

LANDSCAPE IMPACT ASSESSMENT OF SDG2 DEVELOPMENT PROJECTS USING REMOTE SENSING AND UNSUPERVISED CONTROL SITE SELECTION

Hannah Kemper*^a, Theo Renouard^a, Sarah Muir^a, Rogerio Bonifacio^a, Giancarlo Pini^a, Paolo Lucchino^a, Lorenzo Bosi^a

^a United Nations World Food Programme (WFP) Headquarters, Rome, Italy

KEY WORDS: Landscape Impact Assessment, SDG2, BACI, NDVI, remote sensing, Earth Observation, monitoring, food security.

ABSTRACT:

As part of its objective to achieve Zero Hunger under SDG2 the United Nations World Food Programme, in partnership with Governments, NGOs and other UN agencies, supports food insecure communities to increase natural resource availability and improve their management. This is done mostly through the building and rehabilitation of soil and water conservation assets (e.g., small dams, weirs, landscape restoration) and structures that increase productivity (e.g., vegetable gardens, irrigation canals). To adequately monitor these activities around the globe simultaneously, remote sensing was found to be an adequate tool. This study introduces the use of high-resolution satellite imagery, and more specifically NDVI derived from the Landsat series to verify and quantify the impact of such development projects. In total 121 projects in 10 countries and six different climate zones were analyzed using a pre- and post-implementation comparison and a Before-After Control Impact (BACI) study considering randomly selected control sites. Both approaches were found to show robust results throughout the different countries, project types and climate zones. 67% of all projects showed significant improvements in vegetation conditions during the wet seasons only three years after the implementation. Using the proposed workflow based on Python scripting and cloud computing of satellite data, fast and robust analyses can be achieved, while assuring constant data quality.

1. INTRODUCTION

The United Nations World Food Programme (WFP)¹ is the world's largest humanitarian organization focusing on SDG 2: Zero Hunger. As part of its portfolio to improve vulnerable communities' food security and resilience to shocks, WFP, in close collaboration with Governments, UN Agencies and NGOs, builds and rehabilitates assets that improve natural resource availability and management in the most fragile contexts globally. In 2021 alone, WFP has reached more than 8.7 million people across 49 countries with its Food Assistance for Assets (FFA) programmes, whereby communities receive food assistance while participating in the planning, design, and management of such structures. Through these efforts, WFP has rehabilitated over 190.000 ha of farmland, built 3,740 water ponds, planted more than 3,200 hectares of forests, and constructed or repaired 3,400 kilometres of roads (PRO-R WFP, 2022).

Adequate monitoring of FFA activities is a crucial step for the organization to determine on the successful creation and maintenance. Further, this information can support future project decisions and enhance the overall quality of the implementations. The Asset Impact Monitoring from Space (AIMS)² is a WFP internal service that provides satellite-imagery based evidence of FFA interventions and their positive impact on the environment. Robust analysis on environmental conditions is required to monitor a wide range of FFA assets constructed under diverse climatic conditions. The quantification of impacts is a crucial part of this analysis since success of projects should be quantifiable. The overall question of this paper is: Which approach is suitable

to monitor globally the environmental impacts of FFA projects on the landscape and quantify the success of the programmes? Which approach can be a fit to diverse climates and difficult contexts in countries experiencing conflict or war?

2. THEORY

2.1 Remote Sensing for monitoring

The use of satellite imagery for applications in environmental studies is well-established and there is an increasing trend in the monitoring of crop health (Zhang et al., 2019), drainage, soil moisture and the development of blended datasets (Khanal et al., 2020). Remote sensing is recognized for the vast possibilities of large-scale and dynamic observations (Li et al., 2020) and is indispensable for the creation of land-cover and land-use change data (Rogan and Chen, 2004; Brown et al., 2022) with increasingly better resolution. Furthermore, the use of remote sensing for drought monitoring and the assessment of possible impacts is very common and the use of satellite-based drought indicators established in research (West et al., 2019; Bachmair et al. 2016). The increasing number of datasets³, the introduction of new platforms like Google Earth Engine (Gorelick et al. 2017) or the Microsoft Planetary Computer⁴, and the usage of machine-learning based algorithms (Maxwell et al., 2018) that can be used for image classification or data processing (Camps-Valls, 2009) present important developments.

* Corresponding author

¹ <https://www.wfp.org/who-we-are>

² <https://aims-unwfp.hub.arcgis.com/>

³ <https://developers.google.com/earth-engine/datasets>

⁴ <https://planetarycomputer.microsoft.com/>

The 2030 Agenda⁵ for Sustainable Development declared 17 goals with 169 targets for a better future. Several indicators have been described to reliably monitor their progress at a global scale. According to Estoque (2020), 30 of these indicators can successfully be monitored using remote sensing. In fact, further exploration of how Earth Observation can support the monitoring of development projects is realized through several stakeholders. To monitor the successful implementation of FFA assets certainly, the widespread availability and variety of satellite imagery should be considered. Considering the right data sources is an important element in making correct decisions and implementing improved projects (Vorovencii, 2011). For example, time-series analyses (Saroglu et al 2011; Vorovencii, 2011) is suggested as the biggest advantage of remote-sensing based approaches since historical data is available over large areas of the planet, that can describe accurately the conditions at specific time points. Since the implementations of FFA assets focuses on enhanced crop productivity, improved vegetation cover or regrowth of forested areas, the analysis should focus on capturing the vegetation productivity differences during different stages of the crop cycle.

