

Impact Evaluations in Data Poor Settings: The Case of Stress-Tolerant Rice Varieties in Bangladesh

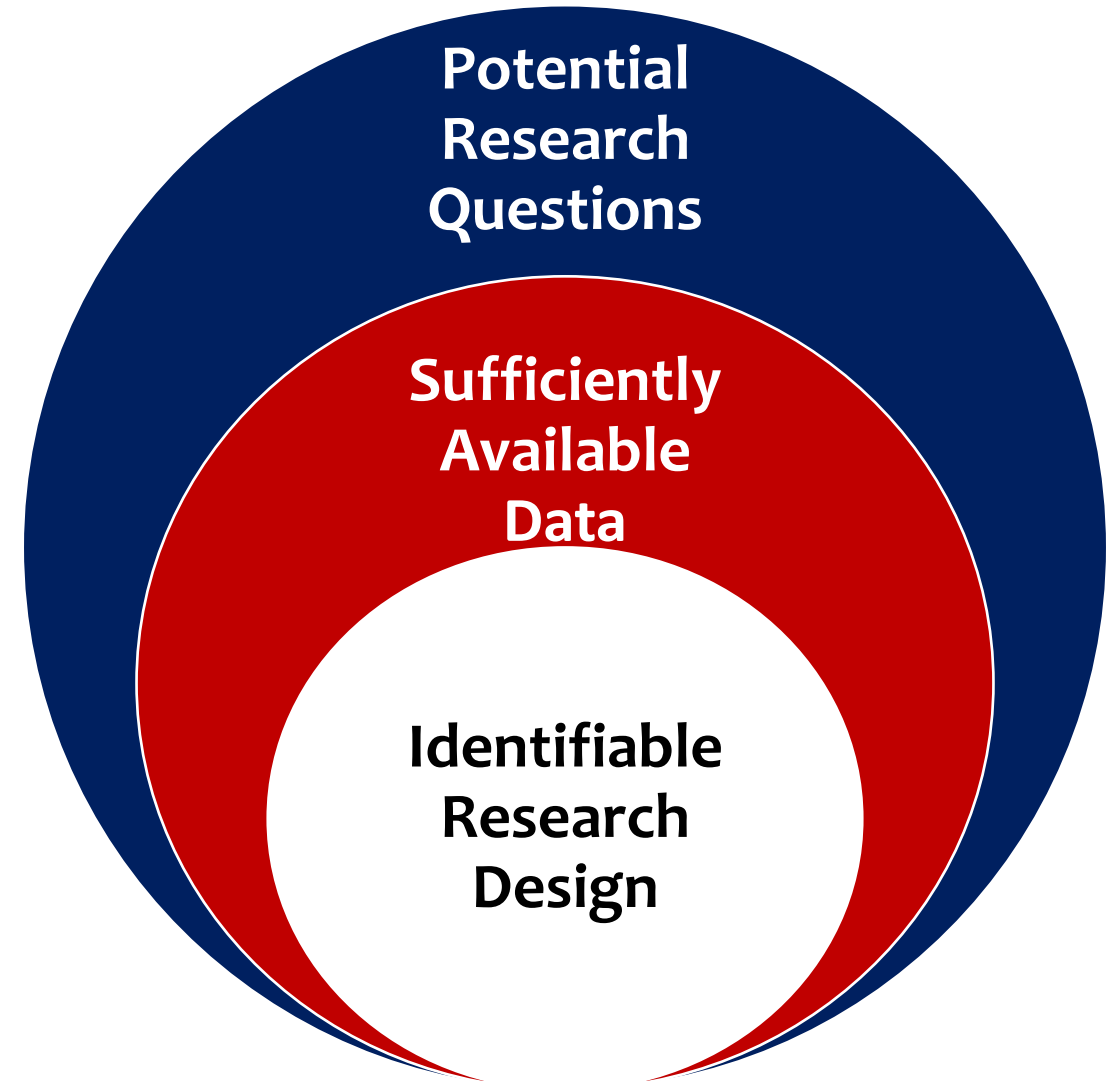
**SPIA Fest
Delhi
2 August 2024**

Jeffrey D. Michler
Dewan Abdullah Al Rafi
Jonathan Giezendanner
Anna Josephson
Valerien Pedé
Elizabeth Tellman



Motivation

- The set of interesting research questions is almost infinite
- The set of questions applied economist can answer is a much smaller subset.
Limited by
 1. Data availability
 2. Plausible identification strategy
- Even if 1. and 2. are met, many technologies we are interested in have:
 - Heterogeneous treatment effects
 - Stochastic treatment effects
- These difficult to measure treatment effects are the source of the new Diff-in-Diff literature



Our Approach

- We use Earth Observation (EO) data where and when we lack household data (**not new**)
 - Complement EO data with 3 waves of household panel data (2014, 2017, 2022)
- Face two hurdles:
 - Lack of ground truth data – Machine Learning (ML) algorithms require training data to identify and classify things like rice and flooding
 - Recency bias in EO data – publicly available high spatial and temporal res data is only ~10 year old
- We develop an innovative approach to reconstructing the past that address both hurdles
 - Employ new approaches to constructing training or ground truth data
 - Deploy recent ML methods (convolutional neural networks) to improve the “resolution” of EO data that is 20+ years old

Pre-Analysis Plan

What

- Describes how a researcher plans to conduct a study and analyze the resulting data
- Pre-specify in a precise way the analysis to be done before examining the data (Olken, 2015)

Why (Janzen and Michler, 2021; Josephson and Michler, 2023)

- To avoid the temptation for data mining
- To avoid post-hoc adjustment of research hypothesis
- Promote ethical, transparent research methods

What have we done?

- Filed a detailed Pre-Analysis Plan in Open Science Foundation at doi.org/10.17605/OSF.IO/YE7PV

