

Using Nationwide and Local Data Sources to Address Decentralized Wastewater Infrastructure Challenges in the Contiguous U.S. and the Caribbean

Tuesday, September 21, 2021, 1:00-2:30 p.m. ET

Decentralized Wastewater Webinar Series

Using Nationwide and Local Data Sources to Address Decentralized Wastewater Infrastructure Challenges in the Contiguous U.S. and the Caribbean

Tuesday, September 21, 2021, 1:00-2:30 p.m. ET

EPA's Decentralized Wastewater MOU Partnership





Puerto Rico Case Study

Paul Fericelli,

Water Enforcement Division, U.S. Environmental Protection Agency



Climate change poses a number of difficult challenges to septic system infrastructure. Rising sea levels, increased precipitation, and warmer temperatures are all adversely impacting these systems and causing septic system failure.

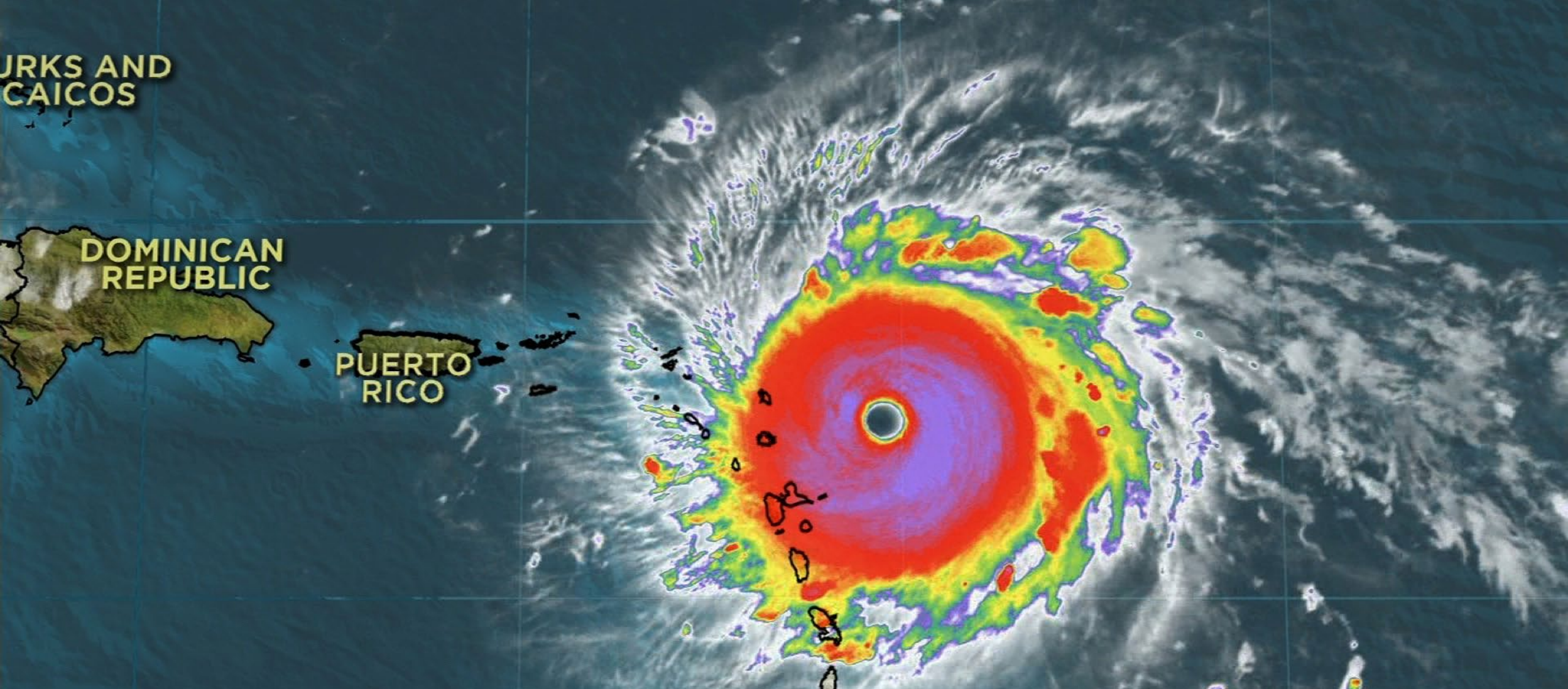
YOU CAN'T FIX WHAT YOU CAN'T SEE



**JRKS AND
CAICOS**

**DOMINICAN
REPUBLIC**

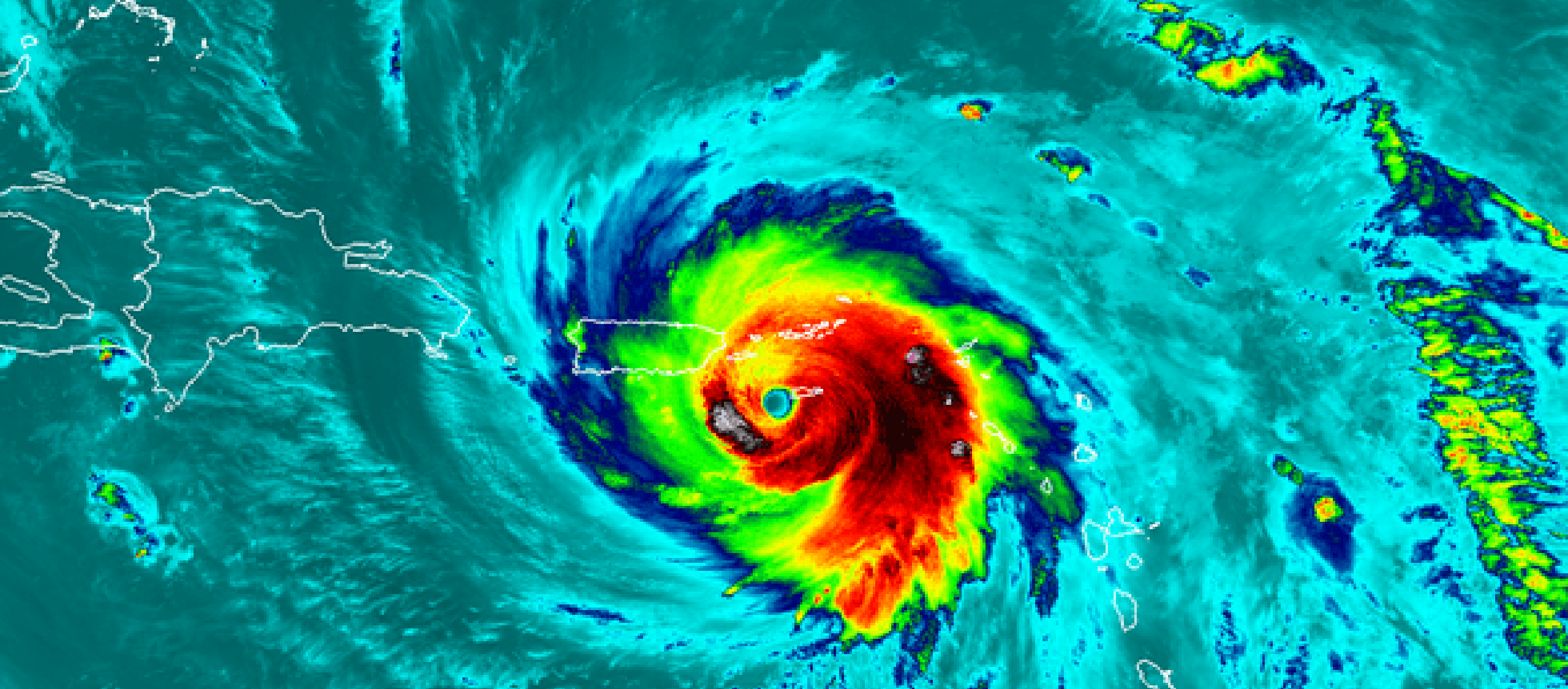
**PUERTO
RICO**



HURRICANE IRMA CATEGORY V | September 5-7, 2017

“Hurricane Irma, one of the most powerful Atlantic storms in a century, tore its way across the U.S. Virgin Islands as a Category 5 storm on Sept. 6, mixing together a deadly combination of winds and rising water.” (USA Today, 09/13/17)

FOR INFORMATIONAL PURPOSES ONLY



HURRICANE MARIA CATEGORY IV | September 19-21, 2017

“Hurricane Maria became the strongest storm to make landfall in Puerto Rico in 85 years when it came ashore” (CNN)

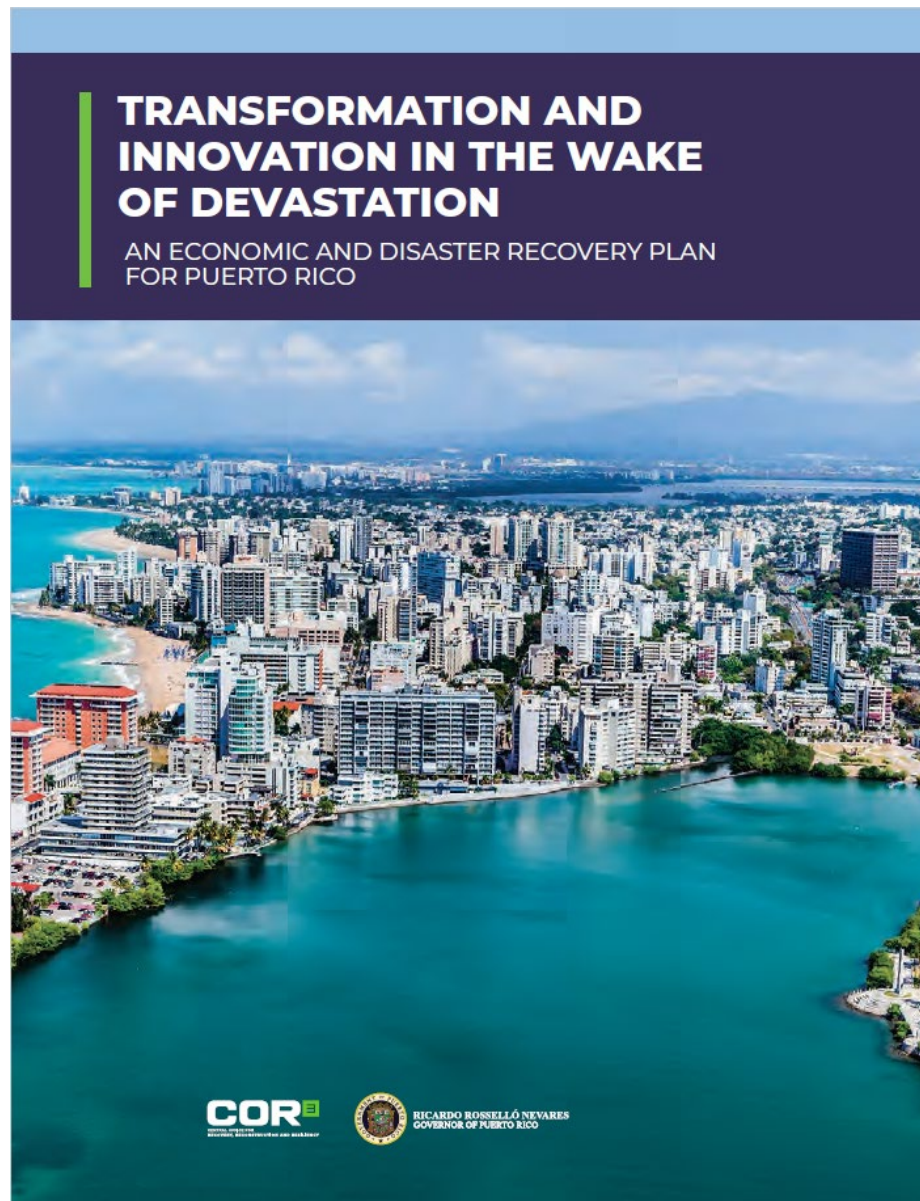
FOR INFORMATIONAL PURPOSES ONLY



After Hurricane Maria

Predisaster Conditions, Hurricane Damage, and Recovery Needs in Puerto Rico

JORDAN R. FISCHBACH, LINNEA WARREN MAY, KATIE WHIPKEY, SHOSHANA R. SHELTON, CHRISTINE ANNE VAUGHAN, DEVIN TIERNEY, KRISTIN J. LEUSCHNER, LISA S. MEREDITH, HILARY J. PETERSON, HSOAC PUERTO RICO RECOVERY TEAM



