

Project evaluation series

**Evaluation of FAO's contribution
to building resilience to
El Niño-induced drought in Southern
Africa**

2016-2017

Annex 1. Methodological approach

**FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2020**

Contents

1	Introduction	1
2	Field data collection	2
3	Quantitative data collection	3
3.1	Identification strategy	3
3.2	Limitations around the causal inference	3
3.3	Propensity score matching and average treatment effect (ATT).....	4
3.4	Sample design and the details of the sample	5
3.5	Qualitative data collection	9
4	Analysis, synthesis and learning phase	10
4.1	Data analysis.....	10
4.2	Validation and utilization of the results of the evaluation.....	11

1 Introduction

1. This annex presents the methodology adopted for the field level data collection and analysis for the evaluation. The four overarching evaluation questions, further elaborated in the evaluation matrix guided the assessment. The evaluation questions (through an evaluation matrix) were in turn used to develop survey questions (both qualitative and quantitative), key informant interview questions as well as to identify or determine data sources that would produce useful and diverse evidence to answer the evaluation questions.

2 Field data collection

2. The field component of the evaluation gathered detailed information at the micro/local level from the direct beneficiaries and non-beneficiaries of the El-Nino response in the three countries. The main source of primary data collection were household survey for the quantitative, and key informant interviews and focus group discussions for the qualitative aspect. This entailed using a number of Rapid Rural Appraisal (PRA) tools and a detailed household questionnaire with sampled beneficiaries and non-beneficiaries of the intervention.

3 Quantitative data collection

3.1 Identification strategy

3. The aim for this adopted methodology was to understand and measure the intermediary impact of the assistance provided to beneficiaries during the 2016/7 agricultural season. In particular, the focus was on assessing the different outcomes between beneficiaries who received the assistance and those who did not. In order to do this, two methodological issues needed to be addressed. The first issue was inherent in measuring the difference between what occurred versus what would have occurred when given a different scenario. Clearly, it is not possible to observe both outcomes for the same individual at the same time. Taking the mean outcome of non-beneficiary households as an approximation is not advisable, since beneficiaries and non-beneficiaries usually differ even in the absence of the assistance. This problem is known as selection bias. The matching approach is one possible solution to the selection problem. It originated in the statistical literature and shows a close link to the experimental context. The basic idea behind this approach, as it relates to this intervention, is to identify within a large group of non-beneficiary households those individuals who are similar to the beneficiaries in all relevant pre-assistance characteristics (vector X). Differences in outcomes between this well selected and thus adequate control group and beneficiaries can therefore be attributed to the assistance. However, this leads to the second issue. Since conditioning on all relevant covariates is limited in case of a high dimensional vector X , Rosenbaum and Rubin (1983) suggest the use of so-called balancing scores. These are functions of the relevant observed covariates or characteristics of non-beneficiaries and beneficiaries such that the conditional distribution of these pre-assistance characteristics is independent of the assistance. This balancing score is the propensity score, i.e. the probability of participating in a programme given observed characteristics. Matching procedures based on this balancing score are known as propensity score matching (PSM) and will be the focus of this analysis. The average treatment effect (ATT) of the El-Niño response was estimated using such matching methods.

3.2 Limitations around the causal inference

4. Having selected the ATT using matching methods as the main evaluation method, there are still a couple of major issues to be addressed. The first is the problematic formation of the counterfactual non-beneficiaries' group due to unobservable variables. The selection of beneficiaries and non-beneficiaries of the assistance was based on certain criteria. Such criterial selection could introduce biases in the estimation of the impacts. For example, if beneficiaries were based on households that have land and are productive, there would be a high possibility of over bias of the impact. Likewise, if the beneficiaries were poor non-productive households, the impact would be underestimated. To overcome this bias, we selected our non-beneficiaries from areas where the assistance was not given. Therefore, we assumed that randomly picking the households would solve the biases.
5. The second limitation is the failure to account for spill overs between beneficiaries and non-beneficiaries. For example, some beneficiaries shared the assistance provided with other non-beneficiaries. By the same token, non-beneficiaries would walk their herds of livestock over long distances to beneficiary villages where boreholes had been provided or where they could purchase livestock feed. To avoid this problem, the questions used to collect information asked simply whether inputs that had been provided through the

assistance were used or not. This allowed the identification of those households that were neither direct beneficiaries, nor indirectly benefited from the use of assistance inputs, but rather relied on the old traditional inputs.

