METHODS FOR PREDICTING WATER SERVICE VULNERABILITIES TO STORM EVENTS IN PUERTO RICO

Terra R. Edenhart-Pepe
Masters of Environmental Studies
Resilience and Adaptation

Spring 2019

Primary Reader: Sarah Willig
Secondary Reader: Dana Tomlin

PREDICTING WATER SERVICE VULNERABILITIES TO STORM EVENTS IN PUERTO RICO

Terra R. Edenhart-Pepe
Concentration: Resilience and Adaptation

Readers: Sarah Willig, Earth & Environmental Science, University of Pennsylvania
Dana Tomlin, PhD, School of Design and Earth and Environmental Science, University Pennsylvania

---

The 2017 hurricane season in Puerto Rico caused wide-spread and lasting damage to the island's infrastructure (Johnson, 2016; NSF, 2018; George Washington University, 2018; NOAA, 2018). Power service interruptions or total outages were reported up to a year after Hurricane Maria (George Washington University, 2018; NOAA, 2018). Cooperative systems, such as the drinking water supply, also experienced widespread failures due, in part, to water treatment and pumping facilities' dependence on a functioning electrical grid. The (Suomi-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) monitors power outages in disaster-affected areas by identifying missing city lights (Cole, et.al., 2017).

This study evaluates the capabilities of four different modeling methods, inverse distance weighted interpolation, ordinary Kriging interpolation, cartographic modeling, and machine learning, to forecast Puerto Rico Aqueducts and Sewers Authority's drinking water service vulnerability to failure in the next storm event. Given the association between electric and water systems, data from the Suomi-NPP satellite can be used as a proxy to estimate water supply failure after Hurricane Maria. Each method is subjected to a trial objective and available data. Model performance is summarized and methodological requirements are derived from characteristics of use and desired outputs. Results suggests that cartographic modeling and kriging are most the appropriate methods for identifying areas vulnerable to water supply interruption.

Index

Chapter 1 | INTRODUCTION

Introduction

Anthropogenic climate change is responsible for an increase in both the frequency and severity of disaster events that damage infrastructure and interrupt utility service for many clients and for long periods of time (IFRC, 2016; Guikema, 2010). Subsequent societal damages can be even more long lasting than storm impacts to infrastructure, and recovery periods can be further delayed when storm events are compounded. Such was the case for multiple Caribbean islands, which experienced aggravated adverse effects from hurricanes Irma and Maria in 2017 (NSF, 2018).

Image credit: The Atlantic

Image credit: The Atlantic

Image credit: WSVN https://sunbeamwsvn.files.wordpress.com/

In the year since Maria, 50 percent of Puerto Ricans say people in their households could not get enough water to drink, according to a new Washington Post-Kaiser Family Foundation survey.

About 2 in 10 say they drank water from a natural source such as a stream or river after the hurricane hit in September 2017. - Washington Post

Image Credit: CNN, Carlos Giusti/AP

Image credit: https://ioneglobalgrind.files.wordpress.com/

Image credit: https://ioneglobalgrind.files.wordpress.com/

Image credit: https://ioneglobalgrind.files.wordpress.com/

Puerto Rico was among those islands that sustained catastrophic storm damage to land, resources, farms, infrastructure, and communities. Hurricane Irma skirted Puerto Rico on September 6, 2017, leaving more than a million people without power (Johnson, 2016), and Hurricane Maria made landfall in Puerto Rico two weeks later on September 20, exacerbating damage and slowing recovery (U.S. Official, 2018; Sullivan, 2018). Cooperative systems, such as the drinking water supply, also experienced widespread failures due, in part, to water treatment and pumping facilities' dependence on a functioning electrical grid (Atienza-Diaz, E., personal interview, February 1, 2019).

Not only is water service failure an infrastructure dilemma but it also causes indirect societal damage. Literature suggests that the prosperity of modern urban environments is dependent upon well-functioning infrastructure networks, with reliable water distribution being critical (Ossi, 2018; Puentes, 2015); generally, U.S. water customers, including those within its territories, expect uninterrupted water supply within service areas (Black and Veatch, 2018). However, service failure is inevitable, especially in the aftermath of storm events. But in the case of Hurricanes Irma and Maria, Puerto Ricans experienced extreme circumstances; more than a year after these storms, the island still experiences regular service interruptions (George Washington University, 2018; NOAA, 2018). The length of failure time and magnitude of the problem could be grounds for categorizing the aftermath of these storms as 'catastrophic,' and the region is regarded as high-risk to future storm damage (George Washington University, 2018; NOAA, 2018). Future preparedness

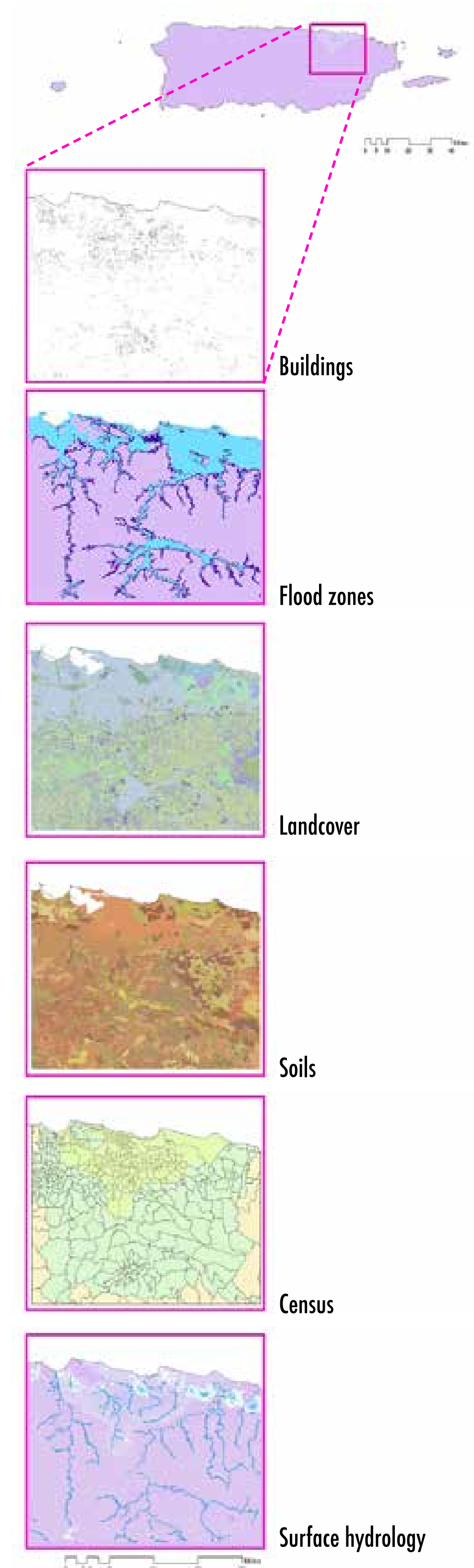


Fig. 1.1. Existing Data

planning may be guided by the discrete areas of higher and lower risk apparent within the island's infrastructure systems and landscape.

This research began nearly a year after Hurricane Maria when aid and recovery efforts were still active. Media headlines regularly reported continued struggles to regain the quality of life previously enjoyed by Puerto Rico's residents, with a particular emphasis

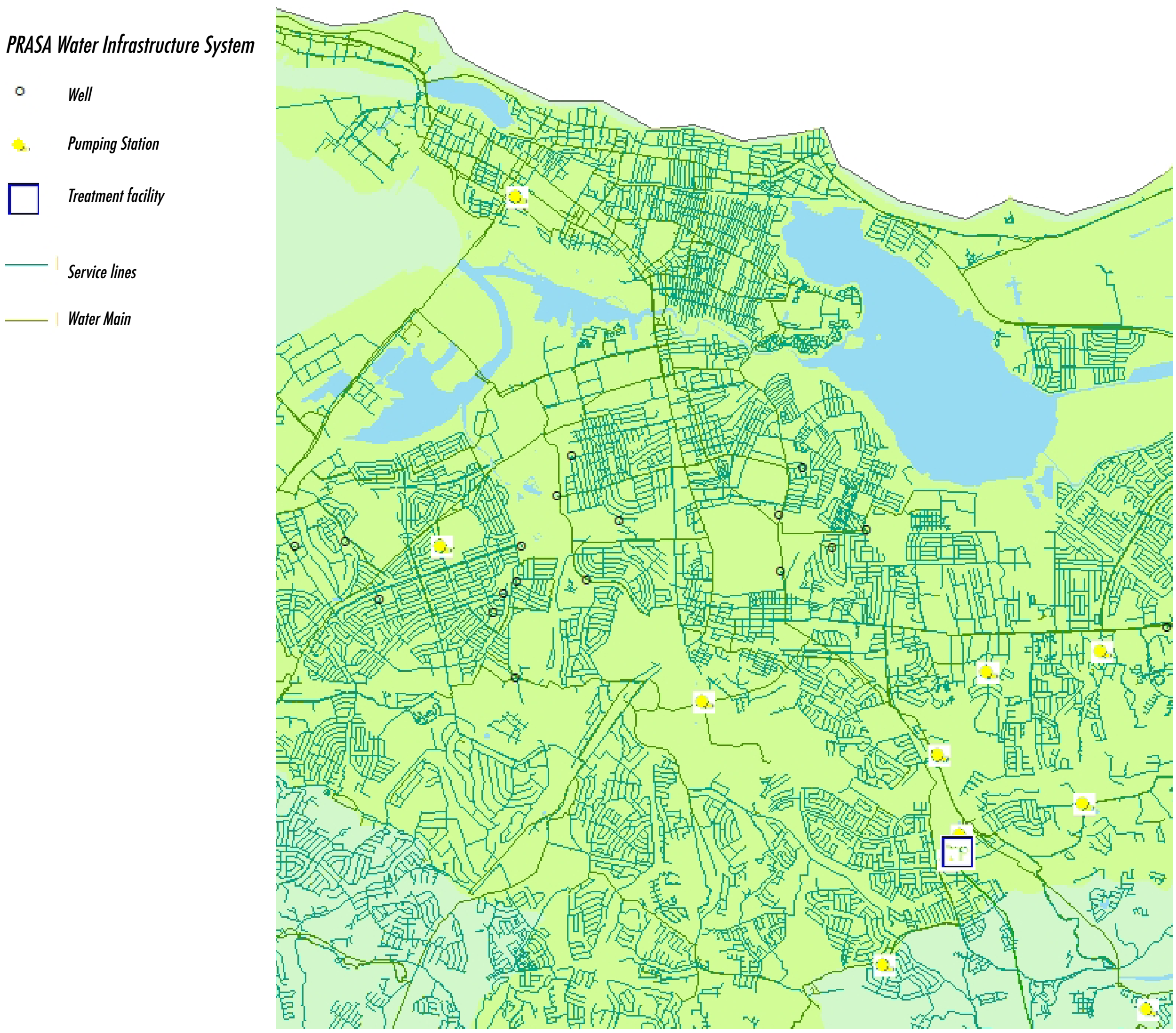


Fig. 1.2. PRASA water infrastructure, retrived from PR government open data portal

on regaining power and water. The flood of media attention on this catastrophe served as motivation to develop a viable methodology for identifying the likelihood of drinking water supply failure, and resulting vulnerability in the next storm event.