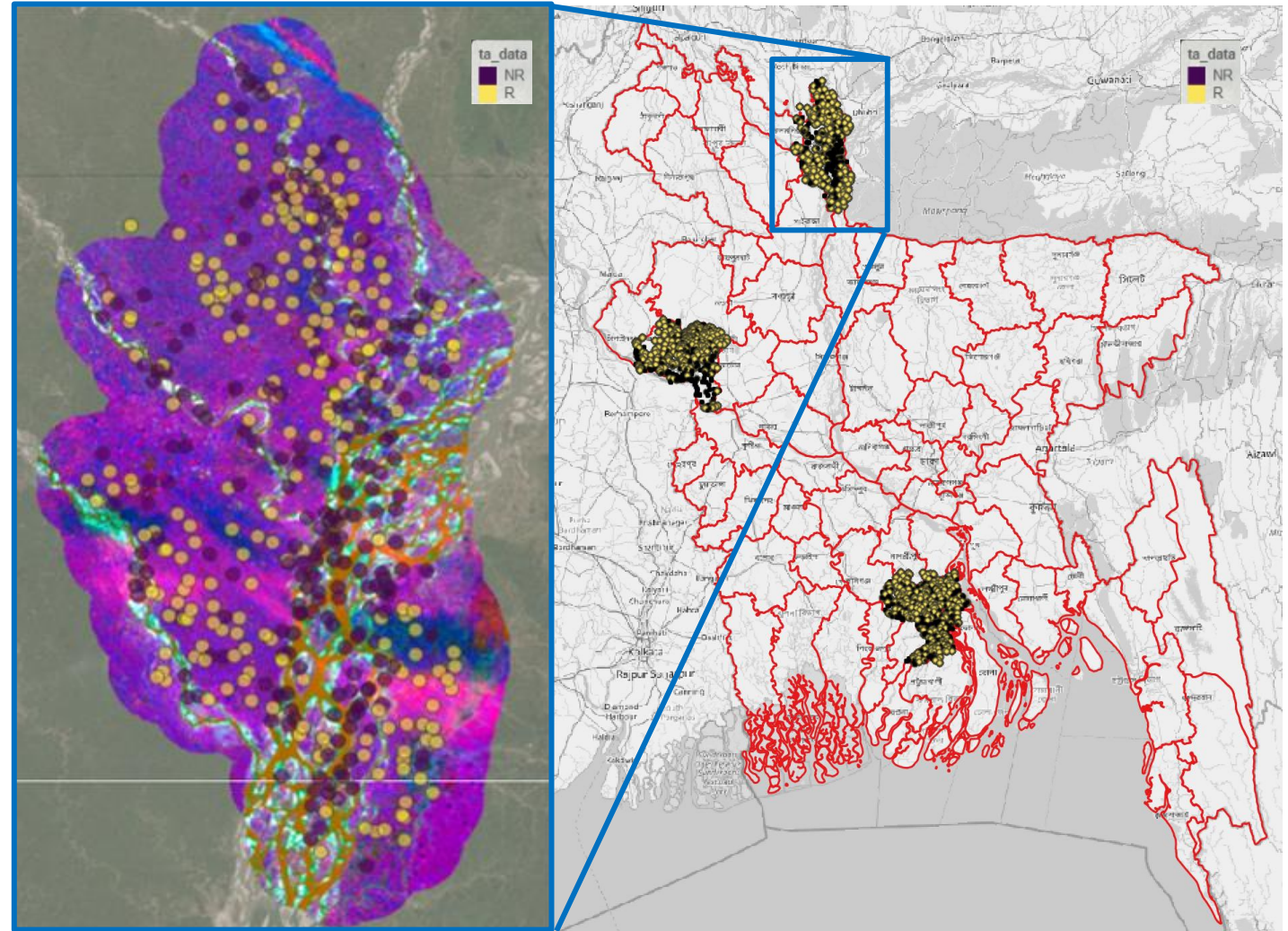
Rice Map Details: Ground Truth

Baseline Rice/NoRice (**RNR**) area map derived from MODIS using agronomic flooding

Random Sampling of RNR pixels → 'Ground Truth' data set

RNR 'Ground Truth' interpreted and confirmed using high resolution Google Earth data

Sampled years: 2002, 2004, 2006, 2009, 2015, 2016, 2018, 2019, 2021
Sampled districts: Barisal, Kurigram, Rajshahi



450 total number of samples

Rice Map Details: Random Forest Generated Maps

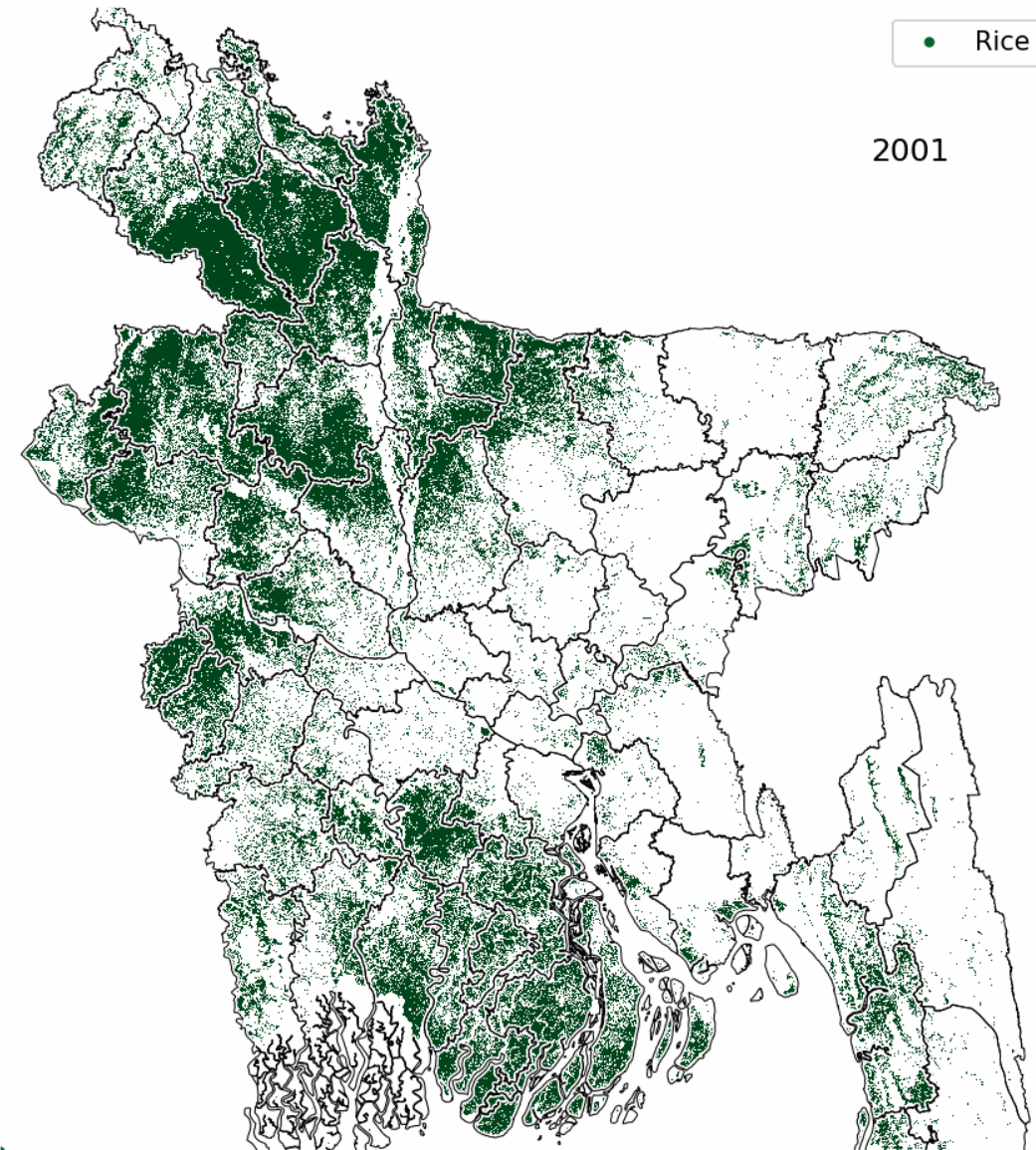
Data and model processed in Google Earth Engine

