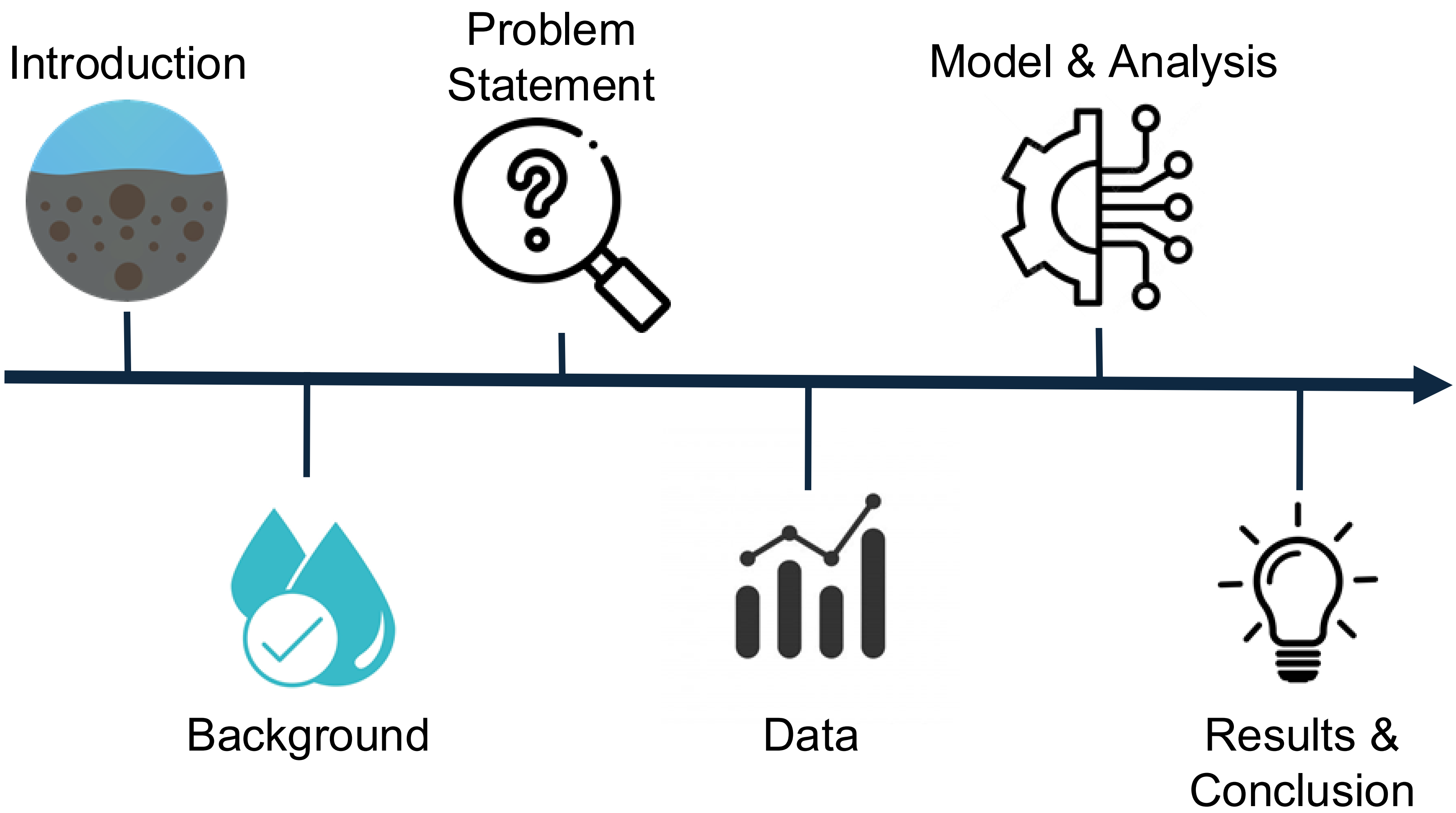
Clear Data, Clearer Waters: An Interactive Dashboard for Visualizing Stream Turbidity Trends

Abiodun Ayo-Bali[1], Payton Davis[2], Sergio Mendoza[3], Galina Shinkareva[4], Srilani Wickramasinghe[5]

Advisor: [1]Dan Bain, [2]Dara Park, [3]Jamie Trammell, [5] Marty D. Frisbee

IOWA
CLEMSON
University of Pittsburgh
Southern OREGON UNIVERSITY
CUAHSI allied for water science www.cuahsi.org
Illinois State University
Purdue University

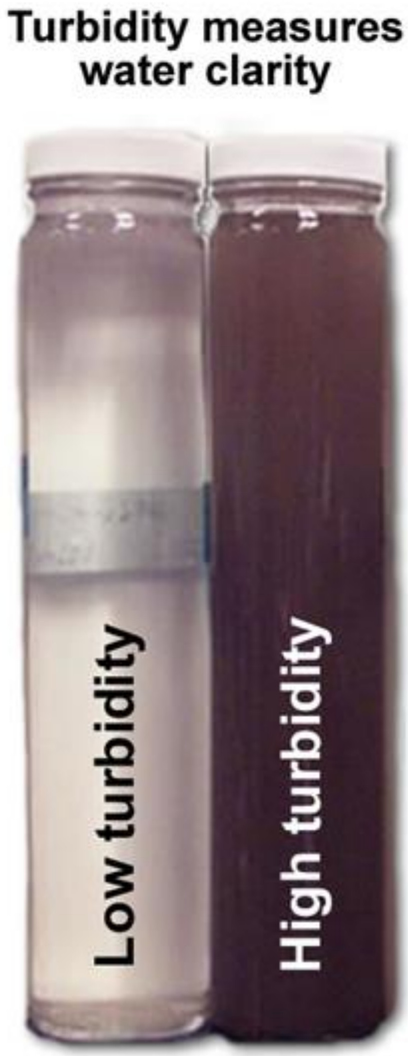
Outline


Introduction
Problem Statement
Model & Analysis
Background
Data
Results & Conclusion

