International Journal of Multidisciplinary Research and Growth Evaluation

International Journal of Multidisciplinary Research and Growth Evaluation
ISSN: 2582-7138
Received: 04-10-2020; Accepted: 07-11-2020
www.allmultidisciplinaryjournal.com
Volume 1; Issue 5; November-December 2020; Page No. 244-262

Analytical Framework for Linking Soil Fertility Parameters with Agricultural Output Efficiency

Sonna Damian Nduka
IMO State University, Owerri, Nigeria

Corresponding Author: **Sonna Damian Nduka**

DOI: https://doi.org/10.54660/.IJMRGE.2020.1.5.244-262

Abstract
This paper advances an analytical framework that links soil fertility parameters to agricultural output efficiency, enabling evidence-based nutrient management under variable biophysical and socioeconomic conditions. The framework integrates harmonized data layers field measurements (pH, soil organic carbon, total nitrogen, available phosphorus and potassium, micronutrients, cation exchange capacity), soil physical properties (texture, bulk density, structure, infiltration), moisture dynamics, and proximal/remote-sensing covariates into a unified, quality-assured repository. Feature engineering derives stoichiometric ratios (C:N:P), base saturation, acidity indices, and water availability metrics, while geostatistical kriging and digital soil mapping close spatial gaps at farm-to-landscape scales. Agricultural output efficiency is quantified via frontier methods data envelopment analysis and stochastic frontier analysis yielding technical efficiency, marginal abatement cost of nutrient stress, and partial factor productivity for N, P, and K. Causal identification leverages directed acyclic graphs, panel fixed effects, instrumental variables (rainfall anomalies, legacy liming), and difference-in-differences around soil health interventions to disentangle fertility effects from confounders such as climate, cultivar, and management. A hierarchical Bayesian layer pools information across sites and seasons, estimating nonlinear response functions and credible intervals for yield elasticities with respect to pH, SOC, and available P. Spatial-econometric components (SAR/SEM) capture neighborhood effects including nutrient runoff and shared management, while machine-learning ensembles with SHAP values provide model-agnostic interpretability and variable importance. Decision analytics convert elasticities and frontier gaps into prescriptive levers lime rates, balanced NPK blends, micronutrient triggers, organic amendments, and residue management optimized under cost, water, and emission constraints. Practical deployment follows a staged roadmap: data audit and calibration; baseline frontier estimation; causal learning with intervention pilots; and operational dashboards that fuse NDVI/EVI, in-situ sensors, and market signals for adaptive recommendations. Key performance indicators include gains in technical efficiency, fertilizer use efficiency, yield stability, profit per hectare, and reductions in nutrient surpluses and nitrous oxide intensity. By unifying measurement, causality, and decision optimization, the framework makes soil fertility actionable for both smallholders and commercial farms, supports climate-smart intensification, and guides sustainable input allocation across heterogeneous agroecosystems. Governance is embedded through data standards, uncertainty audits, and reproducible pipelines, enabling traceable recommendations for extension services, cooperatives, and agribusinesses while aligning with SDG 2, SDG 12, and fertilizer policies.

Keywords: Soil Fertility; Agricultural Efficiency; Stochastic Frontier Analysis; Data Envelopment Analysis; Bayesian Hierarchical Modeling; Digital Soil Mapping; SHAP Interpretability; Nutrient Use Efficiency; NDVI/EVI; Risk-Based Recommendations.

1. Introduction

Agricultural systems continue to exhibit persistent yield and efficiency gaps even where improved seed, irrigation, and mechanization are present, because decisions on nutrient supply, soil amendments, and cropping practices are often made with limited, inconsistent, or non-spatially explicit information about soil fertility. Farmers and agribusinesses routinely confront variability in pH, soil organic carbon, macronutrients and micronutrients, cation exchange capacity, texture, and moisture dynamics across fields and seasons; without a rigorous way to translate this variability into actionable recommendations, inputs are misallocated, fertilizer-use efficiency remains low, and environmental externalities rise. The problem addressed in this work is the absence of an integrated, evidence-based analytical pathway that links measurable soil parameters to farm- and landscape-level output efficiency, separates correlation from causation under confounding climate and management factors, and converts the resulting insights into prescriptive, risk-aware nutrient and amendment strategies (Buenemann, *et al*., 2011, Mamonov, 2019). The objective is to develop a coherent analytical framework that (i) harmonizes multi-source soil, crop, weather, and management data; (ii) constructs spatially resolved fertility indicators and stoichiometric indices that capture binding constraints; (iii) quantifies agricultural output efficiency using frontier methods and elasticity estimates tied to specific soil parameters; (iv) establishes causal attribution of soil-driven gains in productivity and resource efficiency; and (v) operationalizes decision rules and optimization routines that deliver site-specific, economically viable, and environmentally responsible prescriptions. The scope covers smallholder to commercial contexts across diverse agroecological zones and cropping systems, spanning plot-to-landscape scales and in-season to multi-year horizons. It encompasses laboratory and proximal sensing, remote-sensing covariates, digital soil mapping, panel agronomy datasets, and market signals relevant to input and output prices (Assumma, *et al*., 2019, Dur & Yigitcanlar, 2015).

The contributions are fourfold. First, a data architecture and quality protocol that align laboratory assays, field surveys, sensor feeds, and satellite products into a unified, auditable repository with spatial and temporal keys. Second, a measurement and feature-engineering suite that derives interpretable indices such as C:N:P ratios, base saturation, and acidity/alkalinity metrics and propagates uncertainty into downstream analysis. Third, a paired inferential engine that marries efficiency-frontier estimation (e.g., data envelopment analysis and stochastic frontier analysis) with causal tools directed acyclic graphs, panel fixed effects, instrumental variables, and difference-in-differences to isolate the marginal effects of soil parameters on yield and fertilizer-use efficiency under confounding weather, cultivar, and management variation (Monteiro, Martins & Pires, 2018, Reidsma, *et al*., 2011). Fourth, a decision-analytics layer that transforms elasticities and efficiency gaps into prescriptions for lime, balanced NPK and micronutrient blends, organic amendments, and residue management, optimized under cost, water, and emission constraints and delivered through adaptive, spatially explicit dashboards. By integrating measurement, causality, and optimization, the framework enables soil-informed decisions that close yield and efficiency gaps, improve profitability per hectare, and reduce nutrient losses, providing a transparent foundation for extension services, cooperatives, and agribusinesses to scale climate-smart intensification.

Target users (smallholders, commercial farms, extension, agribusiness)

2. Methodology

This methodology integrates agronomic measurement science, geospatial analytics, efficiency frontiers, and causal inference into a single decision pipeline that links soil fertility parameters to agricultural output efficiency. First, multisource data are assembled at plot and farm scales: field surveys of management and inputs, laboratory soil tests, proximal sensors, remote-sensing vegetation indices, and daily weather, alongside market prices to compute profitability metrics. Data engineering follows predictive-analytics and data-quality practices adapted from cross-sector frameworks (e.g., Abass *et al*.; Akinrinoye *et al*.; Atobatele *et al*.), enforcing controlled vocabularies, units, and metadata. Range checks, outlier rules, temporal consistency, and spatial plausibility filters are applied, then missing values are imputed with model-based methods informed by covariates. Security-by-design and robust access controls mirror zero-trust patterns used in digital systems (Ajayi *et al*.), protecting farmer data while enabling reproducible research.

Measurement and feature engineering focus on core soil fertility variables pH, SOC, total N, available P and K, CEC, texture, and moisture and derive agronomic indices including base saturation, acidity/alkalinity flags, C:N:P ratios, buffering/liming need, plant-available water proxies, and salinity/sodicity markers. Remote-sensing layers (NDVI/EVI, thermal) and terrain metrics augment the feature space. Digital soil mapping produces gap-filled surfaces via geostatistics and covariate models, with prediction variances retained for later uncertainty-aware optimization (Buenemann *et al*.; Erb *et al*.). Where biochar or organic inputs are relevant, variable sets reflect their effects on SOC and nutrient dynamics (Agegnehu *et al*.).

Efficiency analysis quantifies how well farms transform inputs into outputs. A two-track specification is used. Nonparametric DEA estimates technical efficiency frontiers with inputs such as fertilizer N/P/K, lime, seed, labor, water, and machinery time, and outputs including yield, gross margin per hectare, and yield stability; metainformation captures production environments and seasons. Parametric SFA provides complementary measures with composed-error models that admit one-sided inefficiency and allow heterogeneity across agroecological zones, cultivars, and management bundles (Hoang). Both approaches generate KPIs such as technical efficiency, partial factor productivities for N/P/K, profit per hectare, and stability metrics. Policy variables (e.g., fertilizer subsidy exposure) are encoded to enable later interpretation in line with literature on subsidies and adoption (Jayne & Rashid; Hemming *et al*.; Ali *et al*.; Arndt *et al*.).

Causal identification distinguishes correlation from impact for soil parameters and management. Directed acyclic graphs formalize assumptions; panel fixed effects absorb time-invariant unobservables; instrumental variables (e.g., rainfall anomalies or distance-weighted access shocks) address endogeneity; and difference-in-differences exploits staged interventions such as lime distribution or advisory rollouts. Spatial econometric models (SAR/SEM) handle neighborhood spillovers and leakage across field boundaries. For predictive interpretability, tree-based and boosting ensembles are fit with SHAP to decompose contributions of soil variables and quantify nonlinear responses and thresholds, while keeping the causal estimates as the basis for prescriptions.

Decision analytics translate elasticities and efficiency gaps into actionable, cost- and risk-aware prescriptions: lime rates to reach target pH, balanced NPK with site-specific ratios and timing, micronutrient packages where deficiencies are detected, and organic amendments (compost, biochar) where SOC is limiting and economics support it. A constrained optimization routine minimizes cost per expected yield gain subject to water limits, labor constraints, and emission or nutrient-loss caps, while propagating uncertainty from DSM variances and model confidence intervals. Rollout is targeted via zonal stratification that overlays prescription surfaces with NDVI/EVI, tenure/market access layers, and extension capacity to prioritize high-return segments.

Implementation uses a modular data lakehouse with versioned datasets, reproducible pipelines, and audit trails borrowed from lakehouse-DevOps practices (Ajayi *et al*.). Dashboards deliver KPIs and maps to agronomists and extension agents, supporting A/B tests and sentinel plots that compare alternative input bundles. A continuous-learning loop ingests new soil tests, yield maps, and price updates to recalibrate DSM models, refresh DEA/SFA frontiers, and update causal estimates; drift detection triggers retraining. Governance codifies standards for data rights, privacy, and sharing; uncertainty audits report intervals for prescriptions and expected gains. This end-to-end method operationalizes integrated soil fertility management by fusing sound measurement, spatial science, economic efficiency, and credible causal evidence into routinized, farmer-facing decisions.