Data:

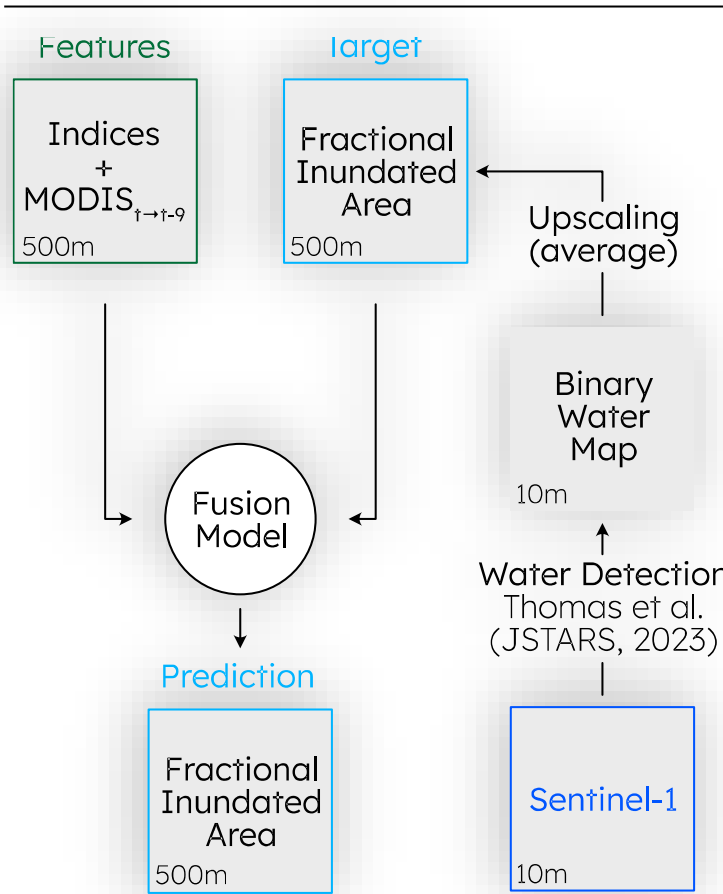
- MODIS Terra 8-days Composite Median value for the rice grow season (\pm June to December, may vary depending on the district)
- FABDEM elevation and slope

Random Forest:

- 70% of data for training, 30% for testing
- .77 accuracy

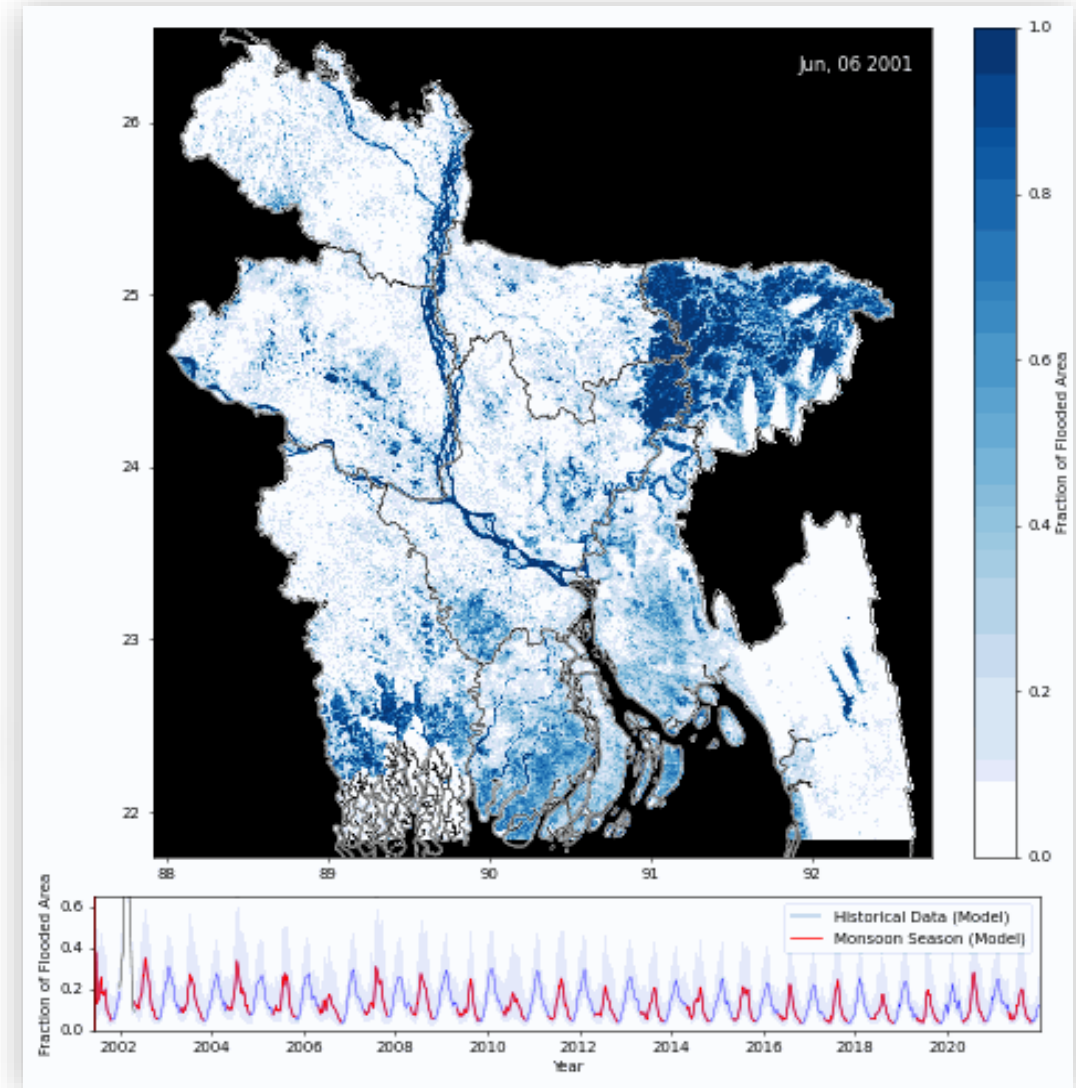


Flood Map Details: Fusion Model



Fusion Model based on recreating Sentinel-1 observed Fractional Inundated Area with MODIS

