Riverscope Rating System

Methodology for Hydropower Risk Similarity Ratings

TMP Systems

Ben Bowie, Luan Mans, Lorna Backhouse, Tin Garcia, Benedict Lever

Riverscope version 1.1

This document provides information on how the underlying indicator values are rated and combined to produce the various Similarity Ratings for Riverscope.

Contents

1. Introduction

Riverscope is a system that combines both a quantitative and qualitative assessment in order to measure social and environmental risk specific to hydropower investments. The system produces a risk report that is valuable to all stakeholders, including business, providing insight into both the costs and benefits of proposed dams in order to improve stakeholder decision-making processes.

Riverscope provides a geographically specific assessment of environmental, social and governance (or ESG) risks based on sub-national data wherever possible and how those risks impact the commercial viability of the project. Its overall method is analogous to approaches developed by criminologists to forecast crime risks based on contextual factors, and in particular to Risk Terrain Modelling (RTM) which provides a methodology to "articulate vulnerable places at the micro-level."[1]

TMP's Riverscope applies an adapted approach, adopting some of the concepts found in RTM and combining them with our own analysis to enable interpretation of geospatial data about political, social and environmental issues concerning hydropower investment. It is designed to assist both companies and investors in emerging markets to prepare a proper assessment for a specific hydropower investment or alternatively, to rapidly assess a portfolio of hydropower investments.

This document is designed to lay out the approaches and methodologies used to develop Riverscope.

The first step of the Riverscope approach is the Rapid Assessment (RA). This assessment relies on geospatial comparative statistical analysis to determine the potential for risk posed by various Environmental, Social and Governance (ESG) factors. This analysis relies on our experience that identified locations with similar macro and micro conditions based on ESG factors, display the same level of risk for new or existing land-based investments.

The RA produces an Overall Similarity Score which is used to predict the potential delays that an investment can experience. These are then investigated through the Expected Delays model which, at the same time quantifies loss through a unique Discounted Cashflow Model (DCM). This DCM works by incorporating delays into the calculation of critical financial metrics.

In the final Financial Model these dam delay statistics are combined with assumptions about cost overruns and different discount rates in order to produce assessments of a project's Levelized

[1] Kennedy, L.W. and Dugato, M., "Forecasting Crime and Understanding its Causes. Applying Risk Terrain Modeling Worldwide", 2018, European Journal on Criminal Policy and Research, https://doi.org/10.1007/s10610-018-9404-3

Cost of Electricity (LCOE), Net Present Value (NPV) and other metrics widely used by investors and other stakeholders.

This comparative statistical analysis is then complemented by a qualitative review via the Deep Dive (DD) process. The DD ensures both the accuracy of the RA while also highlighting any model limitations and, in response, providing oversight into other avenues of data, investigation or analysis that we can use to strengthen or validate the assessment.

This Riverscope Methodology document is split into 3 sections:

i. Riverscope Statistical Model: This section outlines the geospatial statistical analysis that we have used to rapidly identify potential risk. The output of this model is an Overall Similarity Score.

ii. Expected Delays Model: This section outlines the statistical analysis for estimating delay, which is built on work done by TMP Systems and the ODI. The section describes how we determine model inputs and final outputs, or in other words the relationship between the similarity scores and potential delays.

iii. Financial Model: This section describes how the model combines dam delays with a DCM in order to produce an output of a project's LCOE, NPV and other widely used and understood financial metrics. In the process, the section explains how we deal with cost overruns and discount rates.

Riverscope is currently in the early stages of development and we expect to strengthen and improve this methodology over time.

Riverscope is not designed to replace expert insight, nor eliminate the need to invest time and money to understand the human factors that impact an assets' performance. Its purpose is rather to help users to structure this process, to rationalize costs against potential losses and to make due diligence processes more efficient, comprehensive and effective.

2. Riverscope Statistical Model

The Riverscope statistical model is a system that applies a new approach that builds on TMP's Landscope system[2] and tailors it to analyze the political, social and environmental factors which will impact a hydropower asset. It is designed to help companies and investors in emerging markets to prepare a proper assessment for a specific hydropower investment or alternatively, to rapidly assess a portfolio of hydropower project investments.

By using a statistical analysis approach, the model rapidly provides quantitative evidence to indicate levels of potential risk for a hydropower investment. Although hydropower creates significant ESG risks in general, each investment impacts the surrounding communities and environment differently. However, our analysis of ESG factors commonly associated with hydropower investment risks allows us to rate these factors on a unified scale. This provides us with both an intuitive sense of the overall level of similarity to problematic hydropower investments and also gives us insight into the extent to which different ESG factors contribute to that overall similarity.

We define "problematic dams" (hereafter "Test cases") as dams that have been notably problematic, for either being strongly linked to significant environmental impacts or for their impacts on local communities. In most cases, these dams have experienced delays either due to protest action or complications with their financing but were in all cases exposed to incidences of conflict. The "non-problematic dams" (hereafter "Control cases") were identified by leading experts as dams that caused relatively low levels of impact to both their immediate environment and to surrounding local communities.

To understand the issues related to hydropower investments we have reviewed and analyzed over 1,100 dams from the GRanD dam database. [3] By increasing the number of dams we analyze over time, we will be able to improve the statistical basis of Riverscope. For Riverscope v1.1, we have identified 17 indicators at Dam, River and District areas[4].

Each of our indicators are rated on a scale of 0-100, with 0 representing indicator levels rarely or never associated with the Test cases, and 100 representing indicator levels strongly associated with the Test cases and often or always found in places where problematic dam investments have occurred.

---

[2] https://landscope.info/

[3] The data set of cases was not exclusively hydropower dams, but also included multipurpose dams.

[4] That is, we have 17 indicators, some of which figure at multiple levels. So the Global Sediment indicator, for example, operates at Dam, River and District level. See table 1 for detail. More information relating to the Indicators used in Riverscope's model are listed in Appendix I of this document.

These indicator scores are then weighted depending on their location relative to the dam (geospatial importance), i.e. Dam, River or District level[5] and are also dependent on the quality or robustness of the data for that indicator. These scores are then combined to create a Context Factor Rating.

Finally, the scores are weighted for the region within which they fall (Africa, Latin America or Asia) in order to produce a 'Relative Similarity Rating' for each indicator. This reflects the fact that there are discernible differences between average values we see in each region as well as for the performance of dams by region. In Riverscope these 'Relative Similarity Rating' indicator scores are combined to produce a classification for the Overall Similarity of the selected location.

The following sections will focus on the make-up of Riverscope's Statistical Model and how its components function together. At the same time, it will explain the reasoning for our selection of indicators and the way in which we convert these indicators into ratings.

First, we will break down how we developed the bespoke area approach that we use to identify significant indicators for the Dam, River and District areas. Second, we will explain how we went about selecting the 17 sub-national indicators that Riverscope relies on in order to produce a rating for the location in question. Third, we explain the methodology behind normalizing and weighting raw data in order to produce the Context and Relative Similarity Ratings, which are combined to produce the Overall Similarity Rating. Finally, we conclude by summarizing the final model and defining the indicators used as inputs while simultaneously providing a statistical view on the outputs of the model.

2.1. Area approach

The impacts of dams are both wide-reaching and quite uneven. We therefore needed an approach for defining the impact area of a dam that could be universally applied to capture these characteristics using the quite granular data available from our indicators. This is rather challenging because, in practice, every dam is different. We have used three generic and distinct levels of analysis, namely: Dam, River and District. Each level of analysis is supported and can be extended via the DD process to enable tailoring.

The area immediately around the dam generally experiences the most intense impact. We have therefore used a 20km circular buffer around the location of the dam for indicators with the highest weighting since it will generally include the major impacts such as human displacement and deforestation. We do recognize that in some cases the inundation area is larger which we pick up in the DD analysis, where relevant.

[5] For example, impacts at and around the dam are more severely felt by communities and the environment than at the district level

Many hydropower impacts are felt more extensively but often less intensively downstream. Our second tier of indicators therefore considers a 100km stretch downstream of the dam or to the coast, with a 10km buffer on either side of the main tributary. This level of analysis considers things like impacts on sediment flows, fisheries and water quality as well as various social factors.

Finally, dams have wide reaching impacts that need to be considered in cumulative terms with other dams developed in the area. Our third tier therefore analyses indicators at the district level, or Administrative Level 2 (GADM L2),[6] to account for this broader context.

Together, these three tiers help us to roughly quantify the risk of incidence in spatial terms. In other words, they can give us a score that indicates which places are most likely to suffer from environmental and social impacts.

Why 20km radius

We found that based on the Control cases, taking a 20km radius would include the inundation area in addition to encompassing the potential surrounding social and environmental impacts, such as conflict and deforestation, which are often associated with dam construction.

This approach is useful for it provides us with a standardized and practical way to compare a large group of cases.

Why 100km downstream and 10km either side of the main river

The most notable impact commonly highlighted is how the dam has affected the downstream river-dependent population whose livelihoods, particularly farming and fishing revolved around the seasonal flow regimes.[7]

We know there are downstream impacts all the way to the ocean and across the country from large scale dam construction, usually owing to the knock-on impact of a dam reducing the food security of poor people that often rely on large rivers.

Specifically, it has been shown that the biodiversity within the river and in the riverine area are affected from the dam's immediate impact all the way downstream to the ocean.[8] This effect is, of course, limited to the area along the banks of the river and we have found that the majority of people effected are within 10km of the river, so we use this buffer on either side.[9]

However, we need to confine the analysis to a realistic distance that could reflect the immediate social and environmental impacts that will be felt both during construction and immediately after

[6] https://gadm.org/

[7] https://www.tandfonline.com/doi/abs/10.1080/00167223.2016.1258318?src=recsys&journalCode=rdgs20

[8] https://hmr.biomedcentral.com/track/pdf/10.1007/BF02414766

[9] http://www.fao.org/3/y3994e0i.htm

commissioning. We have therefore chosen to limit this distance from the immediate impact of the hydropower investment to 100km and 10km[10] on either side of the river.

We believe this is a conservative approach as some dam impacts stretch far beyond 100km downstream. For example, in the northern Cambodian town of Siem Pang on the Mekong river, some 5,000 people were displaced about 130 kilometers downstream from a completed hydropower project[11].

Why not upstream

The direct downstream effects of dams became our point of focus in part because they receive the most attention from ecosystem managers and researchers. Also, in many cases of hydropower development the natural riverine conditions around the upper reaches of dam reservoirs and further upstream remain largely unchanged.[12] Although our RA does not include upstream impacts beyond the dam site's 20km buffer, our DD analysis accounts for any novel upstream impacts, such as the impacts on fish migrations mentioned[13].

2.2. Approach to model development

The approach we used to develop our model was based on Landscope, which identified the difference in ESG indicators between a Test and Control group of land tenure disputes. By identifying these differences, this approach allows new locations to be categorized according to their level of similarity with either the Test or Control group, which in turn provides an indication as to the level of ESG risk for that location. Similarly for Riverscope, we therefore conducted analyses of the association between the Test and Control cases of dams using ESG indicators that provide sub-national granularity.

Our statistical assessment considered more than 300 indicators from 51 Datasets[14] for ESG significance and resulted in 76 indicators at a regional level and 14 at a national level. After the analysis of the regional and national indicators, 17 indicators were identified at a sub-national level, each showing a direct overall significance between the Test and Control cases at either one or multiple of the Dam, River or District areas.

[10] It has been shown that only 10% of the population lives further than 10 km away from a surface freshwater body. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3110782/

[11] https://www.voanews.com/east-asia-pacific/laos-suspends-new-dam-projects-following-catastrophe

[12] http://www.fao.org/3/y3994e0i.htm#:~:text=Dams%20alter%20aquatic%20ecology%20and,life%20span%20of%20the%20dam

[13] https://pubmed.ncbi.nlm.nih.gov/16705984/

[14] See full list of Databases considered in Appendix V

In terms of the "freshness" of the data, we used the most recently available data from each source. The latest available data ranges from 2006, when the oldest datasets were developed, to 2020, where a majority of the datasets we used were developed from 2016 onwards.

Apart from the "freshness" of the data, the availability of data at the correct granularity was also important to ensure the validity of the model. This was however not a limiting factor as the many databases available to us provided a wide range of data which enabled us to source the correct granularity to use in the model. Although we relied on data that ranged from a national level granularity to a granularity of $1km^2$, we used the highest level (smallest area) of granularity in the models we constructed.

Not all indicators had values recorded for all the dams. In general, if more than 60% of the dams had missing data then the indicator was excluded. Furthermore, if an indicator had missing data for more than 80% of the Test cases or 80% of the Control cases, then it was excluded.

It is important to note at this point that the model developed relies on the amount of data used to determine similarities between locations. In other words, the more data used (in this case the number of Test and Control dams analyzed) the more accurate the model will become. Our sample of Test and Control cases was substantial, but only covers a fraction of all the hydropower dams globally. We recognize that the model will evolve as additional data is included in future but understand that through the verification of the DD process, we are able to justify the outputs of the model.

2.2.1. Sub-national indicators

Our indicators are derived from datasets that show a statistically significant association between our Test and Control dams. The datasets we chose cover a wide range of environmental and social factors from sediment flux, water variability and protected areas to poverty and conflict. For these assessments we compared the average or total values[15] for the area surrounding Test dams with the areas surrounding Control dams.

We identified indicators that would be powerful for predicting risk by using standard statistical techniques to compare the indicator values in the Test dam group with the values for the Control group. This produced a list of indicators that had a statistically significant relationship between the Test and Control cases.

[15] Averages were used for data where scores (typically continuous) or ratings are applied to each location in an area, such as poverty rates or the rate of change in population levels, whereas in cases where data consisted of large numbers of 'real' objects (such as numbers of people or conflict events), counts were used.

We consistently found statistically significant associations between the Test and Control groups in six of the tested datasets that appeared in at least two of the three areas of analysis[16]:

- Oxford Poverty and Human Development Initiative's Multidimensional Poverty Index
- Gridded Population of the World v4
- Modeled Global Suspended Sediment Flux
- Global Drainage Basin Database
- IUCN Red List Species Database
- Blue Water Scarcity Database

There were also four additional datasets included in the model that showed specific significance of at least one of the three areas of analysis:

- Earth City Lights Database
- World Resource Initiative
- Global Land Cover database
- The World Database on Protected Areas

We selected the sub-national indicators for the overall similarity risk model based on the strength of their significance in terms of the impact in variance between the Test and Control cases, alongside an evaluation of which indicators provided the greatest breadth of risk factors.