Introduction/ Problem motivation

Introduction/ Problem motivation

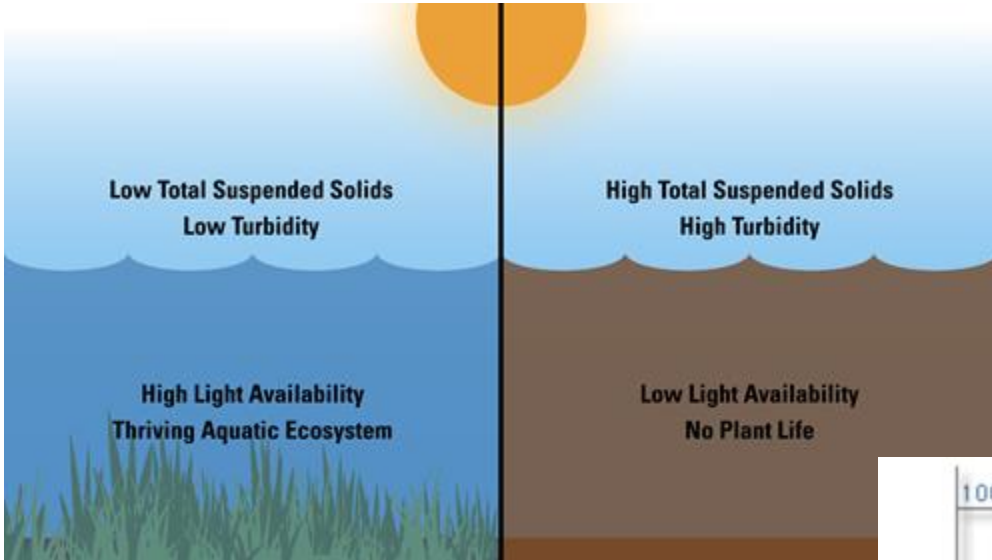
What is Turbidity?

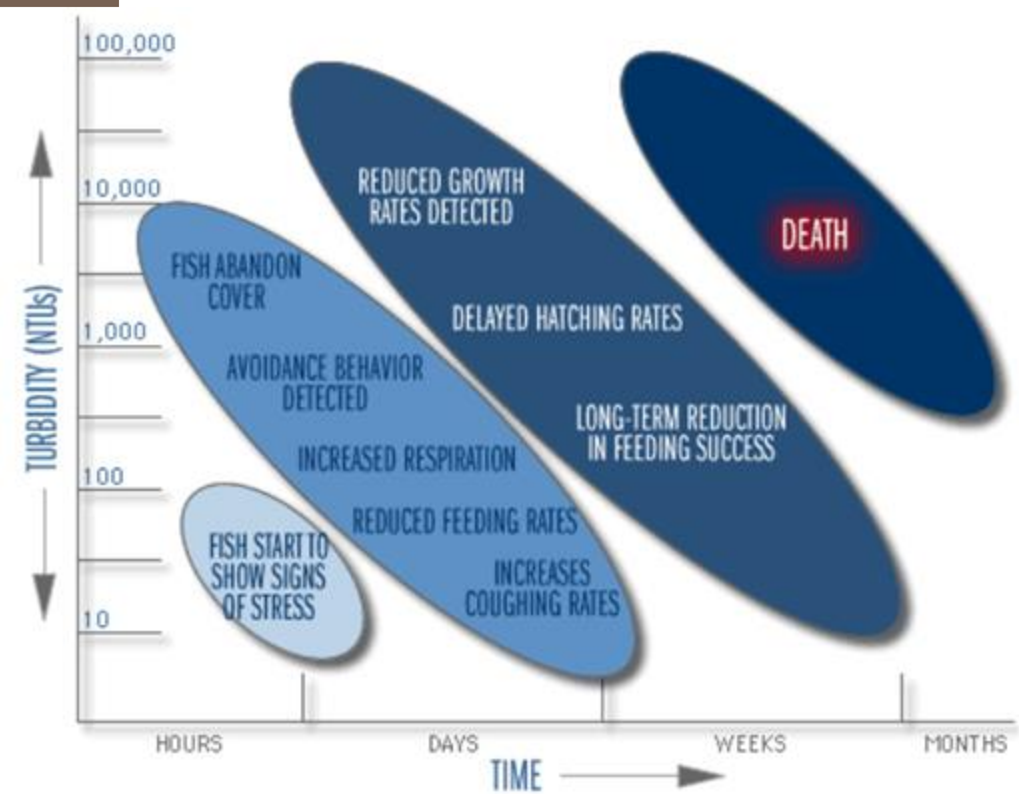


Turbidity is linked to the look of water and therefore the public's perception of water quality.

Sources: https://www.usgs.gov/; https://mrbdc.mnsu.edu/

Turbidity is critical for a thriving aquatic ecosystem





Source: (Newcombe & Jensen, 1996)

Turbidity also affects water utility

- WHO recommendation for drinking water **≤5 NTU**. For chlorinated water <1 NTU
- In the US the turbidity **cannot exceed 1 NTU** at the plant outlet
- Most drinking water utilities aim for levels as low as 0.1 NTU.

Source: https://www.manxtechgroup.com/measuring-turbidity

Problem Statement

- Only a **limited number of USGS stations** record stream turbidity data.
- Manually, filtering out these stations and accessing their daily data-updates is **laborious & time consuming**.

Water samples from the Le Sueur River, Minnesota in 2002

- Our team identified the **lack of an automated, easy to access system** dedicated to stream turbidity, that can expedite the above process.

