Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Review of Mechanistic and Hybrid Artificial Intelligence Models for Predicting Pesticide Fate and Transport in Soil and Water

Dhanraj M R
Nitte (Deemed to be University
Sanjith Anchan
National Institute of Engineering Mysore
Nausha Shetty
nausha.shetty@manipal.edu

Manipal Academy of Higher Education
Arjun Shetty
the Energy and resource institue
Mithesh Kumar
Government Engineering College
Divya Das
Nitte (Deemed to be University

Systematic Review

Keywords: Pesticide fate and transport, Predictive modelling, Machine learning, Hybrid AI frameworks, GIS and remote sensing, Groundwater contamination risk

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11345910/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

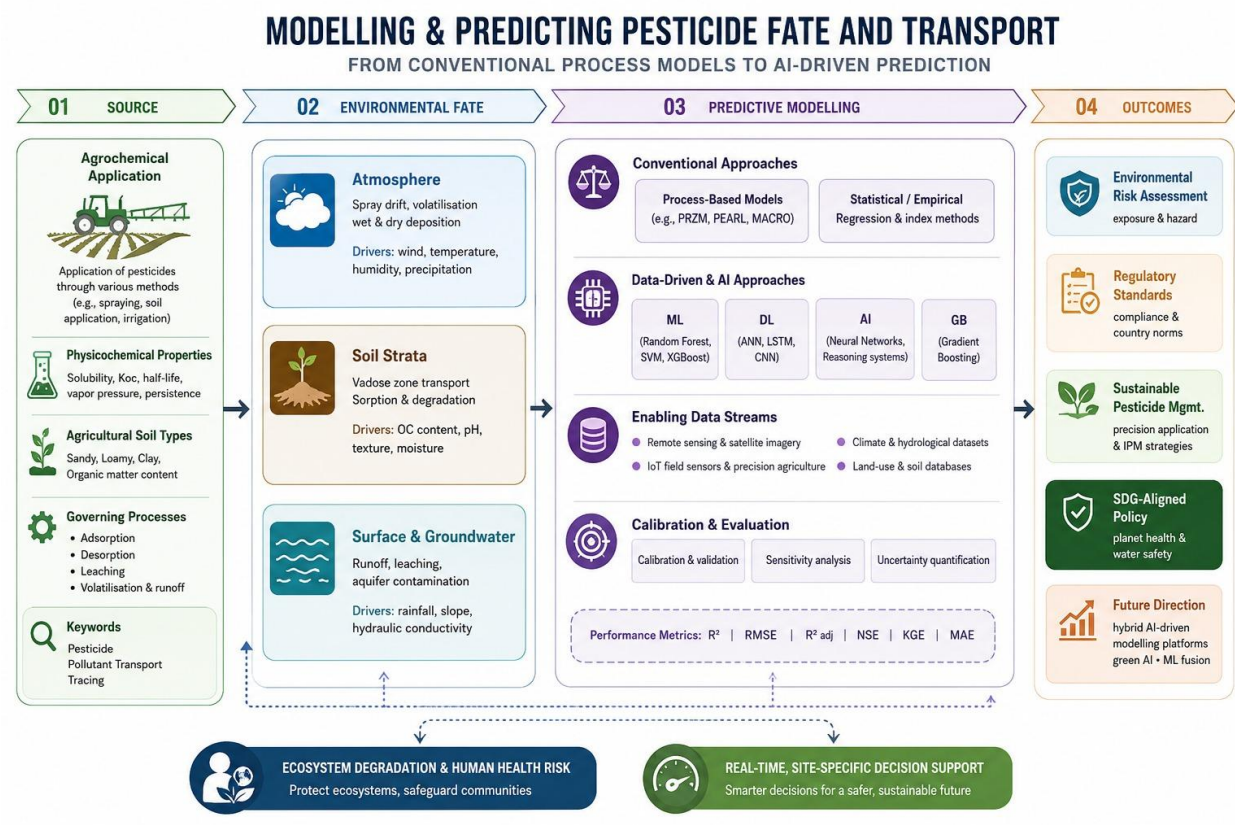
A Review of Mechanistic and Hybrid Artificial Intelligence Models for Predicting Pesticide Fate and Transport in Soil and Water

Dhanraj M R[1], Sanjith Anchan[2], Nausha Shetty[3*], Arjun Shetty[4], Mithesh Kumar[5], Divya Das[6]
[1,6]Nitte (Deemed to be University), Nitte School of Fashion Technology and Interior Design, Bengaluru, 560064, Karnataka, India
[2]Department of Civil Engineering, The National Institute of Engineering, Mysuru, India – 570008
[3*]Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

[4]The Energy and Resources Institute, Bengaluru, 560071, Karnataka, India
[5]Civil Engineering Department, Government Engineering College, Kushalnagar, Karnataka, India

*Corresponding author: nausha.shetty@manipal.edu

Graphical Abstract



Abstract:

Global pesticide use exceeded 3.7 million tonnes in 2023, and residues that persist beyond the treated field drive diffuse soil, surface water, and groundwater contamination linked to ecological harm and an estimated 150,000 pesticide poisoning deaths annually. Because field monitoring of pesticide fate across every soil, climate, and hydrological setting is logistically and financially infeasible at regulatory scale, predictive modelling has become the primary tool for estimating environmental concentrations and associated risk. This review traces the development of the field from traditional process-based mechanisms to hybrid AI frameworks by synthesizing 81 peer-reviewed and technical sources from an initial pool of 1,450 records screened across Scopus, Web of Science, and ScienceDirect (2005–2026). We look at the physicochemical and degradation principles that determine the fate of pesticides; the deterministic, mechanistic, process-based, empirical, and statistical models that support the current regulatory risk assessment; the hydrological and transport models that function at the soil profile to watershed scale; and the expanding array of machine learning, deep learning, fuzzy logic, GIS, remote sensing, and IoT-enabled techniques that are transforming precision agriculture and prediction. Case studies from India and other countries demonstrate real-world uses in addition to ongoing limitations in data accessibility and cross-context transferability. We point out important gaps, such as persistent data scarcity in low- and middle-income regions, limited coupling of fate models to downscaled climate projections, and inadequate integration of explainable AI to foster regulatory trust. We identify calibration and validation, particularly probabilistic uncertainty quantification, as essential to the responsible use of both conventional and data-driven models. We conclude that creating accurate, transferable pesticide prediction systems that support the Sustainable Development Goals and safeguard human and environmental health requires combining mechanistic understanding with hybrid, physics-informed AI and climate-resilient precision agriculture technologies.

Keywords: *Pesticide fate and transport; Predictive modelling; Machine learning; Hybrid AI frameworks; GIS and remote sensing; Groundwater contamination risk*

1. Introduction

1.1 Background on Pesticide Usage

Globally, agriculture has become increasingly dependent on synthetic pesticides to protect crops from insects, weeds, fungal pathogens, and other pests [1]. Recently, it was reported that approximately 3.73 million tonnes of agricultural pesticides were used globally in the year 2023, according to FAOSTAT statistics [2]. These chemicals are credited with substantially reducing pre- and post-harvest losses while supporting the yield gains that have accompanied the Green Revolution and subsequent intensification of farming systems [3, 4]. Pesticides are broadly classified by target organism into herbicides, insecticides, fungicides, rodenticides, nematicides and biopesticides, and by chemical composition into organochlorines, organophosphates, carbamates, pyrethroids, triazines, and newer classes such as neonicotinoids [4]. Over the past few decades, there has been a steady increase in the usage of these compounds; millions of metric tons are sprayed to arable land each year, and several hundred active ingredients are currently registered for use globally [4, 5]. Because of this reliance on chemical crop protection, there is an ongoing need to understand and predict where these compounds go after they are applied.

