

Decision Support via Machine Learning

A Way For Utilities to Do More with the Resources at Hand

4 Nov 2025 (NWWC 2025)

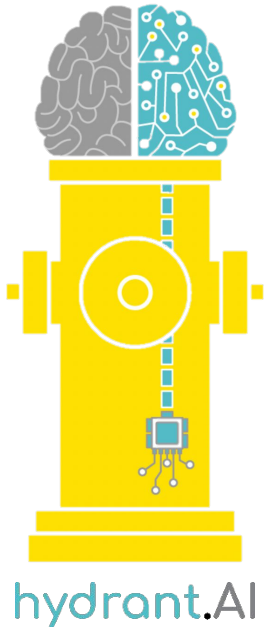
Stan Fong, PhD

CTO @ Digital Water Solutions



Digital Water Solutions (DWS)

DWS brings together experience in artificial intelligence, applied research, software/hardware development and water distribution system know-how to deliver novel solutions to water utilities.

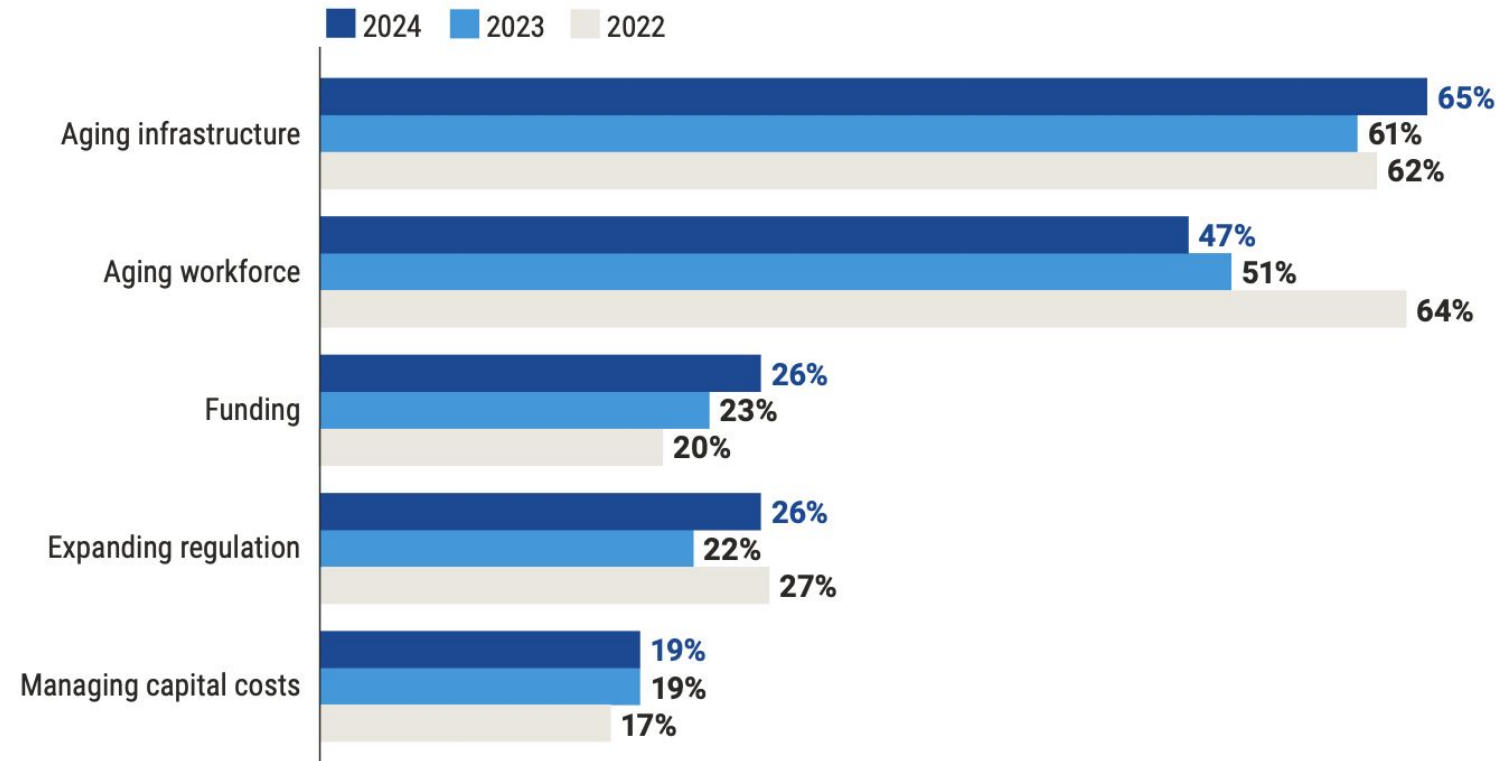


The Case For Data-Informed Decisions

Figure 1

What are the most challenging issues facing the water, wastewater and stormwater industry today? (Select up to three issues)

Source: Black & Veatch



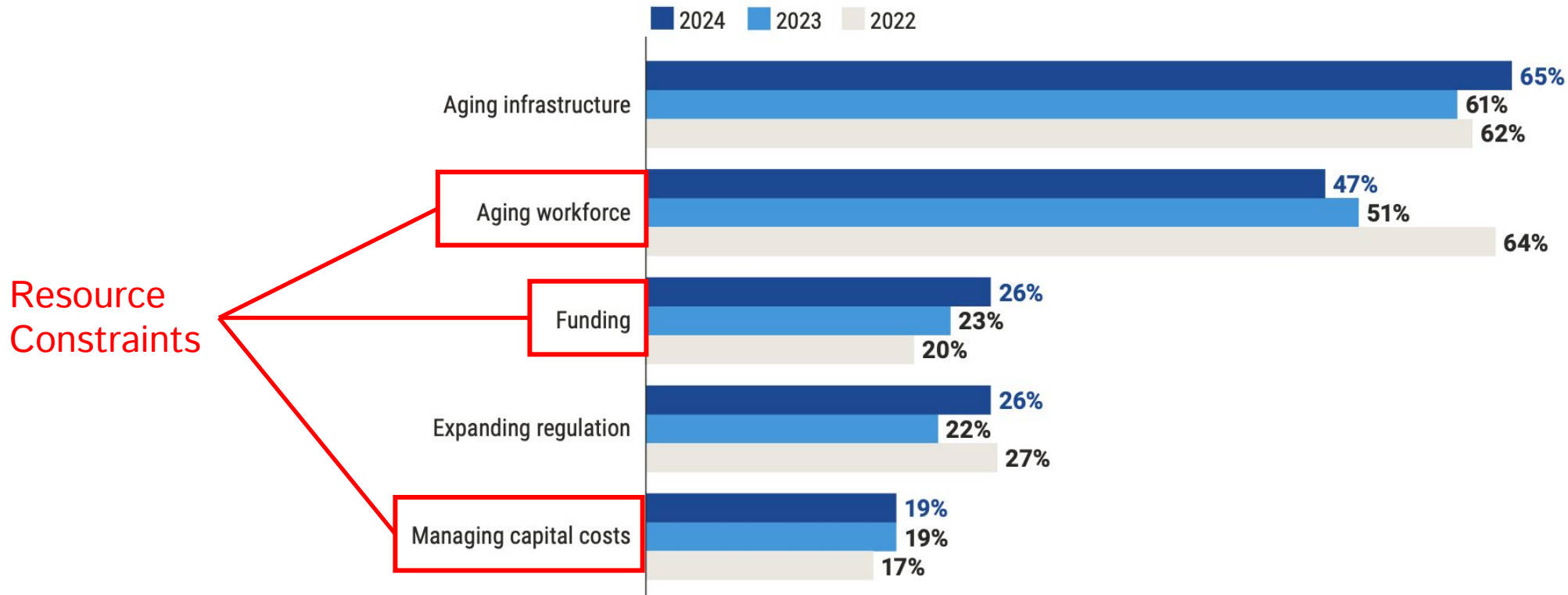
(Black & Veatch 2024)

The Case For Data-Informed Decisions

Figure 1

What are the most challenging issues facing the water, wastewater and stormwater industry today? (Select up to three issues)

Source: Black & Veatch



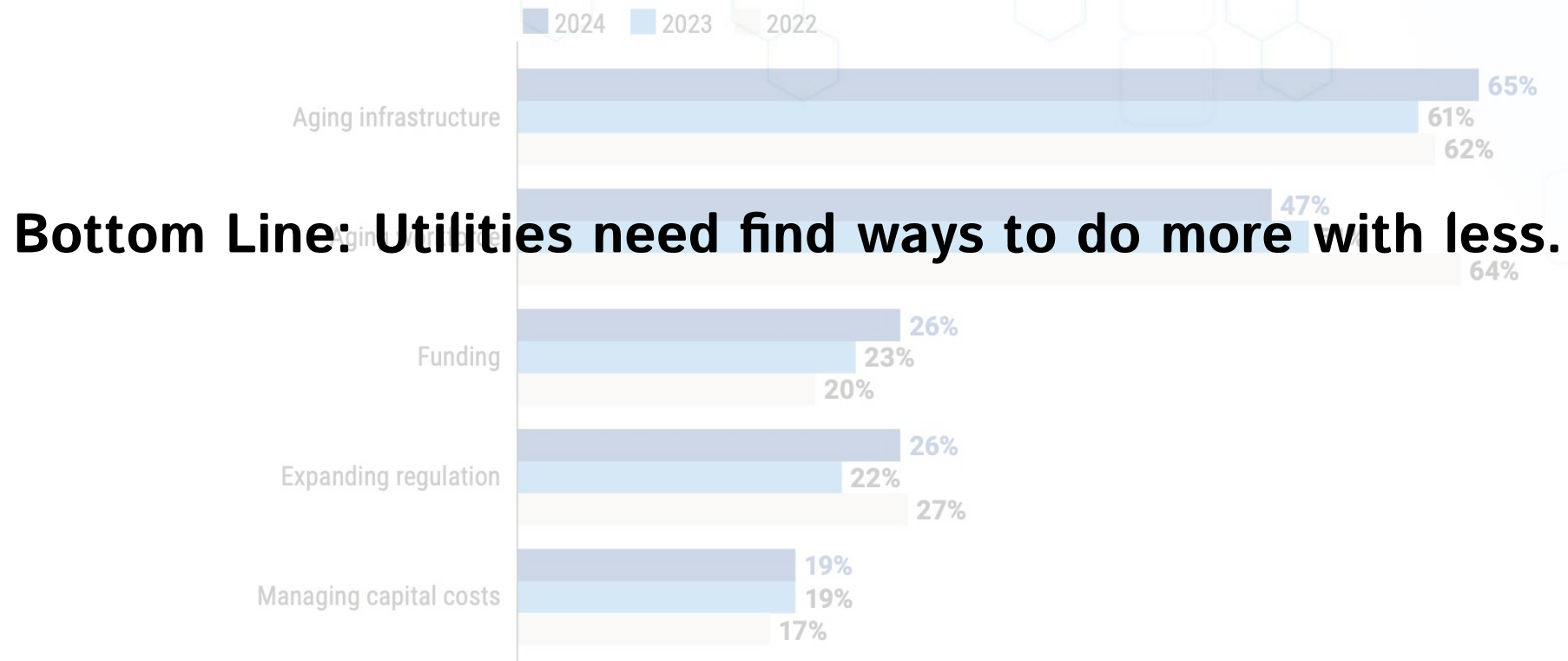
(Black & Veatch 2024)

The Case For Data-Informed Decisions

Figure 1

What are the most challenging issues facing the water, wastewater and stormwater industry today? (Select up to three issues)

