# Data Driven Worker Routing for Irrigation Monitoring

**Sara Masoud**
Assistant Professor
Industrial and Systems Engineering
Wayne State University
Detroit, Michigan, 48201, USA
saramasoud@wayne.edu

**Young-Jun Son**
James J. Solberg Head and Ransburg Professor
Industrial Engineering
Purdue University
West Lafayette, IN 47907, USA
yjson@purdue.edu

## Abstract

Survival and thrift in a globally competitive market strongly depend on the improvement of production quality while monitoring the pricing for grafting nurseries. Real-time monitoring of irrigation quality is a necessary yet cumbersome task for propagation facilities. Even for indoor facilities with automatic irrigation systems, this task remains bothersome due to the large greenhouses. The main goal of this paper is to propose an irrigation monitoring system to increase production productivity by minimizing the time workers spend on responding to irrigation monitoring calls. Here, a two-stage data-driven approach is proposed, where the traveling salesman model is implied on a reduced spatiotemporal network provided by Bayesian analysis to optimize workers routing for irrigation monitoring. The proposed approach is the final piece of a dynamic data-driven simulation-based optimization framework. The results over synthetic data demonstrate a potential reduction of 11.60% in traveled distance per worker per call.

## Keywords
Traveling salesman, Bayesian modeling, Irrigation monitoring

## 1. Introduction
Survival and thrift in a global competitive market strongly depend on improvement of production quality while monitoring the pricing for grafting nurseries. Grafting, a technique that vegetatively joins the roots and the bottom portion of one plant (rootstock) and a tender shoot (scion) from the top portion of another plant, has become an imperative technique for intensive vegetable production due to providing greater yield, more desirable fruit properties, and higher resistant to pests (Lee et al. 2010). Grafting has been increasingly tapped worldwide for cultivation under adverse environs posing abiotic (e.g., low/high temperature, drought/flooding, heavy metal and nutrition) and biotic (e.g., soil borne pathogens, foliar pathogens, arthropods and weeds) stresses to vegetable crops, resulting in expansion of commercial production onto otherwise under-exploited lands (Louws et al. 2010). Compelled by the increase of organic production and restrictions on utilizing synthetic pesticides, grafting has become an even more useful tool to manage soil borne disease issues in grounds of environmental protection.

Greater yield, as one of the most prominent characteristics of grafting, opens a fruitful domain for indoors agroindustry such as greenhouse productions. Greenhouses are Hi-Tech horticulture farms, where advanced technological infrastructure are employed to increase the production efficiency through constant monitoring. The use of computational systems and technology such as sensors in agricultural applications has become viable due to the continuous reduction in their prices (Serôdio 2001). Vegetable cultivation in greenhouses is one of the main branches of agriculture in which the production management must be more controlled, so that the values of a set of parameters have to be monitored and approximated to an ideal value.

The major features that are usually monitored in greenhouses are temperature, humidity, moisture, radiation and CO2 concentration (Hanan 1998). As a result, greenhouse production remains the most intensive agricultural process known in terms of labor and capital (Hanan 1998). Consequently, the design and resource (e.g., labor, robots, space) allocation of greenhouses should be done carefully to sharpen the competitive edge of greenhouse productions by minimizing the price and ensure economic success by justifying the high capital investment. To achieve those goals, greenhouse literatures aimed at increasing the yield by optimizing climatic conditions such as radiation, temperature, moisture and gas distribution and flows (e.g., air, CO2) over the past three decades (Shklyar and Arbel 2004).

Irrigated agriculture is one of the most vital components of total agriculture which supplies many of the fruits, vegetables, and cereal foods consumed by humans. Real-time monitoring of irrigation quality is a cumbersome task for propagation facilities. Even for indoors facilities with automatic irrigation systems, this task remains bothersome due to the large size of the greenhouses. This dissertation aims to reduce the traveling time of the workers through the greenhouse by optimizing the workers' route planning, considering the probabilistic rise of irrigation monitoring request in near future. Here, a two-stage framework is proposed, where the allocation of labor takes place considering the possible rise for irrigation monitoring need with respect to the availability of the workers and their real-time location. Here, a radio frequency identification system is utilized to track workers' location, and radio frequency identification enabled sensors are embedded within the soil to wirelessly report the moisture level.

Given the importance of irrigation scheduling and monitoring systems, the main goal is to reduce the traveling time of the workers through the greenhouse by optimizing the workers' route planning, considering the probabilistic rise of irrigation needs. Real-time monitoring of irrigation is a cumbersome task for propagation facilities even for indoors facilities with automatic irrigation systems due to the large size of the greenhouses. Recent progress in sensors technologies has led to automated irrigation scheduling systems due to reduced sensor price, easy implementation, and low maintenance needs. As a result, several irrigation scheduling methods such as soil water content or matric head monitoring, plant stress monitoring, water balance, computer models or charts (Cahn and Johnson, 2017) have been developed based on sound information systems. These systems can measure various environmental factors such as humidity and soil moisture (Nuñez-Olivieri et al. 2017).

Many studies have shown that the development of decision support systems based on sensor networks could increase the efficiency of irrigation water and as a result increase productivity (e.g., Paris et al. 2018). Several researchers have investigated developing decision support tools for optimizing irrigation strategies. Geerts et al. (2010) utilized historical climate data to derive optimal frequencies of irrigation to avoid drought stress and guarantee maximum water productivity. Garcia-Vila and Fereres (2012) investigate the impact of imposing water quotas or increasing water prices by combining AquaCrop with an economic model. Shang and Mao (2006) developed a simulation-based optimization framework based on crop water production functions to produce the optimal winter wheat irrigation series in North China.

Li et al. (2018) proposed a simulation-based optimization approach for crop optimal irrigation scheduling under uncertainty was developed to maximize the net benefit considering the uncertainties of water pricing. Wen et al. (2017) analyzed the optimal spring wheat irrigation schedules under plastic mulching utilizing a simulation-based optimization model by coupling water balance model, crop water production functions and optimization model.

Although many efforts have been placed for irrigation scheduling, crop water production functions were traditionally obtained from long-term field experiments, which are site-specific, expensive and time-consuming. In addition, AquaCrop simulation is suitable for open field irrigation modelling and fails to model irrigation for greenhouse production. The resulting outcomes of such empiricisms are therefore unlikely to be sufficient and transferable to greenhouse production. At the same time, physical confirmation is mandatory for greenhouse production. To the best of our knowledge, this is the first work to address the routing of workers in greenhouses for irrigation monitoring and our current work is therefore an effort at closing this knowledge gap. The framework increases workers productivity by predicting and clustering the similar irrigation monitoring request into one trip, and then find the optimum route plan that minimizes workers traveled distance while insuring that all the irrigation monitoring requests have been satisfied.