Shows fraction of 500 x 500 meters pixel covered by water every 8 days, from 2002 to 2022



Paper, Data and Code
 Giezendanner et al (2023) *Inferring the past: a combined CNN-LSTM deep learning framework to fuse satellites for historical inundation mapping*, CVPR Earthvision Workshop

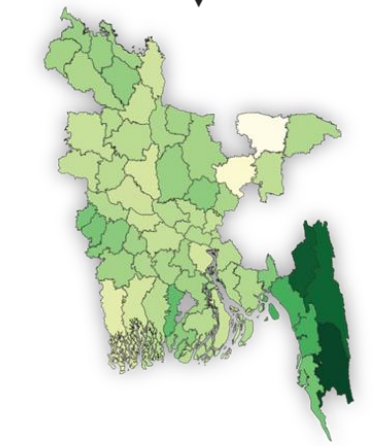
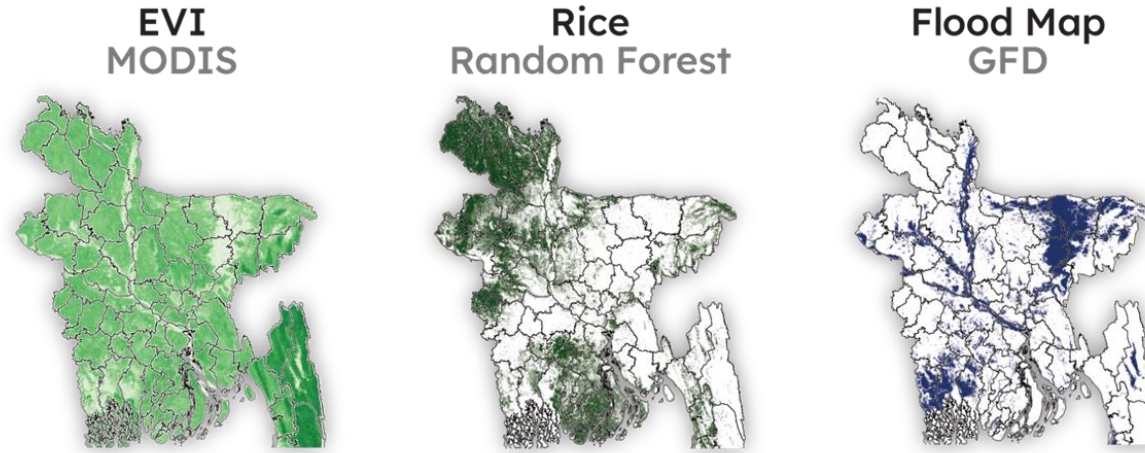
Remote Sensing and Administrative Data

EVI: MODIS median from June to December

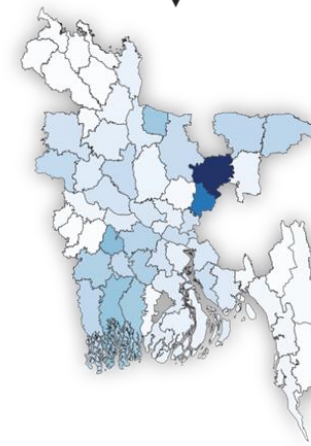
Rice Presence: Random Forest Algorithm based on MODIS

Flood Map: Fractional Inundation Area based on Giezendanner et al. (2023)

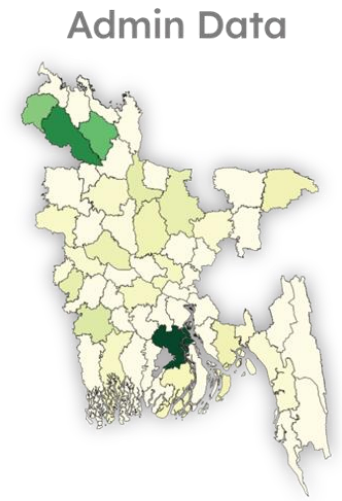
Aman Rice Seeds: Administrative data from government offices in each district