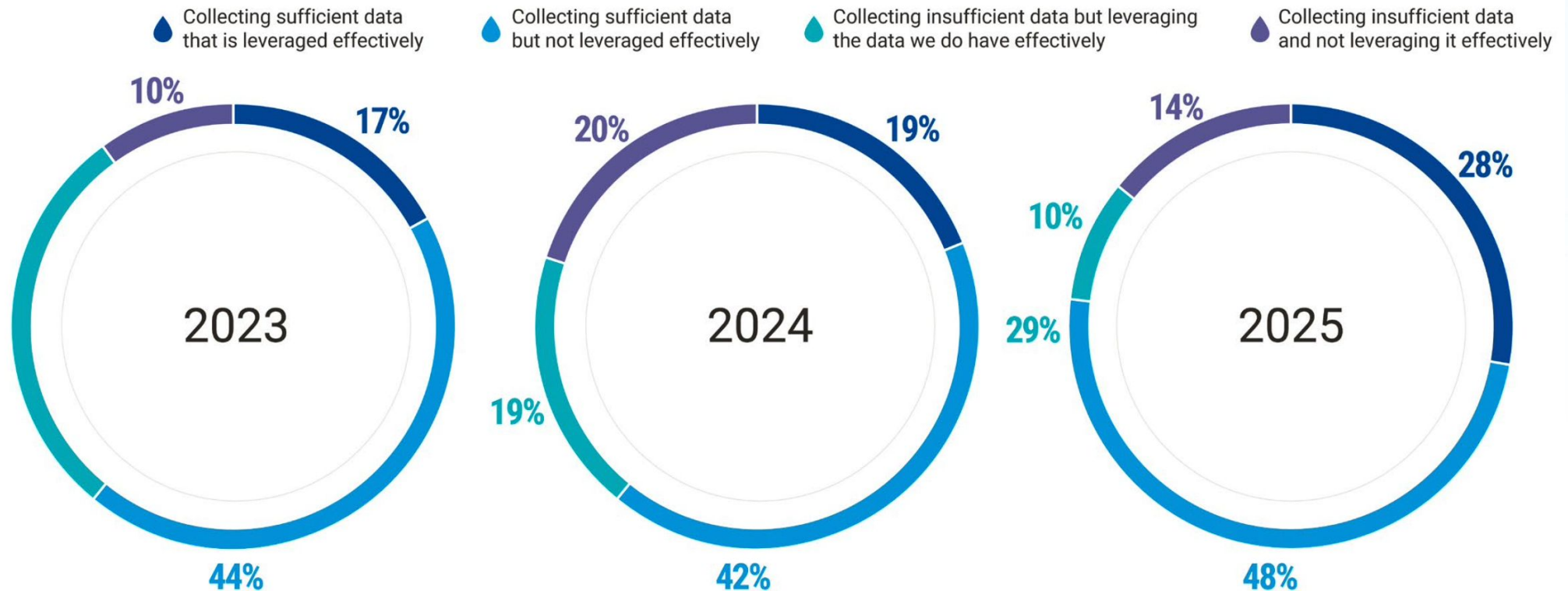
Source: Black & Veatch



(Black & Veatch 2024)

Taking Stock Of What's Available

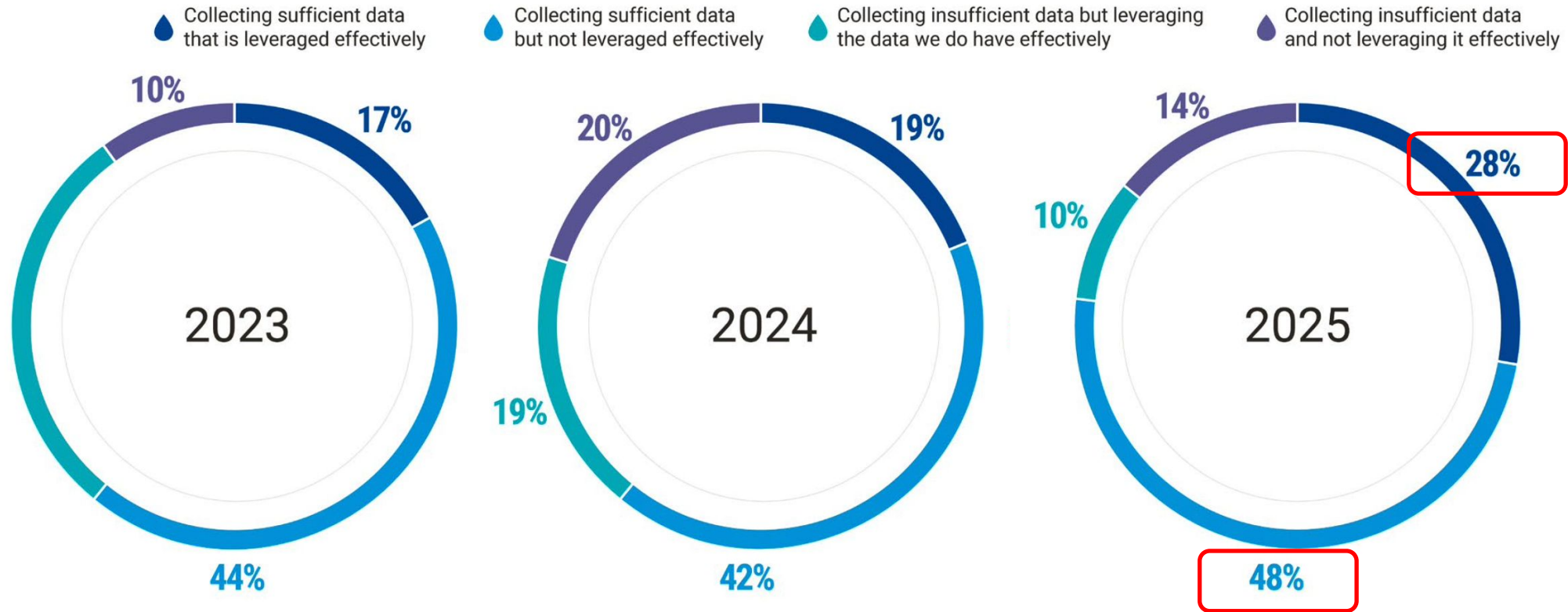
Source: Black & Veatch 2025 Water Report survey



For Most Utilities, The Amount Of Data Is Not A Problem

In 2025, >75% of utilities are collecting sufficient data

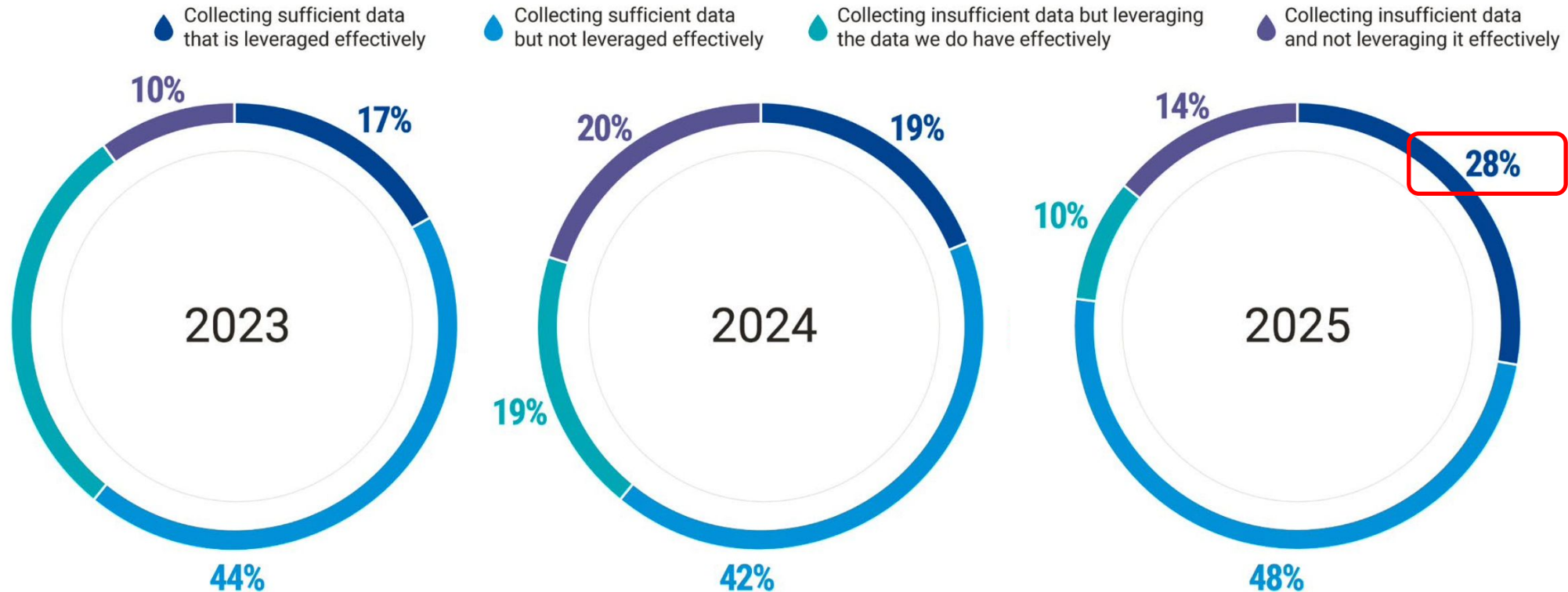
Source: Black & Veatch 2025 Water Report survey



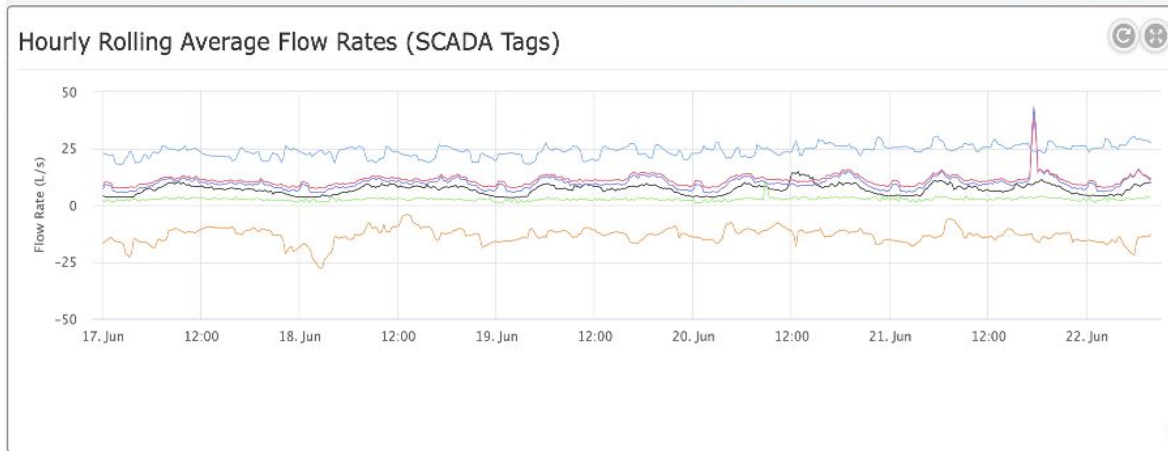
But Rather, Utilization Is The Concern

In 2025, <30% of utilities feel that they are leveraging their data effectively.

