Artificial Intelligence Models Bridging Genotype, Environment and Phenotype in Agricultural Genomics

Gulhayo Rahmonova, Nigora Khujakulova, Khusen Kholov, Laziz Tuychiev, Bisenbay Bekbanov, Gulmira Azamatova, Ilyos Xursandov,

Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. Tashkent, Uzbekistan, ORCID: https://orcid.org/0009-0003-1763-6865 E-mail: rgulhayo75@gmail.com

Associate Professor, Department of Innovative Economics, Karshi State Technical University. Karshi, Uzbekistan, ORCID: https://orcid.org/0009-0001-2936-1972 E-mail: nigorakhujakulova86@gmail.com

Assistant Lecturer, Bukhara State Medical Institute, Bukhara, Uzbekistan ORCID: https://orcid.org/0009-0002-7854-1133, E-mail: husen_xolov@bsmi.uz

DSc, Professor of the Department of Infectious and Pediatric Infectious Diseases of the Tashkent State Medical University, https://orcid.org/0000-0003-2312-8640, l_tuychiev@mail.ru

Researcher, Karakalpak Scientific Research Institute of Agriculture. Chimboy, Karakalpakstan, Uzbekistan, ORCID: https://orcid.org/0000-0003-0881-3009 E-mail: biysenbaybekbanov@gmail.com

PhD in History (Doctor of Philosophy), Acting Associate Professor, Department of History of Uzbekistan, Jizzakh State Pedagogical University. Jizzakh, Uzbekistan ORCID: https://orcid.org/0009-0001-8595-7615, E-mail: g.azamatova1987@bk.ru

Department of Surgical Diseases, Termez University of Economics and Service. Termez, Uzbekistan, ORCID: https://orcid.org/0009-0002-2372-8005 E-mail: ilyos_xursandov@tues.uz

ABSTRACT

Genotype–environment–phenotype (GEP) interactions are central to agricultural productivity, resilience, and sustainability. Genotype determines how organisms respond to environmental variability, shaping phenotypic outcomes such as yield, quality, and stress tolerance. However, the complex, nonlinear, and stochastic nature of genotype–environment (GE) interactions poses major challenges for prediction and decision-making in breeding and agronomic management. Recent advances in artificial intelligence (AI) offer powerful tools to bridge genotype, environment, and phenotype by integrating heterogeneous data sources, modeling high-dimensional interactions, and enabling predictive and inferential analyses across biological scales. This work presents a conceptual and methodological framework for AI-driven GEP modeling in agricultural genomics, encompassing data acquisition, preprocessing, and integration of multi-omics, environmental sensing, phenotyping platforms, and management metadata. We highlight the role of supervised learning and related AI approaches in learning GE relationships, predicting phenotypic outcomes under specific environmental conditions, and supporting genotype selection and deployment strategies. By moving beyond descriptive associations toward causal understanding, AI-enabled GEP frameworks can enhance breeding efficiency, optimize resource use, and improve the adaptability of crops and livestock to diverse and changing environments.

Keywords: *Genotype–environment interaction (GE); phenotype prediction; agricultural genomics; artificial intelligence; machine learning; multi-omics integration; environmental modeling; phenotyping; genomic selection; crop and livestock breeding; data integration; causal inference.*

INTRODUCTION

One of the most widespread childhood ailments in the whole world, dental caries, also referred to as tooth Genotype describes the genetic constitution of an organism; environment refers to external conditions

affecting a living system; and phenotype is the organism's observable characteristics. Genotype, environment, and phenotype (G, E, and P respectively) are fundamentally connected. G determines how environmental forces shape individual phenotype (P|G,E). A crop cultivar's gene variants govern G×E interactions, including gene expression modulation, and these interactions impact P, with consequences for harvest yield and quality. Crop production environments include soil characteristics, temperature, water, solar radiation, diseases, pests, and fertilization and irrigation practices. The magnitude of the G×E interaction varies according to the genetically determined biological functions of the crop, the agricultural-production process, and the stochastic nature of environment variability. Thus, the G×E interaction can be recognized as a mathematical relation among G, E, and P, enabling subsequent G, E, and P predictions and assessments.