TRANSFORMATION AND INNOVATION IN THE WAKE OF DEVASTATION

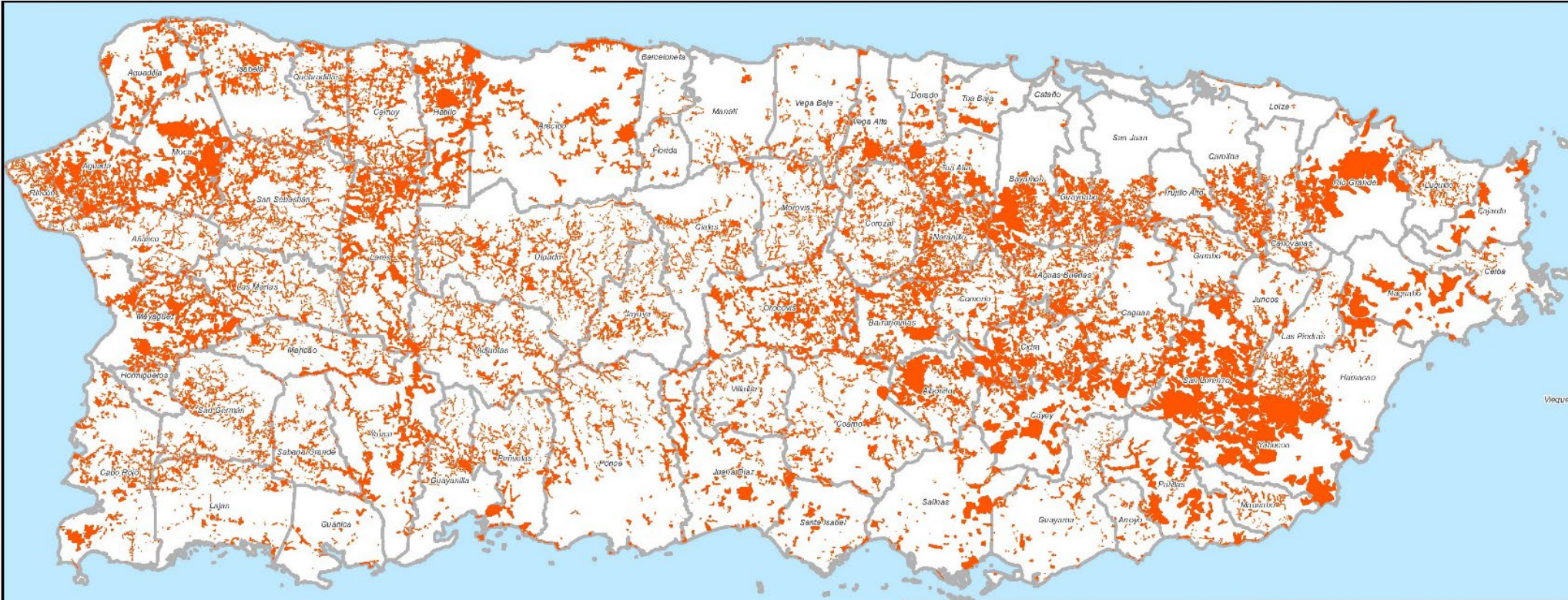
AN ECONOMIC AND DISASTER RECOVERY PLAN FOR PUERTO RICO

Supporting documents for decision making about disaster recovery efforts.

The Starting Point

Existing Baseline Information About Septic Systems in PR Areas in PR with No Connection to Centralized Sewer Systems

Incident Complex - Puerto Rico



About 1.4 million people, or approximately 40% of Puerto Rico's population, have no sewer connection to centralized sewer systems.

Estimated Inventory of Septic Systems:

- Residential (single family) = 550,000 a 600,000
- Commercial = 20,600
- Government = 3,800
- Industry = 230

Source: F. Quiñones, 2011



N

■ Areas with No Sewer Connections
□ Municipalities

DISASTER DECLARATION PROCESS

#1 Incident occurs

#2 Local officials collect initial damage estimates

#3 State requests Joint Preliminary Damage Assessments (PDAs) from FEMA Regional Office

#4 Joint local/state/federal PDAs are conducted

#5 Governor submits disaster declaration request to president through FEMA Regional Office

#6 FEMA reviews request, sends recommendation to president for decision

#7 President makes determination

- The PDAs determine the need for an Individual Assistance declaration, if the damages and thresholds are met. Ultimately the White House makes a Presidential determination on the declaration.
- Through the **Individual Assistance (IA) programs**, FEMA provides direct assistance to individuals and households, as well as state, territorial, tribal, or local governments to support individual survivors.
- After individual survivors **register** for assistance, FEMA is required to verify losses to determine the survivor's eligibility for disaster assistance. Assistance only provides the basic needs for a home to be habitable, functional, safe and **sanitary**.
- FEMA's standard method to verify a loss due to disaster damage and to determine initial eligibility is **an onsite inspection by a FEMA inspector**. Home damage must be disaster-related.

Understanding FEMA Verified Loss (FVL)

- FEMA verified loss determination may vary depending upon an individual's situation.
- The FVL is the total dollar amount for eligible real or personal property items of average quality, size, and capacity.

Item ID	Category	Line Item
5142	2002	Septic tank, replace
5143	2002	Septic tank lid, replace
5144	2002	Septic tank, pump
5145	2002	Septic drain field, replace
5251	2020	Septic System, Mounded
5139	2002	Sewage Ejector Pump, Replace
5740	2009	Washout, fill
5100	2002	Service Call - Utilities
5140	2002	Sewer line, auger
5141	2002	Sewer line, replace



FEMA Assistance to Septic Systems

June 13, 2018
DR-4339-PR FS 038
FEMA News Desk: (866) 366-8807
fema-pr4339prensa@fema.dhs.gov

Fact Sheet

Damaged Septic Tanks and Wells May be Eligible for FEMA Assistance

- Puerto Rico residents whose septic systems or wells were damaged by Hurricane María could be considered for financial assistance under FEMA’s Individuals and Households Program.
- Funding is available to pay for cleaning out septic tanks and repairing or replacing the entire system, as needed. Damaged private wells that are the sole source of water for your home could also be considered for federal funding.
- If you have not yet registered with FEMA, you can register online with any computer, smartphone or tablet. Visit www.DisasterAssistance.gov. Or call the FEMA Helpline at **800-621-FEMA (3362)**. TTY users can call **800-462-7585**. The toll-free lines are open from 7 a.m. to 11 p.m. local time seven days a week.
- If you have applied for FEMA assistance and have not had a home inspection, you should let the FEMA inspector know you have a private well and/or septic system.
- If you have registered and your house has already been inspected, you can call the FEMA Helpline and ask for another inspection.
- Home-repair grants are available to qualified applicants to restore your home to a safe, sanitary and functional condition. You must own your home and it must be your primary residence. Other types of grants are available for both homeowners and renters.
- Grants are not intended to completely restore dwellings to pre-disaster conditions and cannot be used for cosmetic work or repairs already covered by insurance.
- The deadline to register is for FEMA assistance is June 18.

- Develop fact sheets, public service announcements and a communication strategy to share information with survivors.