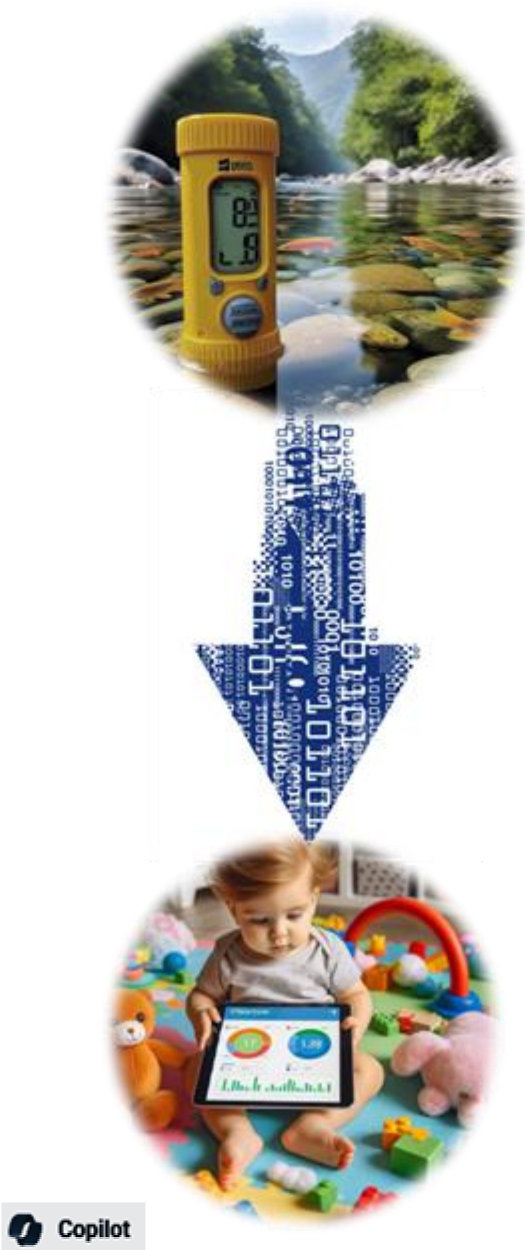
Source: https://mrbdc.mnsu.edu/

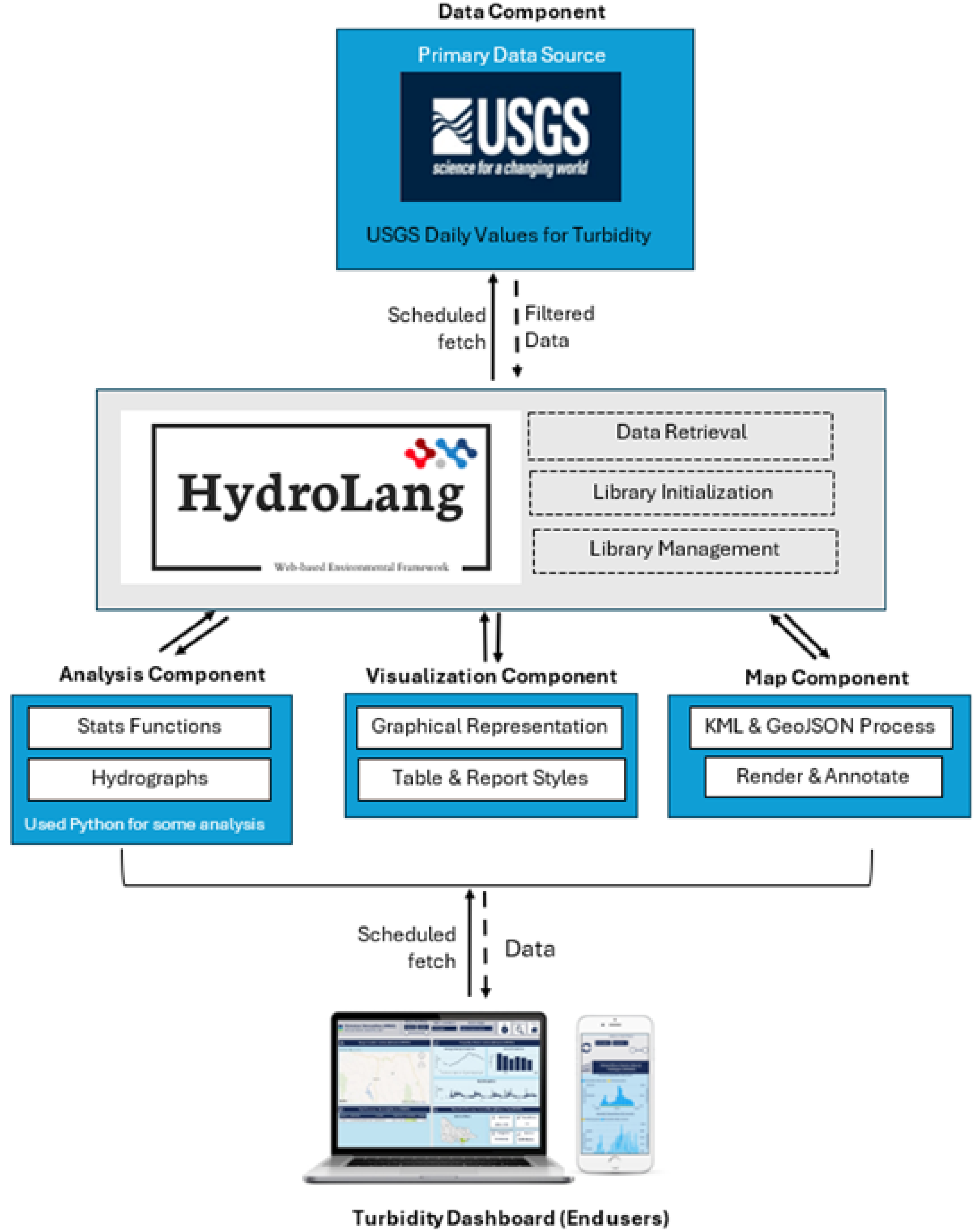
Dashboards: present data in a user-friendly design

- Dashboards are crucial in today's data driven environment (Sarikaya et al. 2019)

- A valuable tool for decision-making & problem solving (Abduldaem & Gravell, 2019)