Purpose

This study evaluates four methods for forecasting vulnerable areas: Inverse Distance Weighted interpolation (IDW), Ordinary Kriging (Kriging) interpolation, cartographic model, and machine learning. Vulnerable areas are those areas with high likelihoods of PRASA drinking water service failure in the next category 4 or higher storm event. The purpose of this paper is to present an overview and evaluation of each method. The disparate methods are best evaluated when tested with real questions and data. The trial question used for evaluation is: *based on the likelihood of failure within the*

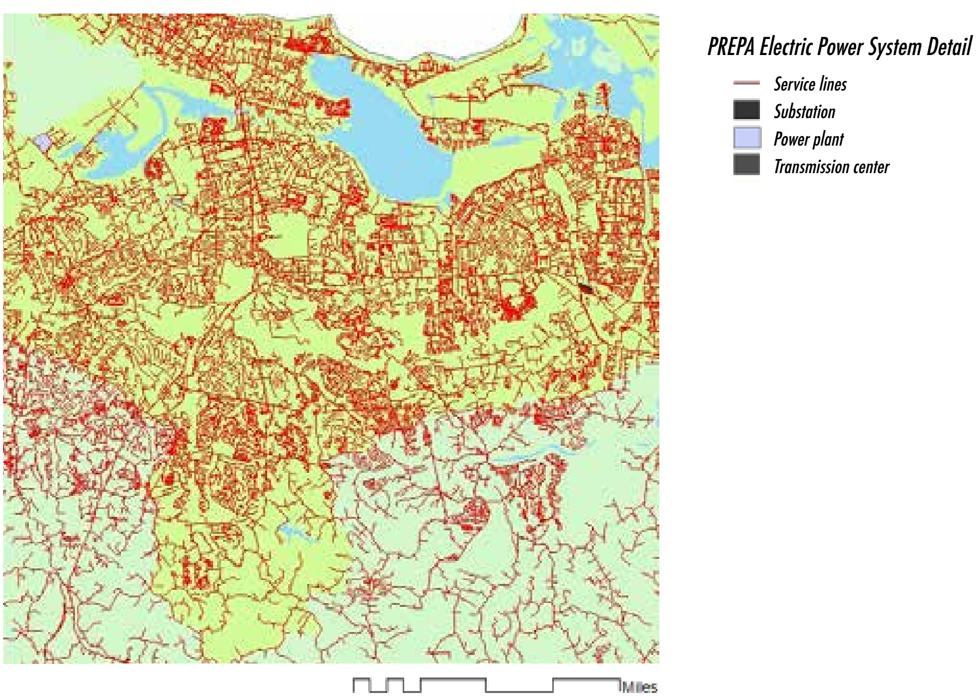


Fig. 1.3. PRASA water infrastructure, retrived from PR government open data portal

Puerto Rico Aqueducts and Sewers Authority (PRASA) drinking water supply system during an eighteen-day period after Hurricane Maria, what are the island's most vulnerable areas in storm events? Methodological requirements are derived from general characteristics of use and desired quality of outputs. Finally, each method is evaluated with respect to the requirements.

Image credit: The Atlantic

Image credit: AZCentral

Image credit: The Atlantic

This study is comprised of standard research paper sections, but the project-specific details of each section are as follows. The methods section explains the four models under evaluation and explicitly outlines the steps taken to forecast failure and vulnerable areas. The results section evaluates each method based on a set of requirements. It is suggested that methods are best used in conjunction. The paper concludes with perspectives on data needs and future study opportunities.

Chapter 2 | METHODS

Methods

"A time when knowing something in any practical sense so often calls for representing that something in digital form, our knowledge of the world around us has become increasingly influenced by the manner in which that world is seen by geographic information systems (GIS)" (Tomlin, 2011). There are many ways to determine a target outcome with geographic information systems, such as the software employed in this study, ArcGIS and R. The four methods considered in this study are IDW and Kriging interpolation, cartographic modeling, and machine learning (fig.2.1). Progression of steps to resolve the target outcome are replicated for all methods and mimic the progression described by Dana Tomlin in Fundamentals of Geospatial

DELIVERABLES:

1. *Interpolation methods based on spatial autocorrelation*

2. *Cartographic modeling to suggest areas of likely failure based on environmental features*

3. *A logistical regression model to mathmatically explore relationship between exogenous features and failure*

4. *Methods evaluation and comparison*

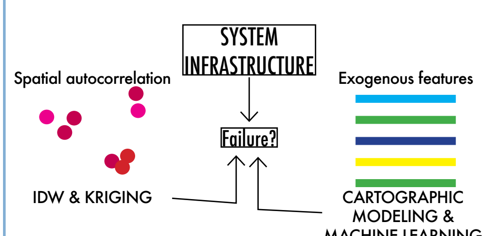


Fig. 2.1

Algorithm Design: observed, understood, confronted, resolved (2012). Each of the models receives either failure data and/or exogenous physical, environmental factors to rate the likelihood of PRASA water service failure and thereby identify which areas are most vulnerable in the next storm event.

The first two models, IDW and Kriging, use a statistical process called interpolation for finding intermediate values from a set of known values (Florinsky, 2016). The second two methods, cartographic modeling and machine learning, incorporate exogenous physical and environmental factors to make prescriptions or predictions, respectively. Cartographic modeling uses a series of descriptive, pixel-based maps (raster maps) in which “each pixel identifies a geographical location and indicates its condition” (Tomlin, 2011). The raster map layers constitute the set of failure related environmental factors. Each layer is subjected to spatial operations and overlaid to identify overlapping water failure indicators. The sum of which represent vulnerable areas (Tomlin, 1990). Machine learning utilizes an algorithm to mathematically perform the task of identifying significant failure-related features (Thrun, 1998). This study uses a machine learning algorithm called a logistic regression to model the probability of failure and identify areas vulnerable to drinking water service failure. The following section outlines procedures to build the training data sets and used in each method.

Monitoring & Analysis Procedures to Determine Failure

Electric power supplied by the Puerto Rico Electric Power Authority (PREPA) plays a significant role in the PRASA distribution system’s water service. Maria compromised 80 percent of the island’s power transmission lines, and approximately 86 percent of the island’s electricity customers were without power (Resnick and Barclay, 2017). PREPA has the benefit of monitoring infrastructure before, during, and following disaster via satellite.

The Suomi National Polar-orbiting Partnership satellite (Sunomi-NPP) was launched by The National Aeronautics and Space Administration (NASA) in collaboration with the National Oceanic and Atmospheric Administration (NOAA) in October, 2011 (Miller, et.al., 2012). The “Suomi NPP carries the Visible/Infrared Imager/Radiometer Suite (VIIRS), an optical spectrum (22 bands spanning ~0.4–13 μm) sensor providing imagery at high spatial resolution (0.375–1.6 km, band dependent) across a 3,000-km-wide swath” (Miller, 2012). VIIRS includes Day/Night Band (DNB) sensing (Lee, 2012) and the unique capability of remotely sensing infrastructure functionality in disaster-affected areas by recording artificial nighttime light (NTL) emissions (Cole, et.al., 2017). These emissions represent human activity in developed areas and change in emission levels “can be observed and monitored in DNB data as darkened or unlit pixels in locations where urban infrastructure (i.e., lit pixels) are expected during normal conditions,” (Cole, et.al., 2017).

Cooperative systems, such as the drinking water supply, also experienced widespread failures when water pumps lost power, leaving residents without water to drink, bathe, wash, or flush toilets. In many cases, both pumping stations and treatment plants ceased operation (PBS, 2018). According to Eli Diaz-Atienza, Executive President at PRASA, determining the condition and performance of a given pumping station in a storm event is related to susceptibility of power interruption and generator condition (personal interview, February 1, 2019). Given water system dependence upon electricity, this study uses DBN data as a proxy to estimate water supply failures after the storm.

PRASA maintains one of the largest systems in the United States, with more than 4,000 facilities and approximately 1.23 million customers, and failures in water service are caused by factors other than electric power interruption, such as, water main leaks, and source pollution (Diaz-Atienza, personal interview, 2019). However, for the purposes of this paper, drinking water service failures are assumed to have occurred in areas with reduced NTL emissions and areas served by PRASA facilities proximate to reduced NTL emissions (the process of estimating emissions reduction and water service failure will be further described in the methods section). While it should be noted that water quality must be good for a system to be considered functioning, this study is focused on drinking water availability, not water quality.

Analysis & Training Data

Given the complexity of raw data collection, due in part to the sensitivity of the VIIRS DNB, many NTL products have been developed. The NASA Black Marble team created a Puerto Rico-specific product series known as the NASA Black Marble High-Definition (HD) product. It combines the standard (500-meter) product with information from several other satellites and data sources, resulting in a 30-meter spatial resolution product with a normalized value of 0 to 1, color-enhancement and cloud removal. The Black Marble HD product makes for timely and quantitative use of NTL (Román, 2018; NASA, 2017).