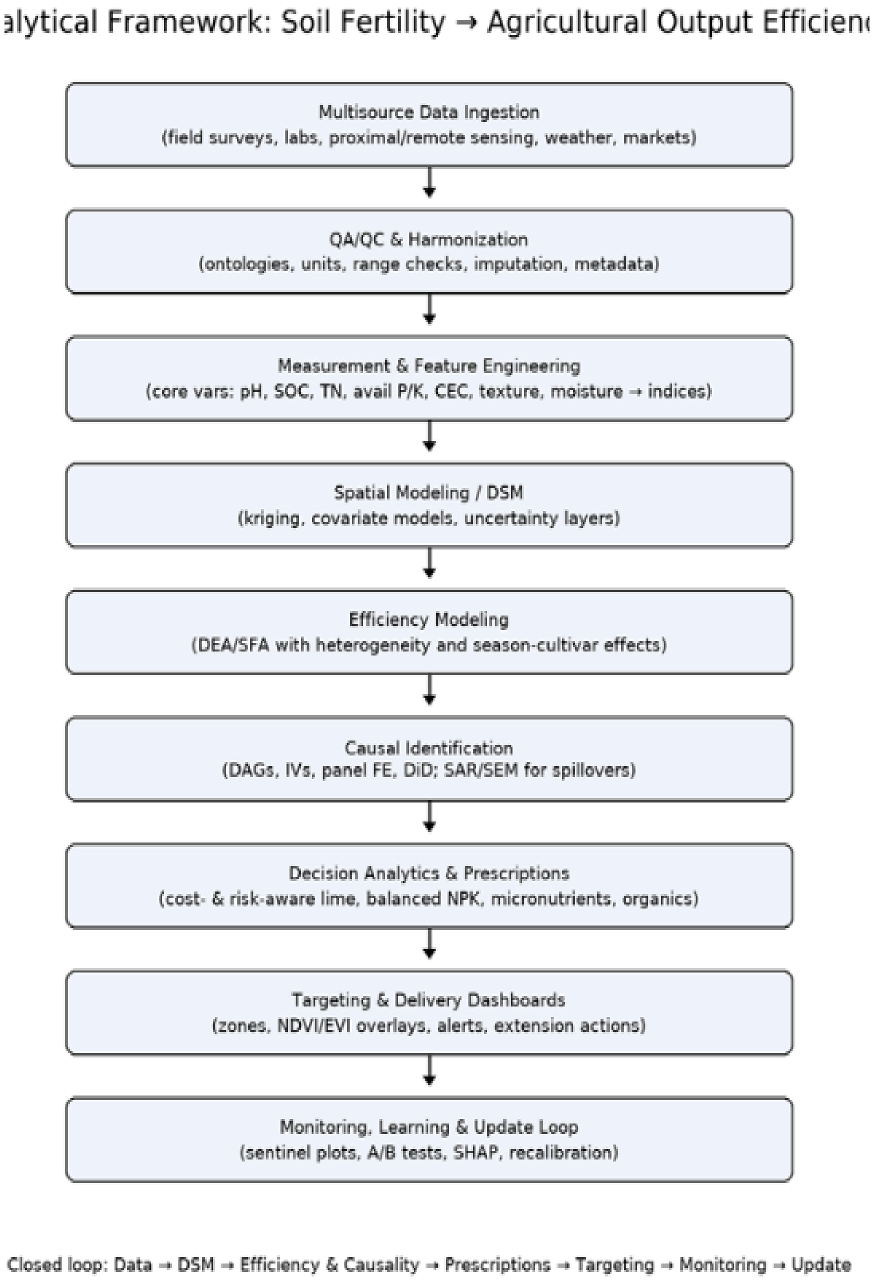

Fig 1: Flowchart of the study methodology

2.1. Background and Conceptual Foundations

The analytical framework rests on biophysical principles that view fields as coupled soil–plant–atmosphere systems in which water, energy, and nutrient fluxes jointly determine crop performance. Roots explore a finite volume of soil whose structure, texture, and pore-size distribution regulate infiltration, drainage, aeration, and mechanical impedance. These physical properties in turn govern the temporal availability of nutrients because diffusion and mass flow carry ions to root surfaces at rates conditioned by water content and temperature. Soil organic matter functions as both a nutrient reservoir and a structural agent: its decomposition supplies nitrogen, phosphorus, and sulfur while improving aggregate stability and cation exchange capacity, which moderates leaching and buffers pH. Microbial communities mediate mineralization–immobilization dynamics, nitrification and denitrification, and the solubilization of phosphorus bound to calcium, iron, or aluminum; their efficiency depends on oxygen status, pH, and available carbon (Dur, Yigitcanlar & Bunker, 2014, Koschke, *et al*., 2012). Canopies intercept radiation and convert it to biomass through photosynthesis, but stomatal conductance and thus carbon assimilation responds to vapor pressure deficit and leaf water potential, linking atmospheric demand to root water uptake capacity. Consequently, yield results from the co-limitation of light, water, and nutrients, with the binding constraint shifting over time as weather and phenology interact with soil properties and management.

Within this continuum, nutrient cycling creates multiple potential bottlenecks. Nitrogen enters via fertilizer, biological fixation, or organic inputs, then partitions among ammonium, nitrate, microbial biomass, and organic pools of varying lability. Losses through ammonia volatilization, nitrate leaching, and nitrous oxide emissions diminish plant-available pools and lower nutrient-use efficiency. Phosphorus cycling is dominated by sorption–desorption and precipitation–dissolution reactions that are highly pH-dependent and strongly influenced by mineralogy and organic ligands; much applied P becomes occluded or sorbed to Fe/Al oxides or precipitated with Ca, rendering plant uptake diffusion-limited (Tang, 2015, Van Westen, 2013). Potassium mobility is intermediate, with exchange sites and interlayer fixation in certain clays buffering solution K. Micronutrients such as zinc, iron, and boron exhibit narrow optimal ranges and strong pH dependence, so mild deviations can cause hidden hunger that depresses yield responses to macronutrients. These dynamics establish the need for a diagnostic approach that identifies the binding constraints among water, pH, organic matter, macronutrients, and micronutrients rather than assuming uniform responses to blanket fertilizer rates.

To translate these processes into operational metrics, the framework adopts efficiency concepts from production economics. Technical efficiency measures how close a farm is to the best-practice frontier the maximum feasible output for given inputs and technology. Allocative efficiency measures whether input mixes are chosen to minimize cost (or maximize profit) given input and output prices, conditional on the technology. Eco-efficiency extends this logic to incorporate environmental burdens, relating output to resource use and emissions (for example, kilograms of grain per kilogram of applied nitrogen and unit of nitrous oxide emitted) (Azumah & Zakaria, 2019, Rashid, *et al*., 2013). In agronomy, a farm can be technically inefficient due to poor timeliness, suboptimal spacing, or pest pressure even if soil nutrients are adequate; it can be allocatively inefficient by overinvesting in N relative to the price of maize and urea while underapplying P or lime; and it can be eco-inefficient by achieving yield gains with excessive N that raises N_2O intensity and nitrate leaching. Distinguishing these dimensions is crucial because soil interventions mostly shift the production frontier and change the marginal product of each nutrient, while management timing and prices dictate movement along the frontier. Figure 2 shows SCP (Structure–Conduct–Performance) framework for farming system analysis presented by Nguyen, *et al*., 2019.

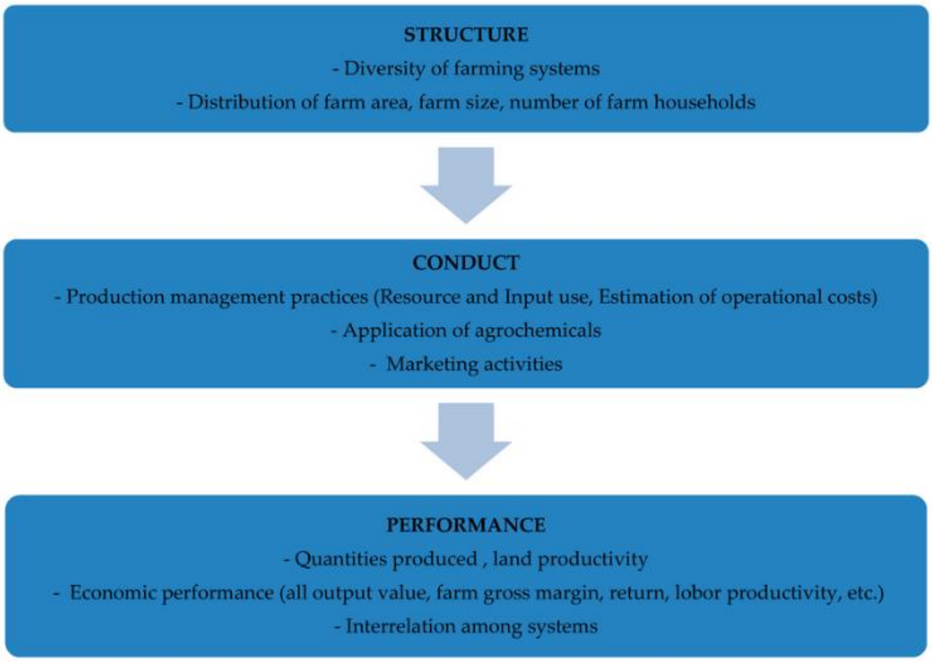

Fig 2: SCP (Structure–Conduct–Performance) framework for farming system analysis (Nguyen, *et al*., 2019).

The conceptual model links soil parameters to output efficiency through three pathways: potential, availability, and delivery. Potential captures the structural and chemical capacity of the soil to support growth texture, depth, bulk density, organic matter, cation exchange capacity, and inherent mineral reserves. Availability translates potential into plant-available pools governed by pH, redox status, sorption, and biological cycling. Delivery represents the spatiotemporal logistics of getting water and nutrients to roots when demand peaks, controlled by rainfall patterns, irrigation, infiltration, root density, and the hydraulic continuity of the profile (Hemming, *et al*., 2018, Jayne, *et al*., 2013). Each pathway is represented by measurable indicators: for potential, texture classes, soil organic carbon, bulk density, and CEC; for availability, Bray or Olsen P, exchangeable K, mineral N, micronutrient tests, and pH; for delivery, infiltration rate, available water capacity, and electrical conductivity proxies for salinity. The framework aggregates these into composite indices such as base saturation, acidity stress index, C:N:P stoichiometric ratios, and water-limited nutrient index that are interpretable and map directly to agronomic levers.

From an efficiency perspective, soil parameters shift both the frontier and the slope of response curves. Liming an acidic soil increases the marginal product of applied P and N by releasing P from Al/Fe complexes and improving root growth, effectively rotating the isoquant inward and raising technical efficiency potential. Building soil organic carbon enhances CEC and water holding, flattening the yield penalty during intra-seasonal dry spells and raising eco-efficiency by reducing N losses and stabilizing yield per unit nutrient (Jayne & Rashid, 2013, Minviel & Latruffe, 2017). Balanced P and K correct hidden constraints that otherwise cause diminishing returns to N, improving allocative efficiency because the cost-minimizing bundle moves toward recommended N:P:K ratios. Micronutrient corrections add small-cost, high-leverage shifts in parts of the landscape where critical thresholds are crossed, often delivering large returns in marginal soils. These causal mechanisms justify modeling strategies that allow interactions and thresholds rather than linear, additive assumptions. Figure 3 shows conceptual framework for organic amendments and plant – soil relationships presented by Agegnehu, Srivastava & Bird, 2017.

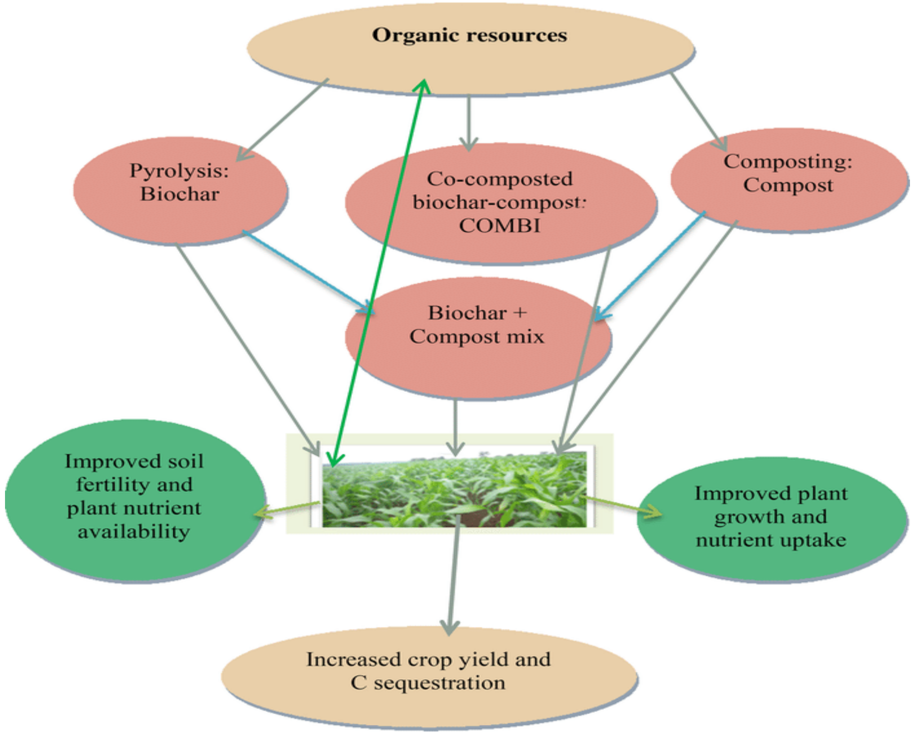

Fig 3: Conceptual framework for organic amendments and plant – soil relationships (Agegnehu, Srivastava & Bird, 2017).

The framework also acknowledges temporal and spatial heterogeneity. Fertility states drift with extraction and replenishment, residue management, and erosion; responses vary by phenological stage, with early-season P affecting root proliferation and flowering, and late-season K affecting grain filling and lodging resistance. Spatially, short-range variability in texture and topographic position creates different water–nutrient dynamics, so prescriptions must be location-specific. The conceptual model therefore treats fields as mosaics of production environments defined by covariates such as elevation, slope, curvature, normalized difference vegetation index history, and electromagnetic induction signals correlated with texture and moisture. Within each environment, response surfaces link soil indicators to output efficiency metrics, and uncertainty is quantified to guide risk-aware decisions (Ali, Rahut & Imtiaz, 2019, Nasrin, Bauer & Arman, 2018).

Integrating economics, the framework connects soil-driven shifts in the frontier to profit and risk. Technical efficiency gains convert to higher expected yields; allocative improvements reduce cost per unit output; eco-efficiency improvements lower emissions intensity and nutrient surplus, which may be monetized through compliance or premium markets. However, uncertainty in weather, prices, and soil measurement implies distributions, not certainties (Arndt, Pauw & Thurlow, 2016, Ricker-Gilbert & Jayne, 2017). The conceptual model therefore includes stochastic dominance and value-at-risk perspectives: a liming program may have a high expected benefit but also variance in response depending on rainfall; balanced NPK may reduce variance by removing binding constraints, stabilizing returns. Prescriptions are therefore optimized for expected profit subject to downside risk constraints and environmental caps, aligning farmer and societal objectives. Figure 4 shows conceptual framework -the relationship between agricultural policies and production practices and diet presented by Bryden, 2012.