© IEOM Society International

## 2. Literature Review

Dynamic Data Driven Applications Systems (DDDAS) is a relatively new paradigm, where the computation and instrumentation aspects of application systems are dynamically integrated (Darema, 2004). This integration can dynamically incorporate instrumentation data into the executing model of the application, which in return enables more accurate and faster modeling and analysis of the behaviors of application systems. As a powerful tool, DDDAS has been perceived to implement effective measurement processes in a variety of application areas (Darema, 2008). In other words, DDDAS can exploit data in intelligent ways to convert them to new capabilities such as decision support systems with the accuracy of full-scale modeling, efficient data collection, management, and data mining (Blasch and Phoha, 2017). DDDAS concept entails the ability to incorporate data into an executing application simulation dynamically. At the same time, DDDAS enables applications to dynamically steer measurement processes. Utilizing dynamic data inputs in the system either in real-time mode or archival, promises the improvement of modeling methods, prediction accuracy of simulations, efficiency of simulations, and the effectiveness of measurement systems.

During the past decades, the DDDAS paradigm has been widely discussed in literature. A false prediction of track as well as magnitude of a storm by meteorologists, and a simulation failure in behavior modelling of wildfire in Los Alamos National Laboratory and its propagation raised the motivation behind DDDAS (Darema et al. 2005). The key characteristic of DDDAS is not only its capabilities to dynamically incorporate data into simulation, but also the empowering of applications to dynamically affect the environmental data (Darema, 2005). DDDAS has been successfully implemented in many different fields such as contaminant tracking (Douglas et al. 2004), natural disaster forecasting (Patrikalakis et al. 2004), shop floor control (Son et al. 2002), supply chains (Celik et al. 2010), patrol control (Khaleghi et al. 2014), and many more. Douglas et al. (2004) introduced numerical procedures for multi-scale interpolation to map sensor data in order to continuously update the simulation of contaminant tracking. Online data acquisition and filtering control framework were proposed by Mandel et al. (2004) to recognize out-of-order data for wildfire control. The challenges of automatically adapting simulations were first brought into attention by Carnahan and Reynolds (2006) in supply chains. To address the mentioned challenges, Carnahan and Reynolds proposed a semi-automated approach to exploit the flexibility and constraints of model abstraction to automate simulation adaptation. Although automatically generalizing adaptation for anticipating all possible ways is difficult, but this problem can be simplified by taking advantage of the flexibilities and constraints of simulation modeling (Parnas, 1979).

While flexibility is required for automatic adaptation in order to find proper alternatives, constraints limits the number of feasible alternatives to be taken into account. Although manual modification of optimization methods is not a requirement for Carnahan and Reynolds proposed approach, but human intervention is required to determine the most likely alternatives. In another instance, Patrikalakis et al. (2004) provided an overview of a DDDAS for rapid adaptive interdisciplinary ocean forecasting, where an Internet-based distributed system enabled remote observations and data assimilation for effective estimation of uncertainties in oceanic fields. A simulation-based shop floor planning and control system was proposed by Son et al. (2002), where the same simulation model covered planning stage (in fast mode) as well as control stage (real-time). Celik et al. (2010) extend Son's framework by proposing a comprehensive system architecture and methodologies such as abnormality detection, fidelity selection, fidelity assignment, and prediction and task generation in real-time DDDAMS based semi-conductor supply chain management. More recently, Khaleghi et al. (2014) proposed an effective and efficient surveillance and crowd control via unmanned aerial and ground vehicles, empowered by agent-based simulation, vision and filtering algorithms, and motion planning by graph search algorithms. As another application, Sang et al. (2018) proposed a DDDAS inventory management approach to optimize automated refreshment systems in a retail environment.

In recent years, not only the grid computing technologies provide advanced computational capabilities, but also time measurement infrastructures such as instruments and sensor systems, data storage technologies, and remote data access have matured. As a result, signal processing techniques are applied to big data analysis to spawn developments in designs, algorithms, architectures, and applications. These improvements empower the capabilities of DDDAS and call for novel application modelling approaches and interfaces to measurement systems, mathematical and statistical algorithms tolerant to perturbations from dynamic data inputs, and systems software to support the dynamic resource requirements of such applications.

Application modelling, mathematical and statistical processing, measurement systems, and systems software design area the four major interactive concepts of DDDAS. The planning updates usually arise from statistical data analysis within the measurement units. New insights have been developed for measurement systems with unobservable data, such as new architectures to emulate data collections for missing data. To complement the modelling and statistical analysis, software methods are paramount for real-time applications. Data-driven simulation models as defined in are designed to be applicable to systems with same structures but different parameters' value. Data-driven modelling approach guarantees low development cost and time in comparison to typical modelling approach due the independency of software utilization and maintenance of model development.

Maximizing productivity is the main goal of irrigation scheduling through providing an optimum water supply by applying the right amount of water to the crops at the right time (Soulis and Elmaloglou, 2018). However, efficient irrigation management is challenging due to factors such as irrigation system characteristics, soil moisture level and crop specifications (Dabach et al. 2013). What makes the irrigation management even more complex for grafting nurseries is the requirement of physical confirmation of smooth irrigation, as abiotic stress have a higher impact on the agronomic growth of greenhouse products (Rodriguez-Ortega, 2017). In this paper, a two stage data driven approach is proposed to minimize the distance traveled by workers for responding to irrigation monitoring calls, where the first stage utilizes Bayesian analyses in order to predict the potential rise of irrigation requests and the second stage in which a TSP model is solved to minimize the travelling of workers. Here, a complete literature review on the core elements of the stages one and two (i.e., Bayesian analysis and travelling salesman problem, respectively) is provided.

## 2.1. Bayesian Analysis

Bayesian analysis is the core element incorporated within first stage of the proposed approach. Bayes and Price (1763) first introduced three essential elements underlying Bayesian statistics as the prior knowledge, the information within the observed data, and the posterior inference, where prior knowledge refers to all knowledge available on model parameters prior seeing the data, the information within the observed data is the observed evidence, which is usually expressed in terms of the likelihood function, and the posterior inference is the combination of the first two elements through Bayes' theorem. Bayesian techniques have the highest impact when data are generated by repeated execution of the same type of random experiment. The main advantage of Bayesian approaches is the fact that they allow researchers to incorporate background knowledge into their analyses and avoid ignoring the lessons of previous studies in contrast to statistical methods (Asendorpf et al. 2013). As a result, the plausibility of previous research findings can be evaluated in relation to new data, which makes the proposed method an interesting tool for confirmatory strategies.