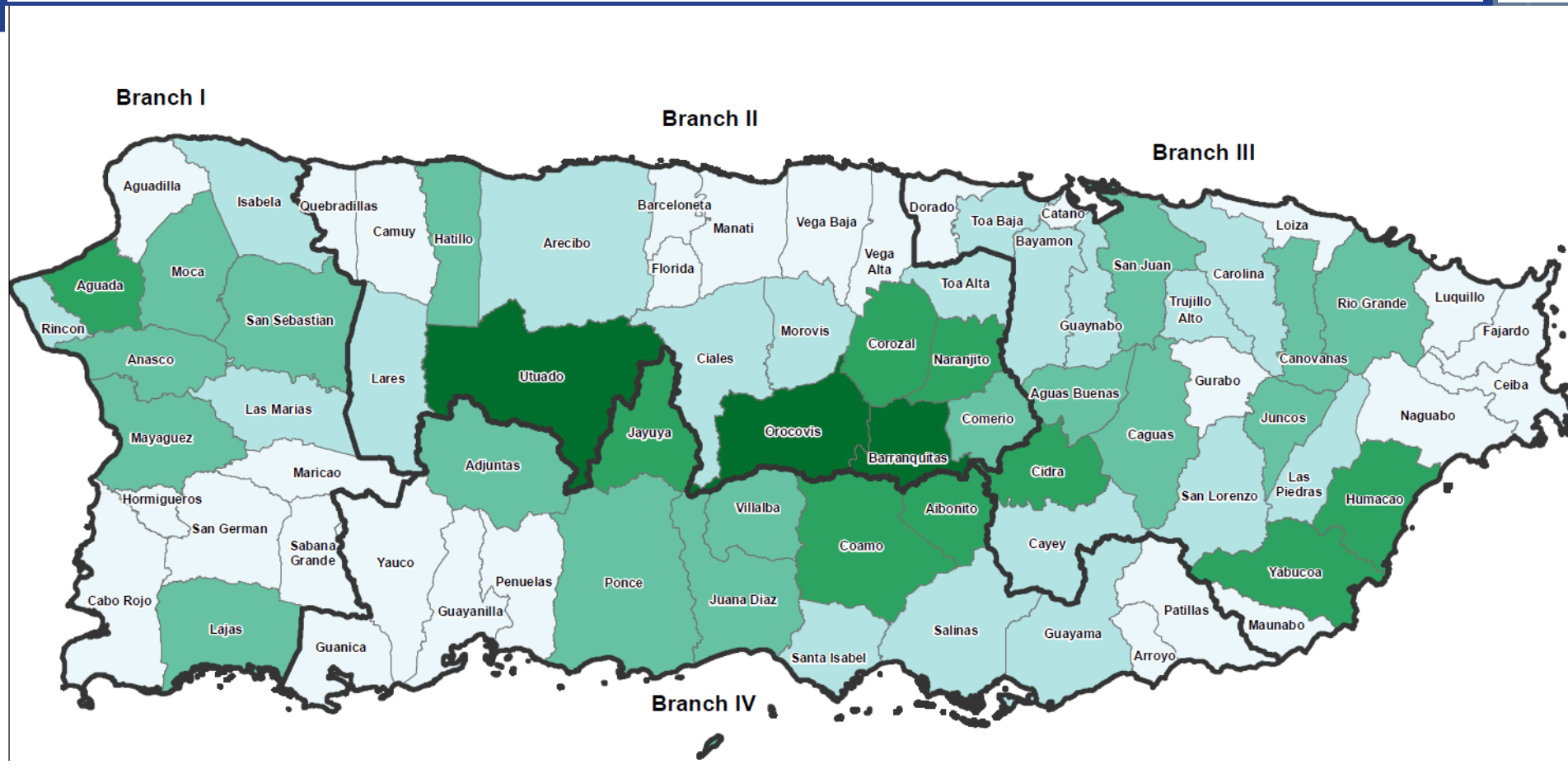


Outreach at radio stations



Outreach at shopping malls

FEMA Verified Loss of Septic System Damages by Municipio

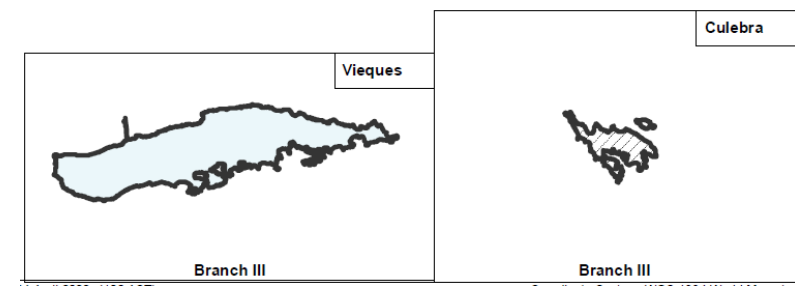


FEMA Verified Loss of Septic System Damages totalized approximately \$1.9 million.

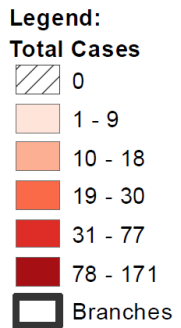
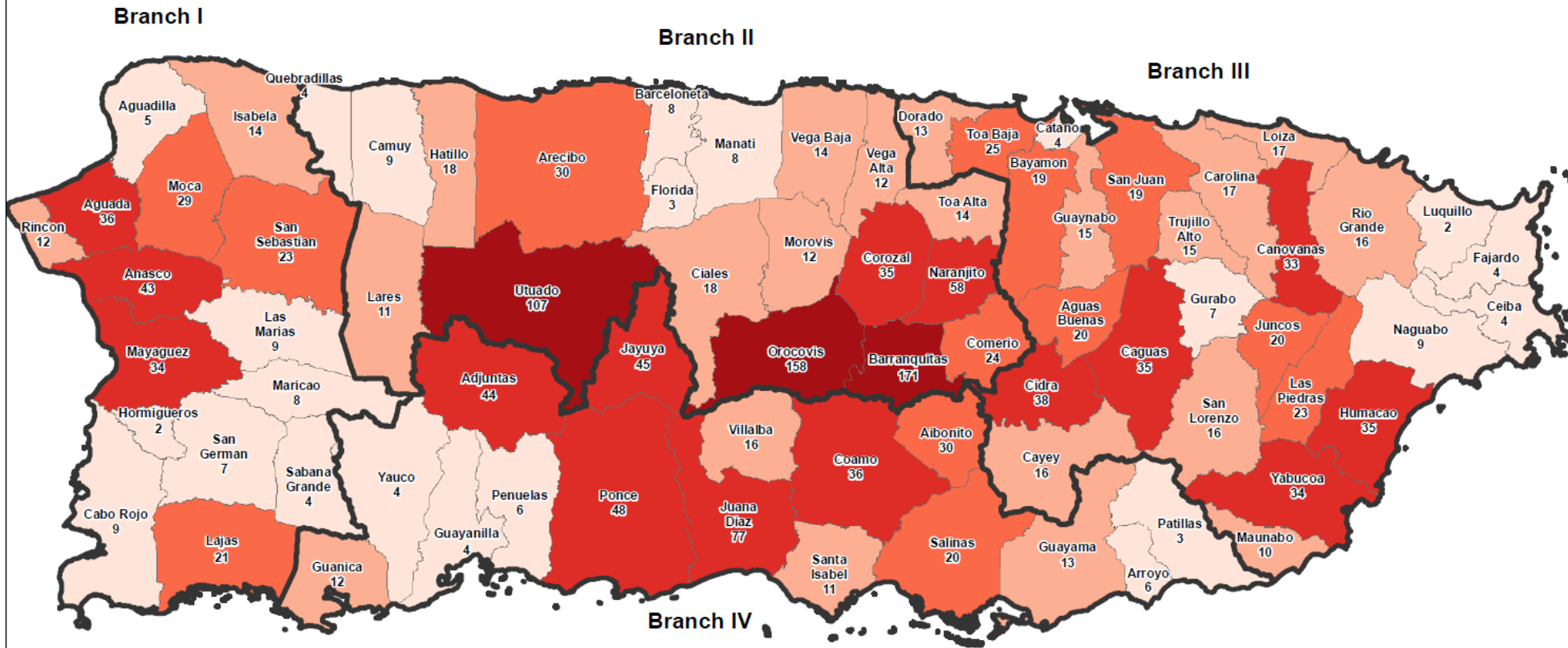
Source: FEMA Individual Assistance, 2020

Legend
FEMA Verified Loss of Septic System Damages

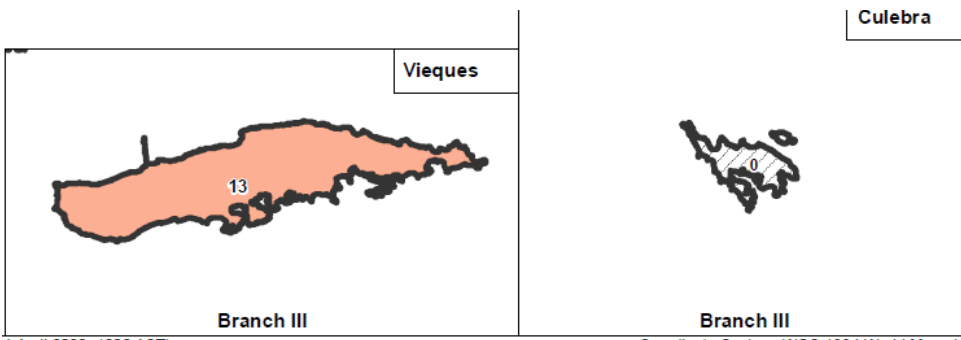
- \$464.34 - \$9,518.12
- \$9,518.13 - \$20,826.16
- \$20,826.17 - \$38,232.32
- \$38,232.33 - \$84,942.94
- \$84,942.95 - \$175,453.02
- No Damage Report
- Branches



Cases with Damage Claims Related to Septic Systems by Municipio



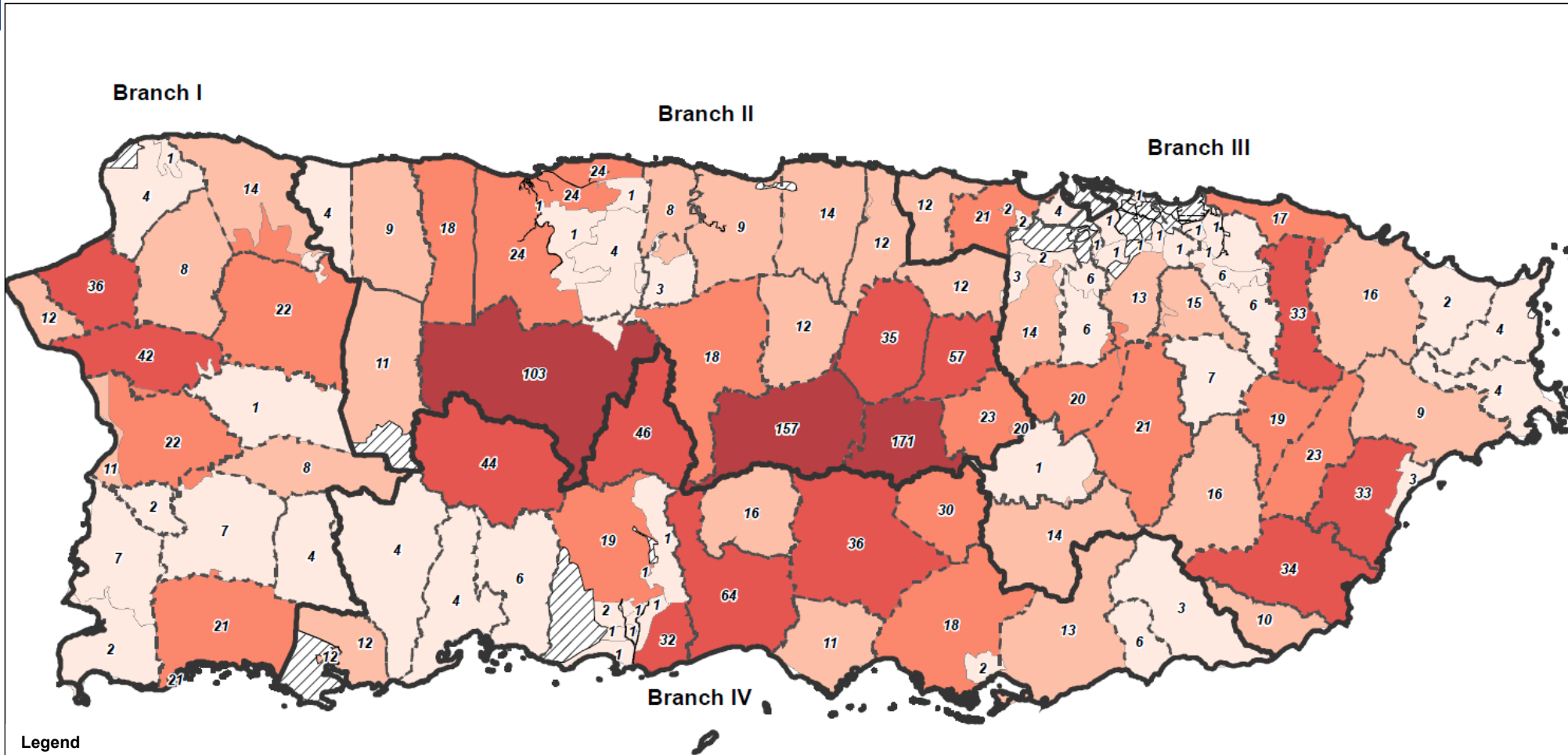
Data Layer / Map Description
 This map illustrates the total cases with damage claims related to septic systems by Municipio.



