Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

EDUCATION IN ANALYTICS NEEDED FOR THE MODELING & SIMULATION PROCESS

James F. Leathrum, Jr.

Department of Computational Modeling and Simulation Engineering
Old Dominion University
Norfolk, VA 23529, USA

Andrew J. Collins
T. Steven Cotter

Department of Engineering Management and Systems Engineering
Old Dominion University
Norfolk, VA 23529, USA

Christopher J. Lynch
Ross Gore

Virginia Modeling, Analysis, and Simulation Center
Old Dominion University
Suffolk, VA 23435, USA

ABSTRACT

The increase in the availability of data has led to organizations asking how to use that data to understand their processes and to plan for their future. This paper discusses integrating analytics with modeling and simulation in a sequence of courses intended to provide organizations the ability to utilize their data to make better-informed decisions. Classical modeling and simulation education has utilized simple statistical techniques to address input data modeling and output analysis. Our premise is that a solid background in analytics provides an analyst not only with the ability to understand the data better but also the ability to improve the quality of a simulation's data inputs and the ability to analyze the increasing amount of data generated by simulations. The paper proposes a set of elements of analytics and presents how they can impact a basic set of modeling and simulation project activities.

1 INTRODUCTION

Data is increasingly becoming a driving force in decision making. Not only is data integral in making organizational business decisions, but it is also regularly used to explain and justify decisions to the public. A current example is the use of data to convey to the public the severity of the coronavirus and the impact of social distancing (Johns Hopkins University of Medicine 2020). This requires handling large data sets, frequently not structured with the intent of supporting the desired analysis (poorly formed, missing data, etc.), analyzing the data, and, most importantly, presenting it in an understandable manner. The data can then be utilized to forecast future trends. Should the processes of the system under consideration be well understood, something not currently true of the coronavirus, the data can provide the backbone of a modeling and simulation (M&S) study, which in itself can generate large quantities of data requiring analysis.

Traditionally, organizations such as the United States of America's Naval Sea System Command (NAVSEA), the sponsor of this effort, are primarily entrenched in spreadsheet-based analysis. Without a background in analytics, their analysts rely on classical statistics due to the lack of access to or awareness of modern tools to support more advanced techniques. A curriculum has been developed that addresses these shortcomings through a set of courses beginning in analytics and finishing with M&S.

Classical M&S education falls short in addressing the current data needs, relying on traditional descriptive and inferential statistical approaches to input data modeling (extracting distributions from historical data) and analysis of output data (Law and Kelton 1991). One area M&S has been aggressively advancing, in the handling data, is data visualization (Vernon-Bido et al. 2015). But the visual results of the simulation are meaningless if the data driving the simulation is not of sufficient quality (Roman 2005). This paper argues that a solid educational background in analytics enables analysts to be better suited to using simulations because the data requirements of simulations continue to rise. The question is how to provide that education. Our solution is a set of courses that begin with a solid foundation in analytics (Data Analytics, Predictive Analytics, Data Modeling, and Data Management) leading into an M&S course. These courses are designed to quickly bring an organization up to speed with future trends in data, focusing on their need to understand the data and the ability to use it in the decision-making process.

After a brief background discussion, this paper begins by presenting a general model of M&S processes and the activities required to support it. Four elements of analytics are described and how they relate to M&S. These elements are then related to the appropriate M&S activities. Examples are provided to demonstrate how advanced analytic techniques, such as machine learning and deep learning, are being used in M&S; these examples are used to further support our argument.

2 BACKGROUND

There has been an exponential growth in data (Gutierrez 2017) and, with this growth, a rise of new techniques to help draw information for this data. These techniques fall under the umbrella term data analytics, sometimes abbreviated to just "analytics" (Foote 2018). Some professional societies and organizations rebranded themselves due to the popularity of analytics; for example, The Institute for Operations Research and Management Science (INFORMS 2019) rebranded its Interfaces journal in an attempt to align the society more closely with analytics (calling it the INFORMS Journal of Applied Analytics); they also created the Certified Analytics Professional (CAP), a certification process for analytics.

The rapid rise of the field has resulted in no definitive definition of data analytics, and, as such, we provide a simple definition: applying algorithmic processes to derive insights from the available data. This definition is an attempt to not confuse analytics with other related fields like analytical methods. Though both analytics and analytical methods derive their names from the word analysis, it should be noted that the two methodologies are very different. Data analytics is concerned with the holistic study of empirical data, whereas analytical methods are concerned with logical deductions about a generalized phenomenon using mathematical proofs.

The techniques and methods relating to analytics can be split into subfields. The National Academic of Sciences splits analytics into three subfields: descriptive analytics, predictive analytics, and prescriptive analytics (Lustig et al. 2010). Descriptive analytics and prescriptive analytics are associated with data analytics. Predictive analytics uses data analytics, combined with advanced techniques, to make predictions about a phenomenon of interest. Lustig et al. define predictive analytics as "the extensive use of data and mathematical techniques to uncover explanatory and predictive models of business performance representing the inherent relationship between data inputs and outputs/outcomes" (Lustig et al. 2010, p. 2).