Given this advantage, the interfaces of applied decision-making have renewed interest in model development and prediction accuracy of Bayesian models. Recently, Bayesian analysis have been applied in different applications such as macroeconomics, finance, Psychopathology, environmental and resource management (Aastveit et al. 2018; Pettenuzzo and Ravazzolo, 2016). This growing body of research is a partial response to the need of improving information flows into policy and decision making at different levels.

## 2.2. Travelling Salesman Problem

The spatial analysis of the proposed irrigation monitoring approach has been modeled as a travelling salesman problem. One of the most popular and hardest problems from the NP set is the Travelling Salesman problem (TSP). As a solution, TSP has been widely applied in a variety of practical fields (Goyal 2010).

TSP can either be defined combinatorial optimization problem, where the minimum Hamiltonian cycle in a graph of cities is of interest, or the decision problem, in which the existence of a weighted hamiltonian cycle in a graph is of question. Although both of these representations belong to NP family, the combinatorial optimization version belongs to the NP Hard set, whereas the decision version belongs to the class of NP complete problems. As result, it can only be assumed that there is no exact algorithm for solving any version of the TSP in polynomial time. Ever since TSP was coined in 1930, many efforts have been made to solve the travelling salesman problem either as a deterministic approach aiming for exact solution, or heuristic approaches to provide a solution with a near minimal cost.

Dantzig et al. (1954) utilized linear programming relaxation techniques to solve the integer formulation through adding well- chosen linear inequality to the list of constraints continuously. The exponential growth in solution time with the increase of the input network inspired researchers to look for heuristic approach to find near optimal solution fast. For example, genetic and ant colony algorithms has been improved, revised and applied to solve TSP as a combinatorial problem (Duan and AI, 2016; Khushboo and Mamta ARORA, 2016).

## 3. Methods

Figure 1 displays the DDDASO framework proposed for real-time planning and control of the facility under study. The framework consists of three main units which are called the real system, measuring unit, and planning unit. The real system consists of the material, workers, management, computation units, and the sensors which are implemented to observe the behaviour of the system under study.

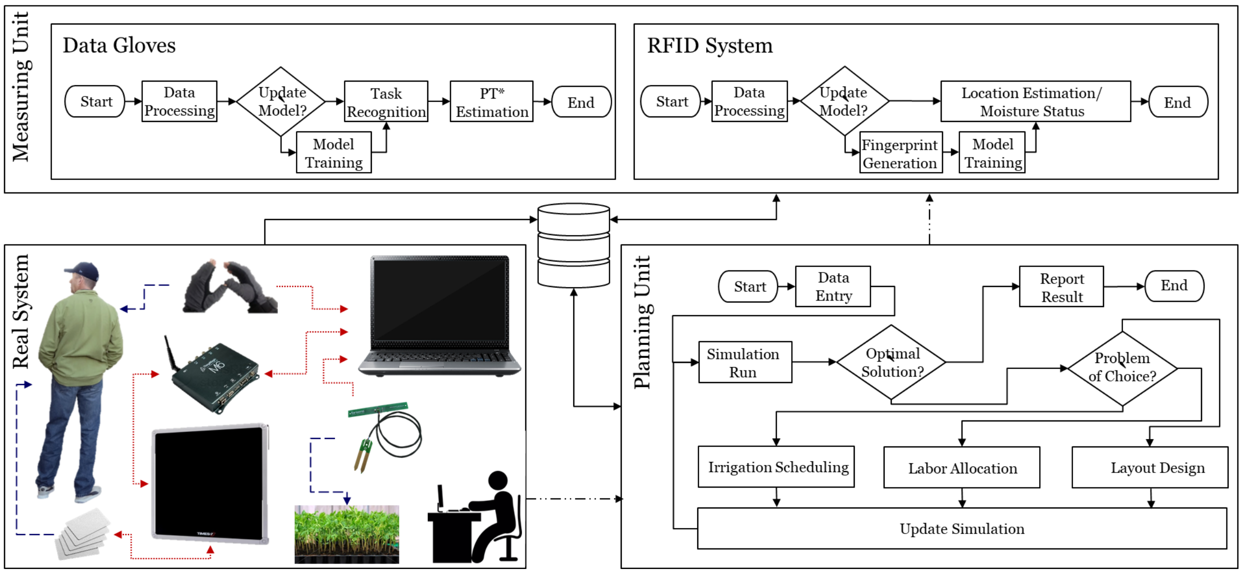


Figure 1. Dynamic data driven adaptive simulation-based optimization framework (Masoud et al. 2019).

The measuring unit handles the streams of sensory data and the estimation of processing time and material handling time of the system under study, details of which are in our previous works (Masoud et al, 2019; Chowdhury et al. 2020; and Chowdhury et al. 2021). In addition, the implementation of DDDAS paradigm in this framework and the fidelity level optimization of the dynamic decision making is discussed here (Masoud et al. 2019). The planning unit relies on the concept of simulation-based optimization where simulation mimics the performance of the system under study through the estimated parameters provided by the measuring unit, and optimization models look through different scenarios to find the optimal layout design (Masoud et al. 2019) and labour management (Masoud et al. 2018). In this work, we will discuss the irrigation monitoring aspect which will finalize this DDDAS framework.

The proposed data driven irrigation monitoring approach is designed to minimize the workers travelling time within the seedling propagation facilities by aggregating potential irrigation monitoring request in a single trip in each time window. The proposed approach relies on the concept of mean time between irrigation monitoring request (MBIMR). MBIMR is defined as the average time between calling for irrigation monitoring, as shown in Figure 2. As displayed in Figure 2, the RFID moisture sensors signals an irrigation monitoring requests whenever the soil moisture level hits a predefined threshold. This threshold is designed considering a $\Delta t_{\text{Lead}}$ to provide enough time for the workers to walk to where the sensor is located and monitor the quality of irrigation. Within the proposed approach, temporal and spatial analyses are the main elements of optimizing irrigation monitoring problem. The temporal analysis covers fitting prior distributions for MBIMR based historical data and employing Bayesian rule to estimate the posterior distributions of the said variable, in order to predict the potential MBIMR request in near future and provide appropriate response in a single trip. Anytime a request for irrigation monitoring is signaled, the status of the system is updated for all the stationed RFID moisture sensors based on their fitted distributions and their last request timestamp.