Artificial intelligence (AI) is involved in multiple stages of the crop and livestock production chain, including genotype and environmental data acquisition, phenotyping data extraction, genotype × environment (G×E) modeling, and the establishment of genotype–environment–phenotype (G×E→P) transition linkages. AI can (i) perform data-handling operations at the genotype (G), environment (E), and phenotype (P) data layers, such as quality control, missing-value imputation, and matrix transformation; (ii) model G×E relationships by learning from G, E, and P data patterns and subsequently predict or evaluate G, E, or P outcomes under specific conditions, thereby enabling G, E, or P → G, E, or P inference tasks; and (iii) support the construction of G×E→P linkages by integrating approaches such as G→P, E→P, G,P→E, or G→E (Hafeez Ullah Khan et al., 2022). Genotype–environment interactions and G×E patterns can be stored as intermediate knowledge within G, E, and P models by establishing direct or indirect relationships among G and P, E and P, G and E, or G alone. Furthermore, genotype information, genotype-derived datasets, and G models from diverse agricultural production systems are essential for identifying key components of the G×E→P transition in breeding programs.

Extra insight will be deducted by going beyond G, E, and P facts, rules, and functions to consider causal relations. The proposed theoretical framework encompasses (1) concerning the role of genotype in genotype–environment and G×E interactions; (2) the directions in which genotype at the organism level may interact with environmental and E variations; and (3) the scope of environment, environmental modality, phenotype, and P interactions relevant to G×E and scientific investigations. Each component of this framework on G×E interactions is applicable to food production, genetic resources, and other crop and livestock systems [table 1].

Table 1. Core Components of the Genotype–Environment–Phenotype (G×E×P) Framework in Agriculture

Component	Definition	Key Elements	Role in Agricultural Systems
Genotype (G)	The genetic constitution of an organism	DNA sequence variants, alleles, markers, structural variants, functional annotations	Determines biological potential and regulates how environmental factors influence phenotype
Environment (E)	External physical, biological, and management conditions affecting growth	Soil properties, temperature, water availability, radiation, pests, diseases, fertilization, irrigation, management practices	Modulates gene expression and phenotypic plasticity; introduces variability and uncertainty

Phenotype (P)	Observable and functional traits resulting from G×E interactions	Yield, quality, biomass, stress tolerance, developmental traits	Represents the measurable outcome of genotype deployment in specific environments
G×E Interaction	The differential response of genotypes across environments	Gene expression modulation, phenotypic plasticity, stochastic environmental variation	Explains why genotypes perform inconsistently across sites and seasons
G×E→P Mapping	Mathematical and biological relationship linking G and E to P	Statistical models, causal relationships, predictive functions	Enables prediction, assessment, and optimization of breeding and management decisions
Temporal & Spatial Context	Time- and location-dependent environmental variation	Seasonal stability, cumulative effects, spatial heterogeneity	Influences long-term productivity and genotype suitability

Conceptual Framework: Genotype, Environment, and Phenotype in Agriculture

Genotype is defined as a genetic constitution of an organism, environment as the external physical and biological conditions affecting growth and reproduction, and phenotype as functional and observable traits or features. Artificial Intelligence (AI) aids in elucidating these parameters in wide-ranging modalities and thus assists in decoding GxE-P mechanisms for agricultural systems in which such interdependence is crucial to reproduction, growth, and productivity (Hafeez Ullah Khan et al., 2022). Elucidation of GxE-P mechanisms is construed here in terms of the characterization of deployment environments, explicit modelling of the GxE-P mapping condition, and the determination of the precision with which G, E, and P can be predicted. Such GxE-P mechanisms govern the expression of a given genotype based on the ambient circumstances, and agricultural breeding selection decisions made in pursuit of superior cultivars-occurs across G, E, and P space. Environments considered are those relatively stable at the seasonal scale and capable of long-lasting cumulative effects over the course of the lifespan of crops and livestock.