FEMA received a total of 1,824 requests for individual assistance with damage claims related to septic systems.

Source: FEMA Individual Assistance, 2020

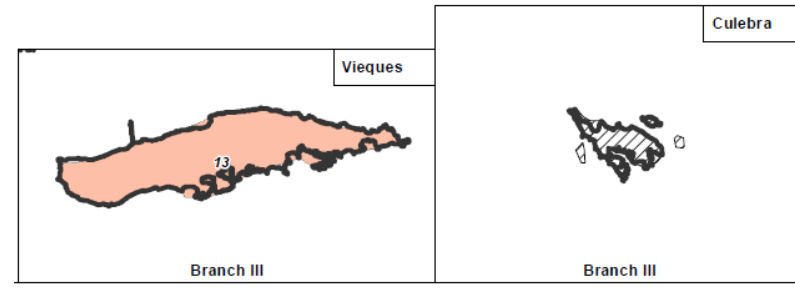
Cases with Damage Claims Related to Septic Systems by Zip Code



Legend
Total Cases by Zip Code

- 0
- 1 - 7
- 8 - 16
- 17 - 30
- 31 - 64
- 65 - 171
- Municipios
- Branches

Data Layer / Map Description
This map illustrates the total cases with damage claims related to septic systems by zip code.

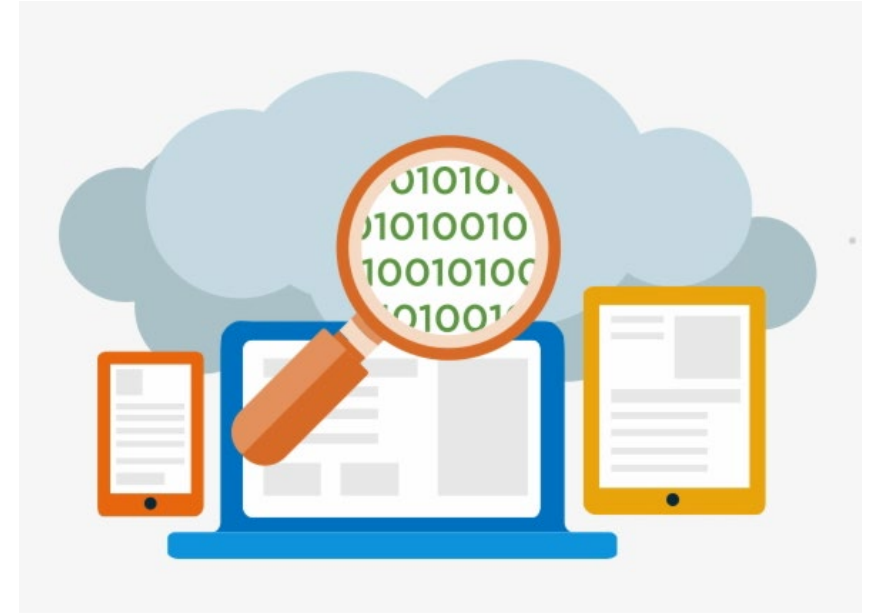


Source: FEMA Individual Assistance, 2020

Determining the Universe of Data Sources Available from Puerto Rico Agencies, Municipalities and Stakeholders that can Support the Development of a Geospatial Tool

Scope of the project:

- Data identification
- Data collection, when available
- Data characterization
- Data inventory
- Develop recommendations to facilitate subsequent analyses using geospatial tools to support decision making about interventions needed to reduce nutrient pollution



Data Characterization

Evaluation Factors Used in Evaluating Existing Data Sources in Puerto Rico

Factor	Criteria
Applicability and utility	Is a document returned in the literature search relevant to the project?
Soundness	Are scientific and technical procedures sound?
Clarity and completeness	Are data clearly and completely documented?
Evaluation and review	Has the information been reviewed?
Uncertainty and variability	Has variability and uncertainty been characterized?



Key Findings from the Analysis of Data Sources in Puerto Rico

- **At least 80 data sources** were identified through meetings with federal and local agencies, municipalities, academia, and private sector practitioners.
- **66 datasets** were characterized determining their quality and usefulness relevant to nutrient reduction analyses.
 - **38 of 66 datasets** are highly relevant for subsequent analyses.
- **The PR Department of Natural and Environmental Resources (local environmental department), the PR Department of Education and the Office of Permits Management** have data with geospatial references for septic systems.
- **Municipal Separate Stormwater Sewer Systems (MS4)** permittees keep logs of possible sources of pollution entering the stormwater system, such logs have geospatial references for septic systems.



Possible Funding Partners to Support Planning and Repair Efforts

HUD

- State Community Development Block Grant
- Community Development Block Grants – Disaster Recovery
- Community Development Block Grants – Mitigation

EPA

- Clean Water State Revolving Fund
- State 319 Grant

FEMA

- Building Resilient Infrastructure and Communities (BRIC)
- Hazard Mitigation Grant Program

USDA

- Housing Preservation Grants



Acknowledgements



CARIBBEAN SEPTIC SYSTEMS WORKGROUP

- **Federal partners:** FEMA, EPA, USDA, USGS, USFWS and HUD
- **Local Partners:** PR Department of Health, PR Department of Natural and Environmental Resources, PR Municipalities, PR Department of Housing, Municipal Revenue Collection Center (CRIM), Office of Permits Management (OGPe)
- **Academia:** Polytechnic University
- **Stakeholders:** San Juan Bay Estuary, Eng. Carl Soderberg, PR Professional College of Engineers and Land Surveyors (CIAPR), RCAP Solutions



septicSMART[™]

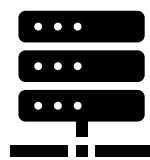
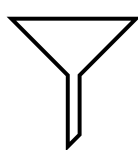
U.S. Environmental Protection Agency



THANK YOU!



Paul Fericelli
fericelli.paul@epa.gov



Developing methods to predict buried sanitation infrastructure across the United States

Nelson da Luz, Emily Kumpel PhD, Jay Taneja PhD

University of Massachusetts Amherst

Contact: Nelson da Luz, ndaluz@umass.edu

Background

- Safe management of human waste is crucial for ensuring human and environmental health
- While on-site wastewater treatment systems (OWTS) are likely used by at least 20% of the US's population, the last census of the prevalence of sanitation systems in the US was in 1990
- As a result, there is a significant gap in our understanding of the number, locations, and density of OWTS across the country
- Asset management techniques have changed over time

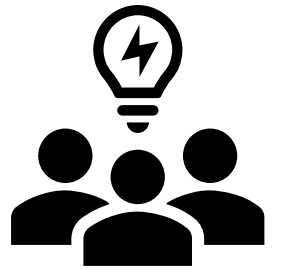


<https://www.plconcrete.net/the-benefits-of-precast-concrete-septic-tanks>

The Big Picture

Knowledge of where septic systems are located on a national scale can allow further research to:

- Identify regions at risk of public health or water quality impairment due to OWTS failure.
- Evaluate feasibility of sanitation infrastructure consolidation
- Understand health and environmental impacts of sanitation technologies
- Target locations for deployment of sustainable OWTS technologies
- Assess equity of access to sanitation infrastructure



How can we find out where septic systems are?

- Geospatial septic system data exists in some locales
 - e.g. Florida, Rhode Island, Delaware, Hawaii
 - Other towns and counties across the country
- Sewer system data exists too (e.g. Massachusetts)
 - assumed to be the inverse of septic system data
- We want to use other indicators to fill in the gaps
- The key to our approach is to fill in the gaps between data that are already available
- **Goal:** National Inventory of Buried Sanitation Infrastructure

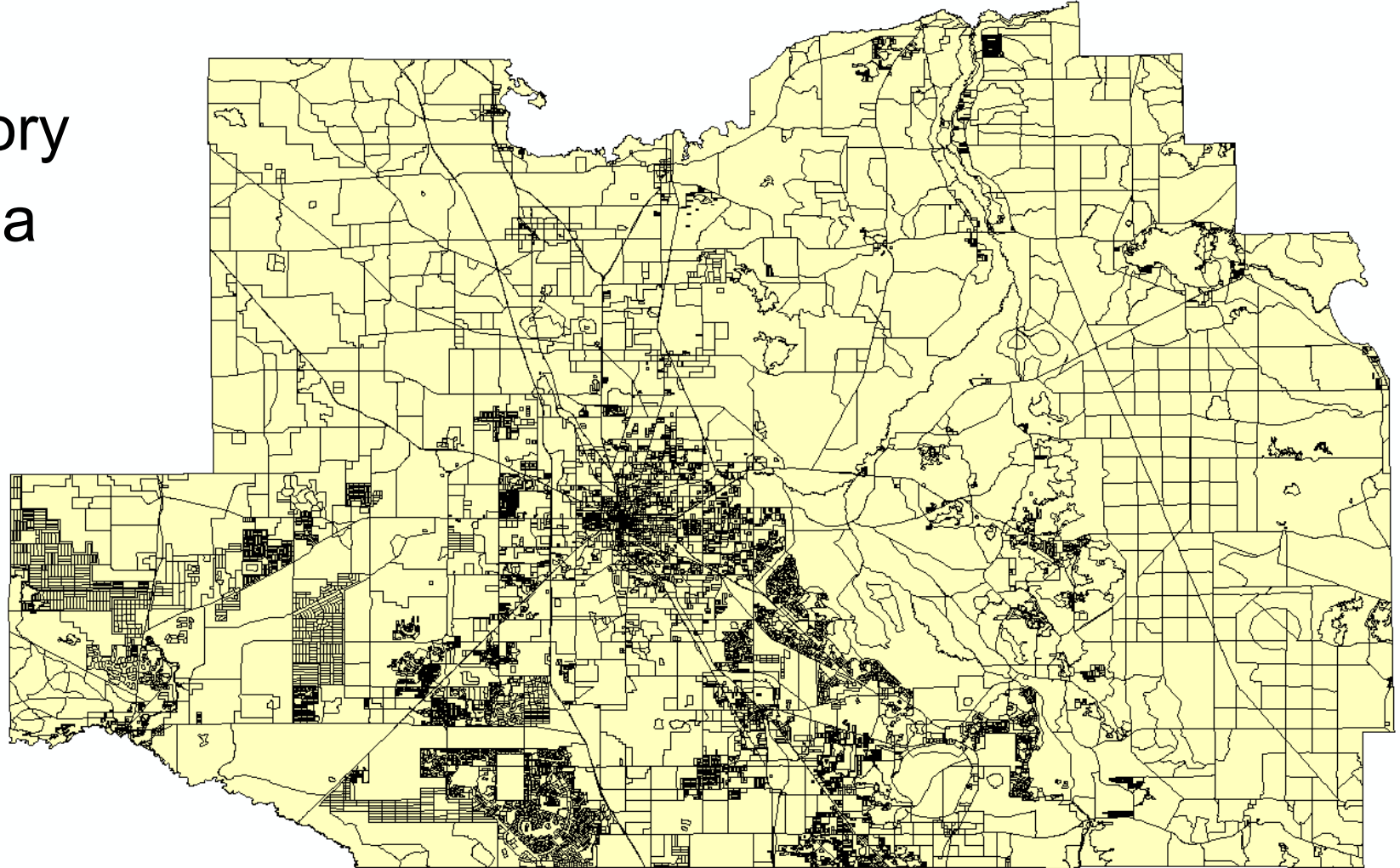
Proof of Concept: Is this idea viable?