NTL data from the Black Marble HD product suite (fig. 2.2) was downloaded from The Federal Emergency Management Agency (FEMA) GIS repository (https://data.femadata.com/). A post-event composite was assembled from the Black Marble product data. The data was acquired over a total of 18 days after the storm (three dates in September and three dates in October. The reference data was pre-event NTL representing normal conditions. Areas where artificial light emissions went from normal conditions to unlit were considered electric power outages. Observed power outages serve as a proxy for water supply system failures and were analyzed in the following ways:

- Failure zones are classified as failed (1) and not-failed (0). This grid of 0's and 1's was saved as a raster map (fig. 2.3) and used in the cartographic modelling method.
- PRASA facilities (pumping stations and treatments facilities), represented by geographically-located points, were given a rating (1 – 10) based on proximity to observed

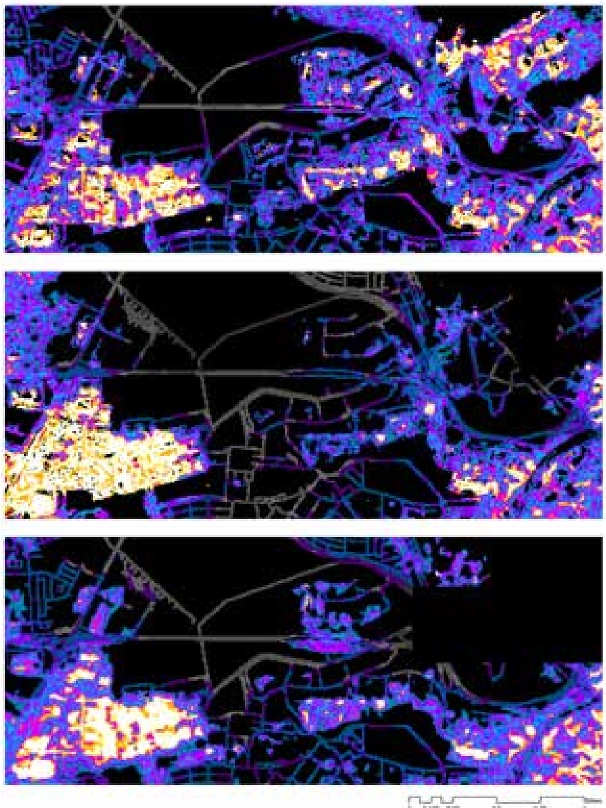
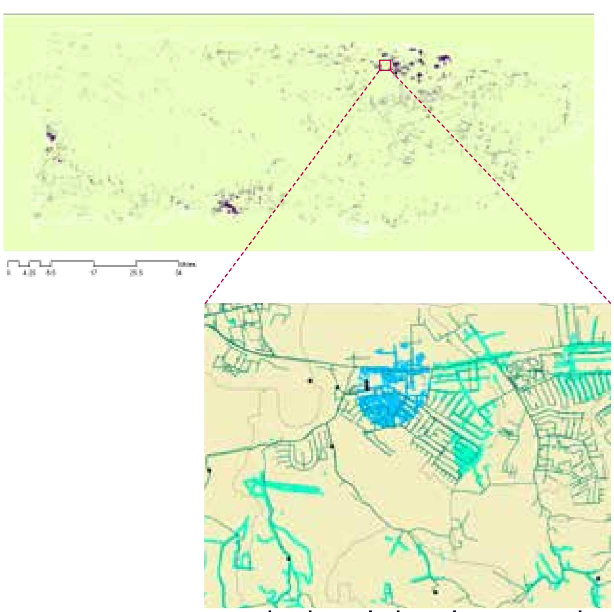


Fig. 2.2. Top to bottom: Pre-event, September, October night time lights (NOAA)



September change (reduction) is represented in blue; October change (reduction) is green

Fig. 2.3. Raster of September & October reduction in light

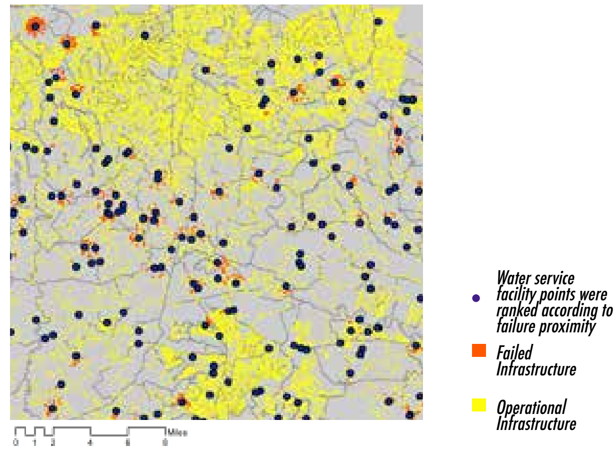


Fig. 2.3. Facility points with proximal infrastructure (data for interpolation)

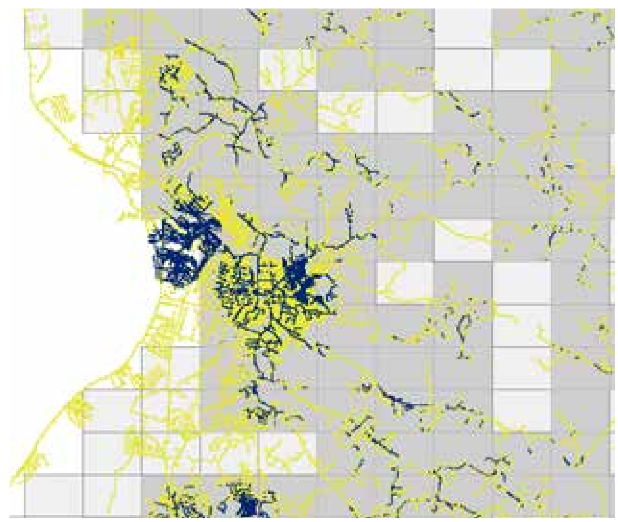


Fig. 2.4. Failure (in blue) within PREPA system is attributed to fishnet (input for machine learning model). The scale of the fishnet doesn't pick up all the detail available in raster, as seen at the edge condition

outage zones (fig. 2.3). The resultant new shapefile, or a map represented by points, lines, or polygons, in which each shape has a geographic location and a table of attributes (Tomlin, 2011), was used as the input for both interpolation methods.

- Lastly, the failure zones were joined to a fishnet, or a sampling tool that creates a feature class comprised of a net of cells (Esri, 2018). The fishnet used for the machine learning model had a cell size of 800 meters (0.497097 mi) (fig. 2.4) and binary values. This grid cell size was selected under the assumption that water service facilities would not be located within half a mile from each other, as well as to capture outage data's granularity. These three files (raster, point, and fishnet) are referred to as the *training data* for each associated methods.

IDW and Kriging Interpolation

Observation of the training data suggested that failure is autocorrelated, meaning that water service failures are correlated to other water service failures. Both interpolation methods showed that a reasonable prediction of drinking water service failure can be based exclusively on failure events themselves. The IDW interpolation operates in the method described by its name "inverse weighted distance", whereby closer things have a greater impact than farther things. This method is executed using a straightforward tool in ArcGIS which took as its input the point data created in the previous step. The second interpolation method, ordinary kriging, employed a gaussian model, a model type best suited for nonlinear problems with an element of uncertainty (Snelson, 2006). To improve forecasting, ArcGIS' Geostatistical Wizard, kriging tool withheld points for cross-validation and iterated 88 permeations of cross-validation.

An important factor of modeling is the ability to measure error, which was done for the IDW and Kriging methods. For IDW interpolation, an estimation of the error was manually calculated using field calculator and map algebra, while the kriging wizard provided error estimates automatically. Both error statistics will be discussed in the subsequent section.

These efficient and easy-to-use methods, begin with the failure outcomes and attempt to extract meaning from the modeled results, enabled a preliminary view of model error and insight into what other factors could be impacting failure.

Cartographic Modeling

Cartographic models can be used to describe of geographic phenomena, such as drinking water service failure, and to solve problems, like locating vulnerable areas are based on a set of criteria (Tomlin, 1990). This process followed the three major steps outlined in Dana Tomlin's "Prescriptive Modeling" chapter from Geographic Information Systems and Cartographic Modeling: state the problem, generate a solution, and evaluate the results.

First, the desired geographic quality was failure. Failure was considered a function of three indicators: flooding, exposure, and erosion. Preliminary observations confirmed that flood and erodibility were correlated with electric failures. It was also believed that wind damage might contribute to failure, although this was not observed. Landscape characteristics related to flooding, exposure, and erosion were subjected to spatial operations, which

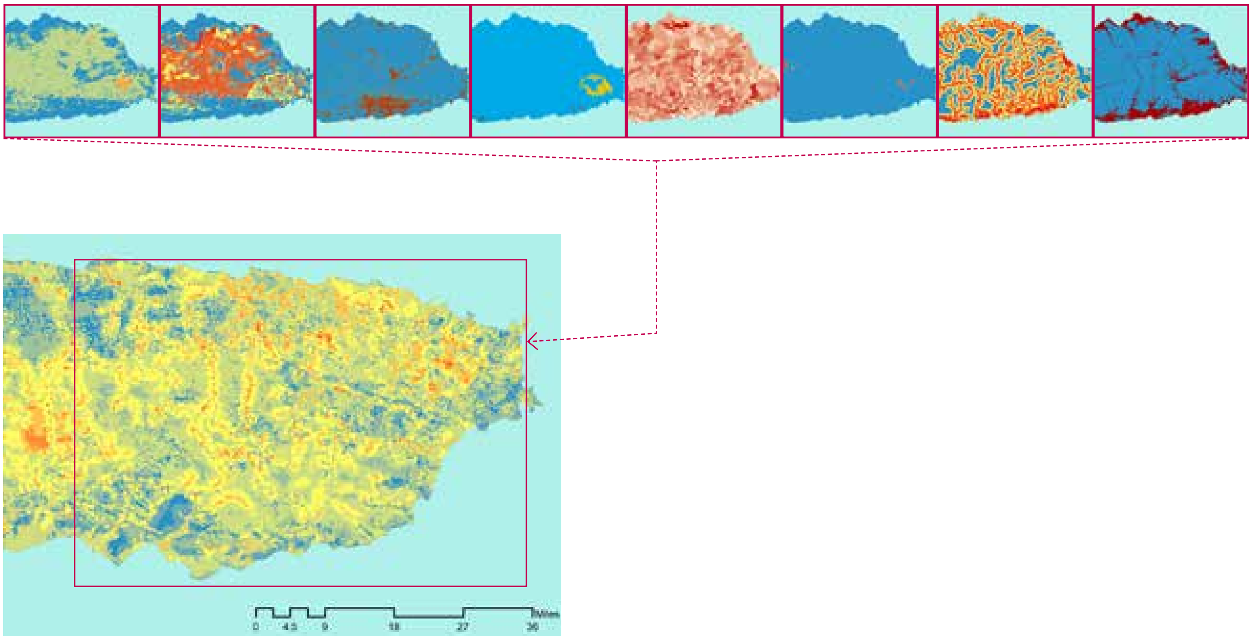


Fig. 2.5. Raster layer inputs to reverse suitability

resulted in the following layers: hydric soils, soil drainage quality, distance to steep slopes, proximity to urban development, quantity of wetlands, population density, weather exposure from high elevations, land cover, distance to surface hydrology, and flood zones (fig. 2.5).

Raster layers were then reclassified into failure factors, such that each grid cell was assigned a rating between 1 – 10, with higher ratings more likely to experience failure. The probability of facility failure was expressed by combining failure factors as a weighted sum overlay. Generally, prescriptive cartographic modeling "must express a range of potential site modifications as a function of existing site conditions and the geographic quality sought" (Tomlin, 1990). The modeling method herein was based on that rule, but ratings varied inversely with suitable water facility conditions. The result was a map in which high value zones correspond to high likelihood of failure, in effect, a "reverse site suitability analysis". Lastly, actual failure was compared to zones of highest failure probability (shown in Results section).