3.3 Propensity score matching and average treatment effect (ATT)

6. Assessing the impact of the response to El-Niño requires making an inference on outcomes that would have been observed for households or communities that received the response had they not benefited. If the yield or livestock number or total harvest for beneficiaries is expressed as Y_b and the yield or harvest or livestock number conditional on not receiving the assistance (non-beneficiary) as Y_{nb} then the benefit of the response is

$$\Delta_y = Y_b - Y_{nb}$$

- Where Δ_y is the difference between the output from beneficiary (Y_b) and non-beneficiary household (Y_{nb}).
- Beneficiary household $= Y_b = (Y|\Delta_y) = 1$ - Given the household received the assistance = 1
- Non-Beneficiary household $= Y_{nb} = (Y|\Delta_y) = 0$ Given household did not receive assistance = 0

$$Y = \Delta_y Y_b + (1 - \Delta_y) Y_{nb}$$

- $E(\Delta_y|\Delta_y = 1)$ What is being measured is the expected output E (of average output) of the change in output given that there is a positive change due to the assistance.
7. For each specific person, only Y_b or Y_{nb} is observed, so Δ_y is not actually observed for anyone. Therefore, while the parameter of interest is the mean benefit of the El-Niño response (**TT**) on the households that actually received the response, we cannot simultaneously observe the outcome of a beneficiary household and the outcome had it not received assistance.

$$TT = E(Y_b - Y_{nb}|X, R = 1) = E(Y_b|X, R = 1) - E(Y_{nb}|X, R = 1)$$

8. Where Y_b and Y_{nb} represents output from a plot of land for beneficiary and non-beneficiary household respectively, and $R =$ El-Niño response package given to the household. Xs are all other explanatory variables (quantities of inputs used, geographical characteristics and household characteristics and endowments). Since the same household cannot be observed having received the assistance and simultaneously not having received the assistance, counterfactuals were established. Certain households have similar characteristics. The main assumption is that when comparing households that received the benefits with households that didn't, but are otherwise similar in all relevant characteristics, eventual differences in outputs in terms of crop production or number of livestock owned after the assistance, can be attributed to the assistance received. Such differences represent the benefits from the assistance. This approach works for observed characteristics, however there may be unobserved characteristics that could affect the output but cannot be controlled, precisely because they are unobserved.

9. To find counterfactuals for the households, propensity scores were constructed from a number of geographical, asset endowment, plot and household characteristics. These selected variables are likely to affect household decisions in agricultural production and productivity. These are also variables that have been widely used in the literature on smallholder farm economics. Plot characteristics used to construct the propensity scores are plot size and quantity of inputs used on the plot. The household characteristics used

are age of household head, education level and wealth indicator of the household. The propensity scores were then used to match plots of similar characteristics given the above variables. In light of the propensity scores, average treatment effects of the treated (ATT) – the average effect of the assistance on the output – were estimated using the Nearest Neighbour Matching method (random draw version).

3.4 Sample design and the details of the sample

10. Household survey data collection was carried out using Open Data Kit (ODK) and ONA. A digital questionnaire was developed and tested in all intervention countries to ensure coherency of terminology and information. Particular attention was paid to ensure collection of survey data also from female household members. Each interview did not go over one hour and focused only on relevant questions to avoid response fatigue.

i. Sample selection

11. Sampling for households to be included in the survey was executed in four stages of multi stage cluster random sampling. The first two stages related to cluster selection.
12. **First**, for each country the main districts and areas where significant dissemination of El-Niño intervention had occurred, in terms of number of projects and numbers of selected beneficiaries, were targeted. In addition to the number of projects, the types of projects (i.e. short-term versus long-term projects) also helped determine the districts that were selected for the survey. The Criteria used for the selection of districts for the in-depth study:
 - i. Integrated Food Security Phase Classification (IPC- Phase 3 and 4)
 - ii. Number of projects in the district
 - iii. Type of projects implemented short term vs. long term interventions
 - iv. Size of the total project budget
 - v. Duration of the project
 - vi. Number of actual beneficiaries
13. **Second**, at least three sub-districts, (wards in Zimbabwe and Lesotho and Extension Planning area (EPA) in Malawi) where significant activities had occurred were randomly selected within each target district. However, two of the sub-districts were areas of intervention, while the third was an area with no intervention, and thus was used as a control. Table 1 shows the selected districts in the three study countries of Zimbabwe, Malawi and Lesotho.