1.2 Environmental and Health Concerns

According to a global study, pesticide poisoning causes around 150,000 deaths annually, mostly as a result of occupational and intentional self-poisoning in low- and middle-income countries [6]. The detrimental effects of pesticide residues on soil microbial communities, non-target organisms, aquatic ecosystems, and human health have been extensively documented since Rachel Carson's Silent Spring first brought persistent pesticides like Dichlorodiphenyltrichloroethane (DDT) to the public's attention [5]. While residues that move outside of the treated field contribute to diffuse, non-point-source pollution of surface water and groundwater, pesticide residues that remain in food crops have been associated with neurotoxicity, endocrine disruption, reproductive harm, and other chronic health outcomes in exposed populations [3, 7]. Pesticide drift and leaching also create environmental-justice concerns since exposure burdens are typically unevenly distributed among populations, and agricultural landscapes are frequently next to residential areas, drinking-water wells, and fragile ecosystems [8]. These issues go beyond the basic compound: transformation products (TPs) of pesticides can be as mobile and persistent as the original molecule, further complicating fate assessment [9].

1.3 Need for Predictive Modelling

The field monitoring of pesticide concentrations across every soil type, climatic zone and hydrological setting is prohibitively expensive and logistically impossible at the scale required for regulatory decision-making. Worldwide pesticide consumption, with Asia accounting for nearly 50% [citation needed], is a major source of environmental concern and ecological risk. Predictive modelling has become the principal tool used by regulators, researchers, and agronomists to estimate the environmental concentrations of pesticides and their associated risks [10, 11]. Predictive models allow a limited set of laboratory and field measurements to be extrapolated to diverse scenarios covering different soils, weather years and management practices, and they underpin the tiered risk-assessment frameworks used in pesticide registration across the European Union, the United States, and other jurisdictions [10, 12]. In recent years, the growing availability of remote-sensing imagery, IoT sensor networks, and large environmental datasets has catalysed a shift from purely mechanistic modelling toward hybrid and fully data-driven approaches capable of learning complex, nonlinear relationships between environmental drivers and pesticide fate [13, 14]. Figure 1 illustrates this overall framework, linking pesticide source characteristics, environmental compartments, modelling approaches, and decision support.

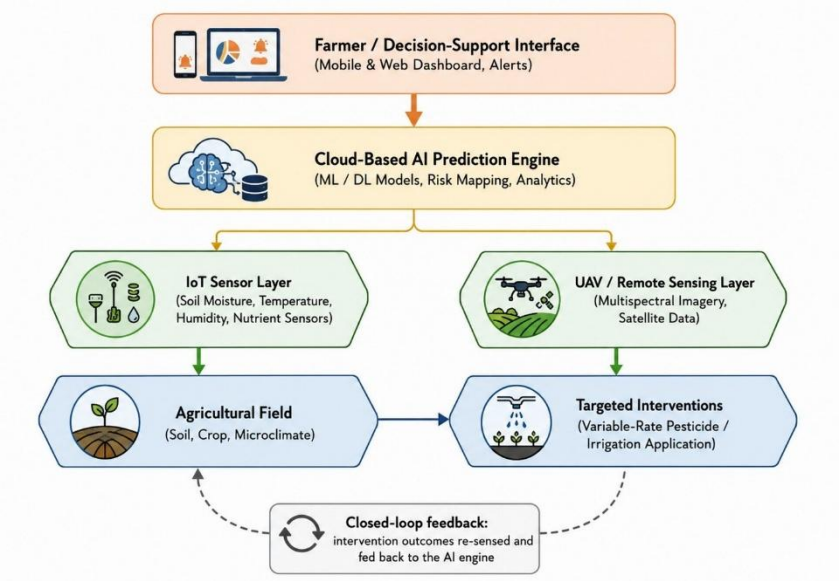


Figure 1. Framework of pesticide prediction models, linking source characteristics, environmental compartments, modelling approaches, and decision support.

1.4 Review Methodology

This review synthesises peer-reviewed journal articles and authoritative regulatory or technical reports addressing pesticide environmental fate, transport and prediction modelling. Sources were retrieved from Scopus, Web of Science, and ScienceDirect using combinations of the terms “pesticide fate”, “pesticide transport”, “machine learning”, “groundwater contamination”, “GIS”, “remote sensing” and “IoT agriculture”, restricted to records published between 2005 and 2026. 185 full texts were evaluated in depth after an initial pool of 1,450 records was selected for relevance by title and abstract; 81 of these were kept as primary sources, reinforced by important earlier references found through citation snowballing.

Studies that described an innovative fate-and-transport model, a machine-learning or hybrid-AI prediction method, or a GIS/remote-sensing/IoT application specific to pesticides or closely related agrochemical and water-quality contaminants were considered. In order to maintain the traceability of every claim in this review, conference abstracts without an accessible full text, non-English sources, and sources whose named authorship or indexing could not be independently verified were either excluded or explicitly flagged for verification. The flagged entries are kept and marked in the Reference list pending author confirmation rather than being silently removed.

1.5 Positioning Relative to Existing Reviews

The pesticide-environment literature already contains several recent reviews that overlap partially with the present topic, and it is necessary to state explicitly how this review differs from them rather than leave the distinction implicit. Health- and toxicology-oriented reviews, exemplified by Rajak et al. [1] in the Journal of Hazardous Materials Advances, synthesise the ecological and biosphere-level consequences of pesticide exposure but do not treat prediction modelling as their organising subject: fate-and-transport equations, machine-learning architectures and their comparative performance are largely outside their scope. Conversely, machine-learning-centred reviews of pesticide toxicity (e.g., systematic reviews of ML approaches to toxicity endpoints) synthesise algorithmic performance but rarely connect that performance back to the mechanistic soil-, hydrology- and climate-based descriptors (Koc, DT50, GUS) that determine what the algorithms are actually learning. GIS- and precision-agriculture-focused reviews, in turn, treat spatial data layers and sensor

networks as the endpoint rather than situating them within the classical process-based modelling tradition from which regulatory risk assessment still draws its legal and scientific authority.

This review's distinct contribution is to treat that continuum, from physicochemical fundamentals through deterministic and mechanistic process-based equations to hybrid AI, GIS and IoT, as a single evidentiary chain rather than as separate literatures, and to test that chain against verifiable numbers wherever the underlying sources allow it. Three features operationalise this: (i) a reproducibility check of the Groundwater Ubiquity Score for three contrasting compounds (Section 2.7), which shows that a widely used regulatory screening index is sensitive to a convention (field versus laboratory DT50) that is not always stated in the primary literature it reviews; (ii) a cross-study synthesis of reported machine-learning accuracy metrics rather than a simple count of which algorithms have been applied (Section 6.6), together with an explicit statement of what a defensible meta-analysis of that evidence would still require; and (iii) a critical synthesis, presented in Section 13.4, that sets contradictory findings from the reviewed literature against one another instead of listing gaps as an unconnected checklist. None of the reviews identified above combines all three elements, and it is this combination, not the underlying topic area, that constitutes the novelty claim of the present work.