- Test methods in a place with actual ground truth data
 - Will let us know how well methods perform
- Once confident with how we did:
 - Apply method other places (doing some spot checks with town or county-level data)
- Then: See how method might work where we don't have consistent data available

Proof of Concept: Marion County, Florida

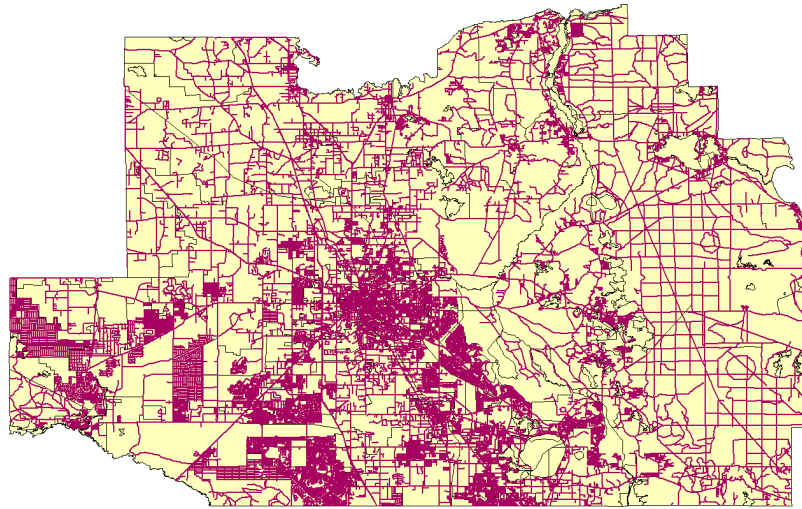
- Florida Water Management Inventory
- Contains city of Ocala
- Mix of Septic and Sewer

Goal: Predict whether an individual parcel has a sewer connection, a septic system, or neither



Data Sources

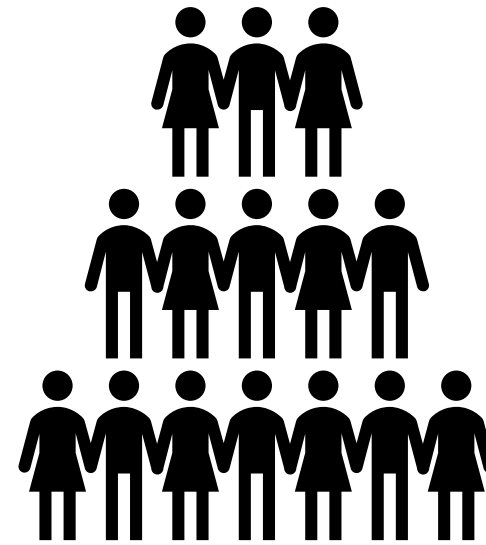
260+ Variables



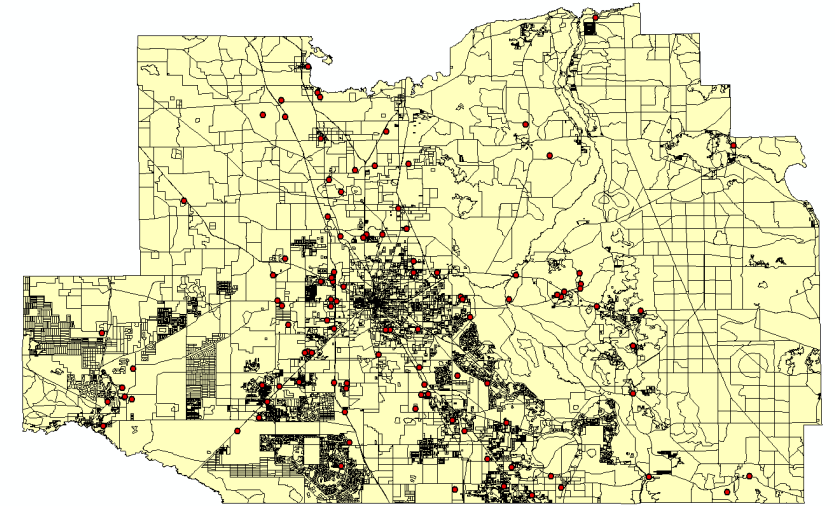
#, Length, Type of Roads



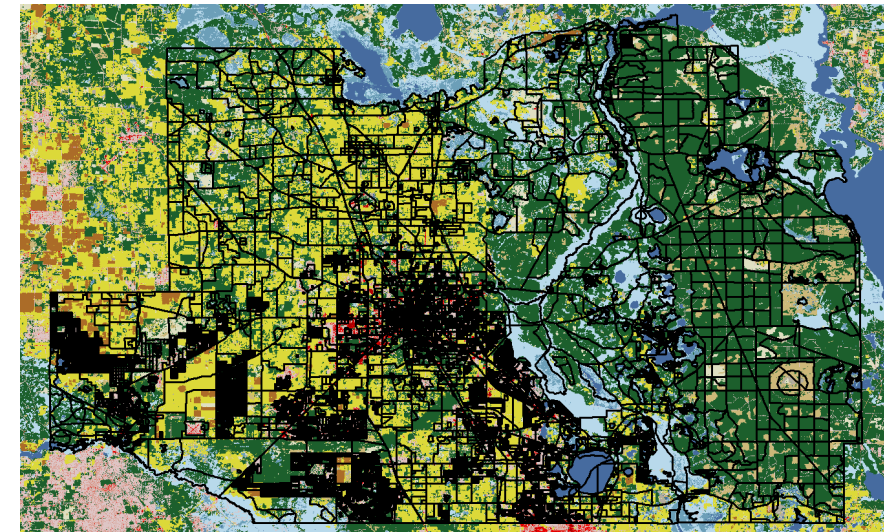
Building and Parcel Properties
#, Area, Surrounding



US Census 5 Year
American Community
Survey Estimates



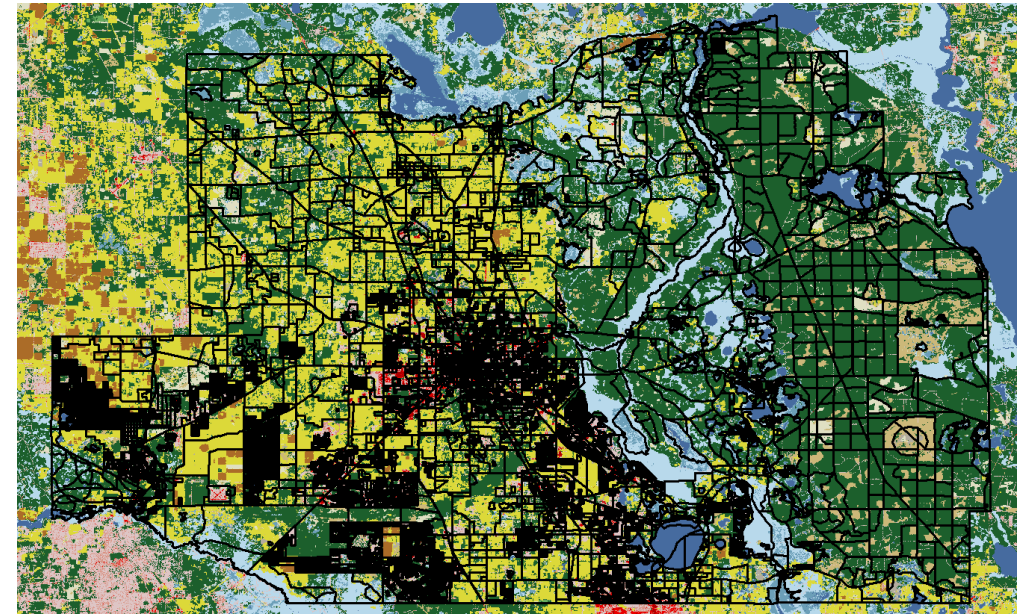
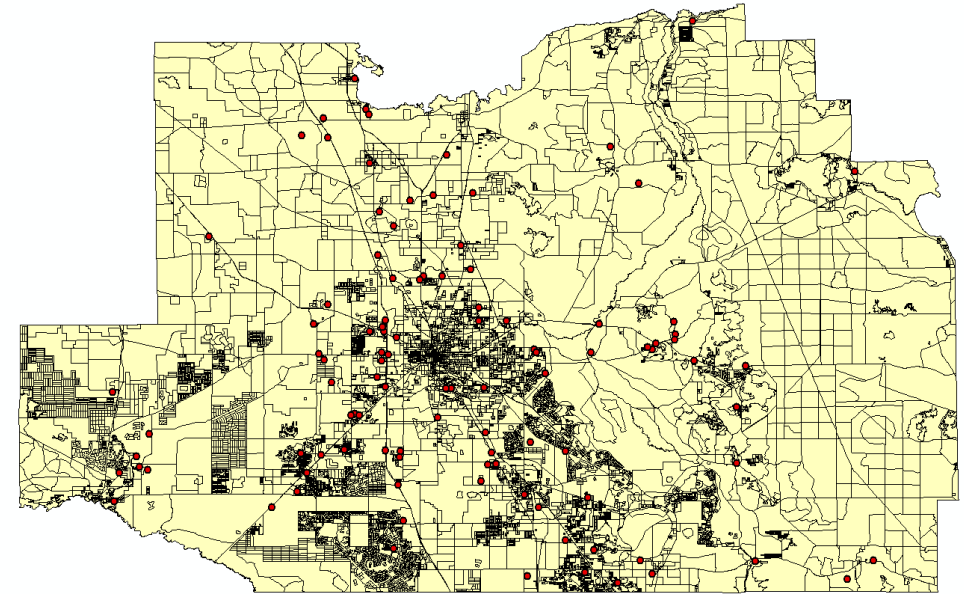
WWTP Distance



Land Cover

Intro to Available Data

- Wastewater Treatment Plant Distance (State Database)
- Parcel Area
- Land Cover (National Landcover Database)



