A Vision on Prescriptive Analytics

Maya Sappelli
TNO
Data Science, The Hague, The Netherlands
Email: maya.sappelli@tno.nl

Maaike H.T. de Boer
TNO and Radboud University
Data Science, The Hague and Nijmegen, The Netherlands
Email: maaike.deboer@tno.nl

Selmar K. Smit
TNO
Modelling, Simulation& Gaming, The Hague, The Netherlands
Email: selmar.smit@tno.nl

Freek Bomhof
TNO
Data Science, The Hague, The Netherlands
Email: freek.bomhof@tno.nl

***Abstract*—In this paper, we show our vision on prescriptive analytics. Prescriptive analytics is a field of study in which the actions are determined that are required in order to achieve a particular goal. This is different from predictive analytics, where we only determine what will happen if we continue current trend. Consequently, the amount of data that needs to be taken into account is much larger, making it a relevant big data problem. We zoom in on the requirements of prescriptive analytics problems: impact, complexity, objective, constraints and data. We explain some of the challenges, such as the availability of the data, the downside of simulations, the creation of bias in the data and trust of the user. We highlight a number of application areas in which prescriptive analytics could or would not work given our requirements. Based on these application areas, we conclude that domains with a large amount of data and in which the phenomena are restricted by laws of physics or math are very applicable for prescriptive analytics. Areas in which the human or human activities play a role, future research will be required to meet the requirements and tackle the challenges. Directions of future research will be in integrating model-driven and data-driven approaches, but also privacy, ethics and legislation. Whereas predictive analytics is often already accepted in society, prescriptive analytics is still in its infancy.**

Keywords–Prescriptive Analytics; Requirements; Applications

I. INTRODUCTION

Prescriptive analytics is one of the big data buzzwords from recent years. Being able to automatically prescribe actions in order to attain some goal would mean a huge step forward in decision support or automatic decision making for any field, especially growing fields like industry [1]. However, the problem of prescriptive analytics is its complexity [2], [3] Nevertheless, there are more and more indications that the increase in computer power allows for more complex calculations. Think only of the field of deep learning in which continuous progress is made on a wide variety of application areas. This suggest that it is time to investigate when and how we can and should apply prescriptive analytics.

In order to assess the feasibility of prescriptive analytics in any application area it is important to understand the complexity of the prescriptive analytics field. In this paper we aim to do so, by analyzing the characteristics of prescriptive problems and how it has been applied so far. We start off by explaining the difference between prescriptive analytics and its brothers descriptive and predictive analytics in Section II. We continue with the challenges and requirements in prescriptive analytics in Section III. In Section IV we use several application domains, such as oil and gas, law enforcement, healthcare and logistics, to explain in which situations prescriptive analytics might be fruitful and in which it will not. This paper ends with a direction of future prescriptive analytics research.

II. DESCRIPTIVE, PREDICTIVE AND PRESCRIPTIVE ANALYTICS

The number of organizations that base their results on data analysis is growing. In the simplest form, the data analysis of organization entails a form of **descriptive** analytics [4]. In this form of analytics, a (typically large) dataset is described quantitatively on its main features with the aim to reduce the amount of data into 'human consumable information'. An example is the extraction of simple statistics, such as average number of products that has been sold per day.

The next step of analytics is **predictive** analytics. In predictive analytics, typically a prediction is made about the future based on information from the past and current situations [5]. An example is the prediction of how many products will be sold in one month, or one year. These predictions are based on correlations and patterns in past data. A simple predictive model can be a linear regression model that assumes that the average number of sales per day decreases each month. More complex models can take into account other aspects that could influence the number of products that will be sold.

In **predictive** analytics, the underlying question is: 'What will happen?' The next step is **prescriptive** analytics: 'What should I do to make this happen?' [4]. This means that **prescriptive** analytics is focused on finding the action that should be taken to optimize some outcome, rather than focused on what will happen if I continue to do the same thing. In the sales example, an example of **prescriptive** analytics would be to prescribe the action or actions that should be undertaken to increase the average number of sales with a certain amount.

Just as descriptive and predictive analytics have tight bonds, **predictive** and **prescriptive** analytics are also strongly connected. One important reason is because **prescriptive** analytics also include predictions to estimate the effect of possible actions. However, note that these are very different kind of predictions. In prescriptive analytics, the prediction of the effect of a (sequence of) actions or interventions is central. This type of prediction deals with more complex situations such as interaction between actions or hypothetical effects for which no historical data is available. Predictive analytics only involves predictions in a single dimension based on the current and historical situations.

