Received: 18 May 2023 | Revised: 18 July 2023 | Accepted: 22 March 2024

DOI: 10.1111/agec.12831

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
IAAE
WILEY

ORIGINAL ARTICLE

Soil quality evaluation for irrigated agroecological zones of Punjab, Pakistan: The Luenberger indicator approach

Asjad Tariq Sheikh[1,2] | **Atakelty Hailu**[1,2] | **Amin Mugera**[1,2] | **Ram Pandit**[1,3] | **Stephen Davies**[4]

[1]Centre for Agricultural Economics and Development, UWA School of Agriculture and Environment, University of Western Australia, Perth, WA, Australia

[2]The UWA Institute of Agriculture, The University of Western Australia, Perth, WA, Australia

[3]Global Center for Food, Land and Water Resources, Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan

[4]International Food Policy Research Institute, Development Strategy and Governance Division (DSGD), Washington DC, USA

Correspondence
Asjad Tariq Sheikh, Centre for Agricultural Economics and Development, UWA School of Agriculture and Environment, University of Western Australia, Perth, WA, Australia.
Email: asjad.sheikh@uwa.edu.au

Funding information
UWA International Research Training Program; International Food Policy Research Institute

Abstract

This article describes the construction of the Luenberger soil quality indicator (SQI) using data on crop yield, non-soil inputs, and soil profile from three irrigated agroecological zones of Punjab, Pakistan, namely, rice–wheat, maize–wheat–mix, and cotton–mix zones. Plot level data are used to construct a soil quality indicator by estimating directional distance functions within a data envelopment analysis (DEA) framework. We find that the SQI and crop yield relationships exhibit diminishing returns to improving soil quality levels. Using the constructed SQI values, we estimate linear regression models to generate weights that could be used directly to aggregate individual soil attributes into soil quality indicators without the necessity of fitting a frontier to the crop production data. For wheat and rice production, we find that SQI is most sensitive to changes in soil electrical conductivity (EC) and potassium (K). The SQI has direct relevance for site-specific decision-making problems where policymakers need to price land resources and conservation services to achieve agricultural and environmental goals.

KEYWORDS
agroecological zones, crop yield, data envelopment analysis, directional distance function, soil attributes, soil quality indicator

JEL CLASSIFICATION
C31, C43, C69, Q15, Q24

1 | INTRODUCTION

Land degradation is threatening food security in developing countries where approximately 1.26 billion people practice agriculture in ways that may not always be sustainable (Barbier & Hochard, 2016; Lal, 2004; Stevens, 2018), with adverse consequences for the environment (Nkonya et al., 2011; Stevens, 2018) and the wellbeing of farmers (Gerber et al., 2014). Sustainable agricultural practices are imperative for maintaining or improving soil quality and ensuring food security (Andrews et al., 2002; Lal, 2004; Stevens, 2018). Monitoring sustainable production requires reliable and accurate data to understand the effects of different agronomic practices on soil quality. Unfortunately, adequate and reliable soil quality information is rarely available in suitable form to allow for the evaluation of the performance of agroecosystems and agricultural practices (Arshad & Martin, 2002). This is mainly because soil quality is a multidimensional concept reflecting an array of physical, chemical, and biological soil attributes (de Paul Obade & Lal, 2016) that cannot be directly measured (Stocking, 2003). Thus, it is beneficial to be able to use sound methods that aggregate various soil attributes into a single summary measure, such as a soil quality index (Arshad & Martin, 2002). The construction of such an index has been the subject of research both in soil science and production economics (Andrew et al., 2004; de Paul Obade & Lal, 2016; Pieralli, 2017).

A three-step procedure has been adopted in the soil science literature to construct soil quality indices (Andrews et al., 2004). The first step involves selecting an appropriate minimum data set (MDS) (Doran & Parkin, 1994; Karlen et al., 1997) incorporating physical, chemical, and biological soil attributes. The MDS can be identified based on expert opinion (Andrews et al., 2004, 2002; Doran & Parkin, 1994; Karlen et al., 2003; Vasu et al., 2016) or by reducing the number of soil attributes using multivariate parametric and non-parametric statistical methods (e.g., principal component analysis, redundancy analysis, and discriminant analysis) (Andrews & Carroll, 2001; Askari et al., 2015; de Paul Obade & Lal, 2016; Mandal et al., 2008; Paul et al., 2020; Rezaei et al., 2006; Shukla et al., 2006; Vasu et al., 2016; Yu-Dong et al., 2013; Zhou et al., 2020). The second step involves translating the MDS into standardized scores (ranging between 0 and 1) using linear transformation methods (Askari & Holden, 2014; Masto et al., 2008; Nabiollahi et al., 2018; Raiesi & Kabiri, 2016; Sharma et al., 2005; Yu et al., 2018; Zhou et al., 2020) or non-linear transformation methods (Andrews et al., 2004, 2002; Askari & Holden, 2014; Karlen et al., 2006; Masto et al., 2008; Nabiollahi et al., 2018; Raiesi & Kabiri, 2016; Yu et al., 2018; Zhou et al., 2020). The final step integrates the dimensionless scores into an index using different aggregation techniques, including additive (Andrews et al., 2002; Askari & Holden, 2014, 2015; Mandal et al., 2011; Nabiollahi et al., 2018; Yu et al., 2018), weighted additive (Andrews et al., 2002; Askari & Holden, 2014, 2015; Nabiollahi et al., 2018; Yu et al., 2018), and max-min objective function-based techniques derived from quality measures (Yakowitz et al., 1993).

However, the transformation and integration methods used in the soil science literature involve ad hoc steps (Hailu & Chambers, 2012) due to the lack of well-defined benchmarks for evaluating soil quality (de Paul Obade & Lal, 2014), and consequently lead to uncertainties in soil quality information. No universal soil quality index construction procedure has been developed (Andrews et al., 2004; de Paul Obade & Lal, 2013). In addition, it can be argued that these ad hoc steps do not capture the effect of soil quality on crop yield appropriately (Hailu & Chambers, 2012) because the scoring employed is not directly related to the observed capacity of the soil to produce outputs. In response to these shortcomings, production economists have pursued approaches that integrate soil attributes in a production function framework to compute soil quality indicators (Hailu & Chambers, 2012).

A production function models the transformation of inputs into outputs, where inputs could be both intermediate and final. Numerous studies have assessed the role of soil quality in agricultural productivity by incorporating qualitative (i.e., farmland slope, soil colour, soil type, and soil depth) (Abdulai & Binder, 2006; Bellon & Taylor, 1993; Chang & Wen, 2011; Di Falco & Chavas, 2009; Fuwa et al., 2007; Sherlund et al., 2002) and quantitative (i.e., soil carbon) (Barrett et al., 2010; Marenya & Barrett, 2009) soil attributes in a production function framework. However, only a few studies have attempted to construct soil quality indexes using production frontier methods. Jaenicke and Lengnick (1999) derived a soil quality measure as a Malmquist index with a multiplicatively separable[1] structure for soil and non-soil inputs. Their soil quality index was constructed as the ratio of two radial output distance function values (with and without soil attributes) approximated using experimental data. Likewise, Pieralli (2017) employed a radial output distance function using cross-sectional field-level data from Kenya to aggregate quantitative soil characteristics into a soil quality measure.

Instead of a radial distance function approach, Hailu and Chambers (2012) used an input directional distance function approach to construct a Luenberger soil quality indicator (SQI) using experimental data. A key advantage of the Hailu and Chambers (2012) approach over Jaenicke and Lengnick (1999) and Pieralli (2017) is the use of a uniform metric or yardstick in a pre-assigned direction to measure the efficiency of a bundle of soil quality attributes. A key limitation of radial distance function is that each decision-making unit (DMU) is evaluated along a direction vector dictated by the mix of inputs and outputs for that DMU leading to inefficiency measures that cannot be compared acrosss DMUs in abosolute terms. This causes variation in evaluation results and inconsistency in rankings (Sun et al., 2013). Another limitation is that a set of weights that is favorable to one DMU may not necessarily be favourable to other DMUs. This may cause one DMU to dominate others, a situation that may be unacceptable (Amin & Toloo, 2007). The advantage of the directional

[1] The radial output distance function (ODF) separates soil attributes and non-soil inputs and outputs; the underlying ODF can be decomposed into two components: soil quality aggregator and ODF without soil quality.

AGRICULTURAL ECONOMICS The Journal of the International Association of Agricultural Economists IAAE WILEY

distance function approach is that the decision maker has the flexibility of specifying in what direction the DMUs will be evaluated. Within this framework, the Luenberger directional distance function evaluates the performance of a DMU by measuring its distance to the boundary of the benchmark production technology along a predetermined direction of measurement (Afsharian & Ahn, 2014).

The disposability assumptions incorporated into the construction of the production frontier are also important. Hailu and Chambers (2012) did not impose a "weakly disposal" (Färe et al., 1993) formulation in the construction of the DEA frontier. Under a weakly disposable formulation, the observation being evaluated is compared against a theoretical frontier that has the same soil quality attributes. Such a formulation can reduce the reference set and inflate estimated efficiency scores (Hailu, 2003). In extreme cases, the reference set may consist of only the observation being evaluated. In the context of soil quality measurement, the value of a better soil attribute vector or bundle does not necessarily translate into a higher value for the constructed soil quality indicator. By contrast, Hailu and Chamber's approach leads to soil quality indicator that are monotonically increasing with soil quality attributes, a feature that is not guaranteed with the radial distance function approach (Hailu & Chambers, 2012).[2]

This study follows the approach by Hailu and Chambers (2012) to construct Luenberger soil quality indicators.[3] It makes two contributions to the literature on the economics of soil quality. First, the study uses plot level non-experimental data on soil attributes across three irrigated agroecological zones that vary in terms of regional characteristics, climate, and cropping patterns. This captures soil quality heterogeneity across farms, using a more realistic empirical modelling with data from a real production environment that farmers face rather than data generated in experiment stations. Most earlier studies used farm-level data without considering the spatial diversity that exists, yet land managers tactically divide farmland into multiple plots to take advantage of differences in soil quality. Second, the study constructs soil quality indicators using the production frontier approach that combines output and input data, including a more comprehensive list of soil attributes compared to prior studies. This provides better insights into the main drivers of soil quality necessary to sustain agricultural production and allows identification of soil attributes that contribute the most to overall soil quality across three agroecological zones in the Punjab, Pakistan.

The rest of the article is organized as follows. Section 2 outlines the methodological framework used in the analysis and describes the study context and survey data. Section 3 presents the empirical results. Section 4 summarizes and concludes the article.

2 | MATERIAL AND METHODS

2.1 | Theoretical framework for soil quality indicators

For a general agricultural production process using a vector of L non-soil production inputs $x = (x_1, x_2, \dots, x_L) \in R_+^L$ and a vector of K soil attributes $s = (s_1, s_2, \dots, s_K) \in R_+^K$ to produce a vector of M outputs $y = (y_1, y_2, \dots, y_M) \in R_+^M$, the production possibility set defines all feasible input and output combinations:

$$T = \{ \left(s \in R_+^K,\ x \in R_+^L,\ y \in R_+^M \right) \mid (s, x) \text{ can produce } y \} \tag{1}$$

In our case, inputs include soil attributes, quantity of nitrogen applied, labor hours, machinery, seed, and other costs, while the output is the quantity of wheat or rice. For soil quality measurement exercises, we represent the production technology T using a directional distance function (henceforth $\vec{D}$) $\vec{D} : R_+^K \times R_+^L \times R^M \rightarrow R$ defined as (Chambers et al., 1996):

$$\vec{D}\left(s, x, y; g_s, g_x, g_y\right) = \sup_{\beta} \{\beta : (s - \beta g_s, x - \beta g_x, y + \beta g_y) \in T,\ \beta \in R_+\} \tag{2}$$

where, $g_s \in R^K$, $g_x \in R^L$, $g_y \in R^M$ represent direction vectors (or translation metrics) along which (s, x) are contracted, and outputs (y) are expanded towards the technology frontier. β is the variable that represents the distance function value. The distance function $(\vec{D})$ inherits its properties from the underlying production technology T (Chambers et al., 1996), including:

(i) For all technologically feasible input–output combinations (s, x, y), $\vec{D}$ will always be non-negative. Therefore, the production technology can be characterized as:

$$\vec{D}\left(s, x, y; g_s, g_x, g_y\right) \geq 0 \Leftrightarrow (s, x, y) \in T \tag{3}$$

(ii) $\vec{D}$ is concave in the input–output vector (s, x, y).

(iii) Monotonicity implies that $\vec{D}$ is non-decreasing in inputs (s, x) and non-increasing in output (y).

[2] See also Ray and Mukherjee (1996) comments on radial distance function based quality indexes.

[3] Both indicator and index refer to a summary measure. Indicator is the term used when the summary measure is defined in terms of differences in directional distances while the index is the term used when it is defined in terms of the ratios of radial distances. In other words, an indicator is an absolute measure that can be aggregated while an index is a relative measure that cannot.