Analytics is not simply statistics; it is a holistic approach that encompasses data mining, data management, and machine learning. As such, we have split it into four aspects of analytics:

- Data Analytics
- Predictive Analytics
- Data Modeling
- Data Management

These topics were chosen as the focus of the study because they relate to the training courses the development team has derived for NAVSEA, which, ultimately, lead to a final course in M&S.

This paper discusses how analytics can be used in the M&S process, justifying the curricular structure the team developed for the training courses, and what are the related education requirements. It is important to note that the reverse is also true, that M&S can be used in the analytics process. For example, Monte Carlo simulation is one of the key methods used in data and predictive analytics (Lustig et al. 2010). M&S involves abstracting a system of interest into a model to be then implemented as a simulation, usually computerized, for the purpose of better understanding the system. Given that both analytics and M&S focus on understanding, it is unsurprising that there has been a two-way connection between these two fields of study and, thus, there is a benefit for considering them together educationally.

It is difficult to determine what research has already been done connecting analytics and M&S. Much in the same way that INFORMS has rebranded its journal, some papers seem to be a rebranding of simulation optimization as data analytics (Miranda et al. 2018). However, we have found several unique attempts to combine M&S with data analytics. Example problem domains include housing, agriculture, and healthcare. For example, Gutiérrez-Andrade et al. (2019) combined simulated annealing, a Monte Carlo approach, with the artificial bee colony algorithm, an optimization algorithm, for simulating residential redistricting. Bouffard et al. (2018) incorporated data analytic concepts into a discrete-event simulation for modeling agriculture supply-chains. Byrum et al. (2016) employed both Discrete Event Simulation (DES) and Monte Carlo simulation in their agricultural models. Schneider et al. (2018) used DES in models to improve bed utilization in hospitals. Section 6 of this paper gives further examples of applications. We have not found any papers that connect analytics to an M&S curriculum, and we believe our paper is the first.

3 M&S CURRICULUM

The M&S curriculum on which this paper is based is targeted to subject matter experts in the business and manufacturing community who are interested in improving their ability to study the impact of processes and scheduling on their workflow. The goal is to move them out of a classic spreadsheet world of analyzing their systems into the modern world of data analytics and predictive analytics to better understand their system from a data perspective. Then they will be introduced to M&S as a means of evaluating alternatives.

M&S is a vast discipline and must be pared down to a subset that is viable within a week-long course. The emphasis is placed on M&S that aligns with the anticipated needs of the target audience, resulting in a focus on stochastic Monte Carlo and discrete-event simulations shown in Figure 1 from Law and Kelton's classic model taxonomy. Continuous simulations are briefly discussed to set a context should they be necessary for a hybrid system, though it is not reasonable to assume they will be capable of building such a system from this curriculum (it is anticipated that not all students will have the necessary background in differential equations). In addition, while agent-based modeling and simulation (not depicted in Figure 1) is an important concept, insufficient time will be available to make them practitioners. A more modern taxonomy can be found in Lynch and Diallo (2015) that identifies model characteristics in two levels of detail to identify appropriate simulation techniques, and this will be presented to the students to give them a more complete view of the field. However, the classic taxonomy from Law and Kelton was utilized in this paper due to its simplicity in identifying the classes of continuous, discrete, and static (Monte Carlo) to easily map to the data analytics techniques supporting them from a curricular viewpoint.

Regardless of the paradigm used, a simplified set of processes is presented in Figure 2 for the lifecycle of a simulation project. The M&S activities, shown in the figure, are where we identify that analytics concepts can be integrated into the M&S process. Input modeling identifies appropriate distributions that represent the modeled system. Data modeling captures the requirements of the data included in the simulation. Experimental designs help to cope with state-space explosions that result from models containing large numbers of components, parameters, and interaction opportunities. Evaluating simulation outcomes requires the use of analytics to provide statistical support for the significance of the outcomes. Visualization pertains to how simulation results can be intuitively conveyed to all parties in support of interpretation and accurate decision support. Discussion in section 5 of this paper will tie these M&S activities to the elements of analytics discussed below.

It is understood that this is an overly simplified, waterfall-based approach. It is assumed that students can apply principles from other courses, for instance, from engineering management or software engineering, to fill in the gaps. They will also be given references to more realistic approaches such as (Aalst 2015; Balci 1986; Balci 1998; Robinson 1997; Sargent 1981; Tolk et al. 2013) to further investigate the concepts. However, Figure 2 is sufficient to highlight how the other components of the overall curriculum impact the culminating M&S course.