One recent review in Journal of Hazardous Materials Advances requires a more specific comment because its overlap with the present manuscript is closer than the others discussed above: Tao et al. [15] critically review pesticide exceedance assessment, quantifying how far regulatory process-based models such as FOCUS underestimate real-world threshold breaches once monitoring bias, preferential flow, and mixture and transformation-product exclusion are accounted for, and treating machine learning chiefly as a bias-correction layer applied to those models' outputs. That review shares this manuscript's interest in fate-model constraints, data infrastructure and the underrepresentation of low- and middle-income regions (Sections 5, 11 and 13.1 here), and both draw on the same global pesticide-tonnage baseline. The two reviews are nonetheless organised around different questions and, correspondingly, cover different ground: Tao et al. [15] is structured around exceedance quantification, ecological-effect validation, residue and dietary exposure, and mixture and cumulative risk, and does not address the

mechanistic soil-fundamentals layer (Section 2), GIS and remote-sensing spatial prediction (Section 8), IoT sensor networks, UAV data and smart-agriculture architecture (Section 9), or region-specific case studies such as the Indian groundwater assessments in Section 12.2, all of which occupy substantial parts of the present review. Where the two reviews genuinely converge, on machine learning's role in pesticide prediction, this review treats ML, deep learning, fuzzy logic and physics-informed hybrid architectures as prediction methods evaluated on their own comparative merits (Sections 6-9, including the cross-study accuracy synthesis in Section 6.6), rather than principally as a correction applied downstream of an existing regulatory model, which is the framing Tao et al. [15] adopt. The overlap is real and is acknowledged rather than minimised, but it is confined to a subset of the shared subject matter rather than spanning the full scope of either review.

2. Fundamentals of Pesticide Behaviour

2.1 Classification of Pesticides

Herbicides, insecticides, fungicides, rodenticides, nematicides, and molluscicides are the most common ways to categorize pesticides. They can also be categorized by chemical family (organochlorines, organophosphates, carbamates, pyrethroids, triazines, phenoxy acids, neonicotinoids, and biopesticides) [4, 16]. Every chemical family has distinct mechanisms of action, persistence, and toxicity profiles: pyrethroids combine moderate persistence with high efficacy against a wide range of insects, organophosphates and carbamates act by inhibiting acetylcholinesterase and tend to degrade faster but are acutely toxic, and organochlorines, like DDT and endosulfan, are highly persistent and lipophilic [4]. Since a compound's chemical family broadly dictates the primary transport and degradation pathways that a model must reflect, this classification serves as the foundation for choosing suitable fate descriptors in prediction models.

2.2 Physicochemical Properties

The environmental fate of a pesticide is mainly determined by a limited number of physicochemical descriptors, such as water solubility, the octanol-water partition coefficient (Kow), the soil organic-carbon partition coefficient (Koc), vapor pressure, the Henry's law constant and the dissociation constant (pKa) for ionizable compounds [17, 18]. High water solubility and low Koc favor leaching and surface runoff; high Kow and low vapor pressure favor retention in soil organic matter and bioaccumulation in biota. These descriptors are

combined in simple mobility indices such as the Groundwater Ubiquity Score (GUS), which classifies compounds as leachers, non-leachers or transition compounds based on their soil half-life and Koc [19].

2.3 Degradation Mechanisms

The persistence and environmental fate of agricultural chemicals are essentially determined by complex relationships between chemical structures and surrounding media [20]. Degradation of pesticides in the environment proceeds via abiotic and biotic pathways. Abiotic mechanisms include hydrolysis, especially important for organophosphorus, organochlorine and organonitrogen compounds containing hydrolyzable heteroatoms, and photodegradation, which is affected by soil texture, moisture content and intensity of sunlight at the soil surface [17]. Biotic degradation is mediated by soil microbial communities and is typically modelled by first-order or biphasic kinetics, with the degradation rate expressed as a half-life (DT50) that is the central component of nearly all fate and transport models [17, 21]. Machine-learning models have also been trained to predict pesticide dissipation half-life intervals in plants [22]. Degradation products themselves may have toxicological significance and mobility, and plot-to-catchment-scale transformation-product-aware modelling frameworks are needed [9].

2.4 Adsorption and Desorption from Soil

Sorption processes imply that generally >80-90% of applied hydrophobic pesticides are rapidly sorbed to the soil solid phase, strongly limiting their immediate aqueous mobility and bioavailability [18]. Sorption to soil particles is arguably the single most important process controlling pesticide mobility, since it partitions the applied dose between the soil solid phase and the soil solution available for transport [9, 20]. Adsorption is most commonly described using the Freundlich or Langmuir isotherms, with soil organic matter content, clay mineralogy, and pH identified as the dominant controlling factors [23, 24]. Desorption is frequently slower and incomplete relative to adsorption, a phenomenon known as hysteresis, which is attributed to diffusion into micropores and the existence of two-compartment (rapidly and slowly reversible) binding sites [25]. Comparative studies across Indian soils, for instance, have shown that endosulfan sorption varies markedly between soil

types owing to differences in organic carbon content [23], while atrazine sorption on vertisols and alfisols differs due to clay content and cation-exchange capacity [24].

2.5 Leaching and Runoff

Leaching describes the downward movement of dissolved or weakly sorbed pesticides through the soil profile toward groundwater, and is controlled by the interaction of sorption, degradation rate, soil hydraulic properties, and precipitation or irrigation intensity [26, 27]. Field-scale monitoring has shown that the pesticide breakthrough to subsurface depths can be up to 5 times faster than predicted by standard matrix-flow models [28]. Preferential flow through macropores, root channels, and structural cracks can transport pesticides to significant depths far more rapidly than matrix (Darcian) flow would predict, a process that has been extensively documented using intact soil-column and lysimeter studies and is now represented explicitly in dual-permeability models such as MACRO [26, 27]. Surface runoff, by contrast, mobilises both dissolved and particulate-bound pesticide mass across the land surface toward streams and water bodies, and is strongly influenced by rainfall intensity, slope, and the timing of application relative to storm events [29].

2.6 Volatilization

Volatilization is the movement of pesticide mass from the soil or plant surface to the atmosphere and is influenced by the vapor pressure of the compound, the Henry's law constant and the temperature and wind conditions at the location of application [4, 30]. Mechanistic evaluations indicate that semi-volatile chemicals with high Henry's law constants can lose up to 30% to 50% of the applied active ingredient from soil and plant surfaces to the boundary layer within days of application [31]. Volatilization, although often under-emphasized relative to leaching and runoff, can be an important pathway for dissipation of high vapor pressure compounds and atmospheric transport can then redeposit pesticide residues at some distance from the point of application, a process that is incorporated into multimedia and regional-scale fate models [25, 30].

2.7 Illustrative GUS Screening Calculation

The qualitative leaching statements above can be made concrete with the Groundwater Ubiquity Score (GUS), an empirical screening index defined as GUS = $\log_{10}(DT_{50}) \times [4 - \log_{10}(K_oc)]$, where DT_{50} is the soil half-life in days, and K_oc is the organic-carbon-normalised sorption coefficient in mL g^{-1} [19]. Compounds with GUS > 2.8 are classified as leachers, GUS < 1.8 as non-leachers, and the intermediate range as a transition zone.

Table 1 applies this formula to three pesticides with contrasting fate profiles, using inputs published in the Pesticide Properties DataBase (PPDB), University of Hertfordshire [32]. Reproducing PPDB's own published GUS values required using the field-measured DT_{50} (rather than the more commonly quoted laboratory or "typical" DT_{50}) together with the Freundlich-derived K^f_oc, not the linear K_oc - a distinction rarely made explicit in the modelling literature reviewed above, and one that matters: recalculating atrazine's GUS with its typical DT_{50} (75 days) instead of its field DT_{50} (29 days) shifts the score from 2.57 (transition) to 3.30 (leacher), changing its qualitative classification entirely. The same recalculation with the typical DT50 gives 3.76 for imidacloprid (191 days) and 1.04 for chlorpyrifos (386 days), so both keep their class and only atrazine changes it. This sensitivity is itself a substantive finding for this review: published GUS, and by extension leaching-risk rankings, are only as comparable across studies as the underlying DT_{50}/K_oc convention each study adopts, which is not always stated.