Machine Learning

According to Ken Steif, of Urban Spatial, there are two ways to predict failure. One way to accomplish this outcome is to ask a [PRASA engineer] to identify how each factor should be weighted, based on experience, then, overlay each scaled factor into a water service failure index (2018), similar to the cartographic model created in the previous method. While local or expert knowledge is undoubtedly essential and overlay is effective, Steif contends that machine learning is the method offering both the efficiency and precision lacking in other methods (Lecture, February 22, 2018). The fourth method evaluated by this paper is a machine learning model based on a logistic regression algorithm. The general process followed these steps: (1) state the problem, (2) import known failure data, (3) check for class bias, (4) create training and test samples, (5) observe important variables, (6) build logit models, and (7) predict on test data.

First, failure was considered a function of physical and environmental factors. Inputs to this model included a fishnet containing known failure points, encoded as *failed* or *not failed/ operational* and failure factors (fig. 2.6). The same failure factors used in the cartographic model were also used in the regression model. However, there were slight differences in the failure factor encoding for this model, due to the rigidity in the regression model method and the unit of analysis (fishnet grid). For example, features were adjusted when attributed to the fishnet, such that the raster layer representing steep slopes was converted to average distance from failure zones to steep slopes within the fishnet grid cell.

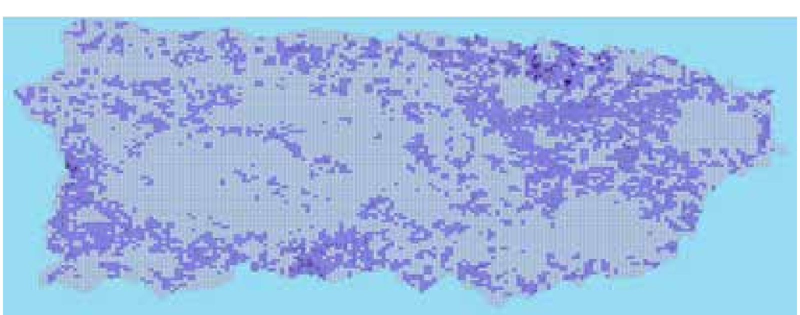


Fig. 2.6. Infrastructure failure across the island is indicated by darker color

The observations were sampled in approximately equal proportions. Ideally, the proportion of failure and non-failure in the Y-variable should be the same. However, given that failure events are limited to areas within the power grid across the entire island, there are many more non-failure events than failure events. Included below is a copy of the code that resamples data proportionally (fig. 2.7).

```
# Create Training Data
df1s <- prmw[which(prmw$ISISNT == 1), ]  # all 1's
df0s <- prmw[which(prmw$ISISNT == 0), ]  # all 0's

set.seed(4305)  # for repeatability of samples
df1_trainr <- sample(1:nrow(df1s), 0.7*nrow(df1s))
# 1's for training
df0_trainr <- sample(1:nrow(df0s), 0.7*nrow(df0s))
# 0's for training
training_df1 <- df1s[df1_trainr, ]
training_df0 <- df0s[df0_trainr, ]
Train <- rbind(training_df1, training_df0)
# row bind the 1's and 0's
# Create Test Data
test_1s <- df1s[-df1_trainr, ]
test_0s <- df0s[-df0_trainr, ]
Test2<- rbind(test_1s, test_0s)  # row bind the 1's and 0's
```

Cross Validation

Accuracy	Kappa
0.6810884	0.1928402

Fig. 2.7. Resampling code

```
prmwModl <- glm(ISISNT ~
          hydro_dis +
          stslp_dis +
          hiel_dis +
          hydsl_dis +
          ag_amt +
          ero_dis +
          wtld_amt +
          urblo +
          urbhi +
          hydro,
        data=Train, family=binomial(link="logit"))
```

Fig. 2.8. Model code

```
> cvFit1  Generalized Linear Model

12674 samples
  13 predictor
   2 classes: '0', '1'

No pre-processing

Resampling: Cross-Validated (100 fold)
Summary of sample sizes: 12547, 12547,
12548, 12547, 12548, 12547, ...
```

Fig. 2.9. Cross-validation

The methods embodied in this project design combines deductive and inductive methods to form a research strategy in which there is a continuous cycle between theory and observation. The project design intentionally promotes a cyclical process of pattern observation, leading to revision and development of theories. This was appropriate for the exploratory nature of the study.

```
glm(formula = ISISNT ~ hydro_dis + stslp_dis + hiel_dis +
hydsl_dis + ag_amt + ero_dis + wtld_amt + urblo + urbhi +
hydro,
        family = binomial(link = "logit"),
        data = Train)

Deviance Residuals:
   Min       1Q       Median    3Q       Max
-2.3041   -0.8383   -0.6665   1.1615   2.2378

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.423e-01  1.029e-01  -4.297 1.73e-05 ***
hydro_dis   -3.839e-05  2.347e-05  -1.636 0.101905
stslp_dis   -1.160e-02  1.805e-03  -6.429 1.29e-10 ***
hiel_dis     4.139e-06  2.662e-06   1.555 0.120003
hydsl_dis   -4.771e-06  6.487e-07  -7.355 1.91e-13 ***
ag_amt      -4.714e-01  1.342e-01  -3.512 0.000445 ***
ero_dis     -5.248e-06  4.268e-06  -1.230 0.218868
wtld_amt    -1.258e+00  3.600e-01  -3.494 0.000476 ***
urblo        9.452e+00  4.751e-01  19.893  < 2e-16 ***
urbhi        7.255e-01  1.421e-01   5.105 3.30e-07 ***
hydro       -3.217e-01  1.081e-01  -2.977 0.002912 **
—
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

   Null deviance: 11311  on 8870  degrees of freedom
Residual deviance: 10360  on 8860  degrees of freedom
AIC: 10382

Number of Fisher Scoring iterations: 4
```

Fig. 2.10. Model Summary

Chapter 3 | RESULTS

Results

Methodological evaluation for each of the four models is an essential, yet challenging part of this project because each method is fundamentally different. The following section seeks to identify a reasonable list of requirements for each of the four models, review the predictions and model evaluations, and summarize preferences and future study improvements.

Requirements

Method requirements combine a subjective assessment of model quality with objective measures of predictive power and other common diagnostics, where applicable. The cumulative list of requirements attempts to provide a scale to measure the validity of each method in relation to each other.

First, functionality (1), or an available range of operations that can be adjusted to suit a particular predictive purpose, and versatility influence a future analyst's decision to match an method with a specific spatial problem. Cartographic modeling manifests functionality; there is a multitude of nuanced ways to express results. Kriging demonstrates functionality by way of the Geostatistical Wizard, which offers a variety of parameters that can be adjusted to improve accuracy. Ease of use and time-effective (2) qualities are self-explanatory but play a significant role in the likelihood that a method will be replicated to address the problem of water service failure. One of the best ways to assess model quality is to gauge the usefulness of visualizations and other output (3) to planning processes. Usefulness is determined, in part, by the output's accessibility. While machine learning provides a variety of model diagnostics, its interpretability can be less evident to decision-makers than a compelling visualization. (4) The magnitude of error measure, or how well or poorly predicted results match actual results, is a basic but essential evaluation method whereby a low mean absolute error (MAE) suggests well-matched results. A model's usefulness will vary with context, situation, and objectives. Therefore, higher levels of generalizability (5) yield a better model. A model's capability (6) is considered advanced if the model outcome surpasses description and achieves prescription or explanation of cause and effect relationships. (7) A model's performance is increased when it has built-in model evaluation methods, such as hold -out (employed in Kriging interpolation) and cross-validation (employed in the machine learning logistic regression model). Finally, models are often evaluated based on a statistic that measures how well you can predict the dependent variable based on the independent variables (Allison, 2014). This statistic is similar to the MAE, but relies on a (8) Receiver Operation Characteristic (ROC) curve and a high area under the curve (AUC) value. These metrics do not measure magnitude of error, but rather the direction of error, which quantifies the model's ability to discriminate between positive and negative classes (Allison, 2014).

Summary of methodological evaluation requirements:

1) High functionality and versatility
2) Ease-of-use & time-effective
3) Output usefulness
4) Low magnitude of error
5) High generalizability
6) Capability of achieving advanced outcomes
7) Capability of including model evaluation

8) Appropriate predictive-power

*Note: R2 values and explicit "goodness-of-fit" evaluations are intentionally excluded from these parameters because the concern here is predictive power, which is easier to compare across methods, and because researchers lack consensus on the best way to calculate R2 for the logistic regression used in this study (Allison, 2014).

Interpolation

Below are the predictions for both interpolation models (fig.3.1 & 3.2).

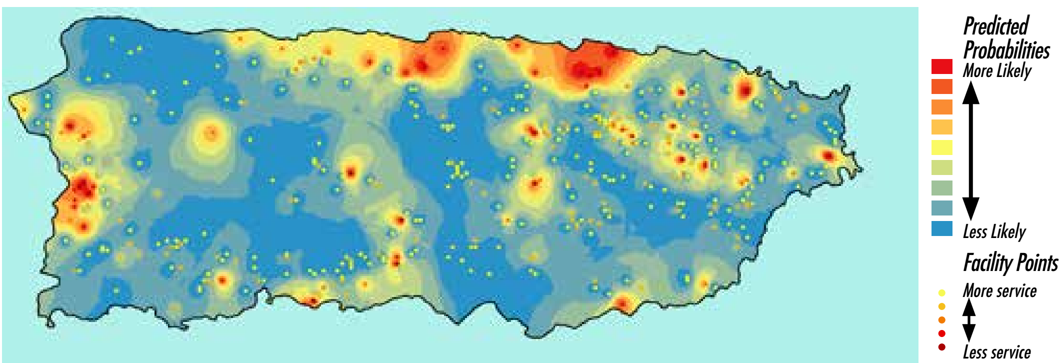


Fig. 3.1. Failure Zone Interpolation: IDW

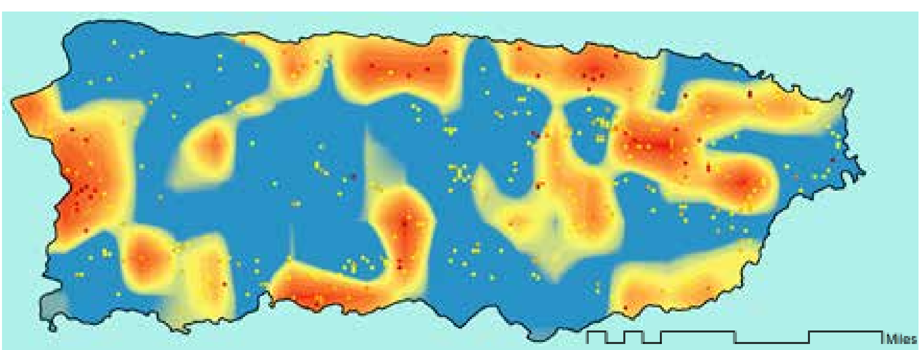
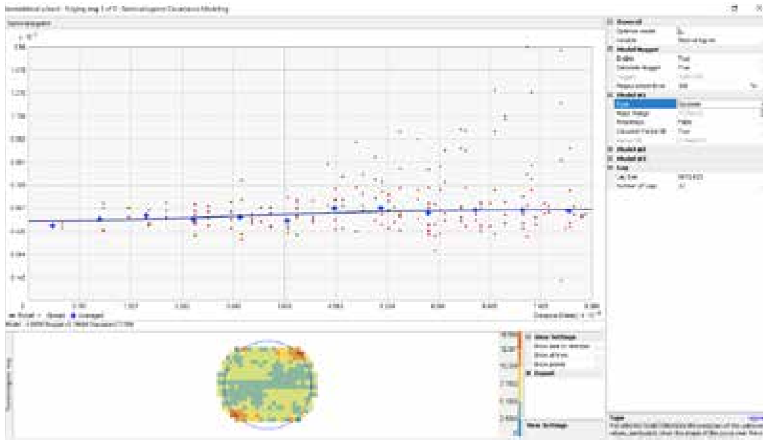


