Site selection strategy for the PACSMAC project

By Adugna Bekele[1]; Gezaghen Berecha[2]; Caleb Gallemore[3]; Janina Grabs[4]; Kristjan Jespersen[5]; Ng'winamila Kasongi[6]; Melkamu Mamuye[7]; Annkathrin Mathe[8]; Guta Regasa Megerssa[9]; Daniel Mwalutolo[10]; Christine Noe[11]; Stefano Ponte[12]; Pily Silvano[13]; Weyessa Terefe[14]; Suyana Terry[15]; Nestory Yamungu[16]

08.06.2023

[1,2,7,9,14]Jimma University (JU), [3]Lafayette College, [4]Escuela Superior de Administración y Dirección de Empresas (ESADE), [5,8,12,15]Copenhagen Business School (CBS), [6,10,11,13,16]University of Dar es Salaam (UDSM)

Cite as: Bekele, et al. *Site selection strategy for the PACSMAC project*. Working paper 2.1, PACSMAC Project, Copenhagen Business School.

Project context

The PACSMAC project is a 5-year collaboration between Copenhagen Business School, the University of Dar es Salaam, Jimma University, Lafayette College, and ESADE Business School. The project aims to investigate how climate change – and the ways actors across the value chain are trying to adapt to or mitigate it – affect coffee farmers' livelihoods and land-use decisions. Work package 1 is dedicated to understanding: 1) How might climate change itself, alongside the mitigation and adaptation efforts intended to address it, affect the governance of coffee value chains originating in Ethiopia and Tanzania? And 2) How do these changes affect the distribution of value along the chain, upgrading opportunities and farmer livelihoods?

Introduction

The Paradoxes of Climate-Smart Coffee (PACSMAC) project, introduced more fully in PACSMAC Project Working Paper 1.3, investigates how Ethiopia and Tanzania's smallholder coffee producers and their respective value chains have been and are likely to be affected by climate change; document patterns of existing adaptive responses on the part of smallholders and their value chains; and explore potential future strategies for promoting equitable resilience in the smallholder coffee sectors in these two countries. Two central methods for collecting data to support this work are a household livelihood survey and focus group discussions in sampled coffee-producing communities in the two countries. This working paper summarizes the site selection strategy developed to identify an appropriate sample of communities in which to conduct the project's fieldwork. After explaining the site selection rationale, the paper presents some basic information on the sampled communities constructed from publicly available geospatial data.

Site Selection Strategy

To select appropriate sample sites in the two countries, the project members adopted a two-stage strategy. First, we selected study regions in the two countries based on levels of coffee production. Here, the objective was to sample the regions in the two countries where the highest levels of smallholder coffee production took place. In Tanzania, an additional factor in region-level site selection was the dominant cultivar. Unlike in the case of Ethiopia, Tanzania has a substantial area of robusta production, creating an additional important variable to consider in site selection. After selecting these regions, we used governmental data on coffee production from Ethiopia and estimates of suitable elevations for coffee cultivation in Tanzania (Tanzania Coffee Board, 2019) to identify only communities in the selected regions engaged in coffee production.

After selecting the study regions and restricting the population to smallholder-coffee-cultivating communities, we were ready for the second stage of sampling. There are a wide range of strategies that can be employed in selecting appropriate sites for a study like ours. For example, were the project's objective to test the effect of community-level treatment variables on a set of outcomes, then an appropriate strategy might be to create a matched sample, selecting samples of treatment and control sites with similar values on a set of potential confounding variables. In PACSMAC's case, however, the bulk of our key expectations have to do with the relationship between household characteristics and adaptation strategies, meaning that our primary "treatment" variables vary within communities, in addition to across communities. Because a key project objective is to identify the range of adaptation strategies smallholder farmers in Ethiopia and Tanzania have adopted or might adopt in the future, as well as the ways that value chains are adjusting to climate change and changes in smallholder strategies in these places, our primary concern at the level of site selection is to ensure that the selected sites are representative of the range of the most important geophysical factors likely to push smallholders to adopt new strategies.

Because the core research questions for the PACSMAC project relate to smallholder farmers' and value chains' responses to climate change, our central concern was to construct a sample of communities representative of the range of experiences of climate change in the study regions. In addition, because there is strong evidence that the impact of climate change on coffee cultivation is likely to be moderated by elevation (Ovalle-Rivera et al. 2015; Bunn et al. 2015), we also designed our sample selection to be representative of the range of elevations in the study regions. We explain the steps of this sampling procedure in more detail in the remainder of this section.

The selection of study sites in Tanzania was based on coffee species and production levels. For the Robusta coffee, we selected the Kagera region (Northwest) because it is the leading Robusta coffee producer in the country. For the Arabica coffee, we selected Ruvuma and Songwe (both in the south), and Kilimanjaro (North). The southern regions were selected because they are currently the leading Arabica producers and recipients of coffee interventions (including climate adaptation interventions). Indeed, it was crucial to include the Kilimanjaro region because of its historical legacy in the coffee sector despite its significant drop in coffee production (Tanzania Coffee Board, 2021). Further, we also selected one representative district from each region following similar criteria used in the region selection. Thus, we selected four districts of Kyerwa (Kagera region), Mbozi (Songwe region), Mbinga (Ruvuma region), and Rombo (Kilimanjaro region).