### 3.1. Temporal Analysis Unit

The temporal analysis heavily relies on Bayesian rule, where prior information and data are combined through a solid and principled decision theoretical framework to provides interpretable answers. To deal with the high computational cost that is associated with Bayesian analysis, the prior distribution selection has been restricted to conjugate families. This restriction helps to avoid the estimation of the marginal distribution, $\int p(z_k \mid x_k) p(x_k \mid z_{1:k-1}) dx_k$ , and provides a closed form formula for estimating the expected MBIMR of each imbedded moisture sensor. Conjugate prior is a family of probability distributions characterized by some parameter θ, where all possible posterior distributions are also members of the same family. This family remains closed under evidence and contains the posterior, which leads to a large amount of algebraic simplicity in the computation of the posterior. The temporal analysis unit keeps track of the distribution of MBIMR for each imbedded soil moisture sensor as well as their latest request timestamp. As a result, whenever a new signal is received, the temporal analysis unit defines a set of soil moisture sensor nodes, whose expected MBIMR are within the $2\Delta t_{\text{Lead}}$ time interval. Then, this set is reported to the spatial analysis unit as a reduced spatiotemporal network to minimize the travelling time of the worker who performs the monitoring.

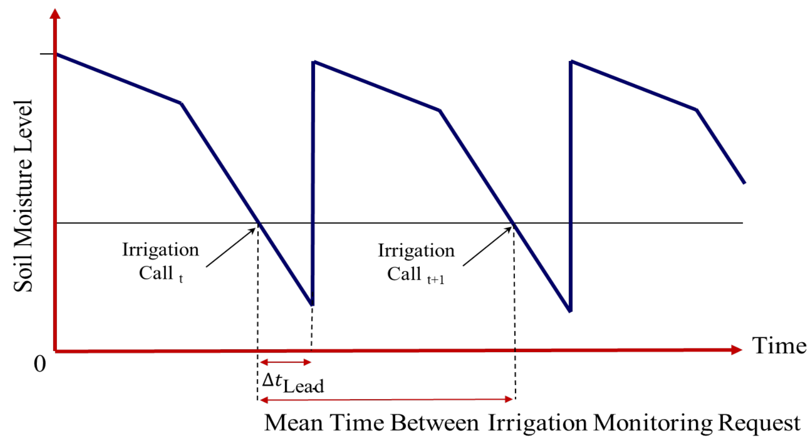


Figure 2. Definition of mean time between irrigation monitoring requests

## 3.2. Spatial Analysis Unit

Given a set of nodes including the coordinates of the imbedded soil sensors reported by the temporal unit and current position of a worker's station, and distance between every pair of the coordinates, the spatial unit tries to find the shortest possible route that visits every node exactly once and returns to the workers station, which can be represented as a travelling salesman problem. The Travelling Salesman problem is one of the most popular problems in operation research and computer science communities. The Travelling Salesman Problem, first formulated as a mathematical problem in 1930, is a problem in combinatorial optimization, where given a list of cities and their pair wise distances, the task is to find the shortest possible tour that visits each city exactly once.

Given a complete graph $G(N_t, D_t)$, the edge distances (i.e., $D_t$) between $N_t$ nodes are defined as $d_{ijt}$ at time $t$ where $i, j = 1 \dots n$. A trip to cover all the nodes of interest can be represented by a cyclic permutation $\pi_t$ of $\{1, 2, \dots, n\}$ where $\pi_t(i)$ represents the city that follows city $i$ on the trip. The goal is to find the optimum permutation $\pi_t$ that minimizes to traveled distance in trip $t$, defined as $\sum_{i=1}^{n} d_{i\pi_t(i)}$. Alternatively, the TSP can be defined by following Miller-Tucker-Zemlin formulation (Desrochers and Laporte, 1991).

$$Min \sum_i \sum_j d_{ijt} x_{ij} \qquad \text{(Eq. 1)}$$

Subject to:

$$\sum_{i, i \neq j} x_{ij} = 1 \qquad j = 1,2, \dots, |N_t| \qquad \text{(Eq. 2)}$$

$$\sum_{j, i \neq j} x_{ij} = 1 \qquad i = 1,2, \dots, |N_t| \qquad \text{(Eq. 3)}$$

$$u_i - u_j + n x_{ij} \leq n - 1 \qquad 2 \leq i \neq j \leq |N_t| \qquad \text{(Eq. 4)}$$

$$0 \leq u_i \leq |N_t| - 1 \qquad 2 \leq i \leq |N_t| \qquad \text{(Eq. 5)}$$

$$x_{ij}, \in \{0, 1\} \qquad \forall\, i, j \in \{1,2, \dots, |N_t|\} \qquad \text{(Eq. 6)}$$

where $x_{ij}$ (i.e., decision variable) defines whether the trip passes through nodes $i$ and $j$, and $u_i$ is a dummy variable. Equation 1 minimized the total distance traveled at time $t$. Equation sets 2 and 3 are defined to require that each node has arrival from exactly one other node and one departure to exactly one other node. Equation 4 enforce a single trip to cover all nodes instead of multiple disjointed trips. To find the exact solution of the formulated problem with $|N_t|$, a comparison of $(|N_t|-1)!$ feasible solutions is required. To evaluate all possible tours is infeasible for even small TSP instances. Although as a NP complete problem, the time required to solve the problem increases rapidly as $|N_t|$ grows, the spatiotemporal network reported by the temporal unit is small enough to be optimized using the existing methods. Here, this problem has been optimized via TSP package provided by Hahsler and Hornik (2019), where routes are weighted as the Euclidean distance between nodes. The Euclidean TSP is a special case of the TSP, where the distance between each two nodes (e.g., i and j) are defined as $dist(x, y) = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}$ for the network under study.

## 4. Data Collection

To test the performance of the proposed irrigation monitoring framework, simulated data have been utilized. The simulated data includes the location of workers within the facility and the water consumption of the plants. The simulation models used in the DDDAS framework has been validated in these works (Masoud et al. 2018; Masoud et al. 2019).

## 5. Results and Discussion

For the design of experiment, the optimal layout design proposed by (Masoud et al. 2019) has been incorporated, where 128 RFID soil moisture are simulated within the system to report the soil moisture level of each growing bench as a number between 0 to 100. The network under study consists of 120 monitoring nodes within the greenhouse area and 30 working stations nodes in the headhouse, respectively. For the mentioned layout, 5 different irrigation patterns have been randomly simulated within the system that resulted in MBIMR fitted distributions displayed in Figure 3.