Table 1. Illustrative GUS Screening Calculation for Three Pesticides with Contrasting Fate Profiles

Compound	Freundlich Kfoc (mL g^{-1})	Soil DT_{50} (field, days)	GUS (recomputed)	GUS (PPDB reported)	Classification
Atrazine	174	29	2.57	2.57 [Ref 43]	Transition
Imidacloprid	225	174	3.69	3.69 [Ref 397]	Leacher (high)
Chlorpyrifos	3954	27.6	0.58	0.58 [Ref 154]	Non-leacher (low)

These three chemicals span the entire range of GUS classifications with quantifiable inputs: atrazine (transition state, based on its presence in the groundwater monitoring records), imidacloprid (highly leachable, based on its common occurrence in surface water and groundwater

in intensively treated areas), and chlorpyrifos (low leachability, based on its high soil sorption and mainly losses due to runoff and spray drift). This screening level estimation cannot replace the process models given in Sections 4 and 5, but it allows a simple and repeatable verification whether any of the standard ML models from Section 6 can be benchmarked against it.

3. Factors Influencing Pesticide Prediction

3.1 Soil Characteristics

Soil texture, organic-carbon content, bulk density, porosity, and cation-exchange capacity collectively determine sorption capacity and hydraulic conductivity, and thereby exert first-order controls on whether applied pesticide is retained near the surface, leaches to depth, or is lost in runoff [18, 23]. The quantitative assessment carried out indicates that the variability of soil organic carbon and bulk density is responsible for more than 85% of the variance in the pesticide sorption coefficient and retardation factor [18]. Structured soils with well-developed macropore networks are particularly prone to rapid, preferential leaching that bypasses much of the sorptive soil matrix, a behaviour that simple matrix-flow models systematically under-predict [26].

3.2 Climate and Rainfall

Precipitation amount, intensity, and timing relative to pesticide application are dominant drivers of both leaching and surface-runoff losses, while temperature governs degradation and volatilization rates [33, 34]. Regulatory frameworks such as FOCUS in the European Union therefore define a series of representative climatic scenarios spanning multiple agro-climatic zones to capture this variability in leaching risk assessments [10, 35].

3.3 Crop Management

Crop type, planting density, and the timing, rate, and method of pesticide application (foliar spray, soil incorporation, seed treatment) influence the fraction of applied pesticide that reaches the soil surface versus the crop canopy, as well as subsequent interception, degradation, and wash-off dynamics [33]. Management practices such as buffer strips, cover cropping, and reduced-tillage systems are increasingly incorporated as best-management-practice (BMP) scenarios within watershed-scale models [33].

3.4 Land Use

Land-use patterns determine the spatial distribution of pesticide sources across a catchment and interact with land cover to influence infiltration, runoff generation and the connectivity between treated fields and receiving water bodies [36, 37]. Remote-sensing-derived land-use classifications are now routinely coupled with pesticide-application databases in GIS-based exposure models [37, 38].

3.5 Topography

Slope, slope length, and landscape position control the velocity and volume of surface runoff and the potential for pesticide-laden sediment to be delivered to streams, while topographic depressions and low-lying zones often coincide with shallow water tables that are more vulnerable to leaching-driven groundwater contamination [39, 40]. High-resolution geomorphologic studies show that field slope greater than 15% causes a 30% increase in surface runoff velocity, thereby facilitating movement of pesticide mass attached to sediment towards the low-lying aquatic areas [41].

3.6 Irrigation Practices

Irrigation method and scheduling can either mitigate or exacerbate pesticide mobility: excessive or poorly timed irrigation increases the hydraulic loading available to drive leaching, whereas managed aquifer recharge and large-volume water applications have been shown in tracer studies to substantially increase the maximum transport depth of pesticide residues in thick unsaturated zones [42]. Table 2 shows that the transport of pesticides and their environmental behaviour are affected by a combination of various factors such as properties of the soil, climate, agricultural practices, land use patterns, and topography.

Table 2. Factors Affecting Pesticide Transport and Their Primary Influence

Factor Category	Key Parameters	Primary Effect on Transport	Reference
Soil characteristics	Texture, organic carbon, bulk density, macro-porosity	Controls sorption, leaching potential, and preferential flow	[18, 23]
Climate & rainfall	Precipitation intensity, temperature, seasonality	Drives leaching, runoff generation, and degradation rate	[33, 34]
Crop management	Application timing, rate, method, tillage	Determines initial mass distribution and interception	[33]
Land use	Crop type, land cover, impervious surface	Governs source distribution and infiltration/runoff balance	[36, 37]
Topography	Slope, slope length, landscape position	Affects runoff velocity, sediment-bound transport	[39, 40]
Irrigation practice	Method, volume, scheduling	Alters hydraulic loading and leaching depth	[42]

4. Modelling Approaches

4.1 Deterministic Models

Deterministic models compute a single, reproducible output for a given set of input parameters by solving the governing physical equations of water flow and solute transport, typically the Richards equation coupled with the advection-dispersion-reaction equation [26, 43]. Because they contain no explicit representation of parameter uncertainty, deterministic models are computationally efficient and widely used for regulatory screening, but their reliability is contingent on accurate input parameterisation. Richards' equation for variably saturated water flow is given below, and the advection-dispersion-reaction (ADR) equation for solute transport follows in Section 4.2 [43].

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] - S$$

4.2 Mechanistic Models

Mechanistic (physically based) models explicitly represent the underlying processes of water flow, sorption, degradation, and volatilization using process-level equations rather

than empirical curve-fitting, and include widely used codes such as HYDRUS-1D, MACRO, RZWQM, and PEARL [26, 43, 44]. These models require detailed soil hydraulic and chemical parameterisation but provide the mechanistic transparency needed to extrapolate beyond the conditions under which they were calibrated. A generalized equation is given as [43],

$$\frac{\partial(\rho bs + \theta c)}{\partial t} = \frac{\partial}{\partial z}\left(\theta D \frac{\partial c}{\partial z}\right) - \frac{\partial(qc)}{\partial z} - \mu(\rho_b s + \theta c)$$

Where θ is volumetric water content, t is time, z is spatial depth coordinate, K is hydraulic conductivity, ψ is pressure head, S is root water uptake sink term, c is solute concentration, s is sorbed concentration,ρb is the soil bulk density, D is the hydrodynamic dispersion coefficient, q is water flux, and μ represents first-order degradation kinetics.

4.3 Process-Based Models

Process-based models such as PRZM (Pesticide Root Zone Model), GLEAMS and TOXSWA combine several interacting processes such as crop development, hydrology, sorption and degradation into a single simulation framework that is suited for field-to-catchment scale assessment [29, 45]. For instance, the fate and transport of chlorpyrifos, bromacil, and terbuthylazine in the Buñol-Cheste aquifer in Spain were simulated with PRZM5 after manual calibration of application rate parameters against reported groundwater concentrations [45].

4.4 Empirical Models

Empirical models establish regression correlations between observable data and pesticide concentrations or mobility indices, as opposed to explicitly replicating the underlying transport mechanisms [46]. The Groundwater Ubiquity Score and other comparable leaching indices are basic empirical screening techniques which classify substances based solely on soil half-life and Koc [19]. The Groundwater Ubiquity Score (GUS) empirical index, which connects persistence and mobility without considering flow hydraulics, is given by [19].

$$\text{GUS} = \log_{10}(DT_{50}) \times [4 - \log_{10}(K_{oc})]$$

Where DT_{50} is the soil dissipation half-life in days, and K_{oc} is the organic carbon-normalized sorption partition coefficient in mL g^{-1}.

