*Proceedings of the 2019 Winter Simulation Conference*
*N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.*

# HYBRID SIMULATION TO SUPPORT INTERDEPENDENCE MODELING OF A MULTIMODAL TRANSPORTATION NETWORK

Tarana Bipasha
Jose Azucena
Basem Alkhaleel
Haitao Liao
Heather Nachtmann

Department of Industrial Engineering
University of Arkansas
4183 Bell Engineering Center
Fayetteville, AR 72701, USA

## ABSTRACT

The inland waterways in the United States (U.S.) are used to transport approximately 20% of America's coal, 22% of U.S. petroleum products, and 60% of farm exports making these waterways a significant contributor to the U.S. multimodal transportation system. In this study, data about natural extreme events affecting inland waterways are collected and used to predict possible occurrences of such events in the future using a spatio-temporal statistical model. We also investigate the waterways disruptions effect on interconnected transportation systems using a simulation tool built on a statistical model. The developed methods are centered on inland waterways but can be used broadly for other local, regional and national infrastructures. A case study based on the Mississippi River and the McClellan–Kerr Arkansas River Navigation System (MKARNS) is provided to illustrate the use of the simulation tool in interdependence modeling and decision making for the operation of a multimodal transportation network.

## 1 INTRODUCTION

The physical distribution infrastructure is critical to national security, economic well-being, global competitiveness, and quality of life in the U.S. (Ellis et al. 1997). The distribution infrastructure, referred to as the transportation network, includes, but is not limited to, the interconnected network of ports and inland waterways, highways, and railroads. The transportation network in the U.S. includes almost 4 million miles of public roads and highways, more than 360,000 interstate trucking companies and 20 million trucks for business, and 1,900 seaports and 1,700 inland river terminals on 11,000 miles of inland waterways carrying grain, chemicals, petroleum products, and import and export goods (USDOT 2017; USACE 2019).

Given that many industries rely on the U.S. transportation network, the economic impacts of possible disruptions affecting the network are potentially enormous. Indeed, such disruptions can cause a cascading effect that can become widespread due to the spatial and temporal distributions of commodity flows (Pant et al. 2015). In fact, even without large-scale disruptions, the Federal Highway Administration (FHWA) estimated the trucking industry losses to be around $8 billion a year due to highway congestion (Herr 2008; Pant et al. 2015). Such losses are expected to increase in the future because of the forecasted increase in the U.S. domestic freight tonnages by 50% in the next 15 years (NCFRP 2010; USDOT 2009). In addition to highway network impacts, railways are expected to experience more significant congestions and breakdowns due to the increased demand on Class I railroads (Cambridge Systematics Inc. 2007).

The U.S. Maritime Administration, an agency of the U.S. Department of Transportation, has called for investment in the domestic waterways for freight movement (USDOT 2011), recognizing the need to reduce

978-1-7281-3283-9/19/$31.00 ©2019 IEEE 

road and rail congestion. The increased use of 25,000 miles of inland waterway freight transport will probably result in less congestion on U.S. roads as well as a reduction in risk of road and rail transport accidents and possibly even reduce emissions of air pollution (Pant et al. 2015). Barge transport is frequently cheaper than rail and truck alternatives, and there are many products which are too large for other transport methods. The nation's waterways are used to transport approximately 20% of America's coal, 22% of U.S. petroleum products, and 60% of farm exports between 38 states summing up the annual weight transported to around 630 million tons (USACE 2019).

Although general freight movements via the inland waterways are expected to increase in the upcoming years due to economic and logistic drivers, current studies addressing the impacts of disruptions on waterways operation and multimodal commodity flow along with the economic analysis are limited. Indeed, one reason for the limited number of studies may be the lack of tools to facilitate research in this area by providing data-driven models. There is an urgent need to protect and coordinate the nation's multimodal transportation infrastructures to support strong economic growth and national security. Without a doubt, inland waterways along with road and rail transport have huge impact on various businesses operations in the U.S., especially in America's heartland along hundreds of miles of the Mississippi River. However, inland water transportation is significantly affected by the weather, current and future waterway conditions, and operation strategies at different locks, dams and ports (Schweighofer 2014). For example, in case of flooding or drought, inland water transport will be constrained by the water levels of dams and ports, and the effects will propagate downstream. In response to such emergency situations, goods on cargo vessels need to be offloaded and re-routed through the available ground transportation system. Since these infrastructures are managed by different governing agencies (USACE 2019), multiple stakeholders need to understand the characteristics of these Interdependent Critical Infrastructures (ICIs) that cross administrative boundaries. Considering the large potential impact and lack of actual data availability, this research will generate simulated data on multimodal transportation systems.

Many studies have investigated modeling and simulation of ICIs through empirical approaches, agent-based approaches, network-based approaches and other approaches (Ouyang 2014). However, only few ones addressed simulation of inland waterways transportation. Bush et al. (2003) developed an iterative technique between optimization and simulation models to check the feasibility of barge routings suggested by the optimization model based on a sampled dataset. Biles et al. (2004) proposed a simulation model of traffic flow in inland waterways with the incorporation of the Geographic Information System (GIS) to improve vessels scheduling. Desquesnes et al. (2018) proposed a simulation architecture of inland waterways based on Markov Decision Process (MDP) and climate projections under uncertainty. All these studies do not consider predicting disruptions in advance based on statistical models, multimodal transportation, and do not allow users to control the lock and dam system to generate different scenarios.