Fig. 3.2. Failure Zone Interpolation: Kriging

IDW Mean error (2.3705)	Kriging Mean error (2.27)

The IDW is easy to use but less functional than the Kriging method. Figure 3.4 illustrates the neighborhood type details and the user interface at which adjustments can be made to scale (as with the IDW method, five nearest neighbors were used). If we make assumptions about scale, then prediction is simple because failure-related factors can be used to predict which areas are likely to experience PRASA service failure in future storms. However, scale assumptions can account for the error displayed in these methods (this will be summarized at the end of the interpolation evaluation). While this is a time-effective analysis, there are numerous subjective processes that may drastically impact its results.

The hotspot map outputs show “a spatial concentration of some distinctive geographic condition” (Tomlin, 2009). In this case, the geographic condition is “failure” or “not-failure.” These maps are popular and relatively intuitive, making them compelling visualization tools. Due to the potentially subjective nature of these models, it may be argued that their usefulness is reduced. These methods are most useful as a graphic illustration and not as a rigorous display of data. However, there are quantifiable accuracy measures researchers can use to assess the confidence of their predictions.



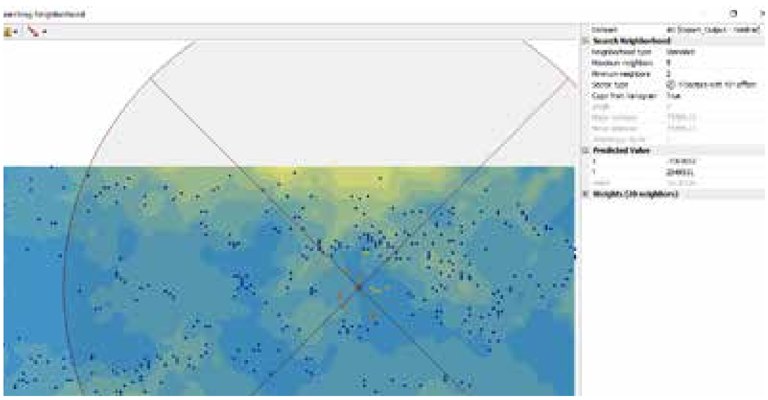
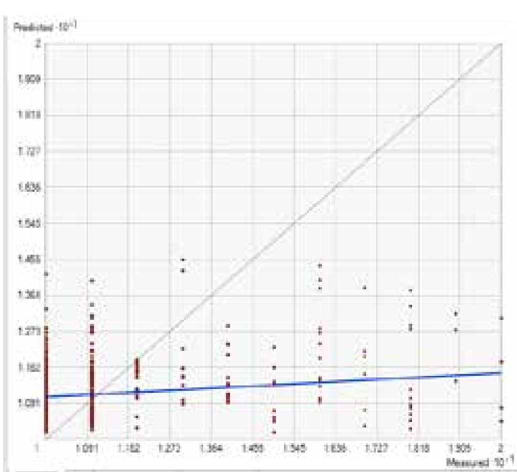


Fig 3.3 & 3.4. (top to bottom) Kriging statistics Semivariogram and model error statistics



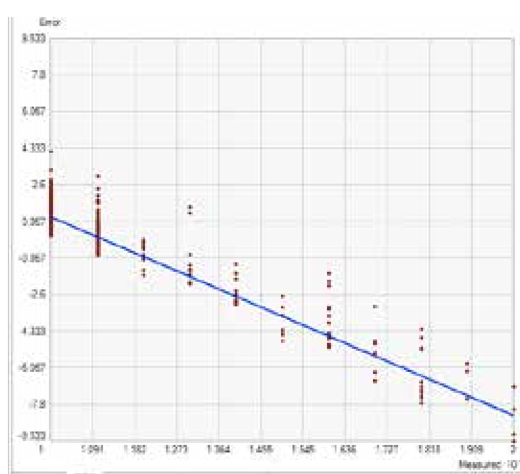


Fig 3.5 & 3.6. Kriging statistics, error statistics

The IDW interpolation can be converted to points in order to calculate MAE via the field calculator. The IDW error is higher than the Kriging MAE, which could result from the wide-ranging error values (.0002 to 10012) or the additional steps performed by the Kriging method. Error can also be mapped (fig. 3.7) to observe relationships between failures and existing site conditions. The IDW interpolation generalizes relatively well across the island's landscapes. However, the error map shows higher levels of error clustered around areas of predicted vulnerability and, geographically, in the island's northeast quadrant. There are exogenous conditions here that are not represented by spatial autocorrelation alone.

One of the downfalls to the interpolation methods is that inferential knowledge is limited in these cases. Neither of these methods answers questions such as "what is the causal effect of distance to surface hydrology on probability of failure?" Nor do these methods answer "to what extent do steep slopes contribute to water service failure?"

Unlike the IDW method, the Kriging method automatically generates model evaluations in the form of a measured value, predicted value, and error. The Semivariogram illustrates a function that describes the underlying spatial autocorrelation pattern (fig. 3.4).

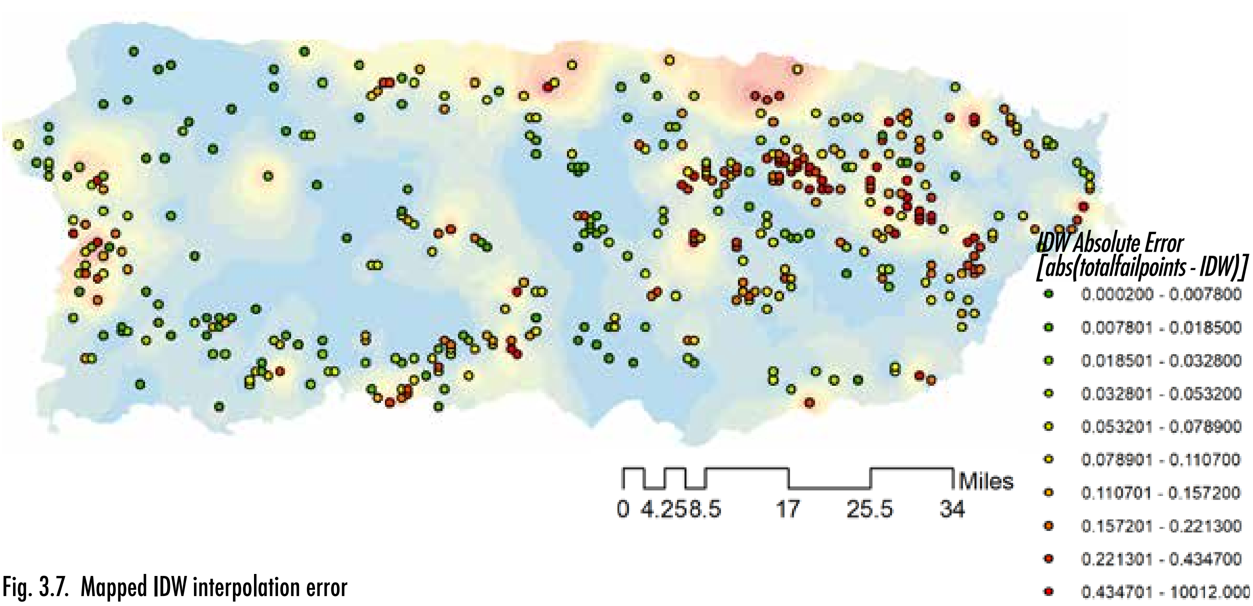


Fig. 3.7. Mapped IDW interpolation error

The error appears to be lower closer to the measured points, as might be expected. The Kriging method can also improve failure forecasting due to cross validation, which predicts for each point in turn (fig.). The final diagnostic plots provided by the Kriging method are predicted values (y-axis) as a function of observed values (x-axis) (figures). The black line shown on the plot would be a perfect prediction; if the measured values matched the observed values, all points would be on the black line. The blue line indicates the function line. The deviation of the two lines is another measure of 'goodness-of-fit.' These evaluation methods exhibit magnitude of error but do not explain granularity of results.

Scale and simplicity are two aspects of the interpolation methods which may account for the prediction error measures. As previously mentioned, interpolation is a method that predicts values from a sample set of known points (ESRI, 2016). The IDW interpolation operates on the assumption that closer variables are weighted more heavily than variables that are farther away (Steif, Lecture. February 22, 2018). In other words, the farther away from the failure points, the lower the correlation of failure to any given point. Following this logic, there is a threshold distance at which autocorrelation ceases. Therefore, analyzing data at a contextually appropriate scale for Puerto Rico is essential for this method. The appropriate scale on an island, like Puerto Rico, could be different than the scale appropriate for an analysis in landlocked regions. Scale also varies across built topographies and natural environments. A trial scale of the five nearest neighbors was used in this method; however, that scale may not be the most appropriate for this analysis.

This method is facilitated by a pattern in which failure begets failure. This could be the result of outdated systems lacking redundancy to protect against failure in a particular area (Neukrug, personal interview, October, 24, 2018). It is worth asking whether prediction based exclusively on spatial autocorrelation yields an overly simplified result. According to Eli Diaz-Atienza, there are many factors contributing to failure (personal interview, February 1, 2019) and predicting solely on other failures leaves those factors out, reducing the model's viability .

Cartographic Modeling

Below is the output from the cartographic model (fig. 3.8) and resolution detail (fig. 3.9). The predictions show an island-wide spatial failure pattern, which is similar to the interpolation methods. The resolution of the cartographic model output is finer than outputs from interpolation methods.