Therefore, agricultural GxE-P is delineated in the form of GxE-contextualised agricultural systems, encompassing multiple dimensionally complex factors-spatial, environmental, and temporal- that affect the attainable aggregate productivity or yield of a regulator and its harvested organs. A further complication arises where productivity arising from the deployment of a given genotype in a specific set of circumstances is the cumulative consequence of multiple environmental factors varying in modality and intensity over time. Thus, genotype and environment broadly pervade the agricultural sector and all species and cultivars deployed therein. The identification of precise genotype and the specification of target GxE circumstances, therefore, have far-reaching implications spanning the agricultural sector broadly.

Data Landscape in Agricultural Genomics

To facilitate genotype–environment–phenotype (GEP) modeling in agricultural genomics, this section charts the data landscape. It inventories relevant omics layers, environmental sensing modalities, and phenotyping platforms, alongside metadata reflecting management practices. A comprehensive GEP approach also demands information on farm-specific conditions, historic site records, and local agricultural statistics. Data heterogeneity poses a major challenge. Genomic, transcriptomic, metabolomic, phenotypic, environmental, and management datasets differ in structure, scale, and observability (Hafeez Ullah Khan et al., 2022). Quality control for (meta)data and common preprocessing operations-such as outlier removal,

normalization, transformation, interpolation, and imputation-are therefore crucial. Data availability plays a pivotal role. Proprietary restrictions, economic confidentiality, and institutional policies limit corporate and public access. Standardization and interoperability of metadata standards, taxonomies, vocabularies, and formats are likewise critical for sharing-both nationally and across protocols, sensors, and platforms. External stewardship and provenance tracking mitigate issues arising from unavailability, changing accessibility, and constraints on transparency or reuse. Data provenance documentation fosters trust, engagement, and successful collaborations.

Within genomic data, markers, structural variants, annotations, functional allele predictions, and protein sequences enhance accessibility and stimulate model development. For transcriptomic and metabolomic data, raw and processed expression measures and profiles accompany annotations. Phenotypic specifications cover traits, traits measured, and time of measurement. Environmental data encompass sources, sensors, practices, and conditions captured, while management principles indicate significant interventions carried out (crop, variety, treatment, rotation, or other details) [table 2].

Table 2. Artificial Intelligence Roles, Data Types, and Methods in GEP Modeling

AI Role	Data Layer(s) Involved	Data Types	Methods / Tasks	Agricultural Applications
Data Handling & Preprocessing	G, E, P	Genomic, transcriptomic, metabolomic, phenotypic, environmental, management metadata	Quality control, normalization, outlier removal, interpolation, imputation, matrix unfolding	Improves data reliability and model readiness
Data Integration	G×E×P	Multi-omics, sensor data, farm records, historical site data	Data fusion, interoperability, metadata standardization	Enables holistic system-level modeling
Supervised Learning	G→P, E→P, G×E→P	Labeled genotype, environment, and phenotype data	Regression, classification, prediction	Phenotype prediction, genotype evaluation, yield forecasting
G×E Modeling	G×E	Genotype × environment interaction matrices	Learning interaction patterns, response surfaces	Identifies environment-specific genotype performance
Inverse Prediction	P→G, E→G	Phenotypic and environmental observations	Genotype inference, suitability assessment	Supports breeding selection and environmental targeting
Knowledge Representation	G, E, P Models	Rules, learned patterns, causal relations	Storage of intermediary knowledge	Reusable insights across crops, locations, and systems
Decision Support	G×E→P	Integrated predictions	Scenario analysis, optimization	Guides breeding programs and