(iv) The translation property indicates that adding a scalar multiple of the direction vectors (g_s, g_x, g_y) to the input–output vectors either reduces or increases the distance function value by the size of that scalar:

$$\vec{D}\left(s - \alpha g_s, x - \alpha g_x, y + \alpha g_y; g_s, g_x, g_y\right) = \vec{D}\left(s, x, y; g_s, g_x, g_y\right) - \alpha \quad (4)$$

Since this study aims to construct a soil quality indicator based on soil attributes, the directional vectors for inputs (g_x) and outputs (g_y) are suppressed or set to zero $\vec{D}(s, x, y; g_s, 0, 0)$, allowing us to translate only soil attributes (s) in the direction of (g_s) towards the best-practice frontier to define an input directional distance function (Hailu & Chambers, 2012) as follows:

$$\overrightarrow{D_s}\ (s, x, y; g_s) = \sup_{\theta}\{\theta\ :\ (s - \theta . g_s\ , x, y) \in T,\ \theta \in R_+\} \quad (5)$$

The translation seeks to find combinations of translated soil attributes (s) that are just sufficient to produce a fixed level of output (y) from a given vector of non-soil inputs (x). Therefore, the directional distance function measures how far the soil vector (s) is from the best-practice production frontier (or isoquant) using the direction vector (g_s) as a yardstick or numeraire (Hailu & Chambers, 2012). Suppose we have two vectors of soil attributes, that is, s' and $\bar{s}$. It is assumed that the vector of soil attributes s' is what is observed on a given plot while $\bar{s}$ could be any reference vector of soil attributes. The reference vector is the benchmark and could simply be the sample average value of soil attributes, or a set of values that are recommended as ideal by researchers, if such a set exists. A Luenberger soil quality indicator (SQIL) for s' is defined as the difference in distances from the production frontier of the soil quality vector s' and the benchmark soil quality vector $\bar{s}$, given other non-soil inputs x' and output y' (Hailu & Chambers, 2012):

$$SQIL\ (s', \bar{s}, x', y'; g_s) = \overrightarrow{D_s}\ (s', x', y'; g_s) - \overrightarrow{D_s}(\bar{s}, x', y'; g_s) \quad (6)$$

If the soil vectors s' and $\bar{s}$ have the same distance from the production frontier in the direction of translation metric g_s, then the soil quality difference is zero. If, then s' has higher quality attributes than $\bar{s}$ and is thus richer; the soil quality indicator will be positive in this case, and vice versa. However, the distance function values and quality indicator will generally depend on the values of non-soil inputs (x) and outputs used to define the reference frontier. Thus, it is natural to generate a second quality measure using a different set of non-soil input (x) and output (y) vectors to define the best-practice frontier. One natural choice for such an alternative measurement is to use the sample average values for non-soil inputs and output, that is, $\bar{x}$ and $\bar{y}$. The second soil quality indicator would then be:

$$SQIL\ \left(s', \bar{s}, \bar{x}, \bar{y}; g_s\right) = \overrightarrow{D_s}\ \left(s', \bar{x}, \bar{y}; g_s\right) - \overrightarrow{D_s}\left(\bar{s}, \bar{x}, \bar{y}; g_s\right) \quad (7)$$

Finally, following Chambers (2002), the two quality indicators (Equations (6) and (7)) can be averaged to define the Luenberger soil quality indicator (SQI):[4]

$$SQI\ \left(s', \bar{s}, x', \bar{x}, y', \bar{y}; g_s\right) = \frac{1}{2}\left\{SQIL\left(s', \bar{s}, x', y'; g_s\right) + SQIL\left(s', \bar{s}, \bar{x}, \bar{y}; g_s\right)\right\} \quad (8)$$

The construction of this quality indicator accounts for the underlying relationship between soil attributes, non-soil inputs, and outputs as it is defined using a production frontier. Soil quality is related to the productive capacity of soil, as implied by the underlying production technology. A positive (negative) value of SQI indicates that the soil quality vector s' is more (less) productive than the benchmark vector $\bar{s}$. A SQI of zero implies no difference between the soil quality vectors $(s', \bar{s})$.

2.2 | Estimation of directional distance functions

The technology frontier against which the directional distance function values in Equations (6) and (7) are measured can be defined using non-parametric methods such as data envelopment analysis (DEA) (Charnes et al., 1978) or parametric methods including stochastic frontier analysis (SFA) (Aigner et al., 1977) and deterministic frontiers (Aigner & Chu, 1968) estimated using mathematical programming (MP) as in Färe et al. (1993) and Hailu and Veeman (2000). The DEA approach was chosen for this study because it does not require ex-ante assumptions about the functional relationship between soil, non-soil inputs and crop output, and because it is easier to implement.

Given a set of N plot-level observations on soil attributes, non-soil inputs, and outputs, we can use DEA to define the first distance function value in Equation (6),

[4] This averaging of the two indicators is similar to the procedure where the Malmquist productivity index is defined as the geometric mean of two alternative productivity indexes defined at two different data points. Luenberger indicators are defined using directional distance functions based on translation and, thus, are absolute measures. Therefore, they are averaged using arithmetic means. Malmquist indexes are defined using radial distance functions based on radial or proportional change and are thus relative measures that are averaged using geometric means.

$\overrightarrow{D_s}(s', x', y'; g_s)$, as follows:

$$\overrightarrow{D_s}\ (s'_n, x'_n, y'_n; g_s) = max\ \beta$$

Subject to:

$$\sum_{n=1}^{N} \lambda_n y'_{nm} \geq y'_{nm},\ m = 1, 2, \ldots, M$$

$$\sum_{n=1}^{N} \lambda_n x'_{nl} \leq x'_{nl},\ l = 1, 2, \ldots, L$$

$$\sum_{n=1}^{N} \lambda_n s'_{nk} \leq s'_{nk} - \beta g_{s'_{nk}},\ k = 1, 2, \ldots, K$$

$$\sum_{n=1}^{N} \lambda_n = 1,\ n = 1, 2, \ldots, N$$

The second distance function value, for the benchmark soil vector, $\overrightarrow{D_s}(\bar{s}, x', y'; g_s)$, is defined similarly:

$$\overrightarrow{D_s}\ (\bar{s}_n, x'_n, y'_n; g_s) = max\ \beta$$

Subject to:

$$\sum_{n=1}^{N} \lambda_n y_{nm} \geq y'_{nm},\ m = 1, 2, \ldots, M$$

$$\sum_{n=1}^{N} \lambda_n x_{nl} \leq x'_{nl},\ l = 1, 2, \ldots, L$$

$$\sum_{n=1}^{N} \lambda_n s_{nk} \leq \bar{s}_{nk} - \beta g_{\bar{s}_{nk}},\ k = 1, 2, \ldots, K$$

$$\sum_{n=1}^{N} \lambda_n = 1,\ n = 1, 2, \ldots, N$$

The above two distance function values measure how far (or how productive) the soil vectors are relative to a production frontier for observed non-soil input (x') and output (y') values. The calculation of the second set of distance function values that compare the soil vectors against the frontier for $(\bar{x}, \bar{y})$ are shown below.

$$\overrightarrow{D_s}\ (s'_n, \bar{x}_n, \bar{y}_n; g_s) = max\ \beta$$

Subject to:

$$\sum_{n=1}^{N} \lambda_n y_{nm} \geq \bar{y}_{nm},\ m = 1, 2, \ldots, M$$

$$\sum_{n=1}^{N} \lambda_n x_{nl} \leq \bar{x}_{nl},\ l = 1, 2, \ldots, L$$

$$\sum_{n=1}^{N} \lambda_n s_{nk} \leq s'_{nk} - \beta g_{s'_{nk}},\ k = 1, 2, \ldots, K$$

$$\sum_{n=1}^{N} \lambda_n = 1,\ n = 1, 2, \ldots, N$$

$$\overrightarrow{D_s}\ (\bar{s}_n, \bar{x}_n, \bar{y}_n; g_s) = max\ \beta$$

Subject to

$$\sum_{n=1}^{N} \lambda_n y_{nm} \geq \bar{y}_{nm},\ m = 1, 2, \ldots, M$$

$$\sum_{n=1}^{N} \lambda_n x_{nl} \leq \bar{x}_{nl},\ l = 1, 2, \ldots, L$$

$$\sum_{n=1}^{N} \lambda_n s_{nk} \leq \bar{s}_{nk} - \beta g_{\bar{s}_{nk}},\ k = 1, 2, \ldots, K$$

$$\sum_{n=1}^{N} \lambda_n = 1,\ n = 1, 2, \ldots, N$$

Given that we do not have externally defined or suggested reference values for soil input values, non-soil inputs, and output, we chose the sample mean values for the reference vectors used above. The direction vector is also defined as the sample mean vector of soil attributes, as is commonly done in the literature using directional distance functions (Hailu & Chambers, 2012). Before estimation, the five non-soil input variables, six soil attributes, and two output variables were normalized by their sample mean values for numerical efficiency. TableA1 shows data are normalised by sample mean values. Therefore, our direction vector becomes a unit vector. The use of the unit vector for direction is equivalent to the use of the average sample direction for the translation (Khataza et al., 2017). This approach is convenient because it allows us to interpret distances to the frontier in terms of the mean bundle of

TABLE 1 Summary statistics of input, output, and soil variables across rice (*Kharif* crop) and wheat (*Rabi* crop) crops.

Wheat crop	Mean	SD	CV (%)
Wheat yield (kg/acre)	1161.96	337.21	29
Non-soil inputs			
Seed rate (kg/acre)	51.85	4.06	8
Nitrogen applied (kg/acre)	56.28	19.75	35
Other variable costs (PKR/acre)	3804.79	1905.24	50
Machinery cost (PKR/acre)	8619.50	2090.88	24
Total labor (h/acre)	52.52	29.12	55
Soil attributes			
Electrical conductivity (dS/m)	5.86	3.74	64
pH	8.80	.25	3
Ammonium	6.78	2.03	30
Phosphate (ppm)	25.19	13.80	55
Potassium (ppm)	169.88	68.25	40
Organic matter (%)	1.45	.33	23
Rice crop			
Rice yield (kg/acre)	1342.72	241.59	18
Non-soil inputs			
Seed rate (kg/acre)	5.66	2.34	41
Nitrogen applied (kg/acre)	60.64	20.60	34
Other variable costs (PKR/acre)	6733.37	3690.72	55
Machinery cost (PKR/acre)	20,030.63	7657.87	38
Total labor (h/acre)	139.60	35.57	25
Soil attributes			
Electrical conductivity (dS/m)	4.48	3.72	83
pH	7.74	.25	3
Ammonium	6.13	2.03	33
Phosphate (ppm)	21.63	11.91	55
Potassium (ppm)	163.54	66.58	41
Organic matter (%)	.59	.34	58

soil quality attributes. The normalisation of variable values by their respective sample mean values is done for ease of interpretation. Otherwise, it has no substantive effect on the nature of the results derived besides scaling. All distance value calculations were done in R using the APEAR package (Hailu, 2013).

2.3 | Study area, data collection, and descriptive statistics

We used farm plot-level data from a household survey conducted between April and June 2019 in the Punjab province of Pakistan. Punjab has over half (about 57%) of the cultivated area in Pakistan and accounts for about 73% of the bulk cereal production in the country (Government of Pakistan, 2018). The survey used a multi-stage, stratified sampling procedure to select farm households. First, three agroecological zones were selected, representing the irrigated areas of Punjab: (i) rice–wheat zone; (ii) maize–wheat–mix zone; (iii) cotton–mix zone (Ahmad et al., 2019; Sheikh et al., 2022a). Second, within the three agroecological zones, we randomly selected one representative district: Hafizabad, Jhang, and Bahawalnagar, respectively. Third, in consultation with the Deputy Director of the Agriculture Extension Department, two *tehsils*[5] were specifically chosen within each selected district: one with a relatively high salt-affected area and the other with relatively normal soil properties. Fourth, in each *tehsil*, two Union Councils[6] (UCs) were selected. Fifth, in each UC, two *mouzas*[7] were selected, with one village randomly chosen from each *mouza*. Twenty-one farm households were randomly selected from each village as sample households. The final sample of 504 farm households were interviewed during the survey. Figure 1 shows the location of the three study districts.

Surveyed farm households either managed or cultivated agricultural plots during 2018–19 for two cropping seasons: *Kharif* (summer season) 2018 and *Rabi* (winter season) 2018–19. The summary statistics are reported in Table 1. Inputs used in both rice and wheat production are total labor (h/acre), seed (kg/acre), nitrogen fertilizer (kg/acre), machinery costs (PKR/acre) that include irrigation cost, and other variable input costs (PKR/acre) such as pesticides/weedicides, farmyard manure, and other chemical fertilizers.

The unique feature of these data is that soil samples were collected from land managers' eligible plots using a zigzag-sampling pattern (SFRI, 2021), with a minimum of 10–30 soil core samples taken to a depth of 30 cm from randomly selected locations. The core samples were mixed to create a composite soil sample for each plot. The samples were then sent to the University of Agriculture, Faisalabad, and analyzed following the procedure mentioned in SFRI, 2021. Six soil attributes, including soil pH, electrical conductivity (EC), soil organic matter (SOM), phosphorus (P), potassium (K), and ammonium (NH_4), were carefully chosen after consultation with professors and soil scientists at the Soil Salinity Research Institute. Each of these attributes is treated as input in an agricultural production process.