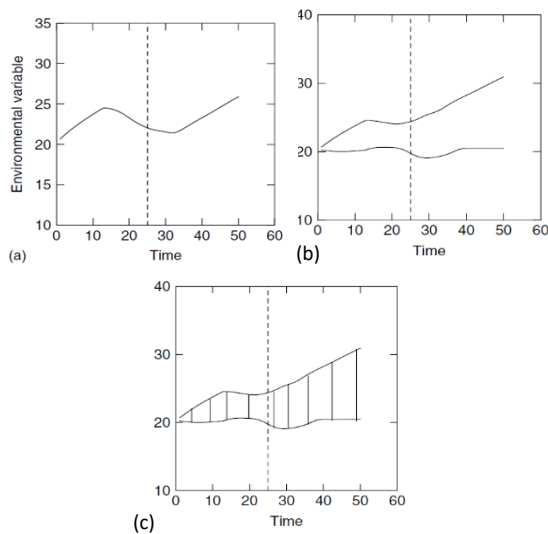


Figure 1. a. simple pre-post comparison, b. Comparison between Impact and Control Site, c. Multiple comparison between pre and post-conditions between Impact and Control Site (Smith, 2013).

For this study the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973) was chosen as a suitable indicator. Using the NDVI in environmental studies is well established and numerous studies show that the indicator is resilient against changing sun angles, topography and shadows and atmospheric conditions (Bunkei, 2007). In past Before-After Control Impact (BACI) studies, indicators based on near-infrared (NIR) and red bands like the NDVI, or the Enhanced Vegetation Index (EVI) have shown reliable results (del Rio-Mena et al., 2021).

2.2 Statistical models for impact assessment

In analysing the impact on the landscape of environmental projects, various analysis models have been dominant, including a simple before-after analysis (Green, 1979) that compares the performance along time within the project area (see Figure 1a). This approach has been criticized thoroughly due to the lack of control measures (e.g., comparing to a site not undergoing any

intervention) or the fact that environmental studies did not account for climatic variabilities despite crop production being highly dependent on the weather conditions in specific years. Results can highlight an improvement in the target metrics, but this improvement cannot be causally attributed to the intervention. Results can therefore be influenced by other trends within the ecosystem e.g., general increase in rainfall causing an overall increase of vegetation.

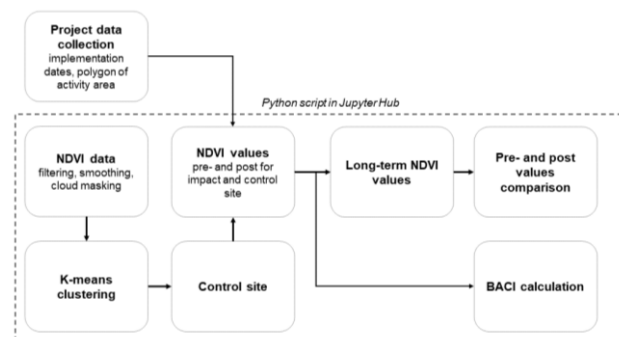


Figure 2. Workflow of applied methodology.

Commonly, the solution to this criticism was the introduction of a control site that would ideally have a similar land cover, be close in space to experience the same weather variability and not subject to anthropogenic changes in the whole research period plus being randomly selected (Meroni et al., 2017).

The Before-After-Control-Impact (BACI) Design includes the selection of one control site for comparison of an indicator before and after the intervention (see Figure 1). This approach was suggested by Eberhardt in 1976 assuming that the start of the impact is known to the analyst. The general idea is to estimate the potential change in the magnitude of variations, in addition to measuring any potential or actual change in the mean of a target variable (Underwood, 1994). By applying this sampling, both annual fluctuations, e.g., crop cycles, as well as inter-annual climate variability are considered (Meroni et al., 2017).

One measure of the BACI analysis is the BACI contrast (del Rio-Mena et al., 2021):

$$BACI\ contrast = (\mu_{CA} - \mu_{CB}) - (\mu_{IA} - \mu_{IB}) \quad (1)$$

with μ being the mean value of an indicator during a specific period, B and A standing for Before and After and C and I for Control and Impact site, FFA asset site in this study

$$relative\ contrast = 100 * (BACI\ contrast / \mu_{IB}) \quad (2)$$

The relative contrast expresses in percent the changes relative to previous conditions, where negative values are associated with an improvement of conditions (Meroni et al., 2017).