				deployment strategies

Artificial Intelligence Methods for genotype–environment–phenotype Modeling

Agricultural genomics is increasingly recognizing the fundamental influence of the environment on the relationship between genotype and phenotype. Modeling the interaction between genotype and environment-referred to as the genotype–environment (GxE) framework-enhances predictive capabilities in breeding and agronomy. Artificial intelligence (AI), which encompasses a multiscale range of data-driven and mechanistic approaches, plays a critical role in bridging genotype, environment, and phenotype. Accurate predictions of phenotype based on genotype and inference of genotype from phenotype or environmental measurements are of high practical value in agriculture. AI approaches enable the integration of a broader range of environmental and phenotypic data than is accessible from conventional genotype-based approaches such as genomic selection and rapid cycling. The ensuing predictive models can guide resource allocation in breeding programs, delineate growth conditions compatible with a given genotype, and indicate whether a genotype is suitable for a particular environment. These capabilities can be leveraged not only for crop and livestock variety evaluation during breeding but also for high-throughput environmental characterization and genotypic characterization of breeding-matched selections in trials. (Hafeez Ullah Khan et al., 2022)

Supervised Learning Approaches

Genotype, environment, and phenotype characterize the fundamental biological entities driving change at the population and ecological scales. Their interactions govern the production of food, feed, fuel, and biobased materials. Definitions begin with genotype: the composition of nucleic acid sequences that directs the assemble of biological structures (Hafeez Ullah Khan et al., 2022). Environment specifies everything external of the genotype, from subsequencing world such as space, time, and temperature, to stimuli that promote irreversible change such as nutrient availability, peril exposure, and chemical metalloids, that modulate content streams flowing from genetic instruction to cellular composition. Thus, environment encompasses an enormous and diverse range of attributes that shape plant and animal growth. Phenotype characterizes material objects and resulting structures produced in the assembly compartment that may be litter piles, fruit, or timber seen on trees and twigs that fall on the ground. Environment-inclusive genotype to phenotype mapping thereby represents a larger, more complex and flexible linkage than that comprising the underlying nucleotide lettering.

Agricultural scientists have long sought to understand and predict genotype–environment (G×E) interactions guiding phenotype, paid to causal feedback components such as phenotypic plasticity occurring during originating from non-genomic perturbation. Yet, G×E linkage and its potential breaches-predictable yet not reconcilable with genotype alone-also prominently appears in biological and chemical engineering and material science since macroscopic substance observations accompany dense and costly microscopy or spectrometry data. Species resemblances suggest simplified mapping methods heavily constrained by particles and material modalities yet extraordinarily complex mapping. Critically, agricultural species extend the challenge from genotype to externally manifested resemblance among breeding materials. Educational needs accumulated in academic, industrial, and government laboratories concern the precise objects, links and means directed in crop and livestock G×E efforts by the artificial-intelligence (AI) community and, its irreducibly sample-sufficient grand challenge.

Artificial intelligence is the art of making machines perform tasks completed by human intelligence. Foundational materials examine outline-driven specification and physically grounded flows for AI transfer to crops, livestock and related recently coined agricultural art. Biophysical structures, assembly-unit formation, and finite-size agent groupings nontrivially enrich ingredients realization within the agricultural domain. Some landscape formulations also interplay with the AI definition.

Bayesian and Probabilistic Methods

Bayesian and probabilistic methods naturally extend classical statistical approaches to accommodate prior information on parameters, representation of parameters as random variables, and explicit modeling of uncertainty. These capabilities align with the issues identified. Agricultural systems can operate in diverse and unpredictable conditions, and extending mechanistic models to coverage space would entail considerable complexity. Probabilistic machine learning (ML) provides an alternative. Unlike deep learning, it tackles uncertainty without large datasets, making it well suited for such applications, even in geoclimatic areas with limited field tests. The appeal of probabilistic solutions entices active research on models for G × E interactions, and preliminary attempts document the feasibility of the endeavor. Early work concentrated on partitioning the response space. Recent efforts in broadening the approach indicate promising directions of exploration.