Imagine that a car seller wants to get some insight in his business. First, he starts with some **descriptive** analytics and calculates the number of sales and profit he made the past 5 years. He also defines some cohorts of interest, such as high end cars and low end cars and the profit he made on each of those cohorts. Then, he moves on to **predictive analytics** and calculates what his expected profit is for the next month, based on the current and past situation. He also calculates a more advance **prediction**, such as the expected number of sales if the supplier prices go up five percent. In the **prescriptive** analytics case, the car seller goes one step further. He wants to increase his total profit by five percent and wants to know what he should do; should he a) find a new supplier, b) stop selling low-end cars, or c) start an advertisement campaign. To answer this question he will most likely use some **predictions** but, since the seller has never done an advertisement campaign before, and has always used the same supplier he needs to rely on other sources of data to calculate the expected effect of strategies a) and c). Moreover, the optimal solution can also be a combination of actions a), b) and c). And, the actions can also interact. For example: an advertisement campaign may not have the same impact when the car seller stops selling low-end cars, because low-end cars might attract precisely those buyers that are attracted through the campaign, making the campaign irrelevant when the low-end cars are not available anymore.

Although a **predictive** analysis itself can also lead to a prescription, this is typically one that is straightforward: 'if the stock prices are expected to go up, I buy'. Although the question behind this **prediction** is an optimization problem (optimizing profit), the optimization itself is not part of **predictive** analytics. An example of a solution to such an optimization problem stems from optimal control theory. This is a mathematical theory dealing with finding a control law to achieve an optimality criterion [6]. On the other hand, a **prescriptive** analysis typically involves two aspects: 1) exploration of possible actions and 2) generation of the prescription. It leads to a complex prescription, such as the prescription of combinations or sequences of actions, which requires more complex predictions. Typically, the decision space for **prescriptive** analytics tends to be larger; multiple situations with many variables, options and constraints are taken into account.

Moreover, the interpretation of the prediction may not always lead to an unambiguous decision. Therefore, an important aspect of prescriptive analytics is the transparency of the method: the algorithm must be able to explain why a certain strategy is prescribed.

This also illustrates the close link between **prescriptive** analytics and **business** analytics. Business analytics has been defined as 'a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving' [7]. Hence, **prescriptive** analytics is a method for automating this manual process that is specifically suited, as all automation methods, when the job is dull, dirty, dangerous, demanding or difficult.

III. Requirements and Challenges

For a problem space to be relevant for prescriptive analytics, there are a couple of requirements that need to be present. The relevance of each requirement is determined by the application domain, but is typically present for interesting applications of prescriptive analytics. An overview of the criteria is presented in Table I, based on [2].

First and foremost the decision space needs to be complex. Complexity can arise from the number of possible actions that need to be evaluated, number of context parameters that need to be taken into account and the influence of a decision on the search space itself. Moreover the **impact** of the decision should be significant. For simple decision spaces, it is most likely sufficient to predict the outcome of the alternative situation compared to the current situation and the analysis is complete. For example, in the stock market example, the impact of the decision will be very limited (except if you are a big player) and hardly influence the new situation.

The same is true if it is not the profit on a single stock that needs to be optimized, but a complete strategy or portfolio. Hence, it is not about optimizing a single action, but a sequence of actions that needs to be executed in a coordinated manner. For such a **complex** situation in which there are multiple variables and multiple interventions that need to be optimized, a prescriptive analytics approach is more suitable for the decision maker to oversee the impact of his decision.

Another important requirement is that the objective is definable, i.e. there is a **clear quantifiable objective**, such as long-term profit. Moreover, this objective often competes with other objectives, making the decision space more complex. For large pension funds for example, the decision to buy or sell will have so many implications that the decision will not only depend on an expected stock price. Another complication that increases the complexity of the decision space even more, are possible **constraints**, for example when limited resources are available.

Finally, the required **data** should be **available**, specifically data on previous actions, decisions and the consequent situation. As predictive analytics can map the current situation to some point in the past and assume that they will progress similarly, prescriptive analytics can map the current situation to a point in the past, and the possible interventions to previous interventions and assume that the response will be similar. Hence, prescriptive analytics requires not only time series as input, but also the actions performed previously.