While we found statistically significant associations with a large number of indicators from these and other datasets, we did not use all of them in our risk model. Firstly, we needed to ensure the model covered all possible risk factors; for example, the Species richness and protected areas were both significant indicators yet not as strongly significant as all the other MPI indicators. To ensure the model covered the greatest breadth of risk factors, the species richness and protected areas were included with some, but not all of the MPI indicators.

Secondly, a large number of indicators (primarily the MPI indicators) are highly collinear, which means there is a large amount of correlation among the indicators themselves, independent of their risk associations to the significance. These indicators show the same variation between the Test and Control cases therefore including them all in a rating model would not increase the power or efficacy of the model and may in fact skew the results.

We conducted extensive testing to reveal the exact extent of this collinearity. Multicollinearity tests were run firstly across all indicators within each area (i.e. Dam, River and District) and then

[16] Please see full citation list in Appendix III

across the three different areas. In each case the Variance Inflation Factor (VIF) of each indicator was calculated to identify which indicators should be included. The process of selecting the indicators and the relevant weightings for each indicator was an iterative process.

To refine our model and improve its accuracy, we also carried out an indicator assessment process. This assessment entailed individual analysis of each indicator within each case (i.e. Test and Control cases) to determine whether the case reflected the expected similarities to their respective Test or Control group. Where a case reflected indicator scores that were statistically different to their respective group (either Test or Control group), then the case was removed and investigated further. We removed a total of 9 cases from the Control group which showed statistically significant similarities to the Test group, and after further investigation proved to be problematic.

2.2.2. Indicators used in the Similarity Rating Model

The indicators that are used in the Similarity Rating Model are drawn from a broad spectrum of social, economic, political and environmental factors. We could have taken the usual ESG approach here, and put all the indicators into environmental, social or governance buckets, but in our view this taxonomy is not well-suited to the reality of the impacts that are experienced by hydropower investments in emerging markets.

Rather, we have thought about the problem from the perspective of where the impacts could occur and tried to determine which indicators describe conditions in each area as outlined earlier (Dam, River and District). Since the model is trying to determine the possibility of risk at these locations, it is important to understand how these risks could impact the hydropower projects so that the analysis can reflect the real risks.

Table 1 overleaf shows the specific indicators in each group that are used in Riverscope's overall similarity rating model.

Table 1: List of statistically significant indicators for Environmental and Social issues. Indicator weightings[17] Dam = 3; River = 2; District = 1.

Indicator	Indicator Description	Dam	River	Distr.
Environmental				
Minimum Percentage Water Scarcity Over the Year (Blue Water Scarcity Database)	WaterStat is the world's most comprehensive water footprint database. The minimum percentage water scarcity indicator is a unique dataset showing blue water scarcity in the world on a monthly basis at high spatial resolution. Blue water, or liquid water, can be compared with green water, in soil moisture and similar. We found that greater scarcity equated with greater risk.			
Species Richness that are Critical, Endangered, Vulnerable (IUCN Red List Species Database)	The IUCN Red List is a critical indicator of the health of the world's biodiversity. The species richness provides an indication of the number of species potentially occurring in a given location. We found that places with high species richness equated with higher risk.			
Global Sediment Flux (Modeled Global Suspended Sediment Flux)	The WBMsed model is a spatially and temporally explicit (pixel scale and daily) global sediment flux model. The indicator uses the modeled values to understand the global flow of sediments within the rivers which reflects the amount of nutrients and minerals that flow downstream to support the growth of life. Our assessment found that greater disruption of sediment flows makes dams more risky.			
Inter-Annual Variability (Aqueduct Global Maps)	Aqueduct's global water risk mapping tool helps companies, investors, governments, and other users understand where and how water risks and opportunities are emerging worldwide. The indicator measures the average between-year variability of available water supply, including both renewable surface and groundwater supplies. We found that higher variability was correlated with higher risk.			
Upstream Drainage Area (Global Drainage Basin Database)	The GDBD is a database made up of six GIS data collections (drainage basin boundary data, river network data, discharge gauging station data, natural lake data, dam lake data, and flow direction data) that store a wide range of information on natural and social sciences. Upstream drainage area provides an indication of the			

[17] The weightings here are averaged for the specific area. For indicator specific weightings see the Appendix IV

	drainage basin into a particular river. The larger the area, the higher the risk.			
Protected Areas (World Database on Protected Areas)	This indicator is part of the most comprehensive global database on terrestrial and marine protected areas. It shows the proportion of the analyzed area covered by a protected area designation (such as national parks or conservation zones), as a percentage of the analyzed area. The higher the proportion of the impact zone that is covered by protected areas, the greater the risk for the dam.			
Percentage Cropland (SEA CCI)	The CCI-LC project delivers consistent global LC maps at 300 m spatial resolution on an annual basis from 1992 to 2015. This indicator draws on the area that is considered as cropland within the dataset. Counter-intuitively, we found that lower percentages of cropland correlated with higher risk, indicating risk in remote areas.			
Drought Severity (WRI)	The indicator considers the average length of droughts times the dryness of the droughts from 1901 to 2008 to develop a drought severity score in a specific area. The more problematic drought, the higher the risk for the dam.			
Social				
Percentage of People Who Are Poor and Deprived in Living Standards: Improved Sanitation (Multidimensional Poverty Index)	A person is considered to have access to improved sanitation if the household has some type of flush toilet or latrine, or ventilated improved pit or composting toilet, provided that they are not shared. We found that higher levels of deprivation were associated with higher levels of risk.			
Percentage of People Who Are Poor and Deprived in Education: Schooling (Multidimensional Poverty Index)	The MPI uses two indicators that complement each other: one looks at completed years of schooling of household members, the other at whether children are attending school. The better the school attendance, the lower the risk.			
Percentage of People Who Are Poor and Deprived in Living Standards: Drinking water (Multidimensional Poverty	A person has access to clean drinking water if the water source is any of the following types: piped water, public tap, borehole or pump, protected well, protected spring or rainwater, and it is within a distance of 30 minutes' walk (roundtrip). Higher levels of deprivation associate with higher risk.			

Index)				
Multidimensional Poverty Index of the country (Multidimensional Poverty Index)	Measures acute poverty: the proportion of people who experience multiple deprivations and the intensity of such deprivations. Higher levels of deprivation, again, associate with higher risk.			
Population Vulnerable to Poverty (Multidimensional Poverty Index)	Identifies a threshold score for the MPI indicators to suggest people are vulnerable to becoming impoverished and should conditions not improve, will fall into severe poverty. Higher levels of vulnerability correlate with higher risk for the dam.			
Population density (GPWv4)	The Gridded Population of the World, Version 4 (GPWv4) consists of estimates of human population (number of persons per pixel), consistent with national censuses and population registers, for the years 2000, 2005, 2010, 2015, and 2020. To our surprise, lower population densities correlate with higher risks.			
Night Lights (Earth City Lights Database)	This image of Earth's city lights was created with data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS). The brightest areas of the Earth are the most urbanized, but not necessarily the most populated. To our surprise, lower light levels correlate with higher risks, suggesting that remote areas are problematic.			
Conflict Number of Explosions & Remote Violence (ACLED)	ACLED collects real-time data on the locations, dates, actors, fatalities, and types of all reported political violence and protest events across the world. This indicator considers all Explosions and incidences of Remote Violence. The more frequent the incident of violence, the higher the risk.			
Conflict Events including Protests, Strategic Developments and Riots (ACLED)	ACLED collects real-time data on the locations, dates, actors, fatalities, and types of all reported political violence and protest events across the world. This indicator considers Protests, Strategic Developments and Riot events. The more frequent the incident of violence, the higher the risk.			

2.2.3. Similarity rating process

Riverscope provides one of five possible classifications for hydropower dam impacts based on the area being analyzed. This is based on the similarity between the indicator profile of the queried location, and of the indicator profile of locations where dams have experienced problems in the past. These 'Overall Similarity' classifications are:

Not Similar, which is when we see that the data in the queried location looks very different from places where ESG risk has been a problem.

Low-level Similarity, when there is a low level of comparison between the specified location and places with ESG risk issues.

Medium-level Similarity, this is when there is a broad similarity between highly problematic hydropower investments and the place you chose, but not to the degree where it's alarming.

High-level Similarity, which means exactly what it sounds like: the queried location looks very much like places where ESG risk has been a problem which resulted in major delays.

Inconclusive, this either means that the data patterns aren't clear, or we don't have enough data to have full confidence in the results.[18]

Riverscope produces these classifications by processing indicator data in five steps:

1. **Indicator Scoring**: calculates scores of 0-100 for each of the 17 indicators in the queried location.
2. **Context Factor Rating**: groups Indicator scores into three baskets – Dam impacts, River impacts, and District impacts and weights them according to their level of impact relative to location of the dam wall.
3. **Relative Regional Weighting:** weights the Context Factor Rating scores for the queried location according to the continent it sits in. These weightings were derived from the Landscope model.
4. **Overall Similarity Rating**: takes the Relative Regional Weighting scores and combines the three location scores to calculate the Overall Similarity Rating score.

[18] In some instances, where the underlying datasets do not cover the area selected by the user and Riverscope lacks sufficient indicator data to make a reliable assessment, Riverscope will return an 'Insufficient Data' output.

2.2.4. Indicator scoring

Sub national indicators

We have developed a replicable, standardized method for scoring underlying indicator values. This can be applied to any range of continuous data, and results in a standardized scoring of between 0-100, where 0 represents the lowest possible association with cases of hydropower that were highly problematic, and 100 the highest. This will make it relatively straightforward to introduce new indicators into the similarity rating methodology, as better data becomes available.

Each indicator is standardized by modelling the distribution of the values. For simplicity, the distributions were limited to Weibull[19], Normal[20], Lognormal[21] and Uniform[22]. For indicators with extreme values, the indicator was first transformed using a natural logarithm with the resulting distribution then modelled. The reason for this was to ensure that extreme values did not mask significant differences between the Test and Control cases. An example of this is population density.

To calculate how a given indicator value should be rated on the 0-100 scale, we group the dams into two groups. We look at the distribution of values in places where problematic dams have occurred, and compare it with the distribution of values in places where dams which have already been built have not resulted in environmental or social issues.

Figure 1, below, illustrates the rating process. To understand the distribution of indicator values we divide the range of indicator values of each group into quartiles (i.e. with a quarter of the total number of values in each). In Figure 1, the boxes represent the 1st and 3rd quartiles and the line within the box, the 2nd quartile (i.e. the median or middle value). The 'whiskers' represent the minimum and maximum values excluding outliers. Outlier values are marked as dots.

In this example, a quarter of the places where problematic dams had not occurred (the 'Control locations' on the left) had an indicator value of less than 2. For places where problematic dams *had* occurred ('Test'), a quarter of all locations had indicator values greater than 3.8. This tells us that indicator values higher than 3.8 should receive a correspondingly higher rating on the 0-100 scale, as they are much more common in dispute locations than in non-dispute locations.

[19] https://www.weibull.com/hotwire/issue14/relbasics14.htm

[20] https://www.statisticshowto.com/probability-and-statistics/normal-distributions/

[21] https://www.statisticshowto.com/lognormal-distribution/

[22] https://www.britannica.com/topic/uniform-distribution-statistics

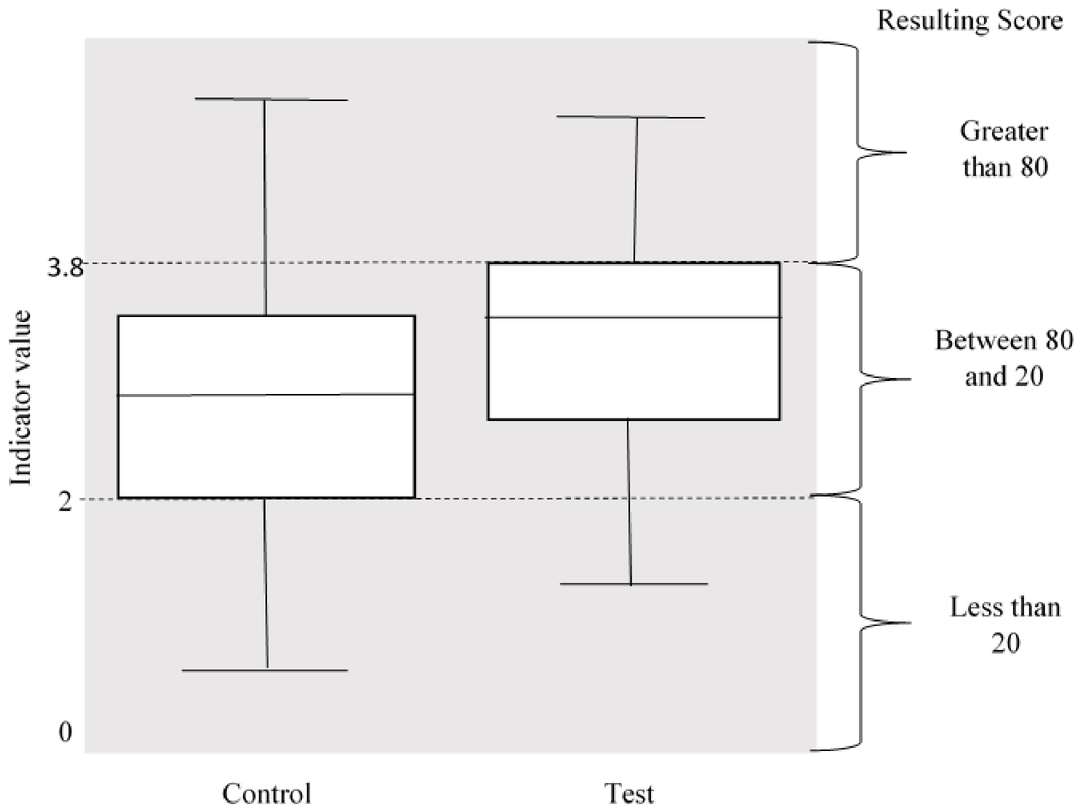


Figure 1: Illustrative example of the application of Riverscope scores to underlying indicator values by comparing project locations where problematic dams have occurred to project locations where they did not

Figure 1 illustrates exactly how the ratings values (to the right of the chart) are applied to indicator values based on the different distributions between control and test groups. There are five 'bins' of ratings values (0-20, 20-40, 40-60, 60-80, 80-100). The dotted lines show how the quartiles of indicator values translate to the thresholds of those bins.

