

Not All Water Consumption Is Equal: A Water Stress Weighted Metric for Sustainable Computing

Yanran Wu, Inez Hua, Yi Ding
{wu2187, hua, yiding}@purdue.edu



Computing is thirsty

- The rise of AI and computing comes with a hidden cost: **water consumption**.

10~50 GPT-3 queries

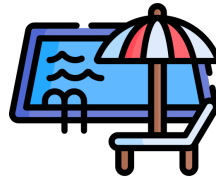
500 mL water



A Google's datacenter

3.7 million m³ water

1,500



Intel

9.6 million m³ water

23,000



Water consumption is high across the entire computing stack.

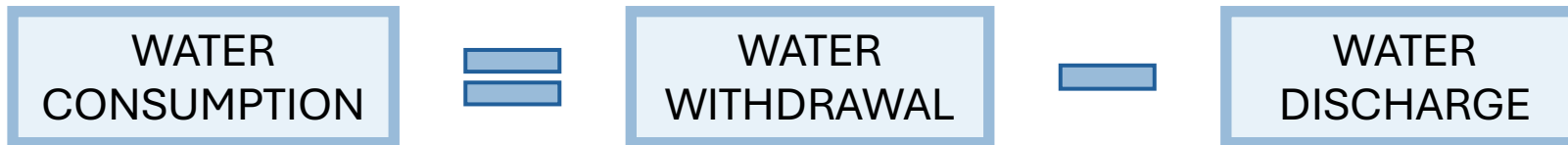
**How much do we understand water
consumption?**

What is water consumption?

- **Water Withdrawal ≠ Water Consumption**

Water Withdrawal: Total water taken from the environment

Water Consumption: Water removed from the environment



- **Water Consumption = On-site + Off-site**

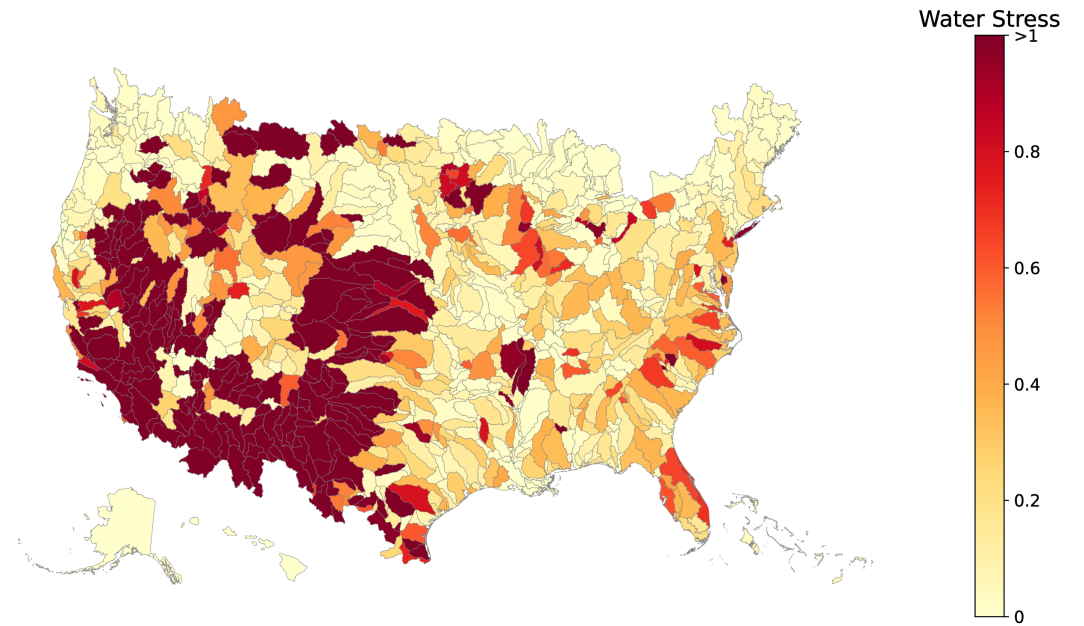
On-site consumption: Loss at facility (e.g., evaporation)

Off-site consumption: Indirect consumption (e.g., power generation)

Water consumption is location-sensitive

$$\text{Water stress} = \frac{\text{Water Demand}}{\text{Water Availability}} \quad [1]$$

Spatially, same amount of water consumption can lead to greater environment burden in arid regions than in water-abundant regions.

