Assessing the impact of farm ponds on agricultural productivity in Northern India

RAMNEEK KAUR, Indian Institute of Technology Delhi, India
KSHITIZ BANSAL, Indian Institute of Technology Delhi, India
DEVANG GARG, Indian Institute of Technology Delhi, India
RAMITA SARDANA, Indian Institute of Technology Delhi, India
SAKETH VISHNUBHATLA, Indian Institute of Technology Delhi, India
SANJALI AGRAWAL, Indian Institute of Technology Delhi, India
SHRUTI KUMARI, Indian Institute of Technology Delhi, India
PARAG SINGLA, Indian Institute of Technology Delhi, India
AADITESHWAR SETH, Indian Institute of Technology Delhi, India

Government welfare schemes such as the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) in India fund the creation of assets for natural resource management in rural villages to support farmers for their agricultural and livelihoods-based needs. With most agriculture in India being rain-fed, structures such as farm ponds, checkdams, trenches and bunds play a crucial role in supporting groundwater recharge and providing critical lifesaving irrigation in times of dry spells and droughts. In this study, we investigate the impact of farm ponds built under the MGNREGA scheme in Northern India as a source of protective irrigation for cropping areas in their immediate neighbourhood. We assess the impact of farm ponds on the following aspects: (i) we study their impact on agricultural productivity for up to five years since their construction, (ii) we separately study their impact in drought years during this period, (iii) we study the extent to which they are able to to reduce the sensitivity to droughts of sites having farm ponds. A causal analysis framework was designed by identifying control sites that did not have farm ponds, and the treatment effect of having farm ponds was computed using the difference-in-differences approach. Remote sensing data was processed to compute changes in vegetation indices around the treated and control locations before and after the construction of farm ponds. Our results indicate that farm ponds were instrumental in improving the agricultural productivity during the monsoon season in general. The impact during the monsoon season in drought years is also positive and significant. Furthermore, farm ponds also facilitated in reducing drought sensitivity during the monsoon season. The impact during the post-monsoon season was found to be lower, and the impact during the summer agricultural season was found to be the least.

CCS Concepts: • **Applied computing → Environmental sciences**; • **Computing methodologies → Model development and analysis**.

ACM Reference Format:
Ramneek Kaur, Kshitiz Bansal, Devang Garg, Ramita Sardana, Saketh Vishnubhatla, Sanjali Agrawal, Shruti Kumari, Parag Singla, and Aaditeshwar Seth. 2024. Assessing the impact of farm ponds on agricultural productivity in Northern India. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS '24), July 8–11, 2024, New Delhi, India.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3674829.3675085

---

1 INTRODUCTION

With 70% of India's rural population being primarily dependent on agriculture for livelihood and 82% of the farmers being small and marginal [13], the availability of water for irrigation becomes critical to ensure sustainable incomes and economic growth of farmers. To this end, several government welfare schemes, such as the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), Prime Minister Krishi Sinchayee Yojana (PMKSY), and Integrated Watershed Management Programme (IWMP), fund creation of assets for Natural Resource Management (NRM) in the rural villages of India. Different types of structures are built for conservation and management of water to support the farmers for their agricultural and livelihoods-based needs. In this work, we study the impact created by NRM structures built under MGNREGA, one of the largest demand-driven work guarantee programmes in the world. The scheme supports creation of assets under different NRM categories such as bunds, check dams and trenches for groundwater recharge, and farm ponds and wells for protective irrigation. Specifically, we evaluate the impact of farm ponds as a source of protective irrigation in their local proximity, i.e., in the immediate neighbourhood of the farm ponds.

Most agriculture in India is rainfed, and therefore, crop health is impacted if there is water stress when temperatures are very high or when there are extended dry spells. Farm ponds can help during such times through protective irrigation of crops during their critical growth stages. Further, even during non-drought conditions, they may help with productivity. In this study, we investigate the impact of farm ponds on seasonal agricultural productivity and support during drought and non-drought years. We consider the standard definition for seasons in a hydrological year (July-June) in India, where the months July, August, September, October fall under the Kharif season (the monsoon season); the months November, December, January, February fall under season Rabi (the post-monsoon season); and the months March, April, May, June fall under the Zaid season (the summer season), in most parts of India. The term cropping intensity refers to the number of crops grown in a hydrological year, i.e., cropping intensity values of one, two and three indicate that only a single crop was grown in a year (rainfed, during the Kharif season), or double crops were grown (during the Kharif and Rabi seasons), or triple crops were grown (Kharif, Rabi, Zaid), respectively.

Evaluating the impact of NRM structures would facilitate planning of prospective assets in a more scientific way through identification of factors that influence their impact. There is evidence that not all structures serve the intended purpose and that there are cases of failure. A similar observation was made by our team during our field visits. For example, during a visit to a community in the Mohanpur block in the Gaya district, Bihar, we saw that a large farm pond had been rendered useless due to poor site selection or inadequate information. The pond could not retain water because of the presence of a 5 km long lineament beneath its surface (Figure 1a). We also found some instances of faulty construction which limits the impact of these ponds (Figure 1b). Contrary to this, we also met farmers who had benefited significantly from the ponds. These farmers reported an increase in crop yield through protective irrigation of crops from the farm ponds during their critical growth stages, and in some cases an increase in cropping intensity from single to double cropping, and that from double to triple cropping. Hence, understanding the impact of these structures is important to assess the planning undertaken for these initiatives.

Traditionally, hydrologists and agricultural economists have been assessing the impact of NRM structures on outcomes such as crop yield, groundwater recharge, and rural livelihoods primarily through localized field experiments, surveys and focused group discussions [21, 30, 34, 37]. Such techniques are limited to small geographical areas ranging from a village to a block since large scale measurements using these techniques can be challenging. Consequently, such methodologies also cannot be used to study the heterogeneity in treatment effects with varying spatial characteristics.

(a) (b)

Fig. 1. (a) Farm pond built over a lineament. Lineaments are faults in the Earth's crust that can aid water infiltration to deeper aquifers, thus lowering the water retention capacity. (b) A poorly built farm pond, on a flat stretch of land with no inlets or outlets. Sections of boundaries were incomplete.

The availability of satellite imagery at high spatial resolutions has enabled the detection of changes in land use and land cover, crop yield, cropping intensity, crop phenology, etc. [7, 17, 19, 22]. Our work uses observational data derived from satellite imagery and other secondary data sources to study the impact of farm ponds built under MGNREGA on agricultural productivity. So far, there are only a few remote sensing based monitoring methods for impact assessment of NRM interventions in India [24, 36]. Moreover, these existing works are based on baseline-endline monitoring of outcome indicators, and do not compute outcome changes that can specifically be attributed to NRM assets.

For example, a study on impact monitoring of water structures [24] proposes a new index for monitoring the impact. The proposed index is based on existing remote sensing based indices which the authors infer as a better proxy for crop yield, based on statistical models built using data on crop productivity. The work includes a case study for the state of Maharashtra, wherein the district with the highest number of MGNREGA works is selected as the treated unit, and likewise, the district with least number of MGNREGA works is selected as the control unit. The proposed composite index is computed at the district level, and the outcomes for two time periods for the treated and control districts are compared to check if a positive difference in number of MGNREGA works led to a positive difference in the computed index values. The case study is based on the assumption that districts within the same agroclimatic zone would have similar climatic and geological conditions. Although this assumption seems reasonable, from our computations we found that even within the same agroclimatic zone, the climatic and geological variables exhibit variations that are quite diverse and may not represent the precise treatment effect of the interventions. For example, Table 1 depicts the variation in the terrain and Rabi precipitation for two districts of Bihar that fall in Agroclimatic Zone III, Bhagalpur and Sheikhpura. These statistics show that even within a district the terrain and seasonal precipitation can vary significantly, and that the precise location of a structure is an important factor to take into consideration in order to compute its effect.