Following district selection, it was necessary to identify appropriate sample communities within each district. This meant that we needed to operationalize indicators for climate change impacts and elevation, our primary exogenous and moderating variables. First, it was necessary to identify the spatial units that we would be using to represent communities. This is not as straightforward a task as it might at first appear, as Ethiopia and Tanzania have rather different jurisdictional structures, with the result that the relevant geographic units in Ethiopia are typically larger than their Tanzanian counterparts. In Ethiopia, we selected the kebele as the appropriate community unit. Kebeles are the smallest formal administrative units having a neighborhood or a localized and delimited group of people. In Tanzania, we selected the village as the corresponding spatial unit responsible for organizing small-holder farmers and where local-level coffee marketing institutions are located. Study villages were then randomly selected at high, middle, and low elevation levels based on coffee farming, high precipitation change, and the existence of coffee marketing societies. Consequently, eight villages were selected in each study district, with three villages at higher and medium elevations, and two at lower elevations. Selection of villages at different altitudinal gradients intended to capture the intensity of coffee production, climate change experiences, and corresponding adaptation strategies by small-holder farmers at such levels.

While there are numerous ways climate change is presently affecting coffee cultivation, and more ways it might do so in the future (see PACSMAC Working paper 1.2), a key factor that emerged in preliminary field interviews is altered rainfall timing and variability. Changes in the level and timing of rainfall can cause several problems for smallholder coffee farmers. This is the case not only for growing coffee cherries on the bush but also for the preparation process. Ill-timed rainfall, as several preliminary interviews pointed out, can wreak havoc on coffee drying. In the longer term, changes in temperature regimes also threaten the expansion of crop diseases and pests.

Based on this initial information, we developed a strategy to identify the range of experiences of changes in precipitation variability across communities in the selected regions. To construct this measure, we collected monthly data from the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) dataset (Funk, et al., 2015). The dataset combines data from worldwide precipitation monitoring stations and remote sensing. The approach uses satellite imagery to create precipitation estimates that are then compared to interpolated estimates based on precipitation-monitoring station data for bias correction. The result is a coarse resolution (0.05°) estimate of monthly precipitation starting in 1980. While these data are invaluable, it is important to note that because they are at a relatively coarse resolution, they may fail to capture some microclimatological variation in the study areas.

To use these data to create an index of precipitation change that would combine information on both differences in the total amount and temporal variability of precipitation, we decided to compare the mean absolute difference between typical monthly precipitation for our study sites at present with the monthly precipitation in those areas for the earliest years available from the CHIRPS dataset. To accomplish this objective, we downloaded monthly CHIRPS estimates for 1981 through 1985 and 2017 through 2021. We then computed the monthly mean precipitation per pixel. We then computed the absolute difference in mean monthly rainfall for each month between the 1980s estimates and those for the most recent five years of data before computing the mean of these absolute values. The result was a high-resolution raster dataset whose pixels include estimates of the mean difference between contemporary monthly precipitation levels and those that were observed 36 years prior. We then aggregated the mean of these values to the community level using zonal statistics (that is, we computed the mean value of all pixels that fell within the borders of each community).

Elevation was our second key variable for site selection. Here, the procedure was somewhat simpler. We downloaded elevation data from the Shuttle Radar Topography Mission at an approximately 90-meter resolution (Farr, et al., 2007). We then computed the mean elevation for each community using zonal statistics.

An important point informing our research design choices is that the literature leads us to expect that climate change and elevation's impacts on coffee production will be interactive. In other words, we expect coffee farmers to face different pressures from the same climatic changes depending on the elevation at which they are cultivating. After calculating these values for each community in the selected regions in each country, therefore, the next step was to generate a sample of sites that would be sufficiently representative of these regions' combinations of elevation and climate change experiences.

To accomplish this objective, we first divided the communities in each country into quintiles (that is, five groups of equal size arranged along the range of a variable) according to their mean absolute monthly precipitation deviation and their elevation. First, we identified all observed combinations of the lower (up to 20th percentile), middle (above 40th and below 60th percentile), and upper (above 80th percentile) values on these two variables across the communities in each country. This generated a total of nine different classes (three quantiles of precipitation deviation times three quantiles of elevation). To ensure that the sampled sites in each country covered all combinations of the selected quintiles of precipitation deviation and elevation, it was necessary to ensure that the final sample of communities in each country included at least one community from each of these nine classes.