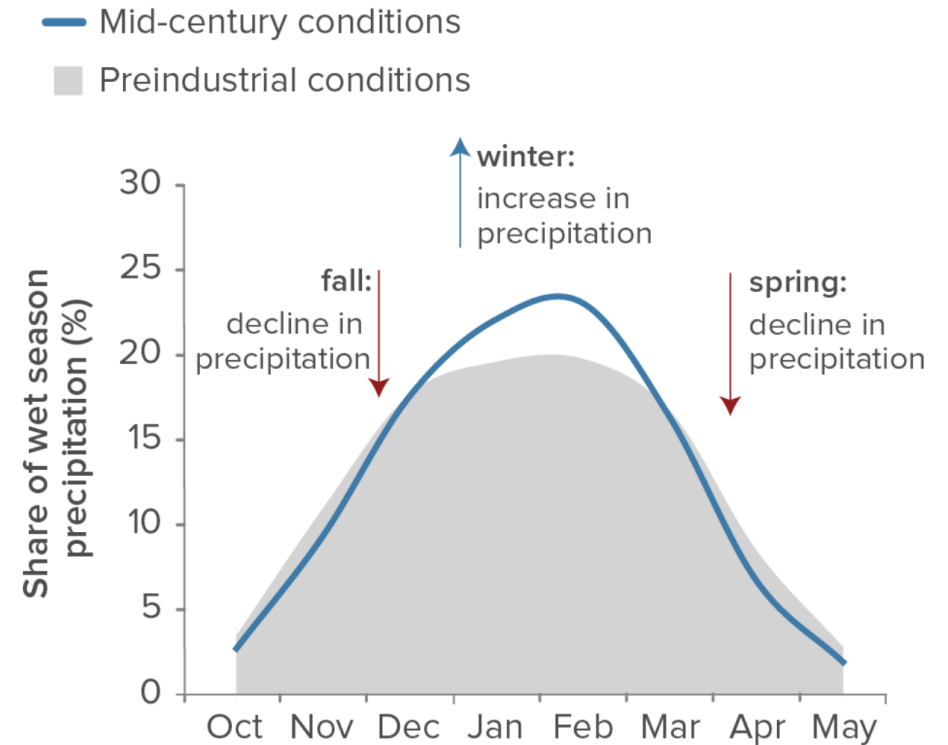


Water stress map of United States

Water consumption is time-sensitive

- Water availability fluctuates over seasons
- Climate change influences water in long-term

Temporally, water supply fluctuates over time. We must account for both the **location** and **time** when analyzing water impact.



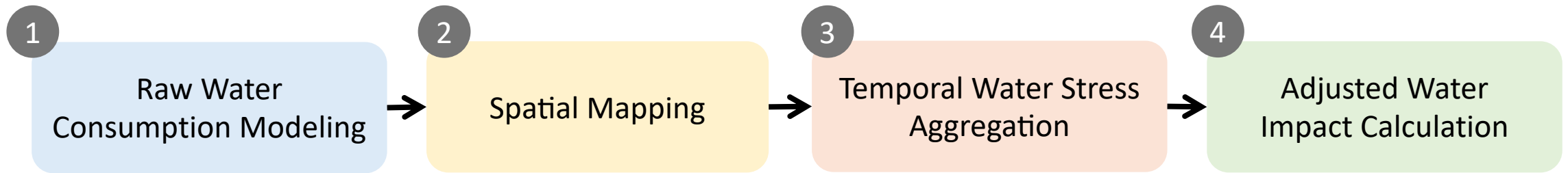
Seasonal Precipitation in California [2]

Our Contributions

- **Present SCARF: Stress-Corrected Assessment of Water Resource Footprint**, a general framework for water impact evaluation in computing.
 - **Incorporate spatial and temporal variations of water stress** to better evaluate water consumption impact.
 - **Introduce Adjusted Water Impact (AWI)** a unified metric combining water consumption with local water stress.
 - **Conduct three case studies** across LLM serving, datacenters, and semiconductor fabs