Soil EC is an index of salt concentration and an indicator of soil salinity—a common problem in irrigated agricultural land in arid zones (Corwin & Lesch, 2005). Generally, the higher the EC, the higher the soil salinity. The optimal EC of soil is crop-specific and depends on the environmental conditions. An increase in EC can result from water

[5] *Tehsil* is an administrative unit of a district.
[6] Union council is an administrative unit of a district.
[7] *Mouza* is an administrative unit of a district, which usually comprises 5–8 villages.

AGRICULTURAL ECONOMICS The Journal of the International Association of Agricultural Economists IAAE

TABLE 2 Summary statistics of soil quality indicators for wheat production by agroecological zone.

Agroecological zone	Min.	First quartile	Median	Mean	Third quartile	Max.	Skewness	Kurtosis
Rice–wheat	−.271	−.081	−.041	−.022	.012	.341	1.149	5.229
Maize–wheat–mix	−.308	−.102	−.041	−.023	.032	.445	.859	4.369
Cotton–mix	−.280	−.064	−.053	−.018	.006	.448	1.412	6.942

TABLE 3 Summary statistics of soil quality indicators for rice crop.

Agroecological zone	Min.	First quartile	Median	Mean	Third quartile	Max.	Skewness	Kurtosis
Rice–wheat	−.178	−.034	−.019	−.001	.006	.438	2.585	12.319
Maize–wheat–mix	−.186	−.024	−.005	.033	.096	.538	1.397	5.512
Cotton–mix	−.127	−.066	.016	.052	.102	.421	.974	2.770

TABLE 4 Soil quality indicator for wheat production and its relationship with actual soil attributes: OLS regression results.

	Rice–wheat zone	Maize–wheat–mix zone	Cotton–mix zone
Soil electrical conductivity (EC)	.027***	.038***	.034***
	(.008)	(.008)	(.005)
Soil pH	.029***	.021**	.015**
	(.006)	(.008)	(.006)
Ammonium (NH_4)	.013*	.037***	.027***
	(.007)	(.011)	(.009)
Phosphorous (P)	.026***	.011	.009
	(.008)	(.010)	(.008)
Potassium (K)	.018***	.032***	.028***
	(.006)	(.010)	(.009)
Soil organic matter (SOM)	.027***	.016*	.017***
	(.008)	(.009)	(.007)
Constant	−.022***	−.023***	−.018***
	(.006)	(.007)	(.006)
Observations	227	192	195
R^2	.417	.467	.448
F statistic	28.35*** (df = 6, 220)	48.48*** (df = 6, 185)	25.02*** (df = 6, 188)

Robust standard errors in parenthesis *$p < .1$; **$p < .05$; ***$p < .01$.

losses through evapotranspiration or a lack of drainage. Using brackish groundwater for irrigation intensifies the problem (Corwin & Lesch, 2005; Ghafoor et al., 2012; Jesus et al., 2015; Kijne, 1996; Qadir & Oster, 2004; Qadir et al., 2014). Soil salinity (EC > 4 dS m^{-1}) negatively affects plant growth (Ashraf, 2009; Katerji et al., 2009; Qadir et al., 2000; Semiz et al., 2014). On average, the soils had EC values above 4 for wheat and rice crops in the sampled agroecological zones. For the analysis, no distinction is made between EC values over the harmless range of 0–4, that is, all are scored as 1. EC values above four are converted into scores ranging from 0 to 1. It is calculated by dividing four by the EC value, where values close to 0 indicate higher EC values (less suitable for production), and those close to 1 indicate lower EC values (more suitable for production). Like all other variables, the new EC variable is normalised by the sample mean value before the estimation for numerical convenience.

Soil pH affects soil nutrient solubility, plant nutrient availability, and organic matter decomposition influencing soil microorganism activities. Nutrient availability for plant uptake varies depending on soil pH. The soil pH can influence many plant characteristics, including height, biomass, and pollen production (Jiang et al., 2016). In general, nitrogen and potassium are readily available at soil pH 6.5—8, whereas phosphorus becomes more available at soil

TABLE 5 Soil quality indicator for rice production and its relationship with actual soil attributes: OLS regression results.

	Rice–wheat zone	Maize–wheat–mix zone	Cotton–mix zone
Soil electrical conductivity (EC)	.015**	.045***	.051*
	(.006)	(.012)	(.024)
Soil pH	.013*	.023*	−.016
	(.006)	(.012)	(.018)
Ammonium (NH_4)	.008	.015	.026
	(.007)	(.020)	(.019)
Phosphorous (P)	.012	.036**	.067**
	(.008)	(.017)	(.028)
Potassium (K)	.014*	.001	.080***
	(.007)	(.015)	(.027)
Soil organic matter (SOM)	.010*	.006	.043
	(.005)	(.013)	(.030)
Constant	−.001***	.033**	.051**
	(.006)	(.011)	(.022)
Observations	203	100	23
R^2	.173	.321	.695
F statistic	9.25*** (df = 6, 196)	11.01*** (df = 6, 93)	36.25*** (df = 6, 16)

Robust standard errors in parenthesis $^*p < .1$; $^{**}p < .05$; $^{***}p < .01$.

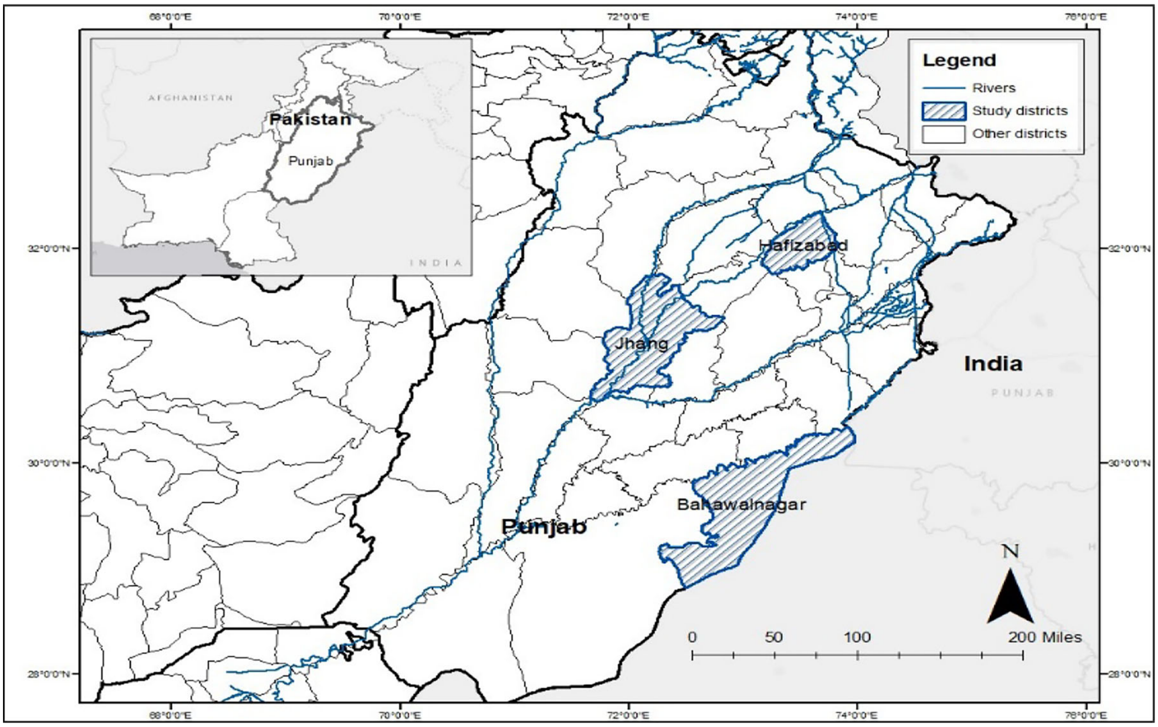

FIGURE 1 Map of the study areas showing the three irrigated districts.
Source: Authors.

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
IAAE

TABLE 6 Soil quality indicator for pooled data series (rice and wheat production) and its relationship to actual soil attributes: OLS regression results.

Soil attributes	Wheat	Rice
Soil electrical conductivity (EC)	.033***	.023***
	(.003)	(.008)
Soil pH	.026***	.012**
	(.003)	(.006)
Ammonium (NH_4)	.017***	.002
	(.004)	(.007)
Phosphorous (P)	.017***	.007
	(.005)	(.009)
Potassium (K)	.023***	.021**
	(.005)	(.008)
Soil organic matter (SOM)	.022***	.027***
	(.005)	(.010)
Rice–wheat zone	.016*	.010
	(.007)	(.013)
Cotton–mix zone	.013	.128***
	(.009)	(.047)
Constant	−.049***	−.007
	(.005)	(.010)
Number of plots	614	326
R^2	.457	.202
F-statistics	84.03***	16.88***
	(df = 8,605)	(df = 8,317)

Robust standard errors in parenthesis $^*p < .1$; $^{**}p < .05$; $^{***}p < .01$.

pH 5.5–7.5. In addition, microorganism activity affected by soil pH can inhibit the conversion of ammonium (NH_4) to nitrate (NO_3), decreasing the nitrogen supply to plants. The average soil pH in our sample was 8.8 and 7.7 for wheat and rice crops, respectively, indicating an alkaline soil. Soil pH on farmland can be managed through amendments, fertilization, tillage practices, improving soil organic matter (SOM) levels, and selecting green manuring crops (McCauley et al., 2017). In the analysis, the soil pH measure is transformed by subtracting each observed pH value from the maximum pH scale value of 14, rescaling it to a new range between 0 and 7. This rescaled pH value is then normalized by the sample mean value before the estimation for numerical convenience.

Troeh and Thompson (1993) identified 17 essential elements for plant growth, including primary macronutrients, such as nitrogen, phosphorus, and potassium. The unavailability of any of these primary nutrients can limit crop yield. The presence of SOM is crucial for fertile soil as it provides essential plant nutrients, influences soil structure, buffers the optimal soil pH range for plant growth (Havlin et al., 2016), and improves water holding capacity and soil aggregation (McCauley et al., 2017), thus serving as a key indicator of soil quality. Moreover, SOM decomposition can increase under elevated temperatures, but adequate soil moisture is critical for stimulating this process. Farmland can serve as carbon sinks, conditional upon having adequate levels of SOM. A reduction in SOM increases soil erosion, reduces carbon sequestration, and implies the need for increased organic or inorganic fertilization to maintain soil quality (McCauley et al., 2017). In temperate regions, SOM often ranges from 3%–4% (Estefan et al., 2013); the average SOM value in our sample was 1.45% and .59% for wheat and rice crops, respectively.

A total of 646 samples were collected from plots belonging to 504 households. However, land managers grew wheat and rice on 614 and 326 plots, respectively. Non-soil inputs and soil attributes were incorporated into the specification of the directional distance function. The mean yields for wheat and rice were about 1162 and 1343 kg per acre,[8] respectively. However, there was considerable variability in wheat and rice yields, with coefficients of variation (CV) values being 29% and 18%, likely due to the high variation in soil attributes and non-soil inputs (Table 1). Machinery costs were high for rice due to the high groundwater extraction costs associated with rice water requirements. Rice seed application rates varied because land managers commonly grow rice in a nursery using about 2–10 kg of seed to avoid unsprouted seeds before transplanting seedlings into a prepared field. The wheat crop had significantly more variation in labor hours used than the rice crop, possibly due to the manual harvesting of wheat (i.e., added labor requirements) to save straw for livestock feed or sell it in the market.

3 | RESULTS AND DISCUSSION

3.1 | Distribution of soil quality indicators across agroecological zones

Given the observations for the ith plot (s', x', y') for the wheat and rice crops, we defined a benchmark vector $(\bar{s}, \bar{x}, \bar{y})$ and direction vector g_s for each agroecological zone (i.e., rice–wheat, maize–wheat–mix, and cotton–mix) to compute directional distance function values in Equations (6) and (7). The sample mean values used as a benchmark vector for each agroecological zone are in Appendix Table A1.

The unit vector $g_s = (-1, -1, -1, -1, -1, -1)$ was used as a direction vector for the six soil attributes (EC, pH, NH_4, P, K, and SOM). Since the data were normalized

[8] 1 hectare = 2.47 acres.

