Using AI/Machine Learning to Evaluate the Distribution of Community Development Aid Across Myanmar

Woojin Jung[a], **Saeed Ghadimi**[b], **Dimitrios Ntarlagiannis**[c], **and Andrew H. Kim**[d]

[a]School of Social Work, Rutgers University; [b]Department of Management Sciences, University of Waterloo; [c]Department of Earth and Environmental Sciences, Rutgers University; [d]School of Social Work, Rutgers University

This manuscript was compiled on September 21, 2023

A critical question in international development is whether donors allocate aid based on stated program objectives such as poverty. However, ground-truth poverty data is usually unavailable at a granular community level where aid interventions take place. This research thus aims to compare granular, remote-sensing-based poverty estimation methods and explores the extent to which poverty and other development characteristics explain aid distribution across communities. This study draws from web-scraped, project-related information on community-driven development projects; 31,300 daytime satellite images; and geospatial attributes of 12,504 villages. First, we compare the performance of four poverty estimation methods: i) spatial interpolation of poverty, using Kriging, ii) nighttime luminosity (NTL), iii) RGB channels, and iv) daytime satellite imagery features. Among these, daytime features extracted by a convolutional neural network outperform other measurements. Next, using the best poverty estimates in the first step, we exploit machine learning (ML) algorithms to predict aid amounts received per village. Our models intend to capture the donor's selection criteria based on need, village capacity, types of projects, ethnicity/state, and accessibility of the village. Overall, our best models explain about 70% of the variance in aid allocation, which leaves major questions regarding what implicit factors dictate the remaining variance in aid targeting. Poorer villages are likely to receive more aid; however, wealth has small relative importance. Instead, village capacity and ethnicity/state better predict aid amount.

community development | poverty | satellite images | aid evaluation | Myanmar

1. INTRODUCTION

With the advent of the Millennium and Sustainable Development Goals (1), poverty elimination has become the central objective of development aid. However, the existing aid-targeting literature reflects that net international aid allocation does not prioritize communities with the highest economic need (2–5). Aid tends to flow to low-income countries but not necessarily to poorer regions within those countries. Yet, the lack of granular poverty and aid measures at a small unit of intervention hampers the accurate assessment of sub-national aid allocation.

A. Research Aim and Outline. This paper explores the question of spatial aid distribution across communities by developing granular poverty measures and machine learning (ML)-based prediction in the context of Myanmar. This research aims to understand which communities in Myanmar are likely to receive more community development aid, given poverty and other characteristics. To do so, it develops poverty measures at the community level and then builds aid prediction models. This paper uses one of Myanmar's largest community development programs, the World Bank's (WB) National Community Driven Development Program (NCDDP), to analyze aid allocation across the country. Myanmar is selected as an extremely data-sparse country receiving a large amount of community development assistance.

B. Motivation. One relevant aid modality in examining the pro-poor nature of subnational aid targeting is community-centered- or driven- development (CCD, CDD). It provides communities with block grants or in-kind support to construct local public goods, infrastructure, and services. CCD involves intra-community level allocation based on village attributes. Thus, it contrasts with other types of aid, such as cash transfers that may only micro-target households regardless of locality. In addition, the evaluation of the need-source match is more relevant to the smaller spatial unit because communities are more likely to have a homogeneous level of wealth than in larger areas (6).

In the literature, the poverty orientation of community development projects is neutral or weakly positive. The central government's allocation across communities is not related to village attributes (7), and CCD does not always target the

Significance Statement

It is challenging to estimate whether area-based interventions are allocated to the communities with the highest poverty in Myanmar, one of the most data-sparse countries receiving a large amount of assistance. We use web-scraped aid data and remote-sensing-based poverty measures to examine which villages are likely to receive more aid. To estimate the poverty level of aid-receiving communities, we first examine the performance of four poverty measures: spatial interpolation of poverty, nighttime luminosity (NTL), Red Green Blue (RGB) channels, and daytime satellite imagery utilizing a Convolutional Neural Network model. Next, we use the most effective measure of estimating poverty to predict aid distribution at the village level. We find that development assistance is more likely to be distributed to villages taking a participatory approach.

Dr. Jung conceived and designed the research; collected the satellite and ground-truth data; performed the deep learning analysis; and wrote the paper. Dr. Ghadimi performed the machine learning modeling and wrote the paper. Dr. Ntarlagiannis collected and analyzed the geospatial data and wrote the paper. Andrew performed the machine learning and feature analysis and wrote the paper.

There is no conflict of interest. There is no financial nor personal interest or beliefs that could affect our objectivity.

[1]corresponding author: w.jung@rutgers.edu

areas that are the poorest, but rather those that are lacking infrastructure (8). Under some conditions, CCD projects are mildly progressive in reaching marginalized localities when poverty maps are drawn from census or survey data (9, 10) or when electoral rewards are in place (11). A recent study (12) highlights that the feasibility of pro-poor aid allocation is hindered by the proximal closeness of communities with differing economic needs as well as a lack of disaggregated data.

We identify several gaps and challenges from the literature in evaluating CCD targeting. While previous studies have considered variance in georeferenced aid location, these studies relied on large administrative units such as states, provinces (3, 13), and districts (5, 14–16). Other sub-national studies tap into already established grid-level data (4, 15, 17), but the large grid size of 50 km by 50 km still misses the nuances in community-level needs. The available data for geolocated aid amounts is also limited. Previous studies (4, 18) have only modeled the binary existence of aid projects (a binary variable) or the number of projects (a count variable) which loses important information in the variation of aid amount.

A significant obstacle in conducting granular analysis is the limited availability of surveyed data on wealth, income, and consumption in subnational regions. Recently, efforts have been made to leverage satellite imagery to develop fine-grained measures of socioeconomic development. Earlier studies estimate regional wealth by the proportion of metal to thatched roofs in satellite images (19). Studies have also found a strong correlation between national-level luminosity and standard measures of economic output such as gross state product (GSP) (20), gross domestic product (GDP), and growth measures (21). More recent work validates new measures of economic outputs by combining multi-spectra imagery and deep learning techniques to predict wealth scores from ground truth surveys (22, 23). To mitigate the limitation in labeled data, previous work transfers model weights trained on the ImageNet dataset (24), a dataset used to classify objects imaged with handheld cameras, to the task of estimating wealth from satellite imagery in Africa (22, 23, 25). We build on this literature with customization to take the previous work to the context of Myanmar.

C. Contribution. In this study, we analyze the distribution of CCD in Myanmar, overcoming the previously discussed challenges. We first gather a novel set of data related to aid, poverty, and village characteristics and then apply ML models to predict WB investment per village. Using villages for our unit of analysis allows aid allocation to be evaluated with more nuance and on its ability to address community-level, sub-national needs. We also geolocate our outcome variable, aid amount, to study aid allocation in great detail. To do so, we retrieve geocoordinates from interactive project maps, which allows the systematic collection of aid data at a high level of spatial resolution. Using a continuous dollar amount improves the variability and distribution of the outcome, which has not been previously attempted.

To the best of our knowledge, this study is the first attempt to leverage ML to predict sub-national transfer in Myanmar. This allows us to study a large set of explanatory variables with multicollinearity and work in high-dimensional spaces. Further, we examine nonlinear modeling combined with explainability techniques that allow policymakers to have a more precise yet intuitive understanding of aid allocation decisions.

The paper proceeds as follows. The next section presents our methodology and data sources. The third section shows our results. The last section concludes and discusses the implications of our work.

2. METHODS AND DATA

This section presents methodologies to measure poverty and predict aid. Subsequently, main datasets for aid and village characteristics are introduced.

A. Methods. Broadly, our approach employs two steps. We first compare the performances of four poverty measures in predicting survey wealth data: i) spatial interpolation, ii) nighttime luminosity (NTL), iii) Red Green Blue (RGB) channels, and iv) Convolutional Neural Network (CNN) feature extractors using daytime imagery.

Our next step is to examine whether poverty and other factors are linked to the amount of aid that each village receives. We include the best poverty estimate from the first step as an input feature in the second step. To predict aid volume, we employ regularized regressions, Support Vector Regression (SVR), and XGBoost. We select various combinations of input features based on aid allocation guidelines for NCDDP.