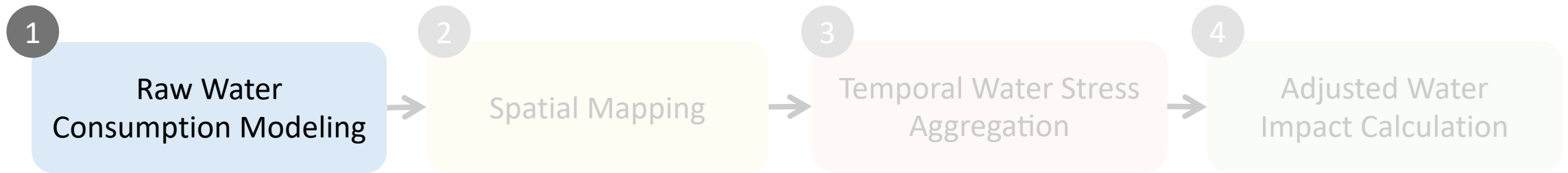
Our method: SCARF

- **SCARF**: Stress-Corrected Assessment of Water Resource Footprint
- **Insight**: Incorporate spatial and temporal variations



Our method: SCARF

- **SCARF**: Stress-Corrected Assessment of Water Resource Footprint
- **Insight**: Incorporate spatial and temporal variations

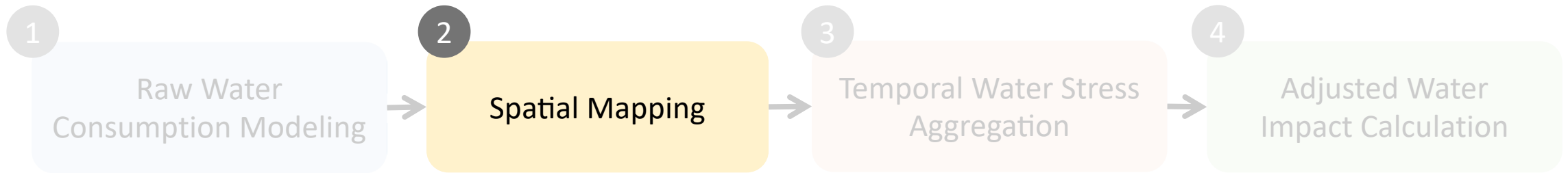


$$\text{Water Consumption} = W_{\text{on-site}} + W_{\text{off-site}}$$

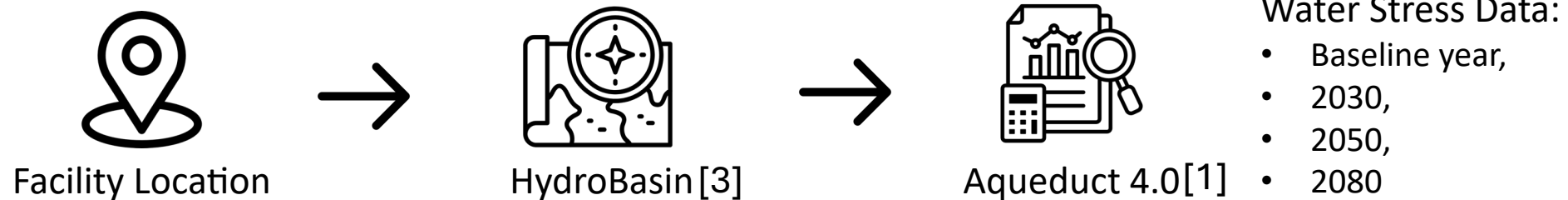
- **On-site water consumption** directly consumed at the facility
- **Off-site water consumption** indirectly consumed through electricity generation

Our method: SCARF

- **SCARF**: Stress-Corrected Assessment of Water Resource Footprint
- **Insight**: Incorporate spatial and temporal variations

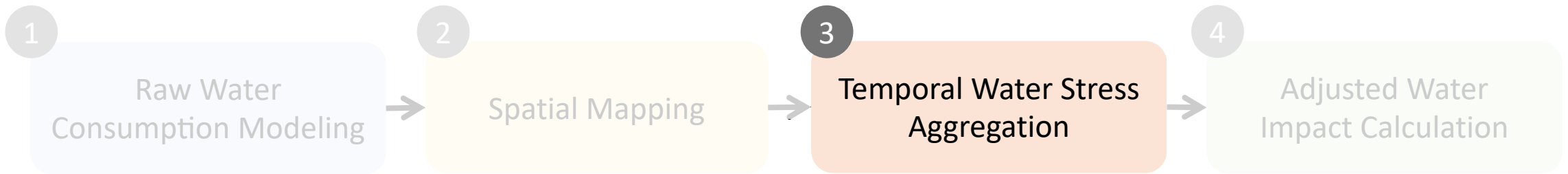


- **We retrieve regional water stress data from Aqueduct dataset.**



Our method: SCARF

- **SCARF**: Stress-Corrected Assessment of Water Resource Footprint
- **Insight**: Incorporate spatial and temporal variations