Another study assesses the impact of MGNREGA dug wells using an integrated approach of field surveys and remote sensing data [36]. Micro-watersheds in six Gram Panchayats were randomly sampled from within the Ratlam district of Madhya Pradesh as treated units. Outcomes for two time periods of the six treated units were compared to assess the impact of the dug wells.

Table 1. Table depicting variation in the terrain and climatic conditions for two districts in the same agroclimatic zone

District	**Elevation (m)**				**Slope (degrees)**				**Rabi precipitation (mm)**			
	Min	**Max**	**Mean**	σ	**Min**	**Max**	**Mean**	σ	**Min**	**Max**	**Mean**	σ
Baghalpur	29.82	75.83	38.76	6.65	0.72	9.22	1.75	0.97	18.32	78.91	34.31	33.66
Sheikhpura	52.36	57.55	54.80	1.56	1.11	2.79	1.71	0.39	21.90	253.02	131.30	105.12

Differently from these studies, we perform a more fine-grained analysis wherein, (i) we track outcome changes across two time periods at each treated location (where a farm pond was constructed) in the study area, (ii) we sample control points as locations where no farm ponds were constructed; further, by taking into account a large enough buffer around other interventions to select control locations, we ensure that these control locations are not influenced by other works in their vicinity, (iii) the matching of control points to treatment points is done by considering an exhaustive list of covariates that affect the outcome variables, (iv) statistical tests are performed to establish causality.

The objective of this study is to build scientific evidence of the impact created by farm ponds built under MGNREGA, which would further facilitate the sanctioning of such structures through government schemes in the future, better planning and implementation of prospective structures, more efficient use of government funds, and higher productivity and thus better livelihoods for farmers. In particular, the contributions of this work are to:

- identify indicators that can be used to assess the impact of the farm ponds
- quantify the impact of farm ponds in terms of change in the selected outcome indicators that can be attributed to the ponds

2 RESEARCH QUESTIONS

The research question that we investigate in this study pertains to the impact of farm ponds on agricultural productivity. Our results for this question indicate that the creation of farm ponds led to an increase in productivity. Going further, we study if this increase in productivity is significant specifically during drought years. The results in this case show that creation of farm ponds led to an increase in productivity during drought years as well. Finally, we delve deeper to understand the impact of farm ponds on sensitivity to droughts. Reducing sensitivity to droughts is an important goal for farmers because it implies steady productivity and resilient livelihood, even during drought conditions. To summarize, we aim to answer the following research questions:

- RQ-1a: Do farm ponds improve productivity when compared with cropping locations without farm ponds?
- RQ-1b: Is such productivity increase specifically significant during drought years?
- RQ-2: Are sites with farm ponds less sensitive to droughts when compared with cropping locations without farm ponds?

For RQ-1a and RQ-1b, positive Average Treatment Effect (ATE) values would indicate a positive impact, implying that the improvement in outcome at treated sites between post-construction and pre-construction years is more than the corresponding improvement at the control sites. Our results for RQ-1a and RQ-1b showed a positive and statistically significant impact of farm ponds in the Kharif season. The impact in Rabi was negative for RQ-1a, and positive for RQ-1b. The impact in Zaid was found to be negative for both RQ-1a and RQ-1b. Although we consider a comprehensive list of covariates in our causal analysis framework, negative ATE values in the Rabi and Zaid seasons are likely due to unobserved covariates such as the source of irrigation (whether a site is supplemented with borewell based groundwater irrigation or not). Sites supported with farm ponds under MGNREGA are likely to belong to smallholder farmers reliant on rainfed agriculture, while control sites may belong to farmers with access to borewell based irrigation. Such control sites are thus more likely to be cropped during the Rabi and Zaid seasons and would not experience as much water stress as compared to the treatment sites which are just reliant on farm ponds, thus resulting in a negative treatment effect during these seasons. During the Kharif season, however, our results validate that farm ponds are able to provide greater benefit than borewell based irrigation.

For RQ-2, to understand the impact of farm ponds on sensitivity to drought, we compute the difference in outcomes between non-drought and drought years. A negative ATE value would indicate a positive impact (less sensitivity to drought), implying that the difference in outcome between normal years and drought years is lower at the treated sites than the corresponding difference at the control sites. Our results for RQ-2 show a positive impact until three weeks of drought incidence for several outcome variables. Beyond three weeks, a negative impact was observed, indicating that farm ponds can reduce sensitivity up to a point, but in cases of more severe drought they are insufficient as compared to control sites which may have access to borewell based irrigation.

3 STUDY AREA

The National Bureau of Soil Survey & Land Use Planning (NBSSI&LUP) developed twenty agro-ecological zones [29] based on an integrated criteria of seasonal rainfall, soil groups, length of growing periods, and delineated boundaries adjusted to district boundaries (refer Figure 2).

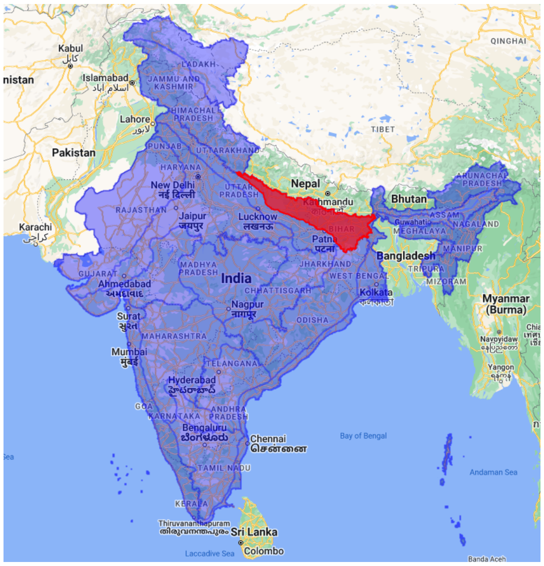


Fig. 2. Agro-ecological zones of India. AEZ 13 is shaded in red.

Agro-ecological Zone (AEZ) 13 (shaded in red in Figure 2), comprising the eastern plains over north-eastern Uttar Pradesh, northern Bihar, and some areas of the Himalayan foothills, was chosen as the study area.

There are 38 districts in this AEZ, spanning an area of approximately 10 million hectares. Our analysis is done on all districts having their centroids inside the AEZ 13 boundary. The climate in this region is characterized by hot and wet summers, along with cool and dry winters. The area receives an average annual rainfall of 1400 to 1800 mm.

The soils in the region consist mainly of level to very gently sloping alluvium-derived soils. Rainfed agriculture is predominant in the area, with crops such as rice, maize, pigeonpea and moong commonly cultivated during the Kharif season. In the Rabi (post-monsoon) season, crops like wheat, lentil, pea, sesamum and occasionally groundnut are grown, relying on residual soil moisture supplemented by one or two protective irrigations during the critical growth stages.

Historically, there has been limited irrigation infrastructure in the area, and there is an opportunity for improvement through surface water sources, as well as groundwater irrigation methods.