Flow Diagram

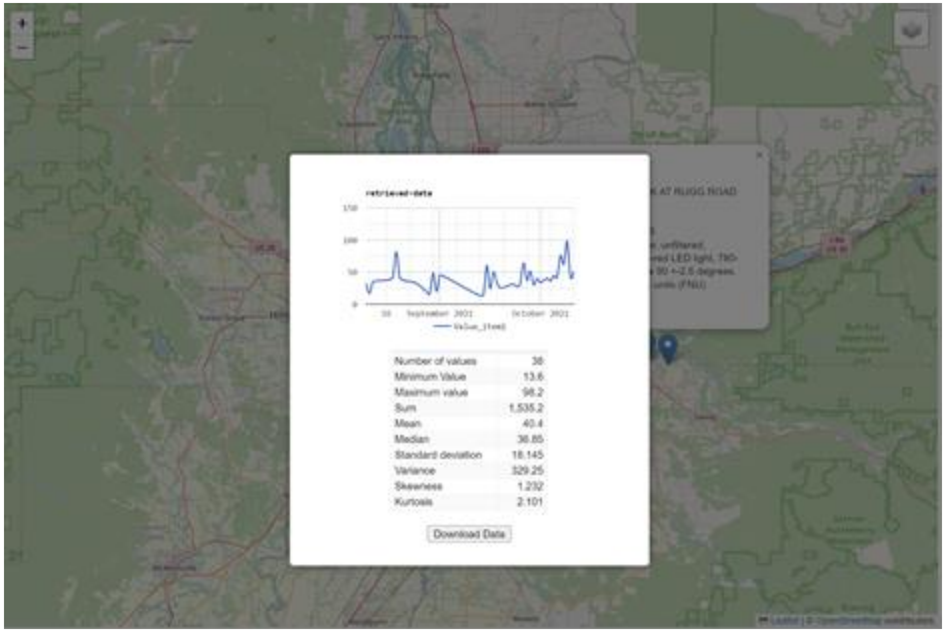




Data

- Turbidity levels (FNU units, variable code 63680)
- Commonly recorded at a fixed interval of 15- to 60-minutes and transmitted to the USGS every hour
- USGS sites within Portland area, OR
- HydroLang to connect to the USGS API and fetch instantaneous turbidity data

Analysis: data preparation



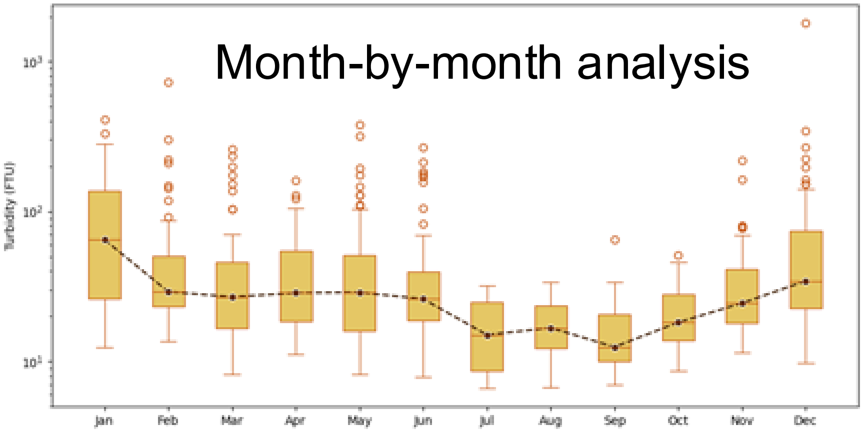
A1 datetime

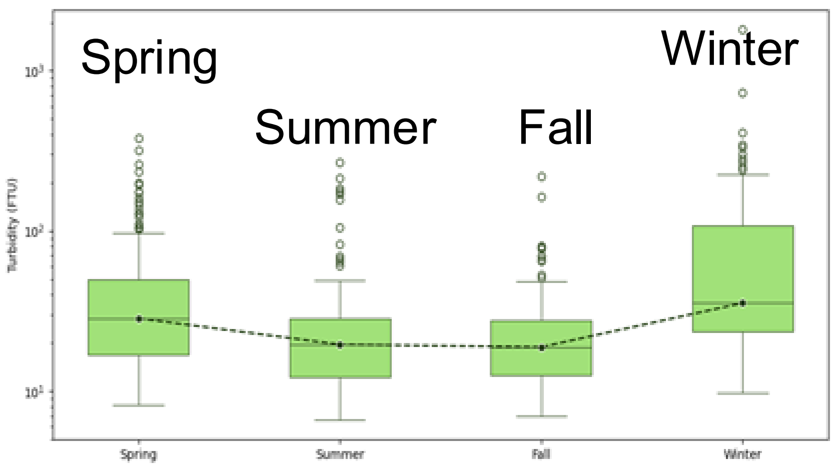
	A	B	C	D	E
1	datetime	20.2			
2	2021-08-1	32.3			
3	2021-08-1	16.6			
4	2021-08-1	34.7			
5	2021-08-1	41.8			
6	2021-08-1	81.7			
7	2021-08-2	41.9			
8	2021-08-2	32.5			
9	2021-08-2	19.8			
10	2021-08-2	16.6			
11	2021-08-3	48.5			
12	2021-08-3	19.8			
13	2021-09-0	45.2			
14	2021-09-1	13.6			
15	2021-09-1	60.5			
16	2021-09-1	24.3			
17	2021-09-1	50			
18	2021-09-1	33.6			
19	2021-09-1	25.1			
20	2021-09-2	31.3			
21	2021-09-2	27.3			
22	2021-09-2	32.6			
23	2021-09-2	64.2			
24	2021-09-2	36.9			
25	2021-09-2	51.4			
26	2021-09-2	30.1			
27	2021-09-3	40			
28	2021-10-0	33.6			
29	2021-10-0	36.8			
30	2021-10-0	40.3			
31	2021-10-0	36.2			
32	2021-10-0	44.2			
33	2021-10-0	42.3			
34	2021-10-0	75.9			
35	2021-10-0	62.1			
36	2021-10-0	98.2			
37	2021-10-1	41.8			
38	2021-10-1	51.3			
39	2021-10-1	undefined			
40					