Table 1: Proposed Districts for in-depth field work

Country	Districts	Long-term projects	Short-term projects	Proposed sample size
Lesotho	Mafeteng	*	*	400
	Quthing	*	*	
	Thaba-Tseke	*	*	
	Mohale's hoek		*	
	Berea		*	
	Maseru	*	*	
Malawi	Chikwawa	*	*	700
	Phalombe	*	*	
	Blantyre rural		*	
	Neno		*	
	Kasungu		*	
	Salima		*	
	Mzimba		*	
Zimbabwe	Chipinge	*	*	700
	Chiredzi	*	*	
	Chivi		*	
	Gwanda		*	
	Lupane	*	*	
	Bulilima	*	*	
	Makoni	*		

14. **In the third stage**, in each selected sub-district, five villages were further selected using Probability Proportional to Size (PPS), size being the number of households in each village. **Fourth**, the required number of households in each sampled village were randomly selected based on village population and survey questionnaires were administered in these selected households. These random samplings of households were carried out based on village rosters of control households, while for the beneficiaries, the sampling was based on the beneficiary list from agricultural extension staff who were involved in the project. The formula below was used to determine the total sample for each district to attain the minimum requirement for a representative sample.

$$S = \frac{Z^2 * (p) * (1 - p)}{c^2}$$

Where:

S = Sample size

Z = Z value (e.g. 1.96 for 95% confidence level)

p = percentage household (Assuming 50%)

c = confidence interval (5%)

15. Table 2 presents geographic coverage that is the actual number of households sampled in the three countries by district and also by the type of household head (male-headed or female-headed). An attempt was made to have an equal distribution of male and female headed households surveyed. Accordingly, the distribution of male and female headed households from the randomly selected households is shown below.

Table 2: Household survey coverage (by district and household head)

Country	District	Female-headed households	Male-headed households	Total
Lesotho	Burea	46	50	96
	Butha Buthe	9	14	23
	Mafeteng	34	47	82
	Maseru	42	81	123
	Mohales' Hoek	22	45	67
	Quthing	18	16	34
	Total	171	254	425
Malawi	Blantyre Rural	18	84	102
	Chikwawa	26	100	126
	Kasungu	24	147	171
	Mzimba	28	153	181
	Neno	13	52	65
	Phalombe	21	56	77
	Salima	36	80	116
	Total	166	672	838
Zimbabwe	Bulilima	31	40	71
	Chipinge	104	157	261
	Chivi	30	71	101
	Gwanda	22	53	75
	Lupane	24	43	67
	Makoni	53	75	128
	Mberengwa	40	103	143
	Total	304	542	846

ii. Household survey structure (modules and information assessed)

16. Household interviews were administered to the household head through a detailed questionnaire. Information collected included household demographics and socio-economic status, agricultural landholding, agricultural input use for crop and livestock production, and specific project activities done by the households. A team of experienced enumerators, were first trained in how to conduct this survey. This was followed by a pre-test of the questionnaire before actually implementing it at household level. The evaluation team closely supervised the collection of high quality data.

iii. Balancing control and treatment groups

17. In order to evaluate the impact of the FAO intervention, the treated and control groups were selected to ensure balance between the two. The treated households were randomly selected from the beneficiary list that was provided by FAO local staff or government extension staff that worked with FAO during the response. The control groups were selected from nearby villages where there were no El-Niño response activities. The main assumption was that these were equally affected by El-Niño and suffered losses like the treated groups, but they did not receive any support from FAO. Care was taken to ensure that they had not received assistance from any other organisations. A question was included in the data collection tool to enquire whether or not the household received any assistance apart from the FAO support.

iv. Indicators of interest (based on the TOC)

