Expansion of the Panama Canal: Simulation Modeling and Artificial Intelligence

Abstract ID: 1228

Abstract

This paper discusses the development of a hybrid model using discrete-event simulation and system dynamics of the Panama Canal Expansion. The different operational processes are modeled using the current set of locks and additional processes have been added to reflect the new sets of locks being built. These new set of locks are for Post-Panamax ships and utilize lateral water saving basins. The differential equations model the volume of the lakes taking into considerations additional type of operations (e.g., power generation, drinking water pumping stations, spillways) and natural processes (e.g., rainfall, evaporation, tributaries). Support vector machines are used to model evaporation based on historical data of parameters such as radiance, humidity, temperature, precipitation, and winds and their directions. In addition, neural networks were developed to map the volumes of the lakes to the respective lake levels. The emphasis is on the prediction of the salt water load in the Gatun and Miraflores Lakes due to the expansion. The simulation model utilized the exchange of water and salt based on the mass balance and salt balance equations of the transit process. The simulation model created is compared against historical data and other models developed by different organizations.

Keywords
Panama Canal, System Dynamics, Neural Networks, Support Vector Machines, Discrete-Event Simulation

1. Introduction
The Panama Canal (Figure 1) currently carries 4 percent of world’s traded goods, and it is an important competitor in some very important shipping routes. For example, the Canal currently handles about 16% of the United States maritime trade, and more than 25% of the containerized trade between North East Asia and the East Coast of the
United States [1]. Within the Republic of Panama the Canal is responsible for the growth of the terminal cities of Panama and Colon (these cities and their metropolitan areas have 70% of the population of the Republic of Panama).

Figure 1: Schematic of the Panama Canal

The Panama Canal Expansion (Figure 2) is a very important project. As stated by the Panama Canal Authority [2]: “The third set of locks project is a plan to expand the Canal’s capacity composed of three integrated components: (1) the construction of two lock facilities – one on the Atlantic side and another on the Pacific side – each with three chambers, each which include three water reutilization basins; (2) the excavation of new access channels to the new locks and the widening of existing navigational channels; and, (3) the deepening of the navigation channels and the elevation of Gatun Lake’s maximum operating level.” The Panama Canal Expansion will open in 2014.

Figure 2: Points 2 (Atlantic side) and 6 (Pacific side) indicate the location of the new set of locks for Post–Panamax ships (adapted and modified from: http://www.pancanal.com/eng/plan/documentos/propuesta)

The Panama Canal expansion through a third set of locks presents new challenges. One challenge is the potential increase in the salinity of Gatun Lake above permissible levels. To understand this operational system and its complexities, we have to both describe and model its current structure [3]. Therefore, we created a basic discrete simulation model of the current Panama Canal operations. This model was enhanced with new information provided by engineers of the Panama Canal Authority of the new set of locks and its impact on operations. A system dynamics model shows the diffusion of the saline ocean water into Gatun Lake. This required the mapping of the
volume of Gatun Lake to its level using neural networks and the mapping of evaporation based on measurable factors such as precipitation, humidity, radiance, salinity, and winds and their directions.

2. Salinity of the Lakes

2.1 Modeling of the Salinity Level
Salinity refers to the mass quantity of dissolved salts per unit of water mass or water volume (1 unit = 1 liter). Seawater’s salinity (S) amounts to 35 parts per thousand (ppt). The chloridity (Cl), which is sometimes used, represents the mass quantity of chloride ions per unit of water mass or water volume. The Cl of fresh water should not exceed 0.2 to 0.25 ppt. This fresh water limit corresponds to a salinity value of 0.4 to 0.5 ppt. The relationship between S and Cl is: 

\[ S = 0.03 + 1.905 \times Cl \] (valid for S and Cl > 1 ppt) [4].