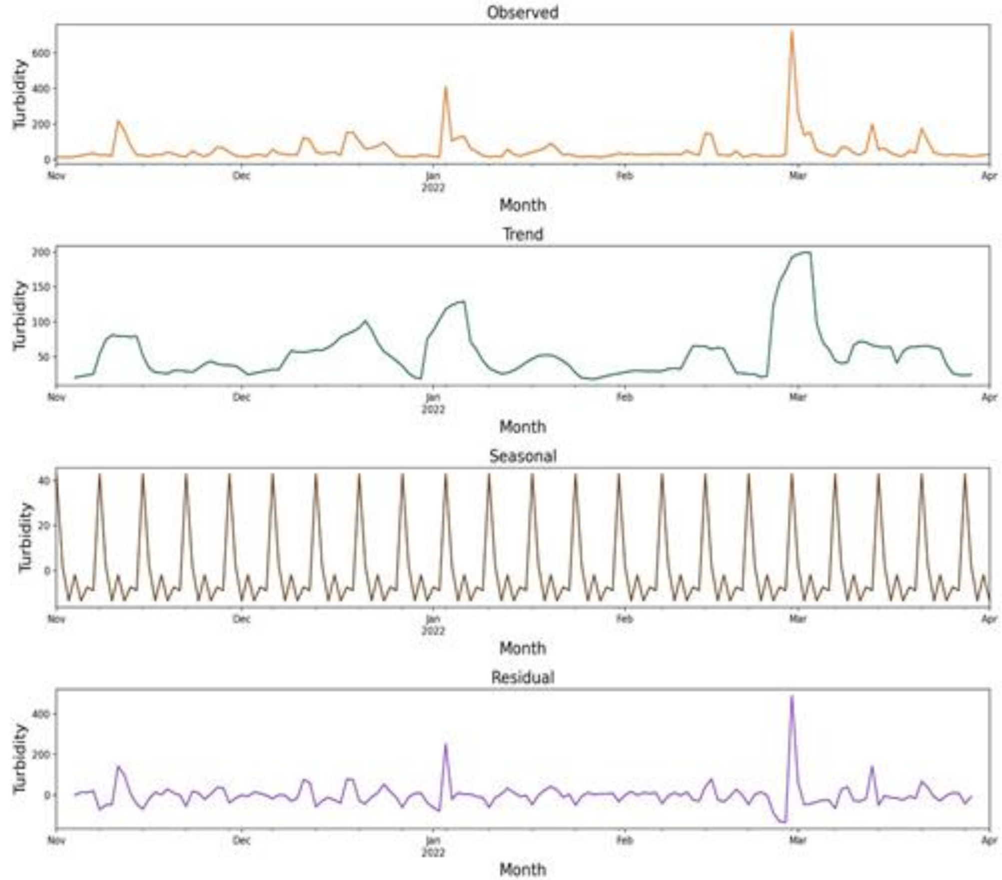
- Handle USGS data misalignment for values and dates.

- Handle missing values:
  - Large periods with missing data were excluded from the analysis.
  - Smaller gaps - interpolation with 'time' parameter to fill NaN values for time series data.