4 METHODOLOGY

We design a causal analysis framework to compute the treatment effect of farm ponds on the selected outcome indicators. The treatment effect is defined as the change in outcome that can be attributed to the intervention. The variables in our framework are as follows:

- Treatment variable: a binary variable that indicates the presence/absence of the intervention, i.e., a farm pond
- Confounding variables/covariates: variables that influence both the treatment and the outcome variables
- Outcome variables: indicators selected to quantify the impact of the interventions

To compute the causal effect of the treatment on the outcome, we employ the Difference-in-Differences (DiD) methodology [40]. The DiD method ensures unconfoundedness by comparing the difference in outcome at the treated sites to the difference in outcome at 'counterfactual sites'[1]. Counterfactual sites refer to locations that did not receive treatment but share similar characteristics with the treated sites in terms of the confounding features/covariates. Propensity score matching is used to match the treated sites to their counterfactual sites. The propensity score is the predicted probability of receiving treatment based on the confounding variables or covariates. Treated sites/assets are matched to counterfactual sites with similar propensity score values (illustrated in Figure 3). The Average Treatment Effect (ATE) is then defined as shown in Equation 1, where $\overline{Difference}_{asset}$ and $\overline{Difference}_{matched\,counterfactual}$ represent the means of differences in outcomes corresponding to the assets and their matched counterfactuals, respectively. The difference in outcome for our study corresponds to the difference in the selected indicators between two time periods, a base year and a target year ($Y_{target\ year} - Y_{base\ year}$, where Y_t represents an aggregated indicator value over a specific time period t).

$$ATE = \overline{Difference}_{asset} - \overline{Difference}_{matched\,counterfactual} \tag{1}$$

4.1 Confounding variables/covariates

We try to minimise the confounding bias in the treatment effect estimates by considering a comprehensive list of covariates (refer to Table 3) which are expected to influence both treatment assignment (site selection for farm ponds in this case) and the selected outcome indicators. The initial selection of covariates for the analysis was done based on our understanding of the domain and upon consulting several implementation organizations to understand the factors they

---

[1] We perform several checks to ensure that the identifying assumptions in causal inference hold true in our analysis, the details of which are documented in our supplementary material: https://bit.ly/3UKlrqn

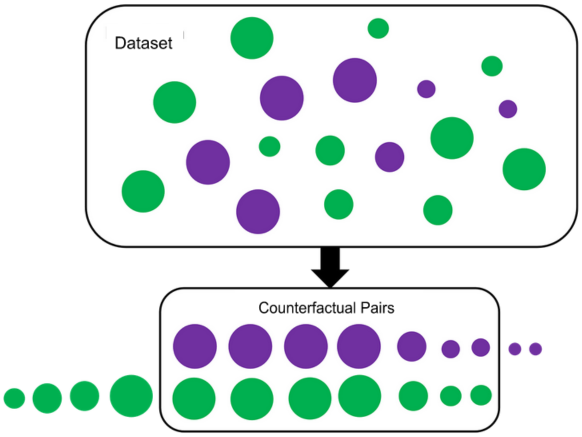


Fig. 3. Propensity score matching illustration

take into consideration when planning farm ponds. We also referred to related studies that employ the DiD approach to assess the effectiveness of environmental policies, for selection of covariates for our study [5, 33, 38, 39]. Standard normalization techniques were used to bring the selected covariates to a uniform scale and normalise the underlying distributions.

Table 2. List of covariates with their description

Variable	**Description**
Elevation (m)	Distance above sea level
Slope (degrees)	Degree of inclination of the topography
Dist closest river (m)	Linear distance of each pixel to the closest river
Dist closest lineament (m)	Linear distance of each pixel to the closest lineament
Dist closest road (m)	Linear distance of each pixel to the closest road
Closest upstream forest (m)	Linear distance of each pixel to closest upstream forest
Proximity water (m)	Linear distance of each pixel to water pixels in vicinity
Drainage density	Length of all water streams divided by area of basin
Flow accumulation	Amount of upslope pixels that flow into each pixel
HSG	Soil classification based on soil's runoff potential
Soil OC	Pixel-level carbon component of organic matter in soil
Soil CEC	Pixel-level measure of negative charges in soil
Soil pH	Pixel-level measure of acidity or basicity of soil
Drought freq (base yr)	Number of drought weeks incident in base year
Drought freq (target yr)	Number of drought weeks incident in target year
Seasonal rain (base yr)	Seasonal precipitation of base year
Seasonal rain (target yr)	Seasonal precipitation of target year

4.2 Outcome variables

An ideal outcome variable to choose would have been a crop yield estimate. Crop yields are influenced by various factors such as the soil quality (nutrient content, pH, texture) and weather (temperature, rainfall, humidity, wind speed),

Table 3. List of covariates with their primary source, normalization technique, mean and standard deviation (σ)

Variable	**Primary source**	**Normalization technique**	**Mean**	σ
Elevation (m)	SRTM DEM	Standard Scalar	90.493	60.657
Slope (degrees)	SRTM DEM	Standard Scalar	2.283	1.863
Distance to closest river (m)	India-WRIS	Power transform	2206.758	2668.179
Distance to closest lineament (m)	Bhuvan	Power transform	2.157e5	4.774e5
Distance to closest road (m)	Geofabrik	Power transform	966.053	1284.339
Distance to closest upstream forest (m)	Dynamic World	Power transform	177.609	5015.719
Proximity to water (m)	Dynamic World	Power transform	28.369	52.488
Drainage density	SRTM DEM	Standard Scalar	3.987e-3	1.148e-3
Flow accumulation	SRTM DEM	Power transform	1.361e3	2.656e4
Hydrological Soil Group	ORNL DAAC	–	–	–
Soil Organic Carbon	HWSD	Power transform	0.766	0.191
Soil cation exchange capacity	HWSD	Power transform	13.087	4.299
Soil pH	HWSD	Power transform	6.911	0.695
Drought frequency of base year	JAXA-GPM	Power transform	–	–
Drought frequency of target year	JAXA-GPM	Power transform	–	–
Seasonal precipitation of base year (mm)	JAXA-GPM	Power transform	647.431	162.177
Seasonal precipitation of target year (mm)	JAXA-GPM	Power transform	624.596	138.923

of which the presence of farm ponds would specifically help counter water stress during low rainfall periods or dry spells. However, the crop species being grown at a particular location, and their precise crop models for yield estimation, are not known. We therefore proxy the crop yield using vegetation indexes. Vegetation indices computed based on remote sensing can be used to derive both quantitative and qualitative assessments of vegetation cover [19].

We consider vegetation indices such as NDVI and GCI (which reflect the greenness of vegetation) as a proxy for crop yield, and NDMI as a proxy for crop water stress. These indices are derived from remote sensing data products, the details of which are presented in Section 5.

- **Normalised Difference Vegetation Index (NDVI):** NDVI is a vegetation index commonly used for tracking crop development dynamics since it measures the photosynthetically active plant biomass. It acts as a direct proxy for plant greenness and, therefore, the crop yield. It is computed using the near-infrared (NIR) and RED spectral bands as follows:
$$\text{NDVI} = \frac{\text{NIR - RED}}{\text{NIR + RED}}$$
- **Green Chlorophyll Index (GCI):** This index calculates the total chlorophyll content of the vegetation. It is sensitive to minor variations in the chlorophyll content and is consistent across various crop species. It is computed using the NIR and GREEN spectral bands as follows:
$$\text{GCI} = \frac{\text{NIR}}{\text{GREEN}} - 1$$

Normalised Difference Moisture Index (NDMI): This index detects moisture levels in vegetation using a combination of NIR and short-wave infrared (SWIR) spectral bands, making it a reliable indicator of water stress in crops. It is computed as follows:

$$\text{NDMI} = \frac{\text{NIR - SWIR}}{\text{NIR + SWIR}}$$

4.3 Field visits to locations in AEZ 13 and AEZ 9

As part of our research, to understand the indicators suitable for the analysis and to validate our preliminary results, we conducted field visits to locations in three districts of AEZ 13 and AEZ 9, viz. Gaya and Jamui in the state of Bihar, and Koderma in the state of Jharkhand. During these visits, we interacted with farmers having farm ponds that were constructed under MGNREGA. Our selection of indicators corresponding to crop yield and crop water stress to study the impact of farm ponds were validated through these farmer interactions. A meeting was also convened with the Programme Officer of MGNREGA to understand their data better. Insights from this interaction were used to process the MGNREGA data (presented in Section 5). Some case studies that strengthen our hypotheses and validate our empirical results are presented in Section 7.