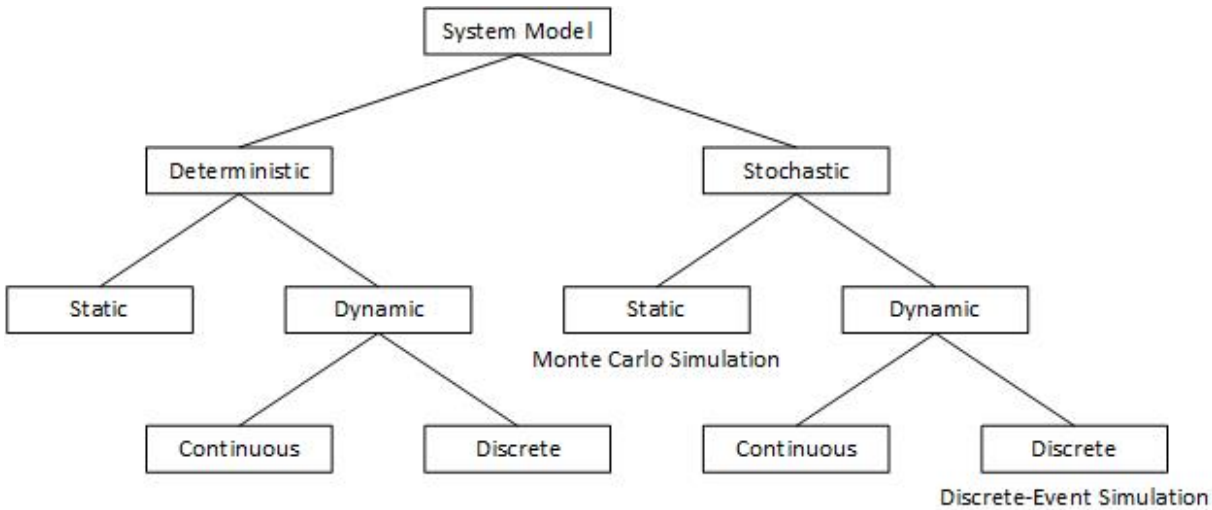

Figure 1: M&S model taxonomy (Law and Kelton 1991).

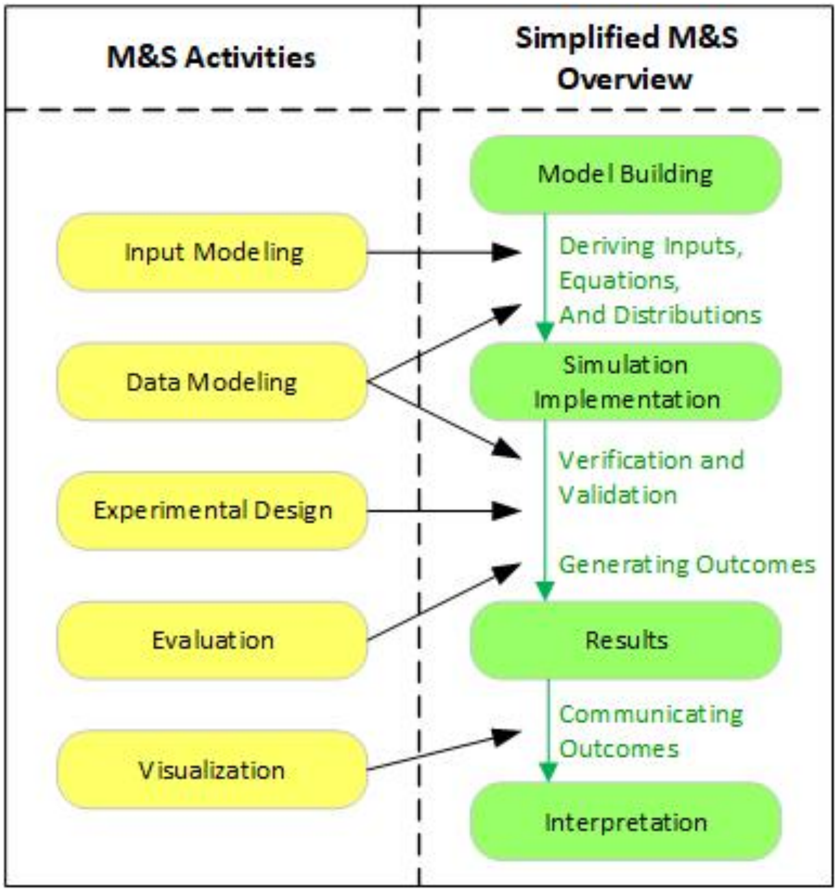

Figure 2: Simplified view of M&S processes and activities involved.

4 ELEMENTS OF ANALYTICS

Having identified four aspects of analytics: data analytics, predictive analytics, data modeling, and data management, these then constitute the set of four courses as a prerequisite to an M&S course in the curriculum. Each of these elements of analytics is described in this section, along with their relationship to M&S. Section 5 then ties these to the M&S activities in Figure 2.

4.1 Data Analytics

As previously noted, there is no universally accepted definition of data analytics. It could be defined as a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions and supporting decision-making, or it could be defined as the use of computer systems to the analysis of large data sets to support decisions. INFORMS (2019) defines it as the scientific process of transforming data into insight for making better decisions. Since there are multiple facets that are used in these definitions, any educational syllabus on data analysis must take into account these facets if it is going to satisfy the potential heterogeneous demands of employers of data analysts, e.g., knowledge of statistics, computer science, and operations research. Unsurprisingly, multiple syllabi exist, which cover different topics.

The INFORMS' Certified Analytics Professional (CAP) syllabus covers problem structuring, data collection, methodology and its selection, Data Modeling, and model development. In contrast to this, Runkler' book on data analytics focuses on visualization, multivariate statistics, and the resultant mathematics. Our own developed syllabus on data analytics focuses on data wrangling (transforming data into a usable format), descriptive statistics, machine learning, visualization, and tool use (RStudio). What is interesting about these syllabi is that each has a clear overlap with any holistic M&S syllabus (e.g., data modeling, visualization, and tool use). Thus, the challenge for introducing data analytics into any M&S syllabus is determining what "flavor" of data analytics will be considered and how much synergistic education exists.