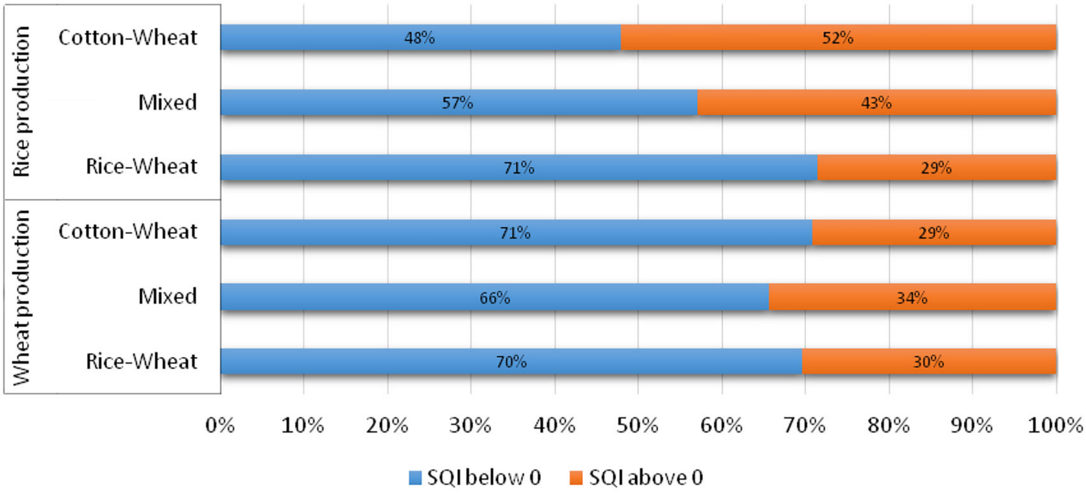

FIGURE 2 Proportion of plots below or above benchmark soil quality for wheat and rice production by agroecological zone.

by sample mean values, the yardstick or direction vector g_s used to measure soil quality is equivalent to the bundle of sample mean values for the soil attributes, that is $(\overline{EC}, \overline{pH}, \overline{NH_4}, \bar{P}, \bar{K}, \overline{SOM})$. Equations (6) and (7) generated two alternative soil quality indicators to rank soil vectors relative to the benchmark soil vector in the context of the production frontier for each agroecological zone. By taking the average of the two indicators (i.e., Equations (6) and (7)), Equation (8) constructs an aggregate measure of overall soil quality (SQI) for each agroecological zone.

In Table 2, the results from the DEA frontier representation show that the SQI for the wheat crop ranged from −.271 to .341 in the rice–wheat zone, −.308 to .445 in the maize–wheat–mix zone, and −.280 to.448 in the cotton–mix zone. The mean values for each agroecological zone were −.022, −.023, and −.018, respectively, and the median values were −.041, −.041, and −.053. The negative sign implies that the plots had lower soil quality than the benchmark soil quality vector on average. As seen in Table 1, there is high variability in most soil attributes, and this is translated into the variability in the constructed SQI values for each agroecological zone. Excess values of particular soil attributes do not necessarily reflect an overall improvement in soil quality or the SQI estimate. On the other hand, the sample mean vector of soil attributes, which is used as the benchmark, is more likely to be balanced in terms of soil attributes. As a result, the sample average soil quality indicator (SQI) value is likely to be negative as plots are, on average, less productive (less balanced) than the benchmark soil vector. This is confirmed by the skewed SQI values in Table 2. Figure 2 also shows that about two-thirds of the wheat plots for each agroecological zone had SQI values below zero, implying that most plots in each agroecological zone had lower soil quality

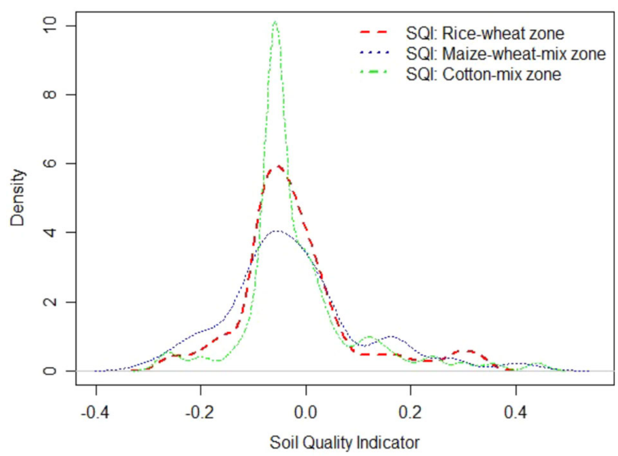

FIGURE 3 Distribution of soil quality indicator (SQI) values for wheat production by cropping zone.

than that represented by the benchmark (average bundle of soil attributes for the sample).

Figure 3 shows the density plots of SQI values for wheat production by agroecological zone. The plots show that the cotton–mix zone has the highest consistency in plot-level soil quality, followed by the rice–wheat and maize–wheat–mix zones. The cotton–mix zone has a higher, narrower, and positively skewed distribution peak for the soil quality indicator but is flatter in both tails than the other zones, indicating less variation in soil quality among plots.

The rice–wheat zone has a relatively low distribution peak compared to the cotton–mix zone, but with similar flatter tails. The maize–wheat–mix zone had a lower peak and wider distribution in soil quality than the other two zones. Based on the consistency in soil quality profiles of sampled plots across agroecological zones, the cotton–mix

AGRICULTURAL ECONOMICS The Journal of the International Association of Agricultural Economists IAAE

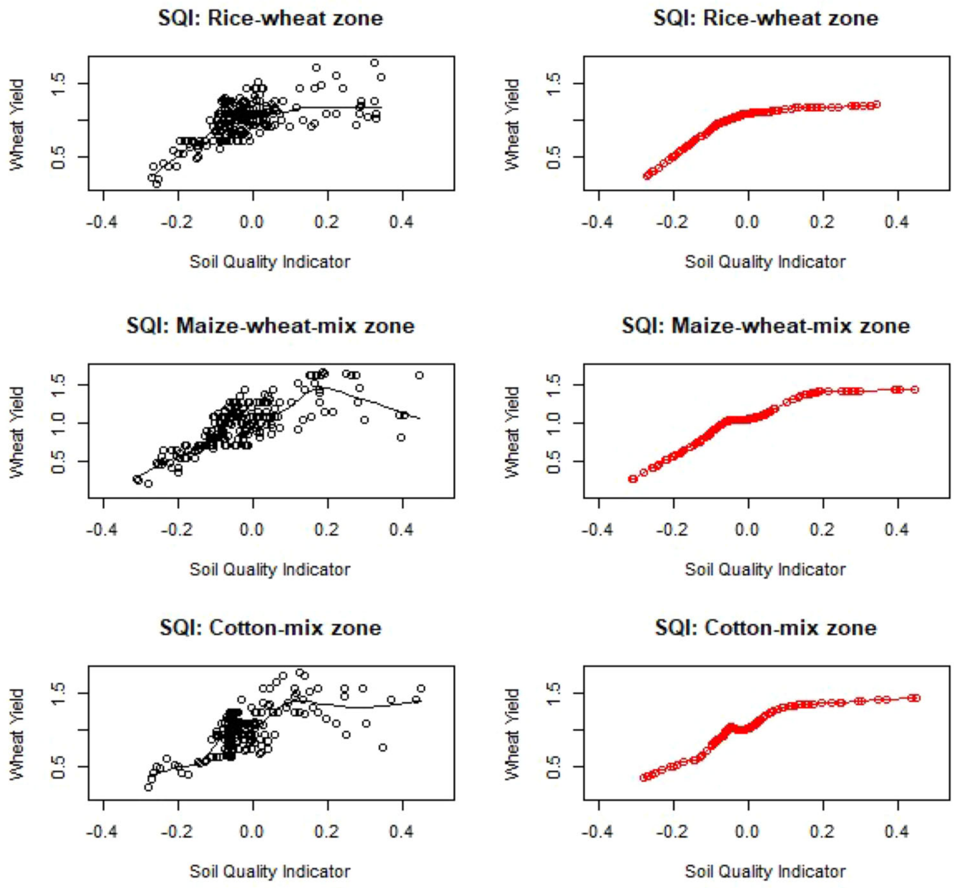

FIGURE 4 Wheat yield–SQI relationship by agroecological zone.
Note: Figures in black are scatterplots of SQI values and wheat yields, while the plots in red are the same values along the fitted curves.

and rice–wheat zones are comparatively more suitable for wheat production.

As Figure 4 shows the estimated SQI values are correlated with reported wheat yield levels for all three cropping systems. A locally weighted scatterplot smoothing (LOESS) tool was used to fit a line to a scatter plot. The fitted curves show that wheat yield increases linearly with increasing SQI up to a point, beyond which the effect of extra soil quality on yield is low or extremely low. This highlights a weaker correlation with yield at higher soil quality levels, which can be explained by factors other than soil quality playing a more significant role in determining yield when soil quality is rich enough to meet crop growth needs.

Overall, the estimated SQI scores for the wheat plots indicate that the plots in the cotton–mix and rice–wheat zones have more consistent soil quality (i.e., less variation in soil quality across farm plots) than the maize–wheat–mix zone. This is consistent with the observation that these two zones have higher net returns from wheat production than the maize–wheat–mix zone (Ahmad et al., 2019; Government of Pakistan, 2018).

For rice production, the estimated SQI values are in Table 3. The values ranged from −.178 to .438 for the rice–wheat zone, −.186 to .538 for the maize–wheat–mix zone, and −.127 to.421 for the cotton–mix zone, with average values of −.001, .033, and .052, respectively. In the rice–wheat zone, about 71% of the plots had degraded soil quality for rice production relative to the benchmark soil quality, compared with about 48% in the cotton–mix zone and 57% in the maize–wheat–mix zone (Figure 2).

The distribution of SQI values for rice production by agroecological zone is in Figure 5. The distribution is narrow with a right skew for plots used for rice production in the rice–wheat zone with similar soil qualities, implying less variability in soil quality and more consistent rice production among plots than the other zones.

The SQI distribution in the maize–wheat–mix zone is positively skewed and relatively wide (diversity) compared with the rice–wheat zone, implying that some plots had less consistent (low variation) soil quality while others were much better than the benchmark soil quality. This indicates that the productive capacity of plots allows rice cropping in the maize–wheat–mix zone. In the

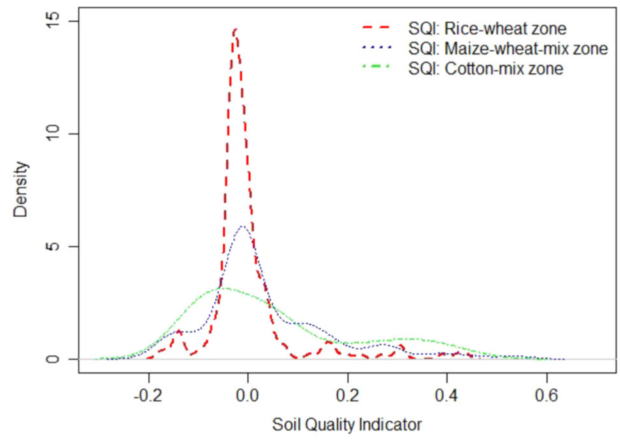

FIGURE 5 Distribution of soil quality indicator (SQI) for rice crop, by agroecological zone.

cotton–mix zone, the distribution of SQI implies that plots have wide-ranging soil qualities and are less consistent for rice cropping.

Figure 6 shows that rice yield did not continue to increase linearly with soil quality. The fitted lines in all three agroecological zones highlight that, under low soil quality, a stronger correlation existed between soil quality and rice yield, as soil quality is a constraining input and contributes more to rice yield. However, beyond a threshold SQI level, a higher soil quality indicator did not contribute much to increasing rice yields, which would be influenced by external non-soil inputs.

Overall, the productive capacity of soils in the rice–wheat and maize–wheat–mix zones is consistent with plot-level soil quality for rice production. However, fewer rice-producing plots in the cotton–mix zone, with variability due to limited data points, led to inconsistency in plot-level soil quality (Figure 5). Thus, it is not clear whether the soils in the cotton–mix zone are suitable for rice cultivation. As expected, the highest net returns from rice production were in Bahawalnagar (cotton–mix zone) followed by Hafizabad (rice–wheat zone), and Jhang (maize–wheat–mix zone)—in line with their suitability for rice cultivation (Ahmad et al., 2019).

Furthermore, we calculated a measure of soil quality by pooling data on soil attributes, non-soil inputs, and crop yields (for wheat and rice production) from the three agroecological zones (see Table A2). The results show that soil quality for wheat production is positively skewed with a short central peak and broader right tail in the distribution (see Figure A1), implying that some plots have higher soil quality than the benchmark; land managers of such plots could be motivated to adopt leguminous crops to maintain soil quality.

The average SQIs reported in Tables 2 and 3 are estimated using the DEA model. Kneip et al. (2015) observed that averaging DEA efficiency scores that are equally weighted may result in substantial bias that may dominate the variance, because the order of the variance can be less than the bias, thus invalidating the application of the standard central limit theorem (CLT). They derived a new fundamental CLT for simple averages of DEA efficiency scores. Simar and Zelenyuk (2018, 2020) have extended the method for generating the new fundamental CLT by Kneip et al. (2015) using Monte-Carlo experiments. The main insight is that when the sample size in DEA analysis increases, this may increase the empirical coverage of the true values by estimated confidence intervals based on the CLT, supporting the CLT results. However, for relatively small sample sizes and high dimensional DEA problems (i.e., many inputs and outputs), the estimated confidence intervals based on the CLT tend to undercover the true values. Therefore, we followed Kneip et al. (2015) and the procedure outlined in Simar and Zelenyuk (2018, 2020) to correct for bias in estimated averages of soil quality indicators and provide a 95% confidence band using Monte-Carlo simulations methods. The results are reported in Appendix Table A3. We do not find bias in the estimated average SQIs. This result aligns with the findings of West et al. (2022), where the authors observed that the bias correction factors were approximated to zero.