This specific requirement of the availability of data and specifically the actions taken is often a big challenge. It is something that is often lacking, either because of (privacy) concerns and legislations or simply because the required data is not monitored. For example, it might be very useful for the car sales-manager in the previous example, to constantly monitor the actions and emotions of employees and customers to derive the best sales tactic. However, this is probably both not desired and hard to record. The data about these employees and customers should thus be collected in a non-intrusive manner, meaning that the individuals that are monitored should not get disturbed.

In the case of little data, simulation models can provide a solution, but this can lead to simulation or knowledge-driven prescriptions, and does not exploit the full capabilities

TABLE I. REQUIREMENTS AND DESCRIPTION FOR PRESCRIPTIVE ANALYTICS

Nr	Name	Description
1	Impact	Is the expected impact worth the effort?
2	Complexity	Is the problem sufficiently complex (i.e. more than one possible action and multiple alternatives).
3	Objective	Does the problem have a clear quantifiable objective that can be optimized?
4	Constraints	Are there boundaries on the decision space that make the problem more complex?
5	Data	Is there data on the possible actions, decisions and the consequent situation?

of prescriptive analytics, hence some hybrid solution is needed. Furthermore, in simulations physical systems are easier to model through the laws of nature compared to the behavior of people. This is because human behavior tends to be less rational or predictable, whereas physical laws tend to be strict. Moreover, a generic 'human simulation model' will not capture the diversity in drives, needs and motivations that are an integral part of an individuals actions.

Another challenge regarding the data is that when (implicitly) including previous decisions to the dataset, the decisions might get towards a certain decision. To illustrate this, we move to an example in policing. Imagine that prescriptive analytics is used to steer surveillance against drugs dealers. Increasing the police surveillance in a specific area of a city will increase the number of catches in that area (and not in other areas). This will cause a prescription algorithm to increase surveillance in that area (since the expected number of dealers caught is the highest there) and creates an infinite loop. This loop for the example of prescriptive policing is visualized in Figure 1.

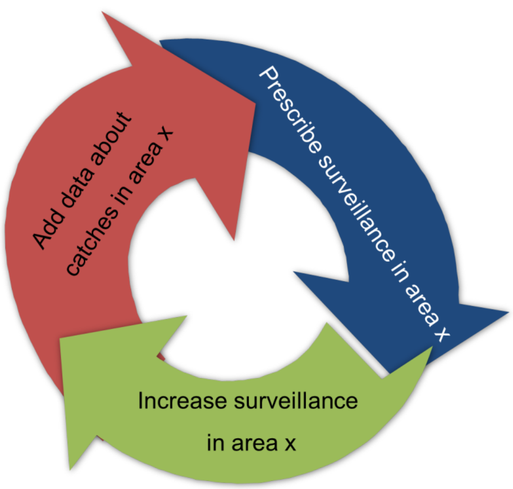

Figure 1. Loop in Surveillance

Besides challenges regarding the data, an important challenge is related to the user and trust. The system should produce decisions that are transparent, explainable and traceable in order for a user to trust and accept the decision of the system. On the other hand, the user should not overtrust the system in cases it provides a suboptimal solution. As explained before in the challenges regarding the data, the system has no creativity and can only produce results based on the data it has seen. The user is needed to validate whether the decision is right in the situation, because humans are able to use their creativity and adaptation capabilities in novel situations. Balancing between those extremes will be a major challenge in any practical application.

IV. Applications

We will elaborate on four application areas to provide an example of how, and in what kind of application areas prescriptive analytics might be useful. We indicate to what extent they meet the mentioned requirements. Note that these areas are merely used as an example, and are not an extensive list of possibilities.

A. Oil, Gas and Offshore

Oil, gas and offshore are a group of domains in which prescriptive analytics can be very beneficial and in which it is already applied [8]. Costs are very high and any reduction, no matter how small, can lead to big profits. Due to the nature of the field, the (sensor) data has nice properties, such as that they are rich, readily available and relatively accurate. The most promising application of prescriptive analytics is, confusingly, predictive maintenance [9]. Three example applications within this domain are:

(a) Predictive maintenance. Any minute a plant or turbine is not working can cost thousands of euros, especially when this occurs unplanned. Hence, maintenance is of major importance to this area. It should not be performed too late; otherwise a breakdown will cost a massive amount of money. But also performing it early is not the solution, since replacement parts are also very costly. Predictive maintenance could come to rescue by suggesting (or prescribing) the ideal moment for maintenance given weather conditions, expected demand, and sensor data indicating the current state of each part. For offshore specifically, the algorithm should also take into account that a bulk replacement might be cheaper than just-in-time.