Though there are overlapping elements, there are practical differences between data analytics and M&S. The computer software packages used are an example. Data analysts tend to use R, a statistical computing language, and Structured Query Language (SQL). Python is another programming language that is used, which has the added benefit that it is also used for simulation projects, i.e., Mesa (Bit.ly/abm-mesa). Any M&S syllabus that includes data analytics would have to include the mastery of some of these computing tools.

4.2 Predictive Analytics

Both predictive analytics and computer simulation attempt to predict the behavior of complex systems, Predictive analytics through applying analytics techniques to historical data to predict future trends based on the data and simulation through modeling the system and then executing the model by use of a simulation. The importance of data analytics to support the model development in M&S has already been discussed, but predictive analytics plays a role as well. Certainly, they can be used together to provide more information in understanding future system behavior. Budgaga et al. (2016) discuss the use of predictive analytic models of scenario variants to provide interactive feedback to improve simulation prediction accuracy. Predictive analytics can also be utilized in the data modeling phase of simulation to improve the input data model (Miller et al. 2013; Bryan et al. 2015). There is also evidence of integrating predictive analytics into the simulation process. Adra (2016) discusses the extension of simulation models to use real-time predictive analytics resulting in improved simulation speeds and accuracy. Ouyang and Nelson (2017, p. 1716) developed a "… two-step method to dynamically predict the probability that the system state belongs to a certain subset and test the …" simulated performance.

Given the potential for leveraging simulation with predictive analytics, the challenge in designing predictive analytics education is balancing between short-term introductory data analytics skills and long-term integrated data-driven, predictive analytics and simulation engineering and scientific capabilities. To create a balance, educators can apply the 80/20 Pareto principle in designing course content. The structure of the Predictive Analytics course developed by our team allocates approximately eighty percent of course content to basic spatiotemporal regression, smoothing methods, and ARIMA forecasting and prediction models, including coding examples in the selected statistical package language (R). Approximately sixteen percent of the course content is allocated to discussing and demonstrating issues in building predictive analytics models (effects of sample bias on model accuracy and precision, data preparation and transformation, probabilistic versus optimization versus predictive simulation models, and validating models against real-world phenomenon behavior). Finally, approximately four percent of course content is allocated to discussion of the advanced topic of predictive analytics and simulation integration, both benefits and challenges. The sequential 80/20 Pareto allocation of predictive analytics training will provide learning goals for students at each skill and knowledge level and lay the foundation in terms of analytics skills, engineering, and science necessary for successful integration into organizational policies, procedures, and practices.

4.3 Data Modeling

Of all the topics discussed, data modeling is the one that has the most overlap with M&S, even appearing in our M&S activities list. Data modeling involves capturing the data requirements necessary to study the problem under consideration. The data requirement capturing might have various formal schemas of varying levels of abstractions (conceptual, logical, and physical) to support information system development and analytics. The ultimate goal of this process is to generate a database that fits the organization's needs. Thus, it is expected that data modeling is already covered in an existing M&S syllabus, as reflected in Figure 2, and does not require elaboration here.

4.4 Data Management

The increased application of simulation in the design of products, processes, and systems enables organizations to identify and correct design problems, optimize performance, and innovate at a lower cost. Correspondingly, predictive analytics, and simulation consume large volumes of data, which currently is not stored and managed for efficient sharing and reuse. Given the volume of data required and the predictive analytics and simulation objects generated to support the development of just a single model, let alone an organization's analytics and simulation needs, it is essential that modelers follow a data management protocol. Organizations must implement formal data management programs to preserve and secure original data sets, intermediate structured and cleaned data sets, intermediate and final models, and graphical objects. A data management protocol must address: (1) acceptable data set formats for the selected statistical analysis software, (2) data set file structures and metadata definitions, (3) statistical analysis object file structures and metadata, and (4) analytics and simulation model specification and related graphical object output file structures and metadata.

The Data Management Association DMBOK2 provides a framework on which an organization can design its data management governance to plan and coordinate the use, archive, retrieval, control, and purge of data sets and model objects. This provides a structure on which our education content is developed. The focus is on metadata, a structured data set that provides information about data sets, statistical objects, analytic and simulation models, and graphical objects. This facilitates storage, access, and retrieval. Three general types of metadata covered are descriptive, structural, and administrative.

In designing analytics training, it is essential to introduce students to data management fundamentals and open source and proprietary data management software available for facilitating organizational data management. Typically, today's statistical computing software provides functions to add metadata to variables, to generate metadata about data sets, and to attach metadata to analysis and graphical objects. Likewise, simulation metadata management software is available to manage simulation structures across the simulation product lifecycle.

5 RELATIONSHIP BETWEEN ANALYTICS AND M&S

Having examined the elements of an analytics course and its potential connections to M&S education, the M&S activities in Figure 2 are revisited to clearly identify how each analytics element supports the M&S activities. It is this relationship that drives the educational curriculum discussion. Analytics fuel many perspectives within the purviews of building models and evaluating simulations. Figure 2 conveys some primary associations between the M&S activities of input modeling, data modeling, experimental design, evaluation, and visualization and their relationships to a simplified high-level representation of the M&S process. The case is now made as to how a strong background in data analytics makes a better simulationist. Figure 3 provides associations that can link analytics to the simplified M&S process view and M&S activities provided in Figure 2. It also highlights the importance as the move is made from a waterfall approach to a more realistic iterative approach.

