Utilizing Python to Evaluate InVest SDR Model in Puerto Rico

By: Kolin Bilbrew

Introduction

- ❖Puerto Rico is experiencing environmental issues due to climate change and flooding
- ❖Reservoirs are built to improve water quality for residents, but are affected by the accumulation of sedimentation
- ❖However, sediment retention can create a service by creating more wetlands and forested areas

Figure 1: Topographic Map of Puerto Rico.

Research Objectives

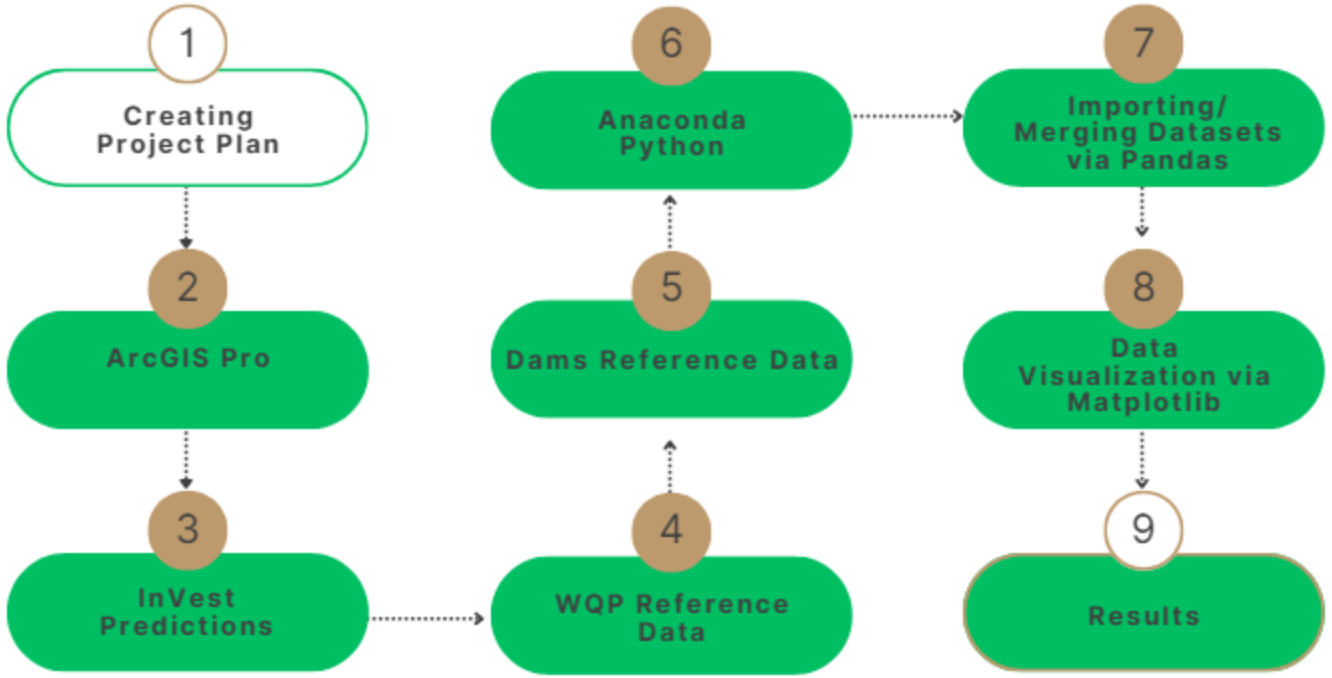
- This study examines different parameters that affects the sediment dynamics and water quality in Puerto Rico in the year 2000.
- Understand the complexities of the environmental issue by visualizing and analyzing the amount of total dissolved and suspended solids based on the extensive data provided.