(b) Where to drill, lift or frack. Within the area of oil and gas, prescriptive analytics can also be used to determine where to drill (or where not to) and with which techniques. Even though at first sight it does not seem to be a very dynamic problem, in practice the dynamics of an (oil) field are very volatile. The technique used for drilling or lifting at one place, affects the performance on other places and is dependent on the type of well.

(c) Automatic drill support. Prescriptive analytics can support horizontal drilling and hydraulic fracturing operations by automatically interpreting real-time sound, video and other forms of data to automatically make real-time adjustment to the parameters of the machines.

For each of the application, the relation to each of the requirements is shown in Table II.

TABLE II. APPLICATIONS IN OIL, GAS AND OFFSHORE IN RELATION TO THE REQUIREMENTS OF PRESCRIPTIVE ANALYTICS (green is requirement met, orange is neutral and red is requirement not met)

	Impact	Complexity	Objective	Constraints	Data
a. Predictive maintenance					
b. Where to drill					
c. Drill support					

TABLE III. APPLICATIONS IN LAW ENFORCEMENT AND JUSTICE IN RELATION TO THE REQUIREMENTS OF PRESCRIPTIVE ANALYTICS (green is requirement met, orange is neutral and red is requirement not met)

	Impact	Complexity	Objective	Constraints	Data
a. Prescriptive Policing					
b. City Planning					
c. Sentencing					

B. Law Enforcement and Justice

Predictive Policing is one of the hot topics within law enforcements all over the world. A typical application is to predict where and when a certain type of crime will occur [10]. Currently, applications exist that use predictive analytics to predict who is most likely to be the victim or the perpetrator, or who is most likely to recidive. Although these applications provide some insight into the dynamics of crime, they are not necessarily of much use directly. Since the objective is to prevent crime from happening, it is more important to know what will be the effect of your intervention; this leads to Prescriptive Policing. This is a challenging field, not only because of the human behavior that is included, but more importantly because of the privacy, bias and other concerns, such as ethical profiling. Furthermore, it is to be expected that explainability of the outcomes is necessary to stand a chance in court. Examples of possible applications of prescriptive modeling in the law enforcement domain are:

(a) Prescriptive Policing. As mentioned above, Prescriptive Policing is applying prescriptive analytics in order to determine the best possible intervention to prevent crime from happening. Problems are all over the ‘requirements spectrum’; constraints are unclear, data is not available or legislated by law, the impact is hard to monetarize. Even the objective is not clear: do you want to prevent crime, or catch criminals?

(b) City Planning & Legislation. From Environmental Criminology it is known that the environment, and specifically the buildings, parks and infrastructure can have a great impact on the actual and the perceived amount of crime [3]. As local government can, more or less, control them, predicting the effect of city planning and legislation could lead to safer cities. Again impact and constraint are unclear, however data is readily available.

(c) Sentencing. Law firms can use algorithms that offer predictions on certain cases and based on how similar cases fared in the same jurisdiction give a prediction how new cases could work out. The small Californian law firm Dummit, Buchholz & Trapp already uses such technology, developed by LexisNexis, to determine in 20 minutes whether a case is worth taking or not [11]. One might even imagine that judges are replaced by prescriptive algorithms that determine the appropriate sentence. This could lead to a more objective and consequent practice. Whether society would accept such developments is, however, highly uncertain.

Table III shows for each application which of the requirements are met.

C. Healthcare

The domain of healthcare is typically well-suited for the application of prescriptive analytics [12]. The impact of decisions in the healthcare-domain is large and the decision space, with patient types and possible treatments is complex. There are also clear trade-offs and constraints in healthcare that cannot be ignored. Especially with additional constraints coming from insurance companies, decision making and optimization is important for hospitals and other healthcare professionals. However, the reason why it has not yet been applied in healthcare often is that it usually requires the modeling of a human action. As explained earlier, this is more difficult than modeling physical systems because their range of choices and actions is much more diverse and less predictable. Examples of possible applications of prescriptive modeling in the healthcare domain are:

(a) Activity planning for the optimization of an individuals well-being. In this example a prescription can look like a sequence of actions that an individual would need to take in order to improve their well-being [13]. Actions can include adapting sleeping behavior, eating behavior, physical activity or a combination. Problems in this example is the effort and privacy issues involved with collecting data on an individuals sleep, eat and activity behavior. Moreover, the objective - increasing well-being - is ill-defined and highly subjective, making it harder to optimize.