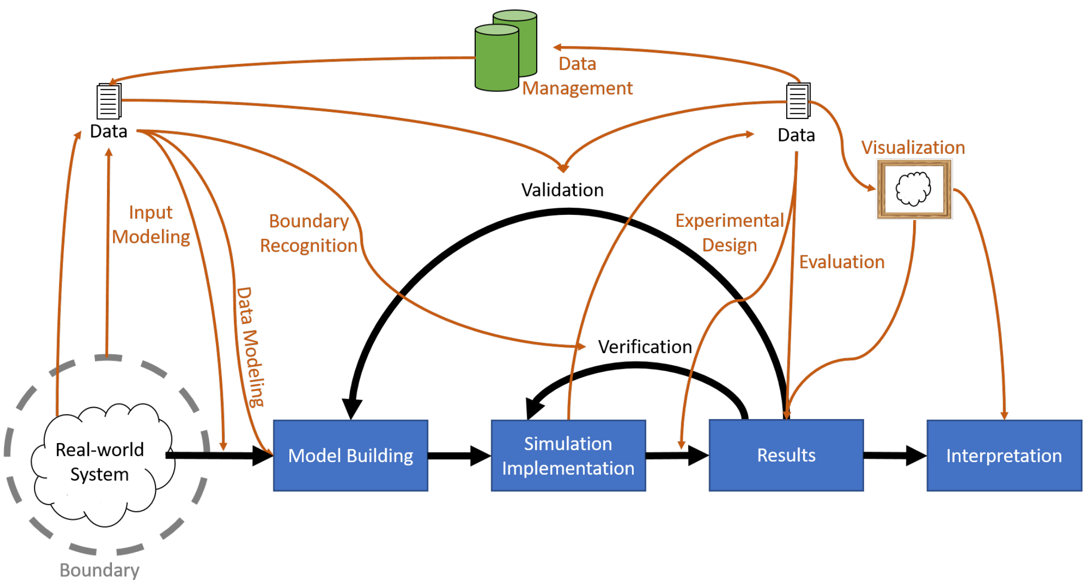

Figure 3: Analytics within the purview of a simplified view of M&S processes and activities.

The simplified M&S process is shown paired with high-level M&S activities in Figure 2. Analytics activities, such as data analytics, predictive analytics, data modeling, and data management, can aid in achieving each of these primary M&S processes. Understanding and properly applying analytics techniques can assist in: (1) filtering the noise from the desired information when assessing and incorporating real-world data into the model building process (Kleijnen 1998); (2) supporting verification and validation by evaluating whether the model is an acceptable representation of the real system or the simulation a correct representation of the model (Lynch 2019); (3) aiding transparency and understanding in the presence of expected and unexpected outcomes (Lynch et al. 2020); and (4) providing traceability of outcome data to the model design through effective data modeling and data management practices (Tolk et al. 2013).

Input modeling takes observations, theory, and prior knowledge and converts it into usable information that informs the simulation structure and initialization. Data comes in many forms, including one-dimensional and multi-dimensional data, as well as data that is specialized for a given context (Feldkamp et al. 2020). These different types contain varied requirements pertaining to how they can be properly analyzed. Analyzing input models is circular as the system provides the characteristics of the input model, while validity often relies on comparing input-output behaviors back to the system. Input models help in connecting the simulation to its intended use by providing the boundaries for the intended execution of a simulation. Characteristics of input models include being time-dependent or independent, univariate or multivariate, or discrete or continuous (Leemis 1995). The characteristics should reflect the modeled system at the necessary level of abstraction to address the modeling question. Exercising a simulation under its input models' boundaries can help to identify sensitive parameters (Vincent 1998; Walrave 2016) and reveal inconsistencies or contradictions with respect to its intended design. Input modeling is an iterative process, and, as such, we believe that data analytics and predictive analytics can provide support in enabling an input modeler to give a better understanding of the system under consideration.

The next activity on the list in Figure 2 is data modeling. As previously mentioned, data modeling is a key component of analytics and M&S. How data modeling is implemented into a simulation's development depends on the scale and resolution of the model developed.

Models may contain very large state spaces as a result of the number of components represented, the number and variety of interactions that can occur between components, and the scale of the simulation runs. In these cases, using good data management principals and methods can greatly ease the development burden. Exhaustive testing of the simulation to identify errors and support credibility is often infeasible. As a result, sampling techniques are relied upon as a route for selecting a wide enough range of configurations from the simulation to still provide statistically significant value for the analysis. Sampling and analysis techniques should be used together to support model credibility and strengthen exploration in the face of increasing combinations of temporal-, spatial-, and network-based interactions (Lynch et al. 2020). Data analytics is just one way to investigate these samples.