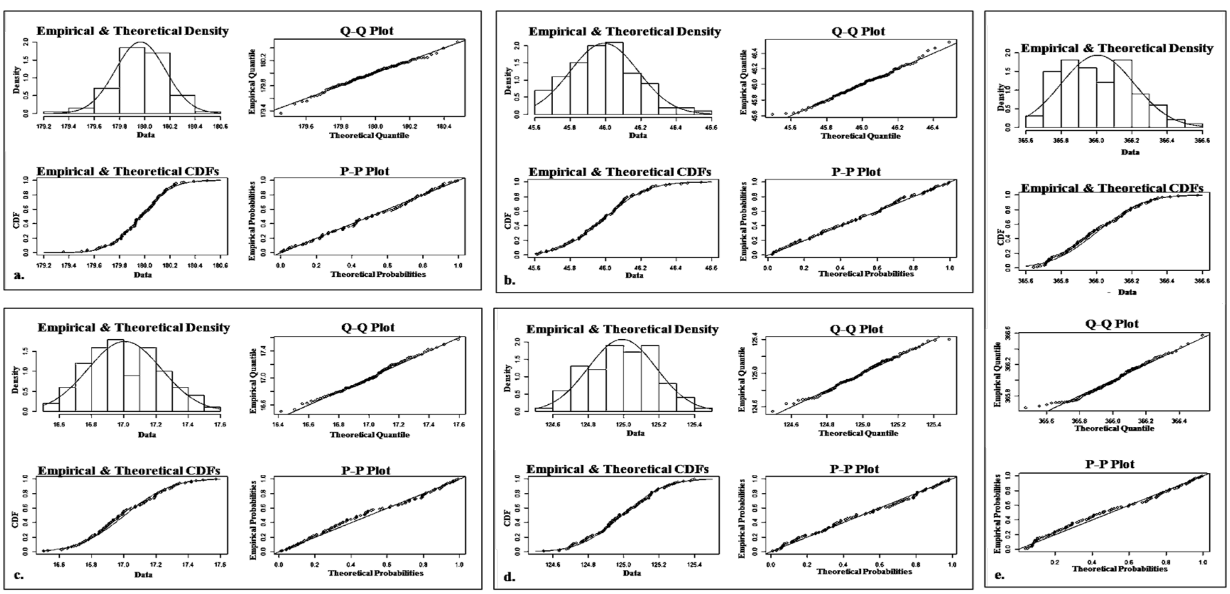

Figure 3. Fitted normal distributions of MBIMR under 5 different randomly simulated scenarios.

Anderson-Darling goodness-of-fit statistics are computed for the fitted normal distribution for all five sets of simulated data. The tests fail to reject the normal distribution at the significance level 0.05 for all fitted distributions due to having p-values of 0.31, 0.18, 0.42, 0.63, and 0.25.

These fitted distributions are utilized as the conjugate prior distributions for the Bayesian analysis. Another important parameter that needs to be defined is the $\Delta t_{\text{Lead}}$. $\Delta t_{\text{Lead}}$ threshold is defined as the travelling time in which the mminimal distance travelled trip to cover all nodes, as shown in Figure 4. The displayed trip in Figure 4 has a length

© IEOM Society International

of 503.16 meters. Considering the minimum humans walking speed of 0.5 meter per second as defined in Chapter 6, $\Delta t_{\text{Lead}}$ is defined as 503.16/0.5 = 1,006.32 seconds, which rounds up 17 minutes.

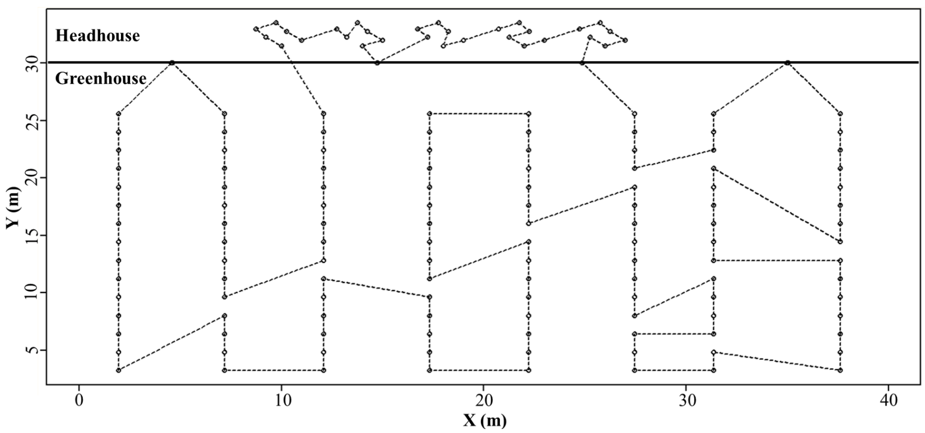


Figure 4. The minimal distance travelled trip to cover all nodes.

By implementing the temporal analysis unit, considering a time window of 34 minutes, the following MBIMR request have been aggregated to be attend in six single trips with a randomly assigned as described in Table 1.

Table 1. The proposed aggregated Trips

| Timestamp | Assigned Worker | Nodes to be Covered | Total Traveled (m) |
|---|---|---|---|
| 1 | 127 | 53, 50, 49, 46, 55, 44 | 79.72 |
| 2 | 133 | 37, 35, 39, 44, 42, 33 | 80.165 |
| 3 | 122 | 3, 6, 11 | 48.81 |
| 4 | 122 | 20, 25, 29, 92, 100, 105 | 131.73 |
| 5 | 142 | 100, 110, 119 | 76.82 |
| 6 | 135 | 61, 63, 72, 76, 87, 90 | 95.80 |

Table 1 describes the nodes that are covered within each trip, the worker assigned to monitor irrigation quality in each trip, and the total distance travelled in each trip. Routes that guarantee the minimum distance travelled for each trip are displayed in Figure 5.