The salinity diffusion was modeled using the exchange of mass transfer [4,5]. This involved the study of the different volumes and salinity gradients of the Panama Canal System: Water systems and Locks. In addition, it was complemented by the data collection using historical data provided by the Panama Canal Authority (and data collected by the research team). For convenience reference, the 6 locks (see Figure 1) are given by the following notation [4,5,6]:

L1: Lowest lock on the Pacific Side
L2: Lock between Miraflores Lake and L1
L3: Lock between Miraflores Lake and Gatun Lake (Pedro Miguel)
L4: Highest lock directly connected to Gatun Lake
L5: Middle lock on Atlantic side
L6: Lowest lock connected to Atlantic Ocean

The Panama Canal has two lakes: Gatun and Miraflores. Water from these two lakes is used for the Panama Canal System to fill the navigation locks. Salt water from the Pacific and Atlantic Oceans gets added to the lakes during the transit of the ships. In addition, water from the lakes is lost to the sea during the same process. The Gatun Lake supplies fresh water to the population of Panama and Colon Cities for drinking purposes. The Miraflores Lake has a level of salinity which is already considered “brackish” water (i.e., Brackish water is water that has more salinity than fresh water, but not as much as seawater) [4].
Derivations of the different formulas from the different steps in the locks were considered (Figure 3). LS (“left side”) in Figure 2 above, represents either the Pacific Ocean, the Caribbean Sea, or the volume of water contained in a lock at a lower level. Similarly, RS (“right side”) represents the volume of water contained in a lock at a higher level such as the lakes Miraflores or Gatun. The initial conditions \((t = 0)\) for the volume of the lock \((V(0))\) and the salinity \((S(0))\) are given by \([3,4]\):

\[
S(0) = S_0 \\
V(0) = V_T - V_L
\]  

(1)

where \(V_T\) is the volume capacity of the lock, \(V_L\) is the volume necessary to reach the next level (RS), and \(S_0\) is the initial salinity of the water inside the lock before the water inside and outside are combined. There are four separate steps in an uplockage or in a downlockage process: (1) the gates on the LS lock open, and the water inside the lock combines with the water pushing the ship into the lock; (2) the gates on the LS lock close and the ship is raised to the next water level; the water of both the LS and RS locks combine and the salinity changes to the new water mix; (3) the gates in the RS lock open, and the water inside the lock combines with the water in the next level (either a lock or a lake); (4) the ship leaves the lock and the water in the lock falls into the lower level (either a lock or the sea); the volume inside the lock is again given by (1).

After the four steps above, the salinity of the water is given by

\[
S(3) = \frac{S(2)[V(2)-V_{RS}]+S_{RS}(V_{RS}+V_S)}{V(3)}
\]  

(2)

where \(S(3)\) is the salinity in steps 3 and 4, since they have the same value. \(V_S\) is the displacement volume of a ship (average). Equation 2 needs to be applied for each chamber using the known historic salinities as well as the estimated salinity value of the lower level water given by \((3)\):

\[
S_N = S(0) + N[S(4) - S(0)]
\]  

(3)

Equation \((3)\) provides the salinity value by lock (e.g., Miraflores, Pedro Miguel, or Gatun locks) for any quantity \(N\) (Number of lockages - supposedly only one ship passes through each lockage, so the number of lockages should be the same as the number of ships).

The six locks have different volumes and geometric characteristics so that ships of different drafts can cross the Panama Canal from the Pacific Ocean to Gatun Lake. The levels shown in Figures 3 and 4 are used with the dimensions shown in Table 1 to calculate the volumes and salinities of the locks when ships cross. In Figure 1 all levels, heights, and depths are referenced to the “Precise Panama Canal Level Reference” (PLD) that matches sea level.

<table>
<thead>
<tr>
<th>Lock</th>
<th>Width (m) × Length (m)</th>
<th>Height over lock (m)</th>
<th>Depth under or over PLD (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>33.53 × 320.34</td>
<td>7.92</td>
<td>-15.85</td>
</tr>
<tr>
<td>L2</td>
<td>33.53 × 326.44</td>
<td>8.53</td>
<td>-6.20</td>
</tr>
<tr>
<td>L3</td>
<td>33.53 × 326.44</td>
<td>9.45</td>
<td>+3.44</td>
</tr>
<tr>
<td>L4</td>
<td>33.53 × 326.44</td>
<td>8.53</td>
<td>+3.96</td>
</tr>
<tr>
<td>L5</td>
<td>33.53 × 320.34</td>
<td>8.83</td>
<td>-4.67</td>
</tr>
<tr>
<td>L6</td>
<td>33.53 × 320.34</td>
<td>8.53</td>
<td>-13.50</td>
</tr>
</tbody>
</table>