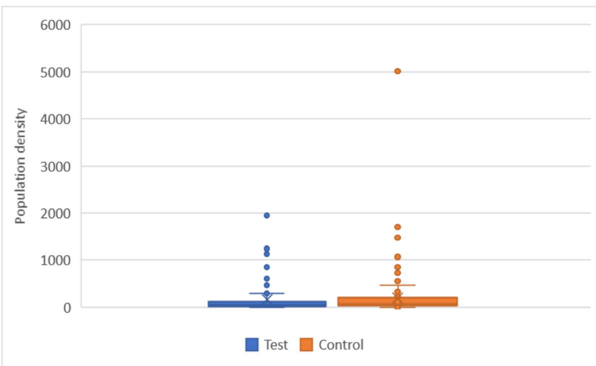
The idea is to ensure that:

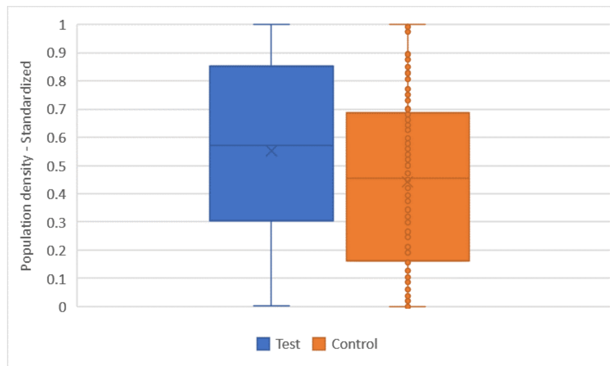
- values less than the first quartile of the places where Test dams did not occur get a rating less than 20;
- values larger than the third quartile of the places where Test dams did occur get a rating larger than 80; and
- values in-between these are distributed evenly using linear interpolation within each bin.

Within each rating bin – e.g. between 0 and 20 – indicator values are distributed evenly. So, in the example in figure 1, if the range of values in that bin is from 0 to ~2, a value in the middle of that distribution (i.e. 1) would get a rating of 10. And a value of 1.5 (halfway between 1 and 2) would get a rating of 15, and so on.

Our testing found that using five ratings bins gave optimal predictive power to the model. With greater than five, the accuracy of the rating (as shown by the difference in rating between control and conflict locations) did not significantly change. But with fewer than five, the differences in ratings started to become less reliable.

The quartile values are calculated from the standardized distribution for each indicator and not from the raw data. This was done to ensure that indicators with extremely large values did not mask the significant difference which exists between the Test and Control cases. The box plot below illustrates this reasoning for the indicator Population density.





Some of the indicators had to be inverted. That is, for some indicators, the Control dams had a higher value than the Test dams. In these cases, the scores are inverted after a thorough investigation of the real impact of that indicator to ensure that their overall score was consistent.

We applied this approach to each sub-national indicator used in the model where there was a proven significant difference between Test and Control dams.

Context factor rating

The context factor rating combines the 17 indicator scores, to provide a single rating that reflects the range of key factors that influence tenure disputes.

The indicator selection process provides a list of n ESG indicators that can be used to model similarity to tenure dispute locations. The overall similarity rating, S, for a location is calculated as a weighted sum:

$$S = B_c + A_c(p_r(\sum_{i=1}^{n} w_i x_i)$$

Where $(w_i)_{i=1}^{n} \in [0,1]$ are fixed weights that sum to 1, A_c is the coefficient depending on the region, chosen, by linear interpolation, to reflect the score relative to the region and $(x_i)_{i=1}^{n} \in [0,100]$ are the indicator scores for each identified ESG indicator at the location.

Weighting of the indicators:

The weights p_r is the weight assigned to the sub-national score.

The weights (w_i) for each indicator were determined from the area in which they occurred, i.e. Dam, River or District to which we then added an assessment of the robustness of the data for each indicator. It was decided that the highest impact is in the immediate vicinity of the dam and hence the indicators applied to the 20km buffer around the dam received the highest weighting. The downstream area received the next highest weighting with the district where the dam is located receiving a weighting a third relative to that of the dam's impact.

In summary, the weightings for the Dam indicators were 3 times greater than the District weightings and 1.5 times greater than the downstream (River) weightings. Within each area the weights were adjusted for the robustness and quality of the data. Full details of the weights given can be found in Appendix IV.

Missing data was handled by adjusting the weights to sum to 1. A dam with a missing value for any indicator was not penalized for insufficient data, rather the weighting was adjusted to omit this value.

Relative regional weighting

We needed to process the indicator scores to reflect the variable quality and consistency of the raw data that goes into them. For example, while the MPI indicators that we use may be theoretically globally consistent and comparable, there will be some differences in the way that the underlying data is collected and reported from country to country.

To account for this issue, the Context scores are weighted according to the continent it is in. Each country therefore has a relative similarity weight, which describes how high the ratings are, on average across the whole country, in comparison to the regional (Africa, Asia, Latin America) average. By combining the queried location's combined Indicator Scores with its country's relative risk weight, we were able to calculate the relative regional weighted indicator scores for the three areas – Dam, River, and District.

Overall Similarity classifications

Riverscope combines three area scores from the Relative Similarity Rating for the specific dam queried in order to arrive at one of five possible characterizations of the queried location.

The specific criteria for each Overall Similarity classification is based on TMP Systems' extensive experience in providing in-depth analysis of tenure disputes, and extensive testing of how different levels of Relative Similarity Ratings and Context Factor Ratings align with the on-the-ground realities in places where those disputes happen.

Riverscope's methodology includes a set of calculations that are used to determine which of the mutually exclusive Overall Similarity categories the location belongs to. The summaries below provide a brief breakdown of the classifications, and the logic behind these calculations:

High-level Similarity, where the data profile is extremely similar to places where that has proved to be problematic in the past. This is proven by the Overall Similarity rating producing a potential risk of delay that can be considered severe. A location is classified as High-level Similarity if:

- The Overall Similarity Rating is above 82 resulting in mean potential delays of over 10 years.
- At least half of the indicators are present.

Medium-level Similarity, when there is a broad similarity between the data profile and places where tenure disputes have occurred, but overall Similarity Ratings are not very high. A location is classified as Medium-level Similarity if:

- The Overall Similarity Rating is between 75 and 81 resulting in mean potential delays of between 6 to 10 years.
- At least half of the indicators are present.

Low-level Similarity, where one or two Context Factors are significantly similar to places where tenure disputes have occurred. A location is classified as Partially Similar if:

- The Overall Similarity Rating is between 45 and 76 resulting in mean potential delays of between 1 to 5 years.
- At least half of the indicators are present.

Not Similar, where the data profile is very different from places where tenure disputes have occurred. A location is classified as Not Similar if:

- The Overall Similarity Rating is below 45 resulting in mean potential delays of less than 1 year.
- At least half of the indicators are present.

Inconclusive, where the data profile is not clear enough for Landscope to provide a reliable conclusion.[23] A location is classified as Inconclusive if:

- The data do not meet the criteria for Highly Similar, Similar, Partially Similar or Dissimilar.

[23] In some instances, where the underlying datasets do not cover the area selected by the user and Riverscope lacks sufficient indicator data to make a reliable assessment, Riverscope will return an 'Insufficient Data' output.

- At least half of the indicators are present.

2.2.5. Final model

We found that 17 sub-national indicators showed the highest level of significance (see Table 1 above or Appendix I for a full description of the sub-national indicators). The indicators provide a suitable range of environmental, social and governance factors that were found to be significant to hydropower projects. The modelling process, as described above, produced a set of Overall Similarity Risk scores that provided Riverscope with statistical evidence that a specific area could experience delays.

	Min Score	Median Score	Max Score
Test	31	71	96
Control	22	48	78

The data quality of Riverscope's final statistical model was statistically adequate and provided reliable results. The model produces a score between 0 and 100 that, from a statistical point of view gives an indication of how likely the project is to encounter problems. This score is not a probability or likelihood. The model included 91 Test cases and 180 Control cases[24]. As shown in the table above, the final average scores for the Test dams range from 31 to 96 with the median score being 71 and 22 to 78 with the median score being 48 for the Control cases.

The overall scores tend to indicate a bi-modal distribution for both the Test and Control cases. This is because there appears to be two sets of cases. Those cases with a stronger social weighting and those with a stronger environmental weighting. That is, some of the Test cases have a higher relative score in the environmental indicators and are similar to the set of Control cases.

Again, we emphasize here that we recognize the limitations of data when using a limited sample and its ability to paint a picture that is accurate. That is why we have developed Riverscope to include the DD, a crucial verification step for the RA.

[24] See full list of Test and Control cases in Appendix II

3. Expected Delay Model

Riverscope's financial model is based on research conducted by TMP Systems and ODI over the last four years.[25] This research looks at the way that ESG risks financially impact a project via delay and slippage. It provides investors with a way to quantify the environmental and social risks associated with the investment.

Our research of social risk over 10 years has demonstrated that ESG risks become financial impacts via the delays that they produce, especially when considering large infrastructure projects. For the delay model, we considered 49 cases of hydropower investments that experienced delays during the feasibility, construction or operational phase. The delays identified in the various phases of the investment are considered for investments that may have been cancelled, are in operation or are currently still being delayed.

3.1. Delay model statistical analysis

In the initial analysis we reviewed 91 cases with delays ranging from 3 days to over 30 years across a combination of the feasibility, construction and operational stages of investment. After reviewing each individual case, we narrowed the delay cases down to 49 based on the credibility of information found for each case.

These 49 selected cases covered a sufficiently wide range of delay cases which we considered to be statistically significant. From this group we found that the number of days delayed follows a lognormal distribution. That is, the distribution is highly skewed, with the majority of dams reflecting delays of between 3 and 4 years. This finding is in line with a paper by Callegari et. al.[26] which analysed the delays of 401 cases that were also found to follow a lognormal distribution.

Our next step was to find any correlation between the Risk scores determined from Riverscope and the number of days delay, which we achieved by using a mixture of correlation testing, clustering techniques and linear regression.

The linear regression yielded the most significant results. A regression of the number of days delay on the 17 subnational indicators revealed a positive relationship with population density and a negative relationship with the percentage of cropland irrigated. In other words, the relationship expressed that the lower the population density, the shorter the dam delay, with the greater the percentage of land irrigated also equating to a lower number of days of delay. None of the other 15 indicators showed a significant relation with the number of days delayed.

---

[25] https://www.odi.org/publications/11283-assessing-costs-tenure-risks-agribusinesses

[26] https://www.sciencedirect.com/science/article/abs/pii/S0301421517308042

We used the modelled distribution to determine static delay values related to the location risk score and found the relationship to be exponential. That is, an increase in Risk score results in an exponential increase in the number of years delay[27]. We calculated the 95% confidence intervals to give an indication of the maximum and minimum expected delays. The Risk scores varied from 0.3 to 1 whilst the number of years delayed ranged from 0 to 23.

For a given location's risk score produced by the Riverscope statistical model, the median delay value is determined with a minimum and maximum value to give a range of expected delays. The minimum expected delay is the median between the median of the distribution and the minimum value weighted by the location risk score. Similarly, the maximum expected delay is the median between the median distribution and the maximum value weighted by the location risk score. The expected delay value is determined as the median of a uniform distribution between the minimum and maximum.

For a risk score of 50, the expected delay would be 491 days (1.5 years approx.) with a minimum expected delay of 0.7 years and a maximum expected delay of 2.3 years. For an extremely high risk score of 90, the expected delay would be 4769 days (13 years approx.) with a minimum expected delay of 5.3 years and a maximum expected delay of 36 years.

[27] The next section further illustrates this relationship and how it was incorporated into our final financial model

4. Financial Model

This section builds on the previous by taking the reader through the various inputs, outputs and processes used to develop our financial model, as well as how we determined each of these.

When considering an investment, investors commonly look at the NPV of the project, which is the difference between the present value of cash inflows and the present value of cash outflows over a period of time as calculated with a DCM. NPV is used in capital budgeting and investment planning to analyze the profitability of a projected investment or project[28]. The most important factors considered within the NPV calculation is the discount factor, the unit purchase price and the lifetime of the investment.

Similarly, a hydropower project's LCOE is a metric used by investors when making a comparison of the energy produced by different energy investors. LCOE can be thought of as the average total cost of building and operating the asset, per unit of total electricity generated over an assumed lifetime[29]. Although from a financier's point of view the LCOE is not used as an indicator of profitability, it functions as an indicator of the affordability of the project for the offtaker (this is an important metric to guide the Power Purchase Agreement (PPA) price).

This section describes how we calculated the various financial metrics and the assumptions we used to produce the final outputs from the financial model. Finally, the section describes how we used the financial metric outputs to conduct a comparison between hydropower and alternative energy solutions, such as solar.

4.1. The Discounted Cashflow Model (DCM)

There are several different ways in which financiers calculate some metrics. For example, NPV is commonly calculated with EBITDA (Earnings Before Interest, Tax Depreciation and Amortization)[30] however, using Net Cash Flow is more accurate since there are several factors that can be identified and not assumed, such as interest[31].

Similarly, we also saw that a project's LCOE was calculated as the present value of the cost of the investment over the expected lifetime of the investment. This was then divided by the present value of the amount of units produced by the project over its lifetime in order to produce the final result.

Finally, we arrived at the following factors (in table overleaf) that had to be either identified or assumed as part of the financial model:

[28] https://www.investopedia.com/terms/n/npv.asp

[29] https://corporatefinanceinstitute.com/resources/knowledge/finance/levelized-cost-of-energy-lcoe/

[30] https://www.investopedia.com/terms/n/npv.asp

[31] https://libn.com/2017/08/02/the-difference-between-ebitda-and-cash-flow/

Master Inputs Descriptions	
Total CAPEX	The initial CAPEX can be added
Size of Project (MW)	The expected size of the project in MW
Capacity factor (%)	Energy Production plants do not operate at 100% capacity. You can indicate the expected capacity factor here. Capacity factor for hydropower ranges from 25-90%[32]
Years of Construction	You can add the initial expected duration of construction
Discount Rate	The Discount rate is the expected lending rate of the financiers from the reserve bank. Also an indication of perceived risk due to the location and investment type
Expected CAPEX Overrun due to Delay (%)	Every year that the project is delayed, we assume there is an expected 10% cost overrun to the initial CAPEX
Interest Rate on CAPEX (%)	You can add the Interest rate that is applied to the CAPEX loan. The interest starts from year 1 of operation.
OPEX per Annum (% of CAPEX)	You can add the OPEX as a percentage of the CAPEX (which includes any cost overruns due to delays): Hydropower O&M typically ranges between 1-4% of CAPEX with large hydro between 2-2.5% of CAPEX[33]
Years of Tax Relief in Years	Here you can add the number of years tax relief. For large infrastructure projects the government often provides a tax relief. We assume this means that the government allows the project to write-off the debt over a number of years
TAX Rate	You can add the percentage tax that is applicable on the revenue. Tax is more complex than this, but this simplified view of calculating tax, which we assume includes the reduction of depreciation, should be sufficient for the analysis
Duration of loan in Years	You can add the number of years for the loan to reach maturation, which is used to understand the amount of interest that accumulated for the cashflow calculation
PPA cost per kwh ($/kWh)	You can add the cost per kilowatt that the investment will be selling the produced power at. This value has a significant impact on the model and should be as close to the actual amount as possible to have realistic results
PPA Price Escalation / Inflation (%)	You can add the percentage inflation here. We assume that the PPA and general costs of operating increase at the same rate. This could be assumed will be either linked or similar to CPI.