EVI
Within rice pixels
Average per district

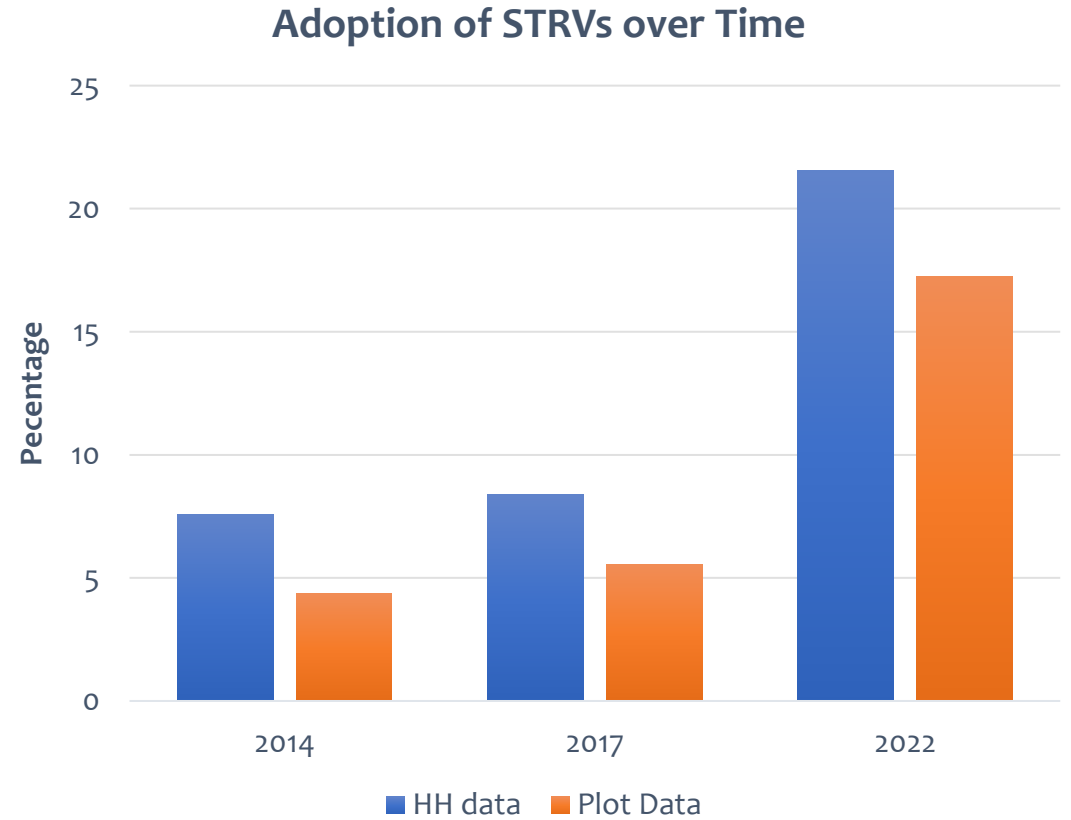
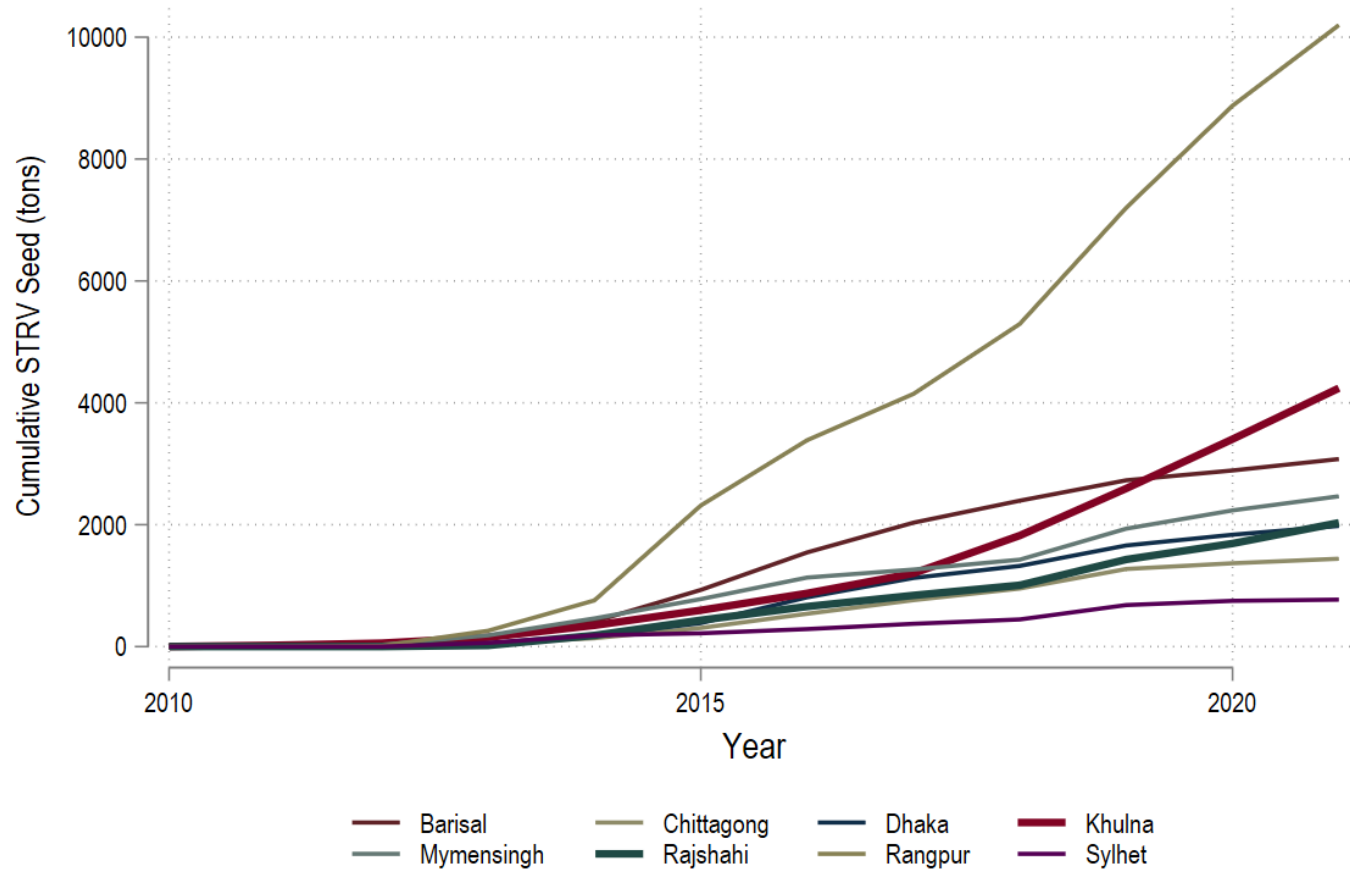


Fraction of Flooded Area
Within rice pixels
Average per district



Aman Rice Seeds
Cumulative per district

Dissemination and Adoption of STRVs



Empirical Models

- Event Study

$$EVI_{dt} = \sum_{j \in T} \alpha_j \cdot 1_t^{t=\tau+j} + \delta rice_{dt} + \varphi FLOOD_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$

- Difference-in-Difference (DID)

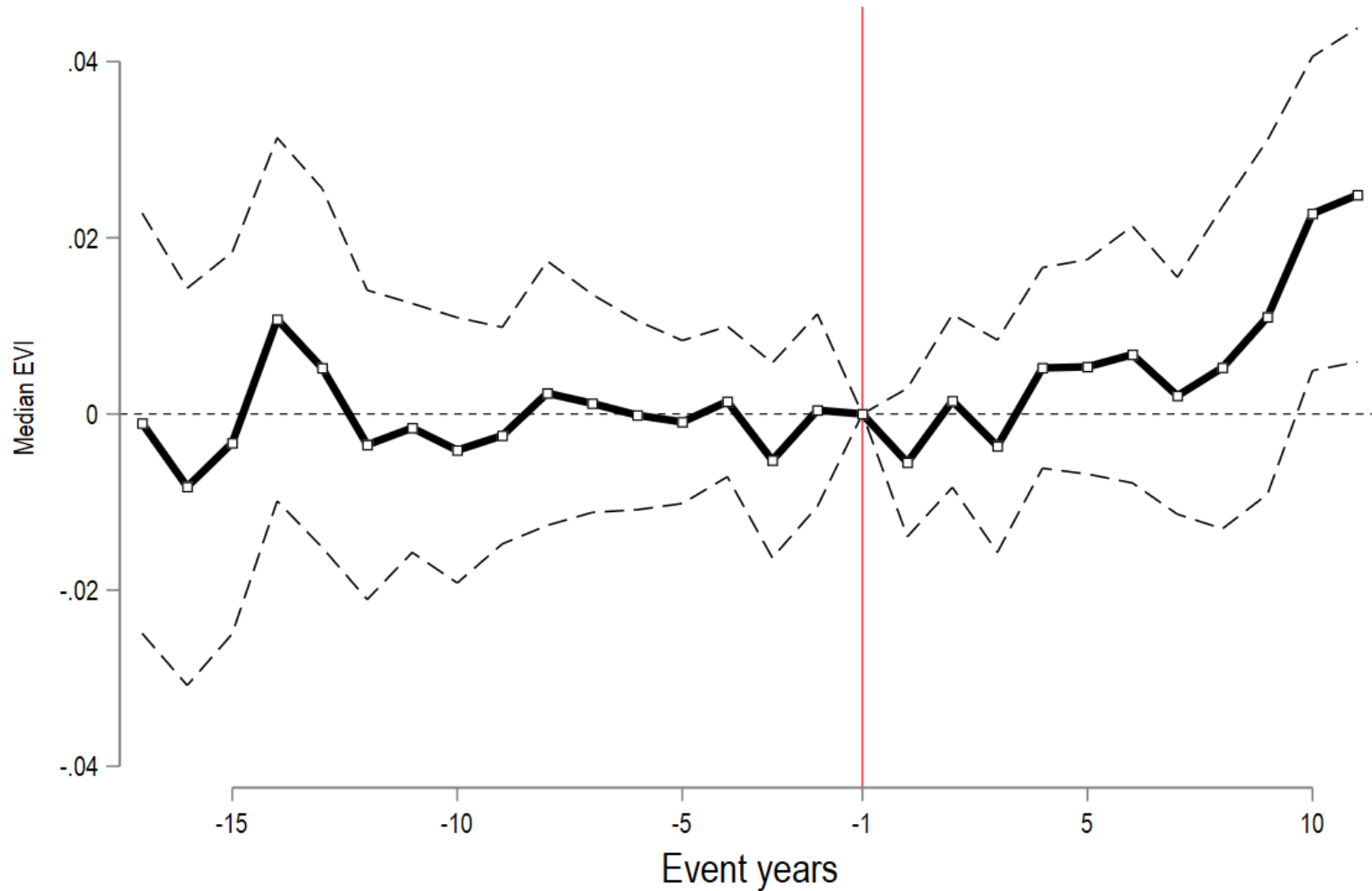
$$EVI_{dt} = \beta \left(1_{dt}^{2010} \cdot 1_{dt}^{flood} \right) + \gamma 1_{dt}^{2010} + \omega 1_{dt}^{flood} + \delta rice_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$