The ultimate goal of this research is to create methods and research application opportunities from which the nation's economic growth and homeland security can significantly benefit, and to provide open-sourced multi-regional multi-industry data-driven statistical models and simulation tools to decision-makers, researchers and other stakeholders in order to build a good understanding of multimodal freight movement processes that combine different data sources. Thus, various data elements from historical events of natural inland waterway disruptions such as floods and droughts in the Mississippi River and the MKARNS were used to develop a spatio-temporal statistical model (Cressie and Wikle 2011; Kyriakos's and Journel 1999) to predict disruptions such as floods at different locations on both rivers, which guide the movement of multi-industry cargo vessels, lock-and-dam system in the area, and decisions regarding other modes of transportation for products shipped from or to inland ports.

Simulated data will be derived from actual data on ICIs. The ICIs related data includes: 1) inland waterway and ground transportation networks (e.g., road type and capacity of road network) (HSIP 2019); 2) locations of dams and locks (Dams 2019; Locks 2019; USDOT 2017); 3) locations of major ports and their top commodities (USACE 2019; USDOT 2017); 4) historical hydrological observation data at ports and locks including water depth, changes in waterways, and the normal capacity of inland water transport (Flood 2019); 5) major types of cargo vessels and barges classified by their capacity and usual transport speed; and 6) weather data covering the studied regions (NOAA Climate 2019). Moreover, the Maritime Transportation Research and Education Center (MarTREC) at the University of Arkansas (MarTREC 2019) provides the Transportation Resource Data Bank (MarTREC DataBank 2019) that compiles rich information, such as freight commodity flow and ports.

The remainder of this paper is organized as follows. Section 2 describes the development of the spatio-temporal statistical model used in this study along with the basic features of the model. Section 3 introduces the simulation tool we developed based on an open-source platform. Section 4 presents a case study to illustrate the capabilities of the tool. Section 5 provides concluding remarks and future research directions.

## 2 METHODOLOGY

A hybrid methodology combining statistical tools and simulations is applied. The statistical modelling is used with two purposes: to map the spatial fluctuations of gage height on a given river across sites, interpolating spatially unobserved points on a river and to forecast the gage height measurements on the sites of interest and anticipate possible interruptions in the flow of vessels. The simulation-based modelling is used to create scenarios for the flow of vessels and trucks taking as input the results from the statistical models. The dynamic interaction of different input parameters and simulation controls allow to estimate various metrics.

### 2.1 Geo Spatial Model

Environmental variables are among the factors affecting the reliability of critical infrastructures. As part of the modeling of the waterway transportation network, modeling relevant variables of the corresponding water bodies, e.g. rivers, becomes central in understanding the processes that affect the availability of the infrastructures of interest. The selected statistical modeling methodology must be capable of making appropriate predictions and estimate variation intervals for relevant variables on the selected sites.

For this purpose, the application of a model capable of capturing the underlying relationship between the selected variables, the spatial correlation among the selected measuring sites and the associated variations in time is one of the performed tasks in this stage of the project. The selected framework is spTimer (Bakar et al. 2015), a Spatio-Temporal Bayesian modeling package using the R language for statistics. The main variable of interest is the Gage Height (GH), a measure of the depth of the water filling the waterways on the measurement sites. The main purpose of this model is to generate data to estimate the GH on unobserved sites of interest. In this context, unobserved sites are selected locations that have no available measurements data and it is necessary to infer it from the observed sites. The model will learn a spatio-temporal mapping for the GH data from the observed sites and generate interpolations for new coordinates of interest along the same rivers.

#### 2.1.1 Data

The data used corresponds to the hourly measurements of GH and lock availability data in 18 different sites. This is equivalent to 18 geo-related time series with 17542 observations each. There are 22,961 missing measurements, representing 7.3% of total observations. Table 1 shows a general statistical summary of the Gage Height measurements. The time window used starts on February 22, 2016 and finishes on February 21, 2018. The observed sites are shown in Figure 1. The sites are classified as connected to the MKARNS (red) or the Mississippi River (green). Selected unobserved locations of interest are marked with “X”.

Table 1: Statistical summary of gage height data.

| Min | 1st Quartile | Median | Mean | 3rd Quartile | Max |
|---|---|---|---|---|---|
| 0.00 | 7.45 | 11.83 | 14.04 | 19.43 | 44.63 |

### 2.1.2 Theoretical Background

To model the GH data, we use a dynamical model depending on an underlying Independent Gaussian Process with Bayesian estimation of parameters (Bakar et al. 2015), using the implementation found in the *spTimer* R package. The hierarchical model used is shown below with description of variables and inputs:

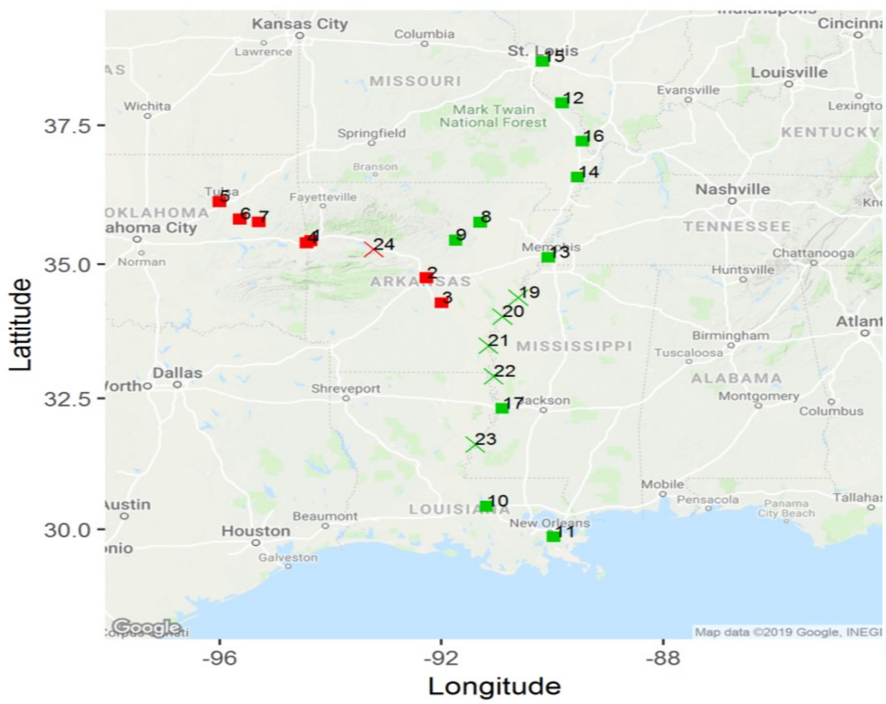


Figure 1: Measurement sites by river.