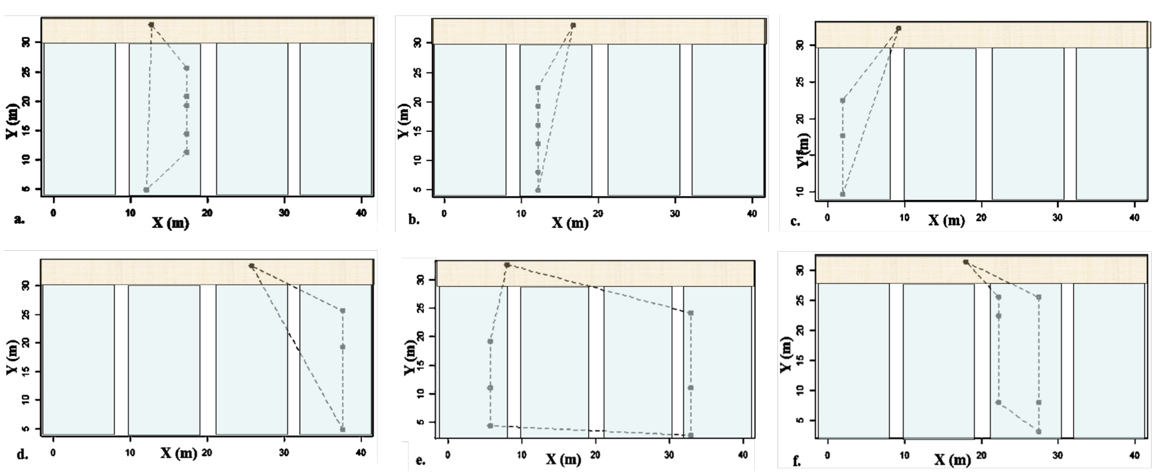


Figure 5. The proposed Categorization of Request

In contrast, we simulate the case where workers are not aware in time, and as a result each irrigation monitoring should be handled in a separate trip. For this case, Floyed-Warshall algorithm (Burfield 2013) is applied to find the shortest trip between the assigned workers and their destination. The results are displayed in Table 2.

Table 2. Separate trips in response to irrigation monitoring requests

| Assigned Worker | Covered Nodes | Traveled Distance (m) | Total Traveled (m) |
|---|---|---|---|
| 127 | 44, 46, 49, 50, 53, 55 | 56.37, 17.36, 26.18, 29.04, 38.27, 44.51 | 211.57 |
| 133 | 33, 35, 37, 39, 42, 44 | 23.11, 29.08, 35.21, 41.42, 50.82, 57.12 | 236.77 |
| 122 | 3, 6, 11 | 24.47, 32.69, 47.55 | 104.72 |
| 122 | 20, 25, 29, 92, 100, 105 | 26.38, 42.26, 55.02, 47.22, 61.06, 73 | 304.951 |
| 142 | 100, 110, 119 | 48.54, 42.71, 65.57 | 156.82 |
| 135 | 61, 63, 72, 76, 87, 90 | 14.5, 20.04, 47.72, 22.28, 50.63, 59.64 | 214.81 |

Table 2 displays the travelling distances, in a case where the workers are not aware of incoming requests, and cover each request in a separate trip based on Floyed-Warshall shortest routes. Each row in Table 2 reports the set of nodes and their associated travelling distance that each worker needs to cover. For example, Worker 122 travels 24.47 m, 32.69 m, and 47.55 m to check the irrigation quality on nodes 3, 6, and 11, respectively which leads to total traveled distance of 104.72 m. By comparing the Tables 1 and 2, it can be noted workers are spending more time on responding to the same irrigation request without utilizing the proposed approach. To normalize the impact of the random assignment of the workers, all workers are assigned to cover the same set of nodes. The normalized results are displayed in Figure 6.

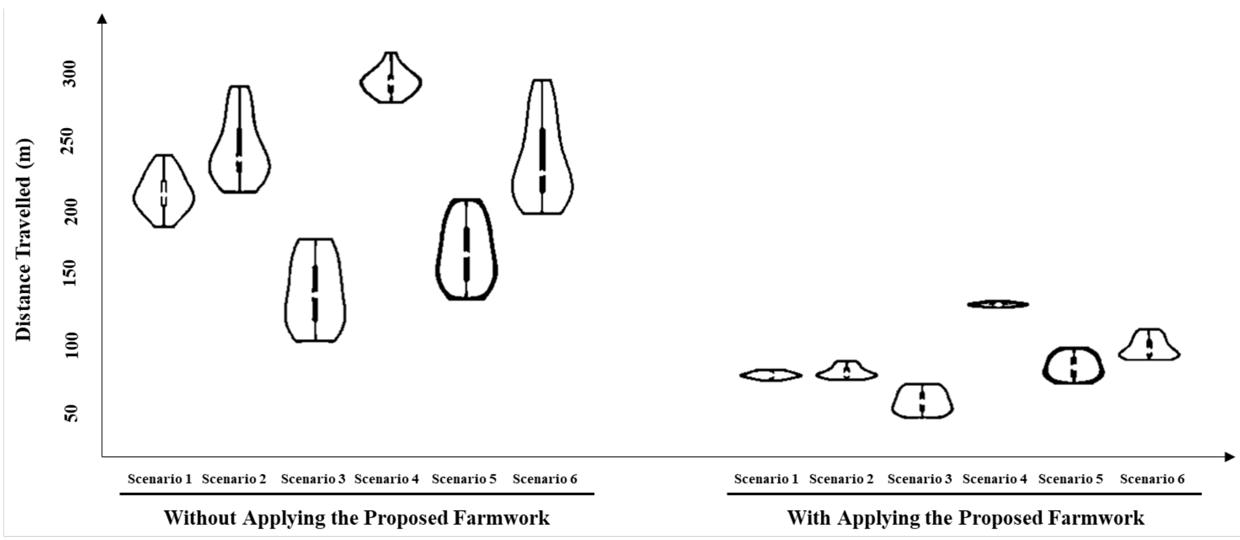


Figure 6. Comparison the response time for each scenario

## 6. Conclusion

Although real-time monitoring of irrigation systems is a cumbersome task for indoor propagation facilities but is necessary for guaranteeing high quality production in vegetable seedling propagation facilities such as grafting nurseries. In this paper, a two stage data driven approach is proposed to minimize the distance traveled by workers for responding to irrigation monitoring calls. The temporal analysis unit (i.e., the first stage) studies MBIMR for the soil moisture in order to predict the potential rise of irrigation requests and clusters the requests that are close timewise as simplified spatiotemporal networks. Then, those spatiotemporal clusters are reported to the spatial analysis unit (i.e., stage two), where a TSP model minimizes the travelling of workers while guaranteeing visiting all nodes reported by stage one. Preliminary experiments with simulated data have been conducted whose results display a potential saving of 25. 18 m (11.60%) in terms of travelled distance per worker per call.