B. Data. Our sample consists of 12,504 villages participating in the NCDDP between 2017-2018. Out of 12,504 villages, approximately 54.0% of them (6,769 villages) have data on the amount of aid received. The main outcome variable is the log of aid per capita for which 4,630 villages had valid amounts. The explanatory variables are remote-sensing-based or spatial poverty measures. We use the Demographic and Health Survey (DHS) as the ground truth wealth data to train these variables. Covariates are community and project characteristics web scraped from the NDCCP management website and the Myanmar Information Management Unit (MIMU). Data from the United Nations Development Programme (UNDP) Myanmar county office is used to identify the history of funding in a given village.

B.1. National Community Driven Development Program. The aid dataset draws from the National Community Driven Development Program (NCDDP), which has been supported by the World Bank since 2012. Among the 12,504 villages participating in the NCDDP between the years 2017-2018, we used 5,650 village locations with block grant data. Aid is actual aid disbursement in Myanmar kyat (MMK), not mere commitments. We use automated web scraping processes to collect project implementation data via the web portal interface.[*] The extracted aid distribution data is similar to the actual project map, validating our scraping algorithm, shown in Figure 1.

B.2. Demographic and Health Surveys. We use the 2015-2016 Myanmar DHS data to provide the ground-truth poverty data in our wealth prediction model. We aggregate a total of 13,260 household-level data to 441 village cluster-level data to use georeferenced information associated with the centroids of the survey clusters. For our analysis, we use the latitude, longitude, and mean wealth data for each of the clusters.

[*] Interactive project maps provide key project monitoring and implementation data in the form of project dashboards. However, some key information, such as the georeferenced locations of project villages are only available as websites' source codes.

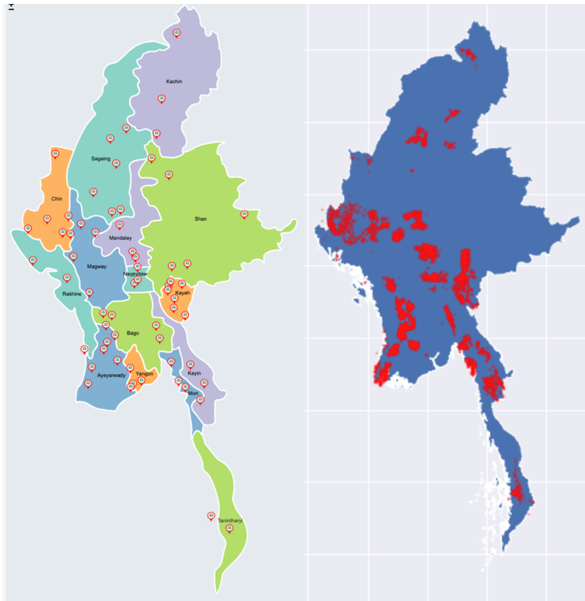

Fig. 1. NCDDP Year 5 Web Portal with village locations (left) and scraped participating village location for NCDDP 2017-2018 or Year 5 (right)

B.3. Daytime Satellite Images. We use daytime satellite images as a feature of wealth prediction, based on previous work in this field (22). A total of about 31,300 images were downloaded from Google Image API in a series of 400 by 400 pixel images (about 1 square kilometer of land area) from the locations of the DHS village clusters as well as the locations recorded as the NCDDP villages. We focus on images in the areas that showed wealth or aid data since these areas are most relevant to our research. The publicly available daytime satellite images used for this study were primarily collected from 2019, while the training data, which is the DHS survey, was collected in 2015-2016.[†]

B.4. Nighttime Luminosity. We also estimate the annual average luminosity values at the center points of project villages. Nightlight Version 1 Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band provides sources for 2015 and 2016 nighttime raster data.[‡] We capture nightlight features at different resolutions while considering any noise effects present in the data (26, 27). We use three measures: high resolution at 5 km by 5 km, low resolution at 10 km by 10 km, and mixed resolution at 2 km for urban and 10 km for rural areas.

B.5. Conflict Data Extrapolation. In order to quantify a conflict measurement for each surveyed aid location, we abstract the measurements of conflict to a single value from the Armed Conflict Location and Event Data (ACLED) from 2015-2019 (Figure 2). The value represents the distance in meters from the center of a village to the nearest area of violent conflict that occurred during the year 2015. Distance has been shown to be an important metric in a previous study where the further away from conflict, the more likely a village is to receive NCDDP aid (18).[§]

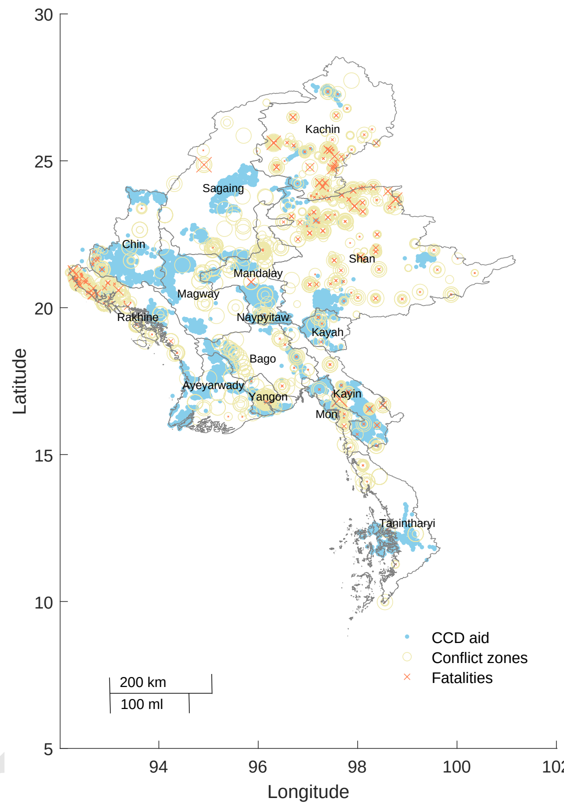

Fig. 2. Map of all recorded conflict areas and aid-receiving villages

B.6. Access to Infrastructure. Geospatial data from the Myanmar Information Management Unit (MIMU) provides spatial context regarding logistical access to and within villages and surrounding infrastructure. We estimate the nearest distance from a village to roads, railroads, dams/lakes, seaports, airports, and low-lying (below 5-meter elevation) areas. We also collect data on hard-to-reach tracts and villages. [¶] NTL data is used to approximate electricity-related infrastructure.

B.7. Other Aid Projects. To model an absence of external funding, we use data shared by the UNDP Myanmar country office. We count the number of aid projects that had been started before 2016 within each township boundary. The data includes aid projects from the United Nations and other international organizations, international and national non-profit organizations (NGOs), community-based organizations, and the Red Cross.

C. Descriptive Statistics. Table 1 displays the main outcome variables and explanatory variables. On average, our sample villages had wealth scores 0.65 below the average wealth of the country's villages (Mean centered wealth at the national

[†] Due to data availability, the timing of data sources between satellite imagery and DHS survey does not overlap in this study.

[‡] The measurement unit for luminosity is a composite cloud-free radiance value estimated in 15 arc-second (approximately 463 m) geographic grids with outliers removed and non-lights set to zero.

[§] We engineered several distance to conflict features using the ACLED-provided categorizations in combination with several geospatial designations. These included looking at violent vs. nonviolent conflicts as well as fatalities from village tract, township tract, and centroid designations. Distance to violent conflicts was selected since it empirically contributed the most to the model and was theoretically aligned with prior research (18).

[¶] To identify such villages, we create a binary variable that distinguishes NCDDP villages located within hard-to-reach tracks, as defined by polygon boundaries. Alternatively, we calculate the distance between NCDDP villages and the nearest hard-to-reach village as point data.