$$\mathbf{Z}_{lt} = \mathbf{O}_{lt} + \boldsymbol{\epsilon}_{lt}, \qquad \mathbf{O}_{lt} = \mathbf{X}_{lt}\boldsymbol{\beta} + \boldsymbol{\eta}_{lt} \tag{1}$$

Let $l$ and $t$ be the two units of time where $l$ denotes the longer unit (month), $l = 1, \ldots, r$, and $t$ denotes the shorter unit (hour), $t = 1, \ldots, T_l$ where $r$ and $T_l$ denote the total numbers of two time units, respectively. Let $Z_l(s_i, t)$ denote the observed point-referenced data and $O_l(s_i, t)$ be the true value corresponding to $Z_l(s_i, t)$ at site $s_i, i = 1, \ldots, n$ at time denoted by two indices $l$ and $t$. From (1), let $\boldsymbol{Z}_{lt} = \left(Z_l(s_1, t), \ldots, Z_l(s_n, t)\right)^T$ and $\boldsymbol{O}_{lt} = (O_l(s_1, t), \ldots, O_l(s_n, t))^T$. Let $N = n\sum_{l=1}^{r} T_l$ be the total number of observations to be modeled.

$\boldsymbol{\epsilon}_{lt} = (\epsilon_l(s_1, t), \ldots, \epsilon_l(s_n, t))^T$ will be used to denote the pure error term, assumed to be independently normally distributed $N(\mathbf{0}, \sigma_\epsilon \boldsymbol{I}_n)$. The spatio-temporal random effects will be denoted by $\boldsymbol{\eta}_{lt} = (\eta_l(s_1, t), \ldots, \eta_l(s_n, t))^T$ and these will be assumed to follow $N(\mathbf{0}, \Sigma_\eta)$ independently in time. The spatial correlation matrix $S_\eta$ is obtained from the general Matérn correlation function (Matérn 1986). This function (2) is well suited to quantify two dimensional correlations based on distance measurements and can be seen as a generalization of a Gaussian radial basis function:

$$\kappa(\mathbf{s}_\mathrm{i}, \mathbf{s}_j; \phi, \nu) = \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(2\sqrt{\nu}\|\mathbf{s}_\mathrm{i} - \mathbf{s}_j\|\phi\right)^{\nu} \mathrm{K}_\nu\left(2\sqrt{\nu}\|\mathbf{s}_i - \mathbf{s}_j\|\phi\right), \phi > 0, \nu > 0 \tag{2}$$

where $\Gamma(\nu)$ is the standard gamma function, $K_\nu$ is the modified Bessel function of second kind with order $\nu$. Let $\boldsymbol{\theta} = (\beta, \sigma_\epsilon^2, \sigma_\eta^2, \phi, \nu)$ denote all the parameters of this model and let $\pi(\boldsymbol{\theta})$ denote the prior distribution that we shall specify later. The logarithm of the joint posterior distribution of the parameters and the missing data, represented by z*, for this GP model (Bakar et al. 2015) is given by:

$$log\,\pi\,(\boldsymbol{\theta},\mathbf{O},\mathbf{z}^{*}|\mathbf{z}) \propto \frac{N}{2} log\,\sigma_{\epsilon}^{2} - \frac{1}{2\sigma_{\epsilon}^{2}}\sum_{l=1}^{r}\sum_{t=1}^{T_l}(\mathbf{Z}_{lt}-\mathbf{O}_{lt})^{\intercal}(\mathbf{Z}_{lt}-\mathbf{O}_{lt}) - \frac{\sum_{l=1}^{r}T_l}{2} log\left|\sigma_{\eta}^{2}S_{\eta}\right|$$
$$-\frac{1}{2\sigma_{\epsilon}^{2}}\sum_{l=1}^{r}\sum_{t=1}^{T_l}(\mathbf{O}_{lt}-\mathbf{X}_{lt}\boldsymbol{\beta})^{\intercal}S_{\eta}^{-1}(\mathbf{O}_{lt}-\mathbf{X}_{lt}\boldsymbol{\beta}) + log\,\pi\,(\boldsymbol{\theta})$$

The spTimer can fit and predict, spatially and temporally, using three models: Gaussian Processes using Gibbs sampling (GP), Auto Regressive (AR), and Gaussian Processes using knot locations for the random effects (GPP). Markov chain Monte Carlo (MCMC) computational techniques are used to calculate estimations. Integration with other R packages that handle large spatio-temporal datasets, prognostics and estimations and graphics is possible. Using the collected GH data, the model was applied using R. Different model configurations were tested. The steps to select the final model configuration and the validation strategy is presented as follows:

1. Load data and transform to a spatio-temporal data frame.
2. Separate data by river to fit an appropriate model for each river: MKARNS, Mississippi River.
3. Separate data in training and validation sets. For MKARNS, sites 1, 2, 3, 5, 6, and 7 are for training. Validation site is number 4. Unobserved site, 24. For Mississippi River, sites 8, 10, 11, 12, 14 ,15, 16, and 17 are for training. Validation site is number 13. The unobserved sites are 19, 20, 21, 22, and 23. These numbering of sites follows Figure 1.
4. Estimate the model using the training set: fit a mean function by river across all sites on a river. To account for seasonality along the year, calculate the river mean across years for any given time point. Use the mean function and dummy variables that encode different seasonality periods as covariates into the model then create predictions for the validation set.
5. Evaluate the MSE of the predictions vs. the real measurements on the validation sites. Table 2 shows a summary of different performance metrics used to evaluate the models.
6. Repeat steps 3-5 for different model configurations. Select the one with the lowest MSE.