Goals:

- Analyze data for decision making to use innovative techniques to create improve sediment retention services.
- Explain the data cleansing and Python skills that that illustrates data organization and visualization

InVest SDR Flow Chart

FLOWCHART

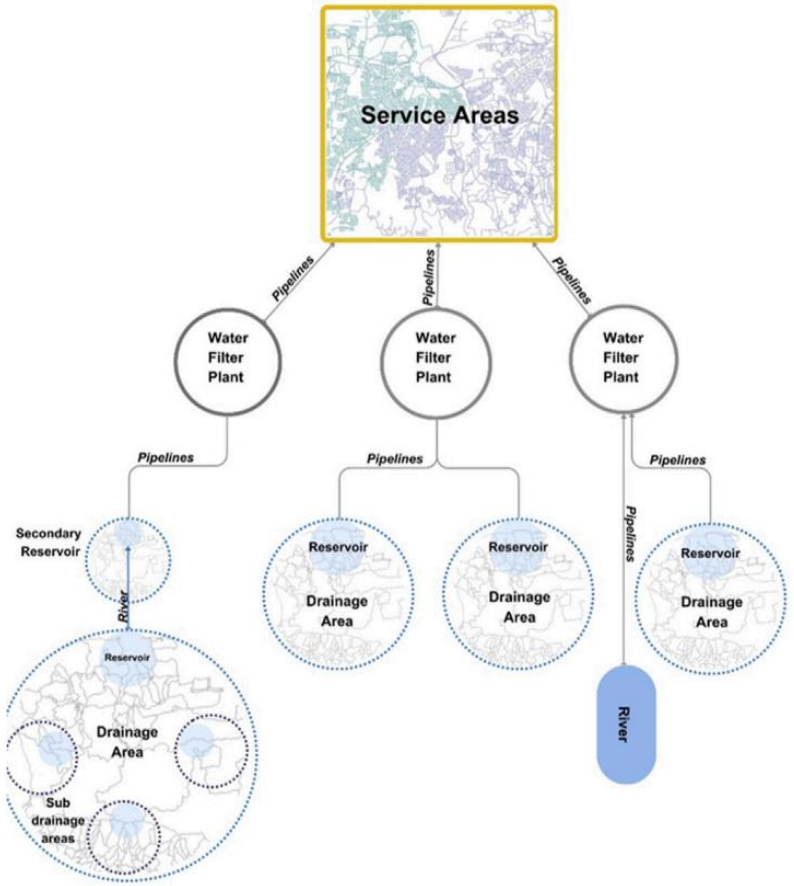


By: Kolin Bilbrew

Methodology: Deciding Parameters

Parameters:

- ❖USGS-MS Identification
- ❖Drainage Names
- ❖Area in square kilometers (SqKM)
- ❖Total Dissolved Solids
- ❖Total Suspended Solids
- ❖Dam Deposition
- ❖ InVest Sediment Exports



(De Jesus Crespo et al., 2023)

Methodology: Importing Packages and Datasets

```
#Script used to automate the calibration of the NDR Model Adapted from Campanhao et al. 2023
import logging
import sys

import natcap.invest.sdr.sdr
import natcap.invest.utils
```

```
import os
import pandas as pd
import numpy as np
```

```
df1= 'C:/Users/kbilbr3/Documents/Puerto Rico Watershed Coding/intercept_dams_wqp2.csv'
intercept_file= pd.read_csv(df1)
print(intercept_file)
```

```
df4= 'C:/Users/kbilbr3/Documents/Puerto Rico Watershed Coding/wqp_data2.csv'
wqp_file= pd.read_csv(df4)
print(wqp_file)
```

```
df3='C:/Users/kbilbr3/Documents/Puerto Rico Watershed Coding/invest_dams_data.csv'
invest_dams_file= pd.read_csv(df3)
print(invest_dams_file)
```

```
df2= 'C:/Users/kbilbr3/Documents/Puerto Rico Watershed Coding/dam_deposition_data.csv'
dam_deposition_file= pd.read_csv(df2)
print(dam_deposition_file)
```