Source: Black & Veatch 2025 Water Report survey



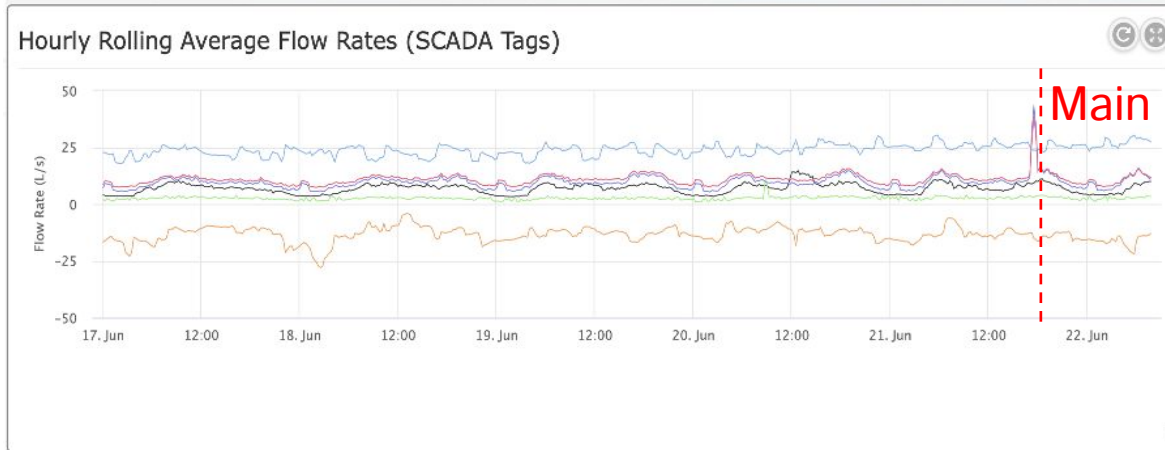
So How Can We Get More Out Of Existing Data?



What's happening underground?

So How Can We Get More Out Of Existing Data?

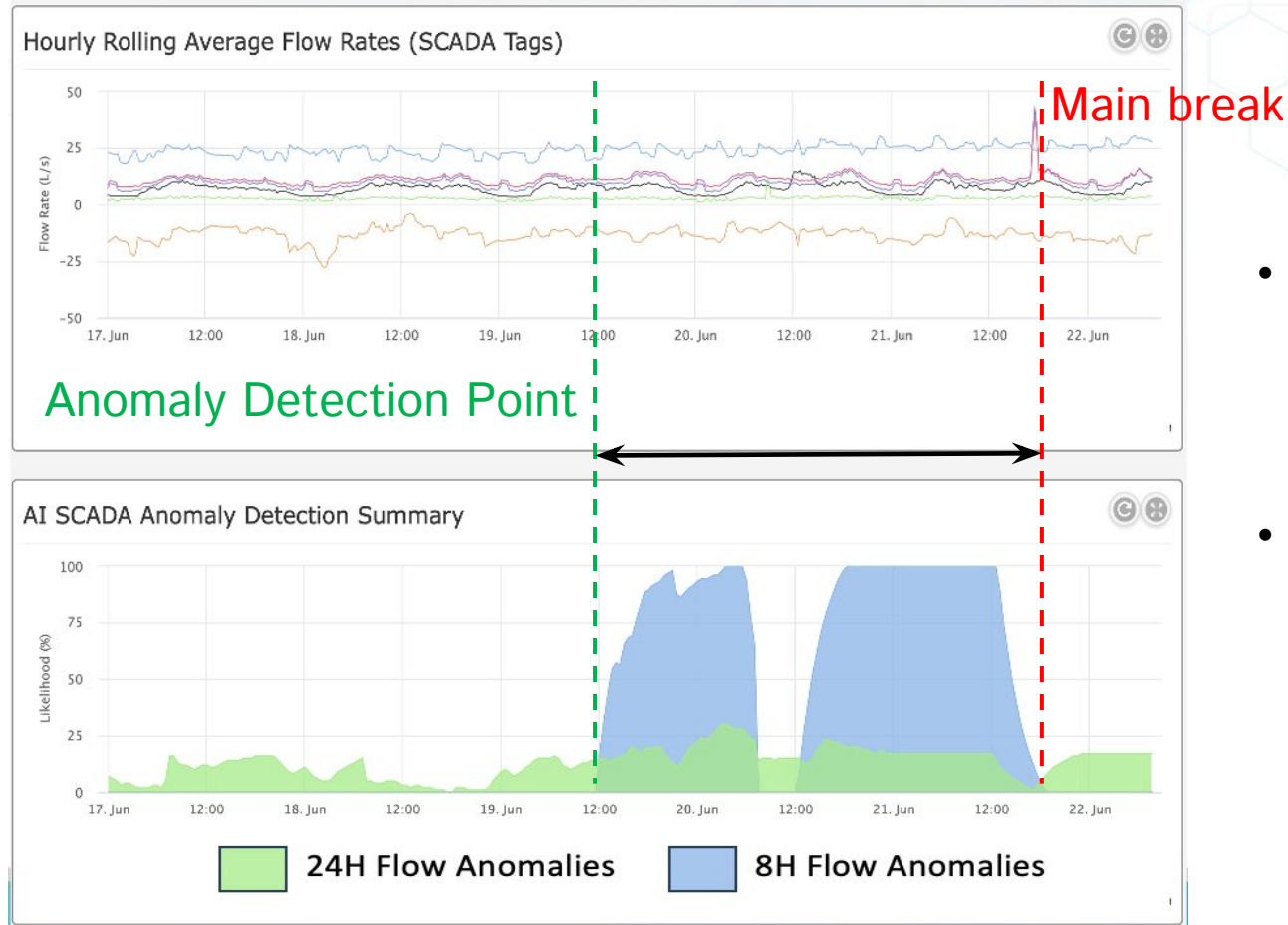
Anomaly Detection



Main break

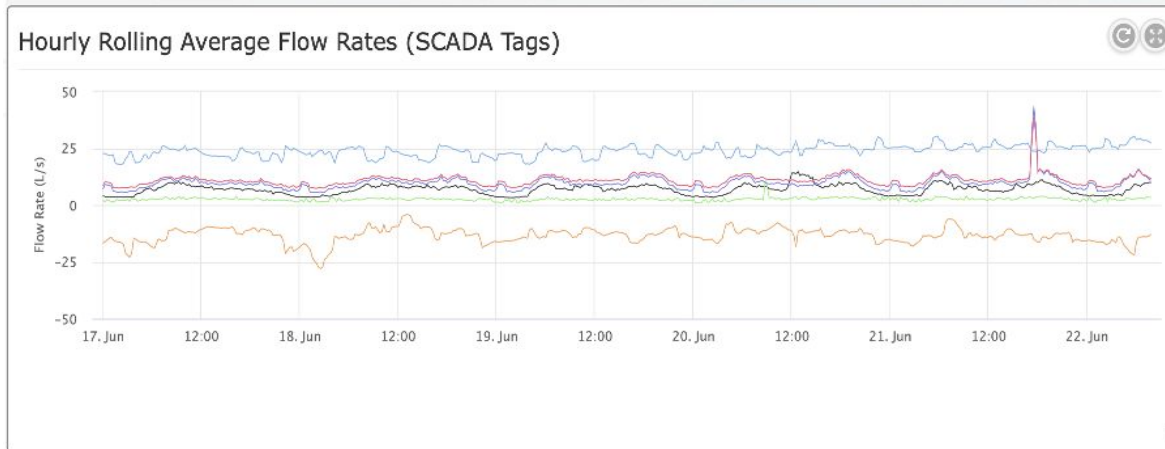
So How Can We Get More Out Of Existing Data?

Anomaly Detection



- With an anomaly detection model trained on the flow data, the utility could have been made **aware** of an issue **up to 2 days before** the main **break** occurs.
- Using ML, we can apply this type of monitoring to any existing time series data.

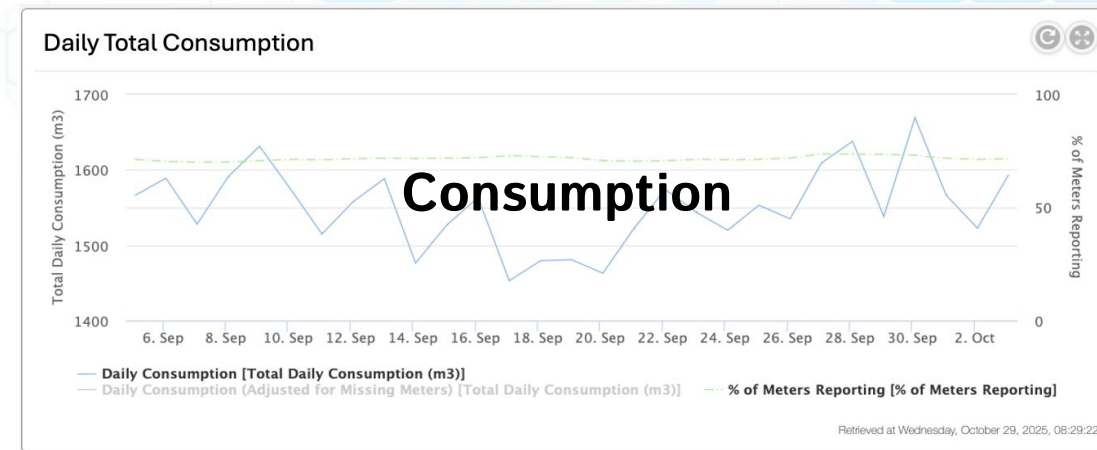
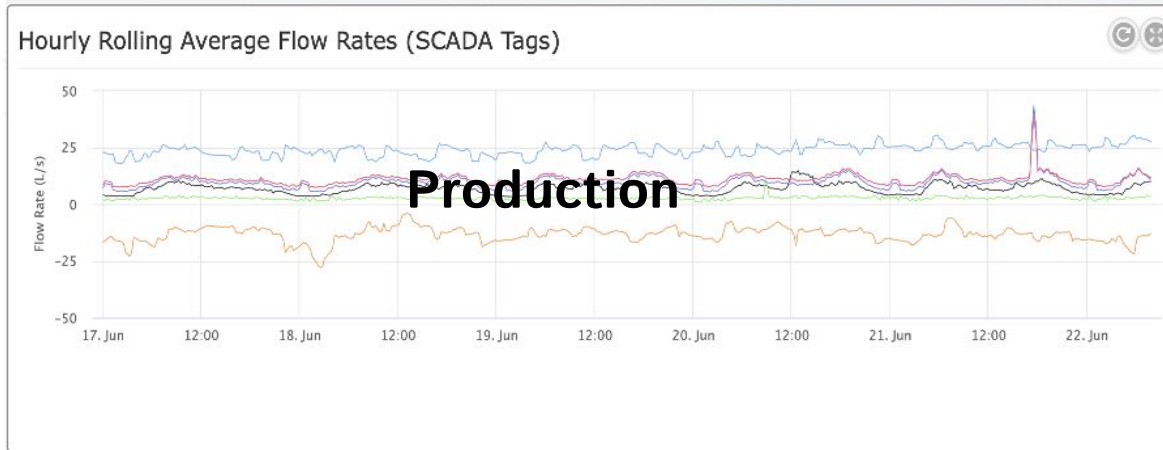
So How Can We Get More Out Of Existing Data?



What's happening underground?