Using this strategy, the MSE for MKARNS is 7.95 and, on the Mississippi River, 43.72. The predicted interpolation is shown graphically in Figure 2, where the red line represents the interpolation and the blue line the observed measurements. Using the river mean function across year, the effects of seasonality are reasonably represented by the model.

Table 2: Selected performance measures for the spatial models for the validation sites.

| Model | MSE | RMSE | MAE | MAPE | BIAS |
|---|---|---|---|---|---|
| Arkansas River | 7.95 | 2.82 | 2.62 | 22.06 | 2.53 |
| Mississippi River | 43.72 | 6.61 | 5.70 | 188.69 | 3.58 |

## 2.2 Time Series Forecasting

To assess future GH levels for the observed sites, an appropriate forecasting method is required and a variety of models were tested. Given desirable performance properties, flexibility, scalability, and capability of detecting and capturing seasonality, Facebook's Prophet framework (Taylor and Letham 2018) was selected and implemented through the R package called *Prophet*.

The Prophet framework models a time series using Generalized Additive Models (GAM), regression models with potentially non-linear smoothers applied to the regressor. In this case, time is the regressor and different functions of the time series are used as components. The series is decomposed into three main components: trend, seasonality and special events. These are denoted by g(t), s(t) and h(t) respectively in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

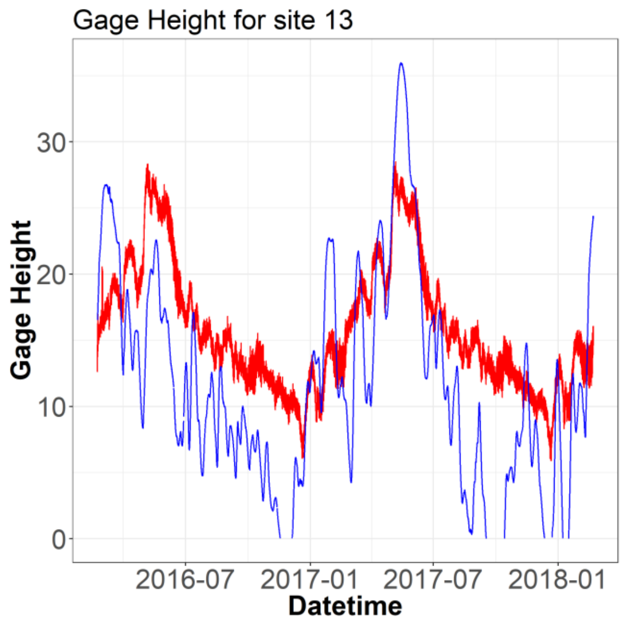


Figure 2: Spatial interpolation for the Mississippi River validation site.

The trend function models non-periodic changes, the seasonality function models periodic changes over a defined time period (day, week, month or year) and the special events identify known irregular disturbances on the time series. The model is fitted to the data using L-BFGS (Taylor and Letham 2018).

For the trend component, a piecewise logistic growth model is used. To account for constraints in the GH levels, maximum and minimum possible values are set. The constraints account for the non-negativity of the measurement and an assumed maximum of 50 feet. The fitting allows for automatic change-point detection in the trend. The adjustments at change-points are computed as:

$$\gamma_j = (s_j - m - \textstyle\sum_{l<j} \gamma_l)\left(1 - \frac{k+\sum_{l<j} \delta_l}{k+\sum_{l\leq j} \delta_l}\right)$$

where $\delta_j$ represent rate adjustments in the time series that occur at a detected change-point $s_j$. The rate at any time is the base rate $k$ plus all the rate adjustments up to that point. $\boldsymbol{a}(t)$ is a vector of indicating variables that take the value of 1 for any time $t$ after the detected change-point. An offset parameter $m$ is used to connect endpoints in different segments. The piecewise logistic growth model is:

$$g(t) = \frac{C}{1+\exp(-(k+\boldsymbol{a}(t)^\intercal \boldsymbol{\delta})(t-(m+\boldsymbol{a}(t)^\intercal \boldsymbol{\gamma}))}$$

where C represents the assigned capacity for each observed site. For the current model, these capacities are fixed and constant across sites, but there is flexibility to vary them if needed.

For the seasonality component, a Fourier series is used to provide flexibility to periodic effects. Given P, the regular period as the number of observations that conform that period, the seasonal effects are smoothed using:

$$s(t) = \textstyle\sum_{n=1} \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right)\right)$$

where the 2N parameters $\boldsymbol{\beta} = [a_1, b_1, \ldots a_n, b_n]$ are estimated constructing a matrix of vectors for each value of *t.* Some smoothing priors are set on these parameters to make estimation easier.

Two complete years of data for each site were used for learning, and a forecast of the next three months, unobserved by the model, is evaluated for validation. For data generation, this model is fitted on the observed sites and used to forecast unobserved time periods. A one year ahead forecast is used as input into the simulation model.