Methodology: Merging Datasets

- Merging datasets through Pandas
- Drainages placed in alphabetically and numbered
- Placed the parameters in an organized code to start data visualization

```
[14]: wqp_invest_dam_deposition_left_join= pd.merge(wqp_invest_dams_left_join, dam_deposition_file, on='Drainage',  how='left')
      print(wqp_invest_dam_deposition_left_join)
```

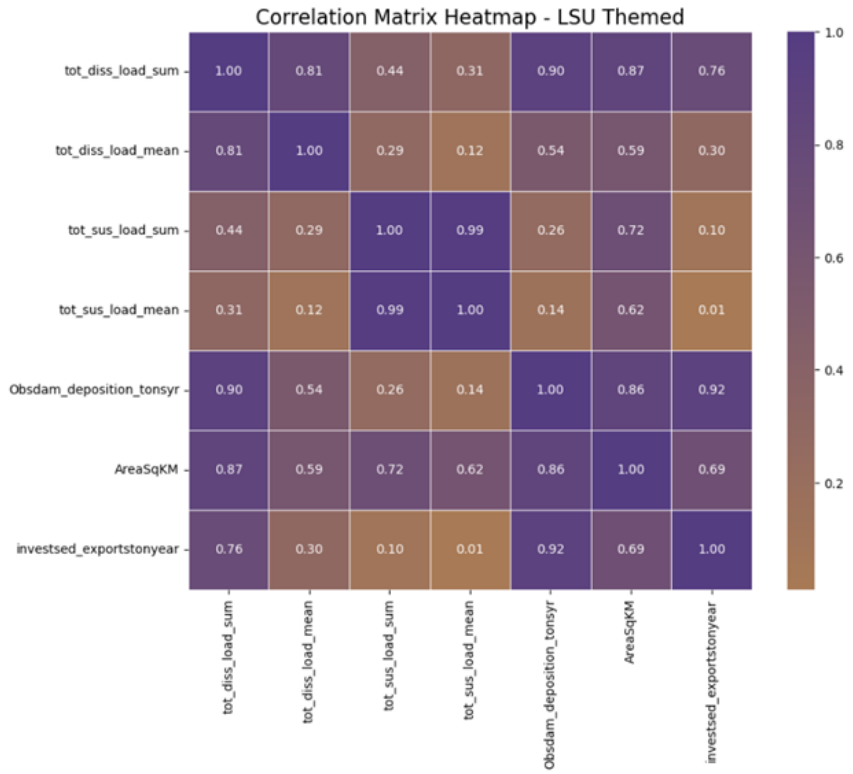
	Drainage	tot_diss_load_sum	tot_diss_load_mean	tot_sus_load_sum	tot_sus_load_mean	AreaSqKM	investsed_exportstonyear	Obsdam_deposition_tonsyr
0	Canovanas	3235.125918	3235.125918	0.000000	NaN	NaN	NaN	NaN
1	Cerrillos	6336.050072	6336.050072	2776.624700	2776.624700	45.09	12441.28	54000.0
2	Coamo	12604.507790	12604.507790	986.472245	986.472245	169.67	45497.53	NaN
3	DosBocas	89018.092834	12716.870405	2023.669041	2023.669041	436.03	487524.37	554400.0
4	Fajardo	5174.149644	5174.149644	3433.621680	3433.621680	27.37	3700.83	1800.0
5	Guajataca	7711.135057	3855.567528	1943.060528	971.530264	60.46	24632.35	124200.0
6	LaPlata	106548.761270	53274.380635	9201.159705	4600.579853	467.45	172625.10	437400.0
7	Loco	40146.203663	20073.101832	0.000000	NaN	21.84	18457.14	41400.0
8	Loiza	60826.169028	15206.542257	30343.503650	30343.503650	537.91	92614.39	244800.0
9	Patillas	6235.443878	6235.443878	983.946059	983.946059	66.65	17512.46	72000.0
10	Portuguez	11844.328640	11844.328640	2212.319178	2212.319178	27.10	2492.13	18000.0
11	RioBlanco	577.384379	288.692189	0.000000	NaN	29.12	328.86	NaN