## References


Aastveit, K. A., Ravazzolo, F., and Van Dijk, H. K. Combined density nowcasting in an uncertain economic environment. *Journal of Business & Economic Statistics*, vol. 36, no. 1, pp. 131-145, 2018.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., Fiedler, S., Funder, D.C., Kliegl, R., Nosek, B.A. and Perugini, M., Recommendations for increasing replicability in psychology. *European Journal of Personality*, vol. 27, no. 2, pp. 108-119, 2013.

Blasch, E., and Phoha, S. Dynamic Data-Driven Applications Systems (DDDAS) Concepts in Signal Processing. *Journal of Signal Processing Systems*, vol. 88, no. 2, pp. 203-204, 2017.

Cahn, M. D., and Johnson, L. F. New approaches to irrigation scheduling of vegetables. Horticulturae, vol. 3, no. 2, pp. 28, 2017.

Carnahan, J. C., and Reynolds, P. F., Requirements for DDDAS flexible point support. *In Proceedings of the 2006 Winter Simulation Conference*, pp. 2101-2108, Monterey, USA, December 3-6, 2017.

Celik, N., Lee, S., Vasudevan, K., and Son, Y. J. DDDAS-based multi-fidelity simulation framework for supply chain systems. *IIE Transactions*, vol. 42, no. 5, pp. 325-341, 2010.

Chowdhury, B. D. B., Masoud, S., Son, Y. J., Kubota, C., and Tronstad, R. A dynamic data driven indoor localisation framework based on ultra high frequency passive RFID system. *International Journal of Sensor Networks*, vol. 34, no. 3, pp. 172-187, 2020.

Chowdhury, B. D. B., Masoud, S., Son, Y. J., Kubota, C., and Tronstad, R. A dynamic HMM-based real-time location tracking system utilizing UHF passive RFID. *IEEE Journal of Radio Frequency Identification*, vol. 6, no. 1, pp. 41-53, 2021.

Dabach, S., Lazarovitch, N., Šimůnek, J., and Shani, U. Numerical investigation of irrigation scheduling based on soil water status. *Irrigation Science*, vol. 31, no. 1, pp. 27-36, 2013.

Dantzig GB, Fulkerson DR, Johnson SM. Solution of a Large-scale Traveling Salesman Problem. *Operations Research*, vol. 2, no. 1, pp. 393410, 1954.

Darema, F. Dynamic data driven applications systems: A new paradigm for application simulations and measurements. *In International Conference on Computational Science*, pp. 662-669, Berlin, Heidelberg, June 6–9, 2004.

Darema, F. Dynamic data driven applications systems: New capabilities for application simulations and measurements. *In International conference on computational science*, pp. 610-615, Berlin, Heidelberg, June 6–9, 2004.

Darema, F, Dynamic data driven applications systems (dddas)–a transformative paradigm, *International Conference on Computational Science*, Berlin, Heidelberg, June 23-25, 2008.

Desrochers, M., and Laporte, G. Improvements and extensions to the Miller-Tucker-Zemlin subtour elimination constraints, *Operations Research Letters*, vol. 10, no. 1, pp. 27-36, 1991.

Douglas, C. C., Shannon, C. E., Efendiev, Y., Ewing, R., Ginting, V., Lazarov, R., Cole, M.J., Jones, G., Johnson, C.R. and Simpson, J., A note on data-driven contaminant simulation, *In International Conference on Computational Science*, pp. 701-708, Berlin, Heidelberg, June 6–9, 2004.

Duan, C., Yang, L., and Liu, Y. Accurate spatial calibration of RFID antennas via spinning tags, *IEEE 36th International Conference* on *Distributed Computing Systems*, pp. 519-528, Nara, Japan, June 27-30, 2016.

Burfield, C. Floyd-warshall algorithm. Massachusetts Institute of Technology, 2013.

García-Vila, M., and Fereres, E. Combining the simulation crop model AquaCrop with an economic model for the optimization of irrigation management at farm level. *European Journal of Agronomy*, vol. 36, no. 1, pp. 21-31, 2012.

Geerts, S., Raes, D., and Garcia, M, Using AquaCrop to derive deficit irrigation schedules. *Agricultural water management*, vol. 98, no. 1, pp. 213-216, 2010.

Goyal, S. A survey on travelling salesman problem. *In Midwest instruction and computing symposium*, pp. 1-9, Eau Claire, USA, April 16-17, 2010.

Hahsler M. and Hornik K, TSP: Traveling Salesperson Problem (TSP). R package version 1.1-7. https://CRAN.R-project.org/package=TSP, 2010.

Hanan, N. P., Kabat, P., Dolman, A. J., and Elbers, J. A. Photosynthesis and carbon balance of a Sahelian fallow savanna, *Global change biology*, vol. 4, no. 5, pp. 523-538, 1998.

Khaleghi, A. M., Xu, D., Minaeian, S., Li, M., Yuan, Y., Liu, J., Son, Y.J., Vo, C. and Lien, J.M., A DDDAMS-based UAV and UGV team formation approach for surveillance and crowd control. *In Proceedings of the 2014 Winter Simulation Conference*. pp. 2907-2918 Savannah, USA, December 77-10, 2014.

Arora K. and Arora M., Better result for solving TSP :GA vs. ACO, *International Journal of Advanced Research in Computer Science and Software*, vol. 6, no. 1, pp. 219-224, 2016.

Lee, J. M., Kubota, C., Tsao, S. J., Bie, Z., Echevarria, P. H., Morra, L., and Oda, M. Current status of vegetable grafting: Diffusion, grafting techniques, automation. *Scientia Horticulturae*, vol. 127, no. 2, pp. 93-105, 2010.

Lenstra, J.K., and RinnooyKan, A.H.G., Some simple applications of the traveling salesman problem, *Operational Research Quarterly*, Vol. 26, no. 2, pp. 717–33, 1975.

Li, J., Song, J., and Shang, S., Simulation-based optimization of irrigation scheduling for maize in arid region to increase field application efficiency. *In 2018 ASABE Annual International Meeting*, pp. 1, Detroit, USA, July 29-August 1, 2018.

Louws, F. J., Rivard, C. L., and Kubota, C. Grafting fruiting vegetables to manage soilborne pathogens, foliar pathogens, arthropods and weeds. *Scientia Horticulturae*, vol. 127, no. 2, pp. 127-146, 2010.

Mandel, J., Chen, M., Franca, L. P., Johns, C., Puhalskii, A., Coen, J. L., Douglas, C.C., Kremens, R., Vodacek, A. and Zhao, W., A note on dynamic data driven wildfire modeling. *In International Conference on Computational Science,* pp. 725-731, Kraków, Poland, June 6-9, 2004.