5 EXPERIMENTAL EVALUATION

We implement our causal analysis framework by utilising the Google Earth Engine API [16] to access, process and visualize satellite imagery hosted on the platform. It offers several useful APIs and tools for visualization and analysis of geospatial datasets.

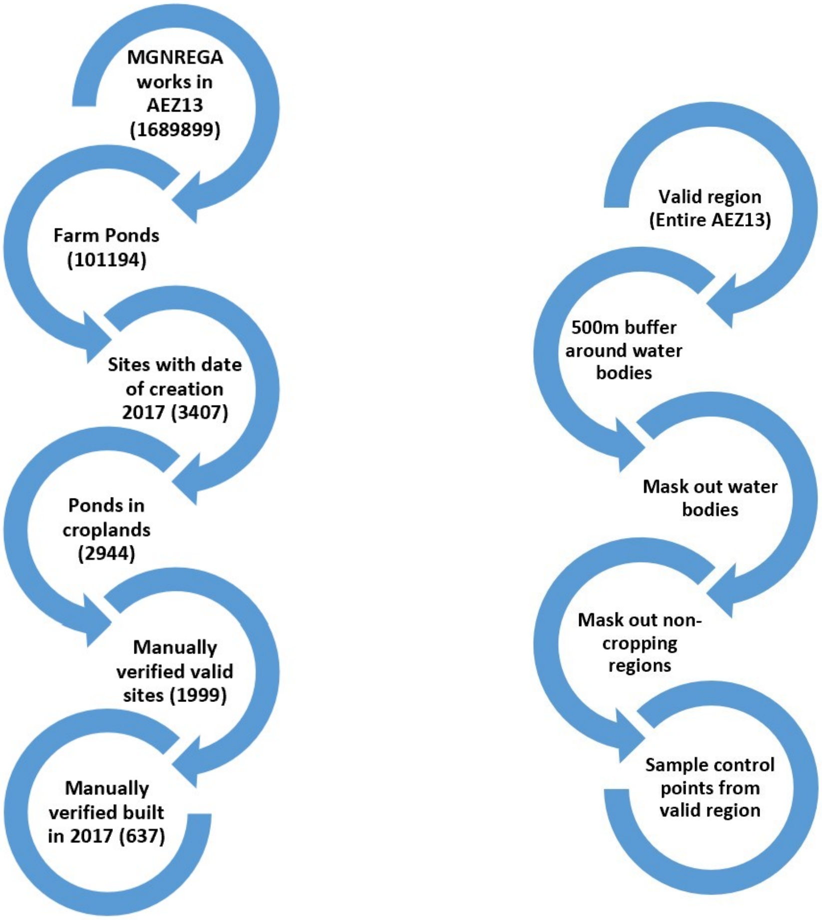
5.1 Identifying treated sites

MGNREGA publishes country-wide geo-tagged data of assets constructed under the scheme [18]. We crawled the data and obtained a total of 1.7 million works done in AEZ 13 until the year 2022. For our analysis, we filtered out the works corresponding to 'farm ponds' in the region (0.1 million in number). Further, we considered farm ponds built in the year 2017 and considered prior years for baseline and subsequent years as target years for the outcome analysis. The year 2017 was selected as the intervention year because we found that most farm ponds in the AEZ were built in this year. Further, we rely on Land Use and Land Cover (LULC) data products, Dynamic World [9] and IndiaSAT [7], to compute several covariates and these product are available only from 2016 onward. We use the 'asset creation date' in the MGNREGA data to filter out works done in 2017. We drop farm ponds with any missing or incomplete data columns.

This process finally gave us 3,407 farm ponds in AEZ 13 that were constructed in 2017. Further, since we are interested in studying the impact of farm ponds on cropping areas, we use the IndiaSAT LULC product to only include farm ponds situated in cropping areas [7]. This left us with 2,944 farm ponds.

A preliminary analysis of this dataset conducted by examining the locations on Google Earth Pro suggested the need for further manual cleaning, since we found that several entries did not have any farm ponds, some were duplicate entries for the same pond, and some entries had pre-existing farm ponds at the location (constructed before 2017). Additionally, there were geo-tagging errors where the tagged locations were offset from the actual locations by a distance ranging from a couple of meters to over 150 meters.

For his manual cleaning, we used the latitude-longitude data to visualize each pond using Google Earth Pro across the available historical imagery, allowing us to remove any multiple reporting for the same pond and update the latitude-longitude values to more accurate centroid locations of the ponds. After this manual cleaning, we were left with a dataset of 1,999 farm ponds that were worked upon (including both construction and maintenance works) under



(a) Treated Sites (b) Control Sites

Fig. 4. Identification of treated and control sites

the MGNREGA scheme in AEZ 13 in the year 2017. A subset of these ponds, 637 in number, were built from scratch in the year 2017. The complete procedure for identification of treated points is shown in Figure 4.

5.2 Identifying control sites

To ensure that the observed impact is attributable to the NRM works and not influenced by external factors, it is crucial to identify suitable comparison locations that share similar characteristics as the asset locations. These sites represent a scenario that would have occurred if no treatment had been given, thus allowing an estimation of the causal effect of farm ponds by comparing the observed outcomes of treatment locations with no-treatment locations. To achieve this, we build a matching method to identify no-treatment locations.

We first mask out all non-cropping areas using the IndiaSAT LULC product. We then further mask out areas close to natural water bodies and other NRM interventions. From our field visits, we found that NRM interventions and natural water bodies substantially influence their surrounding area. To ensure that there is no effect of existing NRM works and natural water bodies on the outcome indicators of control points, we created a 500-meter buffer around such works and water bodies, and masked out these regions. The detection of natural water bodies was done using IndiaSAT. For NRM works in the region, we considered all water related NRM works ever done under MGNREGA identified through their 'Work Type', including specific keywords determined corresponding to ponds, wells, tanks, bunds, trenches, etc.

From the remaining area, we randomly sample control points that are three times the number of treatment points, to allow for a robust matching using the propensity score method described next.

5.3 Computing covariates

We compute the following covariates for propensity score matching.

Topography based (Elevation, Slope, Drainage density, Flow accumulation): These covariates were computed using the SRTM DEM (Shuttle Radar Topography Mission's Digital Elevation Model) data which has a resolution of 1 arc-second (30m) [26]. The corresponding covariate value for a farm pond/control site is taken as the mean of values in a 100 meters buffer around it's coordinates.

Distance to closest river, lineament, road, upstream forest: These covariates were computed using data from India-WRIS (Water Resources Information System) [28] for distance to closest river; Bhuvan [18] for distance to closest lineament; Geofabrik [14] for distance to closest road; and Dynamic World [17] for distance to closest upstream forest. The corresponding covariate was computed at pixel level, and then the linear distance between the site coordinates and nearest pixel of interest was taken.

Soil based (HSG, OC, CEC, pH): Data from ORNL DAAC (Oak Ridge National Laboratory Distributed Active Archive Center) [25] was used to compute the Hydrological Soil Groups (HSG). The Harmonized World Soil Database (HWSD) [12] was used to compute the Organic Carbon (OC) content, Cation Exchange Capacity (CEC), and soil pH. The corresponding covariate value is taken as the mean of values in a 100 meters buffer around the given coordinates.

Drought frequency: The drought frequency of the base and target year covariates represent the number of weeks for which there was either a moderate or severe drought incident in the Kharif season in that particular year. The classification of each week as non-drought vs. mild/moderate/severe drought was done using several metrics such as the duration of dry spells, Standardised Precipitation Index, aridity indicators like the ratio of evapotranspiration to the potential evapotranspiration, cropping area sown, and so on. These calculations were performed based on methods standardised by the Government of India in the Manual for Drought Management [23].