The BACI design itself is a robust tool to understand improvements or deteriorations in a fluctuating surrounding. Nevertheless, there is enough room for critique, including the lack of random selected control sites and a rather manual selection of control areas instead. Further, the fact that usually only one control site is selected for each impact site was strongly criticised by Underwood in 1991.

⁵ <https://sdgs.un.org/2030agenda>

A more asymmetrical model was suggested and the work on several control sites proposed to overcome the issue that the statistical outcomes heavily depend on the few control sites considered. On the other hand, even very well drafted studies fail to have more than a dozen impact sites, which results in reduced statistical power (Wood, 2021). Bootstrapping suggested by Mishra et al. 2023 to generate an increase of sample pairs is a possible solution. In this paper, this challenge was addressed by selecting a control site on matching pre-intervention values at a pixel level, as detailed below.



Figure 3. Half-moon assessed by a technical expert in Dosso, Niger, WFP/Richard Mbouet 2023.



Figure 4. Stone bunds to control the water run-off from the hill in Phalombe, Malawi, WFP/Badre Bahaji 2023.

3. METHODOLOGY

Since a homogeneous response to the intervention within the project areas was expected, the basis for all analysis was polygon based, not pixel based (e.g., del Rio- Mena et al., 2017) in this study.

3.1 Project data

A broad dataset of the AIMS service was used as basis of this research. In total, 121 FFA assets in 10 different countries under different climatic conditions were analysed. The projects were implemented between 2011 and 2021 ranging in size between two and 200 ha. The geographical positions include areas with

Mediterranean climates, semi-arid, deserted and even tropical monsoon climates (consult Koppen Geiger classification in Kottek et al., 2006 for reference).

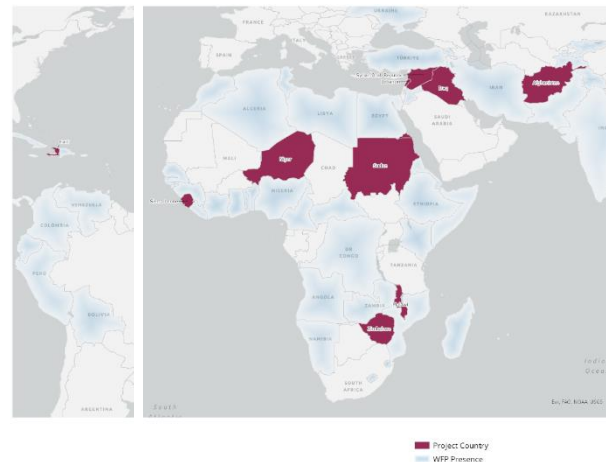


Figure 5. Overview of countries included in this study.

Since the aim is to detect a robust methodology that supports analysis on a global scale, the research area is distributed over several continents. The interventions include irrigation projects, reforestations, soil and water conservation activities to newly created gardens. Under soil and water conservation assets several interventions can be included e.g., half-moons, soil bunds, check dams or terracing. For each project the different interventions undertaken were known, including the exact start and end date.

3.2 NDVI data

The satellite imagery used for this study was taken from the USGS (United States Geological Survey) Landsat Collection 2 Level-2 atmospherically corrected surface reflectance series⁶ and accessed through the Planetary Computer STAC API. Several pre-processing steps were applied to the Landsat time series to ensure data quality using Python. Clouds and cloud shadows were masked using the Landsat Quality Assessment bitmask band as well as surface reflectance outliers. The last pre-processing step is a Whittaker smoother applied on the NDVI computed from the red and NIR bands. The smoother applied is based on a V-Curve optimization and expectile smoothing, as presented by Eilers et al. in 2017. The optimization is carried out pixel by pixel. This is to correct a variety of signal interferences that are mostly due to atmospheric cloudiness and haze, knowing that cloud cover distorts measurements with stronger negative deviations, but it also allows to interpolate data gaps due to cloud coverage.

| Country | Forest | Garden | Irrigation canal | Soil & Water Conservation | Water Pond | Woodlot |
|-------------|--------|--------|------------------|---------------------------|------------|---------|
| Afghanistan | 3 | 2 | 3 | 4 | 2 | 5 |
| Haiti | - | - | 5 | - | - | - |
| Iraq | - | - | 2 | - | - | - |
| Lebanon | 5 | - | 1 | - | 2 | - |
| Malawi | 14 | 3 | 2 | 14 | - | - |

⁶ <https://www.usgs.gov/landsat-missions/landsat-satellite-missions>

| | | | | | | |
|---------------------|---|----|----|---|---|---|
| Niger | - | - | - | 7 | - | - |
| Sierra Leone | - | - | 10 | - | - | - |
| Sudan | 9 | - | - | - | - | - |
| Syria | - | - | 2 | - | - | - |
| Zimbabwe | 1 | 10 | 7 | - | - | - |

Table 1. Countries, where projects were analysed for their environmental impact, WFP 2023