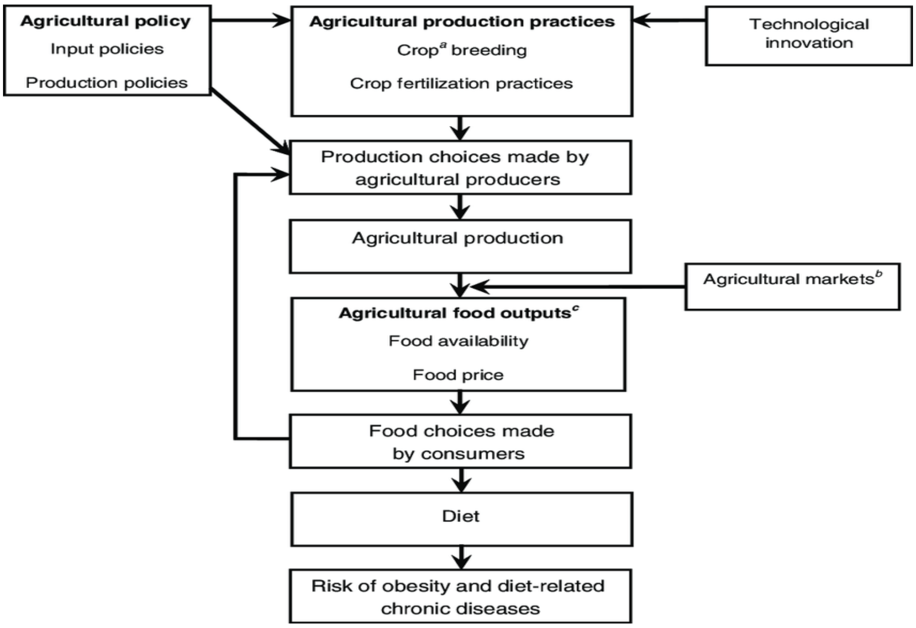

Fig 4: Conceptual framework -the relationship between agricultural policies and production practices and diet (Bryden, 2012).

Operationally, the model produces decision levers that are both agronomically sound and computationally tractable. Lime rates are computed to move pH into crop-specific target ranges while considering buffer capacity; N is split into applications aligned with crop uptake curves and rainfall probabilities to minimize losses; P sources and placements are chosen for low fixation contexts; K is targeted to zones with low exchangeable K and high clay fixation; micronutrients are applied where soil tests and deficiency signatures indicate probability of response above a threshold. Organic amendments are valued for both carbon and nutrient contributions, with mineral N adjusted for mineralization potential inferred from SOC and temperature–moisture regimes. Residue retention and cover crops are considered for their medium-term effects on SOC and nutrient cycling, providing a path to compound efficiency gains across seasons (Luo, *et al*., 2011, Robertson, *et al*., 2018).

To connect back to efficiency metrics, the framework proposes a two-stage analytical map. First, estimate technical efficiency using data envelopment analysis or stochastic frontier analysis across comparable production environments, controlling for inputs and weather. Second, regress efficiency scores or frontier gaps on soil indicators and their interactions, instrumenting for measurement error and confounding where possible. Elasticities derived from these regressions represent how much efficiency improves per unit change in soil parameter, providing the bridge from measurement to prescription. Where panel data exist, fixed effects remove time-invariant unobservables; where interventions occur (e.g., lime rollout), difference-in-differences or synthetic controls identify causal impacts (Nolan, *et al*., 2018, Sharma, *et al*., 2012). These estimates seed an optimization that chooses soil interventions and fertilizer bundles to maximize expected profit per hectare subject to resource, environmental, and risk constraints.

Finally, the conceptual foundations emphasize transparency and learning. Indicators are chosen for interpretability; uncertainty is propagated from lab precision and spatial interpolation through to elasticities and recommendations; and outcomes are monitored to update priors, improving both causal attribution and prescription quality over time. By embedding the soil–plant–atmosphere continuum into efficiency theory and by specifying clear mechanisms through which soil parameters shift frontiers and marginal products, the framework provides a disciplined basis for closing yield and efficiency gaps. It reconciles the agronomic need to correct binding constraints with the economic imperative to allocate scarce resources wisely and the environmental mandate to reduce nutrient losses, turning heterogeneous measurements into actionable, risk-calibrated soil-informed decisions.

2.2. Data Architecture and Harmonization

A robust data architecture for linking soil fertility parameters with agricultural output efficiency begins by enumerating, standardizing, and harmonizing the heterogeneous data sources that collectively determine both soil condition and production outcomes. Field surveys provide geo-referenced plot boundaries, management histories, crop rotations, residue practices, tillage intensity, and in-season operations such as planting dates, seed density, fertilizer blends and timings, liming, irrigation, and pest control. Surveys also capture farmer constraints and objectives that influence input choices (Cowie, *et al*., 2011, Lal, 2019). Laboratory data contribute definitive measurements of soil pH, soil organic carbon, total nitrogen, mineral N (nitrate, ammonium), available phosphorus (Bray or Olsen), exchangeable potassium, micronutrients (Zn, Fe, B, Mn, Cu), cation exchange capacity, base saturation, bulk density, texture, electrical conductivity, and relevant contaminants or salinity metrics. Proximal sensing augments lab depth with dense in-field readings from portable spectrometers (VIS–NIR/MIR), electromagnetic induction for apparent conductivity, penetrometers for compaction, and in-situ moisture/temperature probes. Remote sensing supplies spatial and temporal context: multispectral indices (NDVI, EVI, red-edge chlorophyll proxies), thermal inertia for crop stress, SAR backscatter for moisture and roughness, topographic derivatives (slope, curvature), and land surface phenology indicators that track growth stages. Weather completes the biophysical drivers with gridded reanalyses or station data for rainfall, temperature extremes, solar radiation, vapor pressure deficit, and wind; downscaled forecasts enable in-season decisions. Market data close the loop with input prices (fertilizer, lime, seed, fuel), output prices (farm-gate crop prices), logistics costs, and policy incentives or penalties affecting nutrient use and emissions. Together, these sources define the measurement, exposure, and economic layers required to attribute efficiency changes to soil parameters and to design prescriptions that are both agronomically and financially sound (Eyles, *et al*., 2015, Sokouti, Kaveh & Parvizi, 2017).

Repository design translates this diversity into a coherent, queryable system. The core is a lake–warehouse hybrid: raw assets land in a versioned object store with immutable, time-stamped partitions by source and ingestion date; curated, analysis-ready tables live in a columnar warehouse optimized for spatial and temporal joins. Every dataset carries rich metadata compliant with established schemas: ISO 19115 for geospatial description, OGC SensorML for streaming sensors, and agro-domain dictionaries for soil and crop variables (units, methods, detection limits). A canonical entity-relationship model anchors the warehouse: entities include Farm, Field, Plot, Sample, Observation, Sensor, ManagementEvent, WeatherTile, MarketQuote, and OutputRecord (yield, quality) (Giller, *et al*., 2011, Hoang, 2013). Keys are explicit global field identifiers, sample IDs, sensor IDs and all measurements carry time stamps, depths, and confidence flags. A feature registry documents derived indicators (e.g., base saturation, C:N:P ratio, acidity index, available water capacity), their formulas, inputs, and lineage. Governance policies require that transformations are expressed as version-controlled code with unit tests and data-quality checks so that any analysis is reproducible and auditable.

Quality assurance and control operate at multiple stages. On ingestion, schema validation enforces types and units; range checks catch physically implausible values (negative CEC, pH outside 2–10, nitrate beyond method limits unless flagged as extract artifacts). Cross-variable rules implement mass-balance and stoichiometric plausibility (e.g., base saturation cannot exceed 100% unless accompanied by methodological notes; SOC and bulk density jointly define plausible soil organic carbon stocks per depth). Duplicate detection merges repeated lab records and reconciles barcodes to sample IDs. Spatial validity tests ensure plot polygons are non-self-intersecting, fall within expected farm boundaries, and respect topological rules when combining field blocks. For time series, spike filters and robust smoothing remove electrical noise without erasing agronomically meaningful events (Möller, 2018, Therond, *et al*., 2017). Proximal and remote-sensing signals undergo calibration against lab references using domain-constrained regressions that enforce monotonicity where appropriate (e.g., higher MIR aromatic peaks correlate with higher SOC) and quantify prediction intervals. Sensor calibration schedules, drift diagnostics, and maintenance logs are stored as first-class data to contextualize anomalies.

Missing-data strategy is explicit and tiered by decision criticality. For lab variables, hierarchical imputation leverages pedotransfer functions conditioned on texture, SOC, pH, and landscape position, with uncertainty recorded as variance parameters. For proximal sensing grids with partial coverage, gap-filling uses geostatistical kriging with external drift supplied by covariates (DEM derivatives, satellite indices). Weather gaps are bridged via distance-weighted station blending or bias-corrected reanalysis, and uncertainty envelopes reflect station density and terrain complexity. Management histories with partial recall are completed through rule-based defaults based on local agronomic calendars and cooperative records, but these imputations are flagged and down-weighted in causal inference. Crucially, imputation never overwrites raw data; it generates alternative, labelled columns with provenance so analysts can select observed-only or imputed-enhanced views and propagate uncertainty into models (Kiryushin, 2019, Vanlauwe, *et al*., 2011).

Spatial referencing underpins credible linkage between soil and output. All geometries are stored in a common coordinate reference system appropriate to the region's scale (e.g., UTM zone) with explicit CRS tags to prevent silent reprojection errors. Plot boundaries are captured at a resolution that preserves management-relevant features (headlands, waterways), and positional accuracy is recorded (GPS quality, correction method). A tiling scheme (e.g., Web Mercator tiles or custom equal-area grids) accelerates joins between fields and gridded layers like weather and satellite products. Depth is treated as a spatial dimension for soil: each sample carries upper and lower depth bounds, enabling calculation of depth-integrated stocks and depth-specific responses. Buffering logic handles edge effects when aggregating pixel-level remote-sensing signals to plots, avoiding contamination from adjacent fields. Where yield monitors exist, raw swath data are cleaned for header effects, delays, and speed artifacts, then binned to micro-plots that respect combine width and GNSS accuracy; in their absence, crop-cut quadrats or weighbridge totals are reconciled to plot areas (Erb, *et al*., 2013, Mueller, *et al*., 2011).

Farm-to-landscape aggregation reconciles heterogeneity with statistical power. At farm scale, the repository supports within-field zoning based on clustering of soil and topographic covariates, enabling variable-rate prescriptions and paired-plot experimentation. At landscape scale, fields are grouped into production environments defined by agroecological zones, rainfall regimes, soil types, and market access metrics, allowing fair benchmarking and the estimation of environment-specific response functions. Multi-scale indices for example, a water-limited nutrient index combining field-level available water capacity with seasonal rainfall anomalies are computed to explain interaction effects between soil and climate. Aggregation workflows preserve uncertainty: when averaging soil parameters or model elasticities across plots, weights reflect area, measurement confidence, and relevance to the target crop; error bars shrink appropriately, and analysts can test sensitivity to alternative weighting schemes. Spatial cross-validation folds are constructed by environment rather than random point sampling to avoid overly optimistic generalization (Fairhurst, 2012, Zingore, *et al*., 2011).

Interoperability and timeliness are essential for decision support. The architecture exposes data through standardized APIs (OGC WFS/WMS for spatial layers, REST/GraphQL for tabular endpoints) with role-based access control to protect sensitive farm information. Event-driven pipelines refresh in-season layers weather, remote sensing, sensor streams on schedules matched to agronomic decisions (weekly fertilization windows, irrigation scheduling, top-dressing timing). A feature store caches harmonized, versioned predictors for modeling workflows, while a model registry stores trained models with metadata on training data windows, features, and performance metrics under different environments. Dashboards consume these services to display soil constraints, predicted efficiency gaps, and prescription candidates with uncertainty and expected profit impacts (Akomea-Agyin & Asante, 2019, Awe, 2017, Osabuohien, 2019).

Harmonization culminates in coherent, analysis-ready datasets where rows represent plot–season observations tied to soils, management, weather, markets, and outputs. Each row includes raw and derived soil features, management inputs with timing, weather summaries (growing degree days, rainfall distribution metrics, VPD integrals), remote-sensing phenology descriptors, and price vectors (Awe & Akpan, 2017). Outcome variables include yields, quality grades where relevant, and efficiency metrics derived from frontier methods. Critically, every feature carries provenance and an uncertainty score so downstream models can apply measurement-error corrections, attenuate biased coefficients, or choose robust estimators. Causal identification benefits from this structure: instruments (e.g., rainfall anomalies for N response, legacy liming programs for pH shifts) and fixed effects (field or farmer) are readily merged, and panel structures emerge naturally for farms with repeated seasons.