18. The developed Theory of Change had six main areas of interventions needed in building the resilience of vulnerable small-scale farming households to climate-related events: ***infrastructure rehabilitation and construction; input provision; cash + voucher transfers; training of farmers, training of institutions and generation of information for policy support, decision making and coordination.*** The FAO's 2016/17 El Niño response provided assistance in all of the six areas of interventions (see chapter 2 of the main report for more details). From the developed ToC, the evaluation identified the following key variables which guided the data collection: (i) access to food; (ii) improved and stable crop and livestock production; (iii) factors affecting the adoption of new and improved agricultural technologies; (iv) access to natural resources and alternative income and food sources; and (v) protection of agricultural assets and social networks. **Data was therefore collected on these variables as the principle indicators of interest.**

Construction of the counterfactual

19. After collecting the household and plot data, households' physical asset endowments, community (geographical), household social economic characteristics, and plot level characteristics were used to build the counterfactual. Table 3 presents the variables used for the analysis.

Table 3: Variables used for the analysis

Matching variables	Variable used
Physical assets endowments	Crops or livestock as main source of income
	Land endowment of household
	Wealth index (generated main dwelling house materials)
Community (geographical) characteristics (District Dummies)	Lesotho (Butha-Buthe, Mafeteng, Maseru, Mochale-Hoek, Quithing) (Butha-Buthe = Base)
	Malawi (Chikwawa, Kasungu, Mzimba, Neno, Phalombe, Salima) (Blantyre rural = Base)
	Zimbabwe (Chipinge, Chivi, Gwanda, Lupane, Makoni, Mberengwa) (Bulilima = Base)
	Source of extension services
	Distance to Markets
Social Economic Characteristics	Gender of household head
	Age of household head
	Household size
	Education level of household head
Input use	Inorganic fertilizer
	Quantity of seeds
	Cost of hired labour
	Quantity of organic fertilizer used

3.5 Qualitative data collection

20. Qualitative data was collected at two levels- community and household- through Focus Group Discussions (FGDs) and individual/household interviews supported by the use of selected Participatory Rural Appraisal (PRA) visual tools. The approach supported in-depth discussions on a number of issues including a narrative of the past five years' experience of drought and high rainfall, of the community and households' responses to natural events as well as the positive and negative evolution of their coping strategies. The qualitative data collection used households' livelihood perspectives to evaluate from the households' point of view how effective was the response and what can be done better next time. The use of a livelihoods perspective made it possible to assess some of the cross-sectoral effects and to better characterize the economic activities of vulnerable populations and the distributional impacts of climate change. The qualitative data also enabled gathering information across the sectors of the rural economy that were affected by the El-Niño e.g. agriculture, livestock, natural resources management, labour, small businesses etc. Sector-level assessments was complemented by the livelihoods perspective. The qualitative data collection included key informants interviews, focus group discussions and semi-structured interviews with selected households.
- i. **Key Informants Interviews (KIIs):** A number of key informants were interviewed at the national, district and community level in the sampled districts to get an understanding of the activities that have been carried out and objectives of such activities in response the El-Niño. This included consultations with FAO staff, as well as with representatives from the government ministries and departments, non-governmental organisations (both implementing partners and partners working in the agriculture sector), service-providers (such as agro-dealers) and village chiefs and other authorities (local vets, extension workers etc.).
 - ii. **Community level focus group discussion:** The FGDs was conducted in at least two villages in each selected district¹ with both male and female individuals. A checklist, comprising pertinent study issues, was used to guide the team in their communication with the individuals. The discussions was open-ended and individuals were free to express themselves in a participatory manner. The participatory rural appraisal (PRA) tools used are village resource and socio-economic maps, wealth ranking and agricultural seasonal calendars. The seasonal calendars were done for at least three previous years to get information on trends and changes. This showed the changes in livelihoods over the year in regards to the seasonality of agricultural and non-agricultural workload, food availability, diseases, gender-specific income and expenditure, water, forage, credit and holidays.
 - iii. **Life stories:** some particularly illustrative narratives/life stories from individuals was used in boxes to re-inforce the perspective and voices from the ground. The team summarized verbatim transcripts from interviews at the household level to illustrate specific findings and/or outliers in terms of experiences with the 2015/16 El Niño.