Analysis: basic statistics & seasonal trends

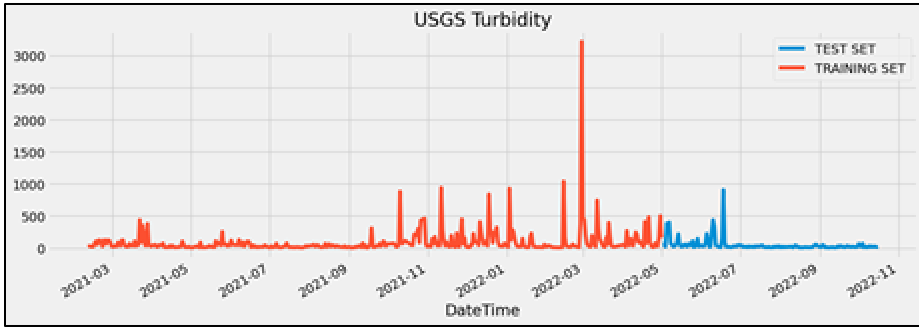


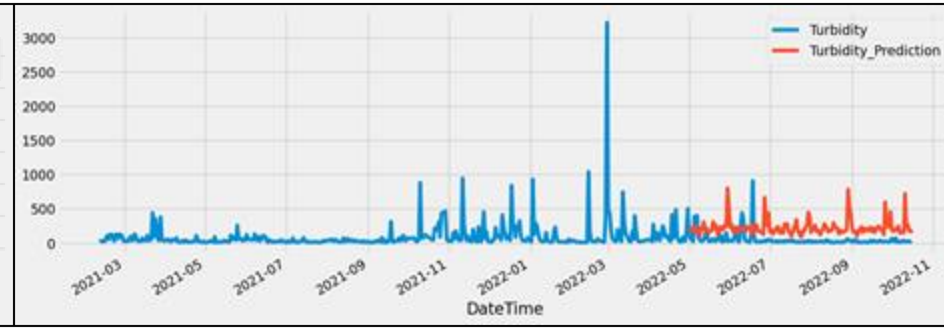


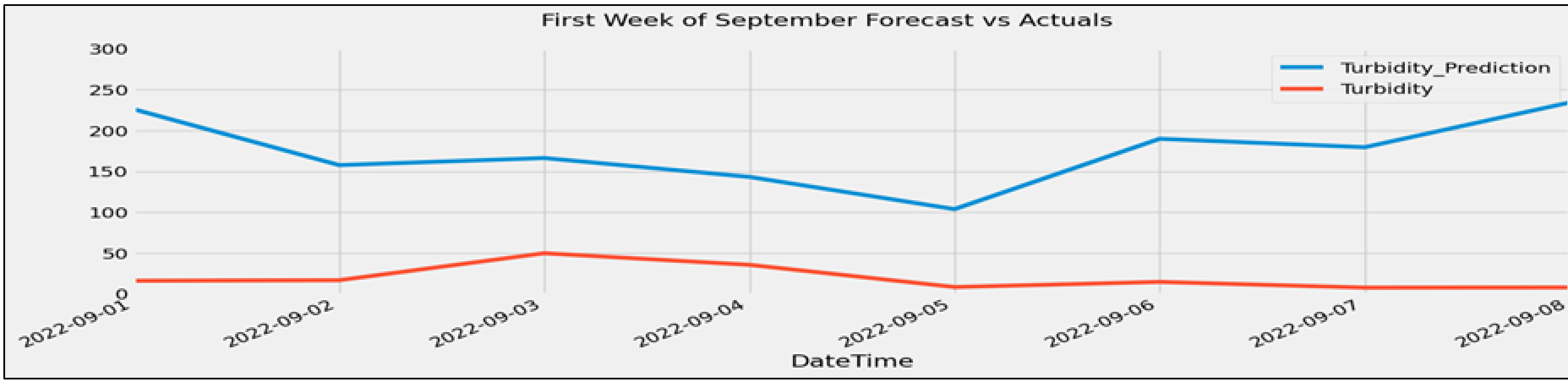


- Basic statistics (median values, min & max)
- Boxplots for quick seasonal trends visualization
- Time series decomposition analysis

Analysis: forecast models (XGBoost)