The architecture is designed to learn. As prescriptions are delivered and outcomes observed, the repository captures treatment assignments, farmer uptake, and realized effects, enabling uplift modeling and refinement of response surfaces. Laboratory calibration sets grow with new regions; proximal-to-lab transfer functions are periodically re-estimated; satellite sensor changes are handled via cross-sensor harmonization. Data-quality dashboards reveal failing sensors, outlier labs, or inconsistent surveyors so corrective action is targeted. Governance committees comprising agronomists, data engineers, and economists review metadata standards, QA/QC thresholds, and missing-data policies quarterly, ensuring practices evolve with evidence (Akpan, Awe & Idowu, 2019, Ogundipe, *et al*., 2019).

In this way, data architecture and harmonization are not back-office chores but the foundation for credible, soil-informed efficiency gains. By knitting together field reality, laboratory rigor, sensor density, environmental drivers, and market context into an auditable, uncertainty-aware repository, the framework equips analysts to separate correlation from causation and equips decision makers to implement prescriptions that are both profitable and sustainable at farm scale and scalable across landscapes.

2.3. Measurement, Feature Engineering, and Spatial Modeling

Measurement, feature engineering, and spatial modeling form the analytical core of any framework that links soil fertility parameters to agricultural output efficiency. Accurate, harmonized, and spatially explicit quantification of soil properties underpins reliable causal inference and predictive modeling. The core variables include soil pH, soil organic carbon (SOC), total nitrogen (TN), available phosphorus (P) and potassium (K), micronutrients (zinc, iron, copper, manganese, and boron), cation exchange capacity (CEC), texture, and soil moisture. Each variable plays a distinct role in nutrient cycling, plant availability, and crop productivity, and their interactions define the effective fertility status of soils (Awe, Akpan & Adekoya, 2017, Osabuohien, 2017).

Soil pH is the master variable governing the solubility and availability of nutrients and the activity of microorganisms. Slight shifts in pH alter the ionization state of essential nutrients phosphorus availability peaks between 6.0 and 7.0, while micronutrients such as iron and zinc become less available in alkaline conditions. pH also affects microbial-mediated processes like nitrification and denitrification, influencing nitrogen use efficiency and emissions intensity. SOC represents both a nutrient source and a structural determinant of fertility: it enhances aggregation, increases water holding capacity, buffers pH, and contributes to CEC. Measuring SOC through dry combustion or mid-infrared spectroscopy provides an index of the soil's capacity to sustain productivity and resilience (Patrick, *et al*., 2019). Total nitrogen complements SOC as a measure of fertility balance, reflecting mineralization potential and long-term nutrient reserves. Available P and K, determined by chemical extraction methods (Bray-1, Olsen, or ammonium acetate), indicate the immediate pools accessible to plants, which are highly sensitive to management, soil mineralogy, and pH.

Micronutrients, though required in small quantities, exert disproportionate influence on yield and nutrient efficiency. Deficiencies in zinc, boron, or iron can constrain growth even when macronutrients are adequate. These elements are quantified using DTPA or hot-water extraction, and their interpretation depends on texture, organic matter, and pH. Cation exchange capacity integrates both the mineral and organic fractions' ability to retain cations (Ca^{2+}, Mg^{2+}, K^+, Na^+) and buffer against leaching losses, serving as a synthetic measure of fertility sustainability. Texture, determined by hydrometer or laser diffraction, classifies the proportions of sand, silt, and clay and thus the soil's capacity for water retention, aeration, and nutrient adsorption (Bankole, *et al*., 2019, Nwokediegwu, Bankole & Okiye, 2019). Soil moisture, captured through probes, time-domain reflectometry (TDR), or satellite-derived indices (e.g., SMAP or Sentinel-1 backscatter), directly regulates nutrient diffusion, microbial activity, and plant uptake. Collectively, these core variables form the empirical foundation for modeling how soil conditions mediate output efficiency.

Feature engineering transforms these raw variables into composite indicators that capture the underlying biogeochemical constraints more effectively. One central derived feature is the carbon–nitrogen–phosphorus (C:N:P) stoichiometric ratio, which reflects the balance of energy and nutrient availability for microbes and plants. Soils with narrow C:N ratios mineralize nitrogen rapidly, improving short-term fertility but risking losses, while wide ratios indicate immobilization and slower nutrient release. The C:P and N:P ratios provide diagnostic insight into phosphorus limitation and the likely responsiveness to P fertilization. Base saturation the fraction of exchange sites occupied by Ca^{2+}, Mg^{2+}, K^+, and Na^+ relative to total CEC acts as a direct indicator of liming need and cation balance; low base saturation implies acidity stress and potential aluminum toxicity. Acidity or alkalinity indices derived from pH, exchangeable acidity, and aluminum saturation quantify stress severity and guide lime or gypsum prescriptions (Atobatele, *et al*., 2019, Filani, Nwokocha & Babatunde, 2019).

Water-related indices integrate physical and climatic properties: available water capacity combines field capacity, wilting point, and bulk density to estimate plant-accessible water reserves per unit depth; the soil moisture deficit index compares actual to optimal moisture during crop-critical growth stages, providing a time-integrated measure of water limitation. Other engineered features include structural stability index (ratio of SOC to clay) that signals susceptibility to compaction or crusting, and nutrient buffer capacity indices that quantify how strongly a soil resists changes in nutrient concentration following fertilization. These derived variables are not arbitrary composites but mechanistic constructs linking soil attributes to plant response functions, making them particularly valuable for causal modeling and optimization.

Standardization and scaling of features ensure comparability across sites and seasons. pH and other variables with known nonlinear response functions are transformed into agronomically meaningful scales such as distance from crop-specific optimal pH to capture diminishing returns near optimum conditions. SOC and nutrient contents are often log-transformed to normalize skewed distributions, while ratios and indices are standardized within agroecological zones to remove climatic bias (Aduwo & Nwachukwu, 2019, Erigha, *et al*., 2019). Principal component analysis or factor analysis can then condense correlated fertility variables into latent dimensions representing fertility potential, chemical balance, and physical condition, providing parsimonious inputs for efficiency or frontier models. Care is taken to preserve interpretability; feature importance and directionality are always traceable to underlying measurable variables.

Spatial modeling bridges measurement gaps and extends field observations across landscapes. Even intensive sampling cannot cover every hectare, so geostatistics and digital soil mapping are employed to predict soil parameters at unobserved locations. Classical kriging uses spatial autocorrelation captured by variograms to interpolate point measurements; the fitted variogram quantifies how similarity decays with distance, capturing spatial structure driven by geology, land use, and topography. For many fertility variables pH, SOC, and available P ordinary kriging performs well where sampling density is sufficient and variation is moderately stationary. However, in heterogeneous terrains or data-sparse regions, covariate-driven models offer stronger predictive power (Bankole & Tewogbade, 2019, Fasasi, *et al*., 2019). Regression kriging combines deterministic relationships with auxiliary covariates (such as elevation, slope, NDVI, or reflectance bands) and residual kriging to exploit both environmental gradients and local spatial structure. Random forest and gradient boosting models, calibrated with these same covariates, are widely used in digital soil mapping to handle nonlinearities and interactions without assuming stationarity.

Covariate selection follows pedological reasoning: topography affects moisture and erosion, influencing SOC and nutrient accumulation; spectral indices relate to organic matter, clay, and iron oxide content; climate variables define weathering and leaching intensity. Models are trained on harmonized datasets where each soil sample is linked to co-located covariates, and cross-validation is conducted using spatially disjoint folds to prevent overfitting due to spatial autocorrelation. Predictions are generated on a regular grid (e.g., 30–100 m resolution), with uncertainty surfaces estimated via kriging variance or ensemble spread. These uncertainty maps are critical for weighting observations in downstream analyses and for guiding additional sampling to reduce uncertainty where it most constrains decision-making (Atobatele, Hungbo & Adeyemi, 2019, Hungbo & Adeyemi, 2019).

Temporal modeling complements spatial prediction. Fertility variables like SOC and pH evolve slowly, while mineral N and moisture vary dynamically. Sequential kriging or spatiotemporal Gaussian processes can integrate time as a third dimension, capturing persistence and seasonal cycles. For proximal sensors and satellite data, time-series smoothing and harmonic analysis extract trend and anomaly components that reveal fertility changes under different managements. This temporal dimension allows the framework to distinguish transient nutrient fluctuations from structural soil improvements or degradation.

Integration across scales is achieved through multi-source fusion. Lab and proximal data provide high-accuracy anchor points, while remote-sensing and terrain covariates extend these across larger areas. Bayesian hierarchical models unify these sources by treating high-resolution lab measurements as ground truth with known uncertainty and allowing lower-resolution sources to inform prior distributions for unsampled areas. This approach preserves consistency between measurement scales and propagates uncertainty transparently. Data fusion also supports near-real-time monitoring: as new sensor or satellite data arrive, predictions update dynamically, providing current estimates of key fertility indicators for adaptive management (Atobatele, Hungbo & Adeyemi, 2019, Hungbo & Adeyemi, 2019).

Ultimately, spatial modeling outputs feed directly into the efficiency analysis. Each grid cell or plot is characterized by a vector of standardized fertility features and derived indices with associated uncertainty. These feed production-frontier and causal models that estimate how deviations from optimal fertility states affect yield and input efficiency. For example, stochastic frontier models can include predicted SOC and available P as explanatory variables while incorporating their prediction variance as measurement error terms, reducing bias. At larger scales, aggregated fertility indices inform policy models assessing nutrient-use efficiency and potential gains from targeted interventions.

The credibility of these analyses hinges on transparent validation. Independent validation points, typically 20–30% of samples held out by region, are used to compute R^2, RMSE, and bias for each predicted variable. Validation is stratified by soil type and land use to ensure broad performance. When digital soil maps are integrated with crop-yield data, spatial cross-correlation checks confirm that patterns of predicted fertility align with observed productivity gradients. Where discrepancies arise, they signal either missing covariates or confounding management effects, guiding refinement (Filani, Nwokocha & Babatunde, 2019, Kamau, 2018).

In essence, measurement, feature engineering, and spatial modeling together create the quantitative backbone for linking soil conditions to agricultural output efficiency. Precise measurements capture the state of fertility; engineered features translate complex interactions into agronomically interpretable metrics; and spatial models fill gaps and generalize across landscapes. The result is a harmonized, uncertainty-aware dataset that faithfully represents how pH, SOC, nutrients, texture, and moisture vary in space and time and how these variations shape yield potential, input efficiency, and environmental performance. By treating soil fertility not as a set of isolated variables but as an interconnected, spatially structured system, the analytical framework enables soil-informed management that is both locally optimized and scalable, providing the empirical foundation for closing yield and efficiency gaps sustainably.

2.4. Efficiency Modeling and Performance Indicators

Efficiency modeling and performance indicators translate soil fertility information into quantitative measures of how effectively farms transform resources into output, revealing where and how soil conditions constrain performance. Within this analytical framework, two principal frontier methods Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) form the backbone of efficiency estimation. Both methods evaluate the distance between observed farm performance and a theoretical production frontier that represents best practice given available technology and resource endowments. DEA, a nonparametric method, constructs the frontier through linear programming envelopes that encompass observed input–output combinations. SFA, a parametric method, fits an explicit functional form (such as Cobb–Douglas or Translog) that separates inefficiency from random noise (Ayanbode, *et al*., 2019, Onalaja, *et al*., 2019). These complementary techniques jointly provide robustness: DEA offers flexibility with minimal assumptions about functional form, while SFA provides statistical inference and the ability to handle measurement error common in agricultural datasets.

In the DEA formulation, each farm or plot is treated as a decision-making unit (DMU) producing output (e.g., crop yield or revenue) from a set of inputs (land, labor, fertilizer nutrients, water, and capital). The basic model seeks to maximize the ratio of weighted outputs to weighted inputs, constrained such that no DMU's ratio exceeds one when applying the same weights. The output-oriented DEA variant identifies how much output could be proportionally expanded without increasing inputs, yielding a technical efficiency score between 0 and 1. In an input-oriented version, the model identifies how much inputs could be proportionally reduced while holding output constant, also producing efficiency scores (Seyi-Lande, Oziri & Arowogbadamu, 2019). In soil-linked efficiency studies, inputs typically include land area, quantities of nitrogen, phosphorus, and potassium fertilizers (and occasionally lime, organic amendments, seed, labor, and water), while outputs are crop yield (tonnes per hectare) or gross value of output. Incorporating soil fertility variables as contextual or non-discretionary inputs allows the frontier to reflect environmental potential rather than penalizing farms for unchangeable constraints. A two-stage DEA can then regress efficiency scores on soil indicators (pH, SOC, available P, texture) to estimate their marginal influence on efficiency gaps.

SFA introduces a stochastic production function of the form:

$$yi=f(xi;\beta)\exp(vi-ui)$$

where (y_i) is output, (x_i) is a vector of inputs, (\beta) are parameters to be estimated, (v_i) represents random shocks (weather, measurement error), and (u_i) represents inefficiency. The inefficiency term is assumed to follow a one-sided distribution (e.g., half-normal or truncated normal), ensuring non-negative inefficiency, while (v_i) follows a symmetric normal distribution. The frontier function (f(x_i; \beta)) can be specified as Cobb–Douglas for simplicity or Translog to allow flexible substitution elasticities among inputs. Soil fertility variables can be incorporated either directly in (x_i) (capturing their productive contribution) or in an inefficiency effects model where inefficiency (u_i = z_i\delta + w_i) depends on soil and management characteristics (z_i). For instance, a model may find that lower SOC or suboptimal pH significantly increase inefficiency, while balanced NPK ratios and adequate base saturation reduce it. SFA thus links agronomic constraints to efficiency gaps with statistical rigor, enabling hypothesis testing on soil effects.