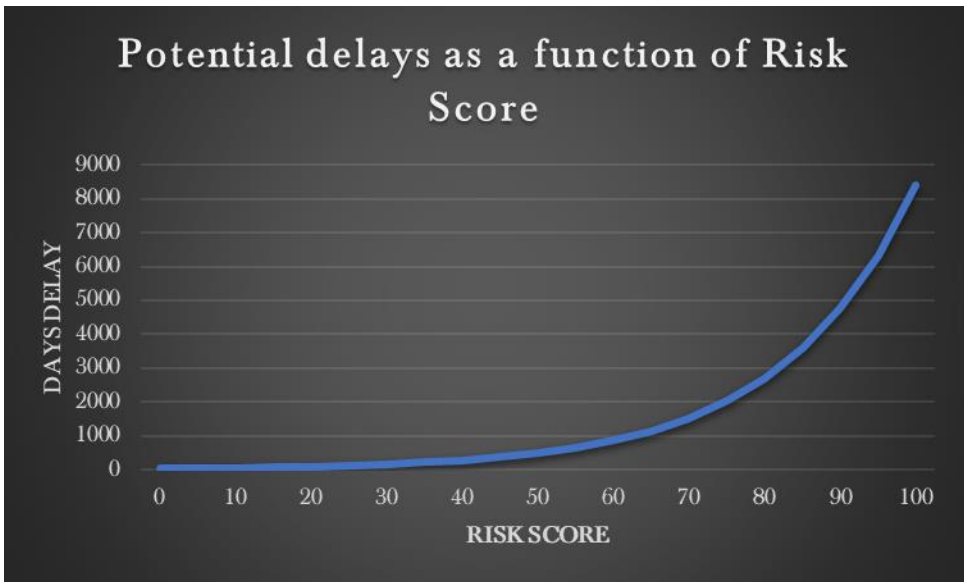
[32] https://www.irena.org/documentdownloads/publications/re_technologies_cost_analysis-hydropower.pdf

[33] https://www.irena.org/documentdownloads/publications/re_technologies_cost_analysis-hydropower.pdf

4.2. Delay Incorporated

The RA score plugs into a DCM to provide an assessment of a project's NPV and the likely LCOE. These metrics are widely used and can be easily compared with alternatives like solar, wind or geothermal.

As mentioned, for Riverscope we looked at 49 dams that have experienced delays which could be attributed to social or environmental factors. This is a statistically significant amount of cases[34], however with an increased number of delay cases we will be able to improve the correlation of the models. Using these delay values, we created a distribution of possible delays (see graphic below).



The RA Score shows the distribution of dams being analyzed. The graph's line correlates to increased risk associated with increasingly lengthy delays i.e. how long the delays would be if a dam planned for this location experienced problems. In other words, we use a statistically robust method to determine ESG risk then show how the receipt of revenues could be delayed by social and environmental problems downgrading the financial viability of a project.

Our approach assumes that a risk score of 0 equates to no delays. The distribution of the number of delays then follows a lognormal distribution. That is, as the risk score becomes larger, the potential delays increase exponentially. The same result was derived in the paper Callegari et. al. published in "Energy Policy" which analyzes the distribution of delays in various energy projects[35].

---

[34] https://www.researchgate.net/post/What_is_the_rationale_behind_the_magic_number_30_in_statistics

[35] https://www.sciencedirect.com/science/article/abs/pii/S0301421517308042

These delays are then combined with assumptions around overruns in spending and different discount rates (or different costs of capital) to produce assessments of a project's LCOE, NPV and other widely used and understood financial metrics.

4.3. Final Financial Model

Upon finalizing the model for the hydropower DCM with delays, we were able to make comparisons against alternative renewable energy investments. This process particularly focused on the impact to the LCOE, relying on an array of delays, capital increases and different assumptions in order to understand how hydropower investment is affected by both initial assumptions about the project and the impact of subsequent delays. Simultaneously these results were then used to see how this compares to other energy alternatives.

The LCOE prices for alternative renewable energy sources were identified by taking current LCOE prices in each specific country of analysis and then extrapolating the expected decrease in LCOE price globally to the LCOE found from existing projects in the country of analysis. We were then able to determine the difference in the LCOE between the hydropower investment in question and the alternative renewable energy investments for potential future commissioning dates. That is, dates when the investment could go online with the expected delays factored into the timeline of the project.

This innovative financial modelling approach has proven to be robust in quantifying an investment's financial losses as a function of its associated ESG risks. The model relies primarily on risk scores determined by Riverscope. These provide potential investment delays that can then be incorporated into a developed DCM. The DCM in turn calculates various useful and common financial metrics, such as NPV and LCOE.

Significantly, these three core elements that make up our financial model can all be enhanced and adapted. We therefore intend to further develop this model and its various components by increasing our sample sizes and introducing additional variables into the DCM. This is ultimately to ensure the model is as representative of an investment as possible.

Appendix I: Indicator details

The indicators were tested individually within each of the three areas to determine if their values were significantly different between the problematic (Test cases) and non-problematic cases (Control cases). A Mann Whitney U test was applied and indicators having a p-value of less than 5% were selected indicating a 95% level of confidence that the values were significantly different between the two sets of cases.

The majority of the indicators have skewed distributions and for the model were transformed before a distribution was fitted to the data. The boxplots below are given for the actual values of each indicator. Hence, due to the skewness of the data the median value is given in the explanations below rather than the average value.

Multidimensional Poverty Index (MPI)

Description: The global Multidimensional Poverty Index (MPI) is an international measure of acute poverty covering 105 countries. It complements traditional income-based poverty measures by capturing the severe deprivations that each person faces at the same time with respect to education, health and living standards.

Provider: Oxford Poverty and Human Development Initiative (OPHI), Oxford Department of International Development

Results

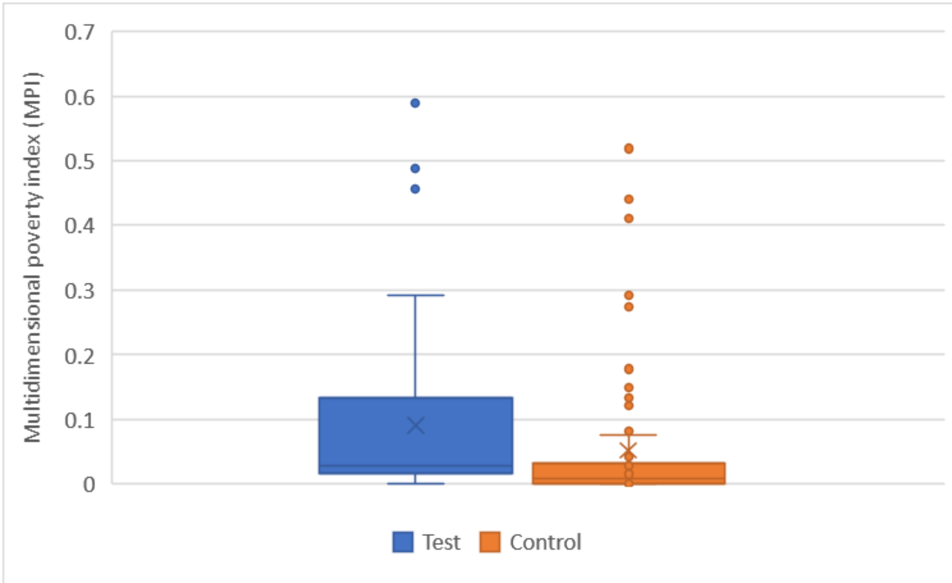
We found that the Test cases tended to occur in areas with a higher percentage of people who are poor and deprived. Except for Child school attendance and nutrition, we found all the indicators showed a significant difference between the Test cases and Control cases in all three areas. This significance was below 1% for all the indicators indicating 99% level of confidence in the significance.

The issue with the MPI data however is that it is not available for all countries. On average it was available for only 64% of the River areas and 69% of the Dam and District areas. However, given its strong significance in distinguishing between the two types of cases, these indicators were included in the model.

Indicators used in the model:

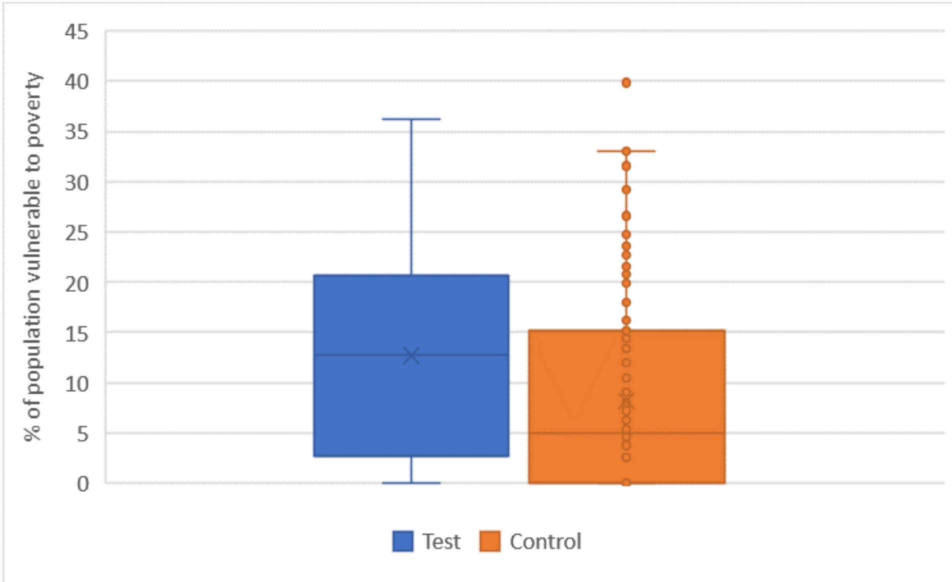
- Multidimensional poverty index
    - Relative poverty measures the ratio of the local level of poverty to the national level, using the overall MPI values for the user's location and the national average. The MPI is a combined measure of the proportion of people in multidimensional poverty, and the intensity of the poverty they experience.

This indicator was applied in the model to the Dam area. The Test cases had a median value of 0.1145 compared to only 0.02 for the Control cases. 42% of the Test cases had an MPI of greater than 0.1 compared to only 16% of the Control cases.



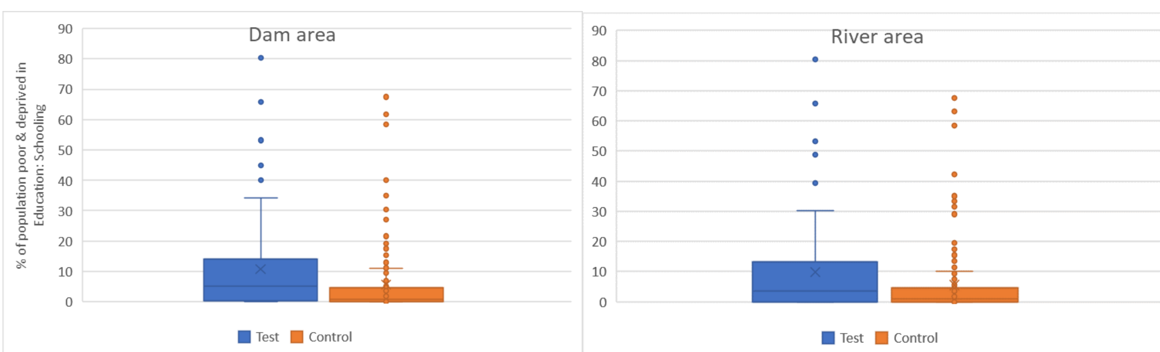
- Population Vulnerable to Poverty
    - The population vulnerable to poverty is defined as the percentage of the population at risk of suffering multiple deprivations—that is, those people with a deprivation score of 20–33% of the weighted indicators.

This indicator was applied in the model at the Dam area. The Test cases tend to be in areas with a percentage of between 2% and 20% compared to the Control cases that occur in areas with percentage of between 0% and 16%. Both sets of cases have a wide spread of values, however the Control dams tend to lie in areas with a lower % of population vulnerable to poverty. There were some Control cases that lie in areas where the percentage of population vulnerable to poverty is higher than 20% however these are only 14% of the cases compared to 32% of the Test cases.



- Percentage of People Who Are Poor and Deprived in Education: Schooling
  - No household member aged 'school entrance age + six years'[36] or older has completed six years of schooling.
  - Any school-aged child is not attending school up to the age at which he/she would complete class eight.[37]

This indicator was applied to both the Dam and River areas. This indicator had a significance value of less than 1% in all three areas. 54% of the Test cases lie in areas where the percentage of people poor and deprived in Education: Schooling is above 5%, compared to only 29% of the Control cases.



[36] This country-specific age cutoff was introduced in 2020. Previously, the age cutoff was 10 years which did not recognize the fact that by age 10 children do not normally complete 6 years of schooling.

[37] Source for official entrance age to primary school: United Nations Educational, Scientific and Cultural Organization, Institute for Statistics database. Education systems [UIS, http://data.uis.unesco.org/?ReportId=163].

- Percentage of People Who Are Poor and Deprived in Living Standards: Improved Sanitation
    - The household's sanitation facility is not improved (according to SDG guidelines) or it is improved but shared with other households.[38]
- Percentage of People Who Are Poor and Deprived in Living Standards: Drinking water
    - The household does not have access to improved drinking water (according to SDG guidelines) or improved drinking water is at least a 30-minute walk from home, round trip.[39]

Aqueduct Global Maps

Description: Aqueduct's global water risk mapping tool helps companies, investors, governments, and other users understand where and how water risks and opportunities are emerging worldwide. The Atlas uses a robust, peer reviewed methodology and the best-available data to create high-resolution, customizable global maps of water risk.