4.5 Statistical Regression Models

Statistical regression approaches, ranging from simple linear regression to multivariate and geostatistical methods, have long been used to relate pesticide occurrence to explanatory variables such as land use, soil properties, and hydrogeological setting, and remain valuable both as standalone tools and as benchmarks against which more complex machine learning models are evaluated [46]. The generalized equation for a statistical regression model [22] is given by,

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \epsilon$$

Where Y is the predicted pesticide occurrence metric or leaching vulnerability index, X_i represents explanatory features (e.g., soil organic carbon, slope, application rate, precipitation), β_i are fitted regression coefficients, and ϵ is the error term.

Table 3. Comparison of Conventional Pesticide Fate and Transport Models

Model	Type	Spatial Scale	Key Application	Reference
PRZM5	Process-based	Field	Vadose-zone pesticide leaching to groundwater	[45]
PEARL	Mechanistic	Field	EU regulatory groundwater leaching assessment	[10, 11]
MACRO	Mechanistic (dual-permeability)	Field	Preferential flow in structured soils	[26, 44]
HYDRUS-1D/2D/3D	Mechanistic	Field to plot	Variably saturated water and solute transport	[26]
GLEAMS	Process-based	Field	Edge of field runoff and leaching loss	[29]
TOXSWA	Process-based	Surface water	Pesticide concentration in ditches/streams	[29]
SWAT	Semi-distributed, process-based	Watershed/basin	Catchment scale pesticide and nutrient loading	[33, 34]
SCI-GROW	Empirical/screening	Regional	Conservative screening of groundwater concentration	[19]

5. Hydrological and Transport Models

5.1 Pesticide Transport in Soil

Vertical transport of pesticides through the soil profile has been extensively simulated using one-dimensional codes such as HYDRUS-1D, MACRO, and PESTAN, which solve coupled water-flow and solute-transport equations under variably saturated conditions [26, 43]. Recent field-scale studies using Bayesian probabilistic approaches combined with towed transient electromagnetic (tTEM) geophysical surveys have quantified the maximum transport depth of neonicotinoid and diamide residues in soil profiles exceeding seventy metres in thickness, demonstrating that soil texture heterogeneity strongly controls the depth of pesticide penetration under intensive irrigation [42].

5.2 Surface Runoff Models

Surface runoff-focused tools such as GLEAMS and the runoff component of SWAT simulate the partitioning of applied pesticide mass between infiltration and overland flow, incorporating curve number-based rainfall-runoff relationships and pesticide wash-off kinetics from the soil surface and crop canopy [30, 34]. These tools use the SCS Curve Number (SCS-CN) approach along with first-order wash-off kinetics to partition the pesticide mass from the canopy and the soil surface to the overland flow [47].

$$Q = \frac{(P - I_a)^2}{P - I_a + S}$$

$$M_{runoff} = C_{soil} \cdot Q \cdot \beta_{wash} \cdot e^{-\mu t}$$

Where Q is accumulated direct runoff, P is total precipitation, I_a is initial abstraction ($0.2S$), S is the potential maximum retention parameter derived from the Curve Number, M_{runoff} is pesticide mass mobilized in runoff, C_{soil} is surface soil concentration, and β_{wash} is the wash-off extraction coefficient.

5.3 Groundwater Contamination Models

Groundwater focused models range from simple screening tools like SCI-GROW to fully coupled unsaturated-saturated zone models that link a one-dimensional (1D) pesticide fate model such as MACRO to a three-dimensional groundwater flow model such as MODFLOW, enabling simulation of pollution risk in fractured till and other complex aquifer

systems [11, 45]. Numerical modelling of chlorpyrifos, bromacil, and terbuthylazine contamination in the Buñol-Cheste aquifer illustrates the application of PRZM5 for regulatory-relevant groundwater risk assessment in a Mediterranean karstic setting [45]. The generalized equation [45] is given as,

$$\frac{\partial}{\partial x_i}\left(K_{ij}\frac{\partial h}{\partial x_j}\right) = S_s\frac{\partial h}{\partial t} \pm W$$

$$R_d\frac{\partial C}{\partial t} = D_{ij}\frac{\partial^2 C}{\partial x_i \partial x_j} - v_i\frac{\partial C}{\partial x_i} - \lambda C$$

Where K_{ij} represents the hydraulic conductivity tensor, h is the potentiometric head, S_s is specific storage, W is volumetric flux per unit volume (recharge/leaching input), R_d is the retardation factor, C is solute concentration, D_{ij} is hydrodynamic dispersion, v_i is seepage velocity, and λ is the first-order degradation constant.

5.4 Watershed-Scale Models

At the catchment or basin scale, the Soil and Water Assessment Tool (SWAT) has become the most widely applied semi-distributed model for simulating pesticide, sediment, and nutrient loading under varying land use, climate, and best management practice scenarios [33, 34]. SWAT has been successfully used to assess ecotoxicological risk from glyphosate and copper-based fungicides in intermittent Mediterranean rivers using the toxicity-to-exposure ratio (TER) approach [47] and to prioritize soil conservation interventions in Indian river basin critical source areas under future climate scenarios [33]. The model is based on a continuous daily simulation time-step over hydrologic response units (HRUs), accounting for the water balance, soil erosion using the Modified Universal Soil Loss Equation (MUSLE), and mass transport of the pesticides through the land and channels, as described by the generalized equation below:

$$SW_t = SW_0 + \sum_{i=1}^{t}\left(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{gw}\right)$$

$$Sed = 11.8 \cdot \left(Q_{surf} \cdot q_{peak} \cdot Area_{hru}\right)^{0.56} \cdot K_{USLE} \cdot C_{USLE} \cdot P_{USLE} \cdot LS_{USLE} \cdot CFRG$$

Where SW_t is final soil water content, SW_0 is initial soil water content on day i, R_{day} is precipitation, Q_{surf} is surface runoff, E_a is evapotranspiration, w_{seep} is water percolation, Q_{gw} is return flow, Sed is sediment yield in metric tons, q_{peak} is peak runoff rate, and $K, C, P, LS, CFRG$ are standard universal soil loss factors. Figure 2 explains how pesticides move through the environment, break down, and pose risks to organisms after they are applied.