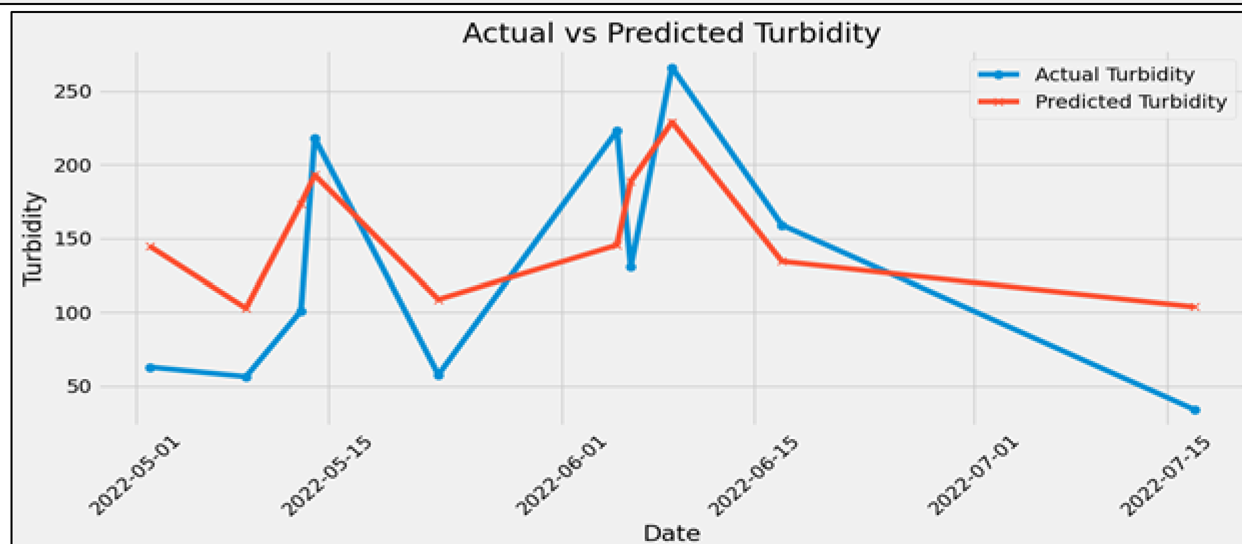






Error Metrics On Test Set

- RMSE error is 51619
- MAE error is 192.70
- MAPE error is 970.78



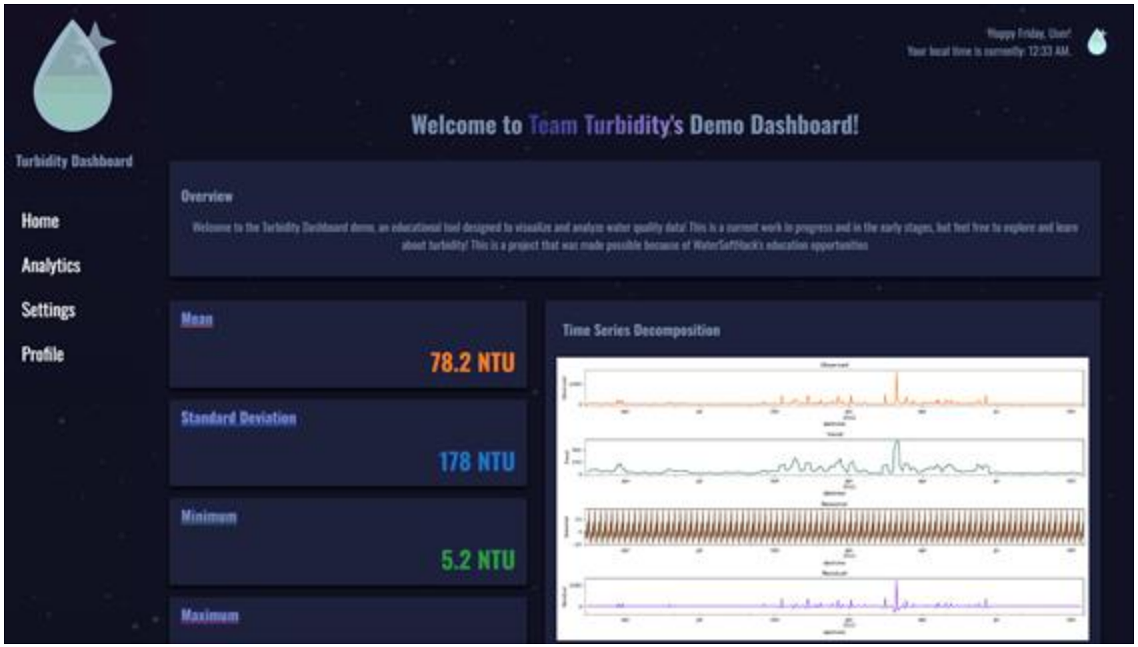
Results & Recommendations

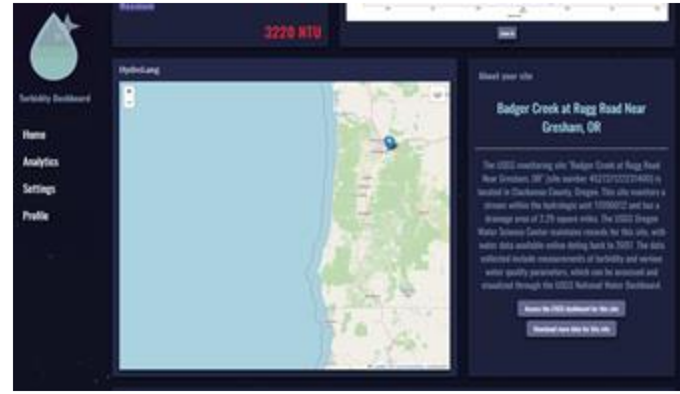
- There are visible trends in seasonal turbidity changes in the Portland area: lower turbidity values at the end of summer (July-September), higher values during winter months (December-January).

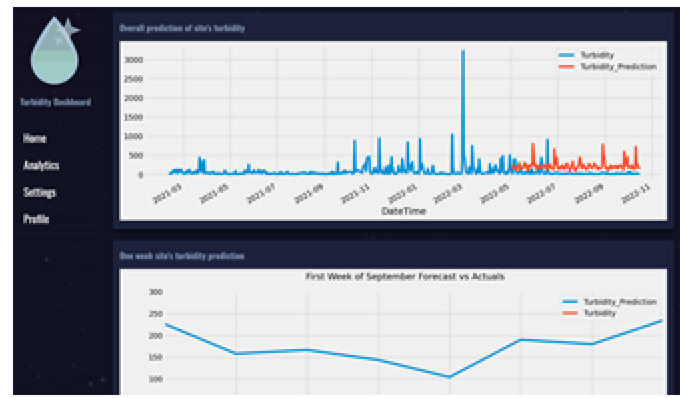
- Additional watershed and stream parameters are needed for a comprehensive analysis of water quality and ecosystem health.

Turbidity Dashboard

Link to Live Demo of Dashboard







Conclusion

USGS turbidity data is sparse and not easily accessible.

The turbidity dashboard demonstrated its high efficiency in time series analysis and forecasting available data.

Future work might focus on expanding the dashboard's functions to include additional water quality parameters.

Challenges

1. USGS Data retrieval
    - ❑ limited, sparse data
    - ❑ inconsistencies in records
    - ❑ not easily accessible

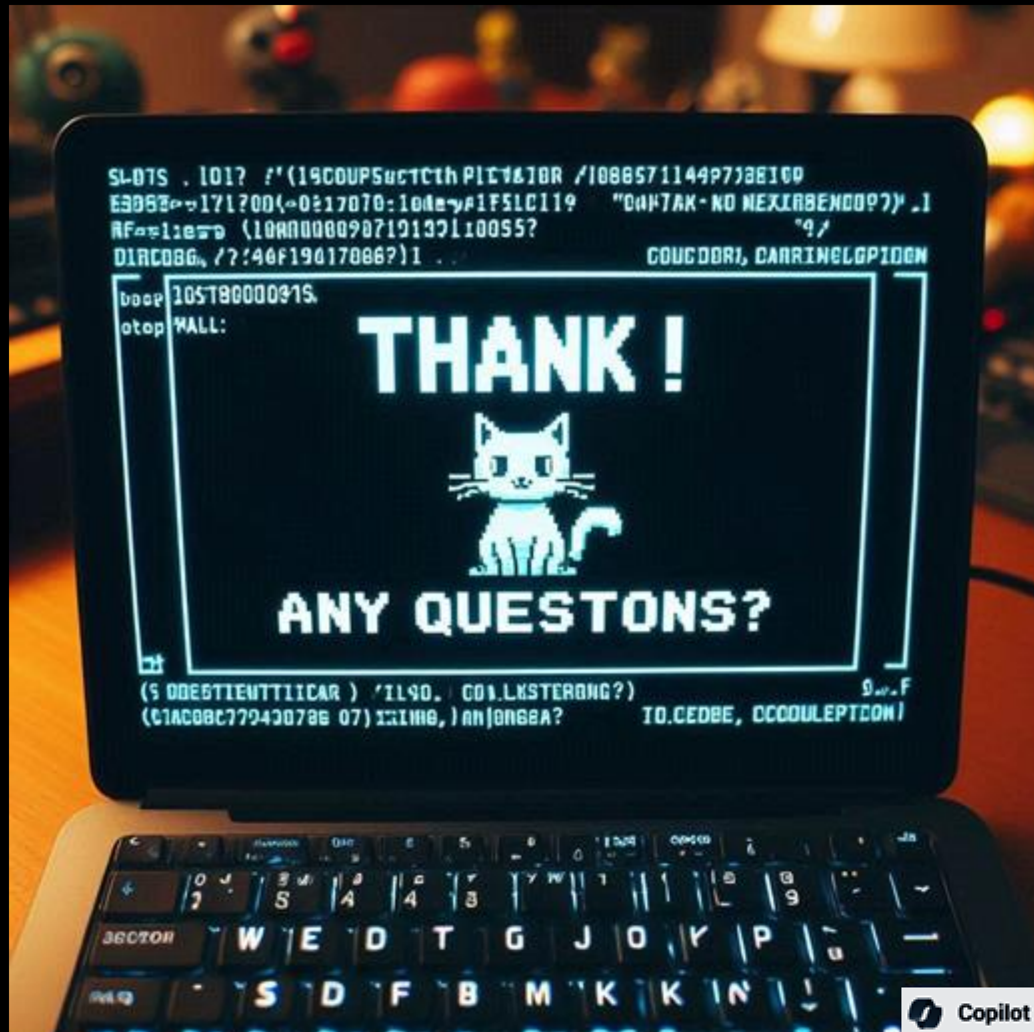
1. HydroLang
    - ❑ Difficulties due to documentation:
        - ❑ limited data sources
        - ❑ only partially extracts data (some stations are not available)