(b) Hospital constraint modeling reduce cost and increase throughput. Another example is on the level of optimizing cost and throughput in a hospital. This problem is highly complex since many people are involved. Especially constraints on availability of employees can make the problem difficult. There are multiple objectives that play a role in this scenario. Not only cost and throughput are important, but patient satisfaction as well. The impact is large, since sending patients home too early is undesirable in the long run. It is likely that hospitals have a sufficient amount of data to work on these prescriptive scenarios. Sir Mortimer B. Davis Jewish General Hospital has been looking into enterprise optimization using prescriptive analytics [14].

(c) Personalized decision support for medical experts. A final example in the healthcare domain would be the prescription of a treatment plan, personalized for an individuals specific situation. The biggest problem in this scenario is the availability of required data and the privacy issues that are involved once aspects, such as sleep, food and activity are involved.

Table IV gives an overview of the relation between the requirements and each application.

TABLE IV. APPLICATIONS IN HEALTHCARE IN RELATION TO THE REQUIREMENTS OF PRESCRIPTIVE ANALYTICS
(green is requirement met, orange is neutral and red is requirement not met)

	Impact	Complexity	Objective	Constraints	Data
a. Activity planning					
b. Hospital constraint modeling					
c. Personalized decision support					

D. Logistics

The logistics domain seems an interesting domain for the application of prescriptive analytics. Although humans are often in the loop, they are typically not the core of the objective that needs to be optimized. This means that a less advanced human model is required, making the application of prescriptive analytics more realistic. Examples of possible applications of prescriptive modeling in the logistics domain are:

(a) Routing of ships and trucks for loading and unloading on docks [15]. A possible example of a routing problem is the case where multiple companies are cooperating. This makes the problem complex and interesting, since there are strong interactions between actions, i.e. if a single driver pauses, than this effects other trucks and ships as well. Moreover, there is a clear objective reducing time and optimizing throughput. Constraints are also clear, for example local speed limits. With the increase in use of GPS trackers, there is a continuous availability of streaming data. The optimization of the activities of all parties involved at a dock can have a large impact on total revenue and throughput.

(b) Dairy farm optimization. In the food industry, farms collect a lot of data about their farm and their animals [16], [17]. This ensures an optimal flow of milk and other animal products. Food industry in general has a large impact on society, however the impact of a single farm might be small. There is an interesting overlap between optimizing the health of animals and optimizing the health of people. Constructing animal models might be less complex than constructing human models, making this a particularly interesting example application. Moreover, optimization in dairy farms could also be focused on optimizing supply-demand flows. This makes the decision space, from cow to retailer elaborate and complex.

The relation of these two applications in the Logistics domain to the requirements of prescriptive analytics is shown in Table V.

TABLE V. APPLICATIONS IN LOGISTICS IN RELATION TO THE REQUIREMENTS OF PRESCRIPTIVE ANALYTICS
(green is requirement met, orange is neutral and red is requirement not met)

	Impact	Complexity	Objective	Constraints	Data
a. Routing on Docks					
b. Dairy farm optimization					

V. Conclusion and Future Research

It is immediately clear from the applications in the previous section that the low hanging fruit of prescriptive analytics is in those areas with lots of data and in which the phenomena can be described with physics or math. Logistics and oil, gas and offshore are just two examples, but automotive, chemistry and additive manufacturing can also benefit from prescriptive analytics.

The areas which are more challenging are areas in which we do not have the data available or in which we have a gigantic number of possible actions or a large degree of freedom. In these circumstances, models and knowledge can come to aid as these can fill in the gaps of missing data [18]. Human behavior is the most typical example here. Healthcare, Law Enforcement, Human Resource Management, Force Protection, Sustainability; each of them has the promise of major (societal) impact. In terms of algorithms, this means adapting or combining current predictive algorithms, self-learning algorithms and knowledge-driven algorithms to deal with the complexity of a prescriptive analytics problem.