Seasonal precipitation: We compute this covariate at the watershed level using the JAXA-GPM (Japan Aerospace Exploration Agency - Global Precipitation Measurement) [3] data. A watershed is a drainage basin or a catchment area from which runoff resulting from rainfall is collected and drained through a single common point [15].

5.4 Computing outcome variables

We use the Landsat 7 and Landsat 8 [27] multispectral imagery covering our study area across multiple time periods to capture the seasonal variations in the selected vegetation indices.

Temporal aggregation: We segment the Landsat imagery into different time periods, representing the distinct seasons of Kharif, Rabi, and Zaid, over the years. This allows us to analyze the vegetation dynamics across various seasons. For each Landsat image within each time period, we calculate the outcome variables (NDVI, GCI, NDMI) using the corresponding bands and methods. Studies show that both mean and maximum values of indices during the vegetation period are good predictors of crop yields [11]. Therefore, to aggregate the time series values of the indices within a season we consider the temporal means of the indices.

Spatial aggregation: During the field visits, we found that farmers from nearby fields use pumps to draw water from the farm ponds. The farmers reported that each pond typically helps provide irrigation in two to three hectares, i.e., in a 100 meters radius. Therefore, we create a buffer of 100 meters around each farm pond and control site and aggregate the outcome variables in this buffered neighbourhood. Within each buffered zone, we calculate the spatial mean of NDVI, GCI, and NDMI values from the corresponding Landsat imagery for each time period. This spatial aggregation enables us to assess the average vegetation condition in cropping regions around the 100 meters neighbourhood of the treated and control locations.

5.5 Propensity score matching

Using the treated and control points dataset, a logistic regression-based classification model is trained. The model is then used to predict the propensity scores (probability of receiving treatment given the covariates) for the treated and control points. The points are then matched based on the k-nearest neighbors algorithm. We perform a one-to-one matching to form treatment-counterfactual pairs.

6 RESULTS AND DISCUSSION

In this section, we present results for the experiments conducted to answer the formulated research questions. We report the ATE for the DiD test and p-values on their distributions.

6.1 RQ-1a: Impact on productivity

In these experiments, we consider the farm ponds that were built in 2017 as the set of treated points. The base year (pre-construction year) for computing DiD is fixed to the Hydrological Year (HY) 2016-'17, and the target years considered (post-construction years) are HYs 2018-'19, 2019-'20, 2020-'21, 2021-'22 and 2022-'23. The outcome for each observation is $Seasonal\ NDVI_{target\ year} - Seasonal\ \ NDVI_{base\ year}$ for the indicator NDVI, and $Seasonal\ GCI_{target\ year} - Seasonal\ GCI_{base\ year}$ for the indicator GCI, which represents the difference in productivity before and after the construction of farms ponds. Table 4 shows the results for RQ-1a. Statistically significant results at $p < 0.05$ in the direction of improvement are highlighted in blue.

It can be seen that the treatment effect of farm ponds on productivity varies across seasons. The effect is positive and statistically significant in the Kharif season. A positive ATE in Kharif indicates that the productivity at the treated sites improved after creation of farm ponds. During Rabi and Zaid, negative ATE values were observed, likely due to unobserved covariates such as the use of borewells at control sites.

Numerous studies based on different regions for various crops such as wheat, rice and cereal, have shown a positive relationship between crop yield and the vegetation indices, NDVI and GCI [1, 2, 6, 8, 20]. One study modelled a linear relationship between NDVI and crop yield, along with other predictors such as Land Surface Temperature indicators, for different crops such as wheat, mercantile corn, soya bean and sugar beet for the region of Vojvodina, Serbia [10]. A study based in the Thanjavur district in Tamil Nadu, India, found an exponential relationship between rice yield and the maximum value in the seasonal NDVI time series [4]. A study on the relationship between cereal yield and NDVI for selected regions of Central Europe [31] found that an increase in NDVI by 0.1 in early spring increases grain yield by about 1.1 to 2.6 tonnes/hectare, with an average gain of 2 tonnes/hectare. As the average grain yield in the study area over the study period was 5 tonnes/hectare, it can be inferred that a 5% increase in NDVI in early spring led to a 40% increase in grain yield for this area. Findings from another study [35] on predicting wheat yield from remote sensing data for Delhi, India, showed that a 5% increase in NDVI leads to 20% increase in yield. Such statistics suggest that ATE values of 2% and 5% for our study area imply a significant increase in crop yield at the treated sites after the creation of farm ponds as compared to the corresponding increase at control sites.

Table 4. Results for RQ-1a

Outcome variable	Season	ATE	P-value
NDVI	Kharif	**0.020869**	**5.447e-27**
	Rabi	-0.0045099	0.00012
	Zaid	-0.008669	1.00483e-13
GCI	Kharif	**0.053639**	**2.13676e-24**
	Rabi	-0.020551	1.663e-9
	Zaid	-0.058739	3.72585 e-05

Overall, the results are coherent with findings from our field visit. Through interactions with farmers, we learnt that farm ponds have the highest effect in Kharif during which time they are used for protective irrigation of crops. The effect is smaller in Rabi as the water stored in the ponds evaporates or percolates into the ground. The effect in Zaid is almost not observed as the ponds are normally dry by that time of the year.

A negative ATE in the Rabi and Zaid seasons indicates that there is more improvement at the control sites during these seasons between the target and base years, than a corresponding improvement at the treated sites. We believe this to be due to unobserved covariates about farm access to other irrigation sources, that the control points identified by us are likely to consist of a mix of croplands with no irrigation sources and those using borewell irrigation. Farms with borewell access may be cropping during Zaid as well, and will show a higher productivity than rain-fed intervention areas which typically leverage farm ponds and do not crop during the summers. These observations are also indicative of valid site selection of farm ponds, that the ponds are indeed being constructed at sites with lack of alternative sources of irrigation.

6.2 RQ-1b: Impact on productivity during drought years

In RQ-1b, we go deeper, and study if the productivity impact seen in RQ-1a is significant specifically during drought years. Similar to the setting in RQ-1a, for RQ-1b we consider the farm ponds that were built in 2017 as the set of treated points. The sets of base and target HYs considered for computing DiD are {2016-'17} and {2018-'19, 2019-'20, 2020-'21, 2021-'22, 2022-'23}, respectively. The treated observations correspond to the treated sites that experienced a drought in the base year. Likewise, the control observations correspond to the control sites that experienced a drought in any of the target years (one observation for each drought incidence). A **frequency threshold** of five was set for classifying years as drought/non-drought years, meaning that years with number of moderate or severe drought weeks greater than or equal to five were classified as drought years, and others as non-drought years. This threshold was determined based on our conversations with field partners who spoke about extended dry spells of over a month as being detrimental to crop health. The outcome for each observation is $Seasonal\ NDVI_{target\ year} - Seasonal\ NDVI_{base\ year}$ for the indicator NDVI, and $Seasonal\ GCI_{target\ year} - Seasonal\ GCI_{base\ year}$ for the indicator GCI. The outcome in this case represents

Table 5. Results for RQ-1b

Outcome variable	**Season**	**ATE**	**P-value**
NDVI	Kharif	**0.031193**	**9.446777e-10**
	Rabi	0.003840	0.313421
	Zaid	-0.009606	0.000699
GCI	Kharif	**0.078515**	**1.861222e-06**
	Rabi	**0.025945**	**0.029126**
	Zaid	-0.024775	0.108344

the difference in productivity before and after the construction of farms ponds during drought years. Table 5 depicts the results for RQ-1b.