Provider: World Resource Institute

Indicators used in the model:

- Inter-Annual Variability
    - Measures the average between-year variability of available water supply, including both renewable surface and groundwater supplies
- Drought Severity
    - the average length of droughts times the dryness of the droughts from 1901 to 2008.

Both Inter-annual variability and drought severity were significant at the 1% level for all the areas, i.e. the dam area, downstream river and the district. The remaining aqueduct indicators were not significant below 10%.

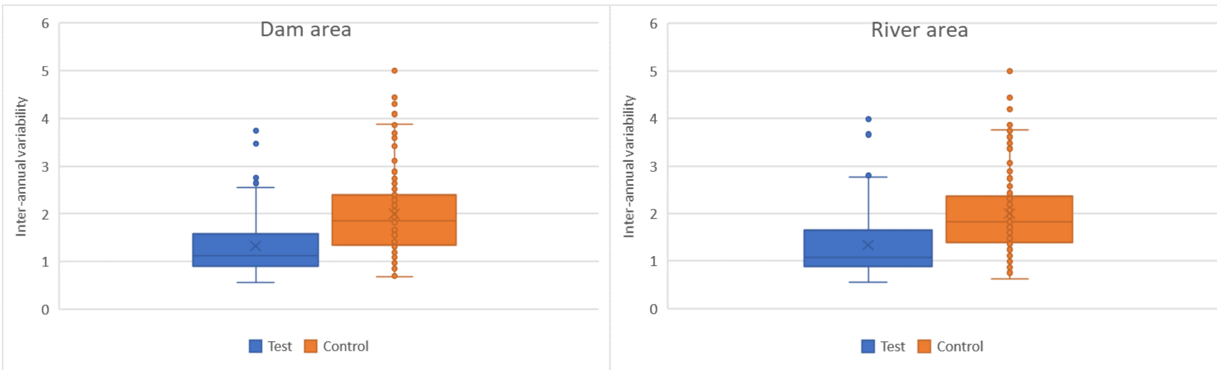
The Interannual variability is an indicator ranging from 0 to 5 with 5 indicating a large degree of variability and 0 little or no variability. Although it was significant for all the areas, it was only

[38] A household is considered to have access to improved sanitation if it has some type of flush toilet or latrine, or ventilated improved pit or composting toilet, provided that they are not shared. If the survey report uses other definitions of improved sanitation, we follow the survey report.

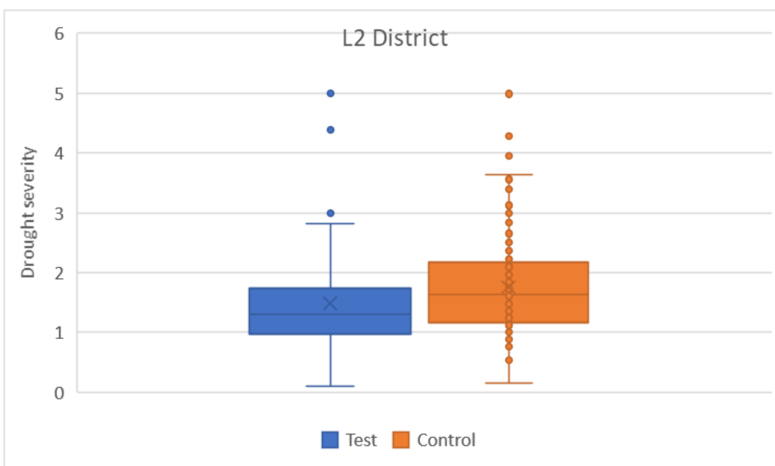
[39] A household has access to improved drinking water if the water source is any of the following types: piped water, public tap, borehole or pump, protected well, protected spring or rainwater, and it is within 30 minutes' walk (round trip). If the survey report uses other definitions of improved drinking water, we follow the survey report.

applied to the Dam and River areas. Likewise the Drought availability indicator was only applied at the District level and not at the Dam and River levels.

The Test cases showed a significantly lower interannual variability than the Control dams. In both the Dam and River areas. The plots below indicate the differences between the Test and Control.



The Test cases showed a significantly lower drought severity than the Control cases. Within the L2 district area the median drought severity for the Test cases was 1.3 compared to 1.6 for the Control cases. A significant number of Control cases had drought severity ratings of more than 2. This can be seen in the boxplot below.



World Database on Protected Areas

Description: The World Database on Protected Areas (WDPA) is the most comprehensive global database on terrestrial and marine protected areas.

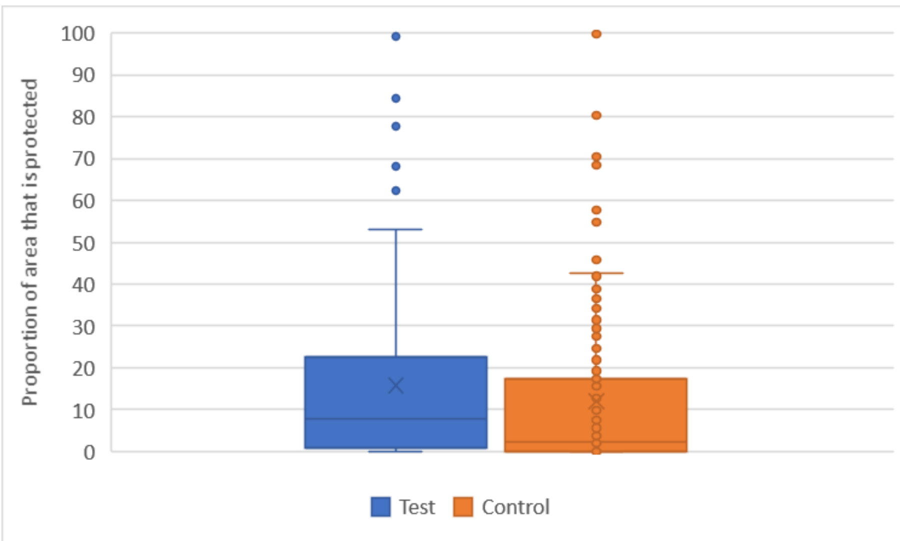
Provider: United Nations Environment Program (UNEP) and the International Union for Conservation of Nature (IUCN)

Indicators used in the model:

- Protected Areas
    - This indicator shows the proportion of the buffer zone covered by some kind of protected area designation (such as national parks or conservation zones), as a percentage of the buffer zone.

The greater the proportion of a district that has Protected land, the more likely a dam built in this area is to experience problems. This indicator was significant within the District area and hence applied to the model within the District area. It was significant at below 1%. Data was available for all areas and the cases covered the full spectrum in that there were cases where dams occurred in areas with no protected area and cases where dams occurred in districts where 99% of the area was protected.

Within the Test cases, 43% of the dams occurred in areas that have more than 10% protected area compared to only 32% of the Control dams.



Modeled Global Suspended Sediment Flux

Description: The WBMsed model is a spatially and temporally explicit (pixel scale and daily) global sediment flux model. It is a component within the Framework for Aquatic Modeling of Earth System (FrAMES), a spatially and temporally explicit multi-scale (local through global) hydrological/biogeochemical modeling scheme.

Provider: Surface Dynamics Modelling Lab

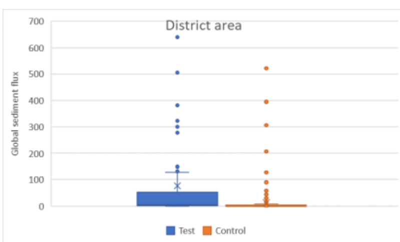
Indicators used in the model:

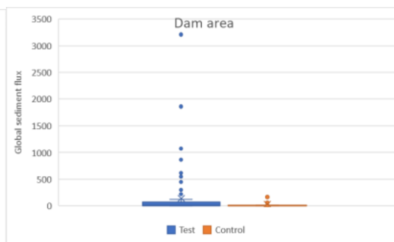
- Global Sediment Flux
    - Quantifying continental sediment flux is a fundamental goal of earth-system science. Ongoing measurements of riverine suspended sediment fluxes to the

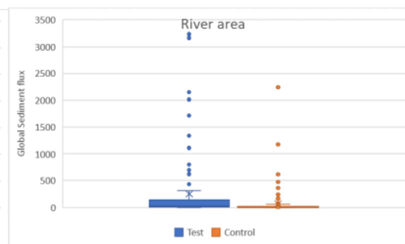
> oceans are limited (<10% of rivers) and intra-basin measurements are even scarcer. Numerical models provide a useful bridge to this measurement gap and offer insight to past and future trends in response to human and environmental changes.

Global sediment flux was significant within all three areas at less than 1%, i.e. greater than 99% confidence. This indicator was applied in the model at all three areas due to its high significance.

Statistic	Test / Control	Dam	River	District
Average	Test Cases	142.24	257.63	77.08
	Control Cases	31.31	76.85	13.67
Median	Test Cases	11.82	28.65	6.50
	Control Cases	1.08	7.22	0.48







IUCN Red List Species Database

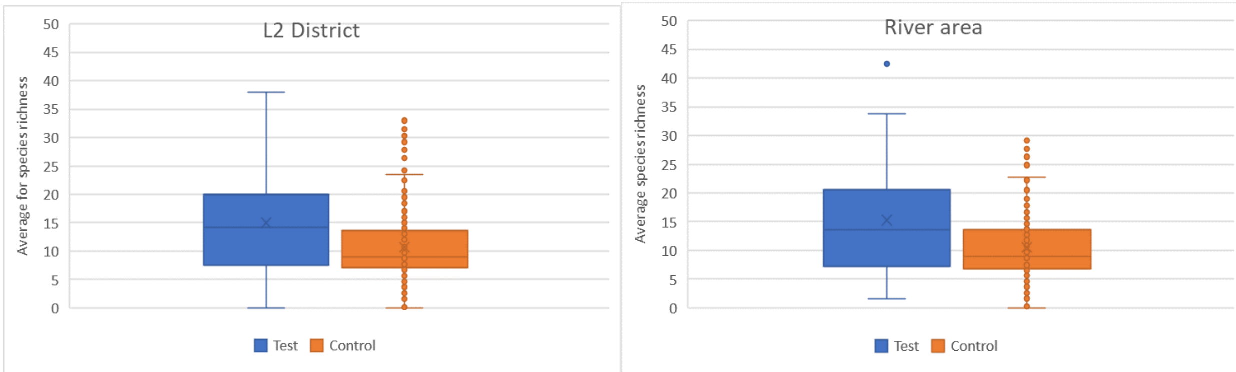
Description: The IUCN Red List is a critical indicator of the health of the world's biodiversity. Far more than a list of species and their status, it is a powerful tool to inform and catalyze action for biodiversity conservation and policy change, critical to protecting the natural resources we need to survive. It provides information about range, population size, habitat and ecology, use and/or trade, threats, and conservation actions that will help inform necessary conservation decisions.

Provider: International Union for Conservation of Nature and Natural Resources

Indicators used in the model:

- Species Richness that are Critical, Endangered and Vulnerable
    - The species richness as determined by analysis by IUCN. The number of critical, endangered and vulnerable species according to the IUCN analysis is used as the indication of poor local behavior and disregard for the environment.

This indicator was significant at 1% for all areas. The Test cases tend to occur in areas with a higher average value for species richness than the Control cases. That is, the Test cases tended to occur in areas with a high proportion of species which are either critical, endangered or vulnerable.



Global Drainage Basin Database

Description: The GDBD is a database made up of six GIS data collections (drainage basin boundary data, river network data, discharge gauging station data, natural lake data, dam lake data, and flow direction data) that store a wide range of information on natural and social sciences.

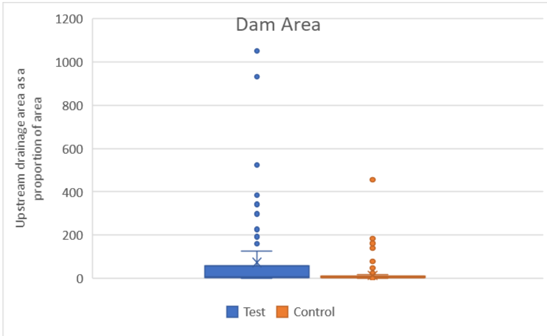
Provider: Center for Global Environmental Research

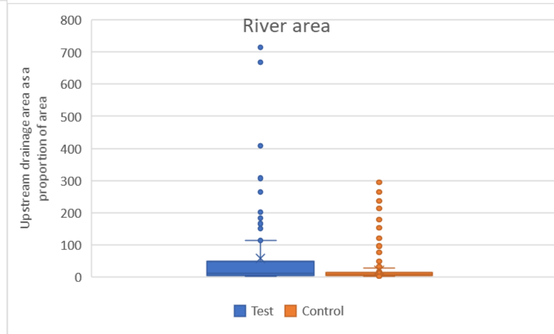
Indicators used in the model:

- Upstream Drainage Area
    - The GDBD is a database made up of six GIS data collections (drainage basin boundary data, river network data, discharge gauging station data, natural lake data, dam lake data, and flow direction data) that store a wide range of information on natural and social sciences.

This indicator was determined as the ratio of the upstream drainage area to the River or Dam area. The greater the upstream drainage area is as a proportion of the area, the more likely a dam is to encounter problems. This indicator was highly significant (less than 1%) both within the Dam and River areas. Therefore, the size of the upstream drainage area as a proportion of area has a significant impact on the Dam and River areas.

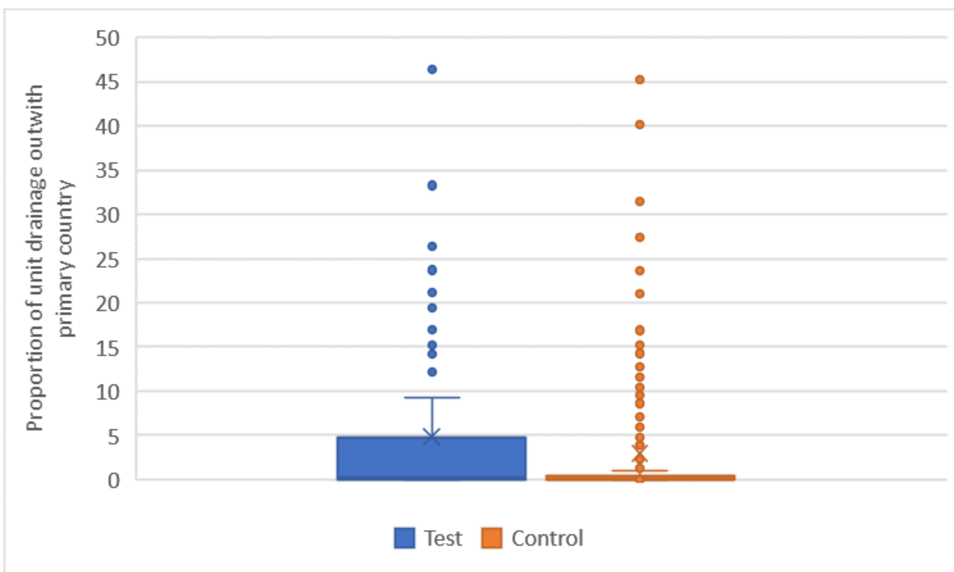
76% of the Control cases occurred in areas where the upstream drainage area was less than 15 times that of the River area, in contrast to only 50% of the Test cases that occurred in these areas. A similar result was found for the Dam area.