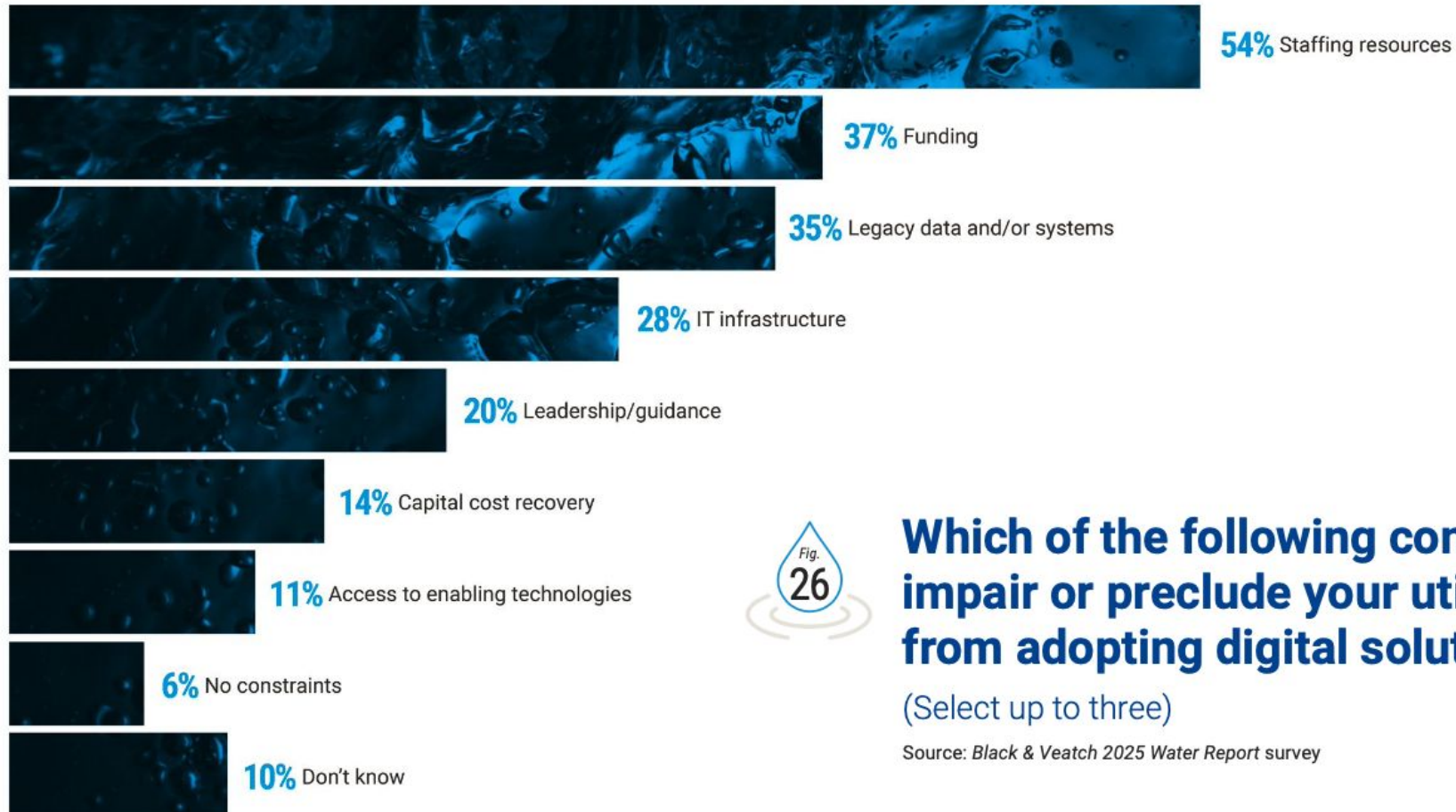
So How Can We Get More Out Of Existing Data?

M36 Water Balance



- Use DS/ML to convert static billing data into real-time data ☐ **estimate NRW in real-time**
- Use ML to detect and/or fill data gaps ☐ **improve billing accuracy**
- Automate billing calculations ☐ **improve efficiency**

What's Standing in the Way?



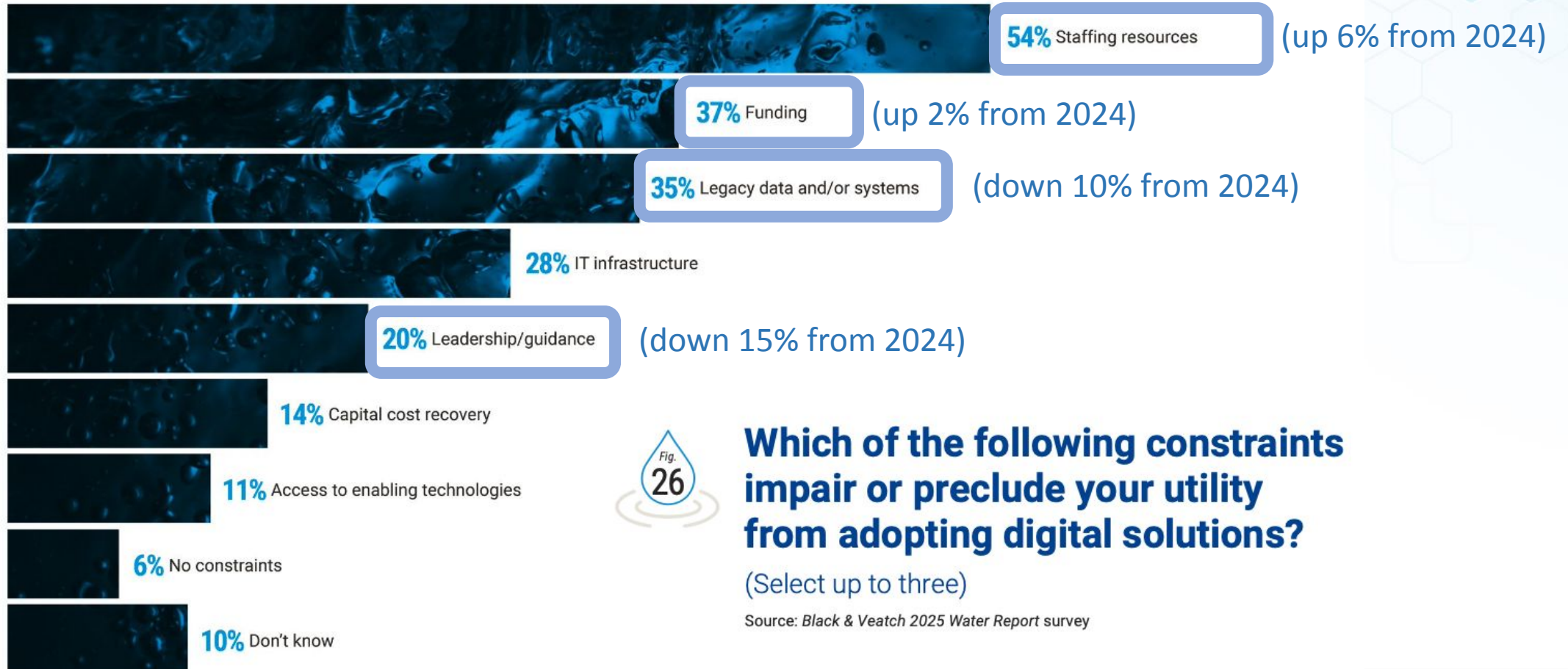
Which of the following constraints impair or preclude your utility from adopting digital solutions?

(Select up to three)

Source: Black & Veatch 2025 Water Report survey

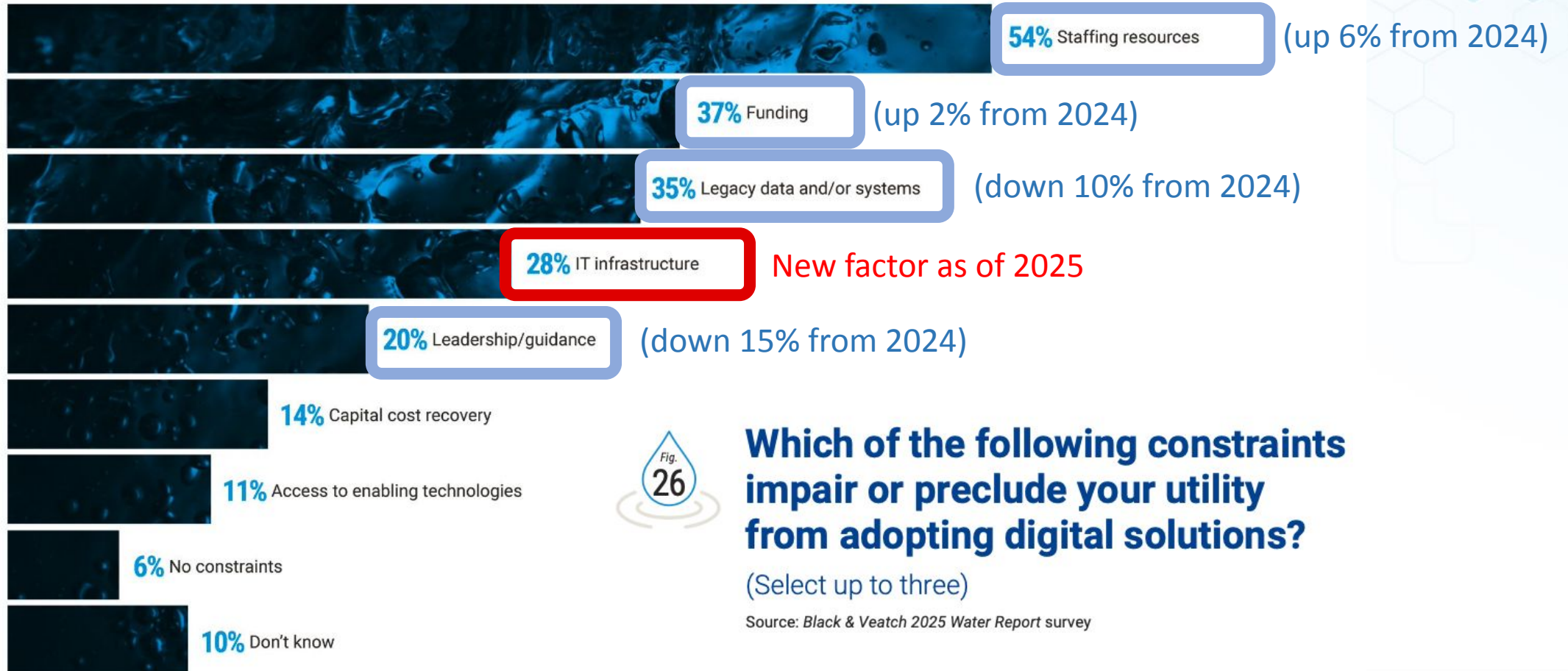
What's Standing in the Way?

How does 2025 compare to 2024?



What's Standing in the Way?

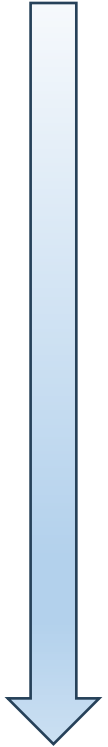
How does 2025 compare to 2024?



From Data to Decision Support

Step

Data



Data Collection

Data Aggregation

Data Science &
Machine Learning

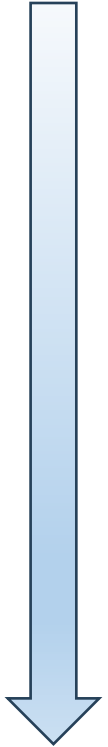
Decision
Support

From Data to Decision Support

Step

Key Questions

Data



Data Collection

- What types of data are available to me?
 - Sensors, SCADA, billing, GIS, etc.
- Am I missing any data that is quintessential to my role?

Data Aggregation

Data Science &
Machine Learning

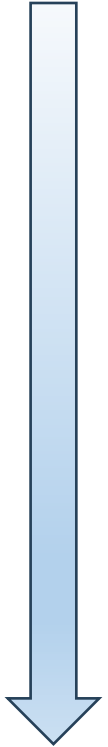
Decision
Support

From Data to Decision Support

Step

Key Questions

Data



Data Collection

- What types of data are available to me?
 - Sensors, SCADA, billing, GIS, etc.
- Am I missing any data that is quintessential to my role?

Data Aggregation

- Where is all this data stored; do I have any data silos?
- Options for exporting data (APIs, CSV)?
- Can I access or export data in real-time?
- Do I have a point or platform to centralize my data?
- Potential IT barriers?

Data Science &
Machine Learning

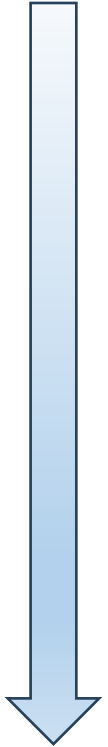
Decision
Support

From Data to Decision Support

Step

Key Questions

Data



Data Collection

- What types of data are available to me?
 - Sensors, SCADA, billing, GIS, etc.
- Am I missing any data that is quintessential to my role?