### 2.3 Spatio-temporal Interpolation

Using the generated forecasts, unobserved time periods can be interpolated in unobserved locations using the output of the temporal forecast as input for the geo spatial model. This methodology is proposed to generate data for the unobserved locations and use it as input for the simulation model.

## 3 HYBRID SIMULATION MODEL

This simulation model is developed using NetLogo, which is an agent-based programming language and simulation platform offered as freeware (Wilensky 1999). NetLogo is also a cross-platform and integrated environment for modeling both simple and complex systems that evolve over time. “Agents” (turtle, link, patch, and observer) are the integral part of NetLogo world and capable of following instructions given by the designers. . Turtles move around in the two-dimensional world, whereas the world contains a grid of patches. Every patch is like a square piece of land. All these agents can operate simultaneously without interfering one another. NetLogo permits users to run the simulation in a browser or desktop application, interact with it, and analyze its behavior under various settings (Tisue and Wilensky 2004).

### 3.1 Overview of the Simulation Model

Our model has been built on four extensions of NetLogo: GIS, R, NW, and CSV. GIS extension provides the ability to load vector GIS (Geographic Information System) data in the form of ESRI shapefiles. We use it to import several maps in our model. Initially, we load the map of the United States as the base of NetLogo environment. Then, we import maps of inland waterways and highways on top of that. However, we focus primarily on MKARNS and Mississippi River during simulation. Figure 3 shows NetLogo’s user interface after opening and setting the basic environment of the model. The graphic window makes the two-dimensional “world” of the model visible. It is divided up into a grid of patches that have coordinates *pxcor* and *pycor*. The basic idea here is creating a NetLogo graph (nodes and links) by importing the GIS maps and creating vessel and truck "agents" which travel along the links. The main components of the program are:

- A map of the United States, drawn on NetLogo in a simplified form. Each state is filled with vivid colors to make them look more distinct. In Figure 3, we see a zoomed version of the map.
- Maps of navigable waterways and highways. Both are made of nodes (turtles with own variables) connected by links. While the waterways/highways are only figurative, the nodes play an active role in the simulation.
- Vessels; these are turtles with their own variables such as current location, destination, distance-traveled, speed, vessel-category, product-weight, product-type, extreme-events, total-delay, etc.
- Trucks; these are also turtles with their own variables such as current location, destination, distance-traveled, speed, product-weight, product-type, etc.
- Ports (turtle breeds) along the waterways. We have shown eight ports (located in Tulsa, Fort Smith, Little Rock, Mississippi, Baton Rouge, Helena, Memphis, and St. Louis) along MKARNS and Mississippi River. In Figure 3, the yellow nodes represent the ports.
- Fifteen locks along MKARNS. They are also made of nodes (turtle breeds) with their own variables such as id and location. In Figure 3, the green crosses on MKARNS depict the locks.
- Twenty-four sites (turtle breeds) along MKARNS and Mississippi River. In each site, we check the Gage Height level and make a decision about whether the vessels will move forward or not. The red nodes in Figure 3 represent the sites.

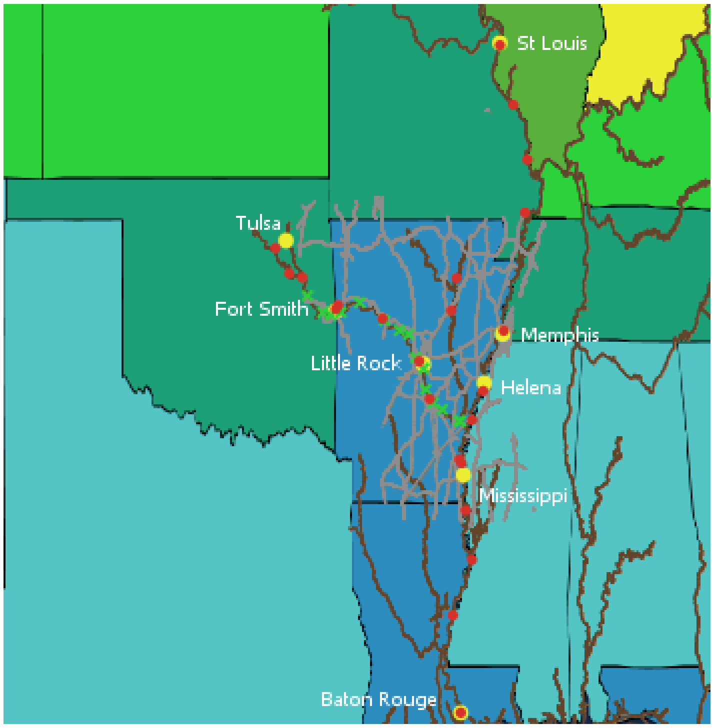


Figure 3: Simulation model interface on NetLogo.

- An algorithm that makes the vessels and trucks move on the waterways and highways, respecting some interaction rules between source and destination, navigation time and speed, and other agents. For example, during the simulation, the vessel always takes the shortest path between its source and destination.

The main assumptions of our model are as follows:

- No random variable is used in the model. Randomization is initiated by the user when they choose different values for lock switches, GH threshold slider and fleet size chooser.
- Vessels are uniformly distributed based on annual demand.
- The speed of vessels varies with its capacity and size. The smallest vessel is the fastest one with an average speed of 9 mph. The medium sized vessel moves at 7 mph where the largest one moves with 5 mph.
- Each vessel and truck carries only one commodity.
- All the vessels and trucks travel only once to their predefined destinations and do not return to their origin ports.