3.3 K-means clustering & control site selection

As previously discussed, the selection of control sites underlies strict criteria. For this study we only consider areas within a 5km buffer of the asset site that are showing similar NDVI patterns (hence vegetation productivity) in the years preceding the implementation of the asset. The selection is done by applying a K-Means clustering (Celik, 2009) over the buffered area. Only important clusters, the most present in the asset area, were kept ensuring that only pixels with similar NDVI characteristics to the asset (i.e., removing constructed and forested areas or bare land if the monitored asset is mainly composed of crops) were

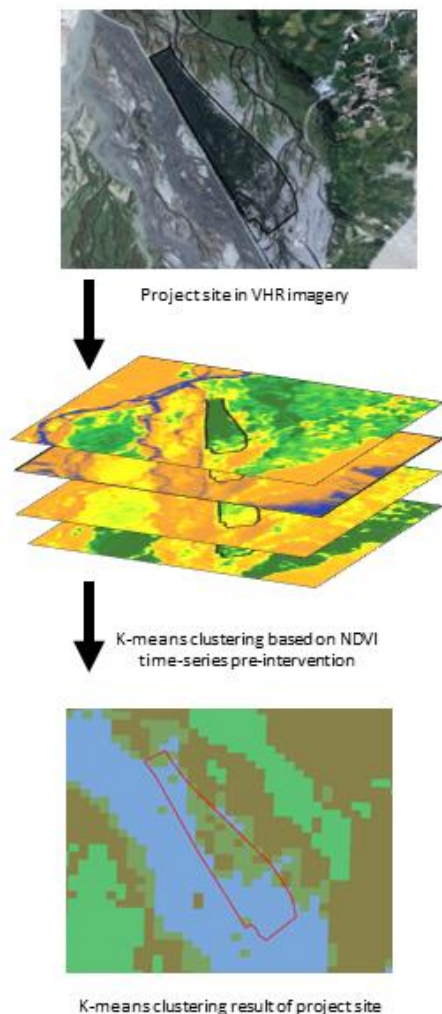


Figure 6. k-means clustering method to identify groups within dataset according to their characteristics, Maxar 2023.

selected for the control site. This control site selection techniques improves the similarity of asset and control areas in terms of confounding variables, thereby reducing selection bias,

and allowing better isolation of the effect of the intervention. Further, bigger areas are included in the sample and not only ones similar in size to the impact site. Very similar weather conditions will apply due to the short distances to the impact site, while similar NDVI patterns reflect a similar landscape type. This can be supported using very-high resolution imagery to compare the land cover at several time slots before the intervention was started. The only constraint is the risk of including areas being under human intervention, yet the size of the control site should make this influence neglectable. Furthermore, it is likely to experience human activity in any form also in areas that were labelled as not being under any project. Overall, this control site selection procedure gives reliable outputs in different climatic and landscape contexts due to its data-based approach, in contrast with more traditional human-based selection methods that are more time-intensive and utilize less data in the process.

3.4 Impact assessments

After retrieving reliable data for both impact and control sites, long-term NDVI values were calculated on the basis of 10 years before the intervention started for the wet and dry season in each specific asset area and control site. Further the values during wet and dry season for impact and control site were calculated for the last three years before and three years after the implementation. Anomalies relative to the long-term average values were calculated for all sites (both impact and control) for six years. A 10% increase of mean NDVI value during dry and wet season in comparison to long-term pre-intervention conditions is used as a standard reference in the corporate results framework in WFP, hence it is useful to understand how this approach compares to the previously presented BACI calculations. The BACI contrast and relative contrast were calculated for both wet and dry season. A two-way ANOVA with interaction model was also applied in

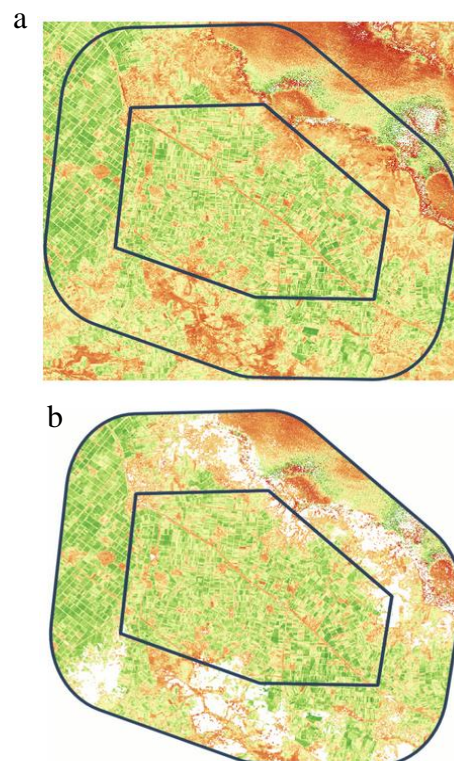


Figure 7. NDVI over an asset and its control site in Syria **a.** Whole area **b.** Analysis ready area after a K-mean clustering selection to keep only pixels with land cover similar to the asset.

the BACI design with intervention (asset / control sites) and period (before / after intervention) as independent factors. The statistical model is as follows:

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, \quad (3)$$

where μ is the overall mean, α_i is the effect of period (i = before or after), β_j is the effect of location (j = control or asset), $(\alpha\beta)_{ij}$ is the interaction between period and location and ε_{ijk} represents the error.