Data Aggregation

- Where is all this data stored; do I have any data silos?
- Options for exporting data (APIs, CSV)?
- Can I access or export data in real-time?
- Do I have a point or platform to centralize my data?
- Potential IT barriers?

Data Science &
Machine Learning

- What do I want to get out of my data?
- New insights? How reliable is my data?
- Ways to improve existing workflows?

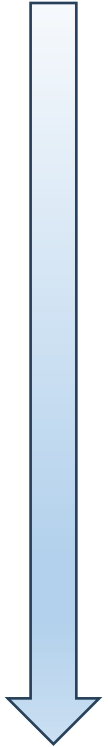
Decision
Support

From Data to Decision Support

Step

Key Questions

Data



Data Collection

- What types of data are available to me?
 - Sensors, SCADA, billing, GIS, etc.
- Am I missing any data that is quintessential to my role?

Data Aggregation

- Where is all this data stored; do I have any data silos?
- Options for exporting data (APIs, CSV)? Extra costs?
- Can I access data in real-time?
- Do I have a point or platform to centralize my data?
- Potential IT barriers?

Data Science &
Machine Learning

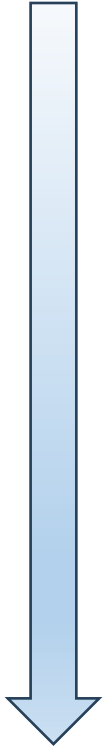
- What do I want to get out of my data?
- New insights? How reliable is my data?
- Ways to improve existing workflows?

Decision
Support

From Data to Decision Support

Step

Data



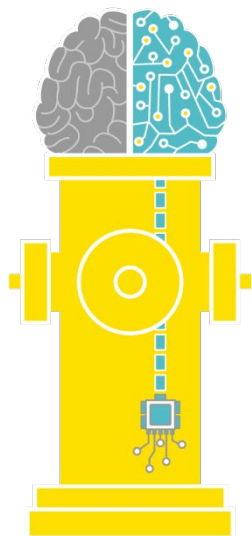
Data Collection

Data Aggregation

Data Science &
Machine Learning

In practice, the most difficult challenges appear during the data aggregation step (can be **up to 80% of the total project time**).

Decision
Support



Case Study

What's happening underground?

Using Existing Data to Better Understand NRW

Utility Details

- 1 DMA (approx. 7000 services)
- Closed pressure system
- 1 Pumping station with 2 pumps
 - SCADA total cumulative flow every 10 mins
- Services are equipped with AMR
- **Utility believes they have an NRW problem in the range of 20-30%.**
- Limited budget for additional CAPEX
- Limited staffing resources

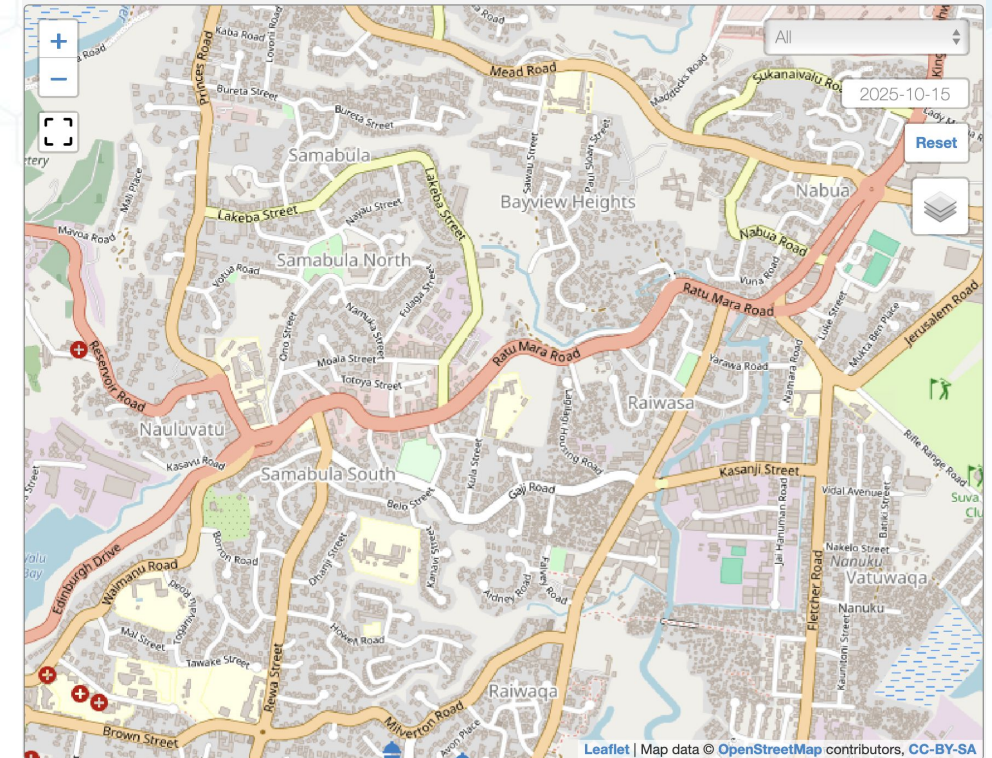
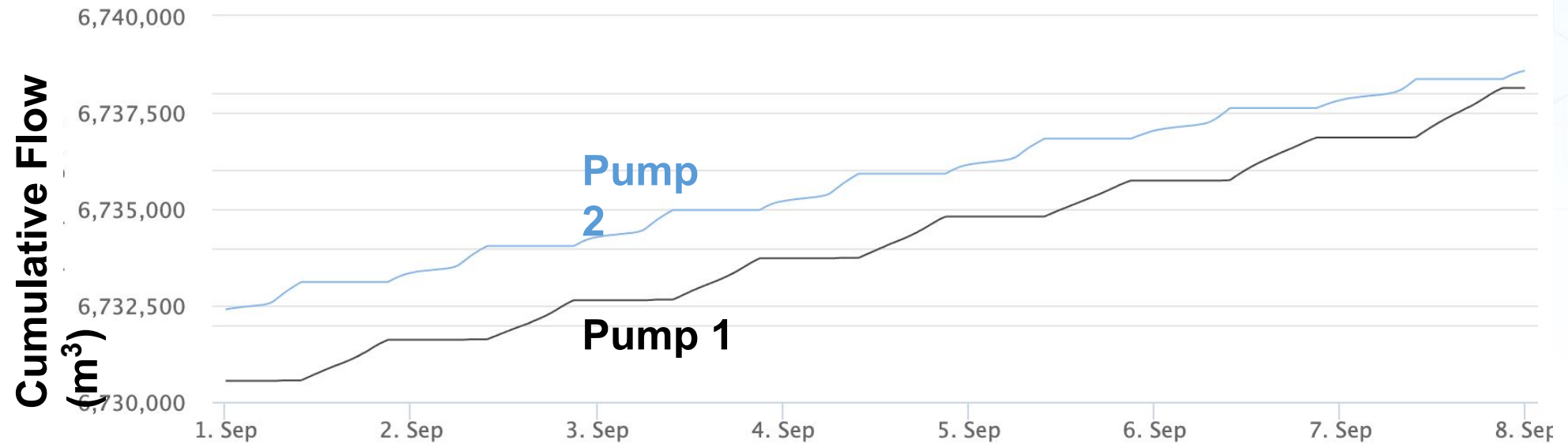


Photo for scale (not the actual utility)