### 3.2 Setup Procedure

Model controls (Figure 4) quickly adjust the settings of the initial environment. They are represented by buttons, sliders, etc. To initiate the simulation after setting up the environment, the user is allowed to input various attributes through sliders, choosers, and switches. The steps are as follows:

- Step 1: First of all, press "Spatial Temporal Analysis" button to run the spatio-temporal model to generate the forecasted Gage Height and Lock availability data.
- Step 2: Press the "Setup Environment" button. This button is used to initialize the model; it is a "once-button" that runs its code once. After this step, all the maps will be drawn on the interface.

- Step 3: Click "Add Ports" button to draw eight circles on the waterways that depict as the ports of our consideration.
- Step 4: Use the "Add Locks & Sites" button to initialize and draw all fifteen locks and twenty-four sites with their variables. Then click "Add Vessels" to draw vessels on the ports based on the data. The vessels are categorized by size and speed into three groups and they carry two types of products: petroleum and crops.

After completing the above steps, the user may set a value for "Simulation-Time" which indicates for how many months the simulation will run; the available options are 3, 6, 9 and 12 months. "Fleet-Size" provides the value of the number of trucks required to carry products from one vessel through the highways when vessels reach to their destinations, and also when they are unable to move for a certain amount of time due to extreme events. The slider "GH-Threshold" may need to be adjusted to a reasonable value. This value acts as the threshold value of Gage Height which is being compared with the hourly value of gage height at each site. We recommend setting a value between 30 and 45 for a better result. We can change the value at runtime. There are fifteen switches each of which acts as controllers to turn on/off a lock. The selection can be changed in runtime too. When one lock is closed, the vessels that are supposed to pass through it will wait nearby and would not move forward until the status changes. After all the settings are done, click the "Start Simulation" button. When this button is pressed, the vessels at each port start moving towards their predefined destination. In the meantime, each vessel checks for any unsafe circumstances at the locks and sites along its route and takes decision accordingly.

## 4 A CASE STUDY

In the simulation, we analyze the system to measure certain representative quantities based on the input parameters listed in Table 3. In addition, Figure 4 represents a sample run and Figure 5 shows plots that were generated during the simulation. We generate an output file at the end of the simulation which presents

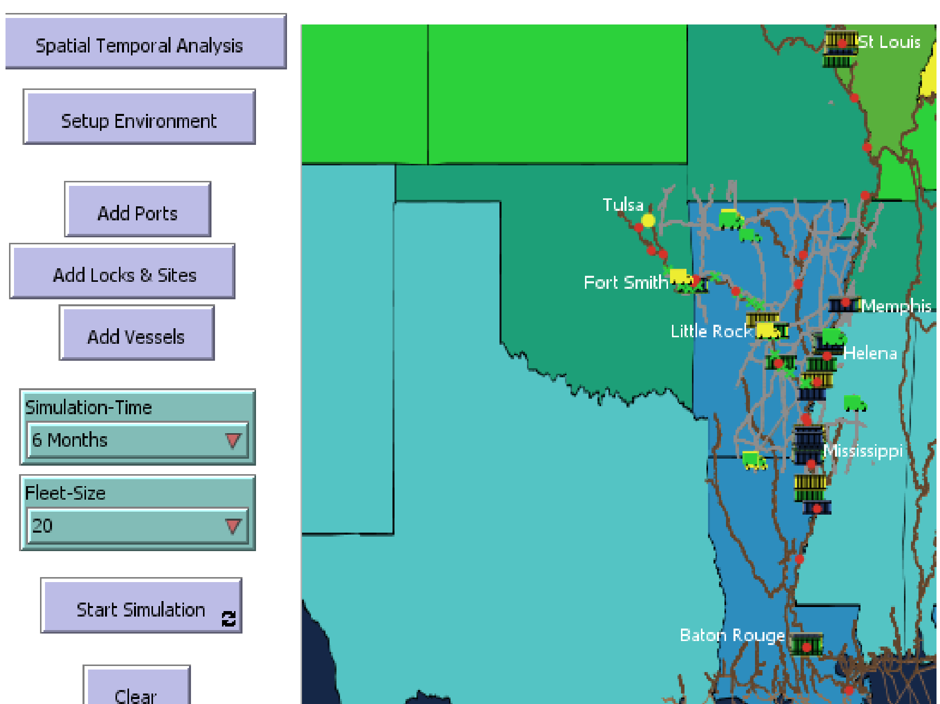


Figure 4: Simulation running on NetLogo.

all the measurements listed above. The simulation time is set for 6 months, fleet size is 20 trucks and GH threshold is 30 feet. Initially, vessels started from the ports of Tulsa, Baton Rouge, Little Rock, Mississippi and Helena, and the locks were all open. The vessels were moving towards their destination ports and extreme events were checked by measuring gage height and lock availability at each site. For example, the vessels passing the sites (with GH more than 30 feet) between the LA-MS route along the Mississippi River were not able to move and had to wait until the GH level falls below 30. Whenever a vessel faces any extreme event and stops moving forward, it turns red in color to broadcast this event. The vessel gets back to its original color when it resumes movement. At some point during the simulation, we decided to close one of the locks, “Robert Kerr”, manually by turning the switch off; which resulted in discontinuation of vessel’s movement through that lock along MKARNS. Due to this extreme event, the average speed drops significantly along TUL–LA route. In Figure 5, the “Avg speed of Vessels” plot depict these changes in speed; the green line TUL–LA represents the average speed of vessels traveling from Tulsa to Baton. If any extreme event occurs and the resulted delay continues for more than 5 days, the cargo from the vessels will be unloaded and transferred by trucks through highways.