Methodology: Data Visualization

- Imported packages such as Matplotlib and Seaborn
- Bar Graph and Correlation Analysis of Normalized Values
- Evaluated data to understand the results of the different parameters

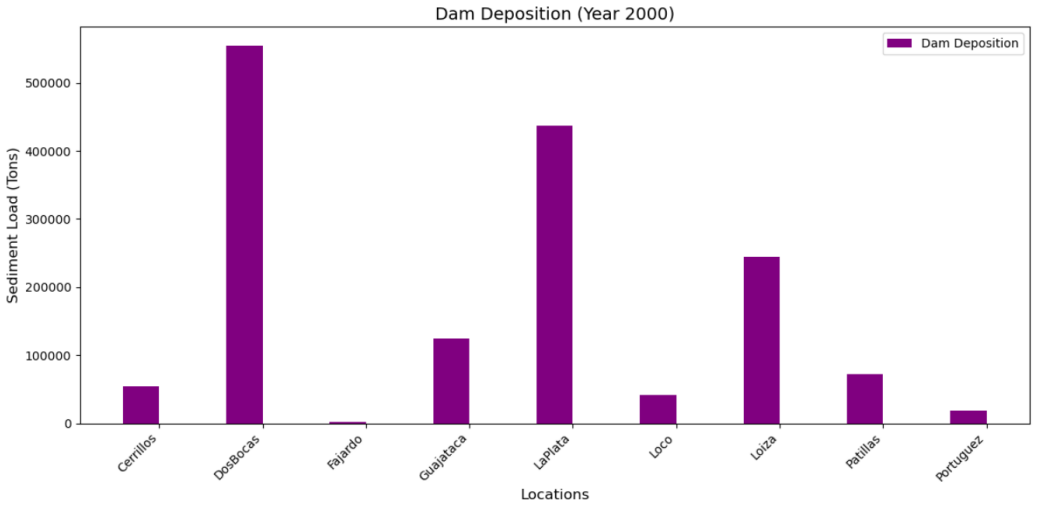
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import matplotlib.colors as mcolors

# Assuming your DataFrame is named 'wqp_invest_dep_data'
# If loading from a CSV file, replace this with the actual file path
# wqp_invest_dep_data = pd.read_csv('your_file.csv')

# Example: Replace this line with your actual data-loading step

data = pd.DataFrame(wqp_invest_dam_deposition_left_join)

# List of relevant columns to convert to numeric
columns_to_convert = [
    'tot_diss_load_sum',
    'tot_diss_load_mean',
    'tot_sus_load_sum',
    'tot_sus_load_mean',
    'Obsdam_deposition_tonsyr',
    'AreaSqKM',
    'investsed_exportstonyear'
]

# Ensure columns are numeric, coercing errors to NaN
data[columns_to_convert] = data[columns_to_convert].apply(pd.to_numeric, errors='coerce')

# Step 2: Compute the correlation matrix for the relevant columns
correlation_matrix = data[columns_to_convert].corr()

# Step 3: Create LSU-themed color palette (LSU Purple and Gold)
lsu_purple = '#553D82'  # LSU Purple
lsu_gold = '#FDB927'    # LSU Gold

# Create a custom colormap (from Gold to Purple)
cmap = mcolors.LinearSegmentedColormap.from_list("LSU_Gold_Purple", [lsu_gold, lsu_purple])

# Step 4: Plot the heatmap with LSU-themed colors
plt.figure(figsize=(10, 8))  # Set the figure size for better readability
sns.heatmap(correlation_matrix, annot=True, cmap=cmap, fmt='.2f', linewidths=0.5, center=0)