After the list of all communities in the selected regions in each country that fell into one of the nine classes described in the previous paragraph was compiled, we summarized some additional data on these communities from openly available geospatial data to provide additional context to aid in site selection:

- Mean forest cover as of 2000, computed from the GLAD dataset (Hansen, et al., 2013), which we used as an indicator of the potential for shade-grown coffee expansion
- Total square kilometers of forest cover loss from 2000 to 2018 as a proportion of total area, computed from the GLAD dataset (Hansen, et al., 2013), which we use as an indicator of pressure on local forests

- Change in land area under crops from 2003 to 2019, as a proportion of cropland in 2003, computed using data from Potapov, et al. (2022), which we use as an indicator of competition for land
- Median travel time to a city of at least 50,000 people, in minutes, computed using data from the Malaria Atlas Project (2015), which we use as an indicator of market exposure
- Proportion of total area falling within protected areas listed in the World Database on Protected Areas (UNEP-WCMC & IUCN, 2021), which we use as an indicator of ecosystem value and competition for land

While it is certainly possible that some or all of these variables might also interact with climate change and elevation, budgetary limitations, as well as the actual distribution of these variables across communities, made it impossible to generate a sample of sites that would include combinations of the lower, middle, and upper quintiles of all these variables (this would be 3^7, or 2,187 different combinations). Nevertheless, we felt it to be important to have these measures available during the site selection process to avoid constructing samples that were inadvertently highly unbalanced on one or more of these variables.

For the final phase of site selection, we took the list of all communities in the selected regions in the two countries, along with data denoting their quintile for each of the variables outlined above. Using this information, we created maps of the possible sites from which a sample could be generated for each region. Figure 1 presents an example of these maps. Maps for each selected region were distributed to the country teams along with the summary dataset. Drawing on the preliminary interviews, their own local knowledge, and taking into consideration budgetary constraints, the country teams selected eight communities in each selected region, for a total of 30 communities in each country, such that the resulting sample included at least one village for each combination of the lower, middle, and upper quintiles for precipitation deviation and elevation in the country and, to the extent possible given budgetary and logistical constraints, maximized variation on the other contextual variables. During fieldwork, it became necessary to adjust the selection again in some instances due to inaccessibility to some sampled areas due to landslides or similar problems or because the team discovered that coffee was not grown in the sampled area. We provide some detail on our final samples in the following section.

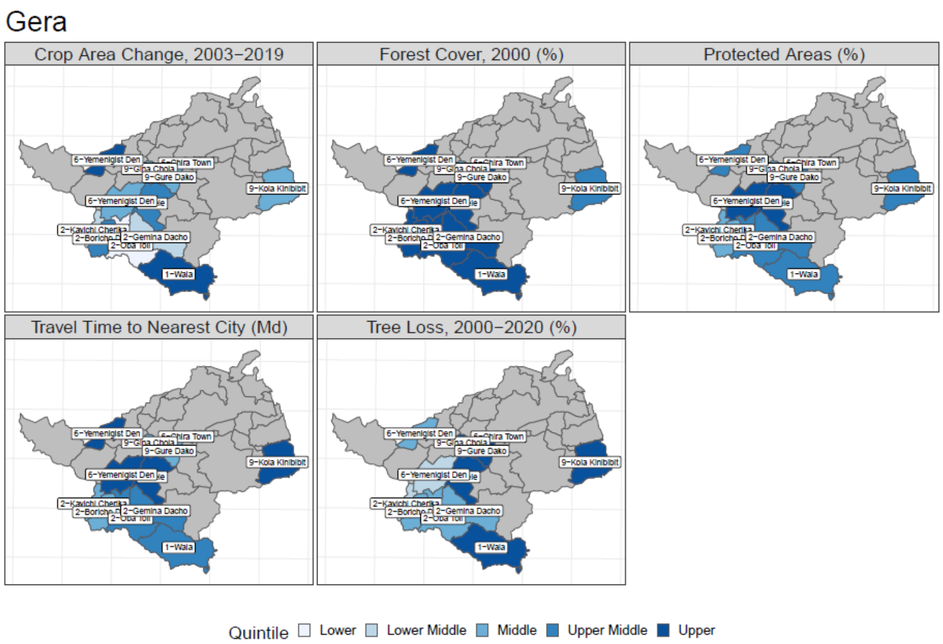

Figure 1. Example of a map of possible sites from which to select a sample in Gera, Ethiopia.

Sample Overview

This section summarizes the key contextual variables we used for site selection across the sampled communities in each country, comparing these values to the values for the population of communities in the selected regions as a whole. While the site selection strategy outlined above did successfully produce representative samples on the most important variables of mean precipitation deviation and elevation, there are some key contextual variables on which the samples differ from the population mean or distribution. While these differences do not necessarily undermine the potential conclusions that we can draw regarding the interaction between elevation and precipitation changes, they could constrain the generalizability of certain findings if the processes in question are closely linked to the unrepresentative variables.