A further indicator derived was the proportion of the upstream drainage area that lies out with the primary country. The hypothesis here was to determine if the area covers multiple countries, do projects tend to encounter a greater number of issues compared to those lying entirely in a single country.

This indicator was applied only at the District area and indicated that the greater the proportion of the upstream drainage area that lies out with the primary country, the more likely a project is to encounter problems. Of the Test cases, 45% occurred in areas where more than 1% of the upstream drainage area was out with the primary country. This is in contrast to only 23% of the Control cases.



Blue Water Scarcity Database

Description: WaterStat is the world's most comprehensive water footprint database. All datasets included come from peer-reviewed research and are based on the Global Water Footprint Assessment Standard. Water footprint statistics can be used to inform public policies or company business strategies, to raise awareness, or as input to research projects.

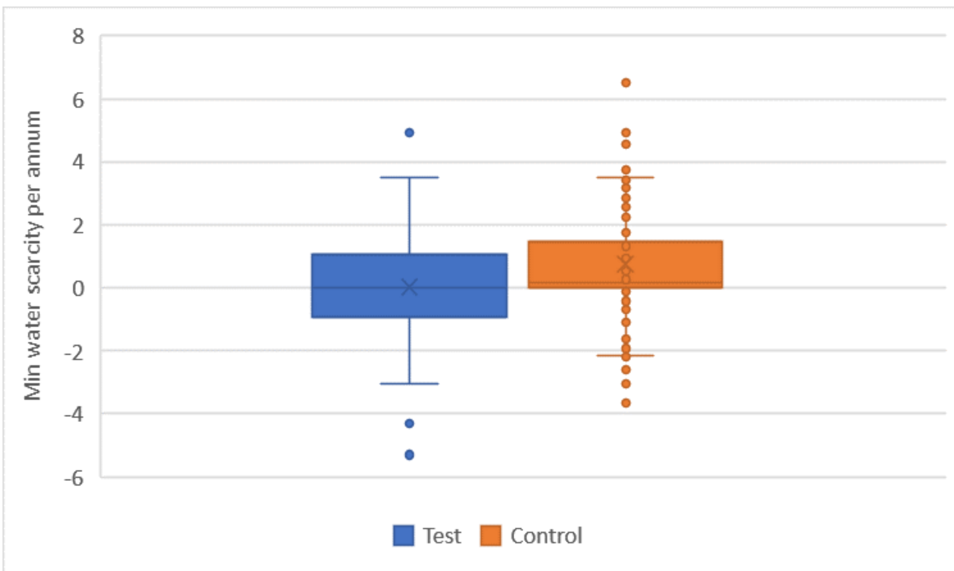
Provider: WaterStat by Water Footprint Network

Indicators used in the model:

- Minimum Percentage Water Scarcity Over the Year
    - Unique datasets showing blue water scarcity in the world on a monthly basis at high spatial resolution.

This database provided us with in-depth data on water scarcity, water availability and natural run-off available monthly. From this we derived numerous indicators including the minimum, maximum, the average, the trend and a seasonality indicator. However, on testing the significance of these indicators at distinguishing between the Test and Control cases, only the minimum water scarcity indicator was significant. This was significant at the 5% level. This indicator was available for 74% of the cases.

The Test cases tended to lie in areas with a low minimum water scarcity compared to the Control cases.



GPWv4

Description: The Gridded Population of the World (GPW) collection, now in its fourth version (GPWv4), models the distribution of human population (counts and densities) on a continuous global raster surface.

Provider: CIESIN, SEDAC and EOSDIS

Results

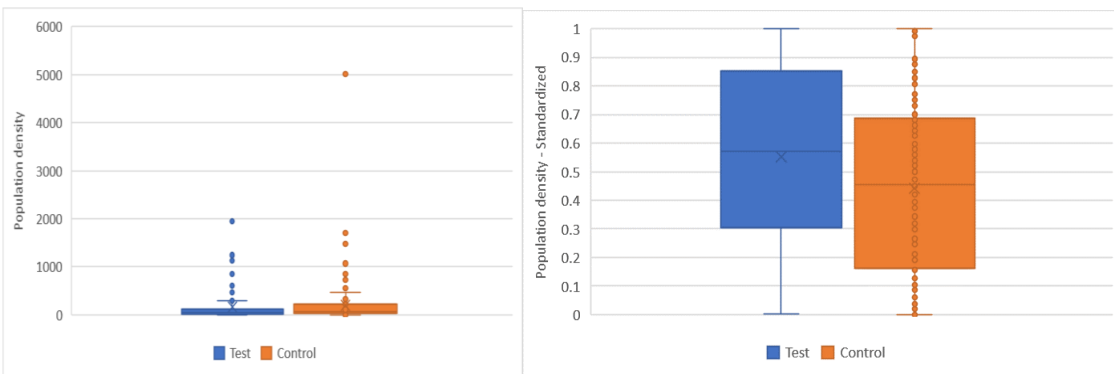
The results of our analysis indicate that the greater the population density of an area, the less likely a project is to experience problems. The Test cases tended to occur in areas with a lower population density compared to the areas where the Control cases occurred.

Indicators used in the model:

- Population density
    - For GPWv4, population input data are collected at the most detailed spatial resolution available from the results of the 2010 round of Population and Housing Censuses, which occurred between 2005 and 2014. The input data are extrapolated to produce population estimates for the years 2000, 2005, 2010, 2015, and 2020.

Data for this indicator was taken from the 2015 estimate. The population density is a highly skewed indicator. That is, there are a few areas with a very high density. The data follows a lognormal distribution and for this reason the data was transformed before standardizing. The plot shows the difference between the actual population density and the inverted standardized values.

This indicator was applied to the River area although it was significant at the 5% for all areas. That is, population density is a significant indicator at all three areas. The Test cases tended to lie in areas with a lower population density than the Control cases.



Earth City Lights Database

Description: This image of Earth's city lights was created with data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS). Originally designed to view clouds by moonlight, the OLS is also used to map the locations of permanent lights on the Earth's surface.

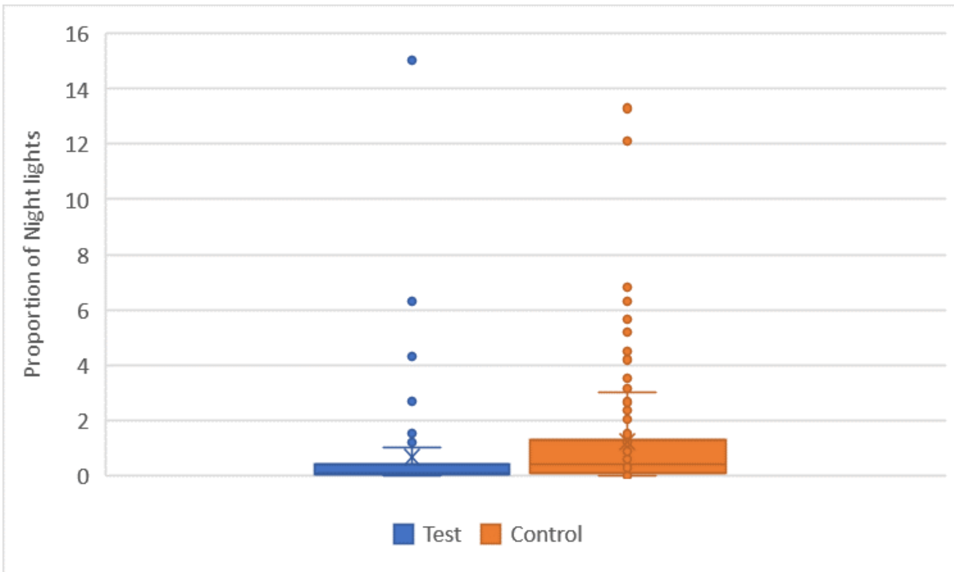
Provider: Defense Meteorological Satellite Program (DMSP)

Indicators used in the model:

- Night Lights
    - The brightest areas of the Earth are the most urbanized, but not necessarily the most populated

This indicator was significant in the River and District areas at 1% and in the Dam area at 5%. In the model the indicator was applied in the River area.

Test dams tended to occur in areas with a low proportion of night lights compared to the Control dams. 50% of the Test dams occurred in areas with less than 8% proportion of night lights, whereas 50% of the Control dams occurred in areas with less than 43% proportion of night lights. Although this indicator is not directly correlated to population density, it does indicate the same trend as the population density indicator.



Global Land Cover database

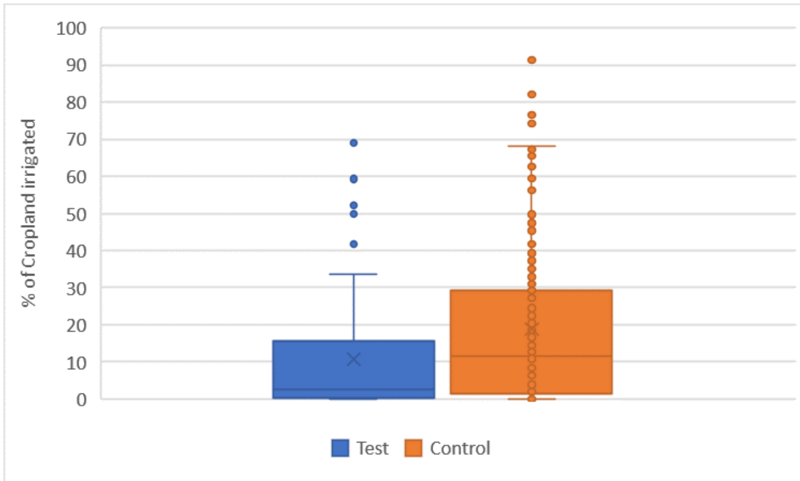
Description: The cropland database is part of the CCI Phase II initiative that focuses on the long-term generation of satellite derived geophysical parameters while considering Landcover data under 2 aspects: stable aspects in the form of land cover maps and dynamic aspects in the form of time series.

Provider: Climate Change Initiative (CCI)

Indicators used in the model:

- Cropland
  - The CCI-LC project delivers consistent global LC maps at 300 m spatial resolution on an annual basis from 1992 to 2015. The Coordinate Reference System used for the global land cover database is a geographic coordinate system (GCS) based on the World Geodetic System 84 (WGS84) reference ellipsoid.

This indicates the % cropland that is irrigated. The Test cases tend not to occur in areas where a higher % of cropland is under irrigation. 72% of the Test cases occurred in areas with less than 15% cropland irrigation compared to 58% of the Control cases.



Conflict events

Description: ACLED records conflict events around the world.

Provider: ACLED

Results

We analyzed events from 1990 onwards that had occurred within the District area. From the data, two indicators were derived.

Indicators used in the model:

- Explosions and remote violence
    - Explosions/Remote violence refers to events where an explosion, bomb or other explosive device was used to engage in conflict. They include one-sided violent events in which the tool for engaging in conflict creates asymmetry by taking away the ability of the target to engage or defend themselves and their location.

The indicator applied here was to the # of events that had occurred since 1990 with regards to explosions and remote violence.

It was found that areas where there were more than 20 of these events tended to have more Test cases than Control cases. This indicator had limited data and only 5 Test cases occurred in these areas compared to only 1 Control case. This is a small sample however the difference is significant and hence the indicator was included.

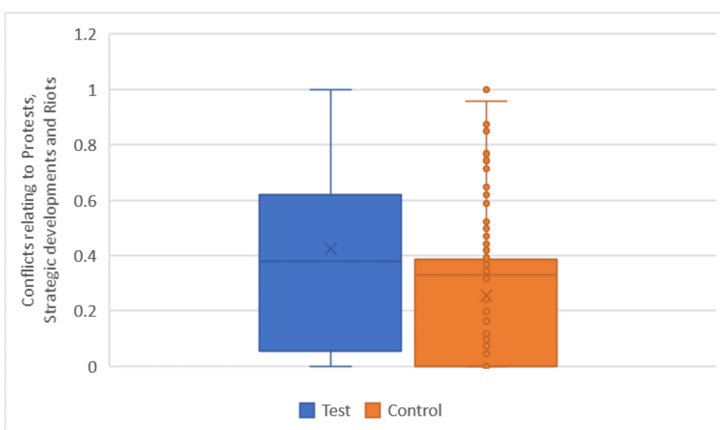
It is important to note here that in terms of the model methodology, if a dam had less than 20 of these events then this indicator was not included in the score calculation. No dam was penalized for not having this indicator.

- Protests, riots, and strategic developments
    - Protests are non-violent demonstrations, involving typically unorganized action by members of society. Riots are a violent demonstration, often involving a spontaneous action by unorganized, unaffiliated members of society. ACLED includes some activity that can broadly be described as 'non-violent' but differs in its role within contexts of disorder. These events, named Strategic developments, include incidences of looting, peace-talks, high profile arrests, non-violent transfers of territory, recruitment into non-state groups etc., and accounts for a small proportion of the total dataset. These common events suggest the context of disorder.

We found the Test cases tended to occur in areas with a greater number of protests and strategic developments but with fewer riots.

A logistic regression was run to determine the influence of the # of protests, # of strategic violence events and the # of riots on a Test or Control outcome. It was found that the # of protests and # of strategic developments had a positive relationship with the # of Test cases yet the # of riots had a negative relationship.

Event type	Weighting
Protests	0.498
Strategic developments	1.19
Riots	-0.616



A Mann Whitney U test was run as was done with all other indicators. This did not reveal any significance in terms of the individual indicators indicating a significant difference between the Test and Control cases. However, given the impact of unrest events on large scale projects it was decided to determine if a combination of events was more significant and hence a logistic regression was run.