Similar to the results for RQ-1a, it can be seen that the treatment effect of farm ponds on productivity during drought years exhibits a seasonal variation. The effect is positive and statistically significant in the Kharif season for both NDVI and GCI indicators. A positive ATE in Kharif indicates that the productivity at the treated sites improved after creation of farm ponds during drought years as well, the magnitude of impact in drought years being higher than the impact in general.

During Rabi, the effect is positive for both NDVI and GCI, and also statistically significant for GCI. Negative ATE values were observed in Zaid, which can again be attributed to the stock of water in the ponds getting exhausted by summers. The control sites on the other hand are likely to be those with access to groundwater irrigation, and thus the cropping intensity at these sites might have increased over the years.

6.3 RQ-2: Impact on sensitivity to droughts

In RQ-2, we study if sites with farm ponds are less sensitive to droughts as compared to cropping locations without farm ponds. Reducing sensitivity to droughts is an important goal for farmers because it implies steady productivity and resilient livelihood, even during drought conditions. Unlike the previous RQs that consider the difference in outcomes post and pre-construction, for RQ-2, we consider the difference in outcomes between non-drought and drought years. We consider the dataset of ponds existing in 2017 as the set of treated points. The base years and target years considered are HYs 2018-'19, 2019-'20, 2020-'21, 2021-'22 and 2022-'23. The observations correspond to unique combinations of target and base years such that the corresponding sites experienced a non-drought (or normal) year in the target year, and a drought in the base year. The outcome for each observation is $Kharif\ NDVI_{target\ year} - Kharif\ NDVI_{base\ year}$ for the indicator NDVI; $Kharif\ GCI_{target\ year} - Kharif\ GCI_{base\ year}$ for the indicator GCI; and $Kharif\ NDMI_{target\ year} - Kharif\ NDMI_{base\ year}$ for the indicator NDMI. To investigate sensitivity to droughts, in addition to indicators NDVI and GCI, we also consider the NDMI indicator (proxy for crop water stress). Furthermore, we vary the frequency threshold parameter to study the effect of drought frequency on the outcome. The results for RQ-2 are shown in Table 6.

For indicators NDVI and GCI, the results are negative and significant until the drought frequency threshold of three. A negative ATE implies a positive impact as the observed gap in productivity at treated sites between normal years and drought years is smaller than the corresponding gap in productivity at the control sites. This supports our hypothesis that farm ponds facilitate in reducing sensitivity to droughts. The results are significant only for lower values of frequency threshold. If the drought persists beyond three weeks, the productivity gap at treated sites drops to values lower than that of control sites. Beyond three weeks of drought, it is likely that the stock of water in the ponds

Table 6. Results for RQ-2

Frequency threshold	**NDVI**		**GCI**		**NDMI**	
	ATE	**p-value**	**ATE**	**p-value**	**ATE**	**p-value**
1	**-0.009436**	**6.18e-16**	**-0.028417**	**5.18e-18**	**-0.011125**	**7.46e-45**
2	**-0.010234**	**5.74e-19**	**-0.030636**	**2.68e-21**	**-0.009389**	**4.00e-34**
3	**-0.007081**	**1.04e-09**	**-0.019066**	**5.13e-09**	**-0.008199**	**1.19e-24**
4	-0.000742	5.57e-01	0.000909	7.97e-01	**-0.009192**	**1.33e-25**
6	-0.000786	5.58e-01	-0.00259	4.87e-01	**-0.016931**	**1.58e-77**
8	0.000061	9.65e-01	-0.000622	8.72e-01	**-0.017001**	**2.09e-76**
10	0.001694	2.48e-01	0.004084	3.27e-01	**-0.011647**	**1.41e-28**
12	-0.003074	1.38e-01	-0.010042	9.19e-02	**-0.006062**	**8.83e-05**
15	-0.003791	5.66e-01	-0.013558	4.49e-01	**-0.028709**	**7.95e-08**

gets exhausted. The treated sites, due to their dependence on the ponds, might not have access to alternative sources for irrigation. The control sites on the other hand possibly have access to groundwater irrigation.

The results for NDMI on the other hand, are negative and significant for all ranges of the frequency threshold, indicating a consistently positive impact. This means that the farm ponds do facilitate in maintaining the moisture content of the crops during drought years in comparison with the normal years. However, beyond three weeks of drought incidence, the moisture content at the treated sites despite being greater than that at control sites, is still not enough to protect the crops (reflected in the NDVI and GCI outcomes) and there arises a need for supplemental irrigation.

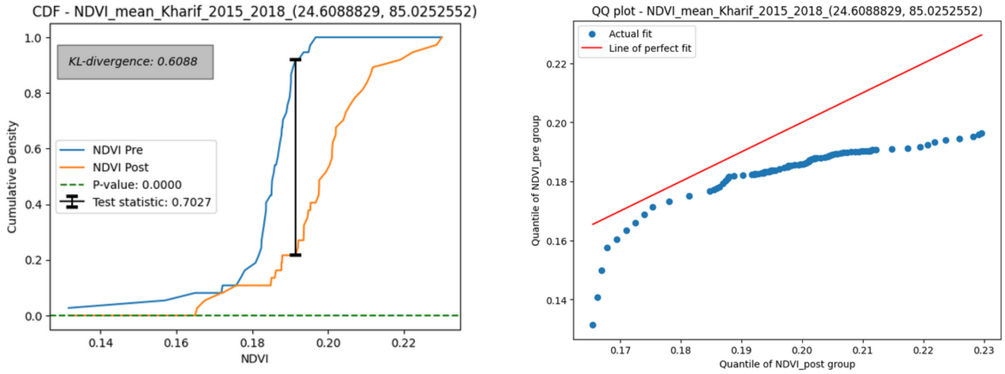
7 CASE STUDIES

While pursuing this research, our team conducted field visits to locations in three districts , viz. Jamui, Gaya and Koderma, aiming to gather feedback from farmers regarding the effectiveness of farm ponds and resulting outcomes. Our observations revealed a mix of positive instances, where the constructed farm ponds had a beneficial impact on the surrounding farms, and negative instances, such as ponds with inadequate water retention capacity and cases of corruption resulting in the absence of constructed ponds. This feedback provided by farmers played a crucial role in validating and strengthening our research findings, emphasizing the significance of engaging with local stakeholders to ensure the accuracy and relevance of our results. Note that in the forthcoming sections, Pre- and Post-NDVI in Figure 5 refer to NDVI values before and after treatment.

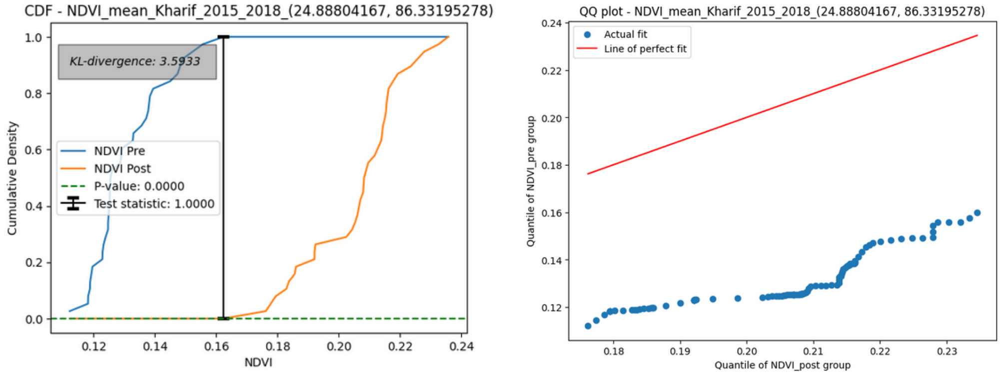
7.1 Gaya

We visited the farm pond built at lat-lon (24°36'32"N, 85°1'31"E) in Gaya (Figure 6a), and show the KS-test and KL-divergence results between the NDVI timeseries before and after the construction of the farm pond. We also plot the quantile-quantile plot to visually check if the NDVI values are from different distributions.