Table 1. Village level participatory design aid distribution selected summary statistics

Measure	Observation	Mean/Prop.	Std	Range
Aid per capita (in MMK)	4,634	25,656	41,971	(0, 2,100640)
Total Aid per Village	6,769	9,705,701	6,498,552	(0,168601000)
Log of aid capita	4,630	9.81	0.79	(4.51,14.56)
Mean-centered Wealth scores	8,096	-0.65	5.61	(-18.69,22.66)
Distance to violent conflict in metres	12,112	36,258	24,580	(2,117062)
Village populations	8,096	602	669	(0,8042)
No. of households	12,086	130	139	(0,2603)
% Social sector project	12,504	18		
% of Burmese village	12,504	53		
% of Chin village	12,504	4		
No. of labour days	8,324	234	384	(0,11603)
Wages paid in MMK	7,014	1,586,691	1,841,155	(0,24460000)
No. of beneficiaries	12,504	592	676	(0,12328)
% of participation	8,843	67	19	(0,100)
No. of committee membership	12,504	14	5	(0,36)
% of female committee membership	8,887	45	13	(0,100)
Satisfaction rate %	274	59	14	(0,100)
No. of Grievances Submitted	12,504	1	2	(0,76)
% Grievances Resolved	12,504	16	36	(0,100)

Note: USD 1= Myanmar Kyat (MMK) 1550 in 2017-2018.
Std: Standard Deviation

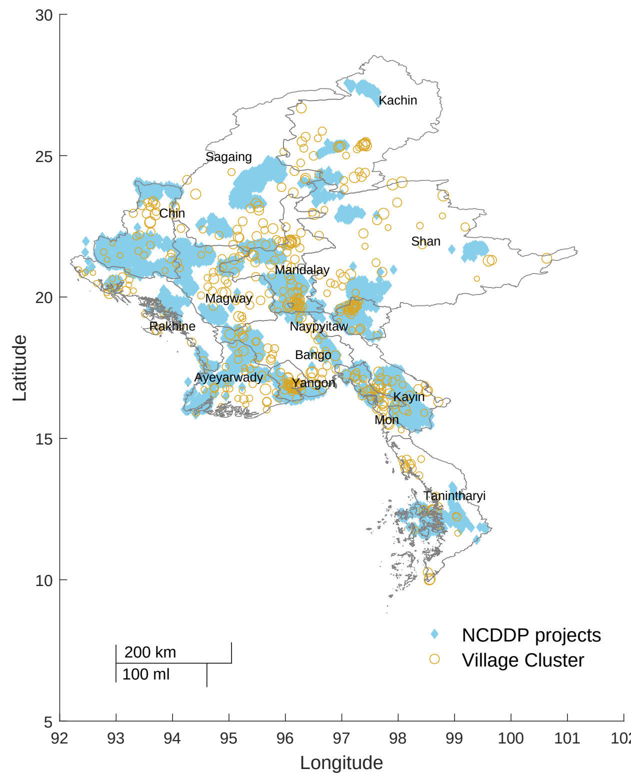

Fig. 3. DHS village clusters and NCDDP projects

level=0, SD=5.61).[‖] Each village consists of 602 individuals living in 130 households and receives a block grant of approximately 17 USD per person. A majority of community members (67%) participated in the project, and a little less than half of the project committee members (45%) were female. The project has received a satisfactory rating from just over half (59%) of the beneficiaries. In terms of complaints, each village submitted one grievance on average, but only 16% of them were successfully resolved. Out of all supported projects, social sector projects make up 18%. Burmese ethnicity makes up slightly more than half (53%) of the project villages.

3. RESULTS OF POVERTY PREDICTION

In the following sections, we report on the results of our study. We begin by comparing the performance of various wealth measures.

A. Wealth Measures. First, we experiment with a few approaches to predict unknown wealth measures for NCDDP villages. Our main objective is to assess how well interpolated values or satellite-extracted features can predict wealth factor scores for DHS village clusters marked with yellow circles in Figure 3. After comparing these models, we find that the CNN-extracted satellite features perform the best. Consequently, we rely on this poverty estimation approach to predict unknown wealth values in blue NCDDP project villages in Figure 3. Detailed results are further elaborated below.

A.1. Four Wealth Prediction Methods. As a baseline, we use Gaussian process regression, also known as Kriging, to spatially estimate DHS wealth factor scores. Second, we extract NTL for DHS village clusters. Our model learned to predict the DHS wealth factor score from NTL intensity. Third, we quantify the intensity of the three color channels (RGB) for each of the daytime images that corresponded to locations surveyed in the DHS data. Utilizing color features is a common method used in image retrieval applications (19, 28). Our model learned to predict DHS wealth scores based on pixel intensity.

Our final method is a deep learning approach using a CNN model. Following the literature, we pre-trained a CNN on ImageNet to identify low-level image features, such as edges and corners, that are common to many vision tasks. This transfer learning approach addresses the limitation in labeled data by transferring the weights of networks trained on ImageNet to downstream tasks. Since our VGG 16 model accepts 224 × 224 pixel images, but our input images are 400 × 400 pixels, we performed data augmentation and normalization on the training set. We then began fine-tuning the full network using the Stochastic Gradient Descent (SGD) optimizer.[**] As the pre-trained CNN has learned a nonlinear mapping from each input image to a feature vector representation, it renders 4,097 features from the satellite images. We then link cluster-averaged wealth from the DHS survey data with the corresponding image features extracted to several regression models. We discovered early in our research that NTL data does not provide a sufficiently accurate measurement of wealth in Myanmar, unlike small African countries in the literature.[††] Due to the lower correlation, we chose not to train the daytime images to classify NTL intensity but to directly predict DHS wealth scores.

A.2. Evaluation of Four Methods. The performance of poverty prediction methods can be evaluated against their ability to predict DHS wealth scores. For each of the four methods, we apply the elastic net (ElNet) regression model with 10-fold cross-validation on 80% of the observations (training set),

‖ The wealth factor score was 0 on average, with a maximum of 22, a minimum of -18, and a standard deviation of 5.

** Most layers were frozen during tuning, and only the last block of fully convolutional layers was tuned. A batch size of 8 was used and the maximum number of epochs was set to 6

†† For instance, Myanmar's R^2 (0.18) is considerably lower than R^2 of other developing countries such as Rwanda (0.74).

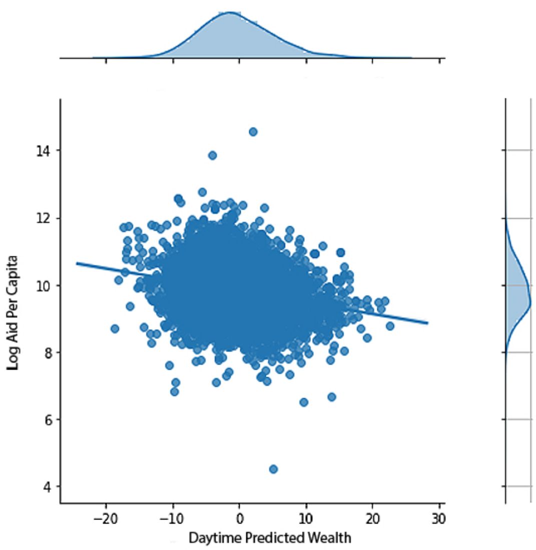

Fig. 4. Regression of Log Aid Distribution against Daytime Satellite Predicted

Note: Regression of Log Aid Distribution against Daytime Satellite Predicted Wealth shows a negative correlation, which is expected if aid goes to less wealthy regions.

Table 2. Wealth Prediction Model Comparison

	CNN with Daytime Image	RGB Avg Pixel Features	Geospatial Interpolation (Kriging)	Nighttime Data
R^2	**0.477**	0.388	0.210	0.180

and then we test its performance on the remaining 20% of the sample (testing set). The CNN-trained daytime features perform the best with R^2 of 0.48, while simple RGB analysis on daytime images score R^2 of 0.39, and spatial interpolation performed next best at R^2 of 0.21. Nightlights explain 18% of the variation in the wealth index. Figure 5 illustrates the distribution of our best model's wealth predictions in Myanmar (CNN-trained daytime features).

A.3. Poverty Prediction. By fitting DHS wealth to CNN-extracted features, we can use this model to input CNN-extracted features of NCDDP villages and predict their wealth. This allows estimation of the wealth of NCDDP villages that have not been surveyed and lack ground truth wealth.

4. RESULTS OF AID PREDICTION

In this section, we examine the variance in aid volumes across different village characteristics, utilizing wealth prediction and other village traits as input features.