Performance indicators derived from these frontier analyses provide multidimensional insights into productivity and sustainability. The first and most fundamental is the technical efficiency (TE) score the ratio of observed output to potential output under the same inputs and technology. A TE of 0.85 means the farm could increase output by 15% with existing resources if operating at the frontier. When soil parameters enter the inefficiency model, improvements in pH balance, SOC, or available P can be quantified in terms of expected gains in TE. Partial factor productivity (PFP) complements frontier scores by measuring output per unit of a single input: nitrogen use efficiency (kg grain per kg N), phosphorus use efficiency, and potassium use efficiency provide interpretable, fertilizer-specific metrics. PFPs reveal nutrient imbalances: high N-PFP but low P-PFP may indicate phosphorus limitation, while the reverse suggests excess nitrogen use. When contextualized with soil fertility maps, PFP diagnostics guide spatially differentiated recommendations raising P rates where available P is below critical levels or reducing N where returns diminish (Akinrinoye, *et al*. 2019, Didi, Abass & Balogun, 2019, Otokiti & Akorede, 2018).

Profitability indicators link agronomic efficiency to economic outcomes. Gross margin or profit per hectare incorporates input costs and output prices, converting technical gains into monetary terms. For example, if improved pH and SOC raise nitrogen efficiency, fertilizer cost per tonne of grain falls, increasing profit margins even before accounting for potential yield gains. Profit/ha becomes a unifying key performance indicator (KPI) that aligns farmer incentives with resource-use efficiency. Another essential KPI is production stability, measured as the coefficient of variation of yield or profit across seasons. Stable performance reflects resilience soils with higher organic carbon and better moisture retention often exhibit lower interannual yield variability. Including stability alongside efficiency ensures that interventions favor both productivity and risk reduction (Akinbola & Otokiti, 2012, Dako, *et al*., 2019, Oziri, Seyi-Lande & Arowogbadamu, 2019).

Aggregating these metrics yields a comprehensive dashboard: (1) technical efficiency (DEA/SFA score), (2) nutrient-specific PFPs, (3) profit per hectare, and (4) yield or profit stability index. Supporting indicators include eco-efficiency output per unit of nutrient surplus or per unit of greenhouse gas emissions and total factor productivity (TFP), which integrates technological change over time. Each indicator is tracked spatially and temporally, enabling comparisons among farms, production environments, and management regimes. When combined with uncertainty intervals from the underlying soil and yield models, this dashboard provides decision-makers with confidence-weighted insights.

Handling heterogeneity is central to ensuring that efficiency estimates reflect true performance differences rather than environmental disparities. Agricultural landscapes vary widely in rainfall, temperature, soils, and infrastructure;

without correction, farms in less favorable zones may appear inefficient merely because of biophysical constraints. The framework addresses this through stratification and hierarchical modeling. First, farms are grouped into production environments defined by combinations of climate (rainfall, temperature), soil type (texture, depth, fertility class), and topography (Seyi-Lande, Oziri & Arowogbadamu, 2018). Separate frontiers are estimated for each environment, allowing technology and response functions to differ. Alternatively, a meta-frontier approach is applied: environment-specific frontiers are estimated first, and a global frontier enveloping all environments is derived; the ratio of environment-specific to meta-frontier efficiency (the technology gap ratio) quantifies how far each zone lags the global best practice due to environmental or technological limitations.

Seasonal heterogeneity is addressed by including time dummies or seasonal-specific models. Efficiency can fluctuate with weather anomalies excess rainfall may lower fertilizer-use efficiency in one year but not another. By estimating separate frontiers for multiple seasons, the framework captures temporal dynamics and identifies whether soil-improvement investments (like liming or organic amendments) stabilize efficiency across years. The inclusion of panel data enhances robustness: fixed-effects SFA models control for unobserved, time-invariant farm characteristics such as managerial skill or micro-topography, isolating the contribution of soil changes and management adaptation.

Cultivar heterogeneity also matters. Different crop varieties exhibit distinct nutrient uptake efficiencies, root architecture, and stress tolerance. Incorporating cultivar dummies or interaction terms between soil parameters and cultivar type allows efficiency modeling to recognize genotype–environment–management interactions. For example, a hybrid maize variety may respond sharply to phosphorus correction, whereas an open-pollinated variety's yield potential is more limited by nitrogen. DEA and SFA frameworks can both accommodate such heterogeneity through separate frontiers or interaction variables, ensuring that recommendations remain cultivar-specific (Ajonbadi, *et al*., 2014, Didi, Balogun & Abass, 2019, Farounbi, *et al*., 2019).

At higher analytical levels, the framework integrates soil fertility–based efficiency diagnostics into sustainability and policy assessment. Spatial aggregation of efficiency scores and nutrient productivity metrics produces maps of nutrient-use efficiency and profit potential, revealing zones of nutrient mining or excessive application. These maps feed into regional nutrient balance analyses, informing fertilizer policy and extension targeting. Over time, tracking changes in average efficiency and nutrient productivity under soil management programs (like integrated soil fertility management or conservation agriculture) indicates progress toward closing yield gaps sustainably. Incorporating uncertainty from soil and yield models prevents overconfidence: efficiency improvements are reported with confidence intervals, enabling risk-based policy decisions (Ajonbadi, Mojeed-Sanni & Otokiti, 2015, Evans-Uzosike & Okatta, 2019, Oguntegbe, Farounbi & Okafor, 2019).

Ultimately, efficiency modeling serves both diagnosis and prescription. Frontier methods diagnose the magnitude and sources of inefficiency, while KPIs translate these insights into actionable targets raising technical efficiency from 0.75 to 0.9, improving N-PFP by 15%, or stabilizing yields by reducing coefficient of variation below 20%. When combined with the soil data architecture and spatial modeling layers, these indicators can be visualized as farm-level dashboards or regional decision maps. Farmers and policymakers can then prioritize interventions that offer the greatest efficiency and profitability gains per unit cost, while minimizing nutrient losses and environmental impact (AdeniyiAjonbadi, *et al*., 2015, Didi, Abass & Balogun, 2019, Umoren, *et al*., 2019).

In summary, the efficiency modeling and performance indicator layer transforms raw soil and management data into interpretable metrics of how effectively resources are converted into outputs. DEA and SFA reveal how soil parameters shift production frontiers and inefficiency distributions; partial factor productivities and profit per hectare translate these shifts into economic and nutrient-use terms; and stability metrics incorporate resilience. Through stratified, environment-aware modeling, the framework ensures fairness and relevance across heterogeneous agroecosystems. This integrated efficiency evaluation provides a quantitative foundation for precision soil management, enabling stakeholders to allocate resources intelligently, design evidence-based policies, and close both yield and efficiency gaps in a sustainable, data-driven manner.

2.5. Causal Identification and Model Interpretability

Causal identification and interpretability are the pillars that transform statistical association into credible inference within the analytical framework linking soil fertility parameters to agricultural output efficiency. The goal is to establish how and to what extent variations in soil conditions such as pH, soil organic carbon (SOC), available phosphorus, potassium, and micronutrients cause changes in productivity, input efficiency, and profitability, rather than merely correlate with them. Because agriculture operates within a dynamic, spatially dependent, and policy-influenced environment, the framework integrates multiple econometric and machine-learning strategies to separate causal effects from confounding and to make model results transparent and actionable (Ajayi, *et al*., 2019, Bayeroju, *et al*., 2019, Sanusi, *et al*., 2019).

Causal reasoning begins with the construction of Directed Acyclic Graphs (DAGs), which map hypothesized relationships among soil parameters, management practices, weather, socio-economic factors, and outputs. A typical DAG positions soil fertility variable as intermediate inputs influenced by natural factors (parent material, rainfall, topography) and human actions (fertilization, liming, organic matter management). Crop yield and efficiency are descendants of these nodes, while unobserved factors such as managerial skill or pest pressure may confound relationships. The DAG formalism clarifies which variables must be controlled or instrumented to isolate soil effects. For instance, rainfall affects both SOC dynamics and yield directly; hence, without proper adjustment, the estimated effect of SOC on yield would be biased. By encoding these dependencies, the DAG guides the empirical strategy indicating where fixed effects, instrumental variables, or difference-in-differences (DiD) designs are required.

Panel fixed-effects (FE) models exploit repeated observations of the same fields or farms over time to control for unobserved, time-invariant heterogeneity soil type, farmer ability, or micro-topography that could confound cross-sectional comparisons. The specification typically takes the form:

$$Yit=\alpha i+\lambda t+\beta Xit+\epsilon it$$

where (Y_{it}) is yield or efficiency, (X_{it}) represents time-varying soil fertility and management variables, (\alpha_i) captures farm-specific effects, and (\lambda_t) absorbs common shocks such as policy or market changes. This approach effectively differences out static unobservables, attributing residual variation to within-farm changes such as soil improvements or nutrient interventions. For example, if a farmer adopts liming and SOC increases over several years, the FE estimator isolates how yield efficiency responds to this internal change rather than to structural differences across farms. Moreover, dynamic panel methods like System GMM can handle potential endogeneity of fertilizer application, recognizing that farmers adjust input use based on prior yields and soil conditions.

Instrumental variable (IV) strategies complement FE when time-varying endogenous variables remain. Endogeneity arises because input decisions depend on expectations of productivity, which are themselves shaped by unobserved soil conditions or farmer knowledge. Instruments must correlate with the endogenous soil or input variable but not directly affect the outcome except through it. In this framework, rainfall anomalies, historical liming programs, or topographic features serve as plausible instruments. Rainfall shocks influence nutrient leaching, SOC mineralization, and soil moisture but not managerial skill, satisfying relevance and exclusion restrictions (Ajayi, *et al*., 2019, Bukhari, *et al*., 2019, Oguntegbe, Farounbi & Okafor, 2019). Legacy liming intensity captured through historical records or distance to lime plants provides exogenous variation in pH that persists over time and affects nutrient availability independently of current farmer decisions. Using such instruments allows consistent estimation of causal elasticities: for instance, quantifying how a one-unit increase in soil pH or SOC, exogenously induced by prior liming or organic amendments, shifts technical efficiency or nitrogen-use efficiency.

Difference-in-differences (DiD) frameworks identify causal impacts of discrete interventions such as integrated soil fertility management projects, fertilizer subsidies, or conservation agriculture programs by comparing changes over time between treated and control groups. The canonical specification:

$$Yit=\alpha+\delta Di+\gamma Tt+\theta(Di\times Tt)+\epsilon it$$

estimates (\theta) as the treatment effect, where (D_i) identifies treated farms and (T_t) indicates post-intervention periods. Within the soil–efficiency framework, DiD isolates how yield efficiency changes after intervention relative to baseline trends, adjusting for temporal shocks. For example, if a liming campaign began in 2018, comparing efficiency changes between participating and non-participating farms before and after implementation attributes differences to the intervention, assuming parallel trends. Combining DiD with FE enhances robustness, and when interventions occur at staggered times, event-study specifications trace dynamic effects, showing how long it takes for pH correction or SOC buildup to translate into efficiency gains.

Spatial econometric models capture interactions among neighboring fields that violate the independence assumption of standard regressions. Fertilizer diffusion, pest migration, shared irrigation, or localized weather anomalies create spatial spillovers in both inputs and outputs. Two canonical forms are used: the Spatial Autoregressive (SAR) model, which includes a spatially lagged dependent variable ((W Y)) to capture outcome interdependence, and the Spatial Error Model (SEM), where spatial correlation resides in the disturbance term ((W \epsilon)). A SAR specification for yield might be:

$$Y=\rho WY+X\beta+\epsilon$$

where (W) is a spatial weights matrix defining neighborhood structure (e.g., inverse distance or shared boundary). The parameter (\rho) measures how much yield efficiency in one field depends on neighboring fields' performance, representing potential learning or input leakage. SEM models, by contrast, address omitted spatially correlated variables like micro-climate or soil type, ensuring unbiased soil-effect estimates. Spatial Durbin extensions combine both, allowing spillovers in both dependent and explanatory variables (such as fertilizer rates or soil moisture). Through these formulations, the framework can detect whether interventions in one area say, localized liming produce positive externalities in adjacent plots via runoff neutralization or social learning, or whether nutrient leaching causes negative cross-boundary effects. Estimating such spatial multipliers refines the economic evaluation of soil-improvement programs, ensuring that aggregate benefits or costs are accurately captured (Ajayi, *et al*., 2018, Bukhari, *et al*., 2018, Essien, *et al*., 2019).