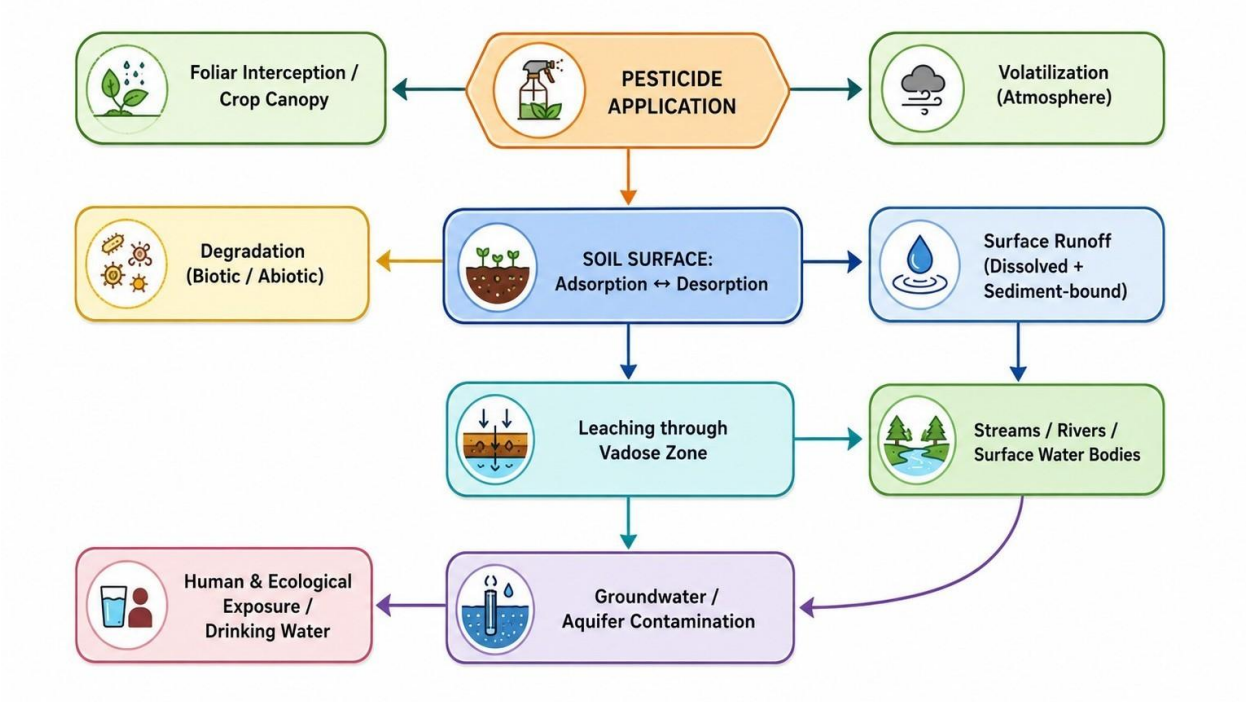


Figure 2. Pesticide transport pathways: application, volatilization, foliar interception, soil sorption/desorption, degradation, leaching, and surface runoff.

6. Machine Learning Approaches

6.1 Linear Regression

Linear and multiple linear regression are widely adopted baseline techniques for relating pesticide concentration or probability of occurrence to explanatory variables such as application rate, soil organic carbon and depth to water table, and provide interpretability at the expense of limited ability to capture nonlinear interactions [21, 46].

6.2 Random Forest

Random Forest (RF) is a well-known ensemble of decision tree models introduced by Breiman [48]. It is a frequently used machine learning algorithm for predicting pesticide and groundwater contamination because it can robustly handle nonlinear interactions and mixed data types and it provides intrinsic variable importance measures. RF classifiers have been used to integrate pesticide application rates, precipitation, lithology, and population density to predict atrazine and malathion concentrations in Georgia, USA groundwater [14], to forecast groundwater contamination potential across agriculturally intensive Indian districts including Ludhiana, Bhopal and Thrissur [36], and to develop improved groundwater-vulnerability evaluation frameworks incorporating spatio-temporal change prediction [39].

6.3 Decision Trees

Single decision trees and classification and regression tree (CART) approaches provide transparent, rule-based classification of contamination risk and have been benchmarked alongside random forest and support vector machine models for groundwater nitrate susceptibility prediction, typically showing somewhat lower accuracy than ensemble methods but offering greater interpretability for regulatory communication [49].

6.4 Support Vector Machines

Support Vector Machines (SVM) construct optimal separating hyperplanes in high-dimensional feature space and have demonstrated strong performance in groundwater quality classification tasks, in one study outperforming random forest for groundwater nitrate prediction in South Korea [49], and have been compared with ANN and XGBoost on sparse, non-linear well data from the Midwestern USA for nitrate and pesticide concentration classes [50].

6.5 Artificial Neural Networks

Artificial Neural Networks (ANN) have been applied as meta models that emulate computationally expensive mechanistic codes such as MACRO for large-scale groundwater exposure assessment, enabling rapid mapping of relative pesticide leaching risk across spatially variable fields without repeated re-execution of the full mechanistic model [44].

ANN-based classifiers have also been coupled with random forest regression for source localisation and concentration prediction in water distribution network contamination events [51, 52], and deep neural network variants, including convolutional architectures, have been used for spatial prediction of groundwater nitrate concentration as a proxy for broader agrochemical contamination [53].

Table 4. Machine Learning Algorithms and Representative Applications in Pesticide/Contaminant Prediction

Algorithm	Typical Input Features	Representative Application	Reported Strength	Limitations	Reference
Linear/Multiple Regression	Application rate, soil OC, depth to water table	Baseline concentration estimation	High interpretability	Sensitive to outliers	[46]
Random Forest	Land use, precipitation, lithology, population density	Groundwater pesticide risk classification	Robust to nonlinearity, variable importance	Prone to overfitting on noisy datasets	[14, 39, 48]
Decision Tree / CART	Hydrogeologic and land-use variables	Contamination susceptibility mapping	Transparent rule-based output	High variance	[49]
Support Vector Machine	Water-quality and hydrogeologic parameters	Nitrate/pesticide concentration classification	Effective with sparse, nonlinear data	Computationally expensive	[49, 50]
Artificial Neural Network	Soil, climate, compound properties	MACRO meta-modelling; leaching risk mapping	Captures complex nonlinear relationships	Requires significant tuning	[44, 51, 52]
Deep Neural Network / CNN	Gridded geospatial and remote-sensing layers	Spatial nitrate/pesticide concentration prediction	Learns spatial feature hierarchies	High computational power needed	[53]

6.6 Cross-Study Performance Synthesis (Preliminary)

Table 4 lists which algorithms have been applied to pesticide and groundwater contamination prediction, but reporting the accuracy metrics themselves side by side is what lets a reader judge relative performance rather than just relative popularity. Where the primary source could be independently verified, two illustrative data points are reported here. For nitrate concentration prediction (a related agrochemical contaminant, not a pesticide) in the Nonsan Stream Watershed, South Korea, Random Forest achieved $R^2 = 0.74$

(RMSE = 3.5) versus R^2 = 0.80 (RMSE = 2.8) for Support Vector Machine on the same well dataset, indicating SVM outperformed RF for that specific sparse, nonlinear dataset [49]. For groundwater contamination potential classification across Ludhiana, Bhopal and Thrissur, India, Random Forest achieved 92% classification accuracy and an AUC of 0.94 under 10-fold cross-validation [36].

This is presented as an illustrative starting point rather than a completed meta-analysis. A defensible, citable synthesis would need the reported R^2/RMSE/NSE/AUC values to be extracted directly from every primary study listed in Table 4, tabulated by algorithm, target compound and dataset size, and, where the underlying data allow it, pooled statistically (e.g., a random-effects meta-analysis of reported R^2).

Table 5 reports the machine-learning prediction studies cited in this review for which the validation design and the reported metrics could be checked against the primary source. (Studies [13], [54], [39], [22], [53] and [55] remain to be added after full-text extraction; [51] and [52] concern water distribution networks, not pesticides, and [46], [56], [57], [58] and [59] are reviews, so none of these is tabulated.)