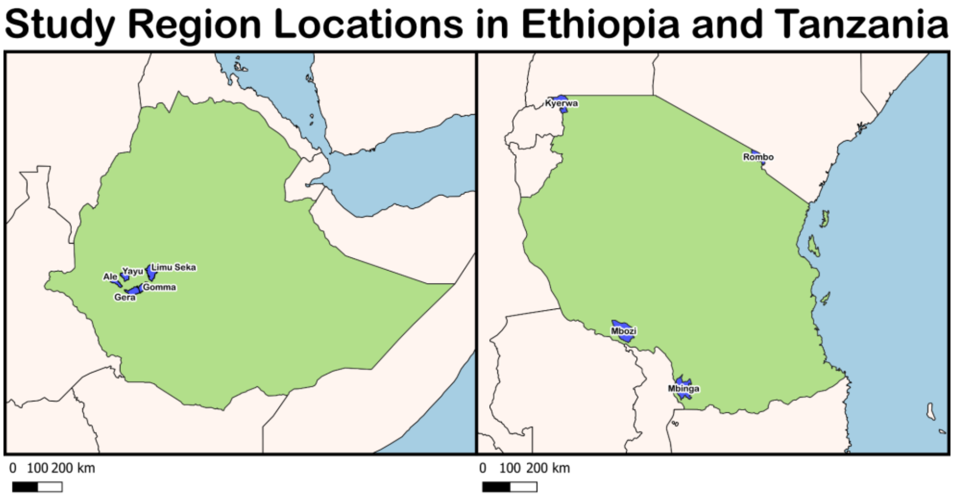

Figure 2. Location of selected study regions in Ethiopia and Tanzania.

Figure 2 presents the selected study regions in Ethiopia and Tanzania (see the Appendix for a complete list of selected kebeles/villages). One important note is that due to budgetary constraints, it became necessary to reduce study region selection in Tanzania from an initially planned six sites to four. This decision came after the initial calculation of quintile groupings for sampling, with the result that the sampled villages in Tanzania are representative of the range of elevation and precipitation deviation observed in all major smallholder coffee regions in the country. In the following, unless specifically noted, the comparisons of the sample to the population for Tanzania refer to the population of villages in the four selected regions in Figure 2, rather than all six villages.

Ethiopia

Figure 3 presents the distribution of the key contextual variables used to inform site selection for the sampled kebeles in Ethiopia and permits comparison with the overall population of villages in the selected regions. The figure highlights the substantial variation in several of the key variables observed across the sampled communities. In the case of the precipitation deviation measure, for example, the most affected communities are experiencing more than twice the deviation of the least affected, a substantial amount of up to around 100 mm of precipitation, on average, per month, as compared to the early 1980s. We also find dramatic variation in forest cover, with some kebeles boasting as much as an 80% forest canopy, while others are nearer 25%.

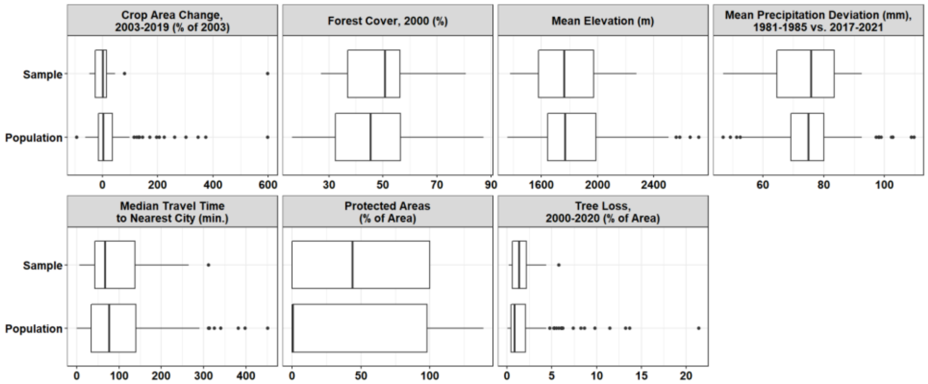

Figure 3. Boxplots of key contextual variables for sampled kebeles in Ethiopia as compared to the overall population of kebeles in the selected study regions. The left-hand side of the box designates the lower quartile of the variable, the line in the middle of the box the median, and the right-hand-side of the box the upper quartile. The whiskers on the box show all observations within 1.5 times the interquartile range of the sides of the box, while the dots show individual outliers.

Some results might at first appear puzzling. For example, how can it be that there have been such dramatic fluctuations in cropland cover in some kebeles?[1] Here, it is important to remember that we measure fluctuations in cropland cover relative to the amount of cropland in the kebele as of 2003. In some cases, this initial number was very low, so over a 100% change in either direction may be dramatic in relative terms, but it is more modest in absolute terms. Similarly, we find cases where 100% or more of the kebele's area is part of a WDPA-listed protected area. The values over 100% reflect cases where ostensible boundaries of multiple protected areas overlap, resulting in double-counting.

Commented [1]: Could this also be because of some of the semi-autonomous ethnic protected areas in Oromia?

Figure 3 also highlights the variables for which the sample is representative of the overall population of kebeles and those areas where it is less so. By design, the sample is most representative of the overall variation in elevation and precipitation deviation, with the median value for these variables in the sample very close to that for the population as a whole. The sample is similarly close to the population distribution of travel time to the nearest city of at least 50,000 people. The sample does have slightly higher forest cover and, as a result, lower crop area change and higher forest cover area change than the population as a whole.