- Two-Way Fixed Effect (TWFE)

$$EVI_{dt} = \beta (seed_{dt} \cdot flood_{dt}) + \gamma seed_{dt} + \omega flood_{dt} + \delta rice_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$

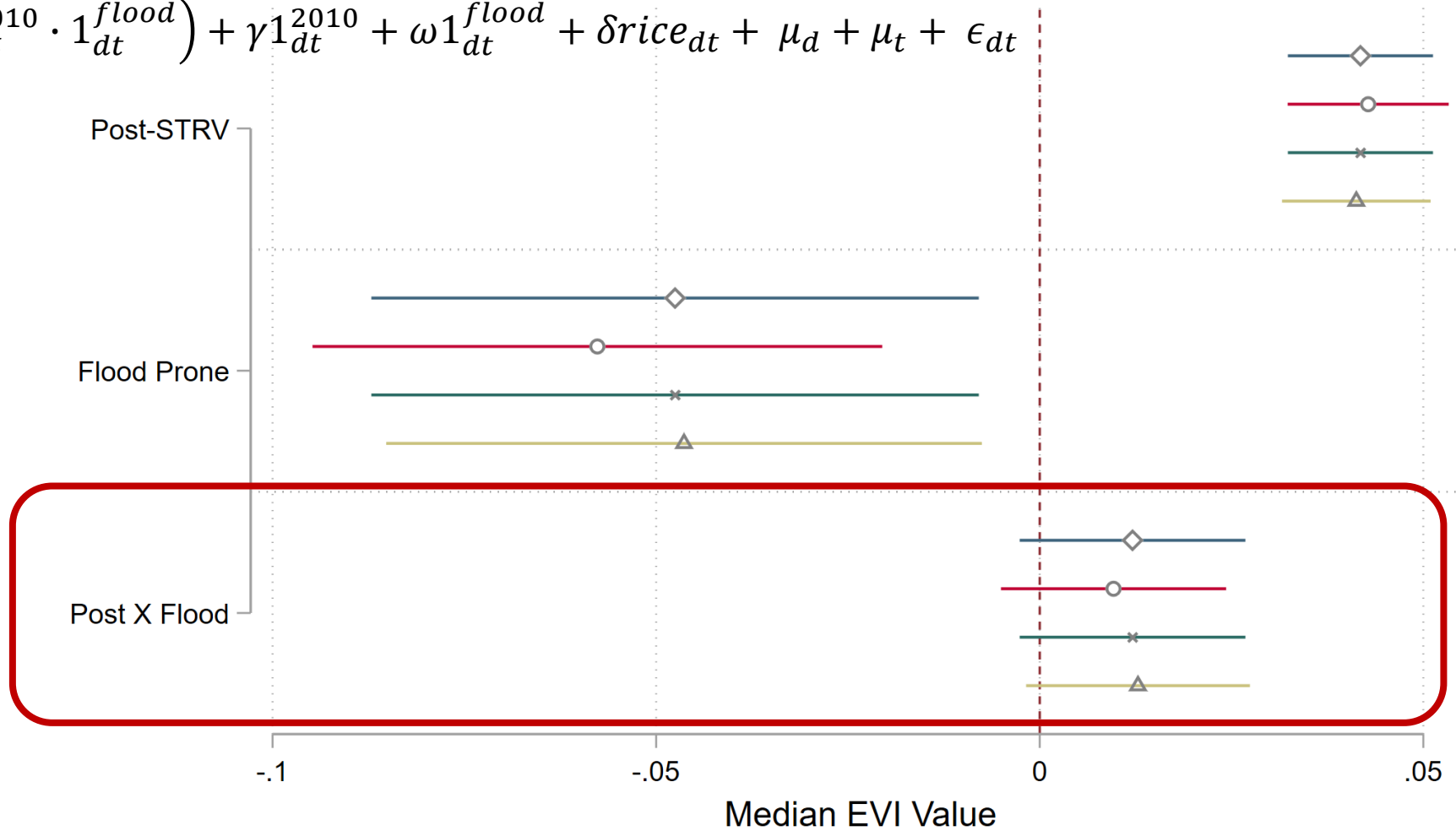
Event Study

$$EVI_{dt} = \sum_{j \in T} \alpha_j \cdot 1_t^{t=\tau+j} + \delta rice_{dt} + \varphi FLOOD_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$



Difference-in-Difference

$$EVI_{dt} = \beta \left(1_{dt}^{2010} \cdot 1_{dt}^{flood} \right) + \gamma 1_{dt}^{2010} + \omega 1_{dt}^{flood} + \delta rice_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$