Microsoft Building Footprints

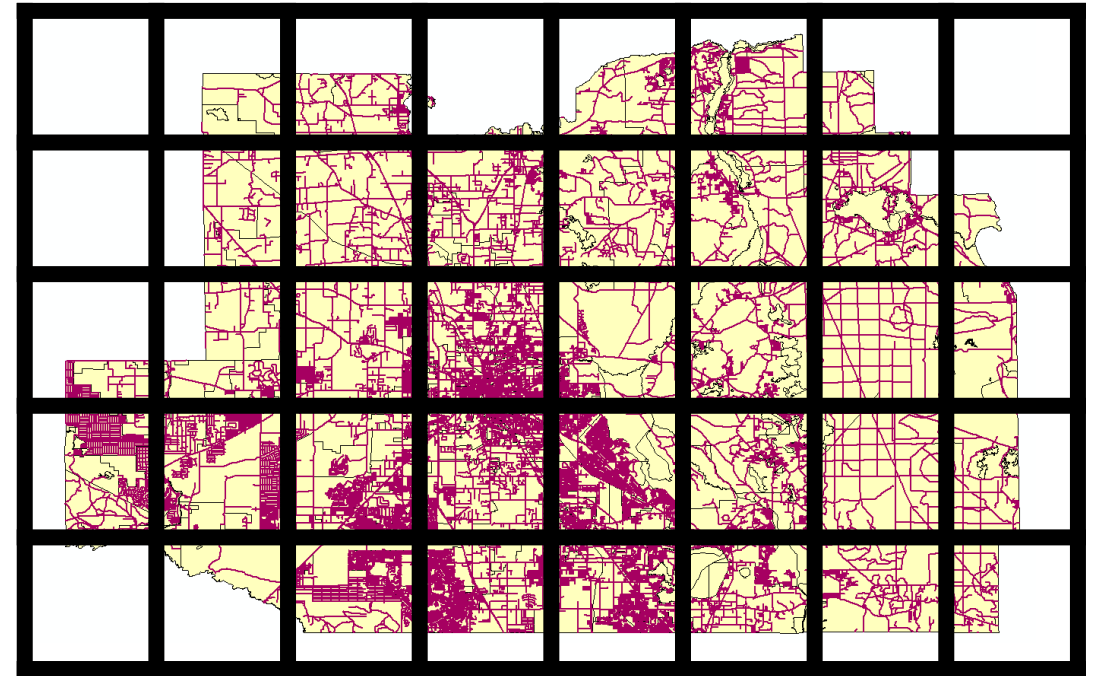
National data layer

- Calculated Variables
- Total Building Area
- Maximum Building Area
- # Buildings Intersecting Parcel
- # Buildings Completely within parcel
- Building Area per Parcel Area
- Non-building area per Parcel Area



Geospatial Processing Technique

- Trying to run these calculations in one shot makes regular computers unhappy
- Break down the dataset into smaller cells
- Run calculations for each cell and then aggregate results
- Ran on the swarm2 high performance computing cluster at UMass



Additional Building Related Properties

- # parcels intersecting the grid cell
- # parcels completely within grid cell
- median parcel area within grid cell (hopefully this will circumvent issues of parcels that fall between 2 or more grid cells)
- # buildings intersecting grid cell
- # buildings completely within grid cell
- total building area within grid cell
- median building area within grid cell

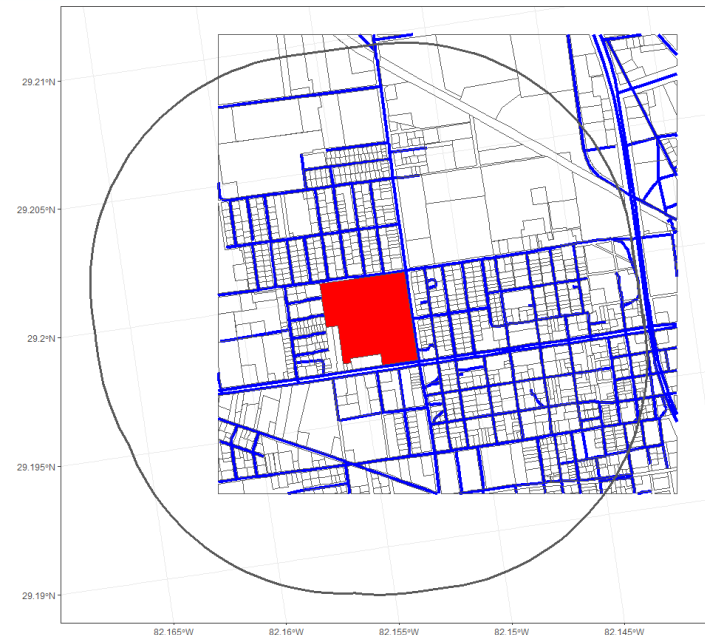
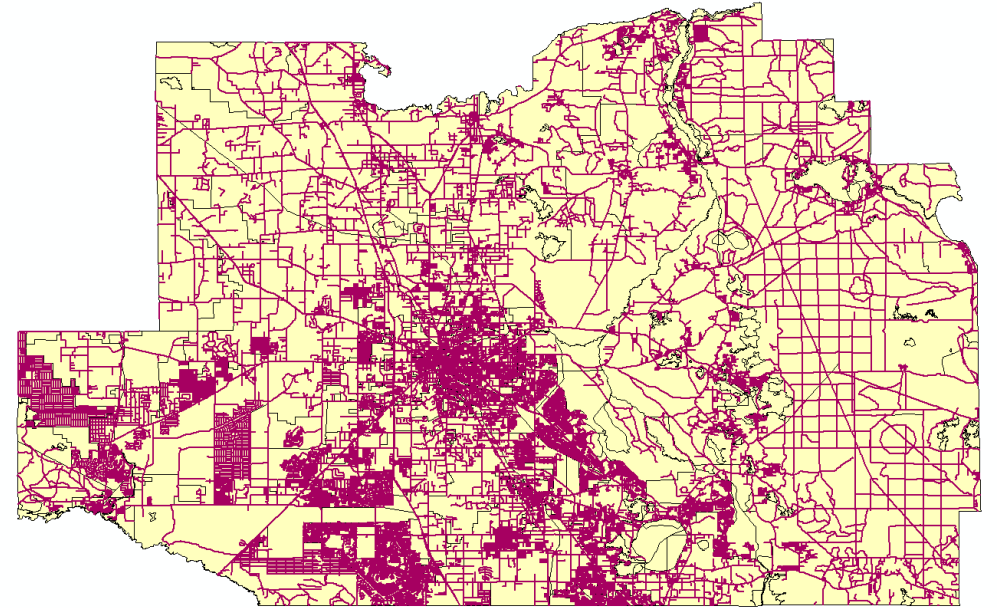


3 Grid Cell Sizes Tested

- 250m
- 500m
- 1000m

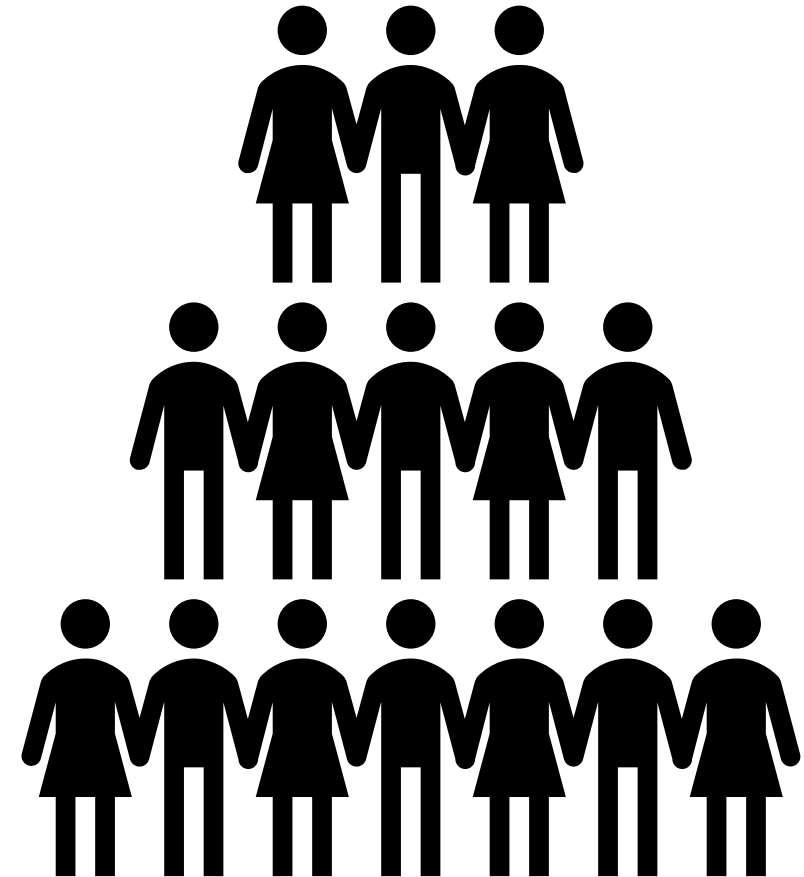
Road Density

- Calculate buffers of various size for each parcel (1km, 2km)
- Calculate length of road falling within the buffer boundary for each buffer size and parcel
- Calculate # of roads falling within buffer boundary for each buffer size and parcel
- Categorize length and number of roads for different road types



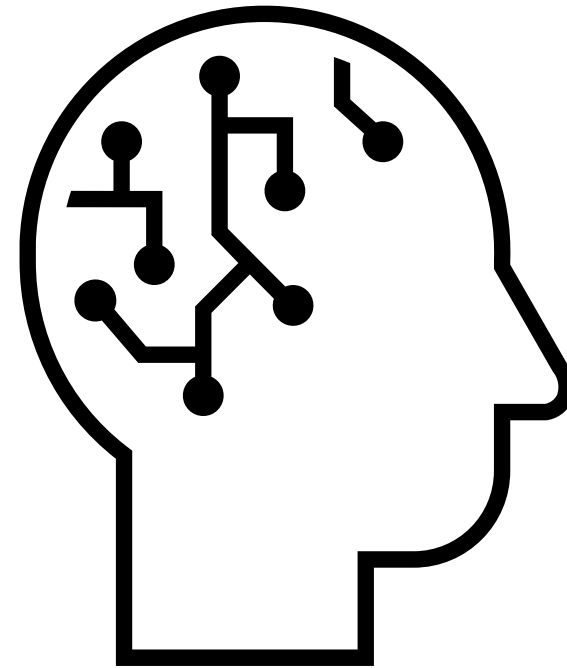
Census Data

- 5 Year American Community Survey Estimates (2019)
- Population
- Commuting Statistics
- Housing Characteristics
 - Rooms
 - Plumbing
 - Year Built
 - Housing Units in Structure
 - Occupancy Type (Owner/Renter)
- Internet Service



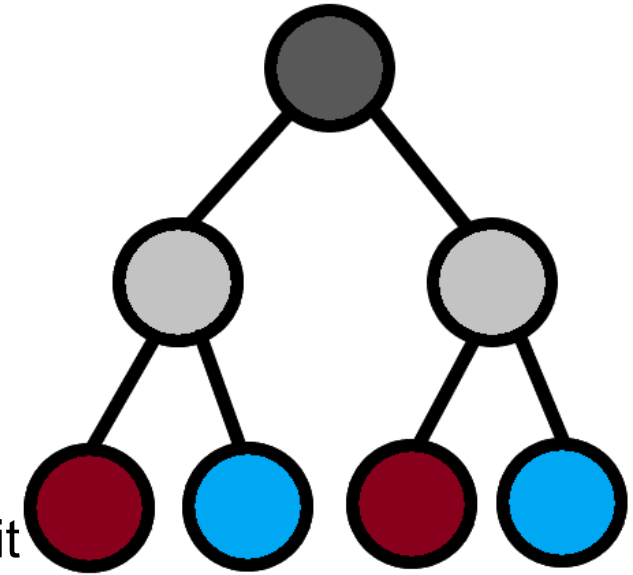
Machine Learning

- Part of the artificial intelligence field
- Focuses on the development of computer programs that can access data and use it to learn for themselves
- Enables analysis of massive quantities of data