A. Aid Allocation Models. We expect that CCD distribution would follow program eligibility criteria to achieve program objectives. The 2015 operational manuals (29, 30) denote that poverty rates are the primary criterion for selecting townships. Additional criteria include the absence of external funding, the capacity of the township, peace and stability, and logistical access. Our prediction models are built based on the NCDDP selection criteria.

- Model 0 (M0) captures the bivariate relationship between poverty and aid allocation.

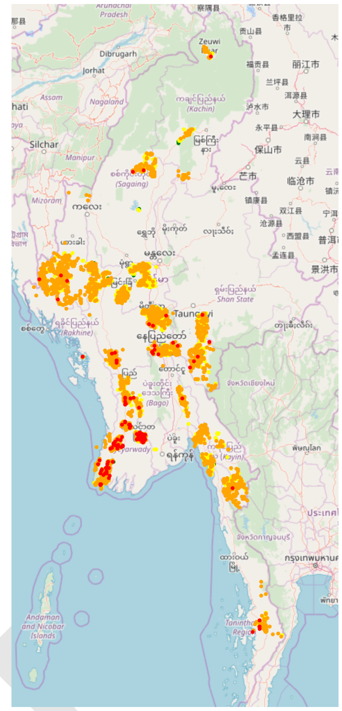

Fig. 5. CNN-based predictions of wealth in areas surveyed in Myanmar

- Model 1 (M1) captures NCDPP program criteria of need such as poverty, the absence of external funding, conflict, and logistical access. Poor villages with some level of peace and stability and lack of aid from other sources are likely to receive more per capita aid (29, 30).
- Model 2 (M2) captures the capacity of communities and local authorities. This is a more qualitative measure. CCD requires the ability to apply for, plan, and manage block grants from both communities and local governments. Communities with a high level of participation (e.g., % of female committee membership, % of grievances reported and addressed) and a solid structure (e.g., village/village tract project support committees) are likely to receive more aid.
- Model 3 (M3) captures the types of aid projects funded. Depending on the project sector, aid volume would change. It accounts for the fact that investment amounts may vary with types of infrastructure projects (e.g., building community centers vs. roads).
- Model 4 (M4) captures states and ethnicity. Given that NCDDP emphasizes the inclusion of minorities, ethnic minorities and struggling states might be encouraged to participate in CCD. Ethnic minority features are also correlated with living on the border states or in conflict areas.
- Model 5 (M5) captures needs relating to logistical accessibility and infrastructure. Additionally, we introduce spatial distance and access variables. NCDDP can add more value if it supports infrastructure in remote villages.

Models 1-3 reflect explicit program criteria. Model 4 is based on the emphasis on inclusion and diversity in NCDDP's language. Model 5 incorporates operational considerations.

B. Choice of Machine Learning (ML) Algorithms. We employ prediction models including ElNet regression, support vector machine regression (SVR), and extreme gradient boosting (XGBoost). ElNet regression offers a balanced approach, striking a harmony between solution sparsity and the prevention of overfitting.

On the other hand, SVR is a variant of support vector machines (SVM) that is used for regression tasks. While SVM is primarily designed for classification, SVR extends the concept to handle continuous target variables. Unlike traditional regression models that aim to minimize the error between predictions and actual values, SVR focuses on limiting the deviations of predictions from a specific margin.

XGBoost is a also powerful distributed gradient boosting approach that builds upon the foundation of the gradient boosting framework and incorporates state-of-the-art ML algorithms. It leverages the Newton-Raphson method by utilizing a second-order Taylor approximation of the loss function which enhances the algorithm's functionality and efficiency.

C. Building and Comparing ML Models. We began by fine-tuning the hyperparameters of various ML algorithms. For ElNet regression, there are two hyperparameters, namely α and λ, which control the coefficients of l_1 and l_2 norm squared regularization terms. In particular, the former is set to $\alpha\lambda$ while the latter is set to $0.5\alpha(1-\lambda)$. We initially tested extreme values (e.g., very high or low α) using a course grid. We then narrowed down the range of values and used a fine grid.

In the case of XGBoost, we defined a hyperparameter space and used Bayesian optimization with the Expected Improvement acquisition function. The hyperparameter γ in this method determines the minimum loss reduction required to make a further partition on a leaf node of the tree.

Similarly, a grid search was used for SVR to iterate through the best combinations of parameters and kernels. The hyperparameter C in this method is used to avoid overfitting. It is a positive number that controls the penalty imposed on observations that lie outside the margin determined by the other hyperparameter ϵ (see Appendix A for the hyperparameters of the best model across algorithms).

We also employed a cross-validation method to train, validate, and test the models. The data was split into two sets, 80% for training and validation and the remaining 20% for testing. The test set remained the same for all algorithms to ensure a fair comparison. We then compared their performances using error metrics, such as mean squared error (MSE) for test sets as well as goodness-of-fit measures, such as R^2.

D. Performance of Regularized Regression ML Models. We start with regression results and then cover other algorithms because regression is the most intuitive method that shows directions and magnitudes of coefficients.

D.1. Performance of Non-Nested ML Models. In this section, we report the results of the non-nested regression models.

M0: Ordinary least squares (OLS) bivariate regression using Wealth (β=-0.184) only explained 6.6% (R^2 =0.066) of the total variance in logged aid allocation. Log aid per capita is larger as villages become less wealthy.

Table 3. Performance of Log Aid per Capita Non-Nested Models

Model	R^2	Adj. R^2	Val. MSE	Test MSE
M0 OLS Bivariate	0.066	0.065	**0.572**	0.617
M1 ElNet α=16, λ=0.0	**0.417**	**0.416**	0.622	**0.578**
M2 ElNet α=0.058, λ=0.1	0.187	0.185	0.756	0.805
M3 ElNet α=0.040, λ=0.1	0.045	0.040	0.963	0.946
M4 ElNet α=0.007, λ=1.0	0.157	0.137	0.870	0.836
M5 ElNet α=0.031, λ=0.0	0.124	0.122	0.903	0.868

M1: NCDDP program criteria of need only explains 41.7% (ElNet; $R^2 = 0.417$) of the total variance. Log aid per capita is larger for villages that are closer to hard-to-reach towns (β=-0.138), less wealthy (based on CNN-extracted satellite features; β=-0.076), further from violent conflict (β=0.044), and have less number of aid projects per township (β=-0.009). Villages with smaller populations (β=-0.569) are also associated with larger aid per capita. However, the strong negative relationship with the population is not surprising given the NCDDP's block grant allocation method, which establishes the floor and ceiling on aid amounts based on population (29, 30). The increases in aid amount are disproportional to the increases in population size and are ordinal by nature (e.g., a population size of 3,001=55-75 million Kyat vs. 6,002=80-90 million Kyat).

M2: Village capacity-related features explain 18.7% (ElNet; $R^2 = 0.187$) of the variation in the outcome variable. Log aid per capita is larger for villages with more participation (β=0.397), a fewer number of committee memberships (β=-0.176), and higher wages (β=0.093). All other features minimally impact prediction.

M3: Economic and social sector projects explain 4.5% (ElNet; $R^2 = 0.045$) of the variance. Only community center projects had a coefficient with a magnitude greater than or equal to 0.100 (β=0.137). Having fewer economic sector projects (β=-0.021) is related to a larger logged aid allocation while having more social sector projects (β=0.014) is related to a larger logged aid allocation.

M4: Ethnic groups and states/regions/territories explain 15.7% (ElNet; $R^2 = 0.157$) of the variation in the outcome variable. Log aid per capita is larger for villages in the north-eastern states of Chin (β=0.140) and Magway (β=0.103), and lower for those in the well-to-do states of Mandalay (β=-0.112) and Nay Pyi Taw (β=-0.105). Additionally, aid size was larger for those who belonged to the Chin ethnic group (β=0.064) and smaller for those in the Burmese ethnic group (β=-0.087). Overall, states/regions/territories appear to be more important in regularized regression prediction when compared to ethnicity (the largest weighted ethnicity, Burmese, is the 5th largest by magnitude).

M5: Finally, geographic and NTL features explain 12.4% (ElNet; $R^2 = 0.124$) of the variation in log aid volume per capita. Log aid per capita is larger for villages further from a seaport (β=0.357), closer to a low-lying area (β=-0.321), are considered hard-to-reach (β=0.259), and are further from mining areas (β=0.095). It is important to note that long distances to the seaport and short distances to low-altitude areas are both associated with larger log aid per capita.