As we can see in Figure 5a, the KL-divergence and KS-test statistics are positive, CDFs of post-NDVI are higher than pre-NDVI, and the QQ plot shows quite different distributions. This indicates that post intervention NDVI values are higher, and affirms our empirical results of a positive impact of farm ponds on agricultural productivity.



(a) Gaya: KS-test for NDVI in Kharif (b) Gaya: Quantile-Quantile plot for NDVI in Kharif



(c) Jamui: KS-test for NDVI in Kharif (d) Jamui: Quantile-Quantile plot for NDVI in Kharif

Fig. 5. KS-test and QQ plots for the two case studies based in Jamui and Gaya districts of Bihar. NDVI Pre and NDVI Post refer to NDVI values before and after treatment.

When visiting this location, we received positive feedback regarding the impact of this farm pond. The farmers reported that the introduction of this pond facilitated an increase in cropping intensity in the surrounding areas. Farmers attested to its direct contribution in enhancing the yield of paddy crops during Kharif, while also benefiting wheat and moong cultivation during Rabi by enabling protective irrigation practices.

7.2 Jamui

We visited the farm pond built at lat-lon (24°53'17"N, 86°19'55"E), picture in Figure 6a. Our analysis indicates a positive KL-divergence and KS-test statistic, higher post-NDVI than pre-NDVI, and starkly different distributions based on the QQ-plot.

During our visit to this location, we observed active utilization of this farm pond for irrigation purposes. Furthermore, we noted the presence of pumps and irrigation channels being utilized for transporting water to fields located at a further distance. The farmers confirmed a positive impact of the pond on their agricultural practices that led to a better crop yield.


(a) (b)

Fig. 6. (a) Visited pond in Gaya - Dried up due to less rain in visit year. (b) Visited pond in Jamui

8 CONCLUSIONS

In this work, we studied the impact of farm ponds constructed under the MGNREGA scheme in the AEZ 13 region of India. We formulated three research questions: (i) to study the impact of farm ponds on agricultural productivity, (ii) to study the impact of farm ponds on productivity specifically during drought years, (iii) to investigate if farm ponds facilitate in reducing sensitivity of agriculture to droughts. Our empirical results across the three questions for the AEZ 13 region show that farm ponds have been instrumental in improving productivity and providing support during drought years, in the Kharif season. The impact in some settings does not persist until the Rabi season though, and the impact during Zaid was found to be the least. Our results are consistent with findings from our field visits wherein we interacted with farmers having farm ponds and validated our methods and results.

A potential direction to improve the estimated treatment effects would be to use better proxies for crop yield. Our study uses NDVI and GCI as an indicator of crop yield. One possible way to to build a more accurate estimator for crop yield is to study crop growth models such as the ones used in existing studies [22], and incorporate data on moisture levels and temperature throughout the crop life-cycle, as high moisture levels at some stage of crop growth alone do not indicate a high crop yield.

Further, we do not consider changes in crop types brought about from the creation of farm ponds. Survey based studies conducted by agricultural economists also consider change in cropping patterns such as shifts from traditional crops to horticulture and cash crops that generate higher incomes (e.g., [34]).

In the future, we plan to extend this study to other AEZs. Farm ponds are used differently in different areas, and therefore they may have varied impacts in the AEZs across India. For instance, in the state of Maharashtra farm ponds are used as water storage structures [32]. Borewells are used to pump water for storage in the ponds that are lined with impervious material to prevent infiltration into the ground. This has led to depleting groundwater levels in the state due to excessive pumping.

Furthermore, the inferences from this study are yet to be augmented with the findings from a landscape-level study wherein we plan to study the combined effect of all NRM structures in a landscape on landscape-level indicators. The current study is based on a site-level analysis wherein we study the effect of farm ponds in their immediate neighbourhood, which is mostly expected to be positive. At a landscape-level, the combined effect of NRM interventions on a landscape could be either positive or negative depending on several factors such as the locations of structures (upland area vs. downstream area), types of structures, behaviour of the individuals of the community (groundwater extraction pattern) over time, and so on.

Acknowledgements: We would like to thank Tower Research, GIZ, and Tarides for supporting the research, and our field partners SUPPORT, 4S, PRADAN, FES, and Gram Vaani, for facilitating site visits and community interactions.

REFERENCES

[1] 2019. Yield prediction in wheat (Triticum aestivum L.) using spectral reflectance indices. *Current Science* 116, 2 (2019), pp. 272–278. https://www.jstor.org/stable/27137836

[2] V. Tare A. K. Prasad, R. P. Singh and M. Kafatos. 2007. Use of vegetation index and meteorological parameters for the prediction of crop yield in India. *International Journal of Remote Sensing* 28, 23 (2007), 5207–5235. https://doi.org/10.1080/01431160601105843 arXiv:https://doi.org/10.1080/01431160601105843

[3] Earth Observation Research Center Japan Aerospace Exploration Agency. 2024. *Global Precipitation Measurement.* https://www.eorc.jaxa.jp/GPM/en/index.html

[4] K. Ajith, V. Geethalakshmi, K.P. Ragunat, S. Pazhanivelan, and Ga. Dheebakaran1. 2017. Rice Yield Prediction Using MODIS - NDVI (MOD13Q1) and Land Based Observations. *International Journal of Current Microbiology and Applied Sciences* 6, 12 (2017). https://www.ijcmas.com/6-12-2017/K.%20Ajith,%20et%20al.pdf

[5] Rodrigo A. Arriagada, Paul J. Ferraro, Erin O. Sills, Subhrendu K. Pattanayak, and Silvia Cordero-Sancho. 2012. Do Payments for Environmental Services Affect Forest Cover? A Farm-Level Evaluation from Costa Rica. *Land Economics* 88, 2 (2012), 382–399. http://www.jstor.org/stable/23272587

[6] Ehsan Asmar, Mohammad H. Vahidnia, Mojtaba Rezaei, and Ebrahim Amiri. 2024. Remote sensing-based paddy yield estimation using physical and FCNN deep learning models in Gilan province, Iran. *Remote Sensing Applications: Society and Environment* 34 (2024), 101199. https://doi.org/10.1016/j.rsase.2024.101199

[7] Chahat Bansal, Hari Om Ahlawat, Mayank Jain, Om Prakash, Shivani A Mehta, Deepanshu Singh, Harshavardhansushil Baheti, Suyash Singh, and Aaditeshwar Seth. 2021. IndiaSat: A Pixel-Level Dataset for Land-Cover Classification on Three Satellite Systems - Landsat-7, Landsat-8, and Sentinel-2. In *Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies* (Virtual Event, Australia) *(COMPASS '21)*. Association for Computing Machinery, New York, NY, USA, 147–155. https://doi.org/10.1145/3460112.3471953

[8] Mohamed Belmahi, Mohamed Hanchane, Nir Y. Krakauer, Ridouane Kessabi, Hind Bouayad, Aziz Mahjoub, and Driss Zouhri. 2023. Analysis of Relationship between Grain Yield and NDVI from MODIS in the Fez-Meknes Region, Morocco. *Remote Sensing* 15, 11 (2023). https://doi.org/10.3390/rs15112707

[9] Christopher F. Brown, Steven P. Brumby, Brookie Guzder-Williams, Tanya Birch, Samantha Brooks Hyde, Joseph Mazzariello, Wanda Czerwinski, Valerie J. Pasquarella, Robert Haertel, Simon Ilyushchenko, Kurt Schwehr, Mikaela Weisse, Fred Stolle, Craig Hanson, Oliver Guinan, Rebecca Moore, and Alexander M. Tait. 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data* 9, 1 (09 Jun 2022), 251. https://doi.org/10.1038/s41597-022-01307-4

