisions” has often been used [1]. This view, that segments processes according to scope and significance, into strategic, tactical and operational, acknowledges the pivotal role of decision making in Supply Chain Management (SCM).

The decision-making process in SCM, as also in the general sense, is crucially about choosing a course of action by assessing alternatives and settling on one “that is most likely to result in attaining the objective” [2]. In this way, the efficacy of the process hinges heavily on the ability to parse and understand uncertainty in the states of the world associated with the alternatives. This notion of uncertainty is at the heart of the Organizational Information Processing Theory (OIPT) that posits uncertainty as the disparity between information processing need and corresponding capacity, and links it to process and organizational performance [3].

In the current environment characterized by supply chains readily embracing digital technologies and transforming themselves into Digital Supply Chains (DSC), decision making under uncertainty presents an apparent contradiction: the abundance of data afforded by digital technologies would lead one to expect DSCs to be exploiting this opportunity to achieve parity in the OIPT sense (between information processing need and attendant capacity) and drive improved performance resulting in a higher utilization of data. However, various studies show that most digital data that is captured is not utilized [4] and less than 1% of unstructured data is analyzed at all [5]. Research in the areas of digital transformation and Value of Information (VOI) offer up some clues to clarify this contradiction.

Digital transformation is about innovating new business models and ways of value creation and capture, focusing on the dual outcomes of customer engagement and integrated digitized solutions. It is perhaps better understood by contrasting with a related term – digitization, which on the other hand, is a narrower technology-centric view [6]. Not surprisingly, supply chains that focus on transformation perform significantly better than peers [6]. On the other hand, a lack of transformation focus leads to unmet expectations and such companies are apt to complain, as have six out of 10 respondents in this survey of 3000 executives, of having more data than they can use effectively [7].

A related line of research inquiry concerns VOI in a big data context. Research into Information Systems (IS) following IS economics tradition highlight the lack of tools to quantify data and the need to address the challenge of "finding a way to quantify the value of information that considers both insightfulness and risks" [8]. The two lines of inquiry are linked, and the convergence is in the fact that supply chains that are transformation focused are more likely to want to justify investments and therefore also want to quantify value of data - and this is where this research aims to contribute.

2 Approaches for Measuring Business Value of Data

2.1 State-of-the-Art

Using a resource-based view, which holds that heterogeneity of organizational resources is a source of value (as it differentiates a firm from competition), Melville et al. [9] argue for consideration of competition and environmental factors to measure value of data as they are seen to impact value. Higher the level of competition or industry concentration, higher is the marginal product and, conversely, lack of competition creates slack resources leading to lower productivity [10]. Environmental factors or external focus, on the other hand, is seen to enhance performance as timely and accurate information regarding a firm's external environment are preconditions for agility [11].

Besides several empirical studies that adopt a general view on the impact of data on value and emphasize the link between data-driven decision making and firm output and productivity (see [12, 13] for representative examples), there are also several studies on particular problem instances. Ketzenberg et al. [14] assessed VOI in the presence of uncertainty around demand, return, and product recovery delivering a key insight that greater the uncertainty, greater is the VOI. Dunke and Nickel [15] incorporated forward-looking information in supply chain planning and proposed an optimization model that utilizes preview of future information with help of lookahead devices (e.g. sensors) to transform an uncertain future into a certain one.

2.2 Need for Further Research

The discussion above points to a wealth of empirical studies and models for specific problems. However, a general-purpose quantitative model with a normative character (elaborated in 2.3) is lacking. In a review of 117 articles on the topic of research contributions in this area, Viet et al. [16] had found that, in a supply chain decisions' context, there is disproportionate attention being paid to inventory whilst other areas have received insufficient attention. They also report that the impact of new and innovative data sources (e.g. sensor data) remains under-explored. In laying out a research agenda for future information systems research, Abbasi et al. [8] call for research on the "value of various data sources and channels in terms of quality of insights, enabling new capabilities, and quantifiable business value."