Probabilistic methods include generalized additive models (GAMs) that permit nonlinear relationships without structure specifications and handle heteroscedastic noise. Bayesian or frequentist fitting accommodates incomplete data; posteriors extend to unvisited environments enabling prediction under unseen conditions. Interpretability remains crucial in breeding scenarios; GAMs provide graphical insights into GE, specify variable importance, and associate G E with biophysical variables observed prior to crop experiments (Hafeez Ullah Khan et al., 2022).

Deep Learning Architectures for GE Phenotyping

Deep learning, the fastest-growing area of machine learning, has shown considerable potential for the agricultural sector, particularly in crop genomic prediction (Jubair & Domaratzki, 2023). This is primarily due to the increasing availability of large, multi-faceted datasets comprising diverse data such as weather, soil, genomic, field management information, and multi-spectral observations collected with unmanned aerial vehicles (UAVs). Governmental and private investments in high-throughput production of multi-omics data are accelerating the generation of multi-omics datasets that complement the structure of different agronomical datasets for crops and livestock and for soil and irrigation. Such datasets are of immediate interest in E–phenotype modeling for agricultural genomics, addressing urgent challenges posed by climate change and providing assessment tools for breeding programs. Multi-scale crop genomic applications range from genomic-selection modeling under variable environments to optimal fertilisation and irrigation management.

Causal Inference and Interpretability

Causal inference and interpretability are essential for linking crop genome data to phenotypic traits using AI. Causal diagrams or graphs allow modelling of complex relationships among traits causally influenced by the genome, exhibited during development in diverse environments. These relationships are integrated into causal models using exposure and counterfactual frameworks, enabling counterfactual reasoning and predictions of trait modifications under environmental changes. Causal models facilitate understanding of the crop organism's reaction to climatic variations, respond to changing climatic conditions, and provide explanatory frameworks for the AI models considered. Causal diagrams can be inferred from physical or physiological knowledge, or identified from datasets using causal discovery methods. Explainable AI frameworks complement phenotypic prediction by providing interpretability, transparency, and accountability to deep-learning black-box models. These approaches generate explanations for model predictions, including feature importance ranking, such as Shapley additive explanation, and visualisation of decoded images providing details on the identified features.

Environmental Modeling and High-Throughput Phenotyping

Agricultural systems significantly influence conformation, development, mortality, and other life-history traits of crops and livestock during their growth cycle. Temporal variations in critical ambient

environmental variables (temperature, radiation, soil moisture, etc.) drive genotype–environment (GE) interaction phenomena, two factors involved in growth and development of agricultural organisms (Hafeez Ullah Khan et al., 2022). Environmental modeling, combining physical simulation of water, heat, and chemical transport (transfer functions) with accumulation functions for variables including biomass and age, is often applied in crop and livestock systems. Sensors can collect indicators of water and energy flows or physically observable phenomena such as soil occupation and surface temperature, enabling a further description of ambient environmental conditions. For many crops and animals, ambient variables are readily observable in field trials at low cost, motivating investment in environmental modeling.

High-throughput phenotyping provides temporal and spatial characterization of dry matter or chemical production and of morphology and structure. Crop and livestock scouting (observations of leaf, flower, and other development phases) concentrates data interpretation, facilitating connections with accumulated environmental description. Required inputs, identified as G × E interactants, thereby become available for artificial intelligence (AI) algorithms.

Genomic Prediction under Variable Environments

Analyzing the effects of environmental variation on phenotypic performance is important for the genetic improvement of agricultural species-crop plants in particular, growing under field conditions. The key to understanding the accumulation of genetic gains across generations is the assessment of selection pressure for the traits targeted by breeding programmes over many environments. The expression of yield -a complex trait that is often the target of genome-wide selection-was closely monitored in breeding trials involving barley over twenty locations in North Dakota. Selection pressure was found to focus on robustness (the ability to perform well under multiple conditions) rather than on optimum or best condition performance, a finding that emphasized the importance of climatic information in the breeding programme (Jubair & Domaratzki, 2023).