Table 1: Physical dimensions of the Panama Canal Panamax Locks - Panamax is a popular term for the size limits for ships traveling through the Panama Canal.
Using the equations of exchange of salinity in the locks, it is possible to set a numerical and differential equations model to define the salinity in Gatun Lake ($S_{GL}$), taking into account the exchange of water (and salinity) in the upper locks of Pedro Miguel and Gatun, and the water contribution by lakes Gatun ($V_{GL}$) and the volumetric inflows of Madden ($V_{Madden}$) and the river tributaries ($V_{trib}$) that flow into these lakes. Madden Lake is a reservoir of water which acts as additional water storage for the canal. This relationship is expressed by the following equation:

$$d(S_{GL})/dt = \frac{V_{GL} \cdot S_{GL} + (V_{Madden} + V_{trib}) \cdot S_{Madden} + (V_{L3} - V_{s}) \cdot S_{L3} \cdot EX_{L3} + (V_{L4} - V_{s}) \cdot S_{L4} \cdot EX_{L4}}{V_{GL} + V_{Madden} + V_{trib} + (V_{L3} - V_{s}) \cdot EX_{L3} \cdot N + (V_{L4} - V_{s}) \cdot EX_{L4} \cdot N}$$ (4)

and

$$d(V_{GL})/dt = (V_{Madden} + V_{trib}) + (V_{L3} - V_{s}) \cdot EX_{L3} + (V_{L4} - V_{s}) \cdot EX_{L4} - \text{Evaporation (t)} + \text{Precipitation (t)} - \text{Panamax Lockages Losses}$$ (5)

$V_{L3}$ and $V_{L4}$ are the volumes of Locks L3 and L4 respectively. $S_{L3}$ and $S_{L4}$ are the respective salinities of L3 and L4 taking into consideration the measured salinity gradients. EX$_{L3}$ and EX$_{L4}$ are the exchange ratios for L3 and L4 respectively. Piecewise linear profiles of Evaporation and Precipitation are added to the calculations of $V_{GL}$[6].

We can provide the same logic and add for the Gatun Lake equations for the new set of locks being built for the expansion of the Panama Canal. The equations for the Miraflores Lake will stay the same. The new Post-Panamax locks are given by the following notation:

L7: Lowest Post-Panamax lock connected to the Pacific Ocean
L8: Middel Post-Panamax Lock on the Pacific Side
L9: Highest Post-Panamax Lock on the Pacific Side directly connected to Gatun Lake
L10: Highest Post-Panamax Lock on Atlantic side directly connected to Gatun Lake
L11: Middle Post-Panamax lock on the Atlantic side
L12: Lowest Post-Panamax lock connected to the Atlantic Ocean

Table 2: Physical dimensions of the Panama Canal Post-Panamax Locks. Post-Panamax is a popular term for ships larger than Panamax that do not fit in the canal, such as supertankers and the largest modern container ships [7]

<table>
<thead>
<tr>
<th>Lock</th>
<th>Width (m) \times Length (m)</th>
<th>Height over lock (m)</th>
<th>Depth under or over PLD (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L7</td>
<td>55 \times 427</td>
<td>8.83</td>
<td>-20.62</td>
</tr>
<tr>
<td>L8</td>
<td>55 \times 427</td>
<td>17.37</td>
<td>-10.67</td>
</tr>
<tr>
<td>L9</td>
<td>55 \times 427</td>
<td>25.91</td>
<td>-2.41</td>
</tr>
<tr>
<td>L10</td>
<td>55 \times 427</td>
<td>25.91</td>
<td>-2.41</td>
</tr>
<tr>
<td>L11</td>
<td>55 \times 427</td>
<td>17.30</td>
<td>-10.49</td>
</tr>
<tr>
<td>L12</td>
<td>55 \times 427</td>
<td>8.68</td>
<td>-18.62</td>
</tr>
</tbody>
</table>
The Exchange Coefficients are different from the Panamax Locks due to the utilization of water saving basins (WSBs). \( E_{X_{L9}} \) and \( E_{X_{L10}} \) are the exchange ratios for L9 and L10 respectively that take into consideration the WSBs recycling [7]. In addition, the losses due to the lockages of the Post-Panamax Locks are reduced by a factor of 64% due to the WSBs [7]. \( V_{PPs} \) is the volume of references for the Post-Panamax ships. Therefore E4 and E5 are modified accordingly. Equations 6 shows the additions to Equation 5.