◇ Cumulative ○ Max × Mean △ Median

Two-Way Fixed Effect

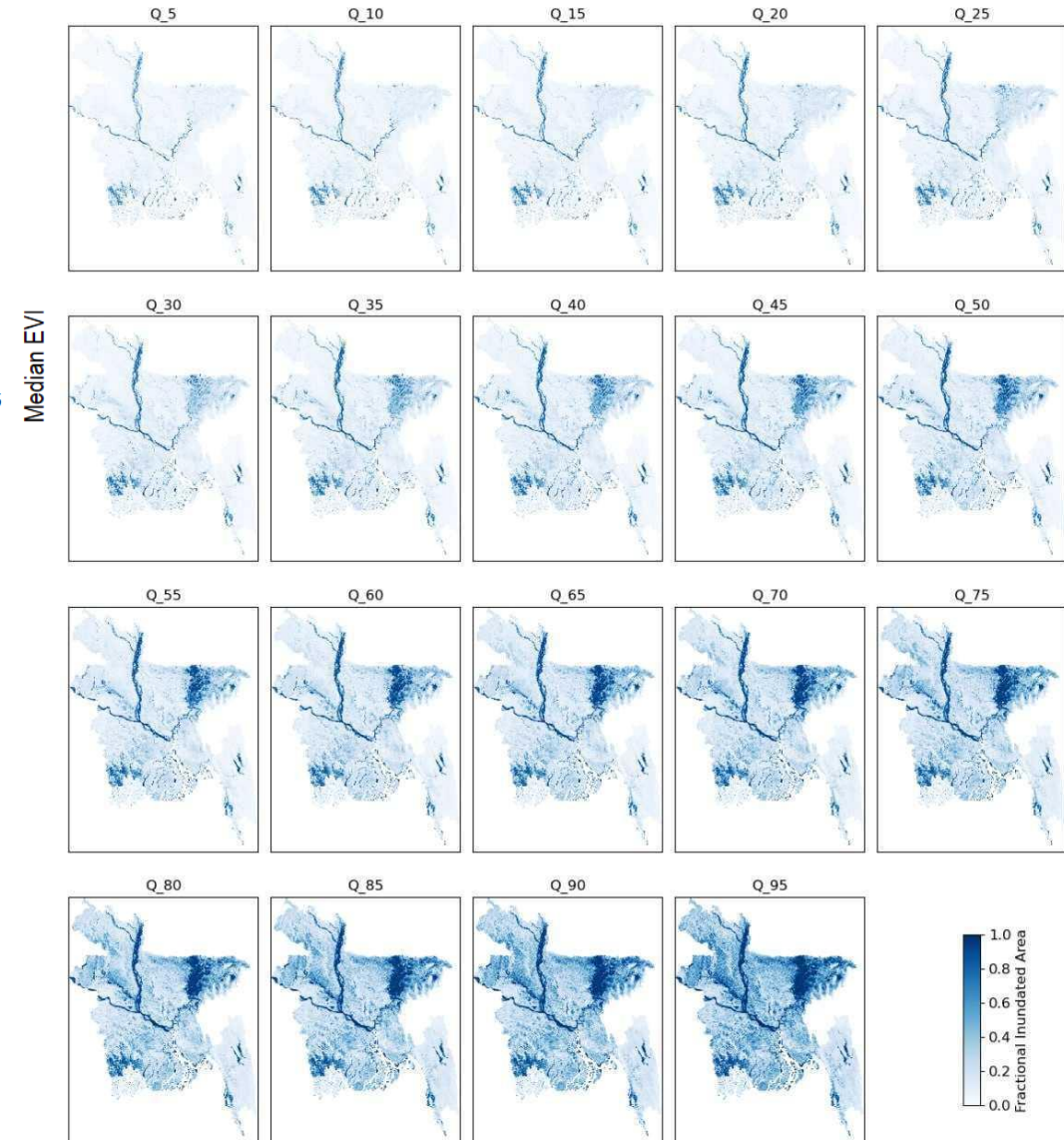
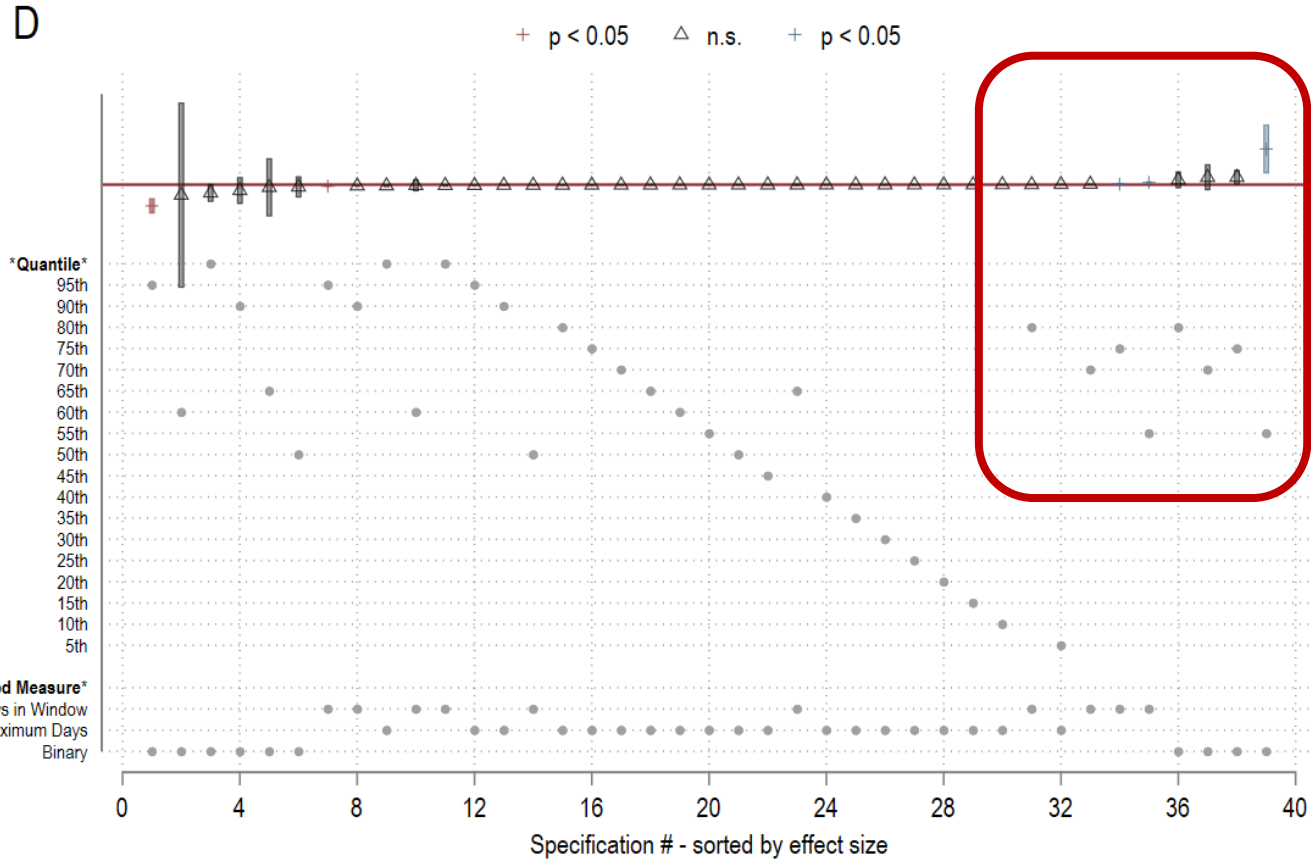
$$EVI_{dt} = \beta(seed_{dt} \cdot flood_{dt}) + \gamma seed_{dt} + \omega flood_{dt} + \delta rice_{dt} + \mu_d + \mu_t + \epsilon_{dt}$$