[10] Jovanović Dušan, Sabo Filip, Govedarica Miro, and Marinković Branko. [n. d.]. Crop yield estimation in 2014 for Vojvodina using methods of remote sensing. *Ratarstvo i Povrtarstvo* 51, 3 ([n. d.]). https://doi.org/10.5937/ratpov51-6712

[11] Sarvarbek Eltazarov, Ihtiyor Bobojonov, Lena Kuhn, and Thomas Glauben. 2023. The role of crop classification in detecting wheat yield variation for index-based agricultural insurance in arid and semiarid environments. *Environmental and Sustainability Indicators* 18 (2023), 100250. https://doi.org/10.1016/j.indic.2023.100250

[12] FAO. 2024. *Harmonized World Soil Database v 1.2.* https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/

[13] Food and Agriculture Organization of the United Nations. 2024. *India at a glance.* Retrieved February 22, 2024 from https://www.fao.org/india/fao-in-india/india-at-a-glance/en/

[14] OSM Foundation. 2024. *Geofabrik by OpenStreetMap.* http://www.geofabrik.de/

[15] Berlin Freie Universität. [n. d.]. *What is a watershed?* Retrieved February 22, 2024 from https://www.geo.fu-berlin.de/en/v/iwm-network/learning_content/introduction_iwm/definition-watershed//index.html

[16] Google. 2024. *Google Earth Engine.* https://earthengine.google.com/

[17] World Resources Institute Google. 2024. *Dynamic World: A near realtime land cover dataset for our constantly changing planet.* https://dynamicworld.app/

[18] ISRO. 2024. *Indian Geo Platform of ISRO.* https://bhuvan.nrsc.gov.in/home/index.php

[19] Xue Jinru and Baofeng Su. 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors* 2017 (05 2017), 1–17. https://doi.org/10.1155/2017/1353691

[20] David M. Johnson, Arthur Rosales, Richard Mueller, Curt Reynolds, Ronald Frantz, Assaf Anyamba, Ed Pak, and Compton Tucker. 2021. USA Crop Yield Estimation with MODIS NDVI: Are Remotely Sensed Models Better than Simple Trend Analyses? *Remote Sensing* 13, 21 (2021). https://doi.org/10.3390/rs13214227

[21] Murali Krishnan and Rabindra Padaria. 2017. Chemical Science Review and Letters Impact and constraints of farm pond based watershed development in the Hassan district of Karnataka. *Chemical Science Review and Letters* 6, 21 (01 2017).

[22] David B. Lobell, David Thau, Christopher Seifert, Eric Engle, and Bertis Little. 2015. A scalable satellite-based crop yield mapper. *Remote Sensing of Environment* 164 (2015), 324–333. https://doi.org/10.1016/j.rse.2015.04.021

[23] Department of Agriculture Cooperation & Farmers Welfare Ministry of Agriculture & Farmers Welfare, Government of India. 2016. *Manual for Drought Management.* Retrieved February 22, 2024 from https://sdma.cg.gov.in/ManualDrought2016.pdf

[24] S Nalgire and P Chinnasamy. 2022. Index-based impact monitoring of water infrastructures in climate change mitigation projects: A case study of MGNREGA-IWMP projects in Maharashtra. *Frontiers in Water* 2, 12 (2022).

[25] NASA. 2024. *Oak Ridge National Laboratory Distributed Active Archive Center.* https://www.earthdata.nasa.gov/eosdis/daacs/ornl

[26] NASA. 2024. *Shuttle Radar Topography Mission.* https://www.earthdata.nasa.gov/sensors/srtm

[27] USGS (United States Geological Survey NASA. 2024. *Landsat Science.* https://landsat.gsfc.nasa.gov/

[28] Government of India. 2024. *India Water Resources Information System.* https://indiawris.gov.in/wris/#/

[29] ICAR-National Bureau of Soil Survey and Land Use Planning. 2015. *ICAR-NBSS&LUP Serving Science and Society.* ICAR-NBSSLUP, 21–22. https://nbsslup.icar.gov.in/wp-content/uploads/2021/E-book/E_BOOK_NBSS&LUP_Part1.pdf

[30] Manoj Panda, Brajesh Jha, Amit Mandal, Vivek Pal, Aakanksha Sharma, Atrayee Choudhury, and Deepak Kumar. 2018. *Rapid assessment of natural resource management component under MGNREGA and its impact on sustainable livelihoods.* Technical Report. Ministry of Rural Development, Government of India.

[31] Ewa Panek and Dariusz Gozdowski. 2020. Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sensing Applications: Society and Environment* (2020), 10028online6. https://doi.org/10.1016/j.rsase.2019.100286

[32] Pooja Prasad, Om P. Damani, and Milind Sohoni. 2022. How can resource-level thresholds guide sustainable intensification of agriculture at farm level? A system dynamics study of farm-pond based intensification. *Agricultural Water Management* 264 (2022), 107385. https://doi.org/10.1016/j.agwat.2021.107385

[33] Carlos Ramirez-Reyes, Katharine R. E. Sims, Peter Potapov, and Volker C. Radeloff. 2018. Payments for ecosystem services in Mexico reduce forest fragmentation. *Ecological Applications* 28, 8 (2018), 1982–1997. https://doi.org/10.1002/eap.1753 arXiv:https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1002/eap.1753

[34] P.V.R.M. Reddy, M. Girija Shankar, B. Janardhan Reddy, Y. Shankar Naik, Yerra Eswara Prasad, and D.V.S.R.L. Rekha. 2020. Farm Pond Impact Analysis of PMKSY-Watersheds Project in Srikakulam District of Andhra Pradesh. *International Journal of Current Microbiology and Applied Sciences* 9, 12 (2020).

[35] V. K. Sehgal, C. V. S. Sastri, N. Kalra, and V. K. Dadhwal. 2005. Farm-Level Yield mapping for precision crop management by linking remote sensing inputs and a crop simulation model. *Journal of the Indian Society of Remote Sensing* 33 (2005). https://link.springer.com/article/10.1007/BF02990002

[36] Bhaskar Sinha, Deep Narayan Singh, Anoma Basu, and Mili Ghosh. 2017. *Application of Remote Sensing in Assessing the Impacts of Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), in Ratlam District, Madhya Pradesh, India.* Springer International Publishing, Cham, 87–95. https://doi.org/10.1007/978-3-319-46010-9_6

[37] Sreedevi TK, Suhas Wani, Raghavendra Sudi, Patel MS, Jayesh T, Singh SN, and Tushar Shah. 2006. On-site and Off-site Impact of Watershed Development: A Case Study of Rajasamadhiyala, Gujarat, India. *On-site and Off-site Impact of Watershed Development: A Case Study of Rajasamadhiyala, Gujarat, India. Global Theme on Agroecosystems Report no. 20, Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropic* 2 (08 2006).

[38] Reinhard Uehleke, Martin Petrick, and Silke Hüttel. 2022. Evaluations of agri-environmental schemes based on observational farm data: The importance of covariate selection. *Land Use Policy* 114 (2022), 105950. https://doi.org/10.1016/j.landusepol.2021.105950

[39] Juan Von Thaden, Robert H. Manson, Russell G. Congalton, Fabiola López-Barrera, and Kelly W. Jones. 2021. Evaluating the environmental effectiveness of payments for hydrological services in Veracruz, México: A landscape approach. *Land Use Policy* 100 (2021), 105055. https://doi.org/10.1016/j.landusepol.2020.105055

[40] Coady Wing, Kosali Simon, and Ricardo A. Bello-Gomez. 2018. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual Review of Public Health* 39, 1 (2018), 453–469. https://doi.org/10.1146/annurev-publhealth-040617-013507 arXiv:https://doi.org/10.1146/annurev-publhealth-040617-013507 PMID: 29328877.