¹ These were the selected villages for the household survey

4 Analysis, synthesis and learning phase

21. The evaluation matrix guided the analysis and synthesis. From the different sources of evidence articulated in the evaluation matrix, the evaluation team answered the four overarching questions outlined in the TOR by deducing and cross-referencing relevant information from different sources around the specific guiding sub-questions or issues. During the analysis phase of the evaluation, the Theory of Change was used as the supporting logic model to determine and describe the results. Based on the Theory of Change, the evaluation team analyzed the evidence on whether/how the FAO activities contributed to building resilience in the following areas:
- i. Stable and diversified access to food
 - ii. Improved and stable crop and livestock production
 - iii. Diversified and stable incomes
 - iv. Protection of agricultural assets and social networks
 - v. Access and utilization of climate and market related information for making decisions
 - vi. Adoption of new and improved agricultural practices and factors affecting or contributing to farmers' decision to adopt
 - vii. Access to natural resources and alternative income and food sources

4.1 Data analysis

22. Data Analysis was done mainly in STATA 14. A number of descriptive statistics were calculated for the variable in table above. These included means, percentages and distribution statistics. These was done for beneficiaries and non-beneficiaries separately. Further to the sampling, the analysis has also shown that these two groups were similar and balancing properties were achieved. The balancing properties of the two groups were based on social economic characteristics and other geographical variables, such as location of the district. The Pscore analysis helped to ensure that only households with similar characteristics were used in evaluating the impacts. All extreme households were not included in the final evaluation of the average treatment effects (ATE) of the FAO intervention. Table 3 shows the number of households that were matched based on a Pscore matching. The households were grouped into blocks of similar characteristics and the balancing properties were achieved by all three countries.

Table 4: Number of households that have been matched by the Pscore matching per country

Pscore	Zimbabwe			Malawi			Lesotho		
	Non-beneficiaries	Beneficiaries	Total	Non-beneficiaries	Beneficiaries	Total	Non-beneficiaries	Beneficiaries	Total
0.1							0	2	2
0.15							9	0	9
0.2	19	7	26	3	8	11	28	9	37
0.4	147	170	317	67	119	186	34	36	70
0.6	104	206	310	130	408	538	29	82	111
0.8	3	14	17	18	77	95	11	65	76
Total	273	397	670	218	612	830	111	194	305

4.2 Validation and utilization of the results of the evaluation

23. The initial analysis of data was aggregated in a series of findings (providing the evidence and answering the questions of the evaluation matrix) and the findings were in turn aggregated to form strategic level conclusions. These draft findings and conclusions were pulled together by the evaluation team in an iterative and collective process of analysis. At the end of this initial process, the team wrote the initial version of the findings and conclusions and highlighted any limitations and gaps in terms of data availability and/or analytical explanations. This initial collection of findings and conclusions was the focus of the second evaluation workshop – during which the evaluation team gathered the same stakeholders as for the inception workshop.
24. The workshop participants were asked to review, validate or challenge the findings and the conclusions. Once this was done, the participants were asked to formulate recommendations for future activities and learning based on the findings and recommendations. They were also asked to make commitments in terms of ensuring a broad ownership, learning and improvement course of strategies, operations and implementation for future crises situations in their respective domains and/or institutions as well as amongst their partners and counterparts.
25. A range of knowledge products to support the learning and dissemination of evaluation results include:
 - i. The evaluation report – following the template and requirements of the FAO office of Evaluation.
 - ii. An evaluation brief or summary – intended for stakeholders who have an interest but only need the main high-lights and learning/action points from the evaluation.
 - iii. A video based on the summary of findings from the range of other process-focused evaluations and reviews. This will be aimed at decision-makers and will be developed jointly with SADC and WFP.

- iv. A video summarizing in English the main findings and lessons learned from the field-based assessments. This video provide the farmers' perspectives and experiences through lessons that can be translated into improved or enhanced strategic and programmatic approaches and portfolios on the part of decision-makers, resource partners and implementing agencies.
- v. Three videos (one per country – Lesotho, Malawi and Zimbabwe) in the local languages, intended for famers and local actors. The videos explain in concrete and simple language the aims of the response activities and the feed-back and experiences of the farmers on the ground.