Geo-Mismatch Challenge

Joseph Ouyang, Hsing-Yu Wen, Caleb Ouyang, David Sun

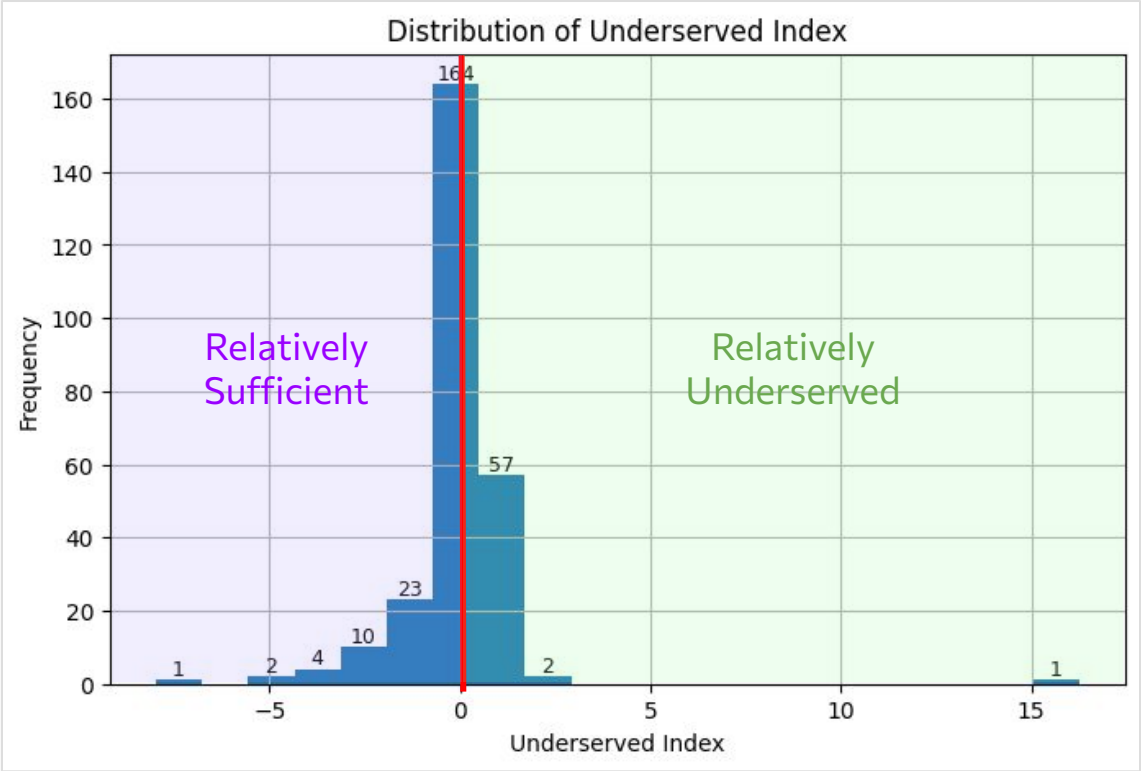
How did we curate our cleaned dataset?

- Define each **country–cluster sample** as the unit of analysis.
- Compute **Need index** (in need ÷ total pop) and **Resource allocation index** (dollars requested ÷ in-need pop).
- **Normalize** both indices across the sample, then compute **Underserved index** = normalized need − normalized allocation.