Design of experiments (DOE) aid in selecting the combinations of factors to utilize for initialization to reduce time requirements and aid evaluation (Kleijnen 1998). Effective application of a DOE allows for engineering a testing solution that accounts for needed sample sizes for statistical evaluation for all of the tests intended for the simulation. Additionally, a pre-specified DOE provides clarity in design, purpose, and in conveying results. While a state-space may be large, remember that "confusion and clutter are failures of design, not attributes of information" (Tufte 1990). DOEs can greatly reduce confusion and clutter of output data. Data analytics, data modeling, and data management all have their place in helping explore and managing these large state spaces.

Simulation evaluations are conducted for numerous reasons: verifying that the simulation construction is correct, validating that the simulation's behaviors adequately represent the real system, conducting what-if analyses, supporting decision making, exploration, and teaching and learning. Across these topics, a primary shared theme is understanding. Exploration is conducted to understand how the system works, interactions are analyzed to understand model behaviors, and simulation components are tested to understand if the simulation is working correctly. Mathematical techniques such as descriptive statistics, sampling statistics, and hypothesis testing aid in describing the interworking of the simulation and in providing evidence when incorrect behaviors occur. Statistical methods help deal with difficulties in experimental error and the complexities of the effects studied (Box et al. 1978; Navidi 2008). Data analytics allows a holistic approach to apply these statistical techniques when a large output dataset has been generated; as mentioned previously, the purpose of a data analyst is to provide a deeper understanding of data, which clearly supports simulation evaluations.

Visualization, as well as other sensory feedback avenues, provides a bridge between simulation occurrences, evaluations, and interpretation (Tufte 1990) and has been shown to increase simulation users' abilities to correctly identify errors (Lynch 2019). For example, showing intervention dynamics plotted within outcome graphics can enhance understanding of the simulation by providing a direct mapping between interventions and outcomes (Walrave 2016). Many frameworks and methodologies exist for properly utilizing visuals to aid understanding of analytical outcomes and simulation behaviors to convey meaning and context without biasing interpretation. Tufte (2006) presents six principles of analytical design in support of this: (1) comparisons, (2) causality, mechanism, structure, and explanation, (3) multivariate analyses, (4) integration of evidence, (5) documentation; and (6) content. The intended level of representation, such as individual values versus aggregated values, can be considered while creating a DOE to further support the identification of parameter combinations. Visual components support the interpretation of data as well as communicating the design of an experiment. We believe that visualization is important to any analytics curriculum and, as such, have included a visualization component in all our analytics courses as well as the M&S course.

This discussion on the M&S activities shows how the elements of analytics can help at all stages of the M&S process. We are not advocating that every M&S project requires analytics, but there are clearly circumstances when it provides support, e.g., when a project involves large input or output datasets.

6 NOVEL ANALYTICS APPLICATIONS TO M&S

While analytics has always been coupled with M&S, a variety of novel analytics applications to M&S are becoming increasingly prevalent. These applications rely on historical data, real-time data collection, and unstructured data sources. Examples include the use of machine learning and deep learning.

The term analytics frequently implies the application of machine learning techniques. Machine learning analysis is well suited for M&S problems in which: (1) some behavioral aspects of the system are difficult to model, but (2) there is enough data to apply machine learning to emulate the behavior. This approach is effective where determining a specific decision rule (i.e. First-In First-Out, Last-In First-Out) for a system is difficult. Recently a manufacturing simulation used neural networks to emulate control rules. In the simulation, a single type of decision was obtained from a neural network and combined with other behavioral aspects that were modeled in the simulation as a discrete, process-oriented workflow (Bergmann et al. 2017).

Additional applications of machine learning analysis within M&S that are similar in purpose to the use of predictive analytic forecasts also exist. A metropolitan emergency room simulation predicts time series data of energy consumption of machines by using deep learning models as input into a traditional discrete-event simulation (Wörrlein et al. 2019). While the setup and the training of these networks is highly complex, the combination of simulation and machine learning predictions brings together the two worlds which are well suited for one another.

Traditional analytics applied to empirical data has a long tradition in model building. However, the use of deep learning, machine learning, and other statistical approaches in modeling building are becoming more prevalent. Furthermore, once built and deployed, analytics are able to process the substantial data resulting from simulations and present results to the user or direct further exploration (Bell et al. 2019). Students in our curriculum are exposed to all of these concepts, and discussion highlights the potentials, but the compressed time frame prevents giving the opportunity to get hands-on experience.

7 CONCLUSIONS

A strong correlation clearly exists between analytics and M&S. However, classical M&S education only gives a cursory coverage of the use of analytics, focusing on input modeling and output analysis at very rudimentary levels. This paper has highlighted how the various elements of analytics can support the M&S process. This makes the case that a strong analytics education makes for a better simulationist. Knowledge of not only data analytics but predictive analytics opens further opportunities to understand and use data better. Data modeling and management create “higher quality” data, thus improving simulation quality.

These relationships between analytics and M&S have been used in the development of an analytics/M&S curriculum in support of NAVSEA, in particular the naval shipyards. However, the curriculum is transferable to numerous application domains, particularly in business and manufacturing. A problematic component of the course development was identifying appropriate example data sets that were meaningful and pertinent to the course participants, particularly difficult when working with secret or proprietary data. However, experience has shown the more relevant the data sets, the better the learning experience. This is seen as the largest hurdle in extending the curriculum to alternative application domains.