Multi-omics Integration for GxE Insights

AI-enabled and high-throughput multi-omics platforms generate extensive soil, meteorological, and diagnostic data toward environmental characterization at field, plant, or organ levels. Accordingly, multi-omics datasets collected from germplasm and environmental-sensor networks spanning diverse agricultural genotypes underpin the inference of genotype–environment–phenotype mapping at multiple scales from genes to crops (Hafeez Ullah Khan et al., 2022). Multi-omics approaches leverage varietal- and temporal-scale multi-omic datasets to unravel the underlying molecular responses of crops and livestock to diverse environments. Strategies include data fusion and multi-omics networks capable of eliciting joint biological insights from concerted analysis of several omics dimensions. Additional opportunities arise from fertilization and management-response analyses based on knowledge from nutrients and management practices, data at relevant scales, and diverse cropping systems-significant contributions for enabling sustainable agriculture through adaptation to climate change and related stressors.

Experimental Design and Validation Strategies

The design and execution of experiments to assess genotype–environment–phenotype (G×E×P) modeling algorithms require careful consideration of data collection and measures of success. In G×E modeling, success metrics should quantify how well the model learning captures the G×E relationship, rather than covariate-driven environmental or phenotypic targets (Hafeez Ullah Khan et al., 2022). Such G×E-focused metrics are readily extendable to G×E×P experiments but given the complexity of full G×E×P modeling, a simpler and higher-value goal for launching crop G×E×P efforts is to establish genotypic rankings under replicate selected environmental stresses. A dataset suitable for a first learning-phase crop G×E×P project would thus determine how well G×E interactions allow the ranking of genotypes with input genomic data and restricted environmental covariates.

To realize such a dataset, define the intended G×E×P crop modeling properties and the expected success metrics. Deploying a model that, provided with a chosen set of covariates from cereal crops, ranks genotypes with respect to simulated environments is a sound first candidate. Ensure that all the data fall within a well-defined field of interest, suitable for dataset cultivation, and that the datasets, sensors, and selected environments remain consistently aligned with all the pertinent environmental covariate metadata. Also gather such metadata for the crop data available from breeders' trials, or either restrict or broaden the crop covariate input set according to still-hypothetical availability.

Ethical, Regulatory, and Societal Considerations

Artificial intelligence (AI) is transforming agriculture and resource management, offering the possibility of more responsive and participatory monitoring and use of resources. It provides new ways to connect the diverse interests of people with different backgrounds and contributes to more responsible resource use through economic models. Environment, the set of external conditions affecting the attributes of crops or livestock, plays a fundamental role in agriculture and is evidently linked to all agricultural practices, such as sowing time, crop rotation, irrigation, fertilization, and pesticide application. Environment also interacts continuously with genotype, controlling the expression of genetic traits, leading to different phenotypes across locations. Agriculture aims to adjust genotype, environment, and practices to achieve optimal and sustainable phenotypes. An accurate understanding of environment–genotype and practice–phenotype relationships is essential. These relationships are complex and still remain poorly understood in crop and livestock production. Consequently, it is common to model the relationship as, with, according to a mechanistic or data-driven model, to assist in further exploration of the exact environment, genotype, or practice that contributes to the observed crop and livestock attributes.

A phenotype is an observable trait of an organism that is influenced by genotype and environmental factors at the time of observation. Genotype is determined at the time of seed purchase in crops. In livestock, breed serves as a rough surrogate for genotype. Data from omics science and tools, such as DNA sequencing, allow one to define genotype at the specific genomic level. The environment of a plant is defined by the combination of location and time. Different sites and periods receive different amounts of light, rainfall, temperature, humidity, and CO_2 concentration, which contribute to a unique environmental signature of each site during its cultivation. Recent studies have indicated that such an environment signature is relevant to the estimation of yield and other economic traits. The measurement of environmental data and advances in covariates science also allow individuals to obtain such an environment signature. Keeping crop or livestock phenotypes fixed, practices can change and a set of practice models can be established similarly. Thus, E, an environmental signature, and P, a set of practices, can be determined, allowing further exploration of the link between genotype and such environment and practice.