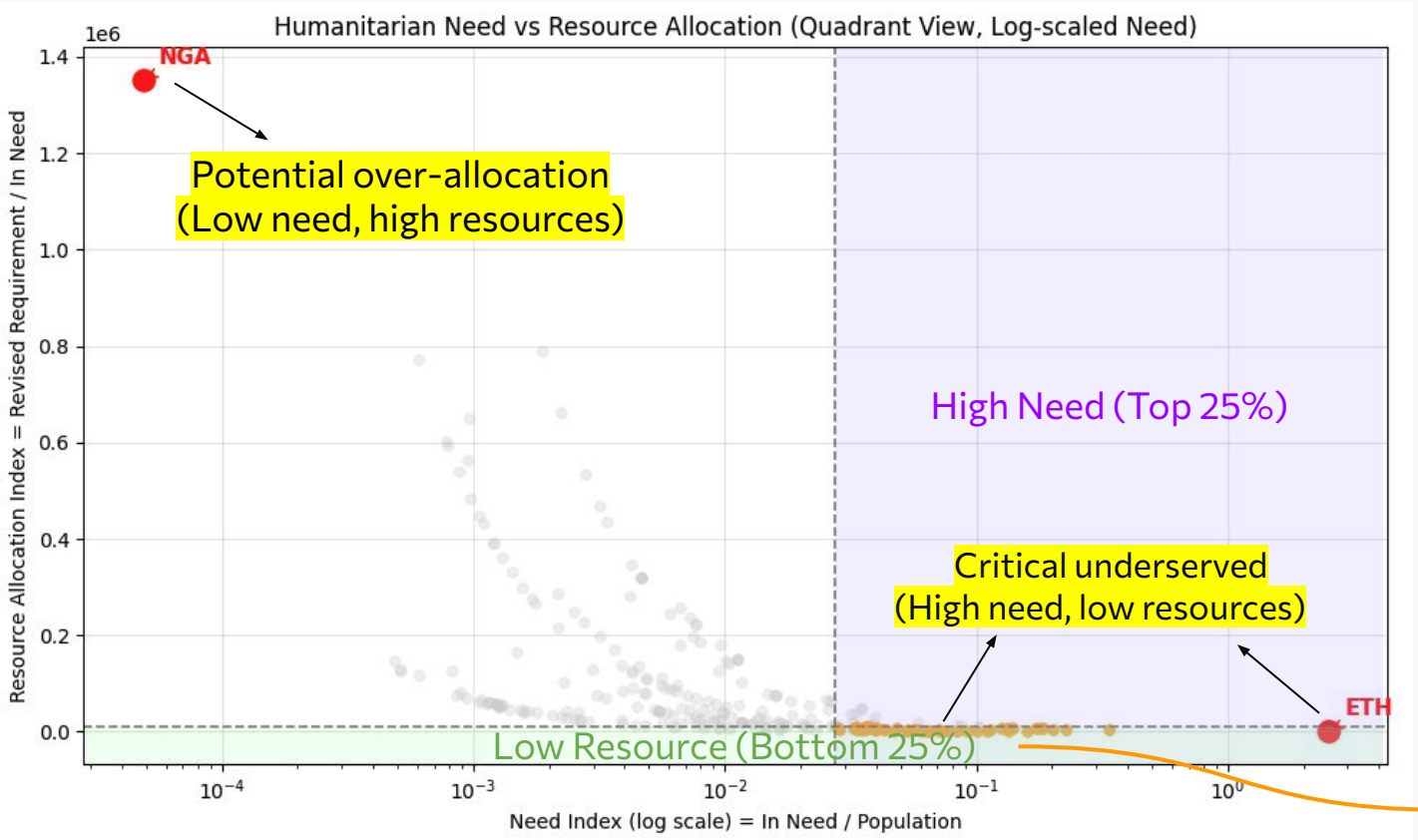
01

Exploratory Data Analysis

Distribution of Underserved Index



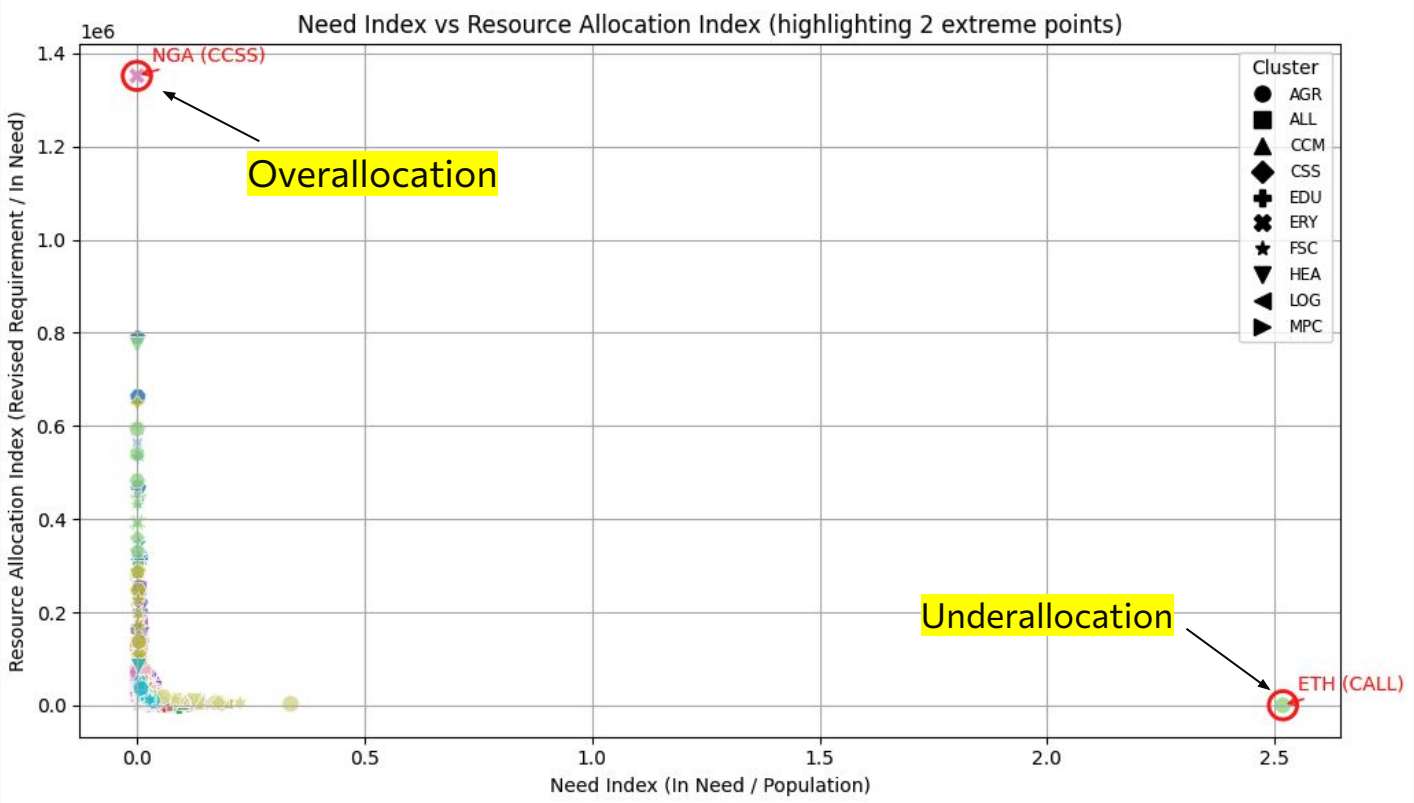
Humanitarian Need vs Resource Allocation(Log-scales)



Underserved pairs

ISO3	Cluster
AFG	ALL
AFG	CSS
AFG	EDU
AFG	FSC
AFG	HEA
AFG	LOG
AFG	MPC
AFG	MS
AFG	NUT
AFG	PRO
AFG	PRO-CPN
⋮	
VEN	PRO-GBV
VEN	SHL
VEN	WSH

