

## Supplementary Online Materials

### Supplementary Methods: Literature search

We adopted a structured process to conduct the literature search for the review. The general format for searches was premised on the Population Intervention Comparator Outcome (PICO) categories complemented by the intuitive knowledge of authors on subject area. The inclusion/exclusion criteria for papers or datasets to be considered for the analysis were: (1) a standard full NOT design with non-limiting N, P and K rates, randomized and replicated nutrient omission trial (replication can be within or across sites) on maize and rice from any part of SSA; (2) must satisfy the minimum data requirement, namely, identifiable site name, study type e.g., on research station or farmers' fields, year/season of implementation, variety, yields and yield components, and pre-planting soil analysis (plot-specific soil data required but not mandatory); (3) trials that provide a minimum set of four treatments, namely NPK (no macronutrient omitted), NK (P omitted), NP (K omitted) and PK (N omitted) to determine response to N, P and K responses; and (4) dataset may be unpublished or published.

Based on the inclusion/exclusion criteria, we identified papers to serve as a ‘gold standard’ of the types of papers that would be appropriate for this analysis. The citation data of these ‘gold standard’ papers were exported to Zotero as RIS UTF-8 or simple text and stored in files. R program (<https://cran.r-project.org>) was loaded with requisite packages: *litsearchr*, *syntheisr*, *igraph*. Then all fields were imported and merged into one dataframe. Duplicate records in the dataframe were removed. We used the *litsearchr* package in R to create search terms by reading into the files. This package finds words and phrases that appear consistently across several papers in the ‘gold standards’. We then specified what proportion (0 to 1) of the paper must have the same words and phrases.

This was followed by the creation of two lists of possible keywords: one for single words and a second for phrases (2 or more words). We then created a document feature matrix, which associates keywords with their respective citations, for every single word and phrase. The next step entailed creating network graphs: one for words and one for phrases. Network graphs plot node strength against the index number of the candidate word or phrase. The network graph generated a total of 35 candidate words and 20 candidate phrases. We further cut down the number of candidate words and phrases by entering the desired cutoff proportions of node strengths. These left us with 15 words (application, control, efficiency, fertilizer, field, maize, rice, management, nitrogen, nutrient, nutrients, response, systems, trials, yield, yields) and 5 phrases (grain yield, maize grain, maize yield, rice grain, rice yield nutrient omission, yield response). We then assigned each of the words to the PICO categories. This was followed by the creation of a matrix with 2 columns and an appropriate length with the first column holding the words and the second our assignments of PICO categories to each word. The matrix was then transformed to a tibble with specific, required column labels. The same procedure was followed for the phrases. The two tibbles were subsequently merged into one tibble, which was subdivided into separate lists for each PICO category. Finally, a Boolean search phrase was generated by combining terms within a PICO category with ‘OR’ and linking terms between PICO categories with ‘AND’. We used the *stemming* option to merge variants of words into one word with a

wildcard character. For example, nutrient and nutrients were merged into nutrient\*. This process produced the following search expression:

```
"((field* OR maiz* OR system*) AND (applic* OR fertil* OR manag* OR nitrogen* OR nutrient* OR trial*) AND (control*) AND (effici* OR respons* OR yield* OR (maiz* grain*)))"
```

However, this search expression did not retrieve all the “gold standard” papers, necessitating a manual modification of the expression based on intuitive knowledge of subject area. Subsequently, the final printed expression that we copied and pasted into the Google scholar search engine is as follows:

```
“((Africa) AND (field* OR maize OR rice) AND (nitrogen* OR nutrient* OR trial*) AND (“nutrient omission”) AND (yield* OR effici* OR respons*))”
```

Consequently, we retrieved 392 articles in Google scholar. After removal of duplicates a total of 380 articles remained. Based on the content of abstracts, we filtered 43 articles for a full text examination. Out of the 43 articles, we ended up with 22 papers that were used in the quantitative analysis. As our interest was in the primary data used in the analysis, we retrieved original datasets from relevant sources (the authors of the papers). Three more relevant papers were manually-searched and included in the database. In addition to published datasets, we also managed to obtain unpublished data from the same sources. The final list of papers assessed was limited to peer-reviewed, journal articles (Table S1). However, some grey literature sources (e.g., a report from IFDC) were initially considered, but the original datasets could not be readily obtained hence were not included in our database.