3.2 | Regression-based weight derivation across agroecological zones

This part of the analysis generates regression-based weights that could be used to aggregate a vector of individual soil quality attributes into an SQI without reference to a production frontier or non-soil inputs and crop output data. Following the approach used by Jaenicke and Lengnick (1999) and Hailu and Chambers (2012) to evaluate the viability or attractiveness of such a method, we generated regression-based weights using the frontier estimated SQI as the dependent variable and the vector of the individual soil quality attributes as the explanatory variables. For the regression, each soil attribute is scaled by its standard deviation (divided) and converted into a unit-free measure. The standardized coefficients are comparable across individual soil attributes and have a straightforward interpretation. Each coefficient describes the effect of a one-standard-deviation change in the soil attribute on the SQI value.

The results for all agroecological zones are presented in Table 4 (wheat) and Table 5 (rice). For wheat production, the R^2 values were 42%, 47%, and 45% for the rice–wheat, maize–wheat–mix, and cotton–mix zones, respectively, indicating that the six constituent soil attributes explain a sizeable portion of the variation in overall SQI. The

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
IAAE

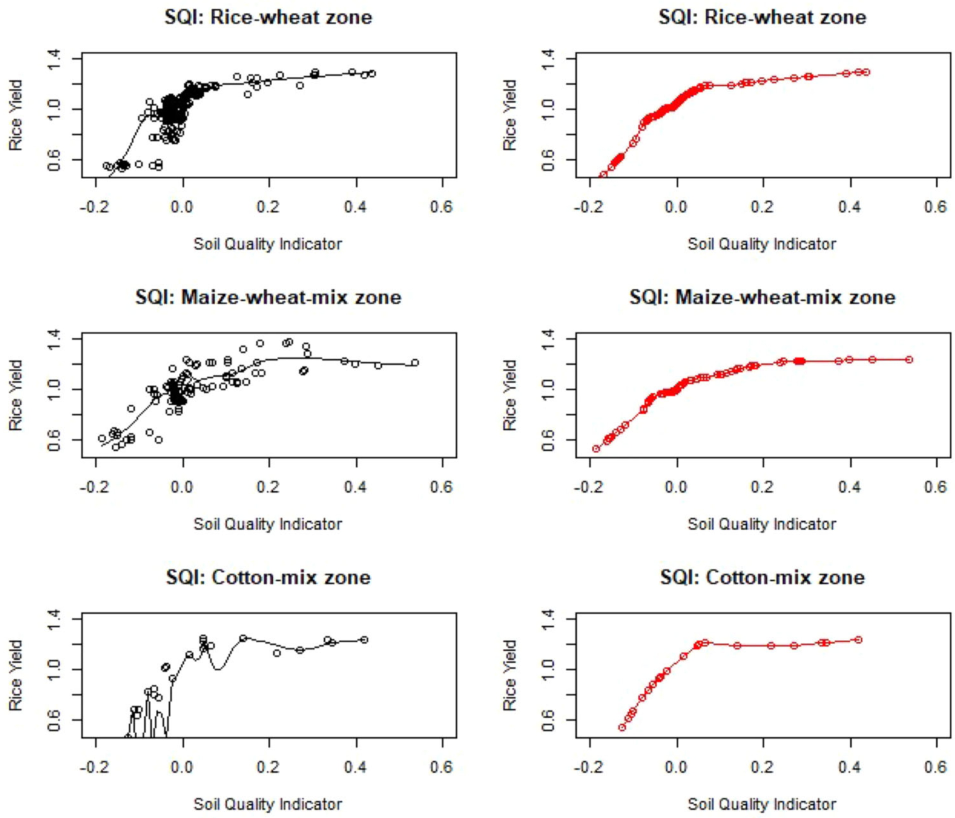

FIGURE 6 Rice yield–SQI relationship by agroecological zone.
Note: Figures in black are scatterplots of SQI values and wheat yields, while the plots in red are the same values along the fitted curves.

coefficient for each individual soil attribute measures the marginal effect that each attribute has on SQI, which can be used as the weight needed to aggregate the individual attributes into the soil quality indicator.

For the rice–wheat zone, the regression weights suggest that the relative importance of soil attributes in defining SQI are as follows: pH > EC > SOM > P > K > NH_4. In the maize–wheat–mix zone, soil EC plays the strongest role, EC > NH_4 > K > pH > SOM. In the cotton–mix zone, EC is also the most important attribute, EC > K > NH_4 > SOM > pH. The contribution of P to SQI in the maize–wheat–mix and cotton–mix zones was statistically insignificant and negligible relative to other soil attributes. However, the rankings ignore the size of the differences in the coefficients and thus might not be the best way to identify the relative importance of soil attributes across zones. If we look at average coefficient sizes across zones, the overall picture is that EC, K, NH_4, and SOM are the most important attributes for wheat production followed by pH.

The results of a similar analysis for rice production are in Table 5. Soil quality appears to be driven by EC and K—the same attributes that were important in wheat production. The R^2 is quite low in the rice–wheat zone (17%). Soil quality is influenced by the combination of soil attributes that contributes to the overall SQI. Moreover, the R^2 is higher in the maize–wheat–mix zone (32%) and is highest in the cotton–wheat zone (69%). The marginal effects of these attributes varies by zone. In the rice–wheat zone, soil EC is the most crucial factor, followed by K, pH, and P. For the maize–wheat–mix zone, EC > P, with the effects of pH, K, SOM, and NH_4 not significant. NH_4, pH, and SOM are not crucial factors in the cotton–mix zone. Overall, EC and K are wheat and rice crops' most influential soil attributes.

In Table 6, we pooled the data across zones and estimated one regression for the wheat SQI series and another for the rice SQI series. As for the individual zone analyses, the same three attributes were identified as the most important in the pooled regressions, namely, EC, K, and SOM for wheat and rice production. pH was also identified as important for wheat and rice production, consistent with the overall picture from the individual zone analysis.

Overall, soil electrical conductivity (EC) was important for both wheat and rice production. Experimental research in the Indus basin of Pakistan's Punjab showed that salt-affected soils produce about 32% and 48% less wheat and rice yields than non-saline soils (Qadir et al., 2014). Approximately 25% of irrigated land in the province is saline, causing about 25–70% land productivity decline on slight to moderate saline soils (Martin et al., 2006). In this context, soil amelioration or sustainable land management practices for salt-affected lands is important for improving or maintaining soil sustainability, protecting the environment, and enhancing agricultural productivity (Sheikh et al., 2022b).

The second-most important soil attribute identified for both wheat and rice production was potassium (K). Land managers using N and P fertilizer often ignore K, accelerating the mining of K resources from soil (Sui et al., 2015; Wu et al., 2014; Xu et al., 2014) and eventually leading to K deficiency that limits sustainable agriculture production (Liu et al., 2009; Wang & Wu, 2015). Potassium acts as a regulator and stimulates the absorption of nutrients, such as N and P (Hou et al., 2019), which is useful for boosting metabolic function and stress tolerance (Wang & Wu, 2015). In K-deficient soils, adequate K application is crucial for improving crop yields. Therefore, applying K fertilizer increased rice yields by 9.8–29.3% (Ye et al., 2020) and increased wheat grain yields (Wang et al., 2020).

In Pakistan, the absence of farming practices that guide the application of inorganic fertilizer to improve soil resources has resulted in the over-exploitation of essential soil nutrients, such as phosphorus and potassium. Unfortunately, these nutrients are both underutilized by land managers (Appendix Table A5). The recommended P to N ratio is 1:2 (Ali et al., 2017); however, the ratio used by land managers in Pakistan was far from optimal and almost stagnant from 1991–92 to 2018–19. In addition, the optimal level for K is yet to be determined (Ali et al., 2017). The ratio of K to N was almost stagnant until 2016–17, gradually increasing to 1:50 in 2018–19 (Table A5). This unbalanced fertilizer use has serious implications for nutrient-use efficiency, excessive soil mining, yields, and environmental sustainability (Concepcion, 2007; Gruhn et al., 2000). Ali et al. (2017) found that land managers applied a lower level of fertilizer on poor quality land than the most fertile land, contrary to the recommended doses for the less fertile land and possibly because the land managers encountered financial and credit constraints.

We used bootstrap regressions with 5000 replications to assess the variability of estimated coefficients and standard errors of the regression models for each sample. The point estimates of the coefficients, as well as their standard errors and 95% confidence interval are reported in Table A4. Our analysis indicates that the estimated coefficients do not change (i.e., no bias) but there is slight variability in the bootstrap standard errors compared to the original standard errors, but this does not affect significance levels. For most of the coefficients, the confidence interval does not include zero suggesting that the soil attributes are significantly related to the soil quality indicator. The distribution of the bootstrapped coefficients is roughly symmetric and centred around the point estimate except for cotton–mix zone and maize–wheat–mix zone for rice crop, suggesting that the estimates remain stable and reliable even in the presence of sampling variability (Figure A2). We note that Banker et al. (2019) demonstrated that the straightforward two-stage DEA + OLS model outperforms the Simar–Wilson model that uses truncated regression in the second stage (Simar & Wilson, 2007).

In summary, the analysis above showed that, in the absence of detailed information on non-soil inputs and outputs, weights of the relative contribution of the six soil attributes can be used to construct SQI estimates for practitioners, policymakers, and development projects. The weights assigned to the wheat and rice cropping enterprises can be translated to similar cropping enterprises and land-use types within the irrigated agroecological zones. However, the weights for wheat cannot be translated to rice and vice versa, due to their distinct environmental requirements for growth and development. Rice thrives in soft, puddled, and water-saturated soil conditions. In contrast, wheat necessitates well-pulverized soil with fine tilth, maintaining a proper balance of moisture, air, and thermal regime (Mahajan & Gupta, 2009).

However, it is important to recognize some limitations of this approach: (1) our results are based on one data set that provides a snapshot of soil quality at a given point in time and does not help to understand soil health dynamics over time; (2) and the SQI results are based on six soil attributes that focus on soil chemical and biological properties but not soil physical properties (i.e., texture, structure, porosity, soil bulk density).

The study highlights some issues related to the use of non-experimental data. For example, we found low R^2 for the regression relating our constructed soil quality indicator to the individual soil quality attributes compared to those in Hailu and Chambers (2012). One plausible explanation is that non-experimental data include more heterogeneity in the soil quality attributes, non-soil inputs and outputs. Another explanation is that the relative importance of noise in non-experimental data is likely to be larger. There are variables defining the production environment that are not incorporated into the analysis, and these variables are likely to vary more in non-experimental data sets than in experimental ones. However, it should also be noted that there are advantages to using non-experimental data. For example, it captures heterogeneity across farms and reflects better the real production environment faced by farmers compared to data generated from experimental stations. Finally, our findings demonstrate some sensitivity of weights to land use types. Thus, further research employing data from more diverse land use types is required to provide more evidence on the robustness of such estimated weights and the extent to which they can be used to estimate soil quality indicators in other agroecological regions. In addition, while the study utilized cross-sectional data to construct the soil quality indicator, future research could explore using longitudinal data to investigate variations in soil quality indicator over time. The studies could incorporate weather variables in their analyses to effectively control for the impacts of climate variability on soil health and crop production.

4 | CONCLUSIONS AND POLICY IMPLICATIONS

Soil health is a crucial concern for land managers and policymakers worldwide. A soil quality indicator (SQI) provides summary information on soil health and is superior to soil health information provided as a vector of individual soil quality attributes that are difficult to evaluate. Numerous studies in the soil science literature have developed soil quality indices using ad hoc weights of different soil attributes. On the contrary, the Luenberger SQI is utilises a production function-based framework to compare the productive capacity of a vector of soil attributes against a benchmark soil quality vector. We used this method to evaluate soil quality for two major crops (i.e., wheat and rice) in three agroecological zones (i.e., rice–wheat, maize–wheat–mix, and cotton–mix) in Punjab, Pakistan.

The SQIs were constructed from six best-practice frontiers estimated using directional distance functions in the data envelopment analysis framework. The results showed that the cotton–mix and rice–wheat zones are more suitable for wheat production than the maize–wheat–mix zone. The SQI and crop yield relationship exhibits diminishing returns to improving soil quality. Besides the frontier method, a simpler regression-based approach was used to generate weights for direct aggregation of soil quality attributes (EC, pH, ammonium, phosphorous, potassium, and organic matter) to compute a SQI. The SQI was found to be most sensitive to soil electrical conductivity (EC) and K for wheat and rice crops. Elevated soil salinity levels and excessive nitrogen, phosphorus, and potassium negatively affects plant growth and development. Therefore, our results implies that the provincial government of Punjab should promote efforts aimed at site-specific land restoration and conservation practices to keep soil EC to acceptable levels to enhance plant growth and productivity. Moreover, there would be benefits from improved extension services and education programs on practices that can enhance the effectiveness and balanced use of chemical fertilizers to overcome NPK deficiency problems.