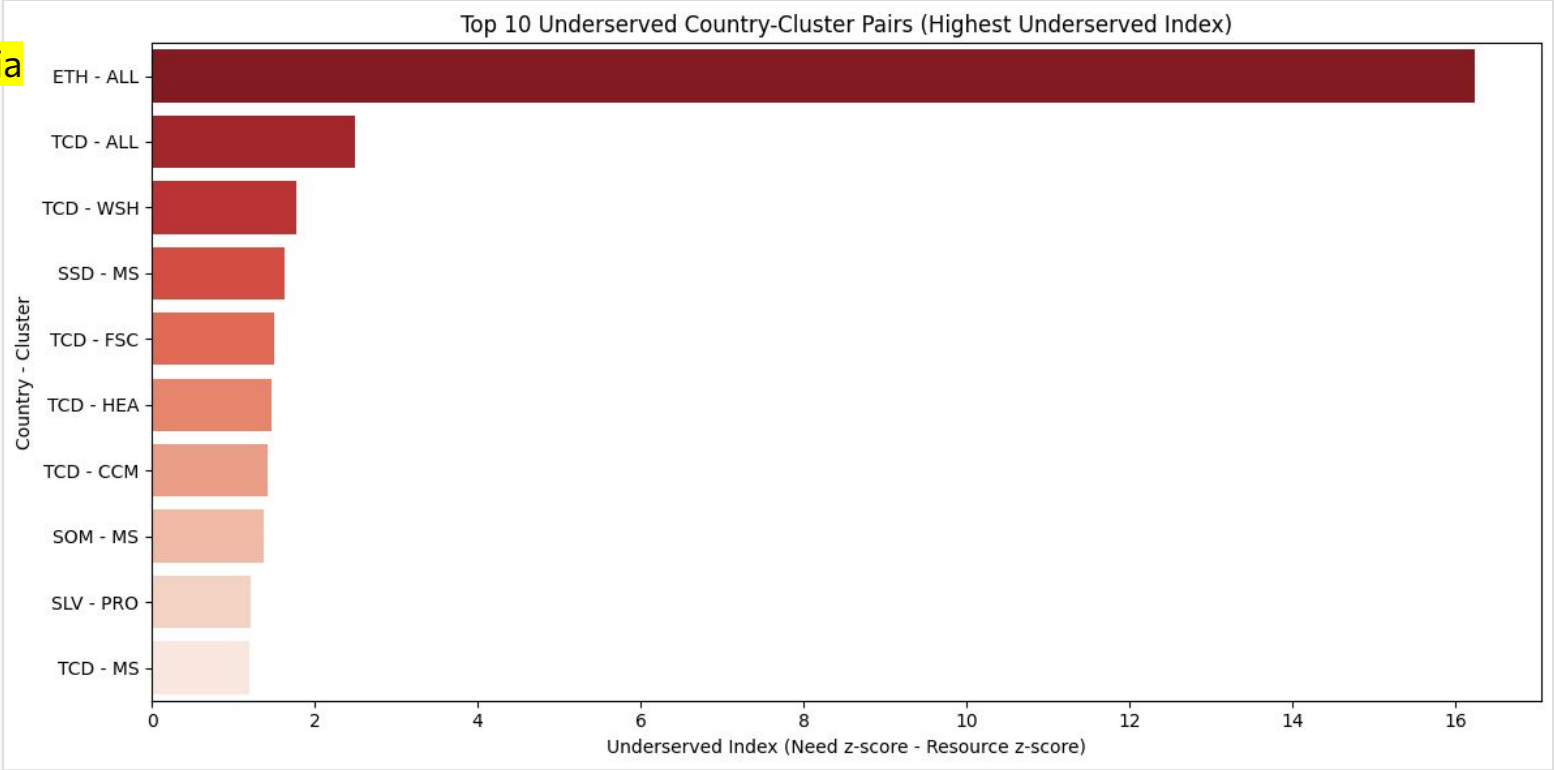
Need Index vs Resource Allocation Index



Top 10 Underserved Country-Cluster Pairs

Ethiopia

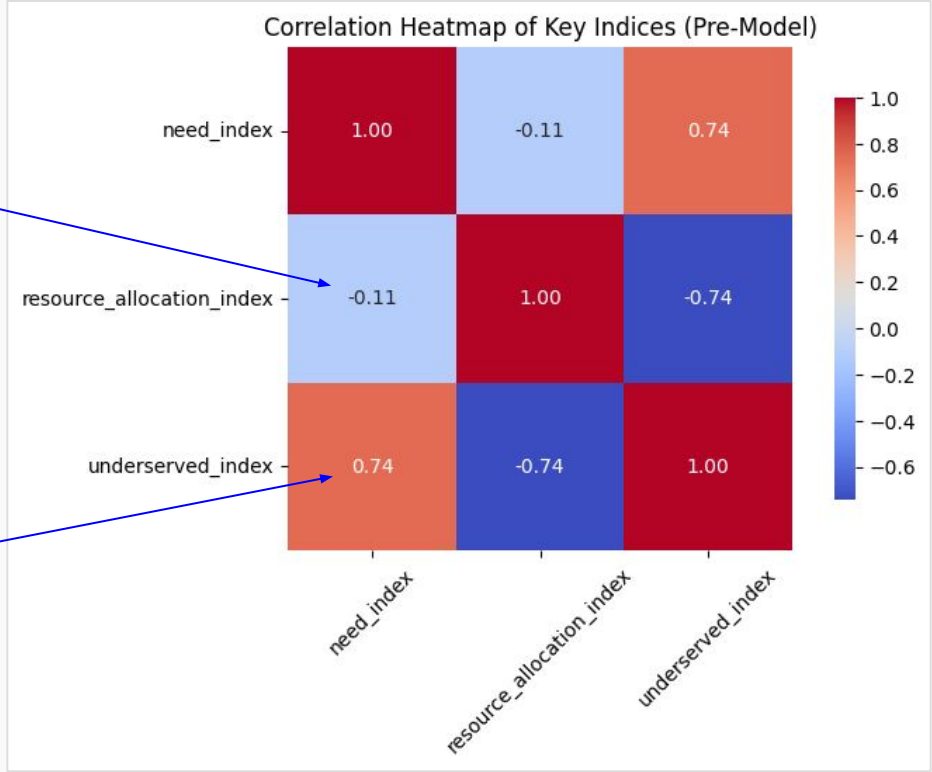
Chad



Correlation Heatmap of Key Indices