[1] We should note here that for presentational purposes, Figure 3 excludes some even more extreme outliers that, if plotted, would make it impossible to visually compare the bulk of the distributions for percentage changes in crop lands.

Table 1 confirms the visual evidence that the sampled kabeles are broadly representative of the population of kebeles in the study regions. It presents t-tests comparing the population to the sample mean for each of our key contextual variables. Here, we find a few key differences that will be important to consider in interpreting the results of the fieldwork. First, the sampled kebeles have statistically significantly higher protected area coverage and tree loss than the population of kebeles in the selected regions. We find no statistically significant differences between the mean value of any of these variables for the sample versus the population as a whole.

Variable	t-statistic	t-statistic p-value	Kolmogorov-Smirnov Statistic	KS-statistic p-value
Crop Area Change, 2003-2019 (% of 2003)	0.25	0.81	0.14	0.63
Forest Cover, 2000 (%)	-0.85	0.4	0.21	0.19
Mean Elevation (m)	0.85	0.4	0.1	0.94
Mean Precipitation Deviation (mm), 1981-1985 vs. 2017-2021	0.36	0.72	0.16	0.47
Median Travel Time to Nearest City (min.)	0.2	0.84	0.11	0.89
Protected Areas (% of Area)	-1	0.31	0.14	0.49
Tree Loss, 2000-2020 (% of Area)	0.57	0.57	0.14	0.66

Table 1. Comparison of key contextual variables for sampled kebeles to the population of kebeles in the selected regions in Ethiopia as a whole. T-statistics compare population to sample mean (i.e., if the t-statistic is negative, this implies that the mean is higher in the sample than the population) and Kolmogorov-Smirnov statistics test the hypothesis that the sampled distribution of each variable differs from the population distribution.

Table 1 also reports Kolmogorov-Smirnov tests comparing the distribution of each variable for the sample and the population, finding, again, no statistically significant differences. Taken together, the results from the two tests indicate that the sampling strategy was effective in generating a set of kebeles that are broadly representative of the study regions on a range of relevant contextual variables.

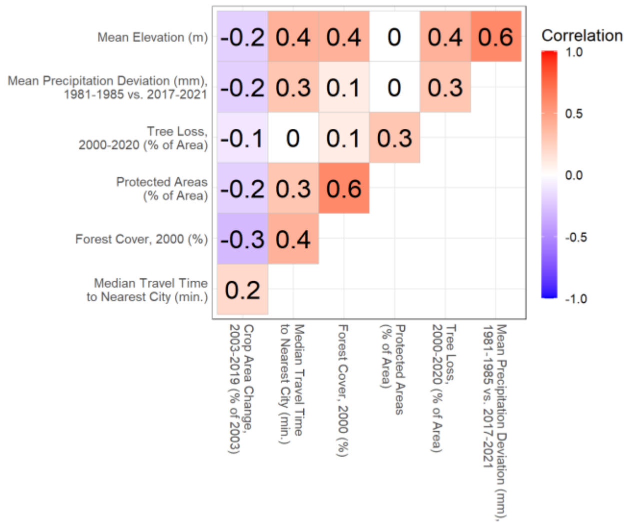

Figure 4. Correlations between key contextual variables for sampled kebeles in Ethiopia.

Figure 4 provides additional context to the characteristics of the sampled kebeles by comparing the correlations between the key contextual variables used to make the sample. The figure demonstrates some clear patterns across elevation levels, with higher elevations, not surprisingly, tending to have less change in crop areas, higher travel times to the nearest city of 50,000 persons or more, higher forest cover, and, in part simply because forests make up a larger percentage of land area, higher forest loss as a percentage of total area. Also not surprisingly, protected area coverage is highly correlated with forest cover.

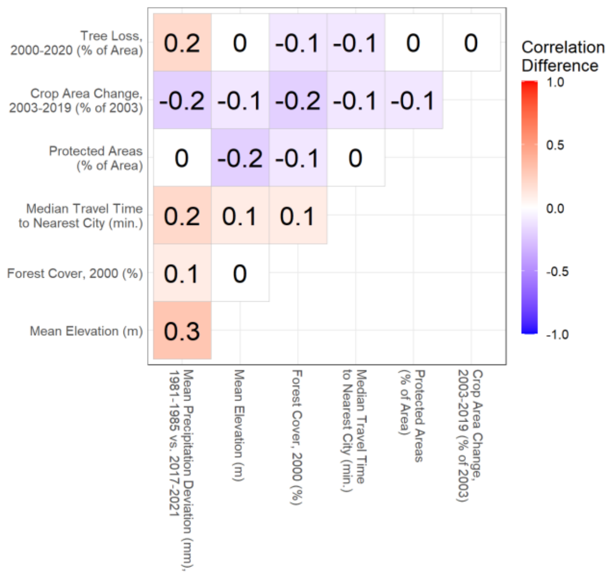

Figure 5. Differences between the correlations between key contextual variables across all kebeles in the selected regions and the correlations between key contextual variables for the sampled kebeles (that is, positive numbers indicate a higher correlation between the variables in the sample than the population, while negative numbers indicate the opposite).