Table 3: Simulation model input and output.

| Model Input | Model Output |
|---|---|
| • Gage Height from a spatio-temporal model<br>• Supply and demand between ports (movement of commodities)<br>• Gage height threshold limit<br>• Lock availability<br>• Vessel distribution at each port<br>• Fleet size<br>• Number of trucks | • Average speed of each vessel category between every two ports (mph)<br>• Number of delays between every two ports<br>• Total time lost due to extreme events (hour)<br>• Total number of vessels delayed and their tonnages<br>• Overall average speeds for the 3 types of vessels (mph)<br>• Number of extreme events and length (time) in MKARNS and Mississippi River<br>• Number of vessels from each category traveled and arrived between every two ports<br>• Average speed of trucks for each product type (mph) |

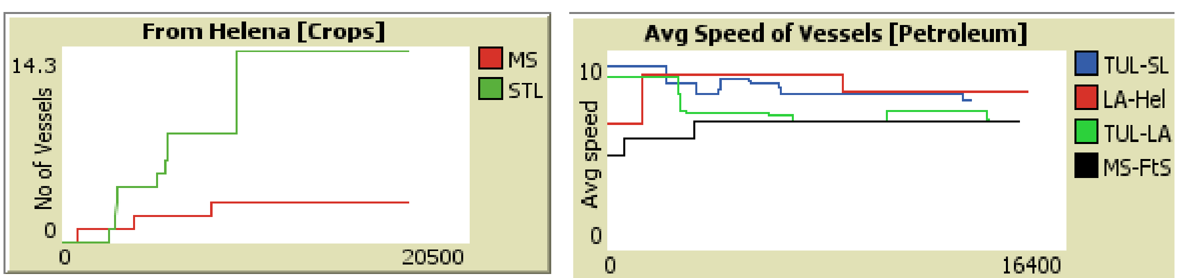


Figure 5: Sample of plots generated during simulation.

Moreover, when a vessel reaches to the destination port, trucks are used to carry its products to the final destinations. Figure 5 (left) shows the number of vessels that were used to carry products (crops) between two ports. At the end of the simulation, we generate an output report summarizing different statistics of vessels (e.g. average speed and number of extreme events) and trucks along with graphical plots (e.g. boxplots) that help the user understand all different aspects of the hybrid model.

## 5 CONCLUSION

In this study, several contributions are made to ICIs risk analysis and economic studies literature. First, a spatio-temporal statistical model was developed to capture extreme natural events causing disruptions in inland waterways and predict them in the future to facilitate commodity flow planning and response actions. The statistical model was developed and tested on the Mississippi River and the MKARNS. Second, we built a simulation tool that captures the effect of inland waterways disruptions on the commodity flow through other ICIs which provides a broad understanding of the multimodal transportation system interdependencies in action. Third, the simulation tool will be available as an open-sourced tool for researchers, decision makers and other stakeholders to push the research in multimodal transportation forward. This work can be extended to include emergency services responses and detailed analysis of ports operations which are the next steps in this research.

## ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation under Grant No. CMMI–1745353.

## REFERENCES

Bakar, K. S. and S. K. Sahu. 2015. "spTimer: Spatio-temporal Bayesian Modelling using R". *Journal of Statistical Software* 63(15):1-32.

Biles, W. E., D. Sasso, and J. K. Bilbrey. 2004. "Integration of Simulation and Geographic Information Systems: Modeling Traffic Flow on Inland Waterways". In *Proceedings of the 2004 Winter Simulation Conference*, edited by R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 331–337. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bush, A., W. E. Biles, and G. W. DePuy. 2003. "Iterative Optimization and Simulation of Barge Traffic on an Inland Waterway". In *Proceedings of the 2004 Winter Simulation Conference,* edited by S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice,1751–1756. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Cambridge Systematics, Inc. 2007. "National Rail Freight Infrastructure Capacity and Investment Study", Technical Report, Association of American Railroads, Cambridge, Massachusetts.

NOAA Climate. 2019. National Oceanic and Atmospheric Administration (NOAA Climate). https://www.noaa.gov/climate, accessed 18th January 2019.

Cressie, N. A. C. and C. K. Wikle. 2011. *Statistics for spatio-temporal data*. Hoboken, New Jersey: John Wiley & Sons.

Dams. 2019. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-dhs-gii.opendata.arcgis.com/datasets?q=*dam&sort_by=relevance, accessed 20th March 2019.

Desquesnes, G., Alves, D., Duviella, E., Lozenguez, G. and A. Doniec. 2018. "Simulation Architecture Based on Distributive MDP for Inland Waterway Management." *EPiC Series in Engineering* 3:555–63.

Ellis, J., D. Fisher, T. Longstaff, L. Pesante, and R. Pethia. 1997, "Report to the President's Commission on Critical Infrastructure Protection". Software Engineering Institute, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

Flood. 2019. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-dhs-gii.opendata.arcgis.com/datasets?q=flood, accessed 19th March 2019.

Herr, P. R. 2008. "Approaches to Mitigate Freight Congestion". Technical Report No. GAO-09-163R, Government Accountability Office, Washington, DC.

HSIP. 2019. Homeland Security Infrastructure Program (HSIP) data. https://hifld-dhs-gii.opendata.arcgis.com/, accessed 19th March 2019.

Kyriakidis, P. C. and A. G. Journel. 1999. "Geostatistical Space-time Models: A Review". *Mathematical Geology* 31(6):651–684.

Locks. 2019. Homeland Infrastructure Foundation-Level Data (HIFLD). https://hifld-dhs-gii.opendata.arcgis.com/datasets?q=Locks&sort_by=relevance, accessed 20th March 2019.

MarTREC. 2019. Maritime Transportation Research and Transportation Center. https://martrec.uark.edu/, accessed 20th March 2019.

MarTREC DataBank. 2019. Maritime Transportation Resource Data Bank, MarTREC, University of Arkansas. https://martrec.uark.edu/research/resource-data-bank.php, accessed 24th February 2019.