The HIGHER a region's need, the LOWER the amount of resources allocated there? (Area of interest)

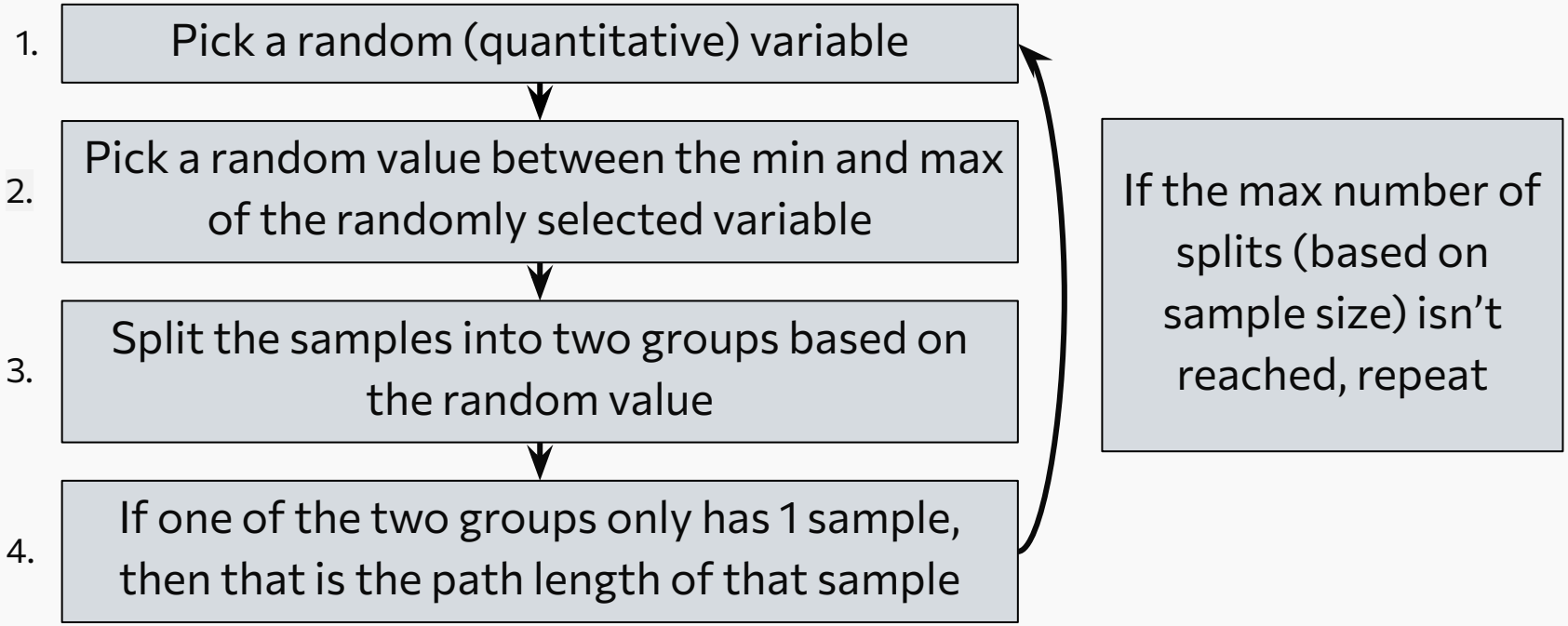
The HIGHER a region's need, the LARGER the gap between resource allocation and need.