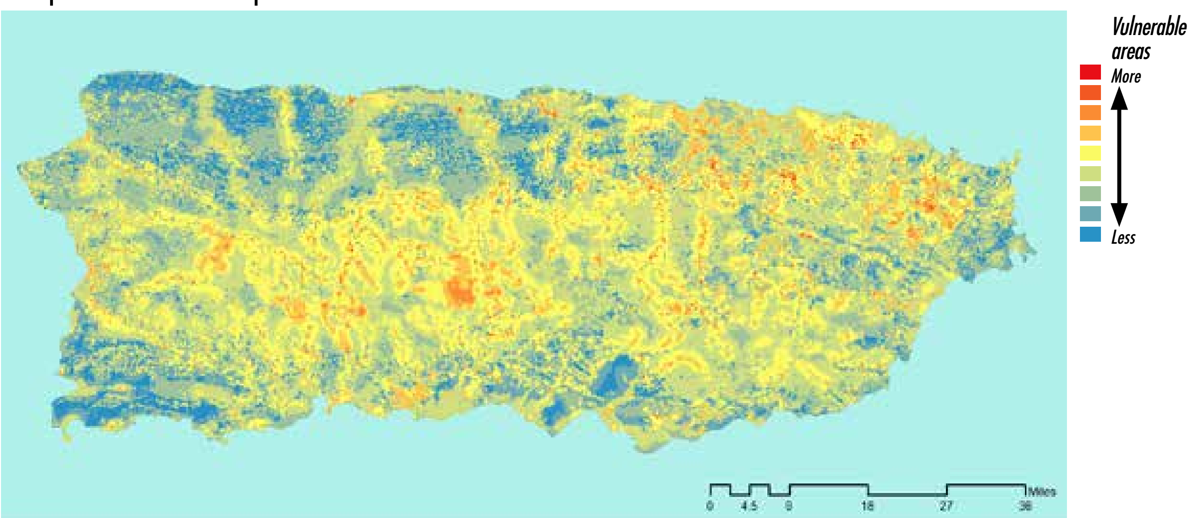


Fig. 3.8. Output from cartographic odeling, using raster overlay, weighted sum tool

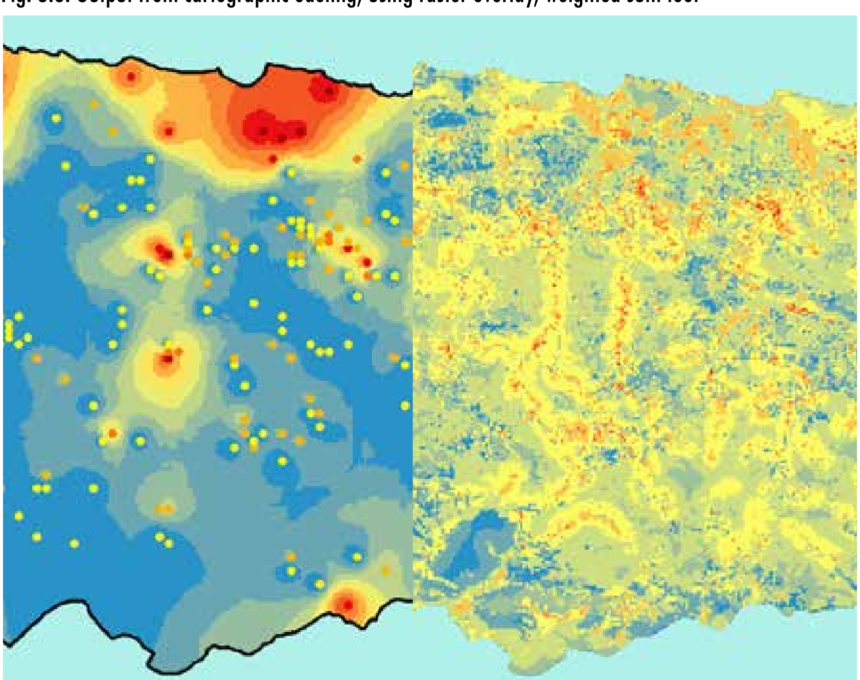
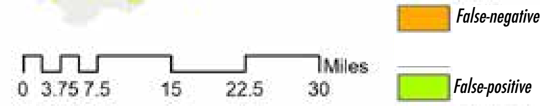


Fig. 3.9. Resolution detail

This maps below illustrate the areas in which predictions were inaccurate (fig. 3.10 - 3.11). Errors, especially false-positive predictions, appear to be clustered in the center of the island and the northeast quadrant, as seen in interpolation methods. The accompanying “confusion matrix” is an adaptation that assess the degree to which the model predicted correctly on known points. Pixel values (counts) provide a measure of magnitude for error.

The cartographic method is the most time-effective and easiest to employ, at a preliminary level, of all four methods. However, this method can be highly nuanced. As

Incorrect	10,065
Correct	105,143



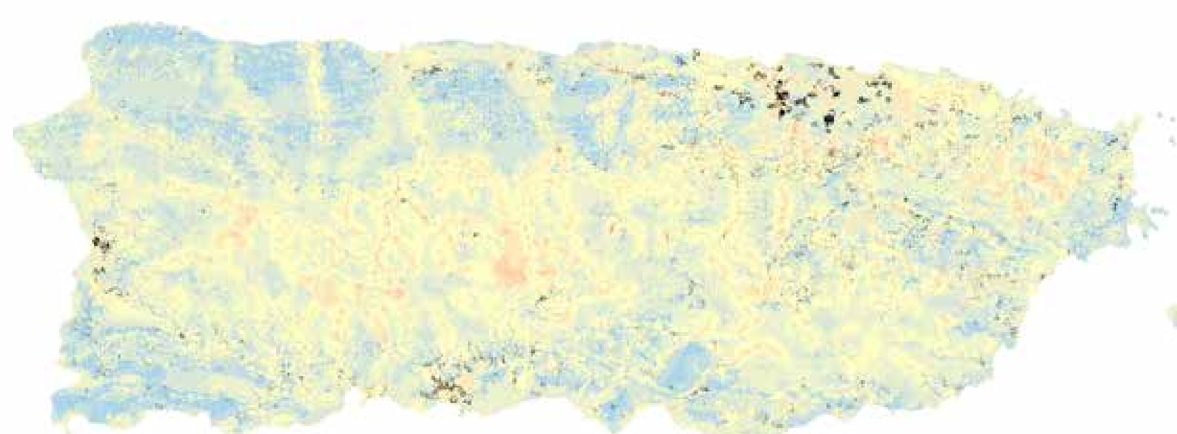


Fig. 3.10 & 3.11 (top to bottom) Incorrect predictions and Actual outages compared to predicted

a result, versatility and functionality are greatly increased with more complex studies. This topic will be addressed again in the conclusion of the Results section. The output's usefulness is entirely dependent upon the sophistication with which input rasters are developed. This method's finer resolution makes it appropriate for different scales, expanding its utility for planning and policy decision-makers. This is also an indication of generalizability. The adapted confusion matrix and map of predicted versus actual failures speaks to a magnitude of error. MAE could be calculated here but is not included in this study. Results might change when input rasters are adjusted to include only those features identified by the regression as significant. A method of this type does not incorporate a method to achieve advanced outcomes, comprehensive model evaluation, or an assessment of predictive power in the same way the machine learning model does.

There are shortfalls not previously listed and next steps for cartographic modeling that are worth noting. One of cartographic modeling's major deficiencies is the subjective quality of the modeling process. An analyst decides decision factors and weights for inputs. Similar to the criticism of interpolation heat maps, cartographic models at once empower users while also exposing the process to high levels of variability based on the analyst. Due to arbitrary inputs, domain or local knowledge limitations, and specific objectives, results can be highly biased and variable. In future, this analysis should be more iterative and complex and should focus on a smaller study area. For example, a cartographic model could be built both before the machine learning model and after. The second iteration could be adjusted based on reported variable significance, and input layers could be carefully engineered to better capture the water system's intricacy. The apparent variability of accuracy across the island makes the case for a smaller scale study. The whole island unit is too broad and contains drastic landscape variation, which inhibits precision.

Machine Learning

The relationship of exogenous factors was first explored using the cartographic modeling method. The spatial reasoning employed during the previous modeling method helped design the geospatial algorithm built in R for this study's machine learning method. The algorithm is a logistic regression and begins with the outcome - a fishnet reference grid of failure (1) and non-failure (0) events. The predictive analysis is "backwards-engineered" from the given outcome. The regression assessed the statistical significance of features related to drinking water service failure and predicted the probability of failure in a future storm event (fig. 3.12).

The machine learning method has the potential to outperform the other methods, according to the basic list of requirements (reference figure 3.18 the end of this section), but this potential is greatly reduced by the lack of data to support the modeling objective. There is great versatility and functionality in this method, but its design and use might be too complicated for the general user, though output usefulness is high due to the versatility and specificity possible with this method.

The requirements of advanced outcome capability and predictive power are both satisfied by this method. Only the machine learning and cartographic modeling methods

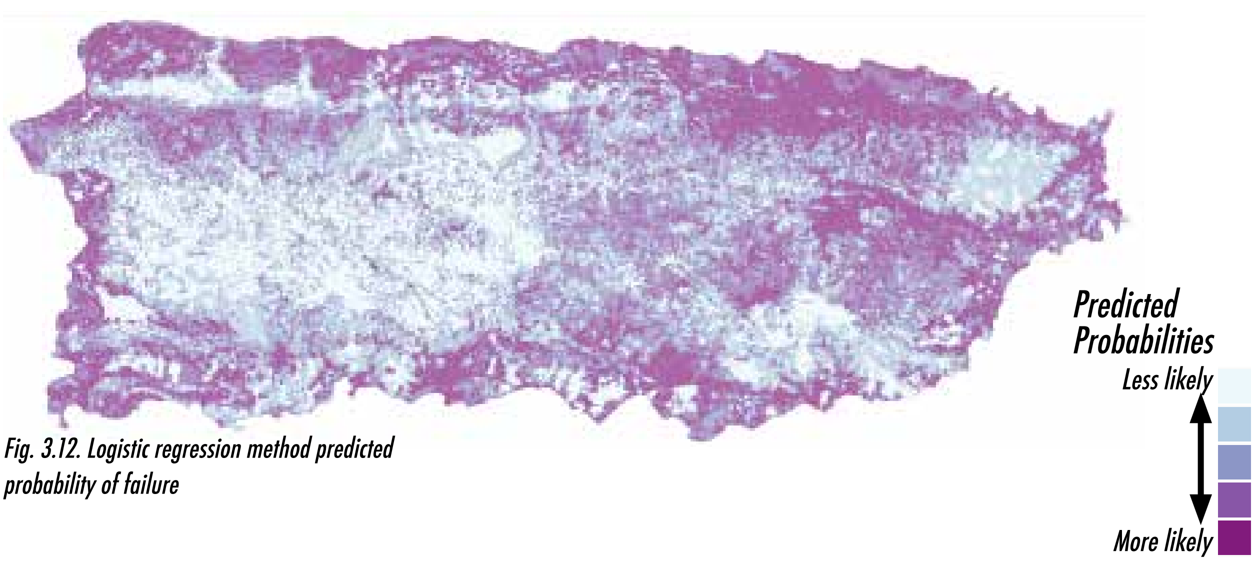


Fig. 3.12. Logistic regression method predicted probability of failure

have the potential for advanced outcomes. As mentioned in the methods section, ten-fold cross-validation was used to iteratively test the model on 100 hold-out test sets to improve its predictive power.