D.2. Performance of Nested ML Models. In terms of variable selection, we compared results from several nested models using

Table 4. Coefficients for Selected Features in Best Performing Log Aid per Capita Regression Models

Coefficient	ElNet M9	ElNet M8.1
Wealth	**-0.046**	-0.043
% of Participation	0.223	**0.225**
Population	-0.492	**-0.511**
Wages	0.190	**0.201**
Dist. to Violent Conflicts (m)	**0.061**	**0.061**
R^2	**0.509**	0.508
adj. R^2	**0.496**	**0.496**

Table 5. Nested Model Description

Model	Description	Name
M6	M1+M2	Need+Capacity
M7	M1+M2+M3	M6+Sector Projects
M8.1	M1+M2+M3+M4-10	M7+Ethnicities (10+ obs)
M8.2	M1+M2+M3+M4-50	M7+Ethnicities (50+ obs)
M9	M1+M2+M3+M4-10+M5	M8.1+Infrastructure
M10	M1+M2+M3+M4-50+M5	M8.2+Infrastructure

various combinations of the 5 non-nested models (see Table 5 for different model combinations). M6 combines the features of M1 and M2. M7 combines the features of M1, M2, and M3. We then create several variations of M8 which combines the first four non-nested models (M1+M2+M3+M4). Once we transformed Ethnic groups into binary variables, we aggregated groups by observations into an "other" category using several thresholds. M8.1 uses a threshold of 10 and M8.2 uses a threshold of 50[‡‡]. M9 adds M5 to M8.1 and M10 adds M5 to M8.2.

The best model (M9) explains just over half of the variance 50.9% ($R^2 = 0.509$). Table 4 shows the coefficients of the key features. The percentage of participation (β=0.223) and wages (β=0.190) are related to larger aid sizes. Population (β=-0.492) is associated with smaller aid sizes.

E. Comparison of Prediction Performance Across ML Algorithms. Our results indicate that the XGBoost algorithm consistently outperforms all other algorithms. The best XGBoost model with the most comprehensive features (M10) shows the highest performance ($R^2 = 0.713$) compared to the best SVR (M7, R^2=0.613), and best ElNet model (M9, R^2=0.509) (Table 6). Even when compared to the best SVR and ElNet models, the XGBoost algorithm outperforms them (see Appendix B for detailed comparisons). Overall, full nested models that incorporate program criteria of need (M1), village capacity (M2), sector projects (M3), ethnic groups and states (M4), and geographic features (M5) perform better than partial models.

E.1. Important Features Across Algorithms. We compare and analyze how the different algorithms utilize the various features. To do so, we consider the best-performing model of each algorithm. ElNet and XGboost algorithms produce coefficient weights and feature importance scores, respectively. ElNet coefficient weights are represented by β while XGBoost uses gain, frequency, or weight to produce its feature importance score. We use gain in this study since it is considered to have more importance for generating predictions (33). Gain scores show the improvement in the model due to the feature (34). For the SVR algorithm, we utilize the Shapley Additive exPlanations (SHAP) method (35) which uses Shapley value concepts (36) from game theory (37).[§§]

Of particular interest is how wealth impacts prediction. In all models, wealth had minimal impact compared to other features. In terms of direction, less wealth is associated with more log aid per capita. The utility of wealth in SVR aid prediction is more driven by avoiding wealthy areas rather than targeting poorer areas.

Next, we analyze which of the non-nested models have the highest predictive power for aid allocation. Overall, when compared to the program criteria of need (M1) features, village capacity (M2) features are consistently more important across algorithms. In particular, participation and wages are consistently more important than M1 features. In addition, XGBoost M10 uses the number of grievances submitted as an important feature, while SVR identifies the number of committee members as an important feature. Of the five program criteria of need features (M1), aid projects per township (XGBoost, SVR) and population (ElNet, SVR) stood out as important features. As previously discussed, the importance of population is not surprising; however, it is notable that population is the 33rd most important in XGBoost (gain=1.004) and has a gain score below the average of all features (mean gain=1.391) in the best model.

Finally, we explore the important features of our best overall model, which is XGBoost with comprehensive features (M10). In this model, ethnic group and state features are used the most in the overall prediction (5 out of the top 6). After aid per township (gain=10.753), the states of Magway (gain=6.202), Thaninthayi (gain=3.717), Yangon (gain=3.301), and Kayin (gain=3.221), and the Dai (Yindu) ethnic group (gain=3.127) make up the top features (Figure 6). For a more detailed analysis of important features across algorithms, see Appendix C.

F. Selection Bias. Our analysis is limited to villages with grant data. Villages with grant data are different from villages without grant data. We ran a t-test for a range of covariates to find nonrandom differences between villages with and without grant data. The analyzed villages were wealthier, more densely populated, and had more participatory practices as shown by the greater percentage/number of committee membership, female committee membership, grievances submitted and resolved, and use of community labor. However, villages with data and those without data are not statistically different with regard to the intensity of conflict, composition of Burmese ethnicity, project satisfaction rate, and percentage of community participation.

G. Robustness Check. In our robustness check, we consider different variations of our target variable as well as the impact

Table 6. Performance Metrics of Best ML Models

Model	R^2	adj_R^2	mean val MSE	test MSE
XGBoost M10	**0.713**	**0.706**	**0.017**	**0.285**
SVR M7	0.613	0.610	0.403	0.383
ElNet M9	0.509	0.496	0.500	0.486

[‡‡] Although we recognize the issues of aggregating less represented groups(31, 32), small sample sizes reduce the chances of balanced representation during cross-validation procedures. Therefore, we experiment with several thresholds.

[§§] It is an extension of Local Interpretable Model-agnostic Explanations approach to black-box ML models (LIME) (38). An added benefit of SHAP values is that we are able to observe how feature values impact the prediction by computing their contribution to the prediction.

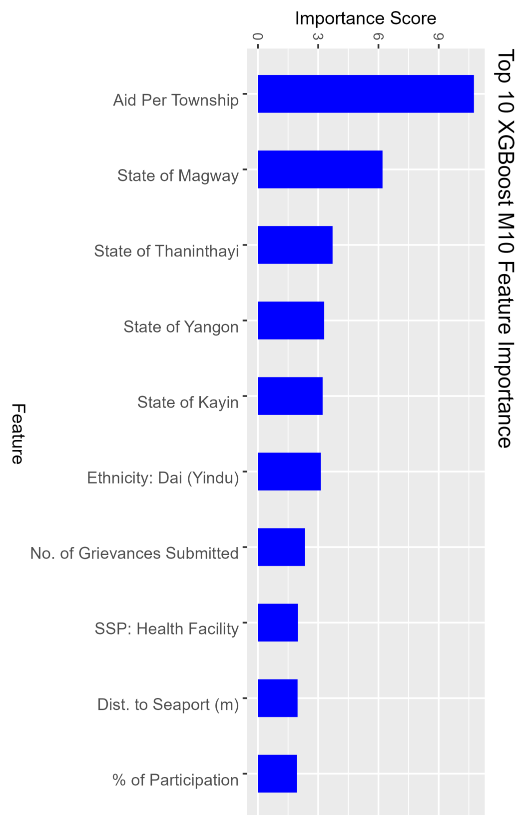

Fig. 6. Top 10 XGBoost M10 Feature Importance (SSP=Social Sector Projects)

of dimensionality reduction.

G.1. Different Target Variations. In addition to logged aid per capita, we predict aid per capita and total aid per village. Both aid per capita and total aid per village perform worse than logged aid per capita. Between the two, total aid per village performs better and explains 30.3% (ElNet; $R^2 = 0.303$) of variance. Similar to the main results, the total aid per village is larger for villages with higher wages (β=0.235). Population (β=0.334) has a positive relationship when using the total aid per village since the floors and ceilings of aid amount increase as the population size increases (29, 30). Detailed results can be seen in Appendix D.