02

Modeling

Isolation Tree



Isolation Tree Outliers at 1% Contamination Level

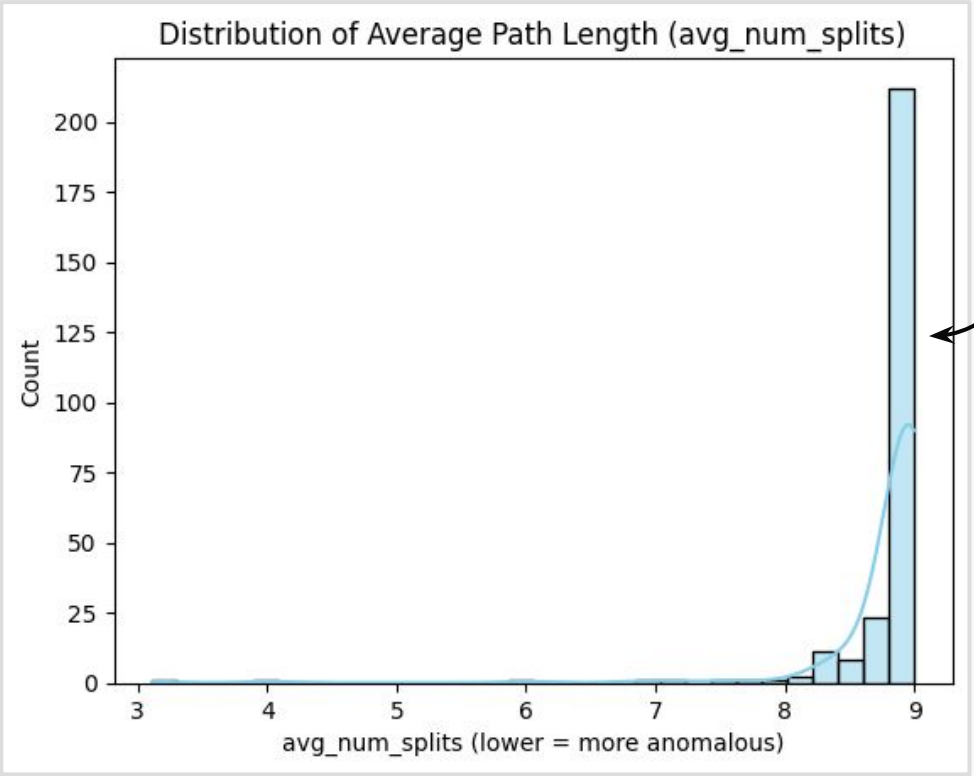
	ISO3	Cluster	is_anomaly
1	ETH	ALL	1
2	TCD	ALL	1
3	NGA	CSS	1