The SQI information generated for agroecological zones in Punjab can be used by policymakers and practitioners to achieve site-specific agricultural and environmental goals. First, SQI can provide the basis for land suitability classification that could aid government departments, such as the Soil Survey of Punjab, to create provincial guidelines for agricultural land use planning and to improve the regulatory framework for large-scale investment in agriculture. In addition, the soil quality information is useful for agencies interested in promoting context-specific biodiversity and conservation programs, ensuring the conservation of the land resources and the sustainability of agricultural production. Second, the SQI could be used as a site-specific evaluation tools by the provincial agricultural department to monitor the outcome of conservation programs and assess possible compensation payments to land managers.

ACKNOWLEDGMENTS

We acknowledge the joint financial support for this research from the UWA International Research Training Program (RTP) scholarship and International Food Policy Research Institute (IFPRI) under contract No. 2019×103.UWA. Dr Davies acknowledges funding from the CGIAR Research Initiative on NEXUS Gains.

Open access publishing facilitated by The University of Western Australia, as part of the Wiley - The University of Western Australia agreement via the Council of Australian University Librarians.

REFERENCES

Abdulai, A., & Binder, C. R. (2006). Slash-and-burn cultivation practice and agricultural input demand and output supply. *Environment and Development Economics*, *11*(2), 201–220. https://doi.org/10.1017/S1355770İ05002779

Afsharian, M., & Ahn, H. (2014). The Luenberger indicator and directions of measurement: A bottoms-up approach with an empirical illustration to German savings banks. *International Journal of Production Research*, *52*(20), 6216–6233. https://doi.org/10.1080/00207543.2014.922712

Ahmad, A., Khan, M., Shah, S. H. H., Kamran, M., Wajid, S. A., Amin, M., Khan, A., Arshad, M. N., Cheema, M. J. M., Saqib, Z. A., Ullah, R., Ziaf, K., ul Huq, A., Ahmad, S., Fahad, M., Waqas, M. M., & Abbas, A., Iqbal, A. (2019). Agro-ecological zones of Punjab, Pakistan.

Aigner, D. J., & Chu, S. F. (1968). On estimating the industry production function. *The American Economic Review*, *58*(4), 826–839. https://www.jstor.org/stable/1815535

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Ali, M., Ahmed, F., Channa, H., & Davies, S. (2017). Pakistan's fertilizer sector: structure, policies, performance, and impacts. In *Agriculture and the rural economy in Pakistan*, 219–272. University of Pennsylvania Press.

Amin, G. R., & Toloo, M. (2007). Finding the most efficient DMUs in DEA: An improved integrated model. *Computers & Industrial Engineering*, *52*(1), 71–77. https://doi.org/10.1016/j.cie.2006.1.003

Andrews, S. S., & Carroll, C. R. (2001). Designing a soil quality assessment tool for sustainable agroecosystem management. *Ecological Applications*, *11*(6), 1573–1585. https://doi.org/10.1890/1051-0761(2001)011[1573:DASQAT]2.0.CO;2

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The soil management assessment framework: a quantitative soil quality evaluation method. *Soil Science Society of America Journal*, *68*(6), 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Andrews, S. S., Karlen, D. L., & Mitchell, J. P. (2002). A comparison of soil quality indexing methods for vegetable production systems in Northern California. *Agriculture, Ecosystems & Environment*, *90*(1), 25–45. https://doi.org/10.1016/S0167-8809(01)00174-8

Arshad, M. A., & Martin, S. (2002). Identifying critical limits for soil quality indicators in agro-ecosystems. *Agriculture, Ecosystems & Environment*, *88*(2), 153–160. https://doi.org/10.1016/S0167-8809(01)00252-3

Ashraf, M. J. B. A. (2009). Biotechnological approach of improving plant salt tolerance using antioxidants as markers. *Biotechnology advances*, *27*(1), 84–93. https://doi.org/10.1016/j.biotechadv.2008.09.003

Askari, M. S., & Holden, N. M. (2014). Indices for quantitative evaluation of soil quality under grassland management. *Geoderma*, *230–231*, 131–142. https://doi.org/10.1016/j.geoderma.2014.04.019

Askari, M. S., & Holden, N. M. (2015). Quantitative soil quality indexing of temperate arable management systems. *Soil and Tillage Research*, *150*, 57–67. https://doi.org/10.1016/j.still.2015.01.010

Askari, M. S., O'Rourke, S. M., & Holden, N. M. (2015). Evaluation of soil quality for agricultural production using visible–near-infrared spectroscopy. *Geoderma*, *243*, 80–91. https://doi.org/10.1016/j.geoderma.2014.12.012

Banker, R., Natarajan, R., & Zhang, D. (2019). Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using data envelopment analysis: Second stage OLS versus bootstrap approaches. *European Journal of Operational Research*, *278*(2), 368–384. https://doi.org/10.1016/j.ejor.2018.1.050

Barbier, E. B., & Hochard, J. P. (2016). Does land degradation increase poverty in developing countries? *Plos One*, *11*(5), e0152973. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4864404/pdf/pone.0152973.pdf

Barrett, C. B., Bellemare, M. F., & Hou, J. Y. (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development*, *38*(1), 88–97. https://doi.org/10.1016/j.worlddev.2009.06.002

Bellon, M. R., & Taylor, J. E. (1993). Folk' soil taxonomy and the partial adoption of new seed varieties. *Economic Development and Cultural Change*, *41*(4), 763–786. https://doi.org/10.1086/452047

Chambers, R. G. (2002). Exact nonradial input, output, and productivity measurement. *Economic Theory*, *20*, 751–765. https://doi.org/10.1007/s001990100231

Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and Distance Functions. *Journal of Economic Theory*, *70*(2), 407–419. https://doi.org/10.1006/jeth.1996.0096

Chang, H. H., & Wen, F. I. (2011). Off-farm work, technical efficiency, and rice production risk in Taiwan. *Agricultural Economics*, *42*(2), 269–278. https://doi.org/10.1111/j.1574-0862.201.00513.x

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Concepcion, R. N. (2007). Sustainable Fertilization Management of Croplands: The Philippines Scenario. *Bangkok: Food and Agriculture Organization. Accessed June*, *20*, 2014.

Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and Electronics in Agriculture*, *46*(1), 11–43. https://doi.org/10.1016/j.compag.2004.1.005

de Paul Obade, V., & Lal, R. (2013). Assessing land cover and soil quality by remote sensing and geographical information systems (GIS). *Catena*, *104*, 77–92. https://doi.org/10.1016/j.catena.2012.1.014

de Paul Obade, V., & Lal, R. (2014). Soil quality evaluation under different land management practices. *Environmental Earth Sciences*, *72*(11), 4531–4549. https://doi.org/10.1007/s12665-014-3353-z

de Paul Obade, V., & Lal, R. (2016). Towards a standard technique for soil quality assessment. *Geoderma*, *265*, 96–102. https://doi.org/10.1016/j.geoderma.2015.11.023

Di Falco, S., & Chavas, J. P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics*, *91*(3), 599–611. https://doi.org/10.1111/j.1467-8276.2009.01265.x

Doran, J. W., & Parkin, T. B. (1994). Defining and assessing soil quality. *Defining soil quality for a sustainable environment*, *35*, 1–21. https://doi.org/10.2136/sssaspecpub35.c1

Estefan, G., Sommer, R., & Ryan, J. (2013). Methods of soil, plant, and water analysis. *A manual for the West Asia and North Africa region*, 3.

Färe, R., Grosskopf, S., Lovell, C. K., & Yaisawarng, S. (1993). Derivation of shadow prices for undesirable outputs: a distance function approach. *The review of economics and statistics*, 374–380. https://doi.org/10.2307/2109448

Fuwa, N., Edmonds, C., & Banik, P. (2007). Are small-scale rice farmers in eastern India really inefficient? Examining the effects of microtopography on technical efficiency estimates. *Agricultural Economics*, *36*(3), 335–346. https://doi.org/10.1111/j.1574-0862.2007.00211.x

Gerber, N., Nkonya, E., & von Braun, J. (2014). Land degradation, poverty and marginality. *Marginality: Addressing the nexus of poverty, exclusion and ecology*, 181–202.

Ghafoor, A., Murtaza, G., Rehman, M. Z., Saifullah, & Sabir, M. (2012). Reclamation and salt leaching efficiency for tile drained saline-sodic soil using marginal quality water for irrigating rice and wheat crops. *Land Degradation & Development*, *23*(1), 1–9. https://doi.org/10.1002/ldr.1033

Government of Pakistan. (2018). *Agricultural statistics of Pakistan*. Pakistan Bureau of Statistics.

Gruhn, P., Goletti, F., & Yudelman, M. (2000). *Integrated nutrient management, soil fertility, and sustainable agriculture: current issues and future challenges*. International Food Policy Research Institute.

Hailu, A. (2003). Nonparametric Productivity Analysis with Undesirable Outputs: Reply. *American Journal of Agricultural Economics*, *85*(4), 1075–1077. https://www.jstor.org/stable/1244787

Hailu, A. (2013). A package for productivity and efficiency analysis in R (apear).

Hailu, A., & Chambers, R. G. (2012). A Luenberger soil-quality indicator. *Journal of Productivity Analysis*, *38*(2), 145–154. https://doi.org/10.1007/s11123-011-0255-x

Hailu, A., & Veeman, T. S. (2000). Environmentally sensitive productivity analysis of the Canadian pulp and paper industry, 1959–1994: an input distance function approach. *Journal of environmental economics and management*, *40*(3), 251–274. https://doi.org/10.1006/jeem.2000.1124

Havlin, J. L., Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (2016). *Soil fertility and fertilizers*. Pearson Education India.

Hou, W., Xue, X., Li, X., Khan, M. R., Yan, J., Ren, T., Cong, R., & Lu, J. (2019). Interactive effects of nitrogen and potassium on: Grain yield, nitrogen uptake and nitrogen use efficiency of rice in low potassium fertility soil in China. *Field Crops Research*, *236*, 14–23. https://doi.org/10.1016/j.fcr.2019.03.006

Jaenicke, E. C., & Lengnick, L. L. (1999). A soil-quality index and its relationship to efficiency and productivity growth measures: two decompositions. *American Journal of Agricultural Economics*, *81*(4), 881–893. https://doi.org/10.2307/1244331

Jesus, J. M., Danko, A. S., Fiúza, A., & Borges, M. T. (2015). Phytoremediation of salt-affected soils: a review of processes, applicability, and the impact of climate change. *Environmental Science and Pollution Research*, *22*, 6511–6525.

Jiang, Y., Li, Y., Zeng, Q., Wei, J., & Yu, H. (2016). The effect of soil pH on plant growth, leaf chlorophyll fluorescence and mineral element content of two blueberries. In XI International Vaccinium Symposium 1180, 269–276. https://doi.org/10.17660/ActaHortic.2017.1180.36

Karlen, D. L., Ditzler, C. A., & Andrews, S. S. (2003). Soil quality: why and how? *Geoderma*, *114*(3-4), 145–156. https://doi.org/10.1016/S0016-7061(03)00039-9

Karlen, D. L., Hurley, E. G., Andrews, S. S., Cambardella, C. A., Meek, D. W., Duffy, M. D., & Mallarino, A. P. (2006). Crop rotation effects on soil quality at three northern corn/soybean belt locations. *Agronomy Journal*, *98*(3), 484–495.

Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., & Schuman, G. E. (1997). Soil quality: a concept, definition, and framework for evaluation (a guest editorial). *Soil Science Society of America Journal*, *61*(1), 4–10. https://doi.org/10.2136/sssaj1997.03615995006100010001x

Katerji, N., Mastrorilli, M., Van Hoorn, J. W., Lahmer, F. Z., Hamdy, A., & Oweis, T. (2009). Durum wheat and barley productivity in saline–drought environments. *European Journal of Agronomy*, *31*(1), 1–9. https://doi.org/10.1016/j.eja.2009.01.003

Khataza, R. R., Hailu, A., Kragt, M. E., & Doole, G. J. (2017). Estimating shadow price for symbiotic nitrogen and technical efficiency for legume-based conservation agriculture in Malawi. *Australian Journal of Agricultural and Resource Economics*, *61*(3), 462–480. https://doi.org/10.1111/1467-8489.12212

Kijne, J. W. (1996). *Water and salinity balances for irrigated agriculture in Pakistan*, 6, IWMI.

Kneip, A., Simar, L., & Wilson, P. W. (2015). When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores. *Econometric Theory*, *31*(2), 394–422. https://doi.org/10.1017/S0266466614000413

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Liu, X., Zhang, W., Zhang, M., Ficklin, D. L., & Wang, F. (2009). Spatio-temporal variations of soil nutrients influenced by an altered land tenure system in China. *Geoderma*, *152*(1), 23–34. https://doi.org/10.1016/j.geoderma.2009.05.022

Mahajan, A., & Gupta, R. D. Eds. (2009). Integrated nutrient management (INM) in a sustainable rice—wheat cropping system. *Dordrecht*: Springer.