This mathematical method allows for a model evaluation. The first step in model evaluation is to determine whether the selected model, in this case a logistic model, is a good fit for the data. There are many ways to determine fit, but popular assessments usually fall into two categories: predictive power and goodness of fit tests (Allison, 2014). This study evaluates the model employed herein based primarily on accuracy and generalizability. Accuracy is defined as the proportion of correct predictions; while generalizability refers to the model's transferability to multiple places. There is often a trade-off between these two characteristics. A model that is highly accurate in a specific location will likely not generalize well to other locations (Steif, March 1, 2018, Lecture). This study emphasizes the model's predictive accuracy rather than its generalizability,

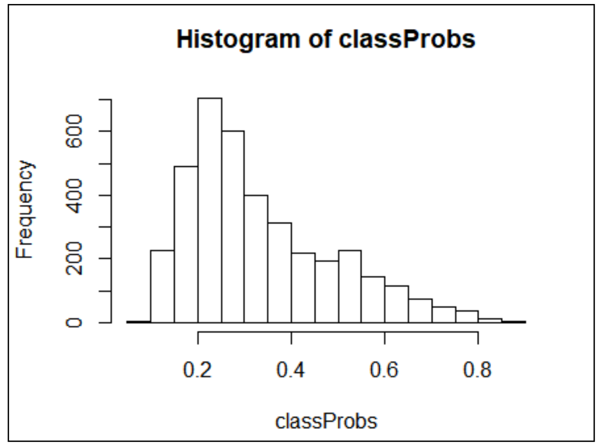


Fig. 3.13. Class probabilities histogram shows the predicted probability of a test set fishnet cell being a failure zone conditional on our model.

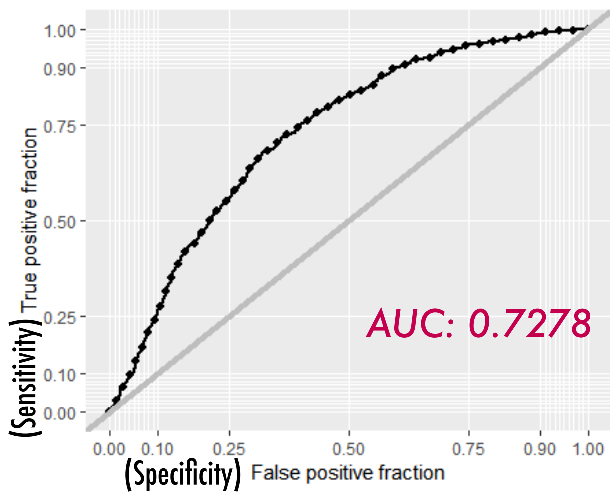


Fig 3.14. Receiver Operating Characteristic (ROC curve)

Confusion Matrix and Statistics

		Reference	
		0	1
Prediction	0	2246	899
	1	284	374

Fig 3.15. Confusion Matrix

due to the drastically different topographies across the island. This method satisfies both the generalizability and predictive power requirements outlined in this section.

The following paragraphs will outline the tests used to evaluate the model. The data indicates which features are statistically significant. Significant factors include: distance to steep slope, distance to hydrology, density of agriculture, density of wetland areas, amount of low-density urban areas, amount of high-density urban areas. Improvements might be made to the model by adjusting or removing features that are not statistically significant.

Figure 3.12 illustrates the ROC curve, which describes the trade-offs between the true positive rate and the true negative rate. As a model is a representation of reality, there are tradeoffs in how reality is represented, such as between accuracy and generalizability. The higher the rate of true predictions, the less generalizable this model would be in other geographic locations. The associated area under the curve (AUC), or concordance sta-

Accuracy : 0.6889
95% CI : (0.6739, 0.7036)
No Information Rate : 0.6653
P-Value [Acc > NIR] : 0.000993

Kappa : 0.2063
Mcnemar's Test P-Value : < 2.2e-16

Sensitivity : 0.29379
Specificity : 0.88775
Pos Pred Value : 0.56839
Neg Pred Value : 0.71415
Prevalence : 0.33474
Detection Rate : 0.09834
Detection Prevalence : 0.17302
Balanced Accuracy : 0.59077

'Positive' Class : 1

Fig 3.16. Confusion Matrix Stats

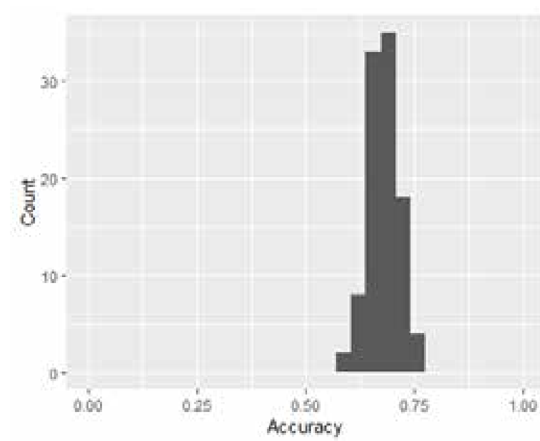


Fig. 3.17. Accuracy Variability Histogram

tistic, describes the classification of variables in the sample of failed or not-failed zones. The AUC (73%) says that this model can predict approximately 27% better than chance (Steif, lecture, February 22, 2018). There is a caveat important to note in relation to this evaluation. The AUC's high value could be due to the large number of correctly predicted non-failure events. However, since the majority of values are non-failure events, the ability to predict non-failure is not exceptional. The ability to predict failure events is the objective, but the AUC value is a coarse test that doesn't capture interesting detail.

The confusion matrix (fig. 3.15) is another diagnostic related to this model's ability to predict between the two outcomes. In the matrix, the "sensitivity", or proportion of actual to predicted failures, is the 374. The true negatives, or ability to predict non-failure or continued operation, is high (2,246). The most concerning component of this metric is the high false positive and false negative values. The errors with the greatest consequence to decision-making would be those that the model predicts as non-failure when, in fact, the actual outcome is a failure or that the model predicts as failure when there is no failure. These two incorrect predictions would cause mis-allocated resources.

Although the machine learning method satisfies many of the requirements, this method is still not advisable at the present time. The major factor limiting this model's predictive power is accurate data. The manual monitoring system precludes real-time data, which is necessary to build an accurate model within the time frame under study. Improvements could be made by adjusting methodologies and features. However, there remains a notable amount of bias and exclusions in the model, and improved data would do much to ameliorate weak model strength.

In addition to the biased predictions for this system using incomplete data, there are other challenges inherent in the machine learning method that makes its use as the primary tool for a vulnerability assessment unadvisable. This method is time consuming and complex. Modeling infrastructure like this may not be the best way to employ a model.

While the development of the system may be built on fundamental and model-able rules, system maintenance incorporates too many variables to model accurately. Additionally, the test question - which areas are most likely to experience water service failure in a storm event? - is difficult to answer accurately given the high level of uncertainty surrounding storm events.

Summary

This paper surveyed four options for forecasting the likelihood that the water supply system will fail, thereby identifying infrastructure areas most vulnerable in the next storm event using spatial autocorrelation and exogenous features. The methods are dissimilar in many ways, but in general, the output from each method was similar. Each method shows comparable spatial patterns representing the most vulnerable areas, often clustered around urban areas. The methods were evaluated on the series of requirements related to usability and output quality previously discussed in this chapter. Their rating on a scale of values between 1- 10 is as follows: Machine learning: 6.56, Kriging: 6.25, Cartographic modeling: 5.75, IDW: 3.57.

These ratings do not accurately reflect the usability and output quality of each method in the context of the trial question. The obstacle of limited data availability is more or less problematic based on the method used. Although a machine learning logistic regression algorithm has the capability to predict vulnerability within the drinking water system, the data inputs required for successful predictions do not exist. Without accurate failure data, a lot of guess-work is necessary to infer failure points or zones, which dramatically increases the margin of error in machine learning. Other methods are more forgiving. Given this limitation, the usability of the cartographic modeling method exceeds the other methods. The adjusted rating scores are as follows: Cartographic modeling: 6.22, Kriging: 6.11, Machine learning: 5.83, IDW: 3.13.

Until a time when real-time data is accessible, it is not feasible to expect analyses to achieve causal explanations, but instead are strictly limited to correlational research. In this preliminary study, there is also no way to 'hold all else equal,' due to the complexities and uncertainty of infrastructure function and actual storm impacts. The most viable method to resolving the trial question would be to create precise raster layers that, when combined, expresses a likelihood of failure. There is a lot of nuance in cartographic mapping that was unexplored in this study, but which could be highly useful for future iterations, such as the incorporation of Kriging results into the cartographic model. Other useful layers would be water demand and water supply. Additional failure indicators should also be suggested by local experts. Combining Cartographic modeling and Kriging and further developing the cartographic model would be the best method for modeling failure and vulnerability.

	Requirement Score (1-10)	Final Score (1-10)
IDW INTERPOLATION	3.57	3.13
KRIGING INTERPOLATION	6.25	6.11
CARTOGRAPHIC MODELING	5.75	6.22
MACHINE LEARNING	6.56	5.83

Fig 3.18. Method Rating

Chapter 4 | DISCUSSION

Discussion

There are numerous ways to evaluate the success of methods applied to the trial question in this paper. While the different methods have a variety of benefits and embedded biases, the given the context and supporting data dictate the use of two methods. The cartographic modeling method is useful because it is fast, empowering, visual, and highly nuanced. The kriging method incorporates some mathematical evaluation and can help build autocorrelation data for input into the cartographic model.

Next Steps

Although the methods rating is not likely to change, there are several aspects of this project design that could be adjusted for improved prediction accuracy and might alter method assessment scores. There are also a wide variety of tangential studies that could evolve from this preliminary study. The following suggestions could be developed as future studies: innovate water distribution monitoring system, improve data collection and open data sharing, study water failure under normal conditions, alter metrics for measuring failure and consumption, alter scale of analysis, regress failure as a function of average failure values, experiment with simulation instead of modeling, and plug into the bigger picture. The following paragraphs elaborate on these suggested next steps.

System innovation

Today, PRASA's distribution system records consumption manually. It takes a full month to complete meter reading, which is why the water utility bills every two months. System innovation is halted due to lack of funding (Diaz-Atienza, personal interview, 2019). In the future, PRASA's monitoring technology could provide near-real-time identification of system components following a disaster event. Once the system is updated and a history of service data is recorded, machine learning models, based on historic failures, supply and demand, and geographic data, will be more accurate, generalizable, and useful for predicting failures. Not only would innovative technology help to increase the amount and quality of available data on water operations, but would increase service transparency. Transparency could assist with public image reparations.