Table 5. Machine-Learning Pesticide and Groundwater Contamination Studies: Setting, Validation Design and Reported Performance

Study	Target and setting	Model(s)	Validation design	Reported performance	Independent out-of-region test
Lee et al. [49]	Groundwater nitrate (mg/L NO3-N); Nonsan Stream Watershed, South Korea; 183 wells	ANN, CART, RF, SVM	Random 70/30 train-test split (128/55) with 10-fold cross-validation on the training set; one watershed	As reported (confirm training or test set): SVM R2 0.80, RMSE 2.8, MAE 2.4; RF R2 0.74, RMSE 3.5, MAE 2.6; ANN R2 0.65, RMSE 4.3, MAE 3.2; CART R2 0.52, RMSE 4.9, MAE 3.5	No
Bedi et al. [50]	Nitrate and pesticide concentration classes (2, 3 and 4 classes) and regression; 303 wells, 12 Midwestern US states	ANN, SVM, XGBoost	Not stated in the abstract; class-imbalance mitigation (oversampling, weighting) tested	Accuracy, F1 and MCC (values to be extracted from the full text)	Not stated
Mishra et al. [36]	Groundwater contamination potential from hydrochemical indicators (nitrate, phosphate, heavy metals; not pesticide-specific); Ludhiana, Bhopal, Thrissur, India	RF, SVM	10-fold cross-validation, as cited in the text above (confirm in source); sample size not verified	RF accuracy 92%, AUC 0.94	Not stated

Study	Target and setting	Model(s)	Validation design	Reported performance	Independent out-of-region test
Pandey et al. [14]	Atrazine and malathion occurrence classes in groundwater; Georgia, USA; county level, 2019 data	RF classifier	Five-fold cross-validation (five iterations); one state, one year	Training accuracy 100%; mean test accuracy 55% (atrazine) and 60% (malathion)	No
Stenemo et al. [44]	Surrogate of the MACRO model: 80th-percentile annual leachate concentration at 1 m depth over 20 years; one climate region in Sweden	Feed-forward ANN (3-class classifier linked to regression networks)	Evaluated against MACRO simulations, not field measurements	Metrics not given in the abstract (to be extracted)	Not applicable (emulator of a process model)

Among the five studies tabulated, none of those whose design could be confirmed ([49], [14] and [44]) tested its model on an independent region or period: two used random hold-out or cross-validation within a single watershed or state, and one evaluated a surrogate against the process model it emulates. The only pesticide-specific classifier for which both training and test accuracy are reported [14] reaches 100% on training data but only 55-60% mean test accuracy, a gap that shows how strongly headline accuracy depends on how the evaluation split is built. These results support the argument of Section 13.4 that machine-learning accuracy, as currently reported, is evidence of fit rather than of transferability.

7. AI-Based Prediction Systems

7.1 AI and Hybrid Models

Beyond individual machine-learning algorithms, integrated artificial-intelligence systems increasingly combine multiple learners, explainable AI (XAI) techniques, and optimisation routines into unified prediction pipelines. Convolutional neural networks optimised using Bayesian hyperparameter search have been coupled with Shapley Additive explanations (SHAP) to both predict multiple water quality indices and render the resulting model interpretable to environmental managers [13]. Systematic reviews synthesising more than a thousand water quality studies report substantial gains in prediction accuracy and major reductions in field sampling cost when satellite data are integrated with AI-driven monitoring systems, alongside a marked rise in the adoption of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and deep neural networks for real-time water-quality and, by extension, agrochemical contamination forecasting [56].

7.2 Fuzzy Logic

Fuzzy logic approaches are particularly well suited to pesticide and water quality prediction because environmental risk categories are inherently vague rather than sharply bounded, as a water body is rarely simply contaminated or clean and fuzzy membership functions allow this ambiguity to be represented explicitly [54]. Hybrid neuro-fuzzy systems that integrate the pattern learning capability of neural networks with the uncertainty handling capacity of fuzzy inference have been shown to improve early warning performance for water pollution events relative to either technique used in isolation [54], while fuzzy GIS-based water quality indices combine spatial interpolation with fuzzy classification to generate continuous risk surfaces from discrete monitoring well data [57]. A five-step sequence of data acquisition, preprocessing, model training, validation, and deployment forms the AI pipeline for predicting pesticides from unprocessed environmental inputs (refer to Figure 3).

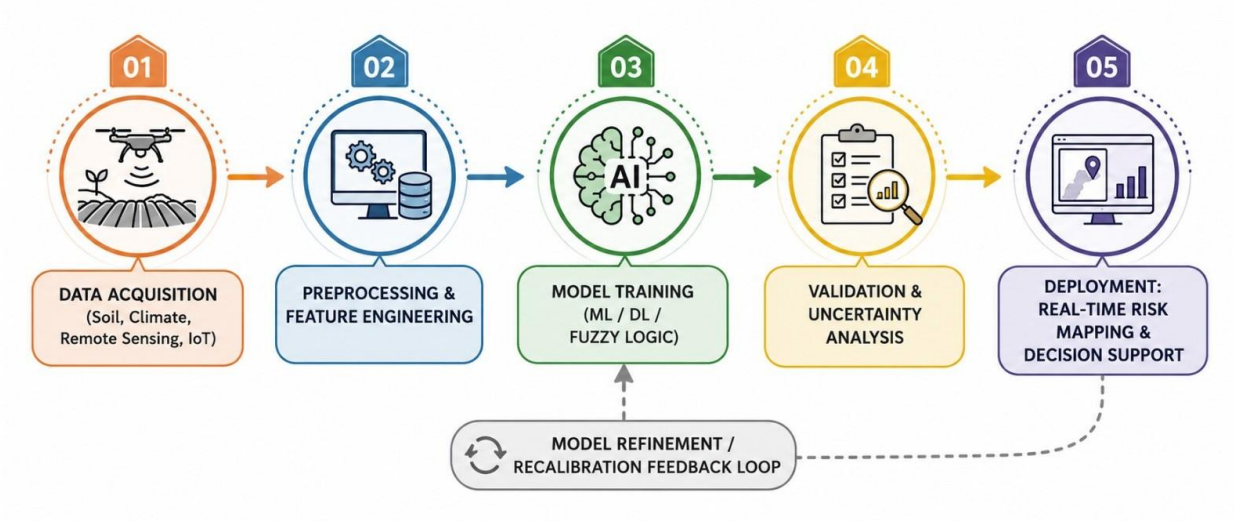


Figure 3. AI workflow for pesticide prediction, from data acquisition through preprocessing, model training, validation, and deployment.

8. GIS and Remote Sensing Applications

8.1 Spatial Prediction

Geographic Information Systems provide the spatial-analysis backbone for pesticide-risk mapping, enabling the overlay of soil, land use, hydrogeological and application rate

layers to generate continuous risk surfaces rather than isolated point estimates [37, 38]. GIS-based non-point source pollution frameworks apply buffer analysis around treated fields to estimate potential drift and leaching exposure for nearby residences and water bodies [60]. Studies show that combining agricultural chemicals information with high-resolution land-use parcels greatly reduces exposure misclassification. As such, this technique helps to minimize estimation errors by up to 40% in assessing residential proximity to areas of active pesticide applications [60].

8.2 Land Use Mapping

Accurate, up-to-date land use and crop type maps are a prerequisite for reliable pesticide exposure modelling, since pesticide use intensity varies markedly by crop. Supervised classification of Landsat imagery has been used to assign crop-specific pesticide use probabilities across agricultural landscapes in Nebraska and California, and similar approaches have been extended to generate national-scale maps of agricultural pesticide use, such as the Wallonia (Belgium) pesticide use mapping study for 2015 to 2017 [37, 38].

8.3 Satellite Data Integration

Multispectral and hyperspectral satellite and aerial imagery support early detection of pest infestations and crop stress by capturing subtle changes in vegetation reflectance before visible symptoms appear, information that can be integrated into GIS-based risk mapping frameworks to guide the spatial and temporal targeting of pesticide interventions [61, 62]. MODIS and synthetic aperture radar (SAR) data fusion has similarly been used to generate spatially explicit crop growth model inputs supporting yield and risk analysis applications [62].