- **We introduce Water Stress Factor (WSF)**

- Time-weighted aggregation
→ capture temporal variation in water stress

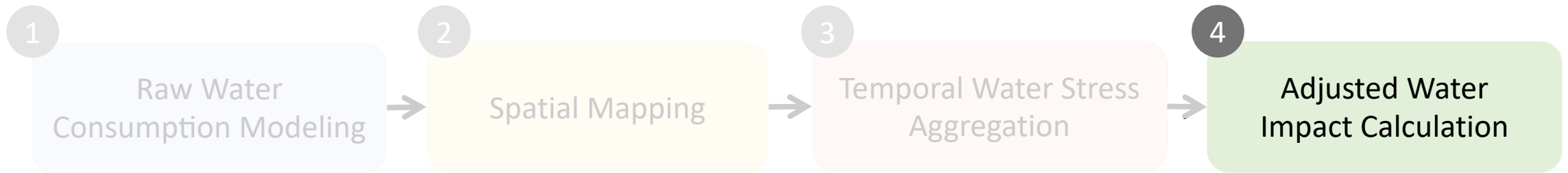
$$WSF_b^{\text{short}} = WS_{t_0,b} \quad WSF_b^{\text{long}} = \sum_{t \in T} w_t \cdot WS_{t,b}$$

- User-defined discount rate (γ)
→ reflect future uncertainty and potential impact

$$w'_t = \frac{1}{(1 + \gamma)^{t-t_0}}, \quad w_t = \frac{w'_t}{\sum_{t' \in T} w'_{t'}}$$

Our method: SCARF

- **SCARF**: Stress-Corrected Assessment of Water Resource Footprint
- **Insight**: Incorporate spatial and temporal variations



- **Our new metric for evaluating water consumption: Adjusted Water Impact (AWI)**

$$AWI = (W_{\text{on}} + W_{\text{off}}) \times WSF_b$$

Evaluation

- Goal:
 - Show the **generalizability** of SCARF.
 - Quantify the hidden **environmental impact** of water consumption
 - Provide **actionable insights** for more water-sustainable deployment
- We apply SCARF across three case studies:
 - ✓ LLM Serving
 - ✓ Datacenters
 - ✓ Semiconductor Manufacturing

Case Study I: LLM Serving

- **Models:** Three Qwen2.5 models — 7B, 14B, 32B [4]
- **Workload:** ShareGPT
- **On-site Water Consumption:**
 - Ramp up QPS, select peak throughput.
 - Measure GPU (H100) power during inference.
 - On-site Water Usage Effectiveness (WUE) from Microsoft reports [5].
- **Off-site Water Consumption:**
 - Estimate total energy from GPU power.
 - Regional electricity water intensity applied [6].
- **Short-term Water Impact Evaluation**

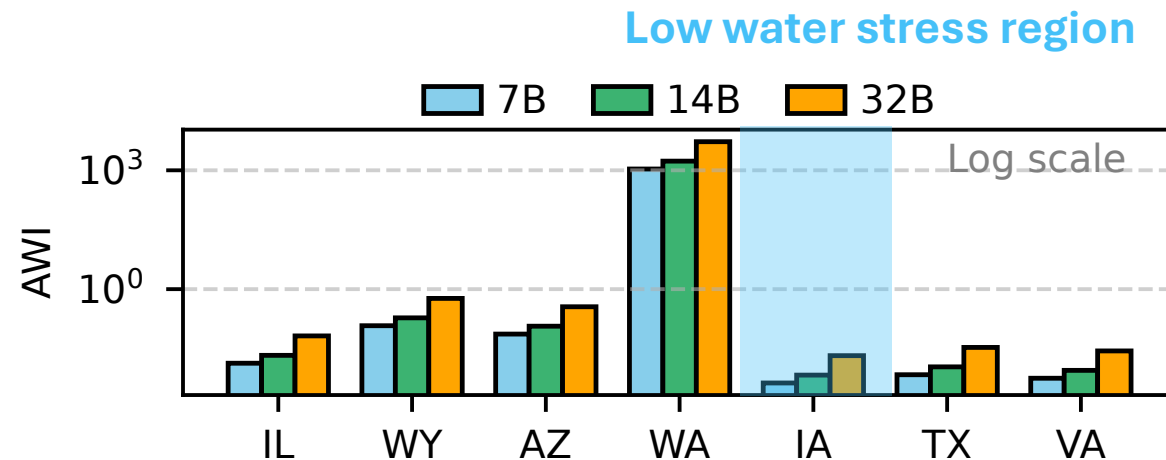
[4] Yang et al. Qwen2.5 Technical Report. arXiv preprint arXiv:2412.15115, 2024.