Agricultural AI is extremely valuable for helping farmers. Crop and livestock phenotypes, including anything that can be directly, accurately, and economically measured of a given crop (or livestock) during the cultivation (or breeding) period, have attracted the broadest and greatest attention in precision agriculture since these attributes directly affect economic benefits. Crop and livestock attributes define the economic value of agriculture; these attributes are very complex, varying with a myriad of practices, environments, installations, and even social and behaviour attributes. Consequently, to optimize practices, environment, and input, other variables are often taken as fixed and the methodology of AI is introduced to model the remaining intricate relationships simultaneously (Hafeez Ullah Khan et al., 2022).

Case Studies and Applications

Bridging Genotype, Environment, and Phenotype in Agricultural Genomics: The Role of Artificial Intelligence

Artificial intelligence (AI) methods have been established in various domains beyond agricultural genomics, facilitating the examination of successful breakthroughs and the lessons they can provide. In molecular biology and bioinformatics, the evolution of protein–protein interaction and gene regulatory network modeling demonstrates the complex interplay between mechanistic and data-driven approaches (Hafeez Ullah Khan et al., 2022). The genomics community has similarly witnessed the widespread adoption of deep learning for the prediction of gene expression, open chromatin, transcription factor binding sites, and DNA methylation states across diverse biological systems, opening opportunities to inform models of genotype–phenotype relationships and garnering broader interest in crop applications. Conversely, the fusion of sensor networks, modeling suites, and optimization algorithms to support demand-driven irrigation scheduling and water allocation on managed agricultural landscapes highlights the benefits of integrating mechanistic understanding and system-level considerations (Ziyaev A. A. et al).

The careful selection of case studies reveals fundamental principles of AI integration that extend beyond specific disciplinary or topical contexts. Far-precursors in agriculture and geospatial sciences demonstrate the feasibility of bridging genotype, environment, and phenotype through high-dimensional data integration and multi-scale spatio-temporal modeling, providing a conceptual model that can be adapted to contemporary G × E problems in crops and livestock. Mechanistic approaches involving environmental control have successfully guided investments in sensor deployment and data processing while resorting to stochastic proxy models to evaluate technology adoption and management practices elsewhere. Existing participatory efforts across the brewing, pest control, and transportation sectors showcase stakeholder engagement techniques that are equally applicable to building trust and aligning success criteria in agricultural G × E applications.

Insights from these case studies guide the articulation of acceptable data, method, and model choices; delineate opportunities for causal exploration; inform the elaboration of device and network attributes; and reveal perspectives on successful validation. For instance, establishing the link between observed data and management practice permits, among other avenues, counterfactual experiments to assess the impact of adopting alternative strategies on production potential. Pragmatic considerations emerge, such as the preference for an explicit empirical specification to facilitate operational implementation previously de-emphasized in gridded-demand settings. The case studies also underscore the augmented uncertainty, challenge, and opportunity associated with operating at the frontier of data return, innovation, and implementation.

Comparative insights across the food–energy–water nexus emphasize the advantages and opportunities inherent in mechanistic representation-either explicitly or implicitly-while simultaneously illustrating how AI can mitigate the complexity of systems requiring diverse temporal and spatial landscape inventorying. Analogous situations where agricultural data distribution aligns with tendencies elsewhere in the food–energy–water domain suggest that machine-learning investigation remains worthwhile despite incomplete, noisy, or distorted observations. The sharing of openly accessible, fully specified publications concerning either the mechanistic or behavioral links between apparent system components has been found to stimulate creativity and ambition in modeling efforts (Azimova S. et al).