Key

ETH - Ethiopia
TCD - Chad
NGA - Nigeria

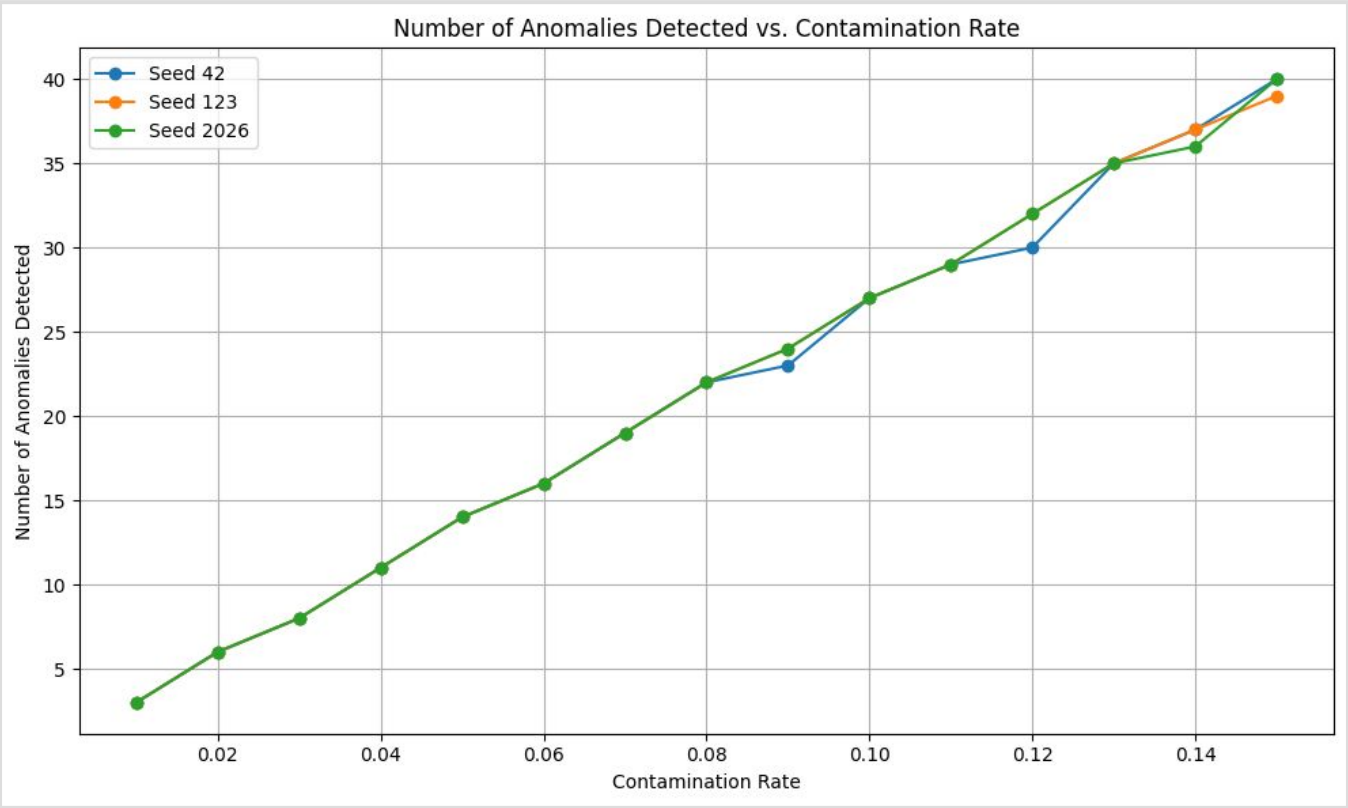
ALL - Overall
CSS - Coordination and Support Services

Histogram of avg_num_splits



The vast majority of country-cluster pairs aren't detected as outliers before the maximum number of splits is reached.

Evaluation of Model Performance



03

Impact

Humanitarian Relevance and Alignment

1

Our **unified humanitarian dataset** combines country-level need, population, and funding, enabling calculation of **standardized indices** (need index, resource allocation index, underserved index).

2

By applying anomaly detection and quadrant analysis, we pinpoint **specific regions** where need far exceeds resources, providing **actionable targets** for UN aid prioritization and resource reallocation.

Potential for Real-World Application

1

The quadrant visualization and underserved index can be integrated into a **lightweight dashboard**, allowing UN field teams to instantly **identify** and **monitor** high-need, low-resource regions—even with limited connectivity.

2

Our methodology supports flexible sensitivity: field teams can **adjust thresholds** (e.g., contamination rate, percentile cutoffs) to match **operational realities**, and the model's outputs are interpretable, enabling **transparent reporting** and **iterative refinement** based on field feedback.