[5] Microsoft. <https://datacenters.microsoft.com/sustainability/efficiency>. 2024.

[6] Siddik et al. Spatially and temporally detailed water and carbon footprints of US electricity generation and use. *Water Resources Research*, 2024.

Case Study I: LLM Serving

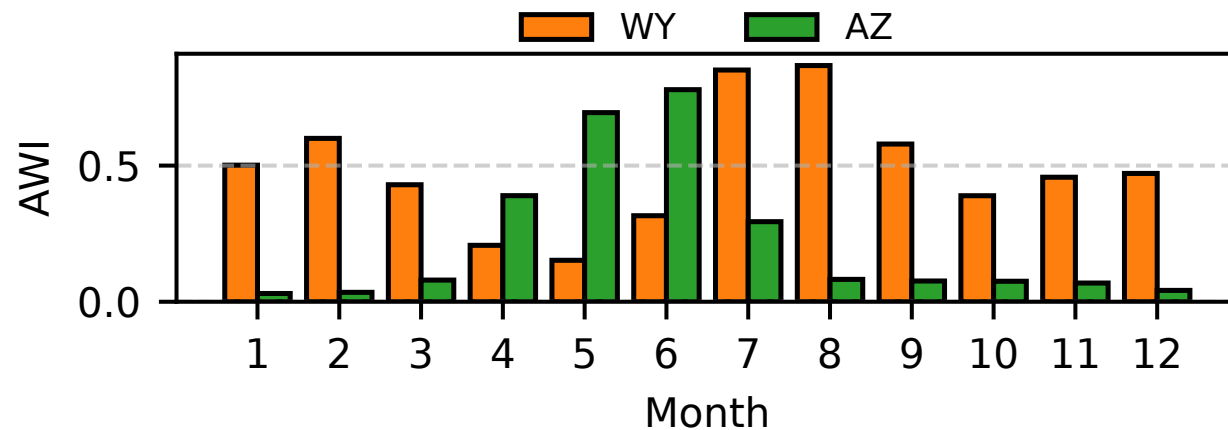
- Evaluate *per-request* Adjusted Water Impact (AWI) of 7B, 14B and 32B Qwen 2.5 model **in different regions**.



The adjusted water impact of deploying LLMs is highly **location-sensitive**. Same workloads can have 1000× differences depending on where they are served.

Case Study I: LLM Serving

- Evaluate *per-request* Adjusted Water Impact (AWI) of 32B Qwen2.5 in **different months** but same regions (AZ vs WY).