Figure 5, however, indicates that the correlations observed in the sample are quite similar to the correlations between these variables in the overall population. It presents the difference between the correlations between key contextual variables presented in Figure 3 and those same correlations in the population as a whole. Reassuringly, the differences are not very large, though the fact that the correlation between elevation and precipitation deviation in the sampled kebeles is about 0.3 higher than in the population as a whole suggests there could be a greater diversity in experiences of these variables across the study areas than the sample fully captures.

Tanzania

Turning to Tanzania, Figure 7 presents the distribution of the key contextual variables for the sampled villages in the country as compared to the population of villages in the target region. As in Ethiopia, there is, by design, substantial variation on the two most important variables - elevation and precipitation deviation. Some of the other variables, however, are more constrained than in the Ethiopian sample. The range of forest canopy cover, for example, is lower in the sampled villages in Tanzania, as, generally, is protected area coverage. As in the Ethiopian sample, we see some dramatic (relative) changes in cropland in some villages, though again from quite low initial levels.

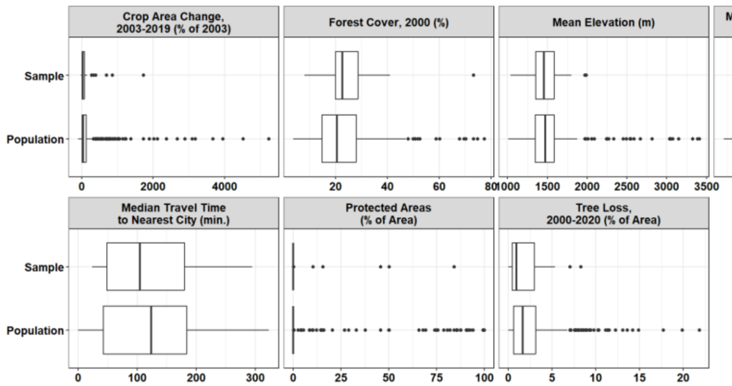

Figure 7. Boxplots of key contextual variables for sampled villages in Tanzania as compared to the overall population of villages in the selected study regions. The left-hand side of the box designates the lower quartile of the variable, the line in the middle of the box the median, and the right-hand-side of the box the upper quartile. The whiskers on the box show all observations within 1.5 times the interquartile range of the sides of the box, while the dots show individual outliers.

Figure 7 also suggests that the sampled villages are broadly representative of the population of villages in the study regions in Tanzania. By design, the distribution of villages by elevation is very similar to the population, though the sample maximum is lower than some substantial outliers in the population, largely due to these region’s inaccessibility to the research team. The final selection also excludes some of the highest and lowest values of precipitation deviation observed across the population of villages in the study regions. These limitations must be considered when generalizing the results of the survey, which may not be fully representative of dynamics in a handful of geographically unique villages in the study regions.

Variable	t-statistic	t-statistic p-value	Kolmogorov-Smirnov Statistic	KS-statistic p-value
Crop Area Change, 2003-2019 (% of 2003)	0.85	0.4	0.12	0.78
Forest Cover, 2000 (%)	-0.96	0.35	**0.24**	**0.058**
Mean Elevation (m)	0.74	0.46	0.098	0.94

Mean Precipitation Deviation (mm), 1981-1985 vs. 2017-2021	-1.6	0.12	0.2	0.17
Median Travel Time to Nearest City (min.)	-0.07	0.94	0.15	0.54
Protected Areas (% of Area)	-0.23	0.82	0.08	0.99
Tree Loss, 2000-2020 (% of Area)	1.5	0.13	0.17	0.32

Table 2. Comparison of key contextual variables for sampled villages to the population of villages in the selected regions in Tanzania as a whole. T-statistics compare population to sample mean (i.e., if the t-statistic is negative, this implies that the mean is higher in the sample than the population) and Kolmogorov-Smirnov statistics test the hypothesis that the sampled distribution of each variable differs from the population distribution. Boldfaced values are statistically significant at the 0.05 level.

Table 2 presents t-tests comparing the population to the sample mean and Kolmogorov-Smirnov tests comparing the sample and population distributions for each of our key contextual variables. As in the Ethiopian case, we find no statistically significant differences in the means of any of our contextual variables between the sample and the population. However, we do find one instance where the Kolmogorov-Smirnov test is statistically significant: the distribution of forest cover in the sample differs substantially from the population. Referring to the boxplots in Figure 7, it is relatively clear that the variation of forest cover in the sampled villages is lower and includes fewer extreme values than that in the population of villages as a whole. Similarly to our potential concerns with the sampled values of elevation noted above, the survey may not fully generalize to outliers on forest cover, which could represent areas with more potential land into which coffee or other crops might expand. As might be expected, there is also considerable overlap in the villages that are population outliers on elevation and forest cover.