| | Cumulative | Max | Mean | Median |
|-----------------------|---------------------------|-----------------------------|------------------------------|------------------------------|
| STRV Availability | 0.139 [0.07,0.20] | 0.034 [-0.15,0.22] | 0.139 [0.07,0.20] | 0.111 [0.08,0.14] |
| Flood | -153.0 [-193.0,-113.0] | -1710.2 [-2424.1,-996.4] | -3518.5 [-4438.4,-2598.6] | -2975.1 [-3659.2,-2291.0] |
| STRV × Flood | -0.022 [-0.05,0.01] | 0.103 [-0.47,0.68] | -0.511 [-1.23,0.21] | -0.331 [-0.82,0.16] |
| Observations | 1,280 | 1,280 | 1,280 | 1,280 |
| Districts | 64 | 64 | 64 | 64 |
| Within R ² | 0.686 | 0.642 | 0.686 | 0.678 |

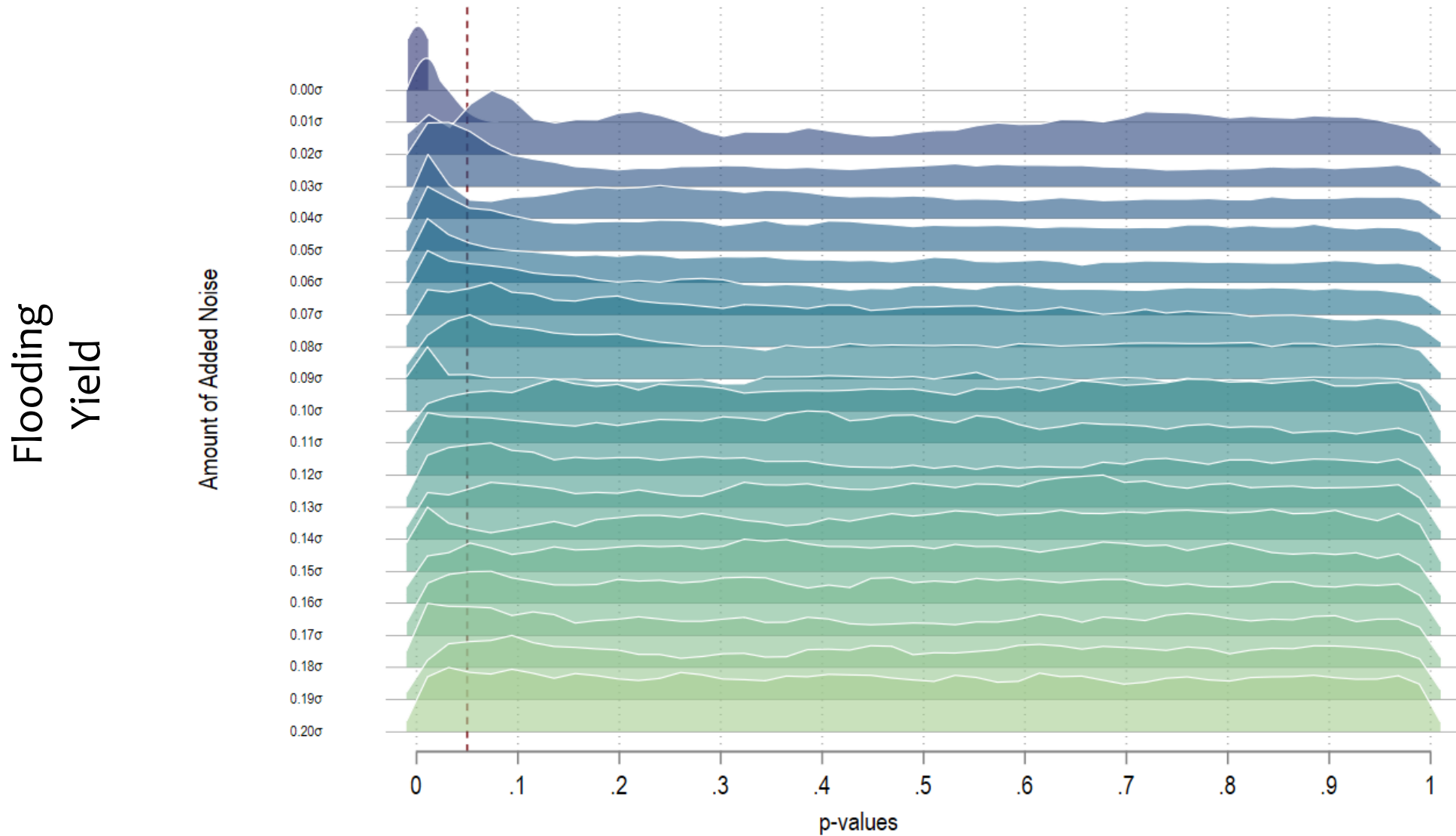
Moving Beyond the Pre-Analysis Plan

- Why are we unable to replicate Emerick et al. (2016) or field trials regarding the efficacy of STRVs?
 - STRVs are not effective outside of controlled settings
 - Insufficient adoption to pick up signal
 - Insufficient flooding of right type to pick up signal
 - Signal is there but EO methods need more refinement to separate signal from noise
- Refining our methods beyond what we pre-specified
 - Improvements on measuring flood days to get at “sweet spot” of flooding
 - Monte Carlo simulations to determine how sensitive results are to mismeasurement/misclassification

Refining Flood Measures



Monte Carlo Simulations – Dar et al. (2013)



Conclusion

Findings

- Flood causes a substantial amount of yield loss
- Dissemination and adoption of STRVs is increasing. Adoption rate was ~8% in 2014 and rose to ~22% in 2022
- Unable to replicate evidence that STRVs reduce yield loss during a flood

Challenges

- New ML methods allow for use of EO data in previously unexplored contexts
- Still a limit to the method's ability to separate signal from noise
- For stochastic technologies, results can be highly sensitive to mismeasurement and misclassification

QUESTIONS ??