This test allows to obtain the statistical significance of changes occurring after the intervention in the asset looking at the test for an interaction effect. Indeed, an interaction effect occurs when an impact will be observed at the after period depending on the value of the location. Hence, this interaction effect informs us if a change is occurring after the intervention but only at one location, ie. the asset site. Therefore, here and after, “p-value” refers to the statistical p-value of the ANOVA interaction test.

Finally, a Pearson’s r was calculated for the NDVI values between impact and control site for pre- and post-implementation timeframes.

4. RESULTS

After running the above-mentioned script, all outputs were compiled. Out of 121 projects analysed, the BACI relative contrast showed improvements for 81 during the wet season of them, representing 67%. Regarding the 10% criteria being currently used in the standard impact assessments, 78% of all projects were showing a substantial increase of vegetation activity in the first 3 years after the implementation. Improvements during the dry season were found in 41 projects considering the BACI criteria and 35 regarding long-term NDVI values.

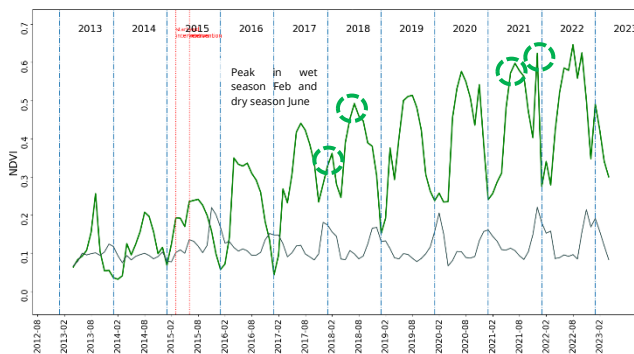


Figure 8. Example of an improved woodlot project in Afghanistan, in an arid climate context. The NDVI of the impact site (green) and NDVI of the control site (grey) in comparison throughout the years. The relative BACI contrast is -156 for the dry season (p-value = 0.015) which shows a clear improvement. The NDVI is up to 400% higher after the implementation than long-term average values.

Regarding the different intervention types, four out of five woodlots showed significant improvement in dry and wet season, meanwhile all other intervention types have a lower number of projects showing vegetation increase during the dry season than the wet season. Soil and water conservation projects show great success during the wet seasons and also great improvements in NDVI of over 40%. Meanwhile Gardens show a mean increase of 25% in wet seasons and even slight reductions during dry seasons. Forestry projects show good results during the dry season, since woody biomass is less dependent on seasonal cycles than crop plants. The only water pond showing an increase in vegetation productivity during the dry season reached improvements of +30% in comparison to long-term average

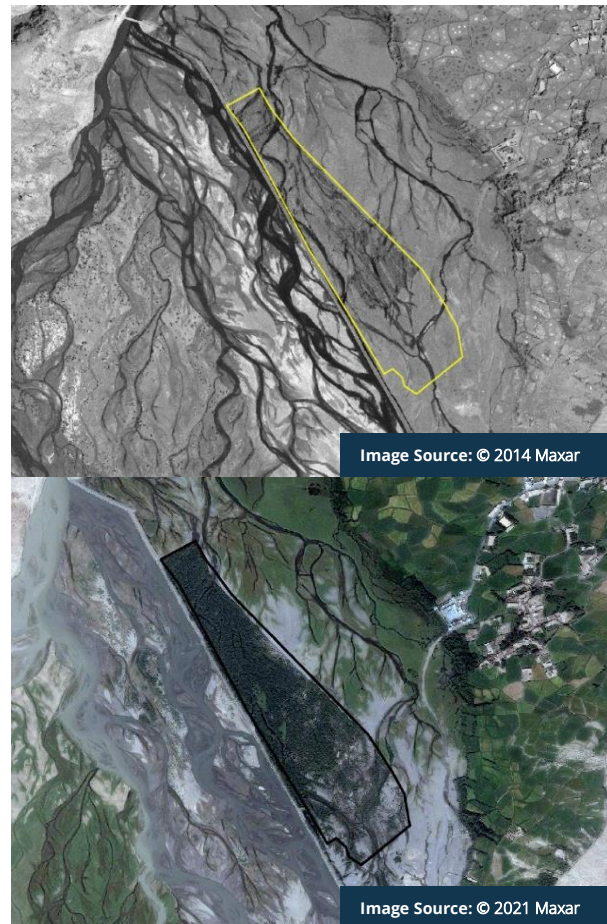


Figure 9. Before (2014) in black-white and after (2021) RGB imagery over the woodlot project site in Afghanistan.