2.3 Model Conceptualization

Before describing the proposed model, it is instructive to go over key model attributes that were considered as prerequisites: (1) Quantitative: the overarching question calls for the ability to measure the incremental value of insights from digital data. This necessitates a quantitative-based model that yields a numerical solution. (2) Predictive: The model must emulate a decision-making process where the performance potential of data-driven insights can be studied. This requires the model to embody predictive or simulative capability. (3) Relevant: Zadeh's principle of incompatibility holds that complexity makes relevance and precision impossible to obtain simultaneously [17]. Therefore, the model needs to be built on a framework that lends itself to strike the right balance. From a performance measurement perspective, it needs to be inclusive (one of the key characteristics of a good performance measurement framework [18]) and not predisposed to any specific supply chain strategy. For instance, both cost (primary focus for efficient supply chains) and agility (primary focus for responsive supply chains) measures need to be supported. (4) Usable: as the key question being addressed most interests supply chain managers, the model should, despite its quantitative rigor, include a graphical component for the decision-making process to be analyzed visually as well.

The proposed model is grounded in the Approximate Dynamic Programming (ADP) methodology [19] (also called reinforcement learning). It is an active field of research that has a long history owing to its evolution from work done in optimal control theory and stochastic approximation (dynamic programming and Markov decision processes). ADP's choice as the model's underpinning is due to its suitability vis-à-vis prerequisites set forth earlier and its effectiveness in addressing the class of problems typical of the supply chain problem domain. One way to justify this claim is by noting the subcomponents of ADP and highlighting structural similarities between ADP and Supply Chain (SC) problems. ADP problem formulation consists of policy, reward, value and model environment. The solution involves an appropriate choice of *policy*, which is a set of endogenous controllable variables (e.g. reorder point in SC) in the face of uncertainty expressed by the *model environment* (exogenous information, for e.g., customer demand in SC) to maximize cumulative *rewards* or *value* (e.g. global perspective in SC). The approximate nature of ADP allows problems involving large state-spaces (typical of SC) to be solved by using an approximation architecture. The approximation architecture or the learning element allows better policies to be adopted as the system learns to interpret the uncertain environment better and develops a more accurate picture of the (delayed) consequence of actions on value. For the proposed model, this last aspect is crucial to modelling the recalibration of uncertainty due to infusion of digital data. [20]

The model incorporates formalisms to represent key elements of uncertainty and digital data. For this research, uncertainty is viewed as an empirical quantity [21] that can be modelled as a probability distribution. Furthermore, a Bayesian view of probability is adopted (other view being frequentist) [21], which is suitable in this problem-context of decision-making where beliefs about states of the world are conditioned on all available information. Quantification of uncertainty is a relatively untapped aspect in stochastic optimization literature [22] but will be an essential component in the model as it impacts policy selection and consequently its predictive ability. In the case of digital data, a semantic model (for example, based on W3C SSN ontology [23]) is adopted that provides similar modelling rigor. Finally, for model visualization, System Dynamics (SD) approach is the primary candidate [24]. SD provides an intuitive representation of causal relationships between variables and their impact on performance.

2.4 Illustration of Model Aspects: Example in the Agricultural Supply Chain

The example pertains to the production and sales of seeds that starts with the production stage (that involves sowing, growing, harvesting, treatment and packaging) and culminates in the sales of seeds to farmers. The problem of estimating yield is the focus of the example and it helps elicit the salient model features.