Classification Trees

- Method where objects are assigned a class depending on meeting certain properties
- “White Box” or interpretable method
 - We can understand what decisions the model has made because it tells us that it used a particular feature and threshold to separate the data
- Summary: take all the "features" that describe each parcel and figure out which features and thresholds are best for separating our dataset into classes



Machine Learning: Two Classification Tree Method

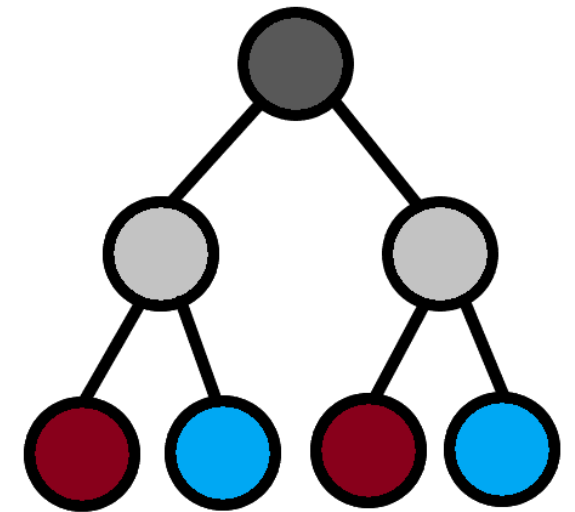
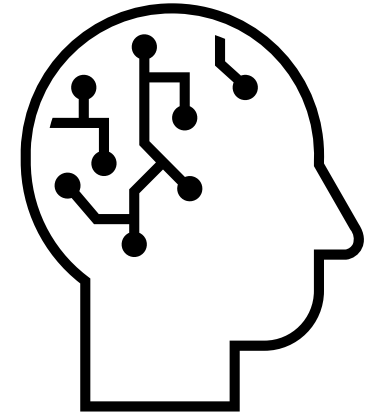
We have to do two classifications!

- So, we use 2 classification trees in sequence

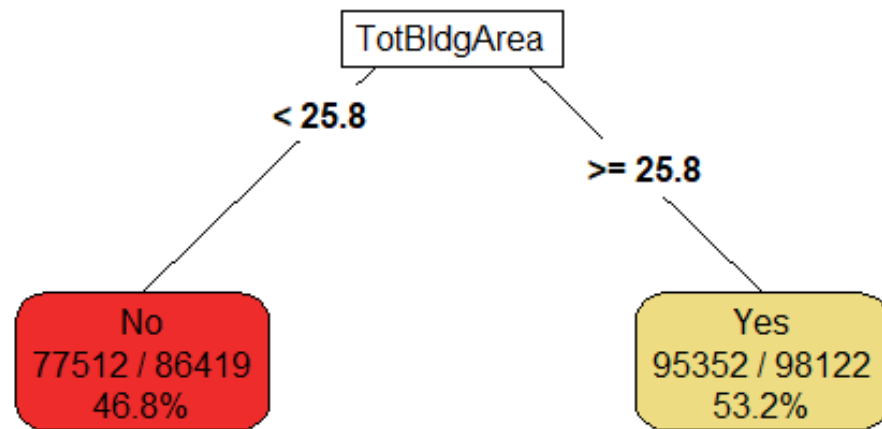
Tree 1: Identify areas where making an estimate of Septic vs. Sewer (SvS) is applicable. (Yes/No)

Tree 2: Categorize areas as Septic or Sewer

Combine results of Tree 1 and Tree 2 to make a prediction



Tree 1: Is selecting Septic or Sewer applicable?



- Tree training data excludes unknown territories
- If Yes: choose Sewer/Septic
- If No: NA (not applicable)

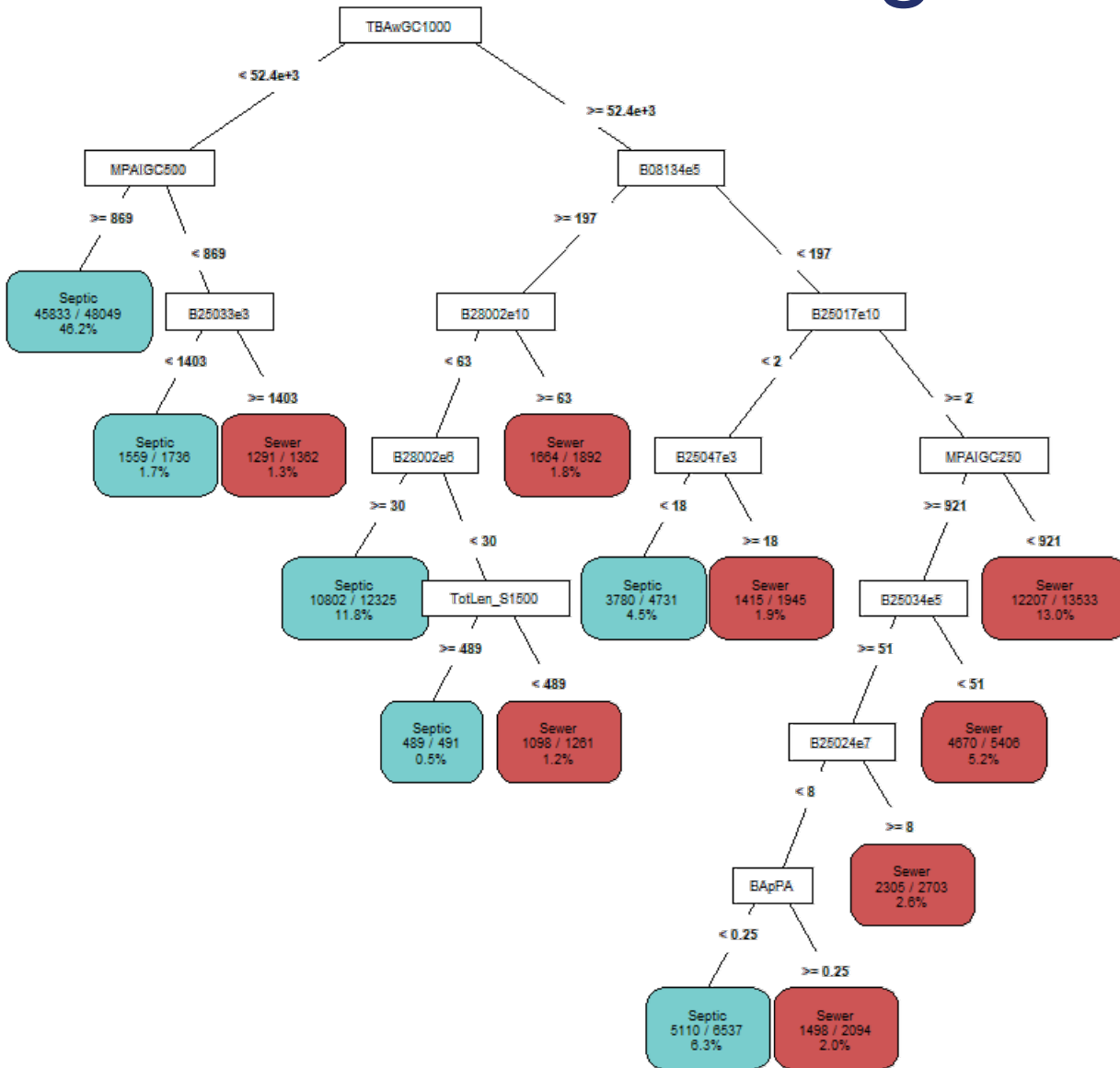
Raw Misclassification Rate: 6.35%

Misclass No as Yes: 3.5%

Misclass Yes as No: 8.6%

		Prediction	
		No	Yes
Actual	No	110957	4008
	Yes	12722	135943

Tree 2: Selecting Sewer/Septic



Tree training data excludes unknown and NonApplicable territories

Raw Misclassification Rate: 9.84%

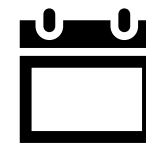
Misclass Septic as Sewer: 5.7%

Misclass Sewer as Septic: 19.1%

		Prediction	
		Septic	Sewer
Actual	Septic	96674	5810
	Sewer	8822	37359

Parameters Used for Tree 2

1. Total Building Area within Grid Cell (1000m)
2. Median Parcel Area intersecting Grid Cell (500m)
3. TOTAL POPULATION IN OCCUPIED HOUSING UNITS BY TENURE BY UNITS IN STRUCTURE: Owner occupied: 1, detached or attached
4. TRAVEL TIME TO WORK: Total: 20 to 24 minutes: Workers 16+ who did not work from home
5. INTERNET SUBSCRIPTIONS IN HOUSEHOLD: Total: Satellite Internet service with no other type of Internet subscription
6. INTERNET SUBSCRIPTIONS IN HOUSEHOLD: Total: Cellular data plan with no other type of Internet subscription
7. S1500 Vehicular Trail (4WD) An unpaved dirt trail where a four-wheel drive vehicle is required.
8. ROOMS: Total: 9 or more rooms
9. Total: Lacking complete plumbing facilities
10. Median Parcel Area intersecting Grid Cell (250m)
11. YEAR STRUCTURE BUILT: Total: Built 1990 to 1999
12. UNITS IN STRUCTURE: Total: 10 to 19
13. Building Area per Parcel Area



Overall Result of 2 Trees

- First used Tree 1 to see if we should apply a Class or NonApp
- Used Tree 2 to apply a Class
- Raw Misclassification Rate: 11.5%

Audience Question: which applications is 11.5% error good enough for, and for which applications is it insufficient?