Data

It is clear from the wake of social and infrastructural damage wrought by Hurricane Maria that the prioritization of system innovation and increased real-time data collection could help support vulnerable water service zones. It is also clear from the modeling exercises included in this study that in order to complete an accurate predictive study of water supply failure using machine learning, real-time data is essential. When appropriate data exists and is readily available, modeling water supply failure could be an important use case for extending machine learning. The essential data list is as follows:

- archived, digital 311 call data reporting problems with water or electrical service
- historic failure points saved as a CSV or XML file
- up-to-date building type or use data
- up-to-date census and building occupancy data (after the diaspora many of the building and census data layers available on open data sources are incorrect)
- population counts (not estimates) at the block or household level
- automated, real-time monitoring data
- accurate facility capacity data as a shapefile or geojson
- accurate service zones data to illustrate redundancy or lack-there-of

- pressure in/pressure out as a shapefile or geojson (in order to accurately estimate consumption volumes, leakages, and tank replenishing)

Normal Conditions

It is likely that all PRASA infrastructure is vulnerable to failure in category 4 or higher hurricanes. Given the variability of output from the test question, I could be assumed that no area within the island is more or less vulnerable than another in storm events. A study that could be more useful is to predict the likelihood of drinking water service failure under normal conditions. Data requirements for this suggested study would be similar to those required for the methods herein. Additionally, archival failure information could be useful to track functional failure under normal conditions across time. Using a longer time horizon, more reliable inferences could be made about cause and effect.

Failure and Reliability Metrics

Failure is a highly subjective concept and quantifying the likelihood of system failure is logistically difficult (Xiao, 2016). There are conflicting theories about how to best represent the significance of failures. Some researchers and engineers prefer estimating reliability as an alternative to estimating failure. Altering the method of quantifying failure would yield a dramatically different study.

A few details pertaining to both failure and reliability are outlined here. The PRASA system includes many pumping stations, treatment plants, tanks, intakes, and miles of pipe. Thus, there are many ways PRASA's water distribution system could fail. In reality, failure results are called “shortfalls” and may be infrequent leakages, partial failures, or total capacity loss (Shamir, 1981). The shortfalls considered in this study were those that could cause total capacity loss, and are not a highly accurate representation of actual function.

Failure can also be measured by the magnitude or scale of failure, such as the total number of failures in a given time period, the duration of periods without failures, and the magnitude or duration of the worst failures (Shamir, et. al., 1981). Failure can be represented by different units, such as units of time or percentage of demand not met (Shamir, 1981, 1987; Hawk, 2003; Tsakiris, 2007; OCW). Generally, the concept of the water utility's reliability is the ability to satisfy the demand of customers in a given area with uninterrupted service over time (Hawk, 2003). There is clearly a relationship between failure and reliability, but the general consensus is that reliability is a more precise measure of system performance. Accurately quantifying failure should be a priority for future studies.

Supply & Demand

Another challenge to be addressed in future studies is how to quantify water consumption. In the industry, values are set by the utility company and based on real data. Consumption fluctuates throughout a given timeframe, but the maximum daily consumption in one distribution area would be determined from the water company's demand record (OCW). Consumption "is often monitored at supply points, where measurements include leakage, as well as the quantities used to refill the balancing tanks that may exist in the system" (Trifunovic, 2006). Although PRASA's records are not publicly accessible, consumption could be approximated from estimated population values and attributed to 2017 population block groups multiplied by an average daily demand value of 100 gallons per person (USGS). 100 gallons of water per person/ day translates to a flow capacity of .06944 GPM and serves as a measure of demand.

Scale

Scales can drastically influence results. Smaller fishnet grid cells could be used for future studies and big data may be augmented with local knowledge. Another method related to scale could be to collect data at a finer resolution and train a model on a selected urban or ex-urban area. Then, the model could be tested on places in Puerto Rico exhibiting similar building topographies, such as San Juan, Ponce, or Aguadilla.

Future use of Spatial Autocorrelation

The methods evaluation herein suggests that the best current method is a combination of cartographic mapping and kriging. For future studies, spatial autocorrelation and exogenous physical, environmental factors could be combined in the logistic regression. The model would then regress failure as a function of average distance to failure. Including spatial autocorrelation into a model to predict failure will improve accuracy.

Simulation

Simulation is “a tool to evaluate the performance of a system, existing or proposed, under different configurations of interest and over long periods of real time” (Maria, 1997), and could be an appropriate alternative for the water service distribution system. This bottom-up method would illuminate emergent phenomena related to failure. Although the water infrastructure system is complex, it may be built upon simple rules. For example, the reason for facility locations and allocation of resources could be a result of economic decisions, and there are many theoretical rules which dictate economic decision-making and development. An algorithm is limited in its ability to identify zones of storm vulnerability due to the complexity and randomness of the system. “Whenever there is a need to model and analyze randomness in a system, simulation is the tool of choice” (Maria, 1997). The element of randomness potentially influences failure in storm events and makes failure simulation a potential use-case.

Bigger Picture

The evaluation of modeling methods is one small piece of a much greater purpose. Three tangential topics related to subject matter broached in this study are: the incongruity of “our right to water” and improper valuation of the resource, repercussions of water distrust, and appropriate steps to preparedness planning.

Water utilities and the public are both guilty of under-valuing water. There is a complex relationship between the public claim to water rights and the willingness to pay for it, but the complexity is exacerbated in Puerto Rico due to a history of water injustice (Black and Veatch, 2018; Melendez, personal interview, September, 1, 2018; Neukrug, personal communication, October, 24, 2018; Rodriguez, personal interview, January 7, 2019). “Puerto Ricans have a unique relationship with water [service]” (Melendez, personal interview, September, 1, 2018). This singular relationship may contribute to how residents conceptualize and value this commodity. Consistent water supply is an expected level of service in US cities, but, utilities pander to public demand for cheap water, thereby jeopardizing their ability to repair or improve systems in crisis (Neukrug, personal communication, October, 24, 2018; Black and Veatch, 2018). PRASA’s stated mission is to deliver the cheapest water possible, which doesn’t sustain the function of a failing system. Perhaps an adjustment of PRASA’s stated mission from a level of service at “the lowest possible cost” to a level of service that has value significant value may be appropriate as well is in order? Although, there is little literature to support this contention, equitable rate increase and involvement of public-private partnerships could be two avenues for supplementing the funding necessary to make updates to the water system.

The second topic that stands out as significant and under-studied is the distrust of water supply safety and reliability. Equitable access to safe drinking water has been a persistent challenge for Puerto Rico for more than three decades, earning the island the undesirable distinction of having the worst drinking water in the nation. Phys.org, a science, research and technology news website, reported four months before the storm that outdated and deteriorated infrastructure, poor implementation of the Safe Drinking Water Act, poor enforcement by the territory, and under-reporting of violations are to blame for the territory’s qualification of worst in the nation (Phys.org, 2017). Then, the storms hit an already vulnerable water distribution system. Not only did the storms transform the built and natural landscape in one fell swoop, they also further entrenched citizens' mistrust of both the quality and reliability of the drinking water distribution system (Melendez, personal interview, September, 1, 2018). An August 8th, 2018 article published in the Guardian, reported that "even by [the summer of 2018], 53 percent [of Puerto Ricans] say they are worried about the quality of water in their homes" (Milman, 2018). The social implication of diffuse distrust of water service safety and reliability may be one of the most challenging problems to repair and should be prioritized among PRASA's missions and future research efforts.

Innovative and integrated preparedness planning would be useful for Puerto Rico. A related study could be to outline appropriate steps for a PRASA-wide preparedness plan. Whether under threat of storm damage or functional failures, Howard Neukrug, former director of the Philadelphia Water Department, suggests that the best way for a water utility to prepare for emergency is to observe the system, understand points of sensitivity, address the problems, and solve the problem (personal interview, October 23, 2018). In other words, preparedness should be focused on the least supported areas. This sentiment is similarly reflected by the increasing demand for vulnerability maps among development agencies and governments, as greater emphasis is placed on scientifically-sound methods for targeting adaptation assistance (de Sherbinin, 2014). According to Eli Diaz-Atienza, PRASA recognizes the increasing cost and impact on people's lives from disasters. This impact continues to grow (personal interview, February 1, 2019).

Chapter 5 | CONCLUSION

Conclusion

A resounding message culled from articles, reports, and interviews is that Puerto Rico is poised for systemic change as the territory faces re-build challenges. The reconstruction process is estimated to take 10 years (USDA, Forest Service, 2017). There are compounded effects and complicated problems of supplying water security to all people in Puerto Rico and some researchers say that it may not be possible to correct the problem (Green, 2018). However, researchers are not all together pessimistic. In fact, many write about predictions and prescriptions to help generate sustainable solutions for current systems (Rosenfeld, 2018). There is also positive interest and gathering momentum in the "re-build movement," evidenced by the increased number of institutional collaborations, conferences, research, and publications in support of Puerto Rico's re-build efforts. For example, symposia like Rutgers' "A Call to Action" in March, 2018 and "A Year After Maria" in October, 2018 drew large, international crowds at both sessions (Community Leadership Center). Additionally, the Center for Puerto Rican Studies, at Hunter College, City University of New York, recently launched an online information clearing house (Center for Puerto Rican Studies). Similarly, HydroShare's Puerto Coverage, supported by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI, 2018) collects and manages Puerto Rico-centric resources to support research.

While the stakes of the game are changing, opportunity can be found in change to mitigate risk through innovative planning and policy (Meyer, David and Kunreuther, Howard, 2018; Santiago, 2018). The 2017 incidents could be an incentive for sustainable growth and development. Many planners and policy-makers have identified this opportunity and are actively negotiating the best way forward, but according to the USDA Forest Service, the effectiveness of future conservation, mitigation, and adaptation action in Puerto Rico are dependent upon post-storm information. Additionally, and especially given climate change and future uncertainty, vulnerable places, like Puerto Rico, may be required to adjust prior expectations for clean drinking water access. In areas vulnerable to increasing environmental disturbances, information is “critical to the design and implementation of on-going recovery work and to longer-term resilience efforts” (USDA, Forest Service, 2017).

It is a matter of time before another disaster hits Puerto Rico, precipitating other destruction, recovery aid, and re-building. Re-build strategies and cultural shifts occur on different timelines, but both immediate needs and long-term change require a baseline understanding of the island’s existing conditions and the best methods for analyzing data. Innovations and planning efforts should be guided by data. However, at this time, research is limited to correlational studies. To the extent possible, interpolation methods should be evaluated for accuracy and cartographic modeling should be framed to support decision-making. Due to the metrics available to assess accuracy and generalizability, machine learning will eventually be at the top of the analyst’s tool kit, but not until the requisite data is available.