Machine learning (ML) ensembles extend causal analysis by uncovering complex, nonlinear response surfaces between soil fertility parameters and efficiency metrics while preserving interpretability. Gradient boosting machines, random forests, and extreme gradient boosting (XGBoost) are used to predict yield or efficiency from high-dimensional soil, weather, and management data. While these models prioritize predictive accuracy, their integration with SHAP (Shapley Additive Explanations) values restores interpretability by quantifying each variable's contribution to individual predictions. SHAP values, derived from cooperative game theory, assign importance scores to each feature proportional to its marginal contribution to the model's output averaged across all possible feature coalitions. Within this framework, SHAP reveals which soil parameters most influence efficiency and how their effects vary across conditions (Akinrinoye, *et al*. 2015, Bukhari, *et al*., 2019, Erigha, *et al*., 2019).

For instance, a SHAP summary plot might show that SOC, pH, and available P are the dominant positive contributors to predicted efficiency, while high exchangeable sodium or extreme moisture deficits exert negative effects. Partial dependence and SHAP dependence plots provide localized response diagnostics, illustrating nonlinearities and thresholds such as diminishing returns to SOC beyond 2.5% or steep efficiency drops when pH falls below 5.5. These diagnostics validate agronomic knowledge and inform site-specific recommendations: areas where SHAP interactions show strong SOC–moisture synergy could prioritize organic amendments under dry conditions, while fields where available P dominates might focus on targeted P placement. SHAP clustering also helps identify distinct soil–efficiency regimes, allowing policies tailored to each regime's binding constraint.

Causal and machine-learning elements are linked through hybrid approaches. The framework can use ML to flexibly model the production function within SFA, or to generate control variables in causal inference models. Double Machine Learning (DML) techniques, for example, use random forests to estimate nuisance components (propensity scores and conditional expectations) while preserving valid inference on the causal parameter of interest say, the marginal effect of SOC on efficiency thus merging predictive power with statistical identification. Spatial cross-validation ensures that models generalize beyond sampled areas, preventing overestimation of performance due to spatial autocorrelation.

Interpretability extends to the policy and management levels through visualization. Causal diagrams, efficiency maps, and SHAP-based feature importance are integrated into interactive dashboards that display both estimated effects and uncertainty. Farmers can view how local soil characteristics influence expected efficiency gains, while policymakers see regional marginal returns to soil investments. Combining DAG-based identification logic with explainable AI tools ensures that recommendations remain grounded in theory, not just correlations.

Finally, the framework embeds validation and sensitivity analysis. Placebo tests, falsification checks, and alternative instrument specifications verify causal robustness. Spatial lag and error models are compared using information criteria and residual diagnostics to ensure proper representation of neighborhood effects. For ML models, permutation tests confirm that top SHAP-ranked features maintain influence under resampling. This layered validation builds trust among agronomists, economists, and policymakers that estimated soil–efficiency linkages reflect genuine causal mechanisms.

In synthesis, causal identification and interpretability transform the analytical framework from a predictive tool into a decision engine. DAGs establish the theoretical scaffolding of cause and effect; panel fixed effects, instrumental variables, and DiD empirically anchor those relationships against confounding; spatial econometrics reveals spillovers that shape aggregate efficiency; and machine-learning ensembles enriched with SHAP provide transparent, nonlinear diagnostics that bridge complexity and intuition. Together they produce credible, interpretable, and actionable insights quantifying how specific soil fertility improvements, whether through liming, organic matter enhancement, or balanced fertilization, causally elevate agricultural output efficiency and resilience across heterogeneous landscapes (Akinrinoye, *et al*. 2015, Bukhari, *et al*., 2019, Erigha, *et al*., 2019).

2.6. Decision Analytics, Prescription Design, and Implementation

Decision analytics, prescription design, and implementation form the operational core of the analytical framework for linking soil fertility parameters with agricultural output efficiency. Once causal and efficiency models quantify the marginal effects of soil variables elasticities of yield or profit with respect to soil pH, soil organic carbon (SOC), available phosphorus (P), potassium (K), or micronutrients the challenge shifts to converting those elasticities and efficiency gaps into site-specific management prescriptions that are agronomically sound, economically feasible, environmentally sustainable, and operationally scalable (Kiryushin, 2019, Vanlauwe, *et al*., 2011).

At its essence, prescription design is a translation exercise: it takes estimated relationships between soil characteristics and output efficiency and converts them into actionable input recommendations. For instance, if model elasticities indicate that yield efficiency increases by 3% per 0.1-unit rise in soil pH, then a liming prescription is formulated to deliver that pH correction based on the soil's buffering capacity, depth, and target crop requirements. Similarly, if efficiency analysis identifies a diminishing marginal return to nitrogen beyond a threshold but a strong positive elasticity for phosphorus or potassium, then a balanced NPK blend is recommended. These prescriptions account not only for nutrient levels but also for their interactions pH correction enhances phosphorus availability, while SOC improvements amplify nutrient retention and water efficiency. Thus, prescriptions emerge as integrated nutrient–soil–water management packages rather than isolated fertilizer recommendations (Möller, 2018, Therond, *et al*., 2017).

Elasticity-driven prescriptions are quantified using standard agronomic algorithms combined with economic optimization. The lime requirement (LR), for example, is derived from the relationship between target and current pH, the soil's buffer capacity, and lime purity. SOC enhancement is addressed through organic amendments compost, manure, cover crops, or biochar whose expected carbon increments are calibrated against soil texture and baseline carbon. Balanced NPK recommendations are fine-tuned with the help of critical nutrient thresholds and nutrient response curves derived from field trials. When micronutrient deficiencies are detected say, zinc or boron leaf tissue diagnostics and soil test correlations guide microdosing or foliar applications, ensuring minimal waste and maximum impact (Eyles, *et al*., 2015, Sokouti, Kaveh & Parvizi, 2017). By embedding these computations within a decision-analytics engine, the framework can dynamically translate model outputs into prescriptive nutrient and amendment quantities per hectare, customized by soil type, crop, and economic conditions.

Optimization within the decision-analytics layer balances multiple objectives: maximizing yield or profit per hectare while respecting constraints on cost, water availability, and environmental emissions. The mathematical problem can be represented as a constrained optimization model:

$$\text{xmax}\pi=\text{Pyf(x;s)}-\text{Pxx}$$

subject to

$$\text{C(x)}\leq\text{Cmax},\text{W(x)}\leq\text{Wmax},\text{E(x)}\leq\text{Etarget}$$

where (x) denotes the vector of management actions (fertilizer rates, lime, organic matter additions, irrigation), (s) represents soil parameters, (f(x; s)) is the production function conditioned on soil state, (C(x)) is total cost, (W(x)) is water use, and (E(x)) denotes emissions or nutrient losses. The optimization identifies the input bundle that maximizes expected profit ((\pi)) while maintaining resource and environmental constraints within allowable thresholds. The model incorporates nonlinearity and diminishing returns via response functions estimated in the efficiency analysis stage. Risk is introduced through stochastic parameters for rainfall, price volatility, and yield uncertainty, enabling risk-adjusted recommendations based on expected utility or downside-risk minimization.

A risk-aware targeting strategy is crucial because farmers operate under uncertainty and resource heterogeneity. Decision analytics therefore integrates probabilistic simulations Monte Carlo runs across rainfall and price scenarios to derive not only expected profitability but also conditional value-at-risk (CVaR) measures. These reveal the likelihood of achieving minimum income thresholds, supporting conservative or progressive input strategies based on farmers' risk tolerance. For resource-limited smallholders, prescriptions may emphasize low-cost incremental interventions partial liming, microdosing of deficient nutrients, or organic residue management whereas commercial farms can adopt precision variable-rate applications (Cowie, *et al*., 2011, Lal, 2019). The framework also incorporates environmental penalties for nutrient surpluses and greenhouse gas emissions, ensuring that eco-efficiency complements profitability. Nutrient-use efficiency targets (kg yield per kg nutrient applied) and emission intensities (kg CO_2e per tonne yield) are explicitly included as optimization constraints or secondary objectives, aligning farm-level decisions with broader sustainability goals.

Implementation follows a phased deployment roadmap that translates analytics into field practice. The pilot phase begins with a subset of farms or regions representing diverse production environments variations in rainfall, soil type, and market access to validate predictive accuracy and economic viability. During this stage, high-resolution soil sampling and laboratory analysis are combined with proximal and remote sensing to benchmark soil and crop conditions. Efficiency gaps identified through modeling are addressed through experimental prescriptions, allowing iterative calibration of fertilizer blends, lime rates, and organic amendments. The pilot's success is assessed through metrics such as yield improvement, input-use efficiency, and profitability relative to control plots, with uncertainty quantified through bootstrapped confidence intervals (Nolan, *et al*., 2018, Sharma, *et al*., 2012).

Following pilot validation, the framework advances toward adaptive, data-driven dashboards that integrate satellite-derived vegetation indices NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index) with in-situ sensor data. NDVI and EVI provide temporal monitoring of crop vigor and biomass accumulation, acting as proxies for nutrient status and water stress. Coupling these indices with ground-based soil moisture, EC (electrical conductivity), and pH sensors creates a feedback loop between observation and prescription. For instance, if NDVI anomalies indicate emerging nutrient stress, the dashboard cross-references real-time weather and soil data to diagnose whether the cause is nitrogen deficiency, water shortage, or disease pressure, triggering targeted advisories. Machine-learning algorithms continuously refine these diagnostics using historical patterns of sensor and yield responses (Luo, *et al*., 2011, Robertson, *et al*., 2018).

The adaptive dashboard serves as the primary interface for decision support at both farm and advisory levels. Farmers or extension agents access spatially explicit recommendations displayed as maps or tables showing optimal nutrient applications, expected yield gains, cost implications, and risk scores. The system highlights critical zones where the marginal benefit of intervention is highest, effectively prioritizing limited resources. Integration with mobile platforms allows users to download prescription maps compatible with variable-rate spreaders or to receive simplified recommendations via SMS or voice for low-connectivity areas. For cooperative or regional planners, dashboards aggregate farm-level data to show nutrient balance maps, efficiency trends, and environmental compliance indicators, supporting evidence-based planning and policy evaluation.

In addition to monitoring, the implementation roadmap emphasizes feedback and learning. Each season's outcomes yield, input use, profit, and NDVI trajectories are fed back into the repository to update model parameters and elasticities. Bayesian updating and machine-learning retraining ensure that prescriptions remain adaptive to evolving soil conditions, climate variability, and technological advances. The system tracks adoption rates, response effectiveness, and barriers to implementation, generating insights into behavioral and institutional constraints. Pilot expansion thus evolves into a continuous improvement loop that scales geographically while retaining site-specific precision (Arndt, Pauw & Thurlow, 2016, Ricker-Gilbert & Jayne, 2017).

The framework's scalability depends on interoperability and governance. Open data standards (e.g., ISO 28258 for soil data exchange and OGC-compliant APIs for spatial data) ensure that diverse data sources labs, satellites, sensors feed seamlessly into the analytics pipeline. Institutional partnerships with extension services, cooperatives, and agribusinesses create local stewardship for data collection and farmer engagement. Economic incentives, such as input credit schemes or carbon payments for SOC increases, can be linked to dashboard metrics, making data-driven soil management financially rewarding. Furthermore, by integrating emissions and water-use monitoring, the system aligns with climate-smart agriculture objectives and global sustainability reporting frameworks (Ali, Rahut & Imtiaz, 2019, Nasrin, Bauer & Arman, 2018).

The final stage fleet-scale deployment operates as a fully automated, adaptive management system. Soil and crop data flow continuously into the analytics engine; efficiency models update elasticities in real time; optimization algorithms compute revised prescriptions; and dashboards visualize results for stakeholders. Artificial intelligence assistants interpret complex outputs into natural language advisories, enhancing accessibility. Over time, the system builds a knowledge graph linking soil fertility evolution, management interventions, and economic outcomes, enabling predictive governance at regional scales.

In essence, the decision analytics and implementation layer transforms the analytical framework from a diagnostic tool into a living decision ecosystem. It operationalizes the scientific understanding of soil–efficiency relationships into practical, economically optimized, and environmentally constrained actions. Through elasticity-driven prescriptions, multi-objective optimization, and adaptive deployment using NDVI, EVI, and in-situ sensors, the framework ensures that each hectare receives the right input, at the right time, in the right amount, and for the right reason. The outcome is a closed feedback system where data inform action, action improves efficiency, and efficiency sustains productivity and ecosystem health delivering a scalable pathway toward data-enabled, soil-informed agricultural transformation (Jayne & Rashid, 2013, Minviel & Latruffe, 2017).

3. Conclusion

The analytical framework presented here demonstrates how soil information can be elevated from fragmented measurements to a disciplined decision engine that improves agricultural output efficiency while advancing environmental stewardship. By integrating a robust data architecture, mechanistic and statistical understanding of the soil–plant–atmosphere continuum, spatial modeling, causal identification, and frontier-based efficiency analysis, the framework links what is measurable in soils pH, SOC, macro- and micronutrients, texture, moisture, and CEC to what matters for farmers and food systems technical efficiency, nutrient productivity, profit per hectare, yield stability, and eco-efficiency. Elasticities estimated from credible identification strategies are translated into prescriptions for liming, balanced NPK and micronutrients, and organic inputs under explicit cost, water, and emission constraints. Adaptive dashboards fuse NDVI/EVI signals with in-situ sensors to close the loop between observation and action. The result is a repeatable, auditable pathway to close yield and efficiency gaps, reduce nutrient losses and greenhouse gas intensity, and allocate scarce capital to the highest-return interventions hectare by hectare, and season by season.