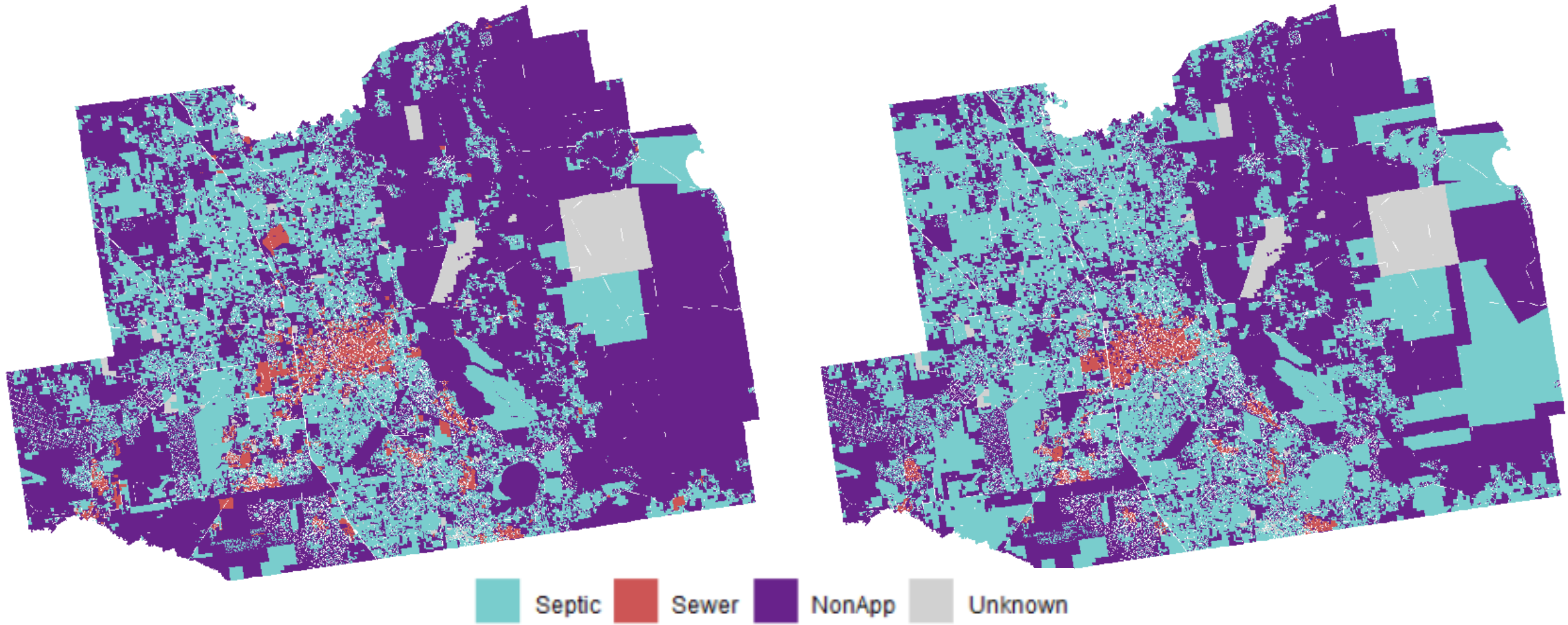
		Prediction		
		Septic	Sewer	NonApp
Actual	Septic	87377	5404	9703
	Sewer	8209	34953	3019
	NonApp	3624	384	110957

Type	% Misclass as Other
Septic	14.7%
Sewer	24.3%
NonApp	3.5%

Known

vs.

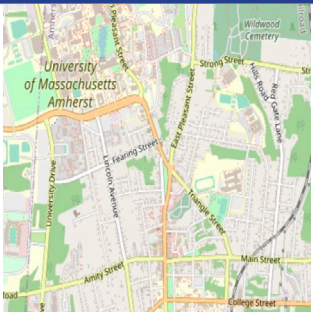
Predicted



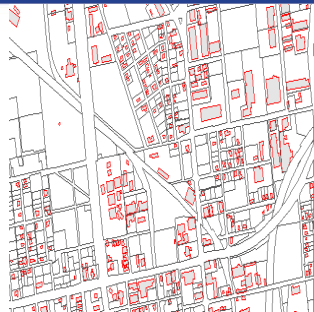
Results are promising for methodology
We overestimate presence of Septic in large parcels
We underestimate presence of Sewer

Moving Forward

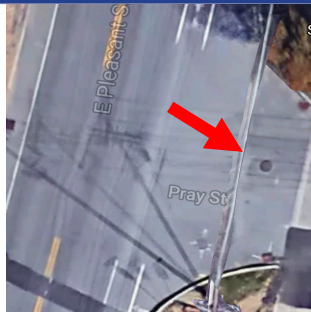
Scaling to the national level



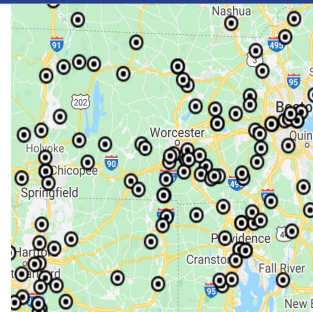
Road networks



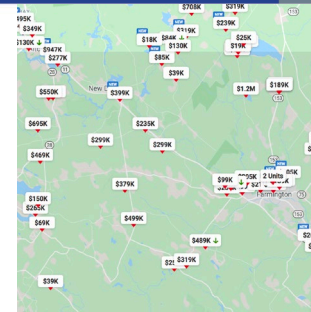
Building footprints



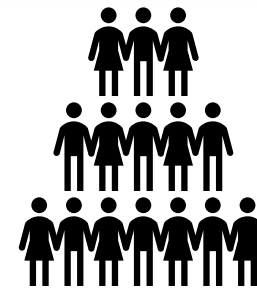
Utility hole covers (via image segmentation)



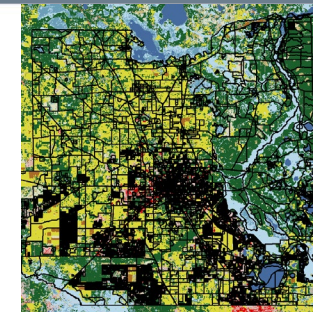
Wastewater treatment plants



Real estate data



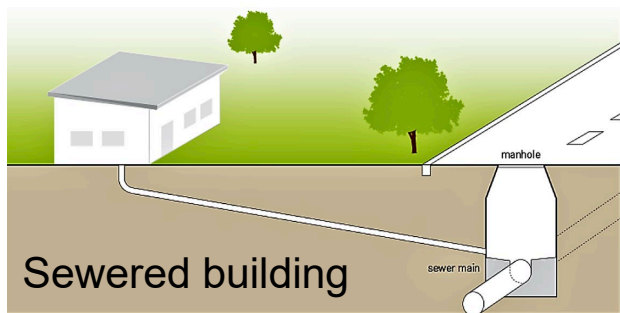
ACS



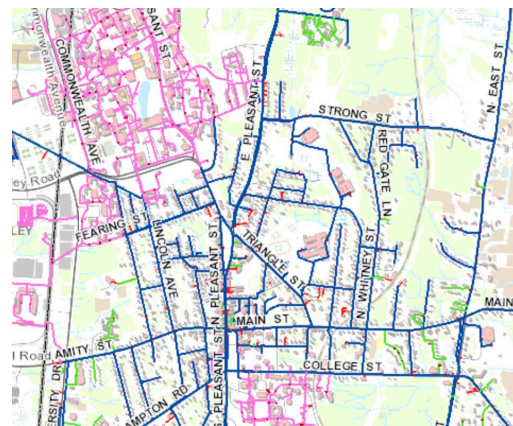
Landcover



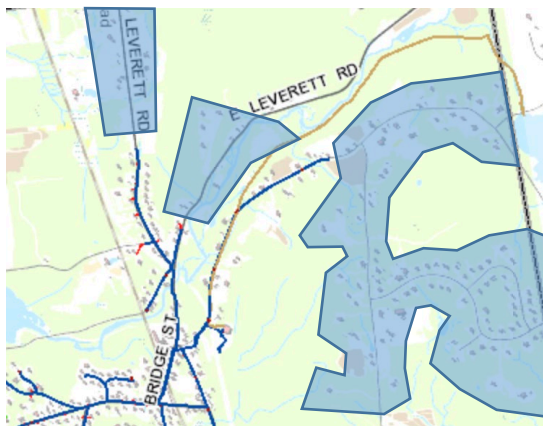
Inferring system coverage via Machine Learning



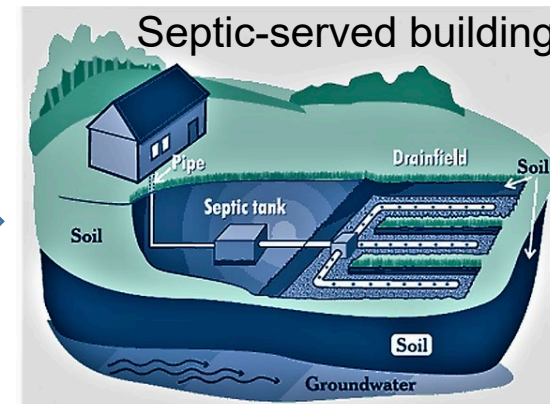
Sewered building



Inferred sewer network coverage



Inferred on-site systems



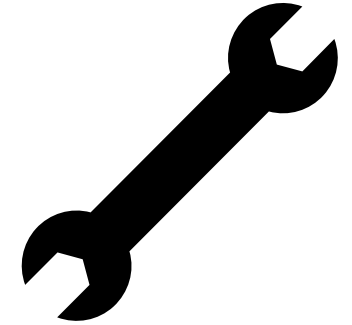
Septic-served building

Next Steps and Improvements

- Iterative Process
 - Several runs before arriving at 2 tree model
 - Additional data has improved model performance from 44.7% misclassification to 11.5% misclassification
- Improve tree performance with
 - Bagging (Bootstrap aggregating)
 - Random Forest
 - Boosting

Neural network techniques

- Able to find more subtle patterns in the data by combining observations across multiple features
- Typically resulting in more accurate predictions than white-box models
- Tradeoff of losing the ability to describe the choices that the model has made.



Next Steps and Improvements

Real Estate Data – Zillow ZTRAX

Utility Hole Covers – Image Segmentation (Machine Learning)

Apply methodology to other Florida counties

Apply methodology to states with Septic data available – DE, RI

Apply methodology to states with Sewer data available – MA



Contact:
Nelson da Luz
ndaluz@umass.edu

UMass
Amherst

Feedback and Questions

Contact:
Nelson da Luz
ndaluz@umass.edu

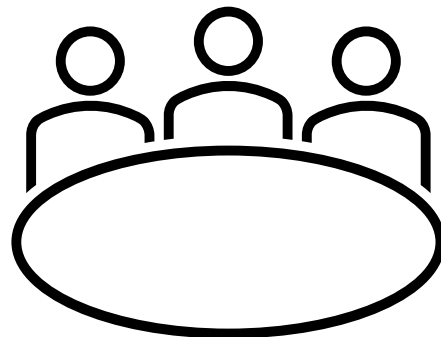
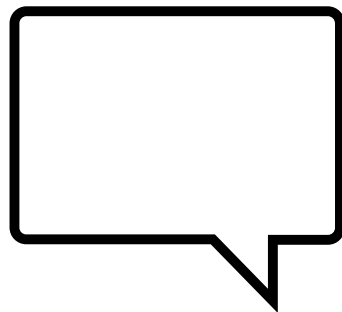
UMass
Amherst

Any questions for us?

Any recommendations on other data we may be able to use?

Questions for the audience:

- Who can use this data and how?
- For what applications is 11.5% error good enough or insufficient?
- What are the risks of depending on permitting data as ground truth?



Thank you!

Contact:
Nelson da Luz
ndaluz@umass.edu

UMass
Amherst