\[
d(V_{Gz})/dt = (V_{Madden} + V_{trib}) + (V_{L3} - V_j) \cdot E_{X_{L3}} + (V_{L4} - V_j) \cdot E_{X_{L4}} - \text{Evaporation (t)} + \text{Precipitation (t)} - \text{Fresh Water Facilities (t)} - \text{Hydropower Plants (t)} + \text{Releases from Water Reservoirs(t)} - \text{Panamax and Post-}
\]

\[
\text{Panamax Lockages Losses} + (V_{L9} - V_{PPs}) \cdot E_{X_{L9}} + (V_{L10} - V_{PPs}) \cdot E_{X_{L10}}
\]

(6)

The different equations were implemented in AnyLogic [8,9]. AnyLogic provides a set of numerical methods for solving differential equations, algebraic-differential equations, or algebraic equations. AnyLogic chooses the numerical solver automatically at runtime in accordance to the behavior of the system. When solving ordinary differential equations, it starts integration with fourth-order Runge-Kutta method with fixed step.

2.2 Prediction of evaporation level through \( \varepsilon \)-SVR

Here we used support vector regression (SVR) for estimating the evaporation level by using as predictors the precipitation, temperature, humidity, salinity, windspeed and wind direction [10,11]. The SVR is the regression version of support vector machines (SVM) a well known supervised learning algorithm used for supervised classification and is based on separation by hyperplanes. In the regression setup we are given a set of \( m \) data point \( \{(x_1,y_1),(x_2,y_2),..., (x_m,y_m)\} \) where \( x_i \in \mathbb{R}^n, i = 1,...,m \) are the vector with the features and \( y_i \in \mathbb{R}, i = 1,...,m \) the output variable. Then under the support vector framework we wish to determine the optimal hyperplane defined by the parameters \( (w, b) \) as the optimal solution of the following convex optimization problem:

\[
\begin{align*}
\min_{w,b,\xi^i,\xi^i^*} & \quad \frac{1}{2} w^T w + C \sum_{i=1}^{m} \xi_i + C \sum_{i=1}^{m} \xi_i^*

\text{s.t} & \quad w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i \\
& \quad y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i^*
\end{align*}
\]

(7)

where \( \xi_i, \xi_i^* \geq 0, i = 1,...,m \). With \( \phi(.) \) we denote the kernel function. Through the kernel function we can generalize the method in nonlinear high dimensional spaces and thus overcome the limitations related to the linear nature of the basic formulation. Each feature was normalized in order to have zero mean and unitary standard deviation. This is a standard preprocessing step that allows all the predictors to have the same weight in the model. The output variable was left as it is since it doesn’t affect the model training. The experiments were performed in 2.7GHz Intel Core i5 with 4Gb of RAM and for SVR implementation libSVM was employed through Matlab (ver 2011b) interface. For the purpose of this problem standard radius basis function (RBF) kernel was used with parameter \( \sigma=0.25 \). The parameter \( \varepsilon \) was set to 0.1 and the cost parameter of SVM C was set equal to 1. At each step there was a sliding window was used in order to select the sample to train the algorithm. The window parameter was set to three. The average per sample square error \( R^2 \) was found to be equal 0.0427 (or 4.27%). The short term prediction results are shown in the following Figure 4:

![Figure 4: SVR’s performance (evaporation)](image)
2.3 Mapping of the Volume to Height (Gatun Lake) using Neural Networks

The modeling of the Panama Canal Expansion includes IF-THEN rules based on the level of the lakes and the respective volumes. In order to facilitate the execution of these rules, we have to develop a mapping from the volume to the height of Gatun Lake. A neural network was developed to perform this mapping. The data was obtained from a comprehensive study of Gatun Lake [6]. The observed data is based upon pre-impoundment surveys from the early 1900’s and include estimation of changes that have occurred due to sedimentation in the last ninety years. The mapping is not straightforward because of Lake Gatun’s irregular shape. The reported bathymetry of Lake Gatun with its numerous bends and small embayments is extremely difficult to match. Therefore, neural networks are good for this mapping [6].

The architecture is of one input (Volume), two hidden-units in one single hidden layer and one output (Height in meters) (Figure 5). It is basically a look-up table with high accuracy and easy to execute during the execution time of the hybrid model.