Within a few years Google, Facebook, IBM and other companies will deliver prescriptive algorithms - 'as a service'. Any company can, and will, move towards a more data-driven approach in decision making. Prescriptive analytics is the Holy Grail in this area; whereas descriptive and predictive algorithms still make you do the thinking, prescriptive algorithms are the first that actually deliver actionable insights. However, although building such algorithms is easy, controlling them and keeping them clear from biases is not. These dynamics are complicated to grasp and require extensive knowledge from both self-learning algorithms and the application area.

Finally, privacy, ethics and legislation are important issues in some of those areas and should not be overlooked. We should not fear a Minority Report; predictions and prescriptions are no forecast, they are just math. We should be careful also to treat them as such. Humans are intrinsically lazy, and will readily comply, even if the prescription is coming from a machine. Hence, machines that are making prescriptions are not that different from autonomous systems, and should therefore be considered equal. If we do not want autonomous weapons, we should also not build prescriptive ones either.

Acknowledgment

We would like to thank the Early Research Program 'Making Sense Of Big Data' of TNO for their financial support.

References

[1] A. Banerjee, T. Bandyopadhyay, and P. Acharya. Data analytics: Hyped up aspirations or true potential? [Online]. Available: http://www.vikalpa.com/pdf/articles/2013/04-Perspectives.pdf (2013)

[2] C. Centurion. Prescriptive analytics in healthcare. [Online]. Available: http://www.riverlogic.com/prescriptive-analytics-in-healthcare (2014)

[3] S. Smit, B. van der Vecht, and L. Lebesque, "Predictive mapping of anti-social behaviour," European Journal on Criminal Policy and Research, vol. 21, no. 4, 2015, pp. 509–521.

[4] D. Delen and H. Demirkan, "Data, information and analytics as services," 2013.

[5] M. A. Waller and S. E. Fawcett, "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management," Journal of Business Logistics, vol. 34, no. 2, 2013, pp. 77–84.

[6] D. E. Kirk, Optimal control theory: an introduction. Courier Corporation, 2012.

[7] M. J. Liberatore and W. Luo, "The analytics movement: Implications for operations research," Interfaces, vol. 40, no. 4, 2010, pp. 313–324.

[8] S. Gupta, L. Saputelli, M. Nikolaou et al., "Applying big data analytics to detect, diagnose, and prevent impending failures in electric submersible pumps," in SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, 2016.

[9] R. K. Mobley, An introduction to predictive maintenance. Butterworth-Heinemann, 2002.

[10] S. Smit, A. de Vries, R. van Kleij, and H. van Vliet, "From predictive to prescriptive policing," TNO report, The Hague, 2016.

[11] S. Thammaboosadee, "Big data and big lawyer," TIMES-iCON2016, p. 193.

[12] W. Raghupathi and V. Raghupathi, "An overview of health analytics," J Health Med Informat, vol. 4, no. 132, 2013, p. 2.

[13] S. Koldijk, W. Kraaij, and M. A. Neerincx, "Deriving requirements for pervasive well-being technology from work stress and intervention theory: framework and case study," JMIR mHealth and uHealth, vol. 4, no. 3, 2016.

[14] P. Troy, A. Berg, J. Richards, G. Pellerin, and L. Rosenberg, "A prescriptive analytics project for maximizing healthcare value generation," 2015.

[15] A. G. Dominguez, "'smart ships": Mobile applications, cloud and bigdata on marine traffic for increased safety and optimized costs operations," in Artificial Intelligence, Modelling and Simulation (AIMS), 2014 2nd International Conference on. IEEE, 2014, pp. 303–308.

[16] G. Leopold. Dairy industry asks: Got big data? [Online]. Available: https://www.datanami.com/2014/10/31/dairy-industry-asks-got-big-data/

[17] R. Schrijver, P. Berentsen, R. A. Groeneveld, A. Corporaal, and T. de Koeijer, "Development of nature-oriented dairy farm systems with an optimization model: the case of 'farming for nature' in 'de langstraat', the netherlands," Agrarwirtschaft, vol. 55, no. 5/6, 2006, pp. 280–289.

[18] M. J. Mortenson, N. F. Doherty, and S. Robinson, "Operational research from taylorism to terabytes: A research agenda for the analytics age," European Journal of Operational Research, vol. 241, no. 3, 2015, pp. 583–595.