It is desirable to now evaluate the benefit of developing the analytic/M&S capability to replace the classic Microsoft Excel based analytics. Merely providing knowledge of and access to alternative tools such as RStudio should greatly enhance analysts' productivity. But extending analytics and M&S knowledge should allow them to consider solutions not previously possible.

ACKNOWLEDGMENTS

This work was supported, in part, by the United States of America's Naval Sea System Command [grant number GS-10F-097CA].

REFERENCES

Adra, H. 2016. Realtime Predictive and Prescriptive Analytics with Real-time Data and Simulation. In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I.Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick 3646-3651. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Aalst, W. van der. 2015. "Business Process Simulation Survival Guide". In *Handbook on Business Process Management 1: Introduction, Methods and Information Systems*, edited by J. Vom Brock and M. Rosemann. 337-370. Berlin: Springer-Verlag.

Balci, O. 1986. "Credibility Assessment of Simulation Results". In *Proceedings of the 18th Conference on Winter Simulation*. Edited by J. Wilson, J. Henriksen, and S. Roberts. 38-44. New York, New York: Association for Computing Machinery.

Balci, O. 1998. "Verification, Validation and Testing". In the *Handbook of Simulation*, edited by J. Banks. New York: John Wiley & Sons.

Bell, D., D. Groen, N. Mustafee, J. Ozik, and S. Strassburger, 2019. "Hybrid Simulation Development–Is It Just Analytics?". In *Proceedings of the 2019 Winter Simulation Conference,* edited by N. Mustafee, K.-H.G. Bae, S.Lazarova-Molnar, M. Rabe, C.Szabo, P. Haas, and Y.-J. Son, 1352-1365. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc..

Bergmann, S., N. Feldkamp, and S. Strassburger. 2017. "Emulation of Control Strategies through Machine Learning in Manufacturing Simulations". *Journal of Simulation* 11(1): 38-50.

Bouffard, S. C., P. Boggis, B. Monk, M. Pereira, K. Quan, and S. Fleming (2018). "Discrete-event simulation modeling unlocks value for the Jansen potash project." *Interfaces* 48(1): 45-56.

Box, G., W. Hunter, and J. Hunter. 1978. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. Edited by R. Bradley, D. Kendall, J. S. Hunter, and G. Watson. Wiley Series in Probability and Mathematical Statistics. New York: Wiley & Sons.

Bryan, C., Wu, X., S. Mniszewski, and K. Ma. 2015. Integrating Predictive Analytics into a Spatiotemporal Epidemic Simulation. In *IEEE Conference on Visual Analytics Science and Technology 2015*, 17-24.

Budgaga, W., Malensek, M., Pallickara, S., and Harvey, N. 2016. Predictive Analytics Using Statistical Learning, and Ensemble Methods to Support Real-time Exploration of Discrete Event Simulations. *Future Generation Computer Systems*, 56, 360-374.

Byrum, J., C. Davis, G. Doonan, T. Doubler, D. Foster, B. Luzzi, R. Mowers, C. Zinselmeier, J. Kloeber and D. Culhane 2016. "Advanced analytics for agricultural product development." *Interfaces* 46(1): 5-17.

Feldkamp, N., S. Bergmann, and S. Strassburger. 2020. "Visualization and Interaction for Knowledge Discovery in Simulation Data." In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Foote, K. D. 2018. "A Brief History of Analytics." https://www.dataversity.net/brief-history-analytics/#, accessed 19th January, 2020.

Gutiérrez-Andrade, M. Á., E. A. Rincón-García, S. G. de-los-Cobos-Silva, P. Lara-Velázquez, R. A. Mora-Gutiérrez and A. Ponsich. 2019. "Simulated Annealing and Artificial Bee Colony for the Redistricting Process in Mexico." *Interfaces* 49(3): 189-200.

Gutierrez, D. 2017. The Exponential Growth of Data. from https://insidebigdata.com/2017/02/16/the-exponential-growth-of-data/, accessed 23rd March 23, 2020.

INFORMS. 2019. Certified Analytics Professional: Candidate Handbook. Catonville, MD, The Institute for Operations Research and the Management Sciences (INFORMS).

Johns Hopkins University of Medicine 2020. Coronavirus Resource Center. https://coronavirus.jhu.edu/, accessed 14th April, 2020.

Kleijnen, J. P. 1998. Experimental design for sensitivity analysis, optimization, and validation of simulation models. In *Handbook of Simulation*, edited J. Banks. 173- 223. John Wiley and Sons.

Law, A. M., and W. D. Kelton. 1991. *Simulation Modeling & Analysis.* 3rd ed. New York: McGraw-Hill, Inc.

Leemis, L. M. 1995. Input modeling for discrete-event simulation. In *Proceedings of the 1995 Winter Simulation Conference*. edited by C. Alexooulos, K. Kang, W. R. Lilegdon, and D. Goldsman, 16-23. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lustig, I., B. Dietrich, C. Johnson and C. Dziekan. 2010. "The analytics journey." *Analytics Magazine* 3(6): 11-13.

Lynch, C. J., and S. Y. Diallo. 2015. "A Taxonomy for Classifying Terminologies that Describe Simulations with Multiple Models". In *Proceedings of the 2015 Winter Simulation Conference*. edited by L. Yilmaz, W. Chan, I. Moon, T. Roeder, C. Macal, and M. Rossetti. 1621-1632, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lynch, C. J., 2019. *A Lightweight, Feedback-Driven Runtime Verification Methodology*. Doctoral dissertation, Computational Modeling and Simulation Engineering Department, Old Dominion University, Norfolk, Virginia.