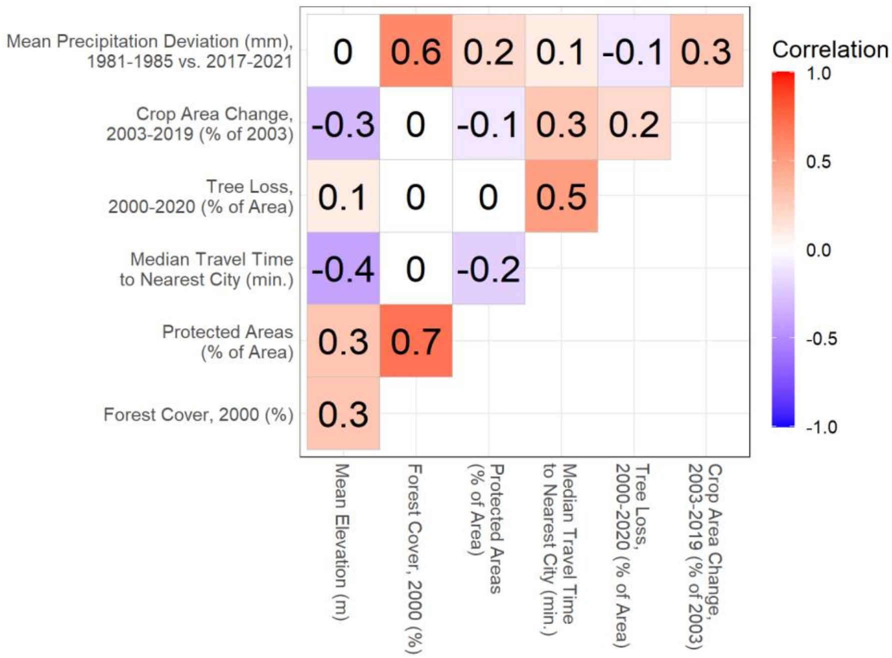

Figure 8. Correlations between key contextual variables for sampled villages in Tanzania.

Figure 8 shows that the correlations between the key contextual variables in Tanzania are generally low, with a few notable exceptions, some of which are quite surprising. First, accessibility to cities of more than 50,000 people in this sample appears to be negatively correlated to forest loss (that is, forest loss, as well as cropland increase, is more common further from these cities). This could in part be due to a slight correlation between proximity to cities and the presence of protected areas, likely driven in particular by the protected areas around Mount Kilimanjaro. As would be expected, areas in higher elevations tend to be less accessible to cities and to have higher forest cover, more areas under protection, and lower rates of change in cropland extent.

Very interestingly, but fortunately for the purposes of inference, we find no correlation between mean elevation and precipitation deviation in the sample. This is a very interesting contrast with the Ethiopian case, where the two variables were strongly correlated. As can be seen in Figure 9, which compares the correlations between the key contextual variables in the population to the correlations in the sample, this (non-)relationship is likely a result of the sampling strategy. The correlation between the two variables is about 0.3 higher in the sample than the population, a relationship that again appears to be driven by a relatively small number of villages that are population outliers in terms of elevation, a group that also accounts for much of the difference in the population and sample correlations between elevation and protected areas and travel time to the nearest city and forest cover. Again, these findings suggest that researchers should take special care when considering generalizing the survey findings to the highest-elevation villages in Tanzania.