Masoud, S., Chowdhury, B., Son, Y. J., Kubota, C., and Tronstad, R. A dynamic modelling framework for human hand gesture task recognition. *arXiv preprint* arXiv:1911.03923, 2019.

Masoud, S., Chowdhury, B. D. B., Son, Y. J., Kubota, C., and Tronstad, R. Simulation based optimization of resource allocation and facility layout for vegetable grafting operations. *Computers and Electronics in Agriculture*, vol. 163, no. 1, pp. 104845, 2019.

Masoud, S., Son, Y. J., Kubota, C., and Tronstad, R. Evaluation of simulation-based optimization in grafting labor allocation. *Applied engineering in Agriculture*, vol. 34, no. 3, pp. 479-489, 2018.

Masoud, S., Chowdhury, B., Son, Y. J., and Tronstad, R. Markov-chain Monte-Carlo Sampling for Optimal Fidelity Determination in Dynamic Decision-Making. *arXiv preprint arXiv:1911.03719*, 2019.

Nuñez-Olivieri, J., Muñoz-Barreto, J., Tirado-Corbalá, R., Lakhankar, T., and Fisher, A. Comparison and downscale of AMSR2 soil moisture products with in situ measurements from the SCAN–NRCS network over Puerto Rico. *Hydrology*, vol. 4, no. 4, pp. 46, 2017.

Paris, P., Di Matteo, G., Tarchi, M., Tosi, L., Spaccino, L., and Lauteri, M. Precision subsurface drip irrigation increases yield while sustaining water-use efficiency in Mediterranean poplar bioenergy plantations. *Forest Ecology and Management*, vol. 409, no. 1, pp. 749-756, 2018.

Parnas, D. L. Designing software for ease of extension and contraction. *IEEE transactions on software engineering*, vol. 2, no. 1, pp. 128-138, 1979.

Patrikalakis, N. M., McCarthy, J. J., Robinson, A. R., Schmidt, H., Evangelinos, C., Haley, P. J., Lalis, S., Lermusiaux, P.F.J., Tian, R., Leslie, W.G. and Cho, W., Towards a dynamic data driven system for rapid adaptive interdisciplinary ocean forecasting. *Dynamic Data-Driven Application Systems.* 1st Edition, Kluwer Academic Publishers, Amsterdam, 2004.

Pettenuzzo, D., and Ravazzolo, F., Optimal Portfolio Choice Under Decision-Based Model Combinations, *Journal of Applied Econometrics*, vol. 31, no. 7, pp. 1312-1332, 2016.

Price, R., Appendix to T. Bayes,'Essay Towards Solving a Problem in the Doctrine of Chances'. *Philosophical Transactions of the Royal Society*, vol. 53, no. 1, pp. 370-418, 1763.

Pröpster, M., März, L., Reinhart, G., and Intra, C., Validation of Line Balancing by Simulation of Workforce Flexibility. *Procedia CIRP*, vol. 33, no. 1, pp. 93-98, 2015.

Rodriguez-Ortega, W. M., Martinez, V., Rivero, R. M., Camara-Zapata, J. M., Mestre, T., and Garcia-Sanchez, F. Use of a smart irrigation system to study the effects of irrigation management on the agronomic and physiological responses of tomato plants grown under different temperatures regimes. *Agricultural Water Management*, vol. 183, no. 1, pp. 158-168, 2017.

Serôdio, C., Cunha, J. B., Morais, R., Couto, C., and Monteiro, J. A networked platform for agricultural management systems. *Computers and electronics in agriculture*, vol. 31, no. 1, pp. 75-90, 2001.

Schwarz, D., and Colla, G. Amelioration of heavy metal and nutrient stress in fruit vegetables by grafting. *Scientia horticulturae*, vol. 127, no. 2, pp. 165-161, 2001.

Sang, H., Takahashi, S., and Gaku, R. Big data-driven simulation analysis for inventory management in a dynamic retail environment. *In Proceeding of the 24th International Conference on Industrial Engineering and Engineering Management,* pp. 687-694, Changsha, China, May 19-20, 2018.

© IEOM Society International

Shang, S., and Mao, X. Application of a simulation based optimization model for winter wheat irrigation scheduling in North China. *Agricultural Water Management*, vol. 85, no. 3, pp. 314-322, 2006.

Shklyar, A., and Arbel, A. Numerical model of the three-dimensional isothermal flow patterns and mass fluxes in a pitched-roof greenhouse. *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 92, no. 12, pp. 1039-1059, 2004.

Son, Y. J., Joshi, S. B., Wysk, R. A., and Smith, J. S. Simulation-based shop floor control. *Journal of Manufacturing Systems*, vol. 21, no. 5, pp. 380-394, 2002.

Soulis, K. X., and Elmaloglou, S. Optimum soil water content sensors placement for surface drip irrigation scheduling in layered soils, *Computers and electronics in agriculture*, vol. 152, no. 1, pp. 1-8, 2018.

Wen, Y., Shang, S., and Yang, J. Optimization of irrigation scheduling for spring wheat with mulching and limited irrigation water in an arid climate. *Agricultural Water Management*, vol. 192, no. 1, pp. 33-44, 2017.

## Biographies

**Sara Masoud** is an Assistant Professor of Industrial and Systems Engineering at WSU. Her research focuses on dynamic data-driven application systems by utilizing machine learning, simulation, and optimization models in different applications such as agro-industry, transportation, healthcare, and manufacturing. She is a member of IISE, IEEE, and INFORMS and has received the IISE Annual Meeting Best Paper Award in Data Analytics and Information Systems Track.

**Young-Jun son** is James J. Solberg Head and Ransburg Professor of Industrial Engineering at Purdue University. His research focuses on a data-driven, multi-scale, simulation and decision model needed for design and control in various applications, including extended manufacturing enterprise, homeland security, healthcare, and social network. He has authored/co-authored over 110 journal papers and 100 conference papers. He is a Fellow of Institute of Industrial and Systems Engineers (IISE), and has received the Society of Manufacturing Engineers (SME) 2004 Outstanding Young ME Award, the IIE 2005 Outstanding Young IE Award, the IISE Annual Meeting Best Track Paper Awards (2005, 2008, 2009, 2016, 2018, 2019), and Best Paper of the Year Award in 2007 from International Journal of IE. He has served as co-Program Chair for ISERC 2007, the General Chair for INFORMS Annual Meeting 2018, and the General Chair for Winter Simulation Conference 2019.