References:

- Aulenbach, D.B., 1967. Water--Our Second Most Important Natural Resource. BC Indus. & Com. L. Rev., 9, p.535.
- Boyd, C.E., 2019. Water quality: an introduction. Springer Nature.
- Omer, N.H., 2019. Water quality parameters. Water quality-science, assessments and policy, 18, pp.1-34.
- Sampedro, Ó. and Salgueiro, J.R., 2015. Turbidimeter and RGB sensor for remote measurements in an aquatic medium. Measurement, 68, pp.128-134.
- Kitchener BG, Wainwright J, Parsons AJ. A review of the principles of turbidity measurement. Progress in Physical Geography. 2017 Oct;41(5):620-42.
- Meyer, A.M., Klein, C., Fünfrocken, E., Kautenburger, R. and Beck, H.P., 2019. Real-time monitoring of water quality to identify pollution pathways in small and middle scale rivers. Science of the Total Environment, 651, pp.2323-2333.
- De Roos, A.J., Gurian, P.L., Robinson, L.F., Rai, A., Zakeri, I. and Kondo, M.C., 2017. Review of epidemiological studies of drinking-water turbidity in relation to acute gastrointestinal illness. Environmental health perspectives, 125(8), p.086003.
- Aram, S., Rivero, M. H., Pahuja, N. K., Sadeghian, R., Paulino, J. L., Meyer, M., & Shallenberger, J., 2020. Multi-Environmental Parameters Dashboard for Susquehanna River Basin using Machine Learning techniques. In 2020 International Conference on Computational Science and Computational Intelligence (CSCI) IEEE. p. 697-700.
- Zainuddin, Z.Q.M., Yahya, F., Gubin Moung, E., Mohd Fazli, B., Abdullah, M.F., 2023. Effective dashboards for urban water security monitoring and evaluation. Int. J. Electr. Comput. Eng. IJECE 13, 4291. https://doi.org/10.11591/ijece.v13i4.pp4291-4305
- Monschein, C., Layman, L., 2024. SoutheastCon 2024 Designing, Implementing, and Evaluating a Municipal Water Quality Dashboard, in: SoutheastCon 2024. Presented at the SoutheastCon 2024, pp. 113–118. https://doi.org/10.1109/SoutheastCon52093.2024.10500297
- Portland Water Bureau (2023) 2023 Drinking Water Quality Report. Available at: https://www.portland.gov/water/water-quality/2023-drinking-water-quality-report (Accessed: 1 August 2024).
- Erazo Ramirez, C., Sermet, Y., Molkenthin, F., Demir, I., 2022. HydroLang: An open-source web-based programming framework for hydrological sciences. Environ. Model. Softw. 157, 105525. https://doi.org/10.1016/j.envsoft.2022.105525
- Erazo Ramirez, C., Sermet, Y., Demir, I., 2023. HydroLang Markup Language: Community-driven web components for hydrological analyses. J. Hydroinformatics 25, 1171–1187. https://doi.org/10.2166/hydro.2023.149
- Hirsch, R.M., Archfield, S.A. and De Cicco, L.A., 2015. A bootstrap method for estimating uncertainty of water quality trends. Environmental Modelling & Software, 73, pp.148-166. https://doi.org/10.1016/j.envsoft.2015.07.017
- Srebotnjak, T., Carr, G., de Sherbinin, A., & Rickwood, C. (2012). A global Water Quality Index and hot-deck imputation of missing data. Ecological Indicators, 17, 108–119. https://doi.org/10.1016/j.ecolind.2011.04.023
- Read, E. K., Carr, L., De Cicco, L., Dugan, H. A., Hanson, P. C., Hart, J. A., Kreft, J., Read, J. S., & Winslow, L. A. (2017). Water quality data for national-scale aquatic research: The Water Quality Portal. Water Resources Research, 53(2), 1735–1745. https://doi.org/10.1002/2016WR019993
- Sprague, L. A., Oelsner, G. P., & Argue, D. M. (2017). Challenges with secondary use of multi-source water-quality data in the United States. Water Research, 110, 252–261. https://doi.org/10.1016/j.watres.2016.12.024
- Reddy, B.K., Ayyagari, K.S., Medam, R.R. and Alhaider, M., 2023. Application of Machine Learning Techniques in Modern Hybrid Power Systems–A Case Study. In IoT, Machine Learning and Blockchain Technologies for Renewable Energy and Modern Hybrid Power Systems (pp. 173-204). River Publishers. https://doi.org/10.1109/CONIT59222.2023.10205638
- Chen, T. and Guestrin, C., 2016, August. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 785-794). https://doi.org/10.1145/2939672.2939785
- Oregon’s extreme weather in 2022 included April snow and a record hot October. The Oregonian. Advance Local Media LLC. Web portal. Date accessed: 8/1/024. URL: https://www.oregonlive.com/weather/2022/12/oregons-extreme-weather-in-2022-included-april-snow-and-a-record-hot-october.html

Acknowledgements

- WaterSoftHack ‘24 - organizers & resource persons
- Team-Turbidity members


THANK !
ANY QUESTONS?
Copilot