G.2. Dimensionality Reduction. In our PCA-based dimensionality reduction, we explore several feature sets using loading thresholds of 0.300 and 0.250 as well as the top 5 of each principal component (PC; see Appendix E for detailed methods and results). Overall, the PCA-based models support that XGBoost M10 is the best-performing model as well as the importance of incorporating features beyond the program criteria of need (M1) (i.e., village capacity, sector projects, ethnic groups and states, and geographic features).

As an additional check, we utilize the L_1 penalty term within the ElNet regression, to guide feature selection. We remove any features that were shrunk to zero. However, there are no notable improvements after doing so (see Appendix F for full results). Overall, L_1 penality dimensionality reduction also supports our main findings: that XGBoost M10 is the best-performing model, as well as the importance of incorporating features beyond the program criteria of need (M1). The details can be found in the Appendix.

5. CONCLUSION AND DISCUSSION

In this section, we present conclusions from our study and discuss important implications for both policy and future research.

A. Conclusion. We find that using features extracted from daytime satellite imagery by the CNN was an improvement over spatial interpolation and RGB or nightlights-based measures. Our findings suggest that daytime satellites, as a far-sensing measurement of wealth, are an effective means of predicting aid distribution with high performance. This remote-sensing-based wealth tends to produce stronger signals than survey-based wealth.

More aid flows to poorer areas as predicted wealth is negatively associated with aid volumes across all models and target outcomes.[¶¶] With better wealth prediction, the association between wealth and aid amount becomes more consistent compared to the previous study (18). However, the mean coefficient of wealth is marginal (β=-0.054). It should be noted that the explanatory power of wealth is weak even if poverty is the most important selection criterion of the NCDDP.

More significant features in regression models are village capacity-related features, such as participation and wage. More participation is associated with larger aid allocation since it may be related to a village's capacity to apply for and manage such aid. Presumably, larger wages are associated with larger aid allocations since more funding would be required to accommodate for the higher cost of labor in that area. When implementing best-performing XGBoost algorithms, the number of aid projects per township is identified as the most important feature, along with states, ethnicity of Dai (Yindu), and village capacity.

In addition, the effect of population size is significant even after normalizing aid by capita; the disbursed amount per person is larger for less populated communities. This is consistent with the cross-country studies, in which population size has a negative relationship and is one of the top predictors of average aid per capita (39) and reflects the block grant allocation method by population size used by the NCDDP (29, 30).

Overall, our results indicate that the prediction of aid size using various implicit and explicit features and algorithms can only be achieved 71.3% of the time (XGBoost M10). Given that we used program selection criteria and reasonable sets of variables that might be related to aid investment at the village level, the finding implies the need to better model qualitative criteria (e.g., willingness or capacity of local authorities) and other unknown but influential factors. For instance, areas known to be more politically active may be more likely to receive aid, as political parties may be inclined to encourage those who are more likely to vote for them. In terms of the aid prediction model, more data for governance and political dimensions, such as election data, can be considered.

B. Discussion. On the policy front, our findings suggest gaps between the explicitly stated aid selection policy and actual implementation. Although the top stated selection criterion is

¶¶ One exception is for M1 using total aid per village but it is very small (0.005).

poverty, participatory components of the project and village capacity seem to matter more than poverty. Interestingly, the SHAP values for our best SVR model suggest that aid selection is driven more by an effort to avoid wealthy areas but is not utilized to target areas of poverty.

Our results also highlight the need for more transparent aid allocation rules. Specifically, program criteria of need alone do not provide a comprehensive picture of aid allocation. Moreover, while other implicit variables can better explain aid distribution, about a third of the variance still remains unknown. Given the sizeable, unexplained portion of aid distribution by proclaimed, operational, and normative angles, decisions on investment size per village should be more clearly stated in the operation document. They should also be included in the monitoring and evaluation document.

This may also reflect a lack of reliable data at the time of the 2017-2018 project and, at the same time, the necessity of developing complementary poverty measures for small areas. The further development of small-area poverty measures may make it more feasible to actually target poverty as current poverty measures may inherently lack variance in poorer areas. For example, asset-based metrics will naturally have more varied profiles for wealthier units of analysis, while poorer ones will become more homogeneous despite having varied experience.

Poverty prediction using remote sensing in Myanmar yields modest prediction compared to African countries with multiple sources of georeferenced poverty data over time. It is a large and topologically and geographically diverse country with relatively small and static wealth data.[***] Historic commercial imagery with a zoom level over 18 would provide a vivid, intuitive, and time-relevant picture of the region's wealth and poverty, bearing in mind the trade-off between accuracy and replicability.

This study promotes evidence-based aid allocation by discussing potential mechanisms behind aid variation for area-based interventions. We show that far-sensing tools, such as daytime satellite imagery combined with ML algorithms, can predict the geospatial distribution of community development projects in a data-sparse country. Similar analyses can be conducted to link nationwide CCD distribution with poverty in other conflict-prone states.

[***](441 georeferenced DHS entries

6. Appendices

In this section, we present more numerical experiment results.

A. Hyperparameter Turning. The hyperparameters obtained through tuning for the best model across algorithms are presented.

Table 7. Hyperparameters of M10 Across Algorithms

Algorithm	Hyperparameters	Value
XGBoost	Boosts	556
	Subsample	0.5
	Min child weight	1
	Max depth	5
	Learning rate	0.046
	γ	0.071
	Colsample by tree	0.7
SVR	C	10
	ϵ	0.1
	Kernel	rbf
ElNet	α	0.008
	λ	0.4

B. Comparison of Best Models. A comparison of algorithms across each best model is presented below.

Table 8. Best XGBoost: M10 (M1+M2+M3+M4_50+M5) Comparison

Model	R^2	adj_R^2	mean val MSE	test MSE
XGBoost	**0.713**	**0.706**	**0.017**	**0.285**
SVR RBF	0.604	0.595	0.435	0.392
ElNet	0.507	0.496	0.504	0.489

Table 9. Best SVR: M7 (M1+M2+M3) Comparison

Model	R^2	adj_R^2	mean val MSE	test MSE
XGBoost	**0.692**	**0.689**	**0.340**	**0.305**
SVR RBF	0.613	0.610	0.403	0.383
ElNet	0.488	0.484	0.520	0.507

Table 10. Best ElNet: M9 (M1+M2+M3+M4-10+M5) Comparison

Model	R^2	adj_R^2	mean val MSE	test MSE
XGBoost	**0.703**	**0.695**	**0.340**	**0.294**
SVR RBF	0.600	0.589	0.432	0.396
ElNet	0.509	0.496	0.500	0.486

C. Supplementary: Important Features Across Models. This section presented further details regarding the features of the best models.

C.1. Wealth. Between XGBoost, SVR, and ElNet, wealth had minimal impact. In XGBoost M10, wealth is the 52nd most important feature out of 67 (importance score=0.708; Figure 7). By comparison, the average importance score is 1.391 (range: 0.189-10.753). In SVR M7, wealth is the 14th most important feature out of 27, and less wealth is associated with more log aid per capita (Figure 8). High wealth values consistently have larger negative SHAP values, while low wealth values have consistently smaller positive SHAP values. Therefore, the utility of wealth in SVR aid prediction is more driven by avoiding wealthy areas rather than targeting poorer areas. Finally, ElNet M9 produced a small coefficient weight (-0.046) for wealth and indicated that less wealth is associated with more log aid per capita.