# Step 5: Display the heatmap with a title
plt.title('Correlation Matrix Heatmap - LSU Themed', fontsize=16)
plt.show()
```

Results (Correlation Matrix Map)



- ❖Area of the drainage area impacts the flow and dynamics of sediment
- ❖Negative correlations with Total Suspended Solids
- ❖Dam deposition is impacted by the sediments and area of the drainage system

Results (Dam Deposition)



- Dos Bocas and LaPlata had high volumes of sediment being deposited
- Canovanas, Caomo, and Rio Blanco did not contain any data
- High sediment deposition in dams hinders water storage volumes

Limitations and Challenges

❖Limitations of Other Factors Within Dataset

- Factors such as turbidity were not used in the project.

- Mapping the drainages via ArcGIS Pro

❖Challenges with Pandas and Merging Datasets

- Organizing the dataset for merging

❖Pivoting and Quick Decision Making

	A	B	C	D
1	OID_USGS_MSDrainage			
2	1USGS-50071000Fajardo			
3	2USGS-50075000RioBlanco			
4	3USGS-50074950RioBlanco			
5	4USGS-50061800Canovanas			
6	5USGS-50056400Loiza			
7	6USGS-50051310Loiza			
8	7USGS-50057000Loiza			
9	8USGS-50055000Loiza			
10	9USGS-50092000Patillas			
11	10USGS-50043000LaPlata			
12	11USGS-50046000LaPlata			
13	12USGS-50106500Coamo			
14	13USGS-50114000Cerrillos			
15	14USGS-50027000DosBocas			
16	15USGS-50115000Portuguez			
17	16USGS-50027250DosBocas			
18	17USGS-50023000DosBocas			
19	18USGS-50021050DosBocas			
20	19USGS-50021000DosBocas			
21	20USGS-50021500DosBocas			
22	21USGS-50020500DosBocas			
23	22USGS-50010500Guajataca			
24	23USGS-50129000Loco			
25	24USGS-50129500Loco			
26	25USGS-50011400Guajataca			
27				

	A	B	C	D
1	USGS_MS	OID_	Drainage	
2	USGS-500	1	Fajardo	
3	USGS-500	2	RioBlanco	
4	USGS-500	3	RioBlanco	
5	USGS-500	4	Canovanas	
6	USGS-500	5	Loiza	
7	USGS-500	6	Loiza	
8	USGS-500	7	Loiza	
9	USGS-500	8	Loiza	
10	USGS-500	9	Patillas	
11	USGS-500	10	LaPlata	
12	USGS-500	11	LaPlata	
13	USGS-501	12	Coamo	
14	USGS-501	13	Cerrillos	
15	USGS-500	14	DosBocas	
16	USGS-501	15	Portuguez	
17	USGS-500	16	DosBocas	
18	USGS-500	17	DosBocas	
19	USGS-500	18	DosBocas	
20	USGS-500	19	DosBocas	
21	USGS-500	20	DosBocas	
22	USGS-500	21	DosBocas	
23	USGS-500	22	Guajataca	
24	USGS-501	23	Loco	
25	USGS-501	24	Loco	
26	USGS-500	25	Guajataca	
27				

Discussion

- In the year 2000, Loiza and LaPlata contained higher sediment loads due to the amount of runoff and erosion.
- Drainages with high sediment loads mitigate erosion and create more vegetation, such as forested areas, to reduce pollutants.
- Different geological characteristics plays a key role in sediment dynamics in drainage systems

References

- De Jesus Crespo, R., Valladares-Castellanos, M., Mihunov, V. V., & Douthat, T. H. (2023). Going with the flow: The supply and demand of sediment retention ecosystem services for the reservoirs in Puerto Rico. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/fenvs.2023.1214037
- Murphy, S. F., & Stallard, R. F. (n.d.). *Hydrology and Climate of Four Watersheds in Eastern Puerto Rico*.
- Yuan, Y., Jiang, Y., Taguas, E. V., Mbonimpa, E. G., & Hu, W. (2015). Sediment loss and its cause in Puerto Rico watersheds. *SOIL*, *1*(2), 595–602. https://doi.org/10.5194/soil-1-595-2015