conditions. Since different climatic conditions have influence on vegetation growth and the increase or decrease of NDVI values, the results were grouped into six different types of climate zones. Projects located in arid climates increased in 60% of the cases the vegetation conditions during the dry season and in 64% during the wet season. Increases generally reach up to 70%, related to generally lower NDVI values in the area. Out of 8 projects

| Intervention type | % of projects improved in dry season | Average increase in NDVI (in comparison to Ita) | % of projects improved in wet season | Average increase in NDVI (in comparison to Ita) |
|--------------------------------------|--------------------------------------|---|--------------------------------------|---|
| Forestry (30 total) | 60% | +7.1% | 80% | +35.4% |
| Garden (15 total) | 53% | -1.1% | 60% | +25.6% |
| Irrigation Canal (32 total) | 43% | +7.7% | 53% | +19.6% |
| Soil & Water conservation (35 total) | 48% | +0.6% | 86% | +41.6% |
| Water Ponds (4 total) | 25% | +30% | 75% | +38.3% |
| Woodlot (5 total) | 80% | +235% | 80% | +163% |

Table 2. Results per asset type.

located in a zone with Mediterranean hot summer only three showed significant vegetation increases in the summer, meanwhile in winter 7 out of 8 showed vegetation increases of 27%. In Semi-arid Steppe climates 52% of the analysed projects showed improvements during the dry season, 59% in the wet season. Improvements were generally 10% higher during the rainy season. Tropical Savana climates of generally higher NDVI values also showed significant improvements. Nevertheless, 78% of 41 projects showed improvements within three years after implementation in the wet season.

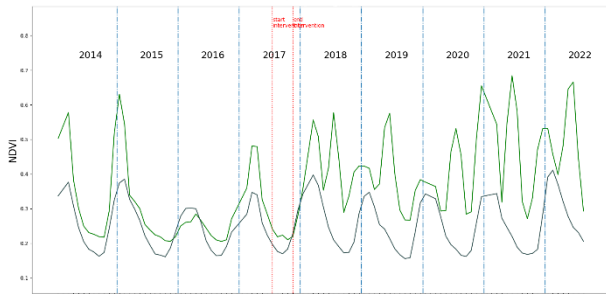


Figure 10. Example of an improved garden creation project in Zimbabwe in semi-arid hot steppe climate. NDVI of the impact site (green) and NDVI of the control site (grey) in comparison throughout the years. The BACI contrast is -36 for the dry season (p-value = 0.005). The NDVI is up to 20% higher post-implementation than long-term average values. Two crop peaks are visible after implementation, meaning an increase in food production, a declared goal of the project.

Table 3. Results per climate zone (based on Koppen Geiger).

| Climate type | Number of improved projects/total number of projects, dry season | Increase of NDVI at improved projects | Number of improved projects/total number of projects, wet season | Increase of NDVI at improved projects |
|---------------------------------------|---|---------------------------------------|---|---------------------------------------|
| Arid Climate hot, BWh | 17/28 ~ 60% | +71% | 18/28 ~ 64% | +72% |
| Mediterranean hot summer climate, Csa | 3/8 ~ 37.5% | +27% | 7/8 ~ 87.5% | +27% |
| Semi-arid Steppe Climate hot, Bsh | 15/29 ~ 52% | +15.5% | 17/29 ~ 59% | +24.8% |
| Tropical Savanna Climate, Aw/As | 24/41 ~ 59% | -8% | 32/41 ~ 78% | +35% |
| Tropical Monsoon, Am | 3/15 ~ 20% | -0.3% | 7/15 ~ 47% | +7.9% |
| Total | 62/121 ~ 51% | +9.9% | 81/121 ~ 67% | +39.3% |



Figure 11: Before and after implementation, grassland was transformed into a garden. VHR imagery supports the positive findings of the NDVI curves.

Projects implemented in Tropical Monsoon climates showed increase in vegetation productivity in 20% of all projects in dry and 47% in wet seasons. NDVI values registered for certain projects decreased slightly due to clearing activities done beforehand, or the fact that the analysis year was particularly dry affecting the growth of young vegetation. Overall, the improvements registered in dry seasons showed NDVI increases of approx. 10%, meanwhile the significant improvements in wet seasons on average are 40% higher than the long-term average conditions.

Comparing the impact methodologies applied, it is visible that the BACI and a 10% NDVI vs. long-term average threshold show similar results in the analysed selection of FFA projects. Meanwhile the BACI methodology found substantial increase in NDVI for more projects in the dry season than the simple pre-/post comparison, the pre-/post comparison found more projects to have a positive impact during the wet season. Results of the two methods differ therefore depending on the season. There are many possible reasons for this behaviour, e.g., the type of intervention analysed or the climatic context of the outliers. Generally, though, none of the groups showed any particularities in comparison to the other projects. Since the investment going into the calculation of both methods has become minor thanks to the data infrastructure provided, a suggestion is to apply both methods simultaneously. In case of major differences in the findings, a more detailed look into the intervention activities and current climatic contexts is needed.

| | >10% increase yes | | >10% increase no | |
|-----------------|-------------------|------------------|------------------|-----------|
| BACI yes | Dry 31 | Wet 73 | Dry 31 | Wet 8 |
| BACI no | Dry 4 | Wet 21 | Dry 55 | Wet 19 |

Table 4. Discrepancies between two proposed methods.