Lynch, C. J., S. Y. Diallo, H. Kavak, and J. J. Padilla. 2020. "A Content Analysis-based Approach to Explore Simulation Verification and Identify its Current Challenges". *PloS One*, 15(5): e0232929.

Miller, J., M. Cotterell, and S. Buckley. 2013. "Supporting a Modeling Continuum in Scalation: from Predictive Analytics to Simulation Modeling". In *Proceedings of the 2013 Winter Simulation Conference,* edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 1192-1202. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc..

Miranda, J., P. A. Rey, A. Sauré, and R. Weber. 2018. "Metro Uses a Simulation-Optimization Approach to Improve Fare-Collection Shift Scheduling." *Interfaces* 48(6): 529-542.

Navidi, W. C. 2008. *Statistics for Engineers and Scientists*. New York, NY: McGraw-Hill Higher Education.

Ouyang, H. and B. Nelson. 2017. "Simulation-Based Predictive Analytics for Dynamic Queueing System". In *Proceedings of the 2017 Winter Simulation Conference,* edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 1716-1727. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Powell, J. H., and N. Mustafee. "Widening requirements capture with soft methods: an investigation of hybrid M&S studies in health care." *Journal of the Operational Research Society.* 68.10 (2017): 1211-1222.

Robinson, S. 1997. "Simulation Model Verification and Validation: Increasing the Users' Confidence". In *Proceedings of the 29th Conference on Winter Simulation,* edited by S. Andradottir, K. J. Healy, D. H. Withers, and B. L. Nelson, 53-59. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Roman, P. 2005. "Garbage in, Hollywood Out". In *Proceedings of the SimtecT Simulation Conference and Exhibition*, May 10th-12th, Sydney, Australia, 1-5.

Runkler, T. A. 2016. *Data Analytics*. Wiesbaden, Springer Vieweg.

Sargent, R. 1981. "An Assessment Procedure and a Set of Criteria for Use in the Evaluation of Computerized Models and Computer-Based Modeling Tools". Final Technical Report RADC-TR-80-409. U.S. Air Force.

Schneider, T., A. P. Luuk Besselink, M. E. Zonderland, R. J. Boucherie, W. B. Van den Hout, J. Kievit, P. Bilars, A. J. Fogteloo and T. J. Rabelink (2018). "Allocating Emergency Beds Improves the Emergency Admission Flow." *Interfaces* 48(4): 384-394.

Tolk, A., S. Diallo, J. Padilla, and H. Herencia-Zapana. 2013. "Reference Modelling in Support of M&S – Foundations and Applications". *Journal of Simulation*. 7(2), 69-82.

Tufte, E. R. 1990. *Envisioning Information*. Cheshire, Connecticut: Graphics Press.

Tufte, E. R. 2006. *Beautiful Evidence.* Cheshire, Connecticut: Graphics Press.

Vernon-Bido, D., A. Collins, and J. Sokolowski (2015). Effective visualization in modeling & simulation. In *Proceedings of the 48th Annual Simulation Symposium*, edited by S. Diallo, and A. Tolk, 33-40. New York, New York: Association of Computing Machinery.

Vincent, S. 1998. "Input data analysis." In *Handbook of Simulation*, edited J. Banks. 55-91. John Wiley and Sons.

Walrave, B. 2016. "Determining Intervention Thresholds that Change Output Behavior Patterns". *System Dynamics Review*, 32(3-4), 261-278.

Wörrlein, B., S. Bergmann, N. Feldkamp, and S. Straßburger 2019. "Deep Learning Based Predication of Energy Consumption for Hybrid Simulation." In *Proceedings of the 18th ASIM Dedicated Conference on Simulation in Production and Logistics*, September 18th-20th, Chemnitz, Germany, 121-131.

AUTHOR BIOGRAPHIES

JAMES F. LEATHRUM, JR. is an Associate Professor of the Department of Computational Modeling and Simulation Engineering at Old Dominion University. He holds a Ph.D. in electrical engineering from Duke University. His email address is jleathru@odu.edu.

ANDREW J. COLLINS is an Assistant Professor at Old Dominion University in the Department of Engineering Management and Systems Engineering. He has a Ph.D. in Operations Research from the University of Southampton. His email address is ajcollin@odu.edu.

T. STEVEN COTTER is a Senior Lecturer at Old Dominion University in the Department of Engineering Management and Systems Engineering. He holds a Ph.D. in Systems Engineering from Old Dominion University. His email address is tcotter@odu.edu.

CHRISTOPHER J. LYNCH is a Lead Project Scientist at the Virginia Modeling, Analysis, and Simulation Center (VMASC) at Old Dominion University. He received his Ph.D. in Modeling and Simulation from Old Dominion University in 2019. His email address is cjlynch@odu.edu.

ROSS GORE is a Research Assistant Professor at the Virginia Modeling, Analysis, and Simulation Center at Old Dominion University. He holds a Ph.D. and MS in Computer Science from the University of Virginia. His email address is rgore@odu.edu.