5. Appendix II: Test and Control Dams

Identifying the Control and Test cases is essential to the makeup of the model. The definition of a Control hydropower project is highly contested since most (if not all) hydropower projects have some kind of negative impact on either the local community or the surrounding environment.

The approach we took started by identifying the Test dams from the Landscope database which have all been identified as Test cases due to delays. For the Control set we used the GRanD database of dams which contains a dataset made up of over 1,100 dams. We then excluded the dams from the Control set that were already identified as Test cases and only considered dams built after 1980 and within the emerging market.

From this a sample was drawn, stratifying within each country and with priority given to dams built for the sole purposes of hydropower. This resulted in a set of 394 dams. Expert advice from leading Hydropower NGO, International Rivers, allowed us to reduce the sample size to 194 dams by identifying dams that were particularly problematic based on their research and experience. Through the modelling process a further 9 Control dams were dropped due to a high similarity between these dams and the Test dams in terms of their indicator values.

Dam Name	Test / Control
Tamalout Dam	Test
Allain Duhangan Hydropower	Test
New Centennial Water Source Project (NCWS)	Test
Brodarevo 1 & 2 Hydropower	Test
Karuma Hydroelectric Power Station	Test
Bujagali hydropower project	Test
Bisri Dam	Test
PT North Sumatra Hydro Energy	Test
Ralco	Test
Changuinola I (Chan 75)	Test
Rogun Dam	Test
Murta Dam	Test
Pubugou Dam	Test
Ituango	Test

Maheshwar Dam	Test
Yali Falls Hydropower Dam	Test
Gumti Hydroelectric Project	Test
Sirindhorn hydropower dam	Test
Panchet Dam and the Damodar Valley	Test
West Seti Hydroelectric Project	Test
Kajbar Dam	Test
Proposed Mega Dam on River Ewaso NgΓÇÖiro (Crocodile Jaw dam)	Test
Stung Cheay Areng hydroelectric dam	Test
Diamer Basha Dam	Test
Hoa Binh Hydropower Dam	Test
Bakolori Dam	Test
Son La Hydropower Dam	Test
Kandadji dam	Test
Daule Peripa	Test
Tarbela Dam	Test
Lower Se San 2 Dam	Test
Condor Cliff dam	Test
La Miel II	Test
Tumarin	Test
Belo Monte	Test
Myitsone Dam	Test
Nam Theun 2 dam	Test
Xacbal Hydroelectric project	Test
Bakun Dam	Test
Hirakud Dam	Test

Three Gorges Dam	Test
Boguchanskaya Hydropower Plant	Test
Yacyreta	Test
Sardar Sarovar Dam	Test
Gilgel Gibe 3 Hydro Power Dam	Test
Aimor├⌐s	Test
Kariba Dam	Test
Lom Pangar Dam	Test
Akosombo Hydroelectric Project	Test
Bui Hydroelectric Dam	Test
Katse Dam	Test
Manantali Dam	Test
Merowe Dam	Test
Santo Antonio	Test
Manso Dam	Test
Hydroays├⌐n	Test
Neltume	Test
El Quimbo	Test
Hidroabanico	Test
El Chaparral	Test
Santa Cruz Barrillas	Test
La Parota	Test
Barro Blanco	Test
Nujiang Dams	Test
Xiaonanhai Dam	Test
Lower Subansiri hydroelectric power Project	Test
Machhakund (or Machkund) Hydroelectric Project	Test

Luhri Hydro project	Test
Srinagar Hydro Electric Project	Test
Kedung Ombo dam	Test
Xayaburi mainstream dam	Test
Murum Dam	Test
Baram Dam	Test
Hatgyi Project	Test
Jalaur River Multi-Purpose Project Phase II (JRMPP)	Test
Daryan Dam	Test
Janna Dam	Test
Medna Hydropower Plant	Test
Rovni Dam	Test
Lukovo Pole Renewable Energy Project (LPREP)	Test
Po ├ºem Hydropower Dam	Test
Le┼í?e Hydropower Plant	Test
Khudoni Hydropower Plant	Test
Trans-Sibirskaya	Test
Nizhne-Zeyskaya Hydropower Plant	Test
Fujian Solid Waste Disposal Company	Test
Minicentral Tranguil	Test
Thoubal Multipurpose Project/Mapithel Dam	Test
Jispa Dam	Test
Jirau	Test
La Barrancosa dam	Test
Foz do Areia	Control
Arcesti	Control
Mtera	Control

Dniestr	Control
Tieshan	Control
Kulekhani	Control
Jabi	Control
Huai Kum	Control
Constitucion	Control
Arroyito	Control
Stikada	Control
Cerro Pelado	Control
Alicura	Control
Shiroro	Control
Racaciuni	Control
Verkhne-Teriberskaya	Control
Chongyong	Control
Luphphlo	Control
Mratinje	Control
Slano	Control
Komani Dam	Control
Baishan	Control
Saguling	Control
Herculane	Control
Ipotesti	Control
Kiambere	Control
Dubrava	Control
Lubuge	Control
Sibinacocha	Control
Gongboxia	Control

Dchar El Oued	Control
Beni Haroun	Control
Pingban	Control
Chalillo	Control
Yayangshan	Control
Hwanggang	Control
Dakrinh	Control
Chahanwusu	Control
Taishir	Control
La Yesca	Control
Dagangshan	Control
Shamkir	Control
Gura Raului	Control
Presidente Jose L. Portillo	Control
Abdelmoumen	Control
Bang Lang	Control
Petresti	Control
Golesti	Control
Khao Laem	Control
Rogojesti	Control
Dadin Kowa	Control
Hapchon	Control
Allal al Fassi	Control
Zeter	Control
Siriu	Control
Wanjiazhai	Control
Maroon	Control

Talaqan	Control
Molla Sadra	Control
Al Wedha	Control
Satpara	Control
Jatigede	Control
Ouljet Es Soltane	Control
Tokwe Mukorsi	Control
Rumela	Control
La Purisima	Control
El Portillo	Control
Gari	Control
Watari	Control
Magaga	Control
Pada	Control
Amapongokwe	Control
Sabaneta	Control
Carrizo	Control
Garde du Loukkos	Control
Mohamed Ben Abdelkrim el Khattabi	Control
Timi Noutione	Control
Y. Gowon	Control
Toussiana	Control
Tiefora	Control
San Lorenzo	Control
Sidi Yakoub	Control
Andres Figueroa	Control
Frasinet	Control

Facau	Control
Klein Maricopoort	Control
Nooitgedacht	Control
Mae Chang	Control
Djoumine	Control
Merdja Sidi Abed	Control
Haaskloof	Control
Roodekopjies	Control
Grassridge	Control
Koos Raubenheimer	Control
Wemmershoek	Control
Loerie	Control
Pongolapoort	Control
Klipfontein	Control
Goedertrouw	Control
Hawane	Control
Ouizert	Control
Keddara	Control
Bou Roumi	Control
Lekhal	Control
El Gallo	Control
Chipembe	Control
Parcovaci	Control
Gurbanesti	Control
Vaal	Control
Mae Ngat	Control
Talimarjan	Control

Botonega	Control
Chilatan	Control
Hammou Ourzag	Control
Halceni	Control
Lebna	Control
Ain Dalia	Control
Hammam Grouz	Control
Dahmouni	Control
Gallito Ciego	Control
Tungujei	Control
Siliana	Control
Bangazaan	Control
Sidi Abdelli	Control
Souani	Control
Ladrat	Control
Gargar	Control
Colonel Bougara	Control
Tahuin	Control
Zavoiu Orbului	Control
Wonnam	Control
Tabsalao	Control
El Ougla	Control
Bocono-Tucupido	Control
Mbindangombe	Control
Erinle	Control
Manyuchi 2	Control
Jibiya	Control

Maroda Tank	Control
El Molini	Control
Mae Kuang	Control
Shahjad	Control
Challawa Gorge Dam	Control
Lam Chang Han	Control
Huai Sam Nak Mai Teng	Control
Kaliasote	Control
Urmil	Control
El Cuchillo	Control
Derivacao Rio Jordao	Control
Shahid Rajai	Control
Taohe	Control
Xiaoshan	Control
Nandoni	Control
Canoas 2	Control
Salto Caxias	Control
Brezina	Control
Fenhe 2	Control
Bhama Asakhed	Control
Zit El Emba	Control
Nina	Control
Baishi	Control
Biri Stage 1	Control
Injaka	Control
Queimado	Control
Quebra Queixo	Control

Bennithora	Control
Kosar	Control
Oued Mellouk	Control
Taksebt	Control
Tilesdit	Control
Nanshahe	Control
Ay-Doghmush	Control
Rompepicos en Corral des Palmas	Control
Itajai Norte	Control
Sahand	Control
Sina Kolegaon	Control
Koudiat Acerdoune	Control
Qingshanzhui	Control
Picachos	Control
Honghuaerji	Control
Francisco J. Mugica	Control
Mahouane	Control
Tabellout	Control
Ouldjet Mellegue	Control
El Zapotillo	Control

6. Appendix III: Model Indicator Citation

Indicator Ref	Indicator Name (Dataset)	Citation
N4	Inter-annual Variability (WRI)	Gassert, F., P. Reig, T. Luo, and A. Maddocks. 2013. “Aqueduct country and river basin rankings: a weighted aggregation of spatially distinct hydrological indicators.” Working paper. Washington, DC: World Resources Institute, November 2013. Available online at https://files.wri.org/d8/s3fs-public/aqueduct_coutnry_rankings_010914.pdf.
N10	Drought Severity (WRI)	Gassert, F., P. Reig, T. Luo, and A. Maddocks. 2013. “Aqueduct country and river basin rankings: a weighted aggregation of spatially distinct hydrological indicators.” Working paper. Washington, DC: World Resources Institute, November 2013. Available online at https://files.wri.org/d8/s3fs-public/aqueduct_coutnry_rankings_010914.pdf.
O1	Multidimensional Poverty Index (MPI) of the country (MPI)	Alkire, S. and Robles, G. (2017). “Multidimensional Poverty Index Summer 2017: Brief methodological note and results.” OPHI Methodological Note 44, University of Oxford.
O4	Population vulnerable to poverty (MPI)	Alkire, S. and Robles, G. (2017). “Multidimensional Poverty Index Summer 2017: Brief methodological note and results.” OPHI Methodological Note 44, University of Oxford.
O7	Percentage of people who are poor and deprived in Education: Schooling (MPI)	Alkire, S. and Robles, G. (2017). “Multidimensional Poverty Index Summer 2017: Brief methodological note and results.” OPHI Methodological Note 44, University of Oxford.

O12	Percentage of people who are poor and deprived in Living Standards: Improved sanitation (MPI)	Alkire, S. and Robles, G. (2017). "Multidimensional Poverty Index Summer 2017: Brief methodological note and results." OPHI Methodological Note 44, University of Oxford.
O13	Percentage of people who are poor and deprived in Living Standards: Drinking water (MPI)	Alkire, S. and Robles, G. (2017). "Multidimensional Poverty Index Summer 2017: Brief methodological note and results." OPHI Methodological Note 44, University of Oxford.
Q1	Upstream drainage area as a proportion of the L2 area (GDBD)	Yuji Masutomi, Yusuke Inui, Kiyoshi Takahashi and Yuzuru Matsuoka, Development of highly accurate global polygonal drainage basin data, Hydrological Processes, 23, 572-584, https://onlinelibrary.wiley.com/doi/10.1002/hyp.7186, 2009
R8	Min % Water scarcity over the year (Blue Water Scarcity Database)	Mekonnen, M.M. & Hoekstra, A.Y. (2016) Four billion people facing severe water scarcity, Science Advances, 2(2): e1500323
AI2	Population density (GPWv4)	Center for International Earth Science Information Network - CIESIN - Columbia University. 2015. Gridded Population of the World, Version 4 (GPWv4): Population Density. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H46T0JKB. Accessed 6 Nov 2015
Y2	Conflict Number of Explosions & Remote Violence (ACLED)	ACLED. (2020). "Armed Conflict Location & Event Data Project (ACLED) Codebook, 2019. Available at: https://acleddata.com/#/dashboard
Y3	Conflict Events including Protests, Strategic Developments and Riots	ACLED. (2020). "Armed Conflict Location & Event Data Project (ACLED) Codebook, 2019. Available at: https://acleddata.com/#/dashboard

	(ACLED)	
S	Total Protected area as a proportion of area (WDPA)	UNEP-WCMC and IUCN (2020), Protected Planet: The World Database on Protected Areas (WDPA) [On-line], [July/2020], Cambridge, UK: UNEP-WCMC and IUCN. Available at: www.protectedplanet.net.
AM1	Night Lights (Earth City Lights Database)	Earth Observation Group, NOAA/NCEI (2017), Version 1 Nighttime VIIRS Day/Night Band Composites, [On-Line], [March/2017]. Available at: https://visibleearth.nasa.gov/images/55167/earths-city-lights
AO2	IUCN Average of species richness for species that are critical (CR), endangered (EN), and vulnerable (VN) (IUCN Red List Species Database)	IUCN 2020. The IUCN Range Rarity Data. Version 2020. https://www.iucnredlist.org. Downloaded on 13/07/20.
AR1	% of Cropland irrigated (SEA CCI)	ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Available at: https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf
AV	Global sediment flux (Modeled Global Suspended Sediment Flux)	Cohen,S., A. J. Kettner, and J.P.M. Syvitski (2014), Global suspended sediment and water discharge dynamics between 1960 and 2010: Continental trends and intra-basin sensitivity, Global and Planetary Change, 115: 44-58, http://dx.doi.org/10.1016/j.gloplacha.2014.01.011.