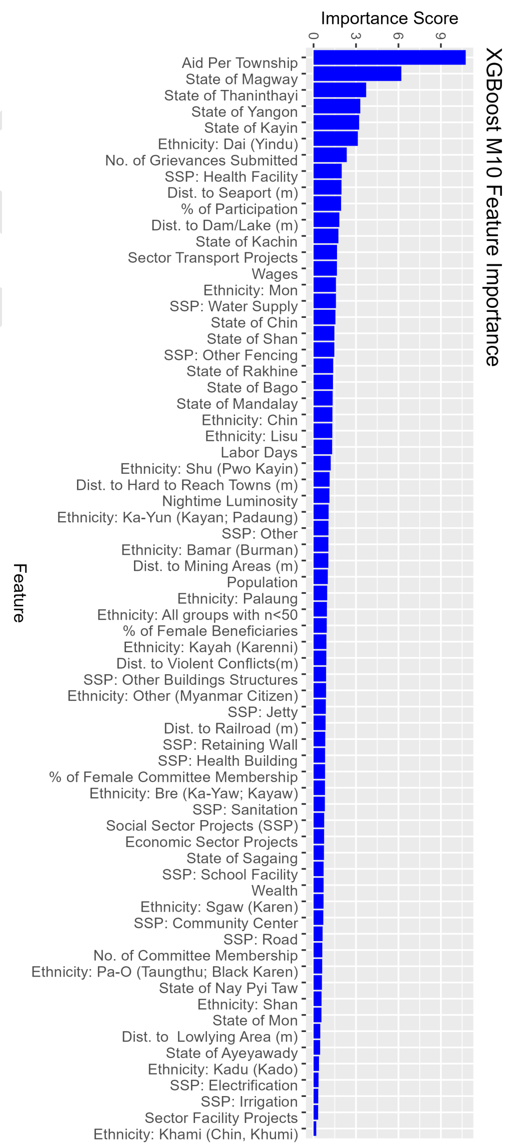

Fig. 7. All XGBoost M10 Feature Importance (SSP=Social Sector Projects)

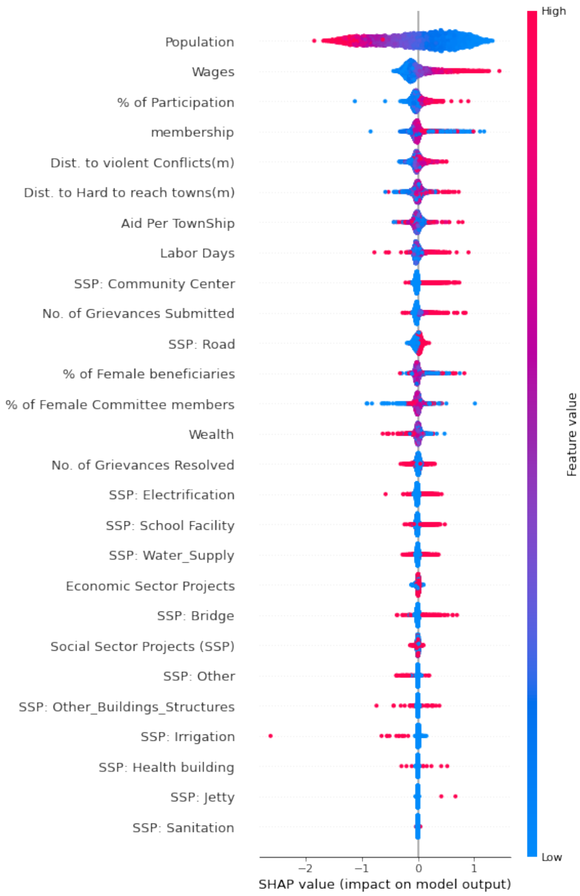

Fig. 8. Shapley Summary plot for the best Model 7

C.2. Predictive Power of Non-Nested Models. The percentage of participation has the 2nd largest magnitude (β=0.223) in ElNet M9, the 3rd most important in SVR M7, and the 10th most important in XGBoost M10. ElNet and SVR indicate that relatively high participation is associated with larger aid predictions in both models. SVR SHAP values indicate that both high and low percentages of participation have similar impacts by magnitude on the prediction. Wage is the 2nd most important in SVR M7 and the 3rd largest weight by magnitude (β=0.190) in ElNet M9. In XGBoost M10, wages are the 14th most important (gain=1.650), and of the program criteria of need (M1) features, only aid projects per township are more important. Both SVR and ElNet indicate that larger wages are associated with more aid.

In addition, the algorithms identify several other important village-capacity features. XGBoost M10 uses the number of grievances submitted as the 7th most important (gain=2.345). SVR identifies the number of committee members as the 4th most important feature. The number of committee members had an interesting behavior; relatively lower membership numbers mostly lead to larger aid predictions, but can sometimes lead to smaller aid predictions. Relatively higher membership numbers typically have no impact or have a marginal impact on the model prediction.

Of the five M1 features, population and aid projects per township stood out as important features. The population is the most important feature in SVR M7 and has the largest feature magnitude in the ElNEt M9 (β=-0.492). Larger village populations are associated with smaller aid predictions. However, this is expected due to the NCDDP's block grant allocation formula that sets aid amount ranges based on population size (29, 30). The use of capita to normalize aid leads to a negative relationship. Interestingly, XGBoost used population as the 33rd most important feature (gain=1.004). Aid projects per township are shared by XGBoost M10 and SVR M7. It is the most important in XGBoost M10 (gain=10.753) and the 7th most important in SVR M7. ElNet does not produce a substantial weight (β=0.038) for it.

Table 11. M1 and Best Regression Performance for all Targets

Target	R^2	Adj. R^2	mean val MSE	test MSE
AidCapita, M1+ElNet α=0.058, λ=0.2	0.197	0.196	1.084	0.285
AidCapita, M7+ElNet α=0.030, λ=0.1	0.288	0.282	1.073	**0.252**
logAidCapita, M1+ElNet α=16, λ=0.0	0.417	0.416	0.622	0.578
logAidCapita, M9+ElNet α=0.030, λ=0.1	**0.509**	**0.496**	**0.500**	0.486
SingleCost, M1+ElNet α=0.023, λ=0.1	0.186	0.185	0.740	0.914
SingleCost, M9+ElNet α=0.046, λ=0.1	0.303	0.283	0.632	0.783

D. Different Target Variations. For the total aid per village target variable, the largest nested model (M9) gives the best performance explaining 30.3% (ElNet; $R^2 = 0.303$). For the features that had a coefficient magnitude greater than or equal to 0.100, the total aid per village is larger for villages with a larger population and higher wages. Additionally, we explore which model explains the most variation. For aid per capita, Model M7 explains 28.8% (ElNet; $R^2 = 0.288$). Aid per capita is larger for villages with smaller populations (β=-0.212).

E. Principal Component Analysis-Based Dimensionality Reduction. We use principal component analysis (PCA), which transforms a dataset of possibly correlated features into a new set of uncorrelated features called principal components (PC). These PCs are linear combinations of the original features and are ordered in such a way that the first PC captures the maximum amount of variance in the data, the second PC captures the second highest amount of variance, and so on. 57 principal components were identified in the most comprehensive PCA. The number of PCs was selected using the explained variance ratio function to retain at least 80% of the variance. We follow a scree plot with the proportion of variance explained to determine a final set of PCs. We then compare several feature sets using a factor loading threshold of 0.300 and 0.250 as well as the top 5 features of each component retaining each unique feature once. Below are the results of the best PCA-based regression and XGBoost results of each of the three feature sets.

Table 12. PCA-Based ElNet Comparison

Selection Method	# of Features	R^2	Adj. R^2	mean val MSE	test MSE
Top 5	95	**0.486**	**0.471**	**0.539**	**0.509**
0.250 Cutoff	89	0.484	0.470	0.548	0.511
0.300 Cutoff	**66**	0.166	0.149	0.863	0.827

F. L_1 Penalty Dimensionality Reduction. This section presents the results of L_1-based dimension reduction models using XGBoost, SVR, and the best regularized regressions (ElNet).

Table 13. PCA-Based XGBoost Comparison

Selection Method	# of Features	R^2	Adj R^2	mean val MSE	test MSE
Top 5	70	**0.710**	**0.701**	**0.331**	**0.287**
0.250 Cutoff	61	0.700	0.691	0.341	0.298
0.300 Cutoff	**48**	0.176	0.160	0.833	0.816

F.1. XGBoost Model Results. XGBoost model results after removing features that were shrunk to zero via L_1 penalty.

Table 14. Myanmar XGboost Statistics for target variable ln_aid_capita with Dimensionality Reduction

name	R^2	Adj. R^2	mean val MSE	test MSE	test MAE
M1	0.594	0.593	0.412	0.403	0.495
M2	0.135	0.133	0.775	0.858	0.739
M3	0.016	0.020	0.989	1.007	0.806
M4	0.067	0.051	0.905	0.925	0.774
M5	0.079	0.077	0.885	0.913	0.770
M6	0.472	0.470	0.384	0.523	0.574
M7	0.457	0.453	0.378	0.539	0.583
M8.1	0.472	0.462	0.372	0.523	0.574
M8.2	0.461	0.453	0.374	0.534	0.579
M9	0.664	0.656	**0.326**	0.333	0.448
M10	**0.688**	**0.680**	0.367	**0.309**	**0.430**

F.2. SVR Model Results. SVR model results after removing features that were shrunk to zero via L_1 penalty.