As a last step the correlation between the impact and control side was calculated to understand if any major changes can be detected. Since the selection of the control site is based on the past NDVI profiles, the correlations should generally be lower after the implementation. In particular, irrigation canals should show a lower correlation between the impact and control site, since it in this case indicates lower seasonality e.g., start of second or third crop cycles in contrast to the surrounding rain-fed plots. In the case of the woodlot project in Afghanistan (see Figure 9) a reduction of the correlation of 0.7 was calculated, which means that the project area stopped following the general landscape trends. The correlation can be a great addition to any findings, yet only the change in correlation is not sufficient to make plausible statements on projects, since other human activities can be the reason for trend changes and not necessarily those are positive as given in the woodlot example.

5. DISCUSSION

Overall, the study succeeded in comparing different impact assessment methods and proposing a control site selection procedure in theoretical and practical approaches. Nevertheless, it is crucial to highlight the eventual shortcomings or opportunities for improvement. Considering the selection of control sites, the suggested unsupervised approach showed fast results at low computation cost, fulfilling all the criteria mentioned in literature. For quality control purposes it could be of great improvement to consider a high-resolution land cover layer as additional criteria. By filtering both project and impact site to be in the same land cover category and showing similar NDVI curves, a very strict additional criteria is applied, hence the quality of the control sites would be higher. Nevertheless, this additional processing step needs to be checked for computation efforts and data availability of reliable land cover data for past years.

From a data infrastructure perspective, the suggested approach in this study was computationally efficient, considering that mostly cloud-computing accelerated the output creation. For this reason, an increase of analysis of post-intervention years is possible and advised for further studies.

While interpreting the results of the two applied impact assessment methods, both showed compelling results. Meanwhile a simple threshold seems to be generous during wet seasons, the BACI approach found more positive impact during

REFERENCES

Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Helm Smith, K., Wall, N., Fuchs, B. Crossman, N., Overton, I., 2016: Drought indicators revisited: the need for a wider consideration of environment and society. *WIREs Water*, 3, 516-536. doi: 10.1002/wat2.1154.

Brown, C., Brumby, S., Guzder-Williams, B., Birch, T., Brooks Hyde, S., Mazzariello, J., Czerwinski, W., Pasquarella, V., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., Tait, A., 2022: Dynamic World, Near real-time global 10m land use land cover mapping. *nature*, 9:251, <https://doi.org/10.1038/s41597-022-01307-4>.

dry seasons. For this reason, it is important to always consider full crop cycles while monitoring these projects and considering the post-implementation weather conditions. As visible in Figure 10 the NDVI improvements during dry and wet season reflect the start of a second crop cycle visible in the graph. The sample projects in Sierra Leone were i.e., affected by drought conditions, showing low responsiveness of crop-related vegetation meanwhile the surrounding forested land was more resistant to abrupt conditions changes. Reducing the results on simple number or labels therefore is not advised, since landscape monitoring considers rather complex relationships.

Considering the need for a fixed criteria though, the thresholds for improvement could be adapted to the climate zone where a project is located. In arid and semi-arid contexts, a NDVI increase of +0.2 can reflect a 300% increase, meanwhile in tropical areas it could reflect only a 25% gain.

Finally, both calculations could be enhanced by regarding other environmental variables depending on the project type. This includes land surface temperature, biomass, or soil quality.

6. CONCLUSION

This study presented a landscape impact assessment of 121 WFP's FFA projects in 10 countries considering the vegetation development post-intervention. Both NDVI thresholds in comparison to long-term average conditions and a BACI assessments have been conducted on the dataset showing compelling results and overall good achievement of the selected projects. A robust control site selection methodology was applied that fulfils different criteria and is computationally feasible. Due to the big sample size and the geographical spread of the projects, a general suitability of both methods could be confirmed. Quick calculation and robust findings confirm their applicability for global monitoring of such development projects.

ACKNOWLEDGEMENTS

This research and publication would not have been possible without the constant support and advice of several people. Thanks go out to the current and previous colleagues of the Asset Impact Monitoring from Space (AIMS) team, including Sayana de Gorostizaga Moxon, Sophia Rosa and Paola di Francesco. Further acknowledgements should be made to the Climate & Earth Observation Unit (RAM-C) unit, the PRO-RL unit, supporting Country Offices and the overall RAM division of WFP. Further we would like to express our gratitude to the US State Department Humanitarian Information Unit for their imagery provision, that is crucial to AIMS activities.