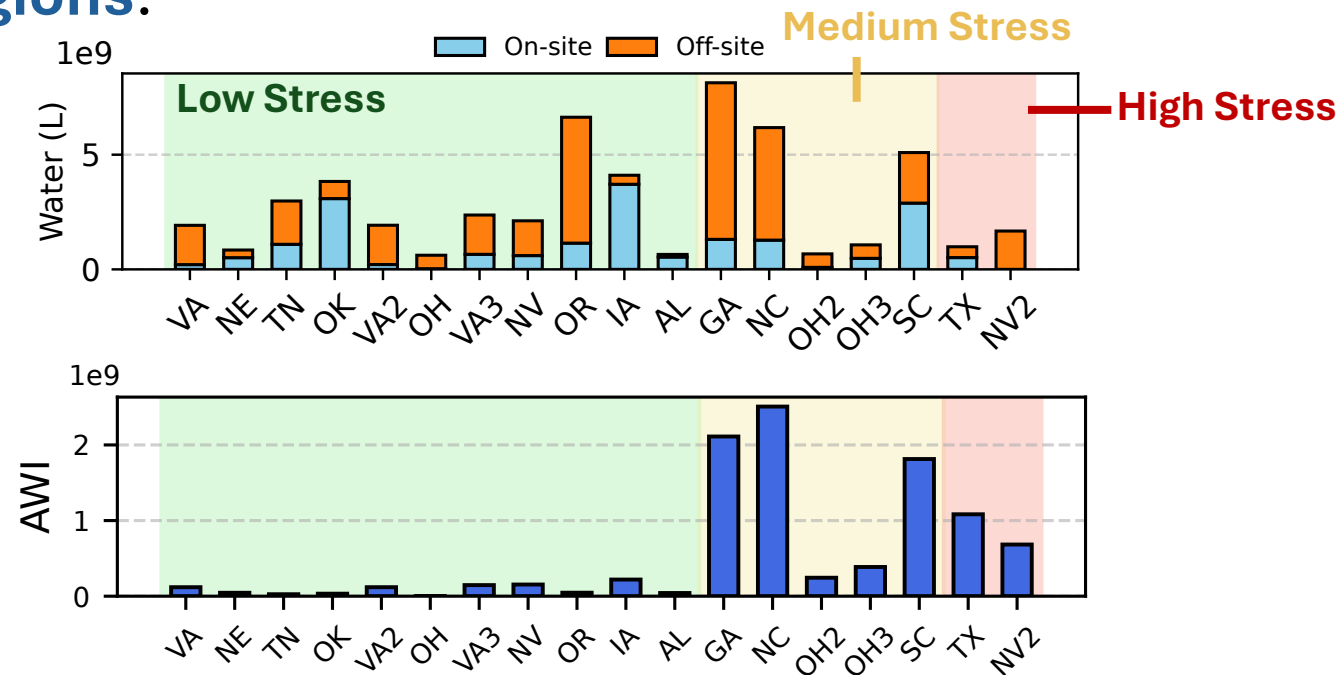
Even in the same location, seasonal changes can significantly affect adjusted water impact. **When we deploy matters**—not just where.

Case Study II: Datacenters

- **On-site Water Consumption:**
 - From Google's public sustainability reports [7].
- **Off-site Water Consumption:**
 - Estimate datacenter power capacity using *DataCenters.com*.
 - Assume 70% average utilization, 24/7 operation.
 - Regional electricity water intensity applied [6].
- **Long-term Water Impact Evaluation**

Case Study II: Datacenters

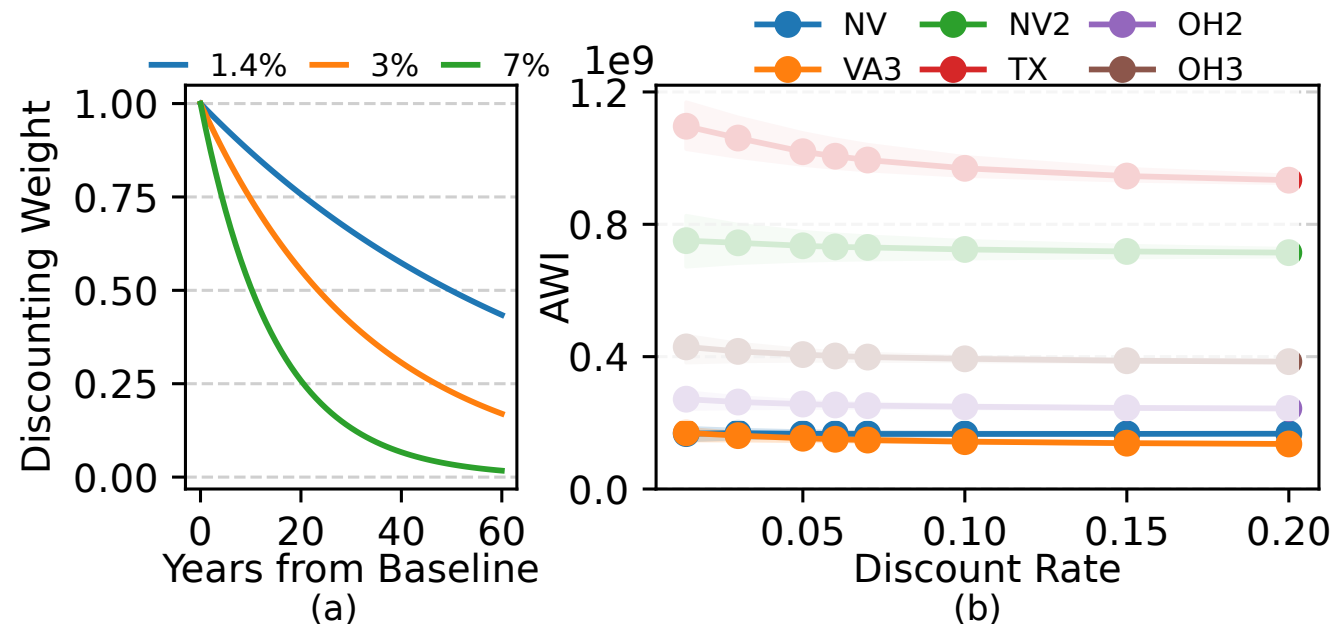
- Evaluate *annual* Adjusted Water Impact (AWI) of datacenters in **different regions**.