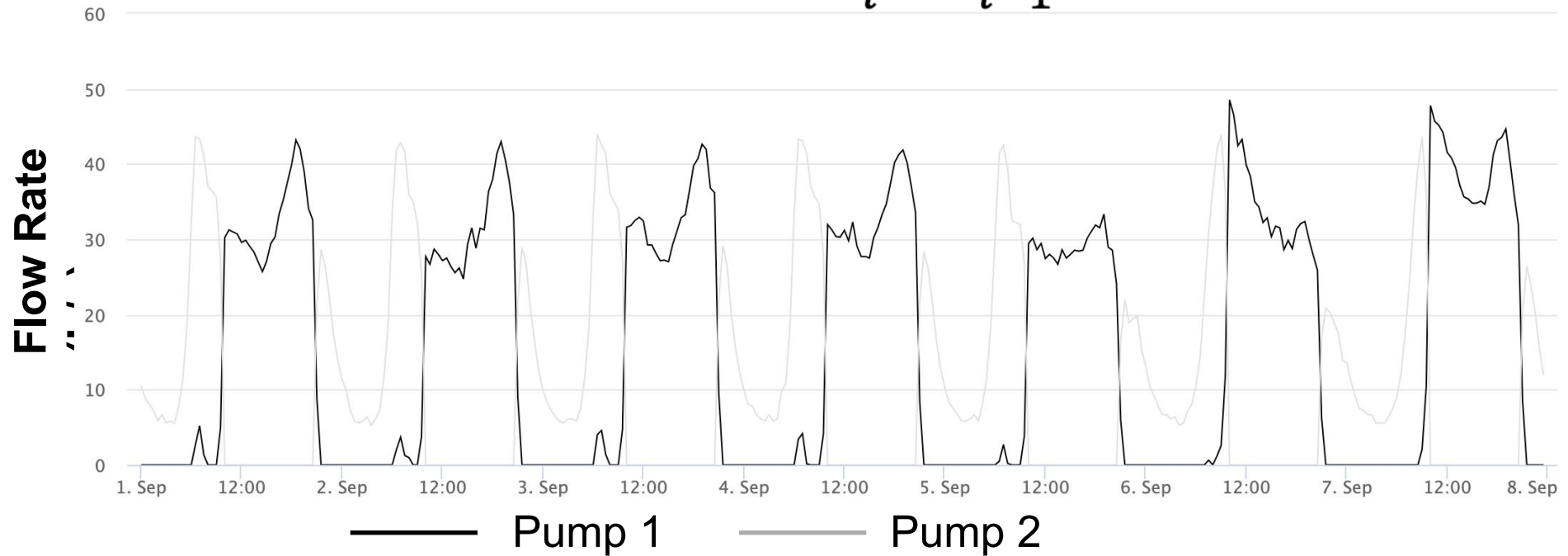
Production: Cumulative Flow Data

SCADA total cumulative flow every 10 mins via an SFTP solution



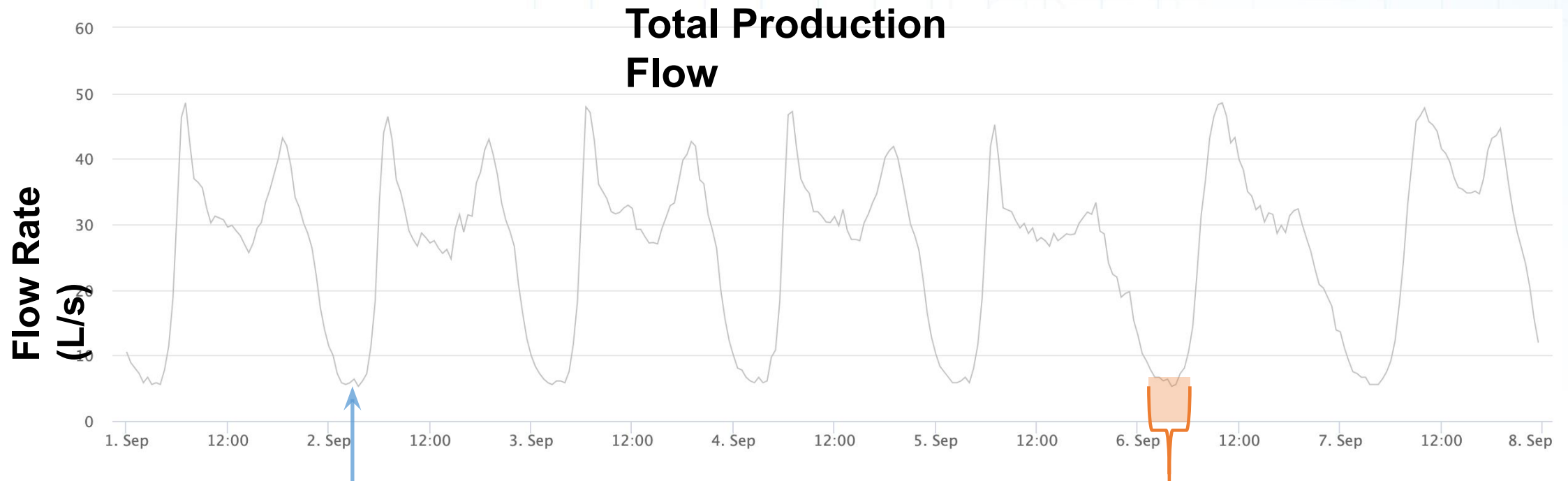
Production: Flow Rates

$$flowRate_i = \frac{totalFlow_i - totalFlow_{i-1}}{t_i - t_{i-1}}$$



Production: Minimum & Total Night Flows

Adding both pumps together gives total production

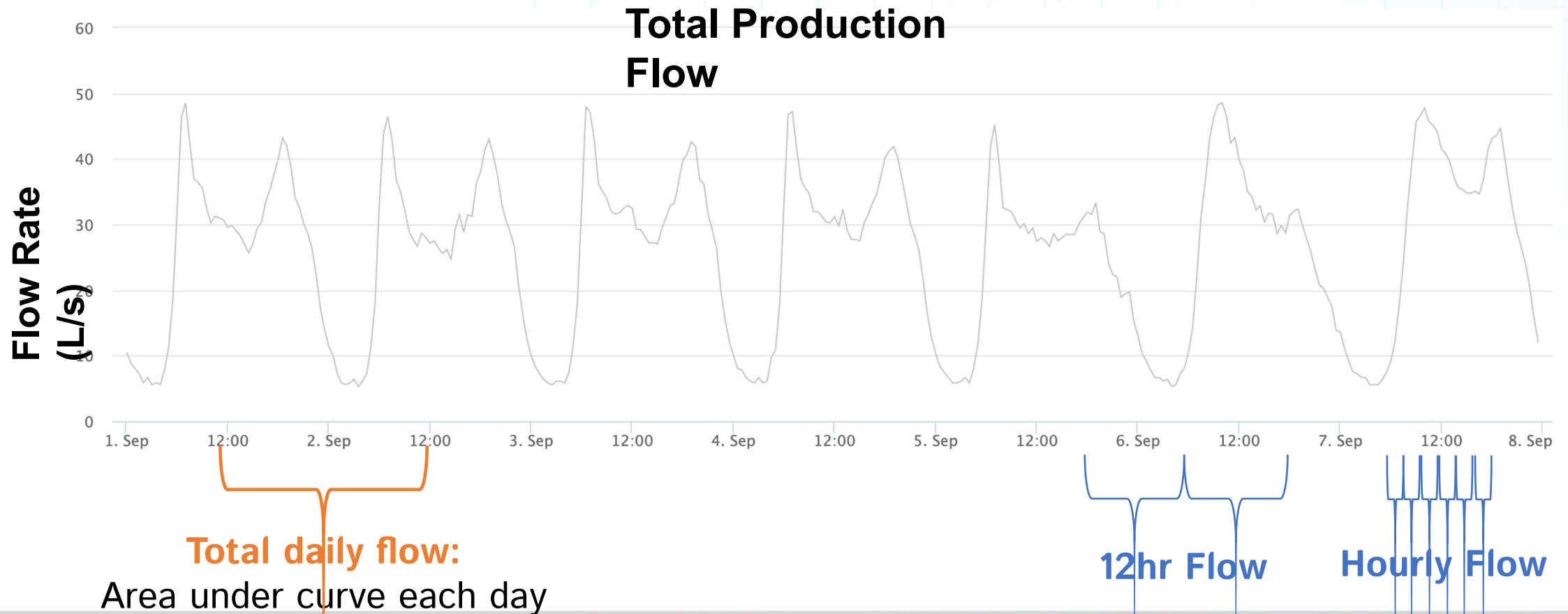


Minimum night flow rate:
minimum value each night

Total night flow:
Area under curve each night

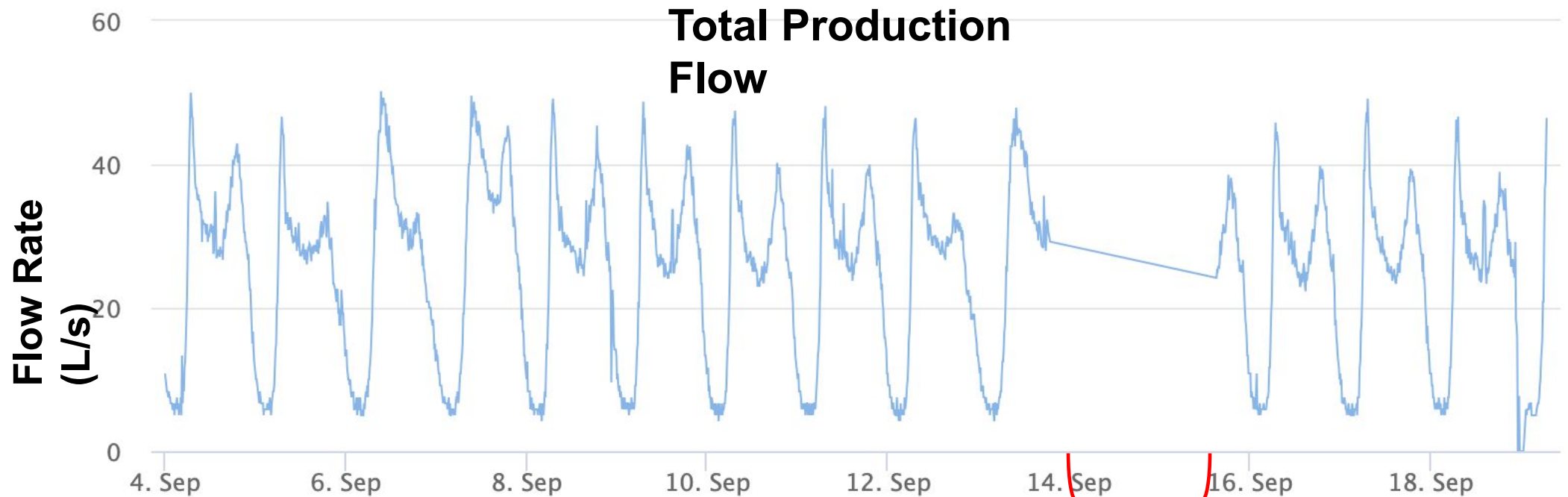
Production: Total Daily, 12H and Hourly Flows

Simple to calculate total flows over different periods...



Production: Dealing With Data Gaps

...but simple calculations can still have their challenges:



How to deal with missing data?

Consumption: Accessing AMR Data

- Utility views AMR data through vendor's dashboard (~7000 meters)
- Dashboard does not allow for export of data in real-time
- API access is available, but at an additional cost
- Limited familiarity with APIs amongst utility staff

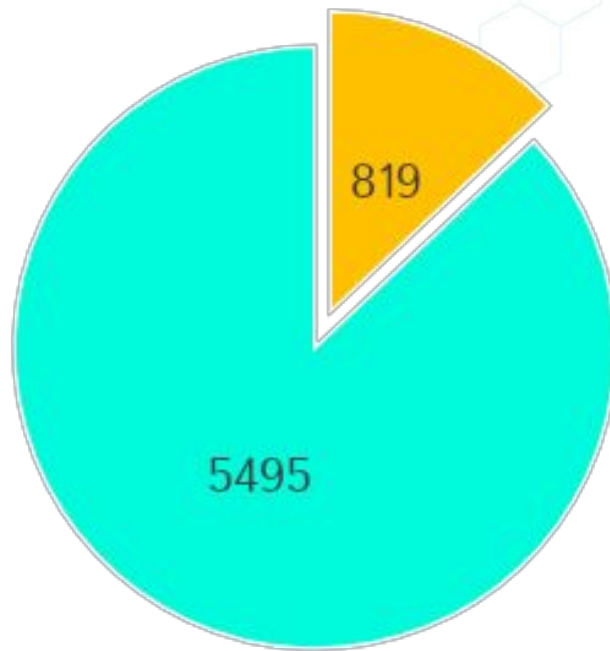


Data Aggregation Challenge

DWS worked with the vendor directly on behalf of the utility to set up necessary access to AMR data.

Consumption: Meter Data Overview

Reporting vs. Non-Reporting Meters (Apr - Jun 2025)

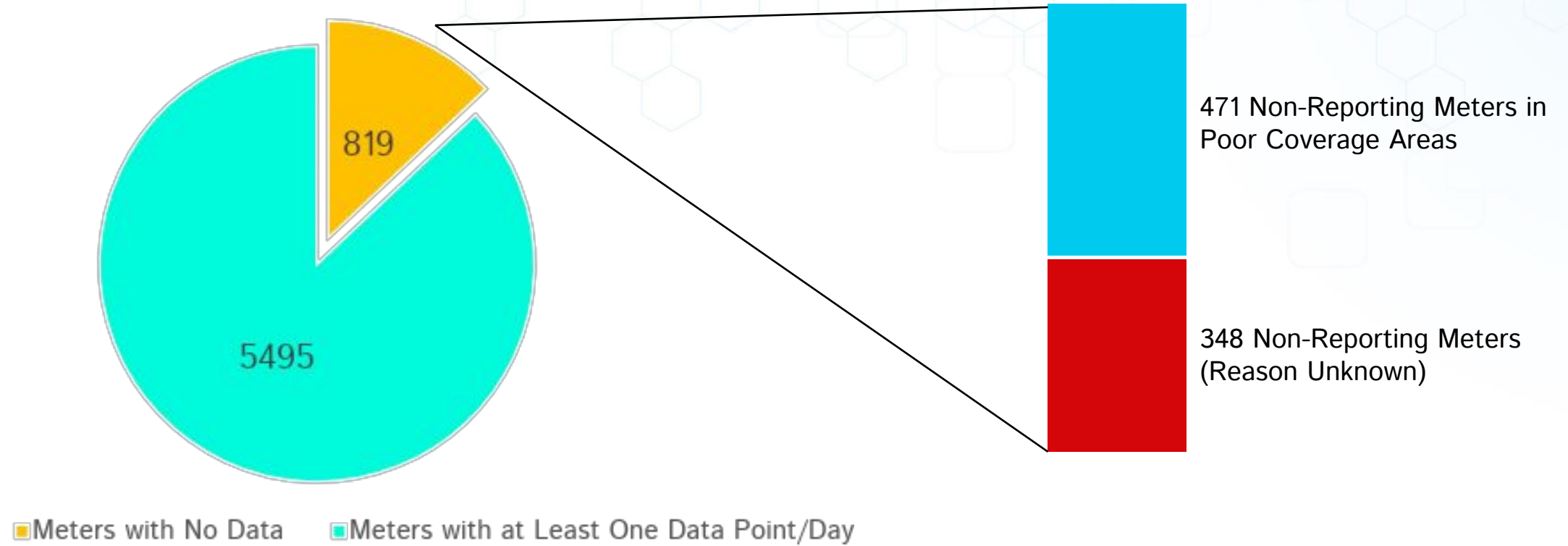


■ Meters with No Data ■ Meters with at Least One Data Point/Day

- Between Apr – Jun 2025, a total of 6314 meters were available to query through the API.
- Inconsistencies in reporting frequency for some meters.