Camps-Valls, G., 2009: Machine learning in remote sensing data processing. *IEEE International workshop on Machine Learning for Signal processing*, 1-6. doi: 10.1109/MLSP.2009.5306233.

Celik, T., 2009: Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering. *IEEE Geoscience and Remote Sensing Letters*, 6:4, 772-776, doi: 10.1109/LGRS.2009.2025059.

- Del Rio-Mena, T., Willemen, L., Vrieling, A., Snoeys, A., Nelson, A., 2021: Long-term assessment of ecosystem services at ecological restoration sites using Landsat time series. *PLoS ONE*, 16, <https://doi.org/10.1371/journal.pone.0243020>.
- Eberhardt, L.L., 1976: Quantitative ecology and impact assessment. *Journal of Environmental Management*, 4, 27–70.
- Eilers, P., Pesendorfer, V., Bonifacio, R., 2017: Automatic smoothing of remote sensing data. *International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 1-3, doi: 10.1109/Multi-Temp.2017.8076705.
- Estoque, R., 2020: A Review of the Sustainability Concept and the State of the SDG Monitoring Using Remote Sensing. *remote sensing*, 12, 1770, doi:10.3390/rs12111770.
- Food Systems and Resilience Service (PRO-R) World Food Programme (WFP), 2022: Building Resilience through Livelihoods and Ecosystem Restoration, www.wfp.org/food-assets
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017: Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Green, R., 1979: *Sampling Design and Statistical Methods for Environmental Biologists*. John Wiley & Sons.
- Khanal, S., Kushal, K.C., Fulton, J., Shearer, S., Ozkan, E., 2020: Remote Sensing in Agriculture – Accomplishments, Limitations, and Opportunities. *remote sensing*, 12, 3783, doi:10.3390/rs12223783.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006: World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*. 15:3, 259-263.
- Li, J., Yanqiu, P., Shaohua, Z., Xiao, R., Sang, X., Zhang, C., 2020: A Review of Remote Sensing for Environmental Monitoring in China. *remote sensing*, 12, 1130, doi:10.3390/rs12071130.
- Matsushita, B., Yang, W., Chen, J., Onda, Y., Qiu, G., 2007: Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. *Sensors*, 7.11, 2636-2651.
- Maxwell, A., Warner, T., Fang, F., 2018: Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, 39:9, 2784-2817. DOI: 10.1080/01431161.2018.1433343.
- Meroni, M., Schucknecht, A., Fasbender, D., Rembold, F., Fava, F., Maucalire, M., Lucchio, L., Goffner, D., Leonardi, U., 2017: Remote sensing monitoring of land restoration interventions in semi-arid environments with a before–after control-impact statistical design. *International Journal of Applied Earth Observation and Geoinformation*, 59, 42-52.
- Mishra, V., Limaye, A., Doehnert, F., Policastro, R., Hassan, D., Yaba Ndiaye, M., Van Abel, N., Johnson, K., Grange, J., Coffey, K., Rashid, A., 2023: assessing impact of agroecological interventions in niger through remotely sense changes in vegetation. *nature*, 13:360, <https://doi.org/10.1038/s41598-022-27242-3>.
- Rogan, J., Chen, D., 2004: Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in Planning*, 61, 301-325. doi:10.1016/S0305-9006(03)00066-7.
- Rouse, J., Haas, R., Schell, J., & Deering, D., 1973: Monitoring Vegetation Systems in the Great Plains with ERTS. *NASA Proceedings of the Earth Resources Technology Satellite Symposium*, 1, 309-317.
- Saroglu, E., Bektas, F., Dogru, A.O., Ormeci, C., Musaoglu, N., Kaya, S., 2011: Environmental impact analyses of quarries located on the Asian side of Istanbul using remotely sensed data. *XXII International Cartographic Conference (ICC)*. 15. http://www2.itu.edu.tr/~dogruahm/Dogru%20et%20al_environmental%20Impact%20Analysis.pdf.
- Smith, E. P., 2013: BACI design. *Ecological statistics. Encyclopedia of environmetrics*.
- Underwood, A.J., 1994: On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecological Applications*, 4, 3–15.
- Vorovencii, I., 2011: Satellite Remote Sensing in environmental Impact assessment: an overview. *Bulletin of the Transilvania University Braşov*, 4 (53), 1, 73-80.
- West, H., Quinn, N., Horswell, M., 2019: Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sensing of Environment*, 232, 111291, <https://doi.org/10.1016/j.rse.2019.111291>.
- Wood, C., 2021: Optimizing landscape-scale monitoring programmes to detect the effects of megafires. *Diversity and Distributions*, 28, 479-492, DOI: 10.1111/ddi.13308.
- Zhang, J., Huang, Y., Pu, R., Gonzalez-Moreno, P., Yuan, L., Wu, K., Wenjiang, H., 2019: Monitoring plant diseases and pests through remote sensing technology: A review. *Computers and Electronics in Agriculture*. 165 104943, 1-14. <https://doi.org/10.1016/j.compag.2019.104943>.