Water stress alone is not sufficient to capture true water impact. High water consumption in medium-stress regions can cause greater impact than expected.

Case Study II: Datacenters

- Evaluate *annual* Adjusted Water Impact (AWI) of Google's datacenters with **different discount rates**.



Different planning perspectives lead to different sustainable decisions.

Case Study III: Semiconductor Fabs

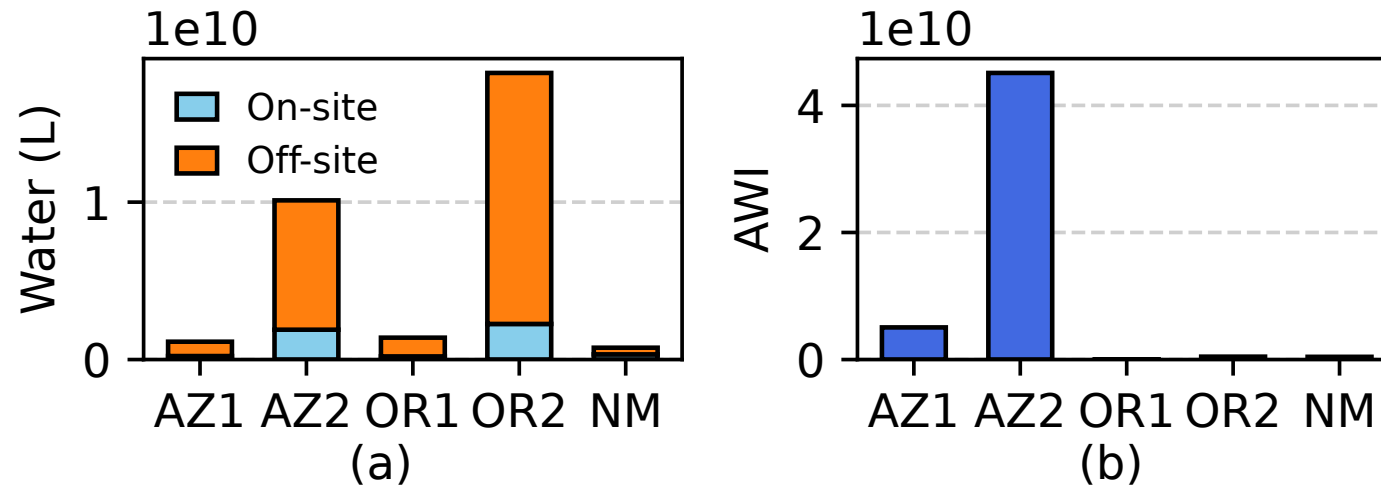
- **On-site Water Consumption:**
 - From Intel's public sustainability report [7-8].
- **Off-site Water Consumption:**
 - Power consumption from Intel's reports.
 - Regional electricity water intensity applied [6].
- **Long-term Water Impact Evaluation**

[7] Intel. <https://csrreportbuilder.intel.com/pdfbuilder/pdfs/CSR-2023-24-Full-Report.pdf>. 2024.

[8] Intel. <https://www.exploreintel.com/>

Case Study III: Semiconductor Fabs

- Evaluate *annual* Adjusted Water Impact (AWI) of Intel's fabs in **different regions**.



Semiconductor fabs consume a large amount of water. Building fabs in water-scarce regions like Arizona causes higher environmental burdens.

Take Home Messages

SCARF quantifies water impact based on spatial and temporal variations in water stress.



- **Region matters.**

Water stress level varies dramatically across regions.



- **Time matters.**

Water stress shifts with seasons and long-term climate change.



- **Amount matters.**

Large water consumption amplifies water impact even in low-stress area.



- **How we value the future matters.**

Different priorities for future can reshape sustainability decisions.