Consumption: Meter Data Overview

Reporting vs. Non-Reporting Meters (Apr - Jun 2025)



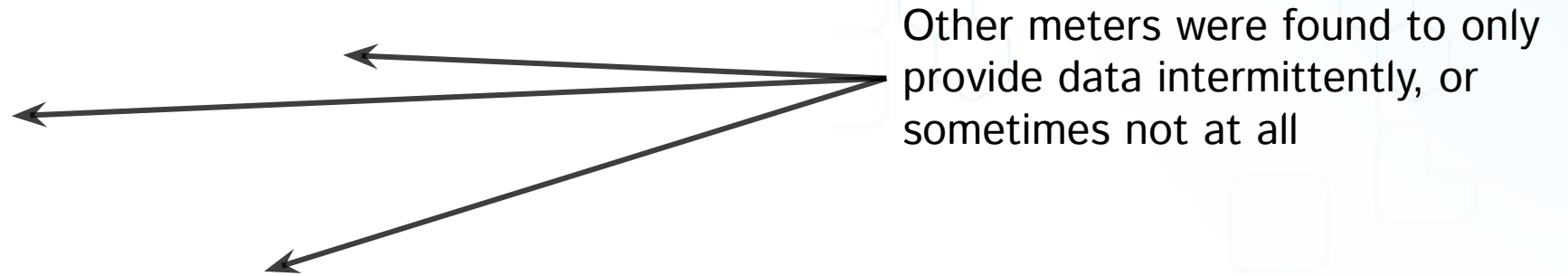
Consumption: Meter Data Reporting

Reported Points/Day

← Some meters consistently provide a new reading each day

Consumption: Meter Data Reporting

Reported Points/Day



Implications On Non-Revenue Water

- On any given day, it is expected that **only 70-74%** of the meters will **report data**.
- Due to the missing meters, the **calculated Total Daily Consumption may not reflect the true total**.

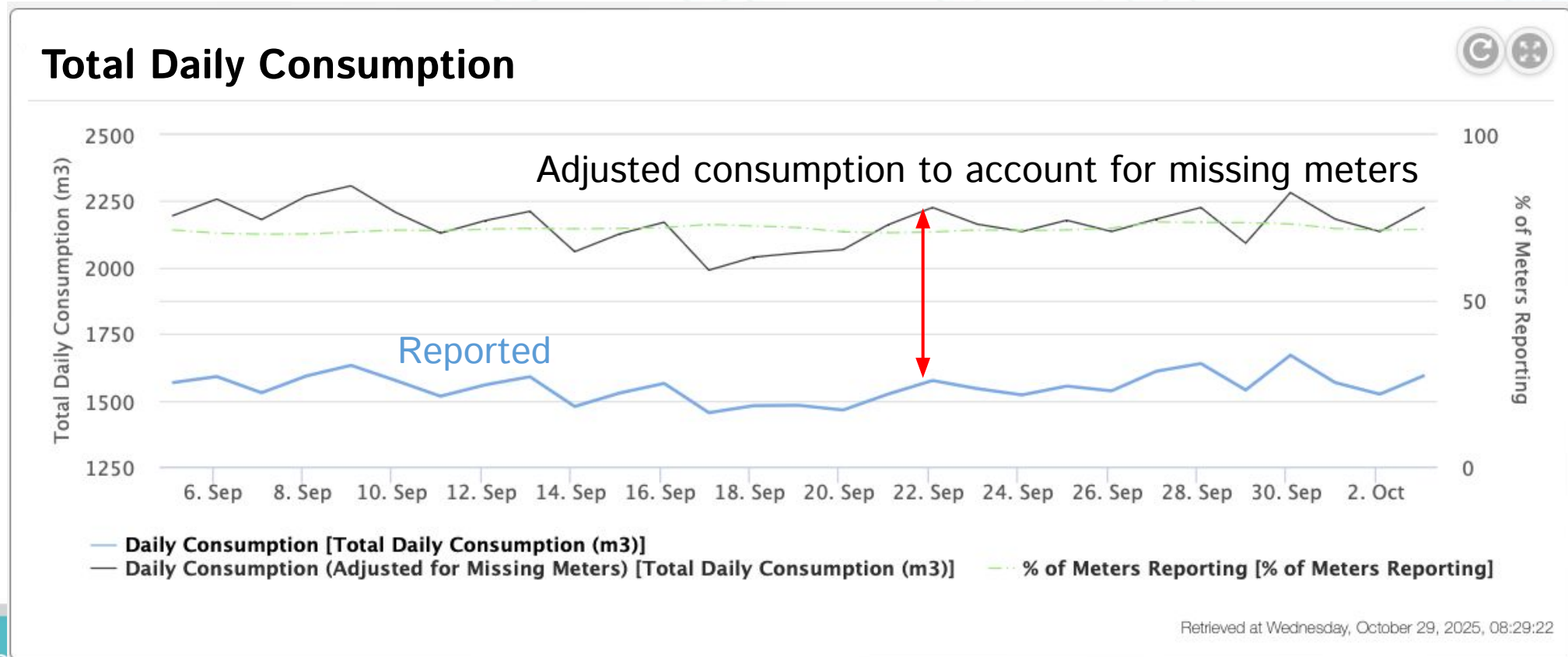
Total Daily Consumption



Retrieved at Wednesday, October 29, 2025, 08:29:22

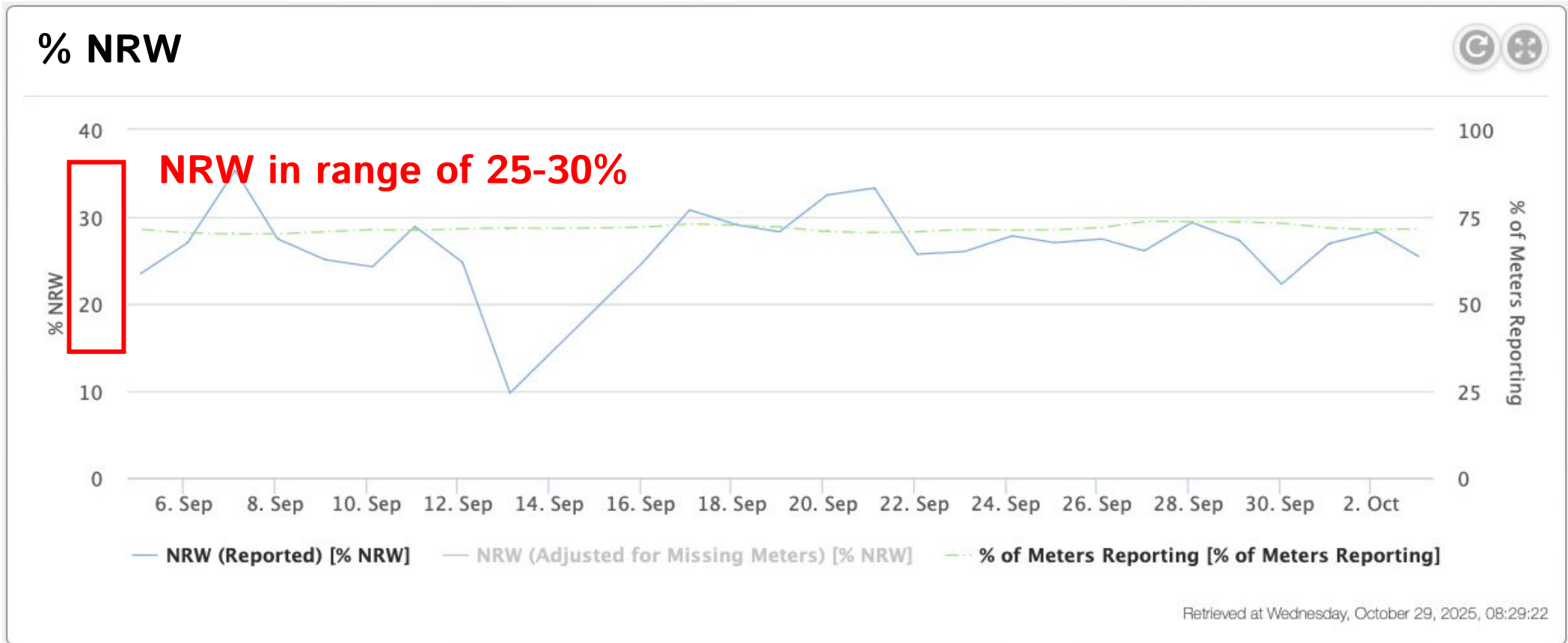
Implications On Non-Revenue Water

- On any given day, it is expected that **only 70-74%** of the meters **will report data**.
- Due to the missing meters, the **calculated Total Daily Consumption may not reflect the true total**.



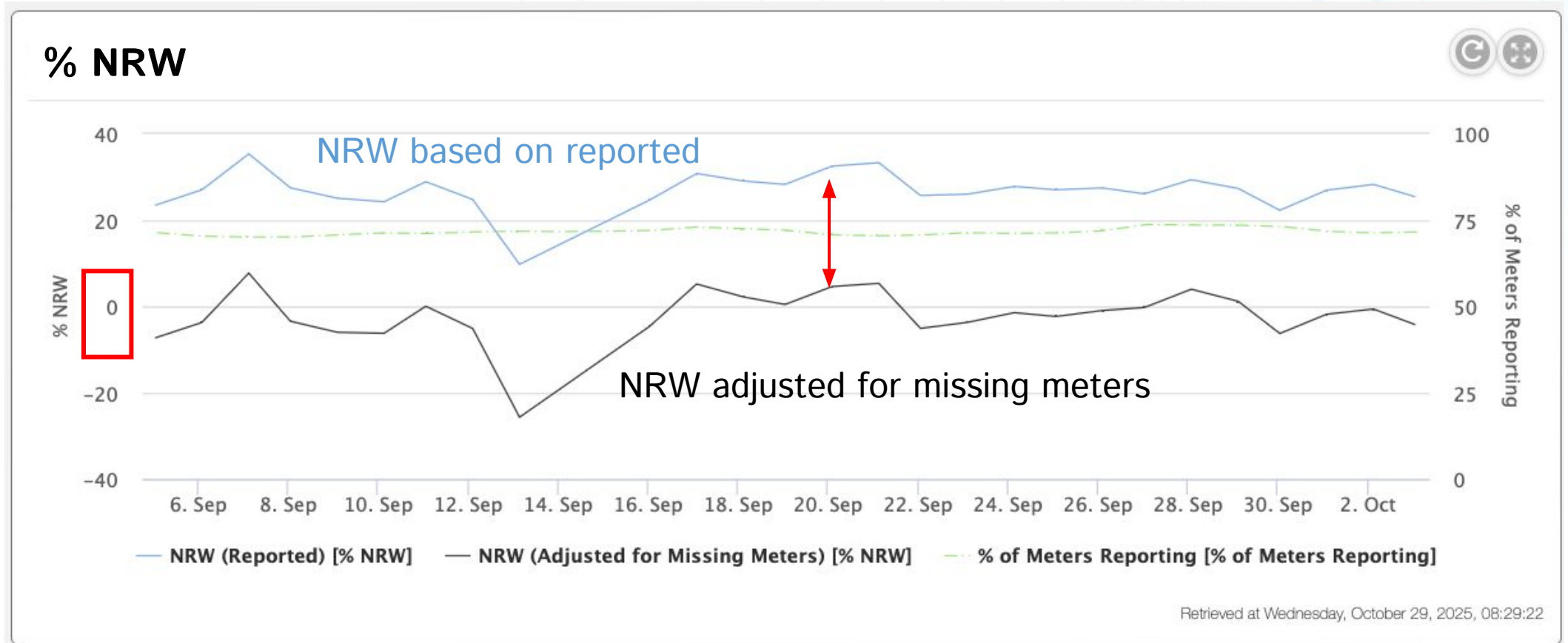
Implications On Non-Revenue Water

NRW calculated based on the reported data:



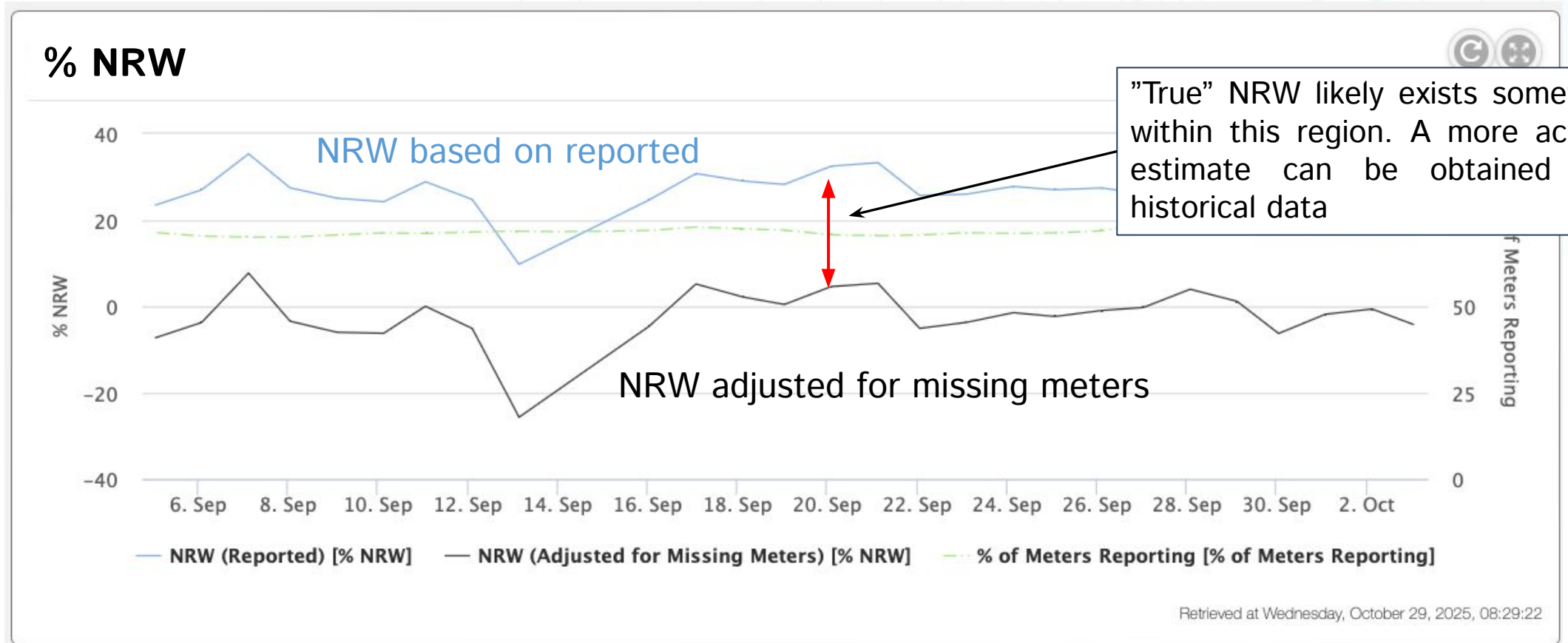
Implications On Non-Revenue Water

NRW after simple adjustment (assuming missing meters have consumption = avg. consumption)



Implications On Non-Revenue Water

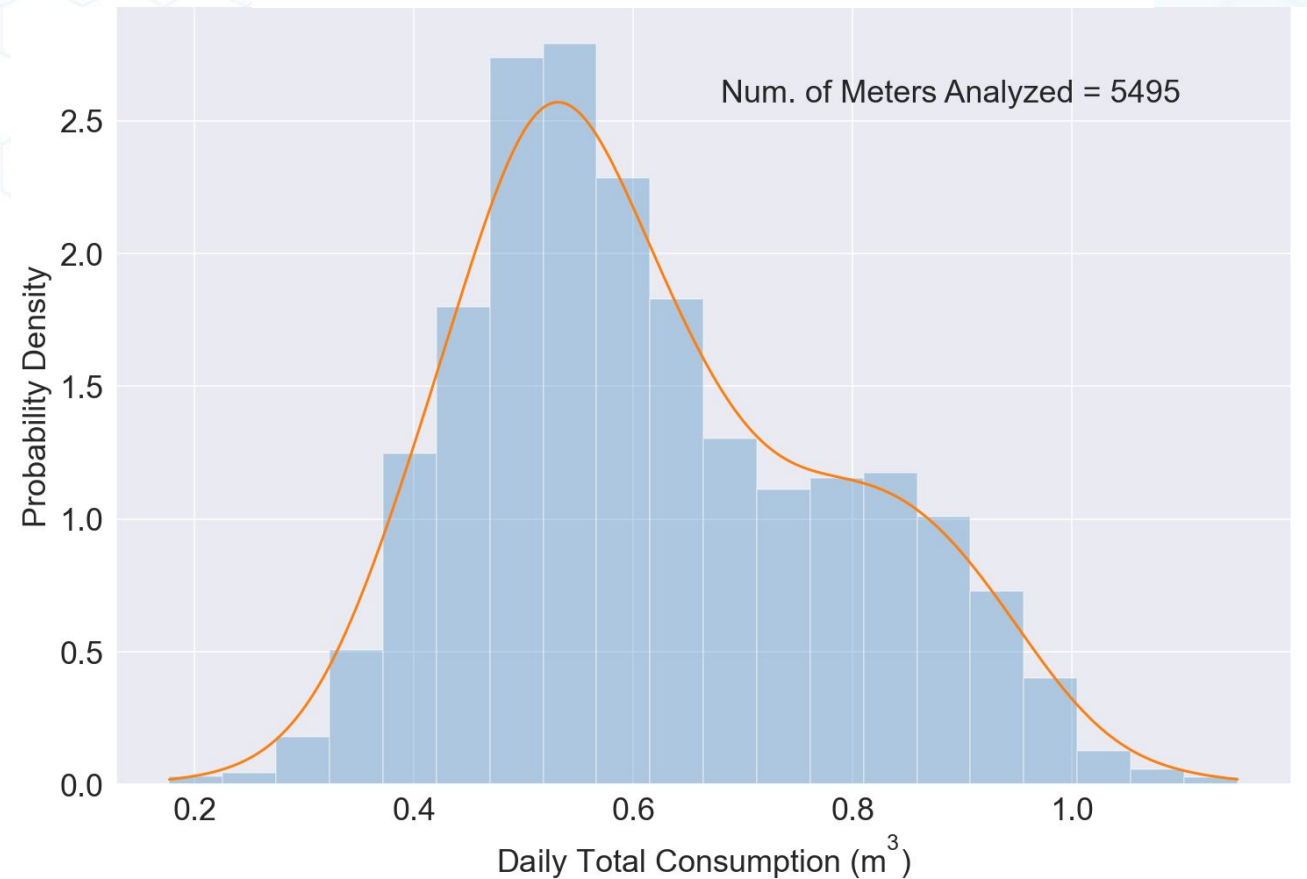
NRW after simple adjustment (assuming missing meters have consumption = avg. consumption)



Building a More Accurate Estimate for Missing Meter Data

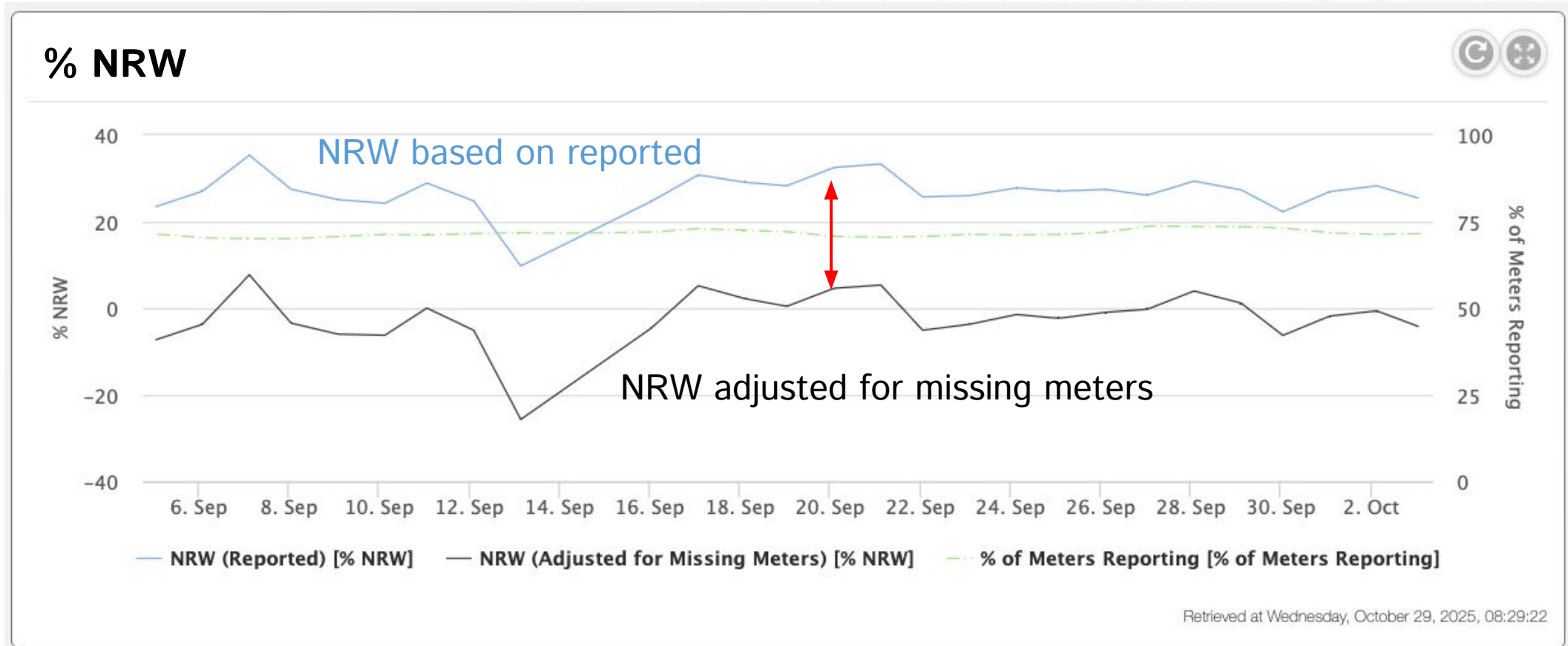
- Historical data not available through API, must be obtained through other means
□ **data aggregation challenge**
- From historical data, we can build a set of statistical distributions corresponding to the daily consumption across all meters.
 - Can create multiple distributions to capture seasonal differences.
- Missing meter data can be sampled from these distributions.

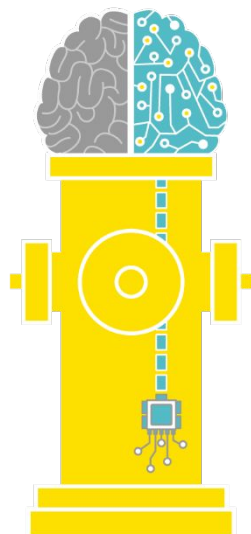
Histogram of Daily Meter Consumption



Implications On Non-Revenue Water

The utility may not have as big of an NRW problem as initially suspected.





Thank you for your time!

Questions?



Stan Fong, CTO
stan.fong@digitalwater.solutions



Barclay Whittaker, Account Executive
barclay.whittaker@digitalwater.solutions

Learn more at <https://digitalwater.solutions> or
come stop by our booth to chat!