7. Appendix IV: Model Indicator Weightings

Level	Indicator ID	Indicator Description	Weight	Positive or negative relationship to disputed cases
Dam	N4	Inter-annual Variability	3,2	Positive
	O1	Multidimensional Poverty Index (MPI) of the country	3,2	Positive
	O4	Population vulnerable to poverty (experiencing intensity between 20–32.9%) - % Population	3,2	Positive
	O7	Percentage of people who are poor and deprived in Education: Schooling	3,2	Positive
	O12	Percentage of people who are poor and deprived in Living Standards: Improved sanitation	3,2	Positive
	O13	Percentage of people who are poor and deprived in Living Standards: Drinking water	3,2	Positive
	Q1	Upstream drainage area as a proportion of the L2 area.	3	Positive
	R8	Min Water scarcity over the year	2,9	Negative

	AO2	IUCN Average of species richness for species that are critical (CR), endangered (EN), and vulnerable (VN)	2,9	Positive
	AV	Global sediment flux	3,1	Positive
River	N4	Inter-annual Variability	2,2	Positive
	O7	Percentage of people who are poor and deprived in Education: Schooling	2,1	Positive
	O12	Percentage of people who are poor and deprived in Living Standards: Improved sanitation	2,2	Positive
	O13	Percentage of people who are poor and deprived in Living Standards: Drinking water	2,2	Positive
	Q1	Upstream drainage area as a proportion of the L2 area.	2	Positive
	R8	Min Water scarcity over the year	1,9	Negative
	AI2	Population density	1,9	Negative
	AM1	Proportion of Night lights	2,2	Negative
	AO2	IUCN Average of species richness for species that are critical (CR), endangered (EN),	1,9	Positive

		and vulnerable (VN)		
	AV	Global sediment flux	2,1	Positive
L2 Region	N10	Drought Severity	1,2	Positive
	Q3	Proportion of unit drainage basin in primary country	0,8	Positive
	Y2	Conflicts relating to explosions and remote violence	1,1	Positive
	Y3	Conflicts relating to Protests, Strategic dvelopments and Riots	1,1	Positive
	O12	Percentage of people who are poor and deprived in Living Standards: Improved sanitation	1,2	Positive
	R8	Min Water scarcity over the year	0,9	Negative
	S	Total Protected area as a proportion of area	1	Positive
	AO2	IUCN Average of species richness for species that are critical (CR), endangered (EN), and vulnerable (VN)	0,9	Positive
	AR1	% of Cropland irrigated	0,7	Negative
	AV	Global sediment flux	1,1	Positive

8. Appendix V: Original Set of Databases

Database	Resolution	Source
World Development Indicators: Deforestation and biodiversity	National	World Bank. Deforestation and biodiversity. Licensed under CC BY 4.0. See http://wdi.worldbank.org/table/3.4# for details.
Total natural resources rents (% of GDP)	National	Estimates based on sources and methods described in "The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium" (World Bank, 2011). Licensed under CC BY 4.0. See https://data.worldbank.org/indicator/NY.GDP.TOTL.RT.ZS?end=2017&start=2017&view=map for details.
Aqueduct Water Stress Country Rankings	National	Luck, M., M. Landis, F. Gassert. 2015. "Aqueduct Water Stress Projections: Decadal projections of water supply and demand using CMIP5 GCMs." Washington, DC: World Resources Institute. Licenses under Creative Commons Attribution International 4.0 License. See https://www.wri.org/resources/data-sets/aqueduct-water-stress-projections-data for details.
Land Mark	National	L. Alden Wily, N. Tagliarino, Harvard Law and International Development Society (LIDS), A. Vidal, C. Salcedo-La Vina, S. Ibrahim, and B. Almeida. 2016. Indicators of the Legal Security of Indigenous and Community Lands. Data file from LandMark: The Global Platform of Indigenous and Community Lands. See www.landmarkmap.org for details.
Internal Displacement Monitoring Centre (IDMC) - Conflict-related	National	IDMC Report on Internal Displacement 2017 Conflict Dataset, 22 May 2017. See http://www.internal-displacement.org/global-figures1 for details.

displacement		
Internal Displacement Monitoring Centre (IDMC)-Disaster-related displacements	National	IDMC Report on Internal Displacement 2017 Conflict Dataset, 22 May 2017. See http://www.internal-displacement.org/global-figures1 for details.
Armed Conflict Dataset (ACD) - Ethnic Power Relations Dataset (EPR-Core)	National	Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." Journal of Conflict Resolution 59(7): 1327–42. See https://icr.ethz.ch/data/epr/geoepr/ for details.
Annual Population Growth (%) - World Development Indicators (WDI)	National	World Bank. Annual Population Growth. Licenced under CC BY 4.0. See https://databank.worldbank.org/reports.aspx?source=2&series=SP.POP.GROW for details.
Human Development Index	National	United Nations Development Program, 2018, "Human Development Indices and Indicators: 2018 Statistical Update," United Nations Development Programme: New York. Data copyrighted under the Creative Commons Attribution 3.0 IGO license. See http://hdr.undp.org/en/content/human-development-index-hdi for details.
Transparency International: Corruption Perceptions Index (CPI)	National	Transparency International, 2018, Corruption Perceptions Index, licensed under CC-BY-ND 4.0. See https://www.transparency.org/research/cpi/overview for details.
Worldwide Governance Indicators (WGI)	National	Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. "The Worldwide Governance Indicators: Methodology and Analytical Issues." Policy Research Working Paper 5430. The World Bank, Development Research Group,

		Macroeconomics and Growth Team. See http://info.worldbank.org/governance/wgi/index.aspx#home for details.
Aqueduct Global Maps 2.1	Local	Aquaduct Global Maps 2.1 Indicators, 2015. Licenced under CC-BY-ND 4.0. See https://www.wri.org/resources/data-sets/aqueduct-global-maps-21-data for details.
Multidimensional Poverty Index (MPI)	Local	Human Development Reports. The 2019 Global Multidimensional Poverty Index (MPI). Oxford Poverty and Human Development Initiative (OPHI), Oxford University. See http://hdr.undp.org/en/2019-MPI for details.
Global Roads Open Access Data Set (gROADS)	Local	Center for International Earth Science Information Network - CIESIN - Columbia University, and Information Technology Outreach Services - ITOS - University of Georgia. 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). See https://doi.org/10.7927/H4VD6WCT for details.
Global Drainage Basin Database (GDBD - Drainage Basin Boundary	Local	Yuji Masutomi, Yusuke Inui, Kiyoshi Takahashi and Yuzuru Matsuoka, 2009. Development of highly accurate global polygonal drainage basin data, Hydrological Processes, 23, 572-584, DOI: 10.1002/hyp.7186. See http://www.cger.nies.go.jp/db/gdbd/gdbd_index_e.html for details.
WaterStat - Water Scarcity Statistics (Average natural runoff; Average blue water availability)	Local	Hoekstra, A.Y., Mekonnen, M.M., Chapagain, A.K., Mathews, R.E. & Richter, B.D. (2012) Global monthly water scarcity: Blue water footprints versus blue water availability, PLoS ONE, 7(2): e32688. See https://waterfootprint.org/en/resources/waterstat/water-scarcity-statistics/ for details.
World Database on Protected Areas (WDPA)	Local	UNEP-WCMC and IUCN, 2020. Protected Planet: The World Database on Protected Areas (WDPA) [Online], September 2020, Cambridge, UK: UNEP-WCMC and IUCN. See www.protectedplanet.net for details.

Anthropogenic Biomes of the World, v2 (2000)	Local	Ellis, E.C., K.K. Goldewijk, S. Siebert, D. Lightman, and N. Ramankutty, 2014. Anthropogenic Biomes of the World, Version 2: 2000. Palisades, NY. NASA Socioeconomic Data and Applications Center (SEDAC). See https://doi.org/10.7927/H4D798B9 for details.
Red list spatial data	Local	IUCN. 2020. *The IUCN Red List of Threatened Species.* Red list version 2020-2. See https://www.iucnredlist.org for details.
Global Roads Inventory Project	Local	Meijer, Huijbregts, Schotten & Schipper (2018): Environmental Research Letters. See https://iopscience.iop.org/article/10.1088/1748-9326/aabd42/meta for details.
Water-Related Intrastate Conflict and Cooperation Database	Local	Bernauer, Böhmelt; Buhaug; Gleditsch; Tribaldos; Berg Weibust & Wischnath (2012): Water-Related Intrastate Conflict and Cooperation (WARICC): A New Event Dataset. International Interactions. https://doi.org/10.1080/03050629.2012.697428.
Armed Conflict Location & Event Data Project (ACLED)	Local	Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. Introducing ACLED-Armed Conflict Location and Event Data. Journal of Peace Research 47(5) 651-660. See https://acleddata.com/?post_type=popup&p=16628 for details.
Social Conflict Analysis Database (SCAD)	Local	Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams. "Social conflict in Africa: A new database." International Interactions 38, no. 4 (2012): 503-511. See https://www.strausscenter.org/ccaps-research-areas/social-conflict/database/ for details.
Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset (GED) - Global instances of political violence	Local	Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, vol.50, no.4, 523-532. And Högbladh Stina, 2019, "UCDP GED Codebook version 19.1", Department of Peace and Conflict Research, Uppsala University.

UCDP/PRIO Armed Conflicts Dataset	Local	Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, vol.50, no.4, 523-532. And, Croicu, Mihai and Ralph Sundberg, 2015, "UCDP GED Codebook version 2.0", Department of Peace and Conflict Research, Uppsala University.
TMP Case Studies social conflict	Local	The TMP Tenure Dispute Database records projects that were subject to a dispute between a company and people local to the project area concerning the use of land or other natural resources. The latest update, v2019.1, has 602 cases and is available for download from the Landscope website: https://landscope.info/about.php.
GlObal geOreferenced Database of Dams (GOODD)	Local	Mulligan, M., van Soesbergen, A. and Saenz, L. 2020. GOODD, a global dataset of more than 38,000 georeferenced dams. Scientific Data 7 (31). See http://globaldamwatch.org/goodd/ for details.
Global Reservoir and Dam Database (GRanD)	Local	Lehner, B., C. Reidy Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J.C. Robertson, R. Rodel, N. Sindorf, and D. Wisser. 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment 9 (9): 494-502. See http://globaldamwatch.org/grand/ for details.
Future Hydropower Reservoirs and Dams (FHReD)	Local	Zarfl, C., A.E. Lumsdon, J. Berlekamp, L. Tydecks, and K. Tockner. 2015. A global boom in hydropower dam construction. Aquatic Sciences 77 (1): 161–170. See http://globaldamwatch.org/grand/ for details.
High Resolution Settlement Layer	Local	Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. High Resolution Settlement Layer (HRSL). Source imagery for HRSL © 2016 DigitalGlobe. See http://www.ciesin.columbia.edu/data/hrsl/ for details.
Global terrestrial Human Footprint	Local	Venter, Oscar et al. (2016), Data from: Global terrestrial Human Footprint maps for 1993 and 2009, v2, Dryad,

maps		Dataset, https://doi.org/10.5061/dryad.052q5.
Population Count - Gridded Population of the World, Version 4 (GPWv4) - 2015 Release	Local	Center for International Earth Science Information Network - CIESIN – Columbia University, 2016, Gridded Population of the World, Version 4 (GPWv4): Population Count. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). See https://beta.sedac.ciesin.columbia.edu/data/set/gpw-v4-populationcount for details.
WorldClim - Global Climate Data	Local	Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978. See https://www.worldclim.org/data/index.html for details.
Columbia Centre for Hazards and Risk Research: Global hazards distribution	Local	Center for Hazards and Risk Research - CHRR - Columbia University, Center for International Earth Science Information Network - CIESIN - Columbia University, and International Bank for Reconstruction and Development - The World Bank. 2005. Global hazards distribution. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H40P0WXQ. And: Dilley, M., R.S. Chen, U. Deichmann, A.L. Lerner-Lam, M. Arnold, J. Agwe, P. Buys, O. Kjekstad, B. Lyon, and G. Yetman. 2005. Natural Disaster Hotspots: A Global Risk Analysis. Washington, D.C.: World Bank. See http://documents.worldbank.org/curated/en/621711468175150317/Natural-disaster-hotspots-A-global-risk-analysis for details.
Global Map of Irrigation Areas (GMIA)	Local	Stefan Siebert, Verena Henrich, Karen Frenken and Jacob Burke (2013). Global Map of Irrigation Areas version 5. Rheinische Friedrich-Wilhelms-University, Bonn, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy. See http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version for details.

Nighttime lights	Local	Earth Observation Group, NOAA/NCEI. Version 1 Nighttime VIIRS Day/Night Band Composites. See https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html for details.
Global Forest Change 2000–2019	Local	Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." Science 342 (15 November): 850–53. See http://earthenginepartners.appspot.com/science-2013-global-forest for details.
IUCN Species Richness Data	Local	IUCN 2020. The IUCN Species Richness Data. Version 2020. See https://www.iucnredlist.org for details. Downloaded on 13/07/20.
IUCN Range Rarity Data	Local	IUCN 2020. The IUCN Range Rarity Data. Version 2020. See https://www.iucnredlist.org for details. Downloaded on 13/07/20.
Terrestrial Biodiversity Indicators	Local	The World Bank, IUCN and Birdlife International. 2019. Distributed under a CC-BY-NC-SA license. See http://wbg-terre-biodiv.s3.amazonaws.com/listing.html for details.
ESA CCI Land Cover Maps	Local	ESA Climate Change Initiative: 2015 Land Cover. Version 2. See http://maps.elie.ucl.ac.be/CCI/viewer/index.php for details.
GlobCover Version 2.3 Land Cover Map	Local	ESA / ESA Globcover 2009 Project. Land Cover Map. Version 2.3. See http://due.esrin.esa.int/page_globcover.php for details.
Forest area (% of land area)	National	Food and Agriculture Organization, 2020. Licenced under CC-BY 4.0. See https://data.worldbank.org/indicator/AG.LND.FRST.ZS for details.
Terrestrial protected areas (% of total land	National	UNEP-WCMC and IUCN, 2020. Protected Planet: Terrestrial protected areas (% of total land area); The World Database on Protected Areas (WDPA)/The Global Database on Protected

area)		Areas Management Effectiveness (GD-PAME), Cambridge, UK: UNEP-WCMC and IUCN. See www.protectedplanet.net for details.
Modeled Global Suspended Sediment Flux	Local	Cohen,S., A. J. Kettner, and J.P.M. Syvitski (2014), Global suspended sediment and water discharge dynamics between 1960 and 2010: Continental trends and intra-basin sensitivity, Global and Planetary Change, 115: 44-58, http://dx.doi.org/10.1016/j.gloplacha.2014.01.011
HdyroSHEDS Drainage Basin	Local	Lehner, B., Verdin, K., Jarvis, A. (2008): New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10): 93-94.
HdyroSHEDS River Network	Local	Lehner, B., Verdin, K., Jarvis, A. (2008): New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10): 93-94.
GHS Urban Centre Database	Local	Florczyk, A., Melchiorri, M., Corban, C., Schiavina, M., Maffenini, L., Pesaresi, M., Politis, P., Sabo, F., Carneiro Freire, S. M., Ehrlich, D., Kemper, T., Tommasi, P., Airaghi D., and Zanchetta, L. 2019: GHS Urban Centre Database 2015, multitemporal and multidimensional attributes, R2019A. European Commission, Joint Research Centre. See https://ghsl.jrc.ec.europa.eu/ucdb2018Overview.php for details.