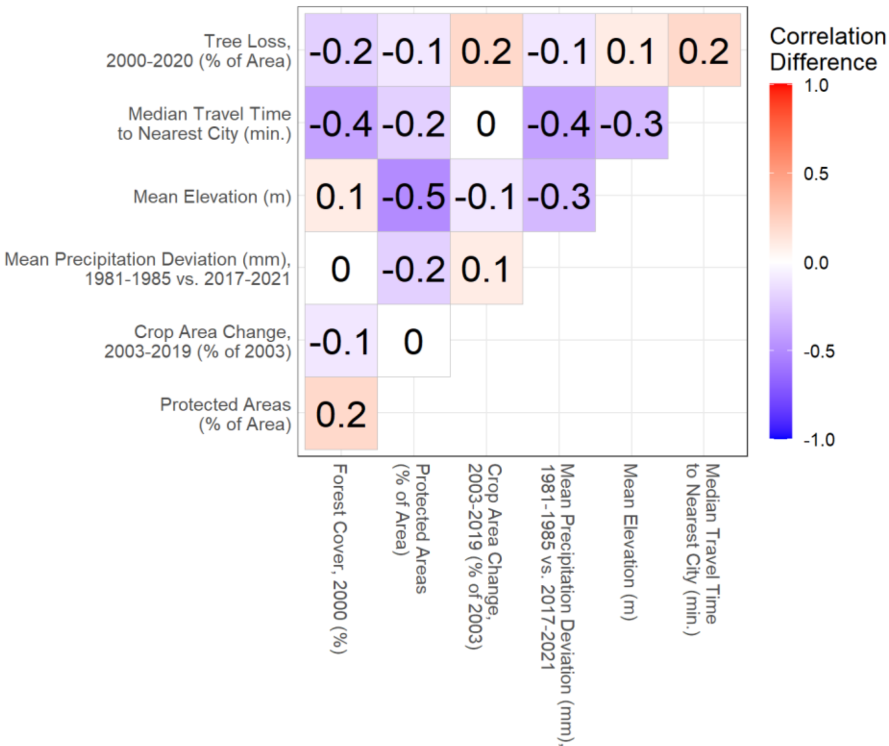

Figure 9. Differences between the correlations between key contextual variables across all villages in the selected regions of Tanzania and the correlations between key contextual variables for the sampled villages (that is, positive numbers indicate a higher correlation between the variables in the sample than the population, while negative numbers indicate the opposite).

Conclusion

This working paper has presented the rationale behind the sample selection strategy adopted by the PACSMAC project. Using a variety of geospatial data, we have demonstrated that the sampling strategy has resulted in a set of sample communities in the two countries that are broadly representative of elevation and precipitation changes in Ethiopia and Tanzania's smallholder-coffee-producing regions as a whole. Furthermore, the samples are also representative of the population of smallholder-coffee-producing communities in the two countries on some other important contextual variables, as well. Still, there are some respects in which correlations between key variables or cases where the sample deviates from the population merit further consideration as the fieldwork and subsequent analysis proceed, particularly in the case of the generalizability of the survey findings to the highest-elevation villages in Tanzania.

Works Cited

Bunn C, Läderach P, Jimenez JGP, Montagnon C, Schilling T (2015). Multiclass classification of agro-ecological zones for Arabica coffee: an improved understanding of the impacts of climate change. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0140490

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342(6160): 850-853.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. Reviews of Geophysics, 45(2). https://doi.org/10.1029/2005RG000183

Funk, C., Peterson, P., Landsfield, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Scientific Data, 2, 150066.

Malaria Atlas Project (2015). Global travel time to cities. Malaria Atlas Project Platform. https://data.malariaatlas.org/

Ovalle-Rivera O, Läderach P, Bunn C, Obersteiner M, Schroth G (2015). Plant phenotypic plasticity in a changing climate. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0124155

Potapov, P., Turubanova, S., Hansen, M. C., Tyukavina, A., Zalles, V., Khan, A., Song, X.-P., Pickens, A., Shen, Q., & Cortez, J. (2022). Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. Nature Food, 3, 19-28.

Tanzania Coffee Board (2019). Tanzania Coffee Industry Profile. https://www.coffeeboard.or.tz/tzcoffee_%20profile.php

UNEP-WCMC and IUCN (2021), Protected planet: The world database on protected areas (WDPA) and world database on other effective area-based conservation measures (WD-OECM). Cambridge, UK: UNEP-WCMC and IUCN. www.protectedplanet.net.

Appendix

Appendix 1: Sampled Kebeles - Ethiopia

District	Kebele
Ale	Kundi
Ale	Jeto Koyami
Ale	Sambe Enole
Ale	Keto Gelecho
Ale	Gumero Abo
Ale	Yobi Mari
Goma	Tesosedecha
Goma	Ketabero
Goma	Koyuseje
Goma	Genjailbu
Goma	Kadimesa
Goma	Getobore
Yayu	Geri
Yayu	Wabo
Yayu	Bondawo
Yayu	Achebo

Yayu	Hamuma
Yayu	Aredin Onigo
Limu seka	Mero Chisa
Limu seka	Sacheni
Limu seka	Atnago Town
Limu seka	Dale Wadera
Limu seka	Gejib
Limu seka	Koma
Gera	Wanija Kerisa
Gera	Sed Loya
Gera	Kola Kinbibit
Gera	Kele
Gera	Genida Chala
Gera	Gure Dako

Appendix 2: Sampled Villages - Tanzania

District	Village
Mbozi	Ipyana
Mbozi	Iyula
Mbozi	Ilomba
Mbozi	Mpito
Mbozi	Halungu
Mbozi	Igamba
Mbozi	Itentula
Mbozi	Nambinzo
Mbinga	Litembo
Mbinga	Mnyangayanga
Mbinga	Maguu
Mbinga	Utiri
Mbinga	Mkumbi
Mbinga	Ukata
Mbinga	Ngima
Mbinga	Buruma

Rombo	Alleni Chini
Rombo	Manda Chini
Rombo	Mengwe Chini
Rombo	Shimbi Kati
Rombo	Machame Aleni
Rombo	Makiidi
Rombo	Mamsera Kati
Rombo	Mamsera Juu
Kyerwa	Kamuli
Kyerwa	Nyakatuntu
Kyerwa	Kikukuru
Kyerwa	Kakerere
Kyerwa	Iteera
Kyerwa	Kibare
Kyerwa	Karukwanzi A
Kyerwa	Murongo