Realizing this value at scale requires governance that treats data quality, transparency, and rights as first-class design criteria. The framework adopts standards for geospatial and agronomic data (e.g., ISO 19115 for metadata, ISO 28258 for soil data exchange, OGC WFS/WMS for spatial services) and codifies FAIR principles findability, accessibility, interoperability, and reusability so that datasets, features, models, and prescriptions are discoverable and reusable with clear provenance. Uncertainty audits are institutionalized: laboratory precision, sensor calibration drift, imputation variance, and prediction intervals from digital soil maps are quantified and propagated through causal and efficiency models to yield confidence-weighted KPIs and prescriptions. Reproducibility is enforced by versioned data pipelines, a feature registry with lineage, and a model registry that stores training windows, hyperparameters, performance by production environment, and deprecation criteria; every prescription references a specific model/version and the evidence behind it. To protect data subjects and encourage participation, granular consent and role-based access control are mandatory, with clear articulation of data ownership and benefit sharing. Farmer-level data remain locally governed; federated learning and privacy-preserving analytics enable model improvement without centralizing sensitive records. Governance bodies comprising agronomists, economists, data stewards, and farmer representatives set thresholds for QA/QC, approve model changes through change-control reviews, and commission external audits that test for bias, leakage, and unintended distributional effects.

Notwithstanding these strengths, limitations remain. Much of the empirical signal in agriculture is observational, and even with DAGs, fixed effects, instrumental variables, spatial econometrics, and DiD designs, residual confounding can persist especially where management quality, pest pressure, or informal knowledge are imperfectly observed. Soil measurements are noisy and temporally uneven: SOC and pH evolve slowly while mineral N and moisture swing quickly; mismatched sampling cadences can blur causal timing. Digital soil maps and proximal sensing, while powerful, can embed spectroscopic or transfer-function biases across soil types and seasons; rigorous spatial cross-validation mitigates but does not eliminate these risks. External validity is another constraint: elasticities and response thresholds learned in one production environment may not port cleanly to others differing in mineralogy, rainfall patterns, or cultivar genetics. Economic optimization layers must grapple with input and output price volatility, credit access, and labor constraints that vary by household and season; purely profit-maximizing recommendations may be infeasible or unacceptable without credit or risk-sharing instruments. Finally, operational scaling faces bottlenecks in extension capacity, digital connectivity, and the availability of calibrated sensors and reliable lab services, particularly for smallholders.

Future work should therefore prioritize longitudinal and adaptive trials that tighten causal attribution and shorten the evidence-to-practice cycle. Multi-year, multi-site experiments that randomize lime, balanced NPK, micronutrients, and organic amendments stratified by soil class and rainfall regime will refine dose–response curves, reveal persistence of effects, and quantify interactions (e.g., pH × P availability × cultivar). Adaptive experiments can exploit the framework's dashboards to embed A/B tests varying timing or placement while ensuring agronomic safety, thereby generating continuous learning at scale. Climate resilience must move from an overlay to a design axis: prescriptions should optimize expected profit subject to climate-stress constraints, explicitly valuing practices that stabilize yields under heat and drought (e.g., SOC-building, water-retention amendments, split N timed to rainfall probabilities). Spatiotemporal models can integrate subseasonal forecasts and weather analogs to adapt in-season recommendations; risk metrics should expand from CVaR on profit to multi-hazard resilience scores that reflect agronomic and market shocks.

Methodologically, two advances are especially promising. First, distributionally robust optimization and double machine learning can better manage nonstationarity and high-dimensional confounding, delivering prescriptions that remain safe under shifting weather and market distributions. Second, federated and transfer learning can accelerate performance in data-sparse regions by borrowing strength from analogous environments while respecting data sovereignty. On the measurement front, harmonizing next-generation sensing hyperspectral satellites, low-cost ion-selective probes, soil DNA for microbial function with established lab baselines will improve attribution of micronutrient and biology-mediated effects. Integrity checks against independent benchmarks (e.g., ring-test labs, reference soils) should be routine.

Scaling the framework also entails institutional innovations. Blended finance and input-credit schemes can tie lending to verifiable, dashboard-tracked soil improvements (e.g., pH correction milestones or SOC gains), aligning incentives across farmers, lenders, and input suppliers. Results-based extension can reward advisors for measured improvements in nutrient-use efficiency and stability, not just input volumes. Policy integration nutrient budgeting, loss caps, or emissions targets can leverage the framework's eco-efficiency indicators to design incentives that are performance-based rather than prescriptive. Lastly, human capacity is the multiplier: training agronomists, data engineers, and extension agents to co-interpret SHAP diagnostics, uncertainty bands, and frontier outputs with farmers will turn analytics into trusted, context-aware action.

In sum, the framework's value lies in its synthesis: it renders soils legible, decisions auditable, and outcomes predictable enough to guide scarce resources toward their highest agronomic and economic returns while reducing environmental externalities. With rigorous governance standards, uncertainty audits, reproducibility, and data rights paired to iterative field learning and climate-aware design, the approach can evolve from promising pilots to durable practice. The path forward is to keep the science honest, the data protected, the models humble, and the prescriptions adaptive so that soil-informed decisions reliably deliver higher efficiency, greater resilience, and more sustainable landscapes at scale.

4. References

1. Abass OS, Balogun O, Didi PU. A predictive analytics framework for optimizing preventive healthcare sales and engagement outcomes. IRE J. 2019;2(11):497-503.
2. AdeniyiAjonbadi H, AboabaMojeed-Sanni B, Otokiti BO. Sustaining competitive advantage in medium-sized enterprises (MEs) through employee social interaction and helping behaviours. J Small Bus Entrep Dev. 2015;3(2):1-16.
3. Aduwo MO, Nwachukwu PS. Dynamic capital structure optimization in volatile markets: a simulation-based approach to balancing debt and equity under uncertainty. IRE J. 2019;3(2):783-92.
4. Agegnehu G, Srivastava AK, Bird MI. The role of biochar and biochar-compost in improving soil quality and crop performance: a review. Appl Soil Ecol. 2017;119:156-70.
5. Ajayi JO, Bukhari TT, Oladimeji O, Etim ED. A conceptual framework for designing resilient multi-cloud networks ensuring security, scalability, and reliability across infrastructures. IRE J. 2018;1(8):164-73.
6. Ajayi JO, Bukhari TT, Oladimeji O, Etim ED. Toward zero-trust networking: a holistic paradigm shift for enterprise security in digital transformation landscapes. IRE J. 2019;3(2):822-31.
7. Ajayi JO, Bukhari TT, Oladimeji O, Etim ED. A predictive HR analytics model integrating computing and data science to optimize workforce productivity globally. IRE J. 2019;3(4):444-53.
8. Ajonbadi HA, Mojeed-Sanni BA, Otokiti BO. Sustaining competitive advantage in medium-sized enterprises (MEs) through employee social interaction and helping behaviours. J Small Bus Entrep Dev. 2015;3(2):89-112.
9. Ajonbadi HA, Lawal AA, Badmus DA, Otokiti BO. Financial control and organisational performance of the Nigerian small and medium enterprises (SMEs): a catalyst for economic growth. Am J Bus Econ Manag. 2014;2(2):135-43.
10. Ajonbadi HA, Otokiti BO, Adebayo P. The efficacy of planning on organisational performance in the Nigeria SMEs. Eur J Bus Manag. 2016;24(3):25-47.
11. Akinbola OA, Otokiti BO. Effects of lease options as a source of finance on profitability performance of small and medium enterprises (SMEs) in Lagos State, Nigeria. Int J Econ Dev Res Invest. 2012;3(3):70-6.
12. Akinrinoye OV, Umoren O, Didi PU, Balogun O, Abass OS. Predictive and segmentation-based marketing analytics framework for optimizing customer acquisition, engagement, and retention strategies. Eng Technol J. 2015;10(9):6758-76.
13. Akinrinoye OV, Umoren O, Didi PU, Balogun O, Abass OS. Evaluating the strategic role of economic research in supporting financial policy decisions and market performance metrics. IRE J. 2019;3(3):248-58.
14. Akomea-Agyin K, Asante M. Analysis of security vulnerabilities in wired equivalent privacy (WEP). Int Res J Eng Technol. 2019;6(1):529-36.
15. Akpan UU, Adekoya KO, Awe ET, Garba N, Oguncoker GD, Ojo SG. Mini-STRs screening of 12 relatives of Hausa origin in northern Nigeria. Niger J Basic Appl Sci. 2017;25(1):48-57.
16. Akpan UU, Awe TE, Idowu D. Types and frequency of fingerprint minutiae in individuals of Igbo and Yoruba ethnic groups of Nigeria. Ruhuna J Sci. 2019;10(1):1-12.
17. Ali A, Rahut DB, Imtiaz M. Affordability linked with subsidy: impact of fertilizers subsidy on household welfare in Pakistan. Sustainability. 2019;11(19):5161.
18. Arndt C, Pauw K, Thurlow J. The economy-wide impacts and risks of Malawi's farm input subsidy program. Am J Agric Econ. 2016;98(3):962-80.
19. Asante M, Akomea-Agyin K. Analysis of security vulnerabilities in wifi-protected access pre-shared key. Int J Comput Sci Inf Secur. 2019;17(5):112-28.
20. Assumma V, Bottero M, Monaco R, Soares AJ. An integrated evaluation methodology to measure ecological and economic landscape states for territorial transformation scenarios: an application in Piedmont (Italy). Ecol Indic. 2019;105:156-65.
21. Atobatele OK, Ajayi OO, Hungbo AQ, Adeyemi C. Leveraging public health informatics to strengthen monitoring and evaluation of global health intervention. IRE J. 2019;2(7):174-93.
22. Atobatele OK, Hungbo AQ, Adeyemi C. Evaluating strategic role of economic research in supporting financial policy decisions and market performance metrics. IRE J. 2019;2(10):442-52.
23. Atobatele OK, Hungbo AQ, Adeyemi C. Digital health technologies and real-time surveillance systems: transforming public health emergency preparedness through data-driven decision making. IRE J. 2019;3(9):417-21.
24. Atobatele OK, Hungbo AQ, Adeyemi C. Leveraging big data analytics for population health management: a comparative analysis of predictive modeling approaches in chronic disease prevention and healthcare resource optimization. IRE J. 2019;3(4):370-5.
25. Awe ET. Hybridization of snout mouth deformed and normal mouth African catfish Clarias gariepinus. Anim Res Int. 2017;14(3):2804-8.
26. Awe ET, Akpan UU. Cytological study of Allium cepa and Allium sativum. J Cytol Genet. 2017;18(2):88-95.
27. Awe ET, Akpan UU, Adekoya KO. Evaluation of two MiniSTR loci mutation events in five father-mother-child trios of Yoruba origin. Niger J Biotechnol. 2017;33:120-4.
28. Ayanbode N, Cadet E, Etim ED, Essien IA, Ajayi JO. Deep learning approaches for malware detection in large-scale networks. IRE J. 2019;3(1):483-502.
29. Azumah SB, Zakaria A. Fertilizer subsidy and rice productivity in Ghana: a microeconomic study. J Agric Stud. 2019;7(1):82-102.
30. Balogun O, Abass OS, Didi PU. A multi-stage brand repositioning framework for regulated FMCG markets in Sub-Saharan Africa. IRE J. 2019;2(8):236-42.
31. Bankole FA, Tewogbade L. Strategic cost forecasting