Once sales projections are made, production is planned assuming a certain yield (using factors like crop physiology). However, this is at best a noisy or imprecise estimate and the reality at harvest time tends to vary widely from projections. One key implication is the planning of treatment and packaging capacity, which is often a bottleneck. If the capacity planned is insufficient, it leads to lost sales and higher than required capacity leads to poor utilization and impinges on profits. However, advances in digital technologies provide the ability to use sensors and the like, which act as lookahead mechanisms and can provide advance insights during the lengthy sow-grow-harvest cycle, which can help revise noisy prior estimates with updated, sharper posterior estimates. The dynamics of interaction are presented in **Fig. 1**. As can be seen from the illustration, relevant sensor data (e.g. weather, water content) that are predictors of yield when captured can be utilized to revise estimates and perform contingency planning in the form of organizing additional subcontracting capacity or shaping demand (promotions) to better match demand and supply. In this way, the proposed model emulates decision making with and without insights from digital data to evaluate the impact on metrics (e.g. backorders, capacity utilization). The key objective is to make the model suitable for assessing investments (for instance by facilitating small-scale experiments) by focusing on the potential for better decision making under uncertainty whereby return on investment can be calculated as a function of incremental value due to insights.

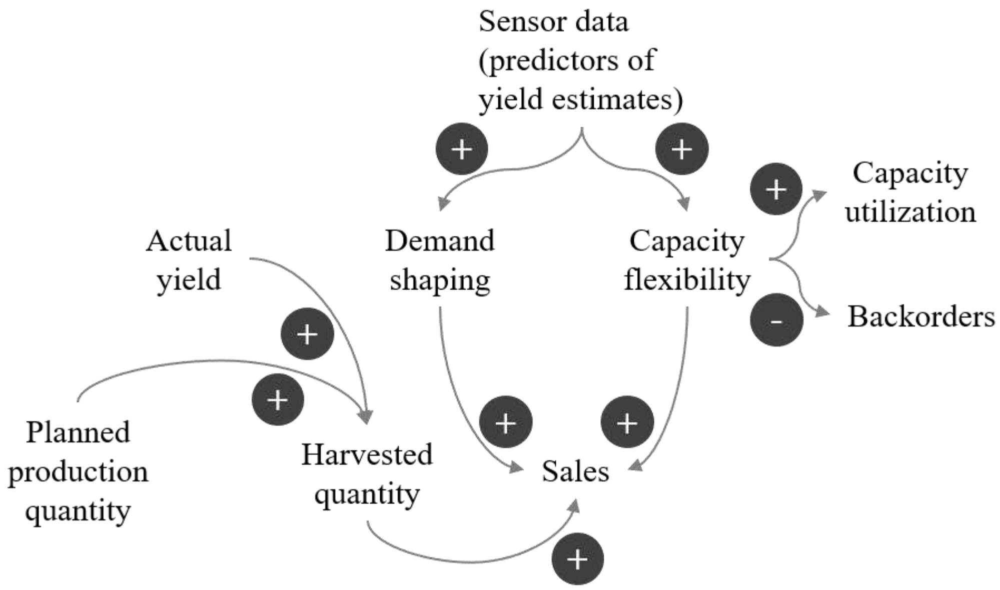

Fig. 1. An example of crop-seeds manufacturing and sales described in the text is illustrated. The increase of a certain measure causes an increase (+) or decrease (-) of the connected measure.

3 Conclusion

An implication of wide adoption of digital technologies by supply chains is the increase in decision-making complexity and uncertainty, which translates to a greater burden on information processing needs and capabilities. This strain is apparent in various studies that show that digital data is heavily under-utilized.

This paper proposed a quantitative-based model that assesses data in terms of its insightfulness, thereby enabling supply chains to address the problem of under-utilization and seeks to provide a means to evaluating digital data based on its moderating influence on uncertainty and its impact on process performance metrics.

The focus of the next stage of research is resolving design decisions pertaining to model conceptualization, which is followed by model development. The third and final stage will be model solving that is supplemented with a case-oriented proof-of-concept.