Mandal, U. K., Ramachandran, K., Sharma, K. L., Satyam, B., Venkanna, K., Udaya Bhanu, M., Mandal, M., Masane, R. N., Narsimlu, B., & Rao, K. V. (2011). Assessing soil quality in a semi-arid tropical watershed using a geographic information system. *Soil Science Society of America Journal*, *75*(3), 1144–1160. https://doi.org/10.2136/sssaj2009.0361

Mandal, U. K., Warrington, D. N., Bhardwaj, A. K., Bar-Tal, A., Kautsky, L., Minz, D., & Levy, G. J. (2008). Evaluating impact of irrigation water quality on a calcareous clay soil using principal component analysis. *Geoderma*, *144*(1-2), 189–197. https://doi.org/10.1016/j.geoderma.2007.11.014

Marenya, P. P., & Barrett, C. B. (2009). Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, *40*(5), 561–572. https://doi.org/10.1111/j.1574-0862.2009.00398.x

Martin, P., Nishida, J., Afzal, J., Akbar, S., Damania, R., & Hanrahan, D. (2006). Pakistan strategic country environmental assessment. *South Asia Region World Bank*, *1*, 1–80.

Masto, R. E., Chhonkar, P. K., Singh, D., & Patra, A. K. (2008). Alternative soil quality indices for evaluating the effect of intensive cropping, fertilisation and manuring for 31 years in the semi-arid soils of India. *Environmental Monitoring and Assessment*, *136*, 419–435. https://doi.org/10.1007/s10661-007-9697-z

McCauley, A., Jones, C., & Olson-Rutz, K. (2017). Soil pH and Organic Matter. *Nutrient Management Module No*. 8. https://store.msuextension.org/publications/AgandNaturalResources/4449-8.pdf

Nabiollahi, K., Golmohamadi, F., Taghizadeh-Mehrjardi, R., Kerry, R., & Davari, M. (2018). Assessing the effects of slope gradient and land use change on soil quality degradation through digital mapping of soil quality indices and soil loss rate. *Geoderma*, *318*, 16–28. https://doi.org/10.1016/j.geoderma.2017.12.024

National Fertilizer Development Centre. (2018). Fertilizer offtake by nutrients by province. http://www.nfdc.gov.pk/Web-Page%20Updating/oftkpnt.htm

Nkonya, E., Gerber, N., Baumgartner, P., von Braun, J., De Pinto, A., Graw, V., Kato, E., Kloos, J., & Walter, T. (2011). The economics of desertification, land degradation, and drought toward an integrated global assessment. *ZEF-Discussion Papers on Development Policy*, (150).

Paul, G. C., Saha, S., & Ghosh, K. G. (2020). Assessing the soil quality of Bansloi river basin, eastern India using soil-quality indices (SQIs) and Random Forest machine learning technique. *Ecological Indicators*, *118*, 106804. https://doi.org/10.1016/j.ecolind.2020.106804

Pieralli, S. (2017). Introducing a new non-monotonic economic measure of soil quality. *Soil and Tillage Research*, *169*, 92–98. https://doi.org/10.1016/j.still.2017.01.015

Qadir, M., & Oster, J. D. (2004). Crop and irrigation management strategies for saline-sodic soils and waters aimed at environmentally sustainable agriculture. *Science of the total environment*, *323*(1-3), 1–19. https://doi.org/10.1016/j.scitotenv.2003.1.012

Qadir, M., Ghafoor, A., & Murtaza, G. (2000). Amelioration strategies for saline soils: a review. *Land Degradation & Development*, *11*(6), 501–521. https://doi.org/10.1002/1099-145X(200011/12)11:6<501::AID-LDR405>3.0.CO;2-S

Qadir, M., Quillérou, E., Nangia, V., Murtaza, G., Singh, M., Thomas, R. J., Drechsel, P., & Noble, A. D. (2014). Economics of salt-induced land degradation and restoration. *Natural resources. Forum*, *38*(4), 282–295. https://doi.org/10.1111/1477-8947.12054

Raiesi, F., & Kabiri, V. (2016). Identification of soil quality indicators for assessing the effect of different tillage practices through a soil quality index in a semi-arid environment. *Ecological Indicators*, *71*, 198–207. https://doi.org/10.1016/j.ecolind.2016.06.061

Ray, S. C., & Mukherjee, K. (1996). Decomposition of the Fisher Ideal Index of Productivity: A Non-Parametric Dual Analysis of US Airlines Data. *The Economic Journal*, *106*(439), 1659–1678. https://doi.org/10.2307/2235206

Rezaei, S. A., Gilkes, R. J., & Andrews, S. S. (2006). A minimum data set for assessing soil quality in rangelands. *Geoderma*, *136*(1-2), 229–234. https://doi.org/10.1016/j.geoderma.2006.03.021

Semiz, G. D., Suarez, D. L., Ünlükara, A., & Yurtseven, E. (2014). Interactive effects of salinity and N on pepper (Capsicum annuum L.) yield, water use efficiency and root zone and drainage salinity. *Journal of plant nutrition*, *37*(4), 595–610. https://doi.org/10.1080/01904167.2013.867985

SFRI. (2021). Soil Fertility Research Institute. Soil and Water Analysis Manual. Punjab, Pakistan https://aari.punjab.gov.pk/system/files/Soil%20Fertility%20Research%20Institute%20Punjab%20Lahore%20.pdf

Sharma, K. L., Mandal, U. K., Srinivas, K., Vittal, K. P. R., Mandal, B., Grace, J. K., & Ramesh, V. (2005). Long-term soil management effects on crop yields and soil quality in a dryland Alfisol. *Soil and Tillage Research*, *83*(2), 246–259. https://doi.org/10.1016/j.still.2004.08.002

Sherlund, S. M., Barrett, C. B., & Adesina, A. A. (2002). Smallholder technical efficiency controlling for environmental production conditions. *Journal of Development Economics*, *69*(1), 85–101. https://doi.org/10.1016/S0304-3878(02)00054-8

Sheikh, A. T., Mugera, A., Pandit, R., Burton, M., & Davies, S. (2022a). The adoption of laser land leveler technology and its impact on groundwater use by irrigated farmland in Punjab, Pakistan. *Land Degradation & Development*, *33*(12), 2026–2038.

Sheikh, A. T., Mugera, A., Pandit, R., Burton, M., & Davies, S. (2022b). What determines the time to gypsum adoption to remediate irrigated salt-affected agricultural lands? Evidence from Punjab, Pakistan. *Soil and Tillage Research*, *217*, 105266.

Shukla, M. K., Lal, R., & Ebinger, M. (2006). Determining soil quality indicators by factor analysis. *Soil and Tillage Research*, *87*(2), 194–204. https://doi.org/10.1016/j.still.2005.03.011

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), 31–64.

Simar, L., & Zelenyuk, V. (2018). Central limit theorems for aggregate efficiency. *Operations Research*, *66*(1), 137–149. https://doi.org/10.1287/opre.2017.1655

Simar, L., & Zelenyuk, V. (2020). Improving finite sample approximation by central limit theorems for estimates from Data Envelopment Analysis. *European Journal of Operational Research*, *284*(3), 1002–1015. https://doi.org/10.1016/j.ejor.202.01.036

Stevens, A. W. (2018). The economics of soil health. *Food Policy*, *80*, 1–9. https://doi.org/10.1016/j.foodpol.2018.08.005

Stocking, M. A. (2003). Tropical soils and food security: the next 50 years. *Science*, *302*(5649), 1356–1359. https://doi.org/10.1126/science.1088579

Sui, N., Zhou, Z., Yu, C., Liu, R., Yang, C., Zhang, F., Song, G., & Meng, Y. (2015). Yield and potassium use efficiency of cotton with wheat straw incorporation and potassium fertilization on soils with various conditions in the wheat–cotton rotation system. *Field Crops Research*, *172*, 132–144. https://doi.org/10.1016/j.fcr.2014.11.011

Sun, J., Wu, J., & Guo, D. (2013). Performance ranking of units considering ideal and anti-ideal DMU with common weights. *Applied Mathematical Modelling*, *37*(9), 6301–6310. https://doi.org/10.1016/j.apm.2013.01.010

Troeh, H. R., & Thompson, L. M.. (1993). *Soils and soil Fertility* (5th ed.). Oxford University Press.

Vasu, D., Singh, S. K., Ray, S. K., Duraisami, V. P., Tiwary, P., Chandran, P., Nimkar, A. M., & Anantwar, S. G. (2016). Soil quality index (SQI) as a tool to evaluate crop productivity in semi-arid Deccan plateau, India. *Geoderma*, *282*, 70–79. https://doi.org/10.1016/j.geoderma.2016.07.010

Wang, Y., & Wu, W.-H. (2015). Genetic approaches for improvement of the crop potassium acquisition and utilization efficiency. *Current Opinion in Plant Biology*, *25*, 46–52. https://doi.org/10.1016/j.pbi.2015.04.007

Wang, Y., Zhang, Z., Liang, Y., Han, Y., Han, Y., & Tan, J. (2020). High Potassium Application Rate Increased Grain Yield of Shading-Stressed Winter Wheat by Improving Photosynthesis and Photosynthate Translocation [Original Research]. *Frontiers in Plant Science*, 11(134),. https://doi.org/10.3389/fpls.202.00134

West, S. C., Mugera, A. W., & Kingwell, R. S. (2022). The choice of efficiency benchmarking metric in evaluating firm productivity and viability. *Journal of Productivity Analysis*, *57*(2), 193–211. https://doi.org/10.1007/s11123-022-00628-1

Wu, L., Cui, Z., Chen, X., Zhao, R., Si, D., Sun, Y., & Yue, S. (2014). High-yield maize production in relation to potassium uptake requirements in China. *Agronomy Journal*, *106*(4), 1153–1158.

Xu, X., He, P., Qiu, S., Pampolino, M. F., Zhao, S., Johnston, A. M., & Zhou, W. (2014). Estimating a new approach of fertilizer recommendation across small-holder farms in China. *Field Crops Research*, *163*, 10–17. https://doi.org/10.1016/j.fcr.2014.04.014

Yakowitz, D. S., Stone, J. J., Lane, L. J., Heilman, P., Masterson, J., Abolt, J., & Imam, B. (1993). A decision support system for evaluating the effects of alternative farm management systems on water quality and economics. *Water Science and Technology*, *28*(3-5), 47–54. https://doi.org/10.2166/wst.1993.0402

Ye, T., Xue, X., Lu, J., Hou, W., Ren, T., Cong, R., & Li, X. (2020). Yield and potassium uptake of rice as affected by potassium rate in the middle reaches of the Yangtze River, China. *Agronomy Journal*, *112*(2), 1318–1329. https://doi.org/10.1002/agj2.20092

Yu, P., Liu, S., Zhang, L., Li, Q., & Zhou, D. (2018). Selecting the minimum data set and quantitative soil quality indexing of alkaline soils under different land uses in northeastern China. *Science of The Total Environment*, *616*, 564–571. https://doi.org/10.1016/j.scitotenv.2017.10.301

Yu-Dong, C. H. E. N., Huo-Yan, W. A. N. G., Jian-Min, Z. H. O. U., Lu, X. I. N. G., Bai-Shu, Z. H. U., Yong-Cun, Z. H. A. O., & Xiao-Qin, C. H. E. N. (2013). Minimum data set for assessing soil quality in farmland of northeast China. *Pedosphere*, *23*(5), 564–576. https://doi.org/10.1016/S1002-0160(13)60050-8

Zhou, Y., Ma, H., Xie, Y., Jia, X., Su, T., Li, J., & Shen, Y. (2020). Assessment of soil quality indexes for different land use types in typical steppe in the loess hilly area, China. *Ecological Indicators*, *118*, 106743. https://doi.org/10.1016/j.ecolind.2020.106743

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sheikh, A. T., Hailu, A., Mugera, A., Pandit, R., & Davies, S. (2024). Soil quality evaluation for irrigated agroecological zones of Punjab, Pakistan: The Luenberger indicator approach. *Agricultural Economics*, *55*, 531–553. https://doi.org/10.1111/agec.12831

APPENDIX

TABLE A1 Sample mean value for each crop by agroecological zone.