framework for SaaS companies to improve budget accuracy and operational efficiency. Iconic Res Eng J. 2019;2(10):421-41.
32. Bankole FA, Dako OF, Onalaja TA, Nwachukwu PS, Lateefat T. Blockchain-enabled systems fostering transparent corporate governance, reducing corruption, and improving global financial accountability. Iconic Res Eng J. 2019;3(3):259-78.
33. Bankole FA, Dako OF, Onalaja TA, Nwachukwu PS, Lateefat T. AI-driven fraud detection enhancing financial auditing efficiency and ensuring improved organizational governance integrity. Iconic Res Eng J. 2019;2(11):556-77.
34. Bayeroju OF, Sanusi AN, Queen Z, Nwokediegwu S. Bio-based materials for construction: a global review of sustainable infrastructure practices. J Bio Mater Constr. 2019;4(2):78-94.
35. Bryden WL. Food and feed, mycotoxins and the perpetual pentagram in a changing animal production environment. Anim Prod Sci. 2012;52(7):383-97.
36. Buenemann M, Martius C, Jones JW, Herrmann SM, Klein D, Mulligan M, *et al*. Integrative geospatial approaches for the comprehensive monitoring and assessment of land management sustainability: rationale, potentials, and characteristics. Land Degrad Dev. 2011;22(2):226-39.
37. Cowie AL, Penman TD, Gorissen L, Winslow MD, Lehmann J, Tyrrell TD, *et al*. Towards sustainable land management in the drylands: scientific connections in monitoring and assessing dryland degradation, climate change and biodiversity. Land Degrad Dev. 2011;22(2):248-60.
38. Dako OF, Okafor CM, Farounbi BO, Onyelucheya OP. Detecting financial statement irregularities: hybrid Benford–outlier–process-mining anomaly detection architecture. IRE J. 2019;3(5):312-27.
39. Didi PU, Abass OS, Balogun O. A multi-tier marketing framework for renewable infrastructure adoption in emerging economies. RE J. 2019;3(4):337-45.
40. Dur F, Yigitcanlar T. Assessing land-use and transport integration via a spatial composite indexing model. Int J Environ Sci Technol. 2015;12(3):803-16.
41. Dur F, Yigitcanlar T, Bunker J. A spatial-indexing model for measuring neighbourhood-level land-use and transport integration. Environ Plan B Plan Des. 2014;41(5):792-812.
42. Erb KH, Haberl H, Jepsen MR, Kuemmerle T, Lindner M, Müller D, *et al*. A conceptual framework for analysing and measuring land-use intensity. Curr Opin Environ Sustain. 2013;5(5):464-70.
43. Erigha ED, Obuse E, Ayanbode N, Cadet E, Etim ED. Machine learning-driven user behavior analytics for insider threat detection. IRE J. 2019;2(11):535-44.
44. Evans-Uzosike IO, Okatta CG. Strategic human resource management: trends, theories, and practical implications. Iconic Res Eng J. 2019;3(4):264-70.
45. Eyles A, Coghlan G, Hardie M, Hovenden M, Bridle K. Soil carbon sequestration in cool-temperate dryland pastures: mechanisms and management options. Soil Res. 2015;53(4):349-65.
46. Fairhurst T. Handbook for integrated soil fertility management. Wageningen: CTA/CABI; 2012.
47. Farounbi BO, Akinola AS, Adesanya OS, Okafor CM. Automated payroll compliance assurance: linking withholding algorithms to financial statement reliability. IRE J. 2018;1(7):341-57.
48. Fasasi ST, Adebowale OJ, Abdulsalam A, Nwokediegwu ZQS. Benchmarking performance metrics of methane monitoring technologies in simulated environments. Iconic Res Eng J. 2019;3(3):193-202.
49. Filani OM, Nwokocha GC, Babatunde O. Framework for ethical sourcing and compliance enforcement across global vendor networks in manufacturing and retail sectors. J Glob Supply Chain Ethics. 2019;12(3):201-18.
50. Filani OM, Nwokocha GC, Babatunde O. Lean inventory management integrated with vendor coordination to reduce costs and improve manufacturing supply chain efficiency. J Lean Supply Chain. 2019;18:19-35.
51. Giller KE, Tittonell P, Rufino MC, van Wijk MT, Zingore S, Mapfumo P, *et al*. Communicating complexity: integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development. Agric Syst. 2011;104(2):191-203.
52. Hemming DJ, Chirwa EW, Dorward A, Ruffhead HJ, Hill R, Osborn J, *et al*. Agricultural input subsidies for improving productivity, farm income, consumer welfare and wider growth in low- and lower-middle-income countries: a systematic review. Campbell Syst Rev. 2018;14(1):1-153.
53. Hoang VN. Analysis of productive performance of crop production systems: an integrated analytical framework. Agric Syst. 2013;116:16-24.
54. Hoyle FC, D'Antuono M, Overheu T, Murphy DV. Capacity for increasing soil organic carbon stocks in dryland agricultural systems. Soil Res. 2013;51(8):657-67.
55. Hungbo AQ, Adeyemi C. Community-based training model for practical nurses in maternal and child health clinics. IRE J. 2019;2(8):217-35.
56. Hungbo AQ, Adeyemi C. Laboratory safety and diagnostic reliability framework for resource-constrained blood bank operations. IRE J. 2019;3(4):295-318.
57. Jayne TS, Rashid S. Input subsidy programs in sub-Saharan Africa: a synthesis of recent evidence. Agric Econ. 2013;44(6):547-62.
58. Jayne TS, Mather D, Mason N, Ricker-Gilbert J. How do fertilizer subsidy programs affect total fertilizer use in sub-Saharan Africa? Crowding out, diversion, and benefit/cost assessments. Agric Econ. 2013;44(6):687-703.
59. Kamau EN. Energy efficiency comparison between 2.1 GHz and 28 GHz based communication networks [dissertation]. Tampere: Tampere University of Technology; 2018.
60. Kiryushin VI. The management of soil fertility and productivity of agrocenoses in adaptive-landscape farming systems. Eurasian Soil Sci. 2019;52(9):1137-45.
61. Koschke L, Fürst C, Frank S, Makeschin F. A multi-criteria approach for an integrated land-cover-based assessment of ecosystem services provision to support landscape planning. Ecol Indic. 2012;21:54-66.
62. Lal R. Carbon cycling in global drylands. Curr Clim Change Rep. 2019;5(3):221-32.
63. Luo Z, Wang E, Sun OJ, Smith CJ, Probert ME. Modeling long-term soil carbon dynamics and

sequestration potential in semi-arid agro-ecosystems. Agric For Meteorol. 2011;151(12):1529-44.
64. Mamonov K. Methodological approach to the integral assessment of the regional lands use territorial development. Geod Cartogr. 2019;45(3):110-5.
65. Minviel JJ, Latruffe L. Effect of public subsidies on farm technical efficiency: a meta-analysis of empirical results. Appl Econ. 2017;49(2):213-26.
66. Möller K. Soil fertility status and nutrient input–output flows of specialised organic cropping systems: a review. Nutr Cycl Agroecosyst. 2018;112(2):147-64.
67. Monteiro J, Martins B, Pires JM. A hybrid approach for the spatial disaggregation of socio-economic indicators. Int J Data Sci Anal. 2018;5(2):189-211.
68. Mueller L, Schindler U, Mirschel W, Shepherd TG, Ball BC, Helming K, *et al*. Assessing the productivity function of soils. In: Lichtfouse E, editor. Sustainable agriculture volume 2. Dordrecht: Springer; 2011. p. 743-60.
69. Nasrin M, Bauer S, Arman M. Assessing the impact of fertilizer subsidy on farming efficiency: a case of Bangladeshi farmers. Open Agric. 2018;3(1):567-77.
70. Nguyen TTN, Tran HC, Ho TMH, Burny P, Lebailly P. Dynamics of farming systems under the context of coastal zone development: the case of Xuan Thuy National Park, Vietnam. Agriculture. 2019;9(7):138.
71. Nolan RH, Sinclair J, Eldridge DJ, Ramp D. Biophysical risks to carbon sequestration and storage in Australian drylands. J Environ Manage. 2018;208:102-11.
72. Nwokediegwu ZS, Bankole AO, Okiye SE. Advancing interior and exterior construction design through large-scale 3D printing: a comprehensive review. IRE J. 2019;3(1):422-49.
73. Ogundipe F, Sampson E, Bakare OI, Oketola O, Folorunso A. Digital transformation and its role in advancing the sustainable development goals (SDGs). J Digit Transform Sustain. 2019;19:48-62.
74. Oguntegbe EE, Farounbi BO, Okafor CM. Conceptual model for innovative debt structuring to enhance mid-market corporate growth stability. IRE J. 2019;2(12):451-63.
75. Oguntegbe EE, Farounbi BO, Okafor CM. Empirical review of risk-adjusted return metrics in private credit investment portfolios. IRE J. 2019;3(4):494-505.
76. Oguntegbe EE, Farounbi BO, Okafor CM. Framework for leveraging private debt financing to accelerate SME development and expansion. IRE J. 2019;2(10):540-54.
77. Onalaja TA, Nwachukwu PS, Bankole FA, Lateefat T. A dual-pressure model for healthcare finance: comparing United States and African strategies under inflationary stress. IRE J. 2019;3(6):261-76.
78. Oni O, Adeshina YT, Iloeje KF, Olatunji OO. Artificial intelligence model fairness auditor for loan systems. J Fintech Regul. 2018;8993:1162.
79. Osabuohien FO. Review of the environmental impact of polymer degradation. Commun Phys Sci. 2017;2(1):1-12.
80. Osabuohien FO. Green analytical methods for monitoring APIs and metabolites in Nigerian wastewater: a pilot environmental risk study. Commun Phys Sci. 2019;4(2):174-86.
81. Oshomegie MJ. The spill over effects of staff strike action on micro, small and medium scale businesses in Nigeria: a case study of the University of Ibadan and Ibadan Polytechnic [dissertation]. Ibadan: University of Ibadan; 2018.
82. Oshomegie MJ, Matter DIRS, An E. Stock returns sensitivity to interest rate changes. J Finance Econ. 2017;12(4):112-30.
83. Otokiti BO. Mode of entry of multinational corporation and their performance in the Nigeria market [dissertation]. Ota: Covenant University; 2012.
84. Otokiti BO. Business regulation and control in Nigeria. In: Book of readings in honour of Professor SO Otokiti. Lagos: Covenant University Press; 2018. p. 201-15.
85. Otokiti BO, Akorede AF. Advancing sustainability through change and innovation: a co-evolutionary perspective. In: Innovation: taking creativity to the market. Book of readings in honour of Professor SO Otokiti. Lagos: Covenant University Press; 2018. p. 161-7.
86. Oziri ST, Seyi-Lande OB, Arowogbadamu AAG. Dynamic tariff modeling as a predictive tool for enhancing telecom network utilization and customer experience. Iconic Res Eng J. 2019;2(12):436-50.
87. Patrick A, Adeleke Adeyeni S, Gbaraba Stephen V, Pamela G, Ezeh Funmi E. Community-based strategies for reducing drug misuse: evidence from pharmacist-led interventions. Iconic Res Eng J. 2019;2(8):284-310.
88. Rashid S, Tefera N, Minot N, Ayele G. Can modern input use be promoted without subsidies? An analysis of fertilizer in Ethiopia. Agric Econ. 2013;44(6):595-611.
89. Reidsma P, König H, Feng S, Bezlepkina I, Nesheim I, Bonin M, *et al*. Methods and tools for integrated assessment of land use policies on sustainable development in developing countries. Land Use Policy. 2011;28(3):604-17.
90. Ricker-Gilbert J, Jayne TS. Estimating the enduring effects of fertiliser subsidies on commercial fertiliser demand and maize production: panel data evidence from Malawi. J Agric Econ. 2017;68(1):70-97.
91. Robertson AD, Zhang Y, Sherrod LA, Rosenzweig ST, Ma L, Ahuja L, *et al*. Climate change impacts on yields and soil carbon in row crop dryland agriculture. J Environ Qual. 2018;47(4):684-94.
92. Sanusi AN, Bayeroju OF, Queen Z, Nwokediegwu S. Circular economy integration in construction: conceptual framework for modular housing adoption. J Circ Econ Constr. 2019;5(2):101-17.
93. Seyi-Lande OB, Arowogbadamu AAG, Oziri ST. A comprehensive framework for high-value analytical integration to optimize network resource allocation and strategic growth. Iconic Res Eng J. 2018;1(11):76-91.
94. Seyi-Lande OB, Oziri ST, Arowogbadamu AAG. Leveraging business intelligence as a catalyst for strategic decision-making in emerging telecommunications markets. Iconic Res Eng J. 2018;2(3):92-105.
95. Seyi-Lande OB, Oziri ST, Arowogbadamu AAG. Pricing strategy and consumer behavior interactions: analytical insights from emerging economy telecommunications sectors. Iconic Res Eng J. 2019;2(9):326-40.
96. Sharma P, Abrol V, Abrol S, Kumar R. Climate change and carbon sequestration in dryland soils. In: Resource management for sustainable agriculture. Rijeka: InTech; 2012. p. 139-64.
97. Sokouti R, Kaveh A, Parvizi Y. Modeling of soil carbon

storage capacity using farm management factors in drylands. Agric For. 2017;63:153-65.
98. Tang Z. An integrated approach to evaluating the coupling coordination between tourism and the environment. Tour Manag. 2015;46:11-9.
99. Therond O, Duru M, Roger-Estrade J, Richard G. A new analytical framework of farming system and agriculture model diversities: a review. Agron Sustain Dev. 2017;37(3):21.
100. Umoren O, Didi PU, Balogun O, Abass OS, Akinrinoye OV. Linking macroeconomic analysis to consumer behavior modeling for strategic business planning in evolving market environments. IRE J. 2019;3(3):203-13.
101. Van Westen CJ. Remote sensing and GIS for natural hazards assessment and disaster risk management. In: Treatise on geomorphology. Amsterdam: Elsevier; 2013. p. 259-98.
102. Vanlauwe B, Kihara J, Chivenge P, Pypers P, Coe R, Six J. Agronomic use efficiency of N fertilizer in maize-based systems in sub-Saharan Africa within the context of integrated soil fertility management. Plant Soil. 2011;339(1):35-50.
103. Zingore S, Tittonell P, Corbeels M, van Wijk MT, Giller KE. Managing soil fertility diversity to enhance resource use efficiencies in smallholder farming systems: a case from Murewa District, Zimbabwe. Nutr Cycl Agroecosyst. 2011;90(1):87-103.