References

1. Silver EA, Pyke DF, Peterson R (1998) Inventory management and production planning and scheduling, 3. ed. Wiley, New York, NY
2. Simon HA (1977) The new science of management decision, Rev. ed. Prentice-Hall, Englewood Cliffs, N.J.
3. Galbraith JR (1974) Organization Design: An Information Processing View. Interfaces 4(3): 28–36
4. Manyika J, Chui M, Bisson P et al. (2015) Unlocking the potential of the Internet of Things. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world. Accessed 21 Jan 2019
5. DalleMule L, Davenport TH (2017) What's Your Data Strategy? https://hbr.org/2017/05/whats-your-data-strategy. Accessed 12 November 2018

6. Ross JW, Sebastian IM, Beath CM et al. (2017) Designing Digital Organizations - Summary of Survey Findings. https://media-publications.bcg.com/MIT-CISR-Designing-Digital-Survey.PDF. Accessed 12 Nov 2018
7. LaValle S, Lesser E, Shockley R et al. (2011) Big Data, Analytics and the Path From Insights to Value. https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/. Accessed 28 Jun 2018
8. Abbasi A, Sarker S, Chiang RHL (2016) Big Data Research in Information Systems: Toward an Inclusive Research Agenda. Journal of the Association for Information Systems 17(2)
9. Melville, Kraemer, Gurbaxani (2004) Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value. MIS Quarterly 28(2): 283
10. Melville N, Gurbaxani V, Kraemer K (2007) The productivity impact of information technology across competitive regimes: The role of industry concentration and dynamism. Decision Support Systems 43:229–242
11. Tambe P, Hitt LM, Brynjolfsson E (2012) The Extroverted Firm: How External Information Practices Affect Innovation and Productivity. Management Science 58:843–859
12. Brynjolfsson E, Hitt LM, Kim HH (2011) Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? SSRN Journal
13. Yu W, Chavez R, Jacobs MA et al. (2018) Data-driven supply chain capabilities and performance: A resource-based view. Transportation Research Part E: Logistics and Transportation Review 114: 371–385
14. Ketzenberg ME, Laan E, Teunter RH (2006) Value of Information in Closed Loop Supply Chains. Production and Operations Management 15:393–406
15. Dunke F, Heckmann I, Nickel S, Saldanha-da-Gama F (2018) Time traps in supply chains: Is optimal still good enough? European Journal of Operational Research 264:813–829
16. Viet NQ, Behdani B, Bloemhof J (2018) The value of information in supply chain decisions: A review of the literature and research agenda. Computers & Industrial Engineering 120:68–82
17. Lovell B (1995) A Taxonomy of Types of Uncertainty. Dissertations and Theses
18. Beamon BM (1999) Measuring supply chain performance. Int Jrnl of Op & Prod Mnagemnt 19(3): 275–292
19. Powell WB (2014) Clearing the Jungle of Stochastic Optimization. In: Newman AM, Leung J (eds) Tutorials in operations research: Bridging data and decisions. INFORMS, Hanover, Md, pp 109–137
20. Sutton RS, Barto AG (1998) Reinforcement learning: An introduction / Richard S. Sutton and Andrew G. Barto. Adaptive computation and machine learning. MIT Press, Cambridge, Mass., London
21. Morgan MG, Henrion M, Small MJ (1990) Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis / M. Granger Morgan and Max Henrion with a chapter by Mitchell Small
22. Powell WB (2017) A Unified Framework for Stochastic Optimization. https://castlelab.princeton.edu/wp-content/uploads/2017/09/Powell-UnifiedFrameworkStochasticOptimization_July222017.pdf. Accessed 30 Jan 2019
23. Ganzha M, Paprzycki M, Pawlowski W et al. (2017) Streaming semantic translations. In: 2017 21st International Conference on System Theory, Control and Computing (ICSTCC): Proceeeings : October 19 - 21, 2017, Sinaia, Romania. IEEE, [Piscataway, New Jersey], pp 1–8
24. Duggan J (2006) A Comparison of Petri Net and System Dynamics Approaches for Modelling Dynamic Feedback Systems