3. System Dynamics and Discrete-Event Model

Different interviews with different subject-matter experts (SMEs) to learn about the operations were performed in order to obtain the different distributions and times of the discrete-event processes. A discrete-event model was developed in AnyLogic [3,8,9], with the respective animations, Queues, Switches, Java Classes, and the Enterprise (i.e., discrete-event) Library. The Switches were complemented with Java statements to capture the logic of assignment of locks and the schedule of the Panama Canal. In addition, animation was added in order to support the visualization and validation by subject matter experts (Figure 6).
After the discrete-event model of the Panama Canal using the Panamax locks was built and validated, meetings with personnel of the Panama Canal Authority and projections of future transits of Post-Panamax ships were also obtained. This supported the addition of the processes and rules to model the new Post-Panamax locks. The projections utilized for Post-Panamax traffic in the future are shown in Table 3.

Table 3: Daily (projected) traffic for the new Post-Panamax lane. Panamax Plus vessels are of Panamax size with more than 12.04 m of draft. Post-Panamax can be containers or tankers with more than 14 m of draft [5]

<table>
<thead>
<tr>
<th>Type of Vessels</th>
<th>Years: 2015-2020</th>
<th>Years: 2020-2025</th>
<th>Years: 2025-2030</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panamax Plus</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Post-Panamax</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total Daily Traffic</td>
<td>4</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

4. Hybrid Model

AnyLogic has capabilities for hybrid modeling. We combined the system dynamics model with the discrete-event model into one model using the capabilities of the Active Object Class from AnyLogic [8,9]. The Active Object Class from AnyLogic may have multiple concurrent activities that share object local data and object interface. Activities can be created and destroyed at any moment of the model execution. An activity can be described by a Java function or by a hybrid statechart. The currently active set of equations and triggers is defined by the current simple state and all its containers. Therefore, we can have a discrete-event process and the continuous process running concurrently, sharing information, and influencing their behavior. The discrete-event model feeds the number of lockages at specific discrete even times and the set of differential, algebraic equations, IF-THEN rules, the neural network and the SVR are running in continuous time modeling the salinity influenced by the lockages (Figure 7).

5. Validation and Prediction

The hybrid model was validated using actual transits and salinity field measurements:
- Actual data of Transits (Years: 2000 – 2009)
- Salinity Level for the Gatun Lake from field measurements (Years: 2003 – 2009)
The operational model was developed mainly with interviews with the Panama Canal pilots and data obtained for each event. Our validation was performed with actual data of the transits for the current Panama Canal. In addition, the simulation model was executed generating 100 sets of independent replications (using different random numbers). We compared the simulation with the system by constructing a confidence interval (Table 4). The validation of the Expansion was just performed by using subject matter experts of the Panama Canal Authority. They agreed that the model approximates and obtains the projected transits of the Post-Panamax locks.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Simulation Output</th>
<th>Historical Data (2000 – 2009)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit Rate/Year</td>
<td>13416 [13173, 13659]</td>
<td>13328 [12985, 13671]</td>
</tr>
</tbody>
</table>

The Salinity level output of the hybrid model was used to compare with those from the real system for the current situation (no Post-Panamax locks). The values from the real system were provided by field measurements conducted by the Panama Canal Authority in several points of Gatun Lake. In addition, these values and the output of the models were discussed by experts in the different engineering fields. Figure 8 provides the comparison of the hybrid model against the actual measurements.

In addition, the output in salinity of our model is compared against two other models for salinity of the current Panama Canal. The model developed by the Army Corps of Engineers predicts the salinity in the years 2003 – 2009 to be stable at a value of 0.032 ppt [4]. The Army Corps of Engineers did not consider precipitation, evaporation, hydropower plants, fresh water facilities, and spillways flows. The other model was commissioned by the Panama Canal Authority to the company Delft Hydraulics (http://www.wldelft.nl/) [5]. This model does not consider
precipitation and evaporation and uses a constant transit rate of 36 ships/day for the current Panama Canal. It is visually clear that the hybrid model has a higher quality than the previous models.

6. Conclusions and Further Research
This research provides a unique example for applying hybrid modeling. Hybrid modeling can benefit organizations with complex systems by providing them with a modeling environment which takes into account the internal and external changes taking place in their systems where continuous and discrete variables are present.

Currently, we are refining the model and adding more animations. Different scenarios with the Post-Panamax locks are being simulated. We will report in future papers the results of these developments.

Acknowledgements
This work was sponsored by the Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT) of the Republic of Panama under contract/project number FID09-079.

References