Agroecological zone	Crops					
	Wheat			Rice		
Rice–wheat	$\bar{s} = \frac{\sum_{i=1}^{227} s^i}{227}$	$\bar{x} = \frac{\sum_{i=1}^{227} x^i}{227}$	$\bar{y} = \frac{\sum_{i=1}^{227} y^i}{227}$	$\bar{s} = \frac{\sum_{i=1}^{203} s^i}{203}$	$\bar{x} = \frac{\sum_{i=1}^{203} x^i}{203}$	$\bar{y} = \frac{\sum_{i=1}^{203} y^i}{203}$
Maize–wheat–mix	$\bar{s} = \frac{\sum_{i=1}^{192} s^i}{192}$	$\bar{x} = \frac{\sum_{i=1}^{192} x^i}{192}$	$\bar{y} = \frac{\sum_{i=1}^{192} y^i}{192}$	$\bar{s} = \frac{\sum_{i=1}^{100} s^i}{100}$	$\bar{x} = \frac{\sum_{i=1}^{100} x^i}{100}$	$\bar{y} = \frac{\sum_{i=1}^{100} y^i}{100}$
Cotton–mix	$\bar{s} = \frac{\sum_{i=1}^{195} s^i}{195}$	$\bar{x} = \frac{\sum_{i=1}^{195} x^i}{195}$	$\bar{y} = \frac{\sum_{i=1}^{195} y^i}{195}$	$\bar{s} = \frac{\sum_{i=1}^{23} s^i}{23}$	$\bar{x} = \frac{\sum_{i=1}^{23} x^i}{23}$	$\bar{y} = \frac{\sum_{i=1}^{23} y^i}{23}$

Note: s is soil-related inputs, x is non-soil inputs, y is crop yield (output), and i is sample observation.

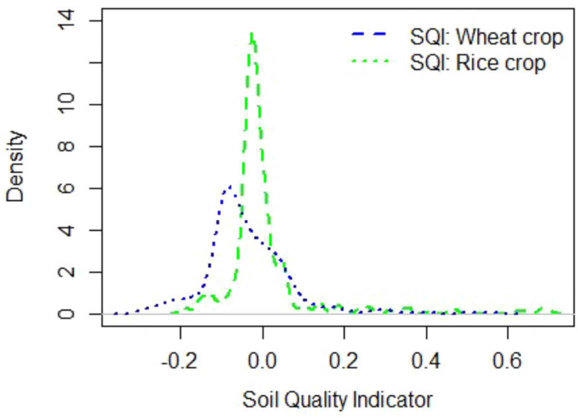

FIGURE A1 Distribution of soil quality indicator for wheat and rice crops.

TABLE A2 Summary statistics of soil quality indicator by crop across the 3 agroecological zones.

Crop	Min.	First quartile	Median	Mean	Third quartile	Max.	Skewness	Kurtosis
Wheat	−.305	−.096	−.056	−.039	.011	.569	1.574	9.434
Rice	−.201	−.036	−.018	.008	.008	.703	3.355	17.071

TABLE A3 The mean of soil quality indicators and their 95% confidence band by crop and agroecological zones.

Crop	agroecological zones	Mean SQI	Bias corrected SQI	Bias	SQI (SD)	SQI_low	SQI_high
Wheat	Pooled	−.039	−.039	.000	.107	−.039	−.038
Rice	Pooled	.008	.008	.000	.121	.007	.009
Wheat	Rice–wheat	−.022	−.022	.000	.111	−.023	−.021
Wheat	Maize–wheat–mix	−.023	−.023	.000	.133	−.025	−.022
Wheat	Cotton–mix	−.018	−.018	.000	.111	−.020	−.017
Rice	Rice–wheat	−.001	−.001	.000	.086	−.002	.000
Rice	Maize–wheat–mix	.033	.033	.000	.131	.030	.036
Rice	Cotton–mix	.052	.052	.000	.158	.041	.062

Note: The procedure follows jackknife bias correction, where original sample is split into two subsamples (50:50) and efficiency scores are estimated for each sample using DEA. Monte-Carlo simulation uses 3000 replications.

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists

TABLE A4 Soil quality indicator for wheat production and its relationship with actual soil attributes: Bootstrap regression results.

Variables	Wheat crop			Rice Crop			Pooled data	
	Rice–wheat zone	Maize–wheat–mix	Cotton–mix zone	Rice–wheat zone	Maize–wheat–mix	Cotton–mix zone	Wheat	Rice
Soil electrical conductivity (EC)	.027***	.038***	.034***	.015***	.045***	.051	.033***	.023***
	(.008)	(.008)	(.005)	(.006)	(.013)	(.037)	(.003)	(.008)
	(.010–.043) †	(.022–.055) †	(.025–.043) †	(.004–.026)	(.019–.071) †	(−.021–.123) †	(.027–.039) †	(.008–.038) †
Soil pH	.029***	.021***	.015***	.013*	.023*	−.016	.026***	.012**
	(.006)	(.008)	(.006)	(.007)	(.013)	(.026)	(.004)	(.006)
	(.017–.041) †	(.006–.037) †	(.004–.027) †	(−.000–.025) †	(−.003–.048) †	(−.067–.034) †	(.019–.033) †	(.000–.025) †
Ammonium (NH_4)	.013*	.037***	.027***	.008	.015	.026	.017***	.002
	(.007)	(.011)	(.009)	(.007)	(.019)	(.029)	(.004)	(.007)
	(−.001–.027) †	(.016–.059) †	(.010–.045) †	(−.006–.021) †	(−.023–.054) †	(−.032–.083) †	(.009–.025) †	(−.011–.015) †
Phosphorous (P)	.026***	.011	.009	.012	.036**	.067**	.017***	.007
	(.008)	(.010)	(.008)	(.008)	(.017)	(.032)	(.005)	(.009)
	(.010–.042) †	(−.008–.031) †	(−.006–.024) †	(−.004–.028) †	(.004–.069) †	(.004–.130) †	(.008–.027) †	(−.011–.025) †
Potassium (K)	.018***	.032***	.028***	.014*	.001	.080**	.023***	.021***
	(.006)	(.010)	(.009)	(.008)	(.017)	(.037)	(.005)	(.008)
	(.005–.030) †	(.012–.052) †	(.010–.045) †	(−.001–.029) †	(−.031–.034) †	(.008 -.153) †	(.013–.033) †	(.005–.037) †
Soil organic matter (SOM)	.027***	.016*	.017***	.010*	.006	.043	.022***	.027***
	(.008)	(.009)	(.007)	(.005)	(.013)	(.055)	(.005)	(.010)
	(.011–.044) †	(−.003–.034) †	(.004–.030) †	(−.001–.020) †	(−.019–.031) †	(−.064–.149) †	(.012–.031) †	(.007–.047) †
Rice–wheat zone							.016**	.010
							(.007)	(.014)
							(.002–.029) †	(−.016–.037) †
Cotton–mix zone							.013	.128***
							(.009)	(.047)
							(−.004–.031) †	(.036–.221) †
Constant	−.022***	−.023***	−.018***	−.001	.033***	.051*	−.049***	−.007
	(.006)	(.007)	(.006)	(.006)	(.011)	(.027)	(.005)	(.010)
	(−.033–−.010) †	(−.037–−.009) †	(−.030–−.007) †	(−.012–.010) †	(.011–.054) †	(−.002–.105) †	(−.058–−.040) †	(−.028–.013) †
Observations	227	192	195	203	100	23	614	326
R^2	.417	.467	.448	.173	.321	.695	.457	.202

Note: Number of bootstrap replications: 5000. Bootstrap standard errors in parenthesis *$p < .1$; **$p < .05$; ***$p < .01$.
†95% confidence interval.

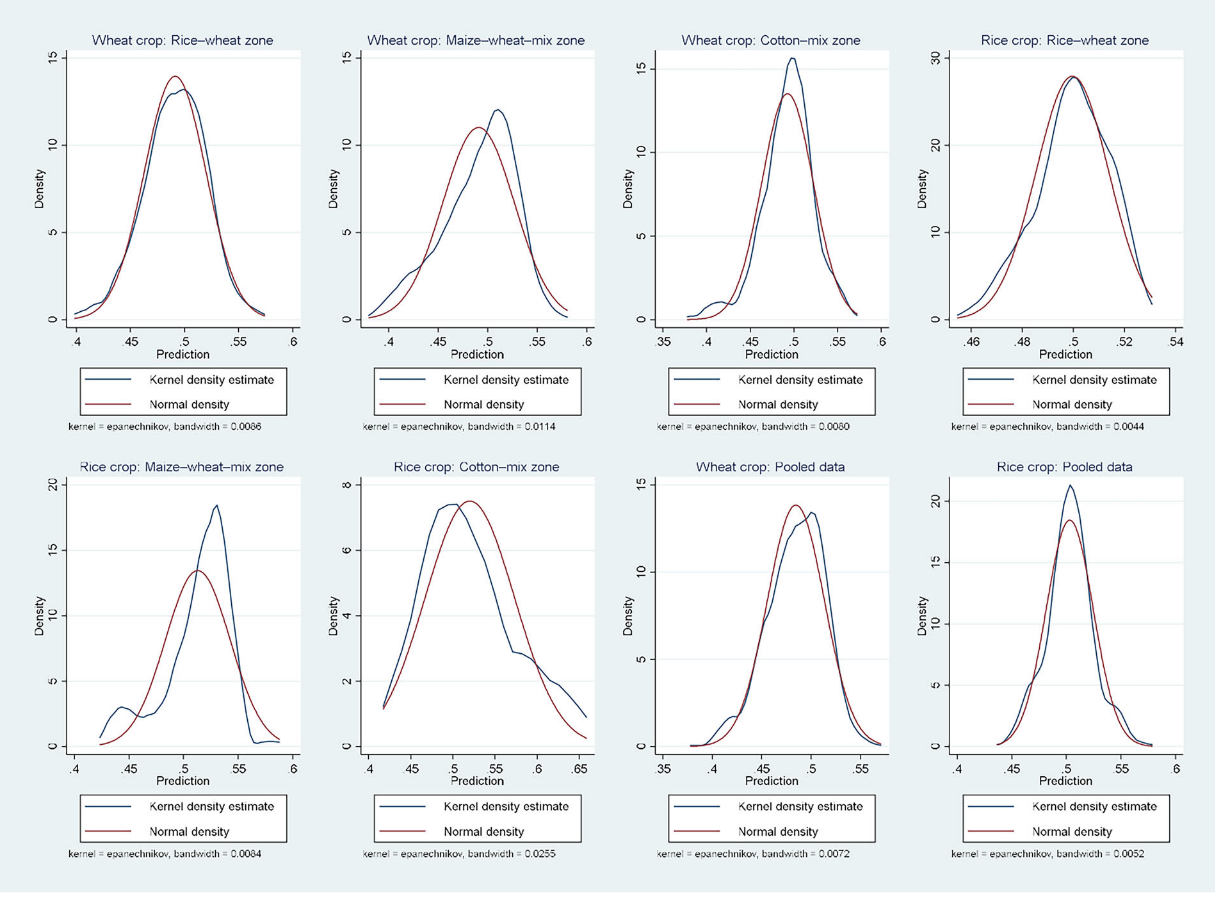

FIGURE A2 Distribution of the bootstrapped coefficients.

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
IAAE

TABLE A5 Fertilizer consumption by nutrients for Punjab province.

Years	N ('000 nutrient tonnes)	P ('000 nutrient tonnes)	K ('000 nutrient tonnes)	Total	N (%)	P (%)	K (%)	P/N	K/N
1991–92	1028	284	15	1327	77.5	21.4	1.1	.28	.01
1992–93	1150	343	17	1510	76.2	22.7	1.1	.30	.01
1993–94	1146	326	16	1487	77.1	21.9	1.1	.28	.01
1994–95	1233	312	12	1557	79.2	20.0	.8	.25	.01
1995–96	1445	372	20	1837	78.7	20.3	1.1	.26	.01
1996–97	1377	309	7	1694	81.3	18.2	.4	.22	.01
1997–98	1390	392	19	1801	77.2	21.8	1.1	.28	.01
1998–99	1406	316	18	1740	80.8	18.2	1.0	.22	.01
1999–00	1500	416	11	1927	77.8	21.6	.6	.28	.01
2000–01	1561	481	18	2060	75.8	23.3	.9	.31	.01
2001–02	1537	426	14	1977	77.7	21.5	.7	.28	.01
2002–03	1578	470	15	2063	76.5	22.8	.7	.30	.01
2003–04	1752	488	15	2255	77.7	21.6	.7	.28	.01
2004–05	1912	636	23	2571	74.4	24.7	.9	.33	.01
2005–06	2049	601	19	2672	76.7	22.5	.7	.29	.01
2006–07	1785	683	31	2499	71.4	27.3	1.2	.38	.02
2007–08	2016	434	20	2470	81.6	17.6	.8	.22	.01
2008–09	2080	454	18	2552	81.5	17.8	.7	.22	.01
2009–10	2515	615	17	3147	79.9	19.5	.5	.24	.01
2010–11	2231	548	24	2804	79.6	19.5	.9	.25	.01
2011–12	2181	452	16	2649	82.3	17.1	.6	.21	.01
2012–13	1988	537	15	2540	78.3	21.1	.6	.27	.01
2013–14	2164	623	17	2804	77.2	22.2	.6	.29	.01
2014–15	2252	714	23	2989	75.3	23.9	.8	.32	.01
2015–16	1772	718	13	2503	70.8	28.7	.5	.41	.01
2016–17	2537	930	31	3498	72.5	26.6	.9	.37	.01
2017–18	2354	925	37	3316	71.0	27.9	1.1	.39	.02
2018–19	2323	810	39	3173	73.2	25.5	1.2	.35	.02

Source: NFDC (2018).

Note: N = nitrogen; P = phosphorus; K = potassium.