Matérn, B. 1986. *Spatial variation*. Berlin: Springer-Verlag.

National Cooperative Freight Research Program. 2010. "Current and Completed Projects". Transportation Research Board.

Ouyang, M. 2014. "Review on Modeling and Simulation of Interdependent Critical Infrastructure Systems". *Reliability Engineering and System Safety* 121:43–60.

Pant, R., Barker, K. and T. L. Landers. 2015. "Dynamic Impacts of Commodity Flow Disruptions in Inland Waterway Networks". *Computers and Industrial Engineering* 89:137–149.

Schweighofer, J. 2014. "The Impact of Extreme Weather and Climate Change on Inland Waterway Transport," *Natural Hazards* 72(1):23–40. https://doi.org/10.1007/s11069-012-0541-6 accessed 19th March 2019.

Taylor, S. J. and B. Letham. 2018. "Forecasting at Scale". *The American Statistician* 72(1):37–45.

Tisue, S. and U. Wilensky. 2004. "NetLogo: Design and Implementation of a Multi-agent Modeling Environment". In *Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction*, *Reflexivity and Emergence*, October, Chicago, IL,7-9.

USACE. 2019. Navigation Data Center. http://www.navigationdatacenter.us/, accessed 27th March 2019.

USDOT. 2009. "Freight Facts and Figures 2009". U.S. Department of Transportation. Federal Highway Administration, Office of Freight Management and Operations, Washington, DC.

USDOT. 2011. "America's Marine Highway Report to Congress". Technical Report, U.S. Department of Transportation. Maritime Administration, Washington, DC.

USDOT. 2017. "Geospatial at the Bureau of Transportation Statistics". U.S. Department of Transportation. Washington, DC. https://maps.bts.dot.gov/arcgis/home/index.html, accessed 20th March 2019.

Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/ accessed 27th March 2019.

## AUTHOR BIOGRAPHIES

**TARANA BIPASHA** is currently pursuing her Masters in Computer Science, under the supervision of Dr. Susan Gauch, with the Department of Computer Science and Computer Engineering, at the University of Arkansas in Fayetteville, USA. She received her BSc. degree in Computer Science and Engineering from Bangladesh University of Engineering and Technology. Her research interests include natural language processing, machine learning and cloud computing. Besides simulation modeling and analysis, she is also working on social network analysis based on literary text. Her email address is ttbipash@uark.edu.

**JOSE AZUCENA** is a Ph.D. student in the Industrial Engineering Department at the University of Arkansas, Fayetteville, US. He received a Business Engineering B.S. degree from ESEN, Santa Tecla, El Salvador, in 2014. His main research interests are centered around prognostics and health management, systems reliability, machine learning, and deep learning. Email address: jch051@uark.edu.

**BASEM ALKHALEEL** is a Ph.D. candidate in the Department of Industrial Engineering at the University of Arkansas. He received the B.S. degree in industrial engineering from King Saud University, Riyadh, Saudi Arabia, in 2012, and the M.S. degree in industrial engineering from Texas A&M University, College Station, TX, USA, in 2015. His main research interests are centered around systems reliability, networks optimization and critical infrastructures resilience. His email address is baalkhal@uark.edu.

**HAITAO LIAO** is a Professor, and John and Mary Lib White Endowed Systems Integration Chair in the Department of Industrial Engineering at University of Arkansas - Fayetteville. He received a Ph.D. in Industrial and Systems Engineering from Rutgers University in 2004. His research has been sponsored by the U.S. National Science Foundation, Department of Energy, Nuclear Regulatory Commission, Oak Ridge National Laboratory, and industry. The research findings of his group have been published in IISE Transactions, European Journal of Operational Research, Naval Research Logistics, IEEE Transactions on Reliability, IEEE Transactions on Cybernetics, The Engineering Economist, Reliability Engineering & System Safety, etc. In 2014, he served as Chair of INFORMS Quality, Statistics and Reliability (QSR) Section, and President of IISE Quality Control and Reliability Engineering (QCRE) Division. He served as Associate Editor for Journal of Quality Technology and IEEE Transactions on Reliability, and currently serves as Associate Editor for IISE Transactions on Quality and Reliability Engineering. He received the U.S. National Science Foundation CAREER Award in 2010, the IISE QCRE William A.J. Golomski Award for three times, 2013 QCRE Track Best Paper Award, Stan Ofsthun Best Paper Awards in 2015 and 2019, and the prestigious 2017 Alan O. Plait Award for Tutorial Excellence. His email address is liao@uark.edu.

**HEATHER NACHTMANN** is the Associate Dean for Research in the College of Engineering, Earl J. and Lillian P. Dyess Endowed Chair in Engineering, and Professor of Industrial Engineering at the University of Arkansas. Dr. Nachtmann serves as director of the Maritime Transportation Research and Education Center (MarTREC), a U.S. Department of Transportation University Transportation Center and the Mack-Blackwell Transportation Center. Her research includes economic and operations analysis of inland waterways, intermodal network optimization, and vulnerability assessment and emergency preparedness for transportation networks. Dr. Nachtmann has published more than eighty peer-reviewed publications and led over ten million dollars in research grants as principal investigator. Dr. Nachtmann teaches in the areas of operations research, engineering economics, and cost analysis. Dr. Nachtmann is a Fellow of the American Society for Engineering Management and Institute of Industrial and Systems Engineers, a member of the Arkansas State Highway and Transportation Department Research Advisory Council, and co-editor-in-chief of the Engineering Management Journal and an associate editor for The Engineering Economist. She received her Ph.D. from the University of Pittsburgh. Her email address is hln@uark.edu.