Table 15. Myanmar SVR Statistics for target variable ln_aid_capita with Dimensionality Reduction

name	R^2	Adj. R^2	mean val MSE	test MSE	test MAE
M1	0.568	0.567	0.432	0.428	0.507
M2	0.211	0.209	0.725	0.782	0.694
M3	0.036	0.032	0.960	0.955	0.778
M4	0.151	0.137	0.871	0.841	0.722
M5	0.150	0.148	0.856	0.842	0.732
M6	0.597	0.595	**0.407**	0.400	0.494
M7	0.608	0.605	0.414	0.389	0.482
M8.1	**0.620**	**0.613**	0.418	**0.377**	**0.466**
M8.2	0.614	0.608	0.413	0.382	0.468
M9	0.613	0.604	0.432	0.384	0.468
M10	0.597	0.588	0.428	0.400	0.479

F.3. Regression Model Results. Regression model results after removing features that were shrunk to zero via L_1 penalty.

Table 16. Myanmar Regression Model Statistics for target variable ln aid per capita with Dimensionality Reduction

name	R^2	Adj. R^2	mean val MSE	test MSE	test MAE
M1	0.415	0.414	0.621	0.580	0.605
M2	0.461	0.460	0.533	0.534	0.578
M3	0.045	0.040	0.958	0.947	0.772
M4	0.156	0.136	0.866	0.836	0.725
M5	0.123	0.121	0.902	0.869	0.735
M6	0.476	0.474	0.529	0.519	0.573
M7	0.488	0.483	0.519	0.508	0.565
M8.1	0.508	**0.496**	**0.500**	0.488	**0.551**
M8.2	0.505	0.495	**0.500**	0.491	0.552
M9	**0.510**	**0.496**	**0.500**	**0.486**	**0.551**
M10	0.505	0.494	0.506	0.490	0.554

ACKNOWLEDGMENTS. This material is based upon work supported by the National Science Foundation under Grant No. DGE-1747486. We are grateful to Dr. Joshua Blumenstock for his support and advisory role on this project.

1. UG Assembly, Transforming our world: The 2030 agenda for sustainable development (a/res/70/1). united nations (2015).
2. H Öhler, P Nunnenkamp, Needs-based targeting or favoritism? the regional allocation of multilateral aid within recipient countries. *Kyklos* **67**, 420–446 (2014).
3. RC Briggs, Does foreign aid target the poorest? *Int. Organ.* **71**, 187–206 (2017).
4. RC Briggs, Poor targeting: A gridded spatial analysis of the degree to which aid reaches the poor in africa. *World Dev.* **103**, 133–148 (2018).
5. H Öhler, M Negre, L Smets, R Massari, Ž Bogetić, Putting your money where your mouth is: geographic targeting of world bank projects to the bottom 40 percent. *PloS one* **14**, e0218671 (2019).
6. M Gerell, Smallest is better? the spatial distribution of arson and the modifiable areal unit problem. *J. Quant. Criminol.* **33**, 293–318 (2017).
7. E Galasso, M Ravallion, Decentralized targeting of an antipoverty program. *J. Public economics* **89**, 705–727 (2005).
8. RS Chase, Supporccting communities in transition: The impact of the armenian social investment fund. *The World Bank Econ. Rev.* **16**, 219–240 (2002).
9. C Paxson, NR Schady, The allocation and impact of social funds: spending on school infrastructure in peru. *The World Bank Econ. Rev.* **16**, 297–319 (2002).
10. S Wong, What have been the impacts of world bank community-driven development programs? cdd impact evaluation review and operational and research implications. *Wash. DC: World Bank* (2012).
11. A de Janvry, H Nakagawa, E Sadoulet, Pro-poor targeting and electoral rewards in decentralizing to communities the provision of local public goods in rural zambia. *Univ. California, Berkeley* (2009).
12. JA Harris, DN Posner, Does decentralization promote poverty alleviation? evidence from kenya's constituencies development fund. (2021).
13. JF Marineau, MG Findley, 'the swarm principle': A sub-national spatial analysis of aid targeting and donor coordination in sub-saharan africa. *Stability: Int. J. Secur. Dev.* **9** (2020).
14. R Burgess, R Jedwab, E Miguel, A Morjaria, G Padró i Miquel, The value of democracy: evidence from road building in kenya. *Am. Econ. Rev.* **105**, 1817–1851 (2015).
15. R Hodler, PA Raschky, Regional favoritism. *The Q. J. Econ.* **129**, 995–1033 (2014).
16. T Masaki, The political economy of aid allocation in africa: Evidence from zambia. *Afr. Stud. Rev.* **61**, 55–82 (2018).
17. RC Briggs, Why does aid not target the poorest? *Int. Stud. Q.* **65**, 739–752 (2021).
18. W Jung, Mapping community development aid: Spatial analysis in myanmar. *World Dev.* **164**, 106124 (2023).
19. B Abelson, KR Varshney, J Sun, Targeting direct cash transfers to the extremely poor in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14. (Association for Computing Machinery, New York, NY, USA), p. 1563–1572 (2014).
20. T Ghosh, et al., Shedding light on the global distribution of economic activity, the open geography journal, 3, 148–161 (2010).
21. JV Henderson, A Storeygard, DN Weil, Measuring economic growth from outer space. *Am. economic review* **102**, 994–1028 (2012).
22. N Jean, et al., Combining satellite imagery and machine learning to predict poverty. *Science* **353**, 790–794 (2016).
23. C Yeh, et al., Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nat. communications* **11**, 2583 (2020).
24. J Deng, et al., Imagenet: A large-scale hierarchical image database in *2009 IEEE conference on computer vision and pattern recognition.* (Ieee), pp. 248–255 (2009).
25. S Ganguli, J Dunnmon, D Hau, Predicting food security outcomes using convolutional neural networks (cnns) for satellite tasking. *arXiv preprint arXiv:1902.05433* (2019).
26. CN Doll, JP Muller, JG Morley, Mapping regional economic activity from night-time light satellite imagery. *Ecol. Econ.* **57**, 75–92 (2006).
27. A Bruederle, R Hodler, Nighttime lights as a proxy for human development at the local level. *PLOS ONE* **13**, 1–22 (2018).
28. T Gevers, Color-based retrieval in *Principles of visual information retrieval.* (Springer), pp. 11–49 (2001).
29. Government of the Republic of the Union of Myanmar, Community-driven development project operations manual (2015) (accessed: 05.25.2019).
30. Government of the Republic of the Union of Myanmar, Community-driven development project operations manual (2018) (accessed: 05.25.2019).
31. P Yoon, et al., Alternative methods for grouping race and ethnicity to monitor covid-19 outcomes and vaccination coverage. *Morb. Mortal. Wkly. Rep.* **70**, 1075 (2021).
32. AJ Umaña-Taylor, MA Fine, Methodological implications of grouping latino adolescents into one collective ethnic group. *Hispanic J. Behav. Sci.* **23**, 347–362 (2001).
33. J Dong, et al., Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with xgboost in different regions of china. *Eng. Appl. Artif. Intell.* **117**, 105579 (2023).
34. AYM Abdullah, et al., Spatio-temporal patterns of land use/land cover change in the heterogeneous coastal region of bangladesh between 1990 and 2017. *Remote. Sens.* **11**, 790 (2019).
35. SM Lundberg, SI Lee, A unified approach to interpreting model predictions. *Adv. neural information processing systems* **30** (2017).
36. LS Shapley, A value for n-person games. (1953).
37. MJ Osborne, A Rubinstein, *A course in game theory.* (MIT press), (1994).
38. MT Ribeiro, S Singh, C Guestrin, " why should i trust you?" explaining the predictions of any classifier in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* pp. 1135–1144 (2016).
39. N Qian, Making progress on foreign aid, (National Bureau of Economic Research, Inc), NBER Working Papers 20412 (2014).