Challenges and Future Directions

Genomic and environmental data availability is expanding rapidly, yet agricultural systems remain largely underserved by data-intensive methods, especially in model-assisted breeding. AAGxE–phenotype frameworks have proven successful in academia and medicine but are underexplored in crops and livestock. A collaborative, interdisciplinary effort is needed to identify relevant datasets, establish key research questions, and quantify the prospective economic value of AIGxE approaches in breeding and management (Hafeez Ullah Khan et al., 2022).

Conclusion

Artificial intelligence can play a vital role in agricultural genomics by bridging genotype, environment, and phenotype. A conceptual framework for genotype–environment–phenotype (G×E×P) interactions underpins two hypotheses. First, the expression of genotype-dependent phenotypes frequently exhibits distinct genotype–environment (G×E) and environment–phenotype (E×P) links that are also discernible across species and in both crops and livestock (Hafeez Ullah Khan et al., 2022). Second, agricultural genotype, environment, and phenotype data from breeding trials and agronomic observations constitute a sufficient landscape to connect genotype, environment, and phenotype across both species with artificial intelligence. Characterizing G×E×P interactions thus helps address broader issues such as predicting product quality, fostering resilience to climate change, and supporting transitional agricultural practices (Sasmakov S. A. et al). Artificial intelligence contributes to understanding G×E×P interactions through mechanistic and data-driven perspectives. Mechanistic models formulate genotype, environment, and phenotype relationships based on physical knowledge and can incorporate or calibrate biophysical or biogeochemical components; they benefit from extrapolation to unexplored regions not represented in the training dataset. In contrast, data-driven approaches exploit available data to elucidate relationships without explicitly encoding physical mechanisms. Hybrids, combining mechanistic and data-driven, strike a balance between these extremes. In agricultural genomics, bridging genotype, environment, and phenotype renders G×E×P elucidation of substantial value. Data relating projects for perennial grain crops and livestock production across multiple species and environments indicate broad transferability of G×E×P insights beyond these domains.

References:

1. Hafeez Ullah Khan M., Wang S., Wang J., Ahmar S., Saeed S., Ullah Khan S., Xu X., Chen H., Akhter Bhat J., Feng X. Applications of Artificial Intelligence in Climate-Resilient Smart-Crop Breeding // [Electronic resource]. - 2022. - Available at: ncbi.nlm.nih.gov (accessed: date).

2. Jubair S., Domaratzki M. Crop genomic selection with deep learning and environmental data: A survey // [Electronic resource]. - 2023. - Available at: ncbi.nlm.nih.gov (accessed: date).

3. Abdurakhmanov J. et al. Cloning and expression of recombinant purine nucleoside phosphorylase in the methylotrophic yeast Pichia pastoris // Journal of Advanced Biotechnology and Experimental Therapeutics. - 2023. - Available at: https://doi.org/10.5455/jabet.2023.d153

4. Aakansha Soy. (2025). Human–Animal–Environment Interfaces: Ecological and Conservation Perspectives. Journal of Animal Habitat and Environmental Interactions, 2(1), 20–27.

5. Ziyaev A. A. et al. Synthesis of S-(5-aryl-1,3,4-oxadiazol-2-yl) O-alkyl carbonothioate and alkyl 2-((5-aryl-1,3,4-oxadiazol-2-yl)thio) acetate, and their antimicrobial properties // Journal of the Turkish Chemical Society. Section A: Chemistry. - 2023. - Available at: https://doi.org/10.18596/jotcsa.1250629

6. Azimova S. et al. Study of the immunogenicity of combination of recombinant RBD (Omicron) and nucleocapsid proteins of SARS-CoV-2 expressed in Pichia pastoris // The Open Biochemistry Journal. - 2023. - Available at: https://doi.org/10.2174/011874091X2737716231122102205

7. Sasmakov S. A. et al. Expression of recombinant PreS2-S protein from the hepatitis B virus surface antigen in Pichia pastoris // VacciMonitor. - 2021. - Vol. 30, No. 1. - P. 27–32.