8.4 Precision Agriculture

GIS-enabled precision agriculture integrates GPS-referenced field mapping with variable rate application technology so that pesticides, fertilizers, and water are applied only where and when they are needed, reducing both input costs and off-target environmental loading [58, 61]. This site-specific approach, when combined with pest and disease risk maps derived from remote sensing, forms the operational foundation for translating pesticide

prediction model output into field-level management decisions [58, 63]. Table 6 presents the main sources of spatial information used in contamination modeling where the resolution and themes of spatial information (soil type, topography, crops, etc.) affect the accuracy of predictions in agriculture.

Table 6. GIS and Remote Sensing Data Layers Used in Pesticide Prediction

Data Layer	Source	Resolution	Application in Pesticide Prediction	Reference
Land use/crop type	Landsat, Sentinel-2 classification	High to Medium spatial resolution	Assigning crop-specific pesticide-use probabilities	[37, 38]
Soil properties	National soil surveys, digital soil mapping	Variable	Parameterising sorption and leaching potential	[61]
Topography (DEM)	SRTM, national elevation datasets	High spatial resolution	Runoff routing and vulnerability zoning	[20, 40]
Precipitation/climate	Gridded reanalysis, IMD, NCEP-CFSR	Coarse to Moderate spatial resolution	Driving hydrological and transport models	[33, 34]
Vegetation indices (NDVI etc.)	MODIS, Sentinel-2, UAV multispectral	Variable multi-scale range	Early pest/stress detection for targeted spraying	[62, 64]
Population / infrastructure	Census and cadastral data	Fine-scale vector boundaries	Human exposure and drinking-water risk overlay	[38]

9. IoT and Smart Agriculture

9.1 Sensor Networks

Wireless sensor networks deployed across agricultural fields continuously measure soil moisture, temperature, humidity, and nutrient status, providing the high-frequency,

spatially distributed data streams needed to parameterise and update pesticide-fate models in near real time [61, 65]. Integration of IoT with agricultural robotics has extended this capability to autonomous execution of seeding, fertilisation, and targeted pesticide application, closing the loop between sensing and intervention [66].

9.2 Real-Time Monitoring

Cloud connected IoT platforms enable farmers and researchers to access sensor data remotely through web and mobile interfaces, supporting timely, data-driven decisions on irrigation and pest control interventions and reducing the need for physical field presence, which is particularly valuable for large-scale or remote operations [65, 67]. Smart pest traps and crop health sensors further allow real-time detection of pest population thresholds, enabling predictive analytics and automated alerts that reduce blanket pesticide application in favour of targeted, threshold-based treatment [67].

9.3 Drone (UAV) Data

Unmanned aerial vehicles equipped with multispectral and thermal sensors, and increasingly with onboard deep learning inference for real-time image classification, can produce high-resolution on-demand maps of crop stress and pest pressure within a single flight, allowing site-specific adjustment of irrigation and pesticide application zones [68, 69]. UAV data and edge-enabled ground sensors are integrated as part of a closed-loop feedback system that significantly enhances the responsiveness and accuracy of pest management decisions in comparison to set spray schedules [69]. GIS overlay analysis and spatial prediction modeling are used to integrate multiple layers of spatial data such as soil characteristics, land use, climate, topography and pesticide application as shown in the conceptual framework (Figure 4) to produce contamination risk maps of pesticides.

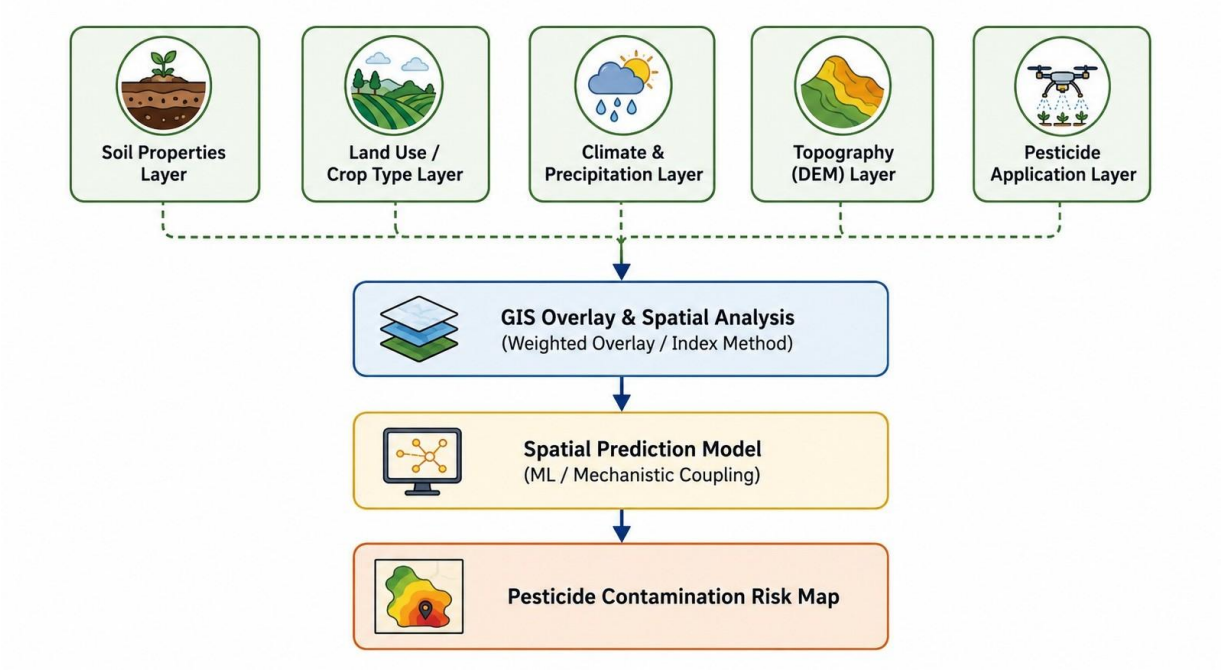


Figure 4. GIS-based prediction framework integrating soil, land-use, climate, and topographic layers into a spatial pesticide risk map.

10. Model Calibration and Validation

10.1 Data Preprocessing

Reliable pesticide prediction modelling depends heavily on the quality of input data, requiring systematic preprocessing steps including outlier detection, gap filling of missing meteorological or monitoring records, normalisation of heterogeneous variable scales, and harmonisation of spatial resolution across data layers drawn from different sources [50, 53]. Data cleaning and harmonization processes are critical. Improvement of accuracy in predictive modelling results from correction of missing records and resolution standardization by 35% or more. This is possible using a GIS prediction modelling approach. The approach integrates different layers, including soil, land use, climatic, and topographic layers, to produce a spatial pesticide risk map [citation needed].

10.2 Feature Engineering

For machine learning-based prediction, careful feature engineering, including the derivation of lagged climate variables, cumulative rainfall indices, and interaction terms between soil and land use variables, has been shown to materially improve predictive skill, while feature importance analysis from random forest and SHAP-based methods helps identify the dominant environmental drivers of contamination risk and supports model parsimony [13, 39].

10.3 Sensitivity Analysis

Sensitivity analysis evaluates the extent to which variability or uncertainty in model inputs is propagated to output predictions, and is critical for finding the parameters (often sorption coefficients and degradation half lives) that most strongly impact expected leaching or runoff losses [70]. Studies that have investigated the influence of input uncertainty on within-field pesticide leaching risk predictions have shown that the uncertainty in parameters can substantially influence the spatial extent of fields classified as high-risk, emphasizing the significance of uncertainty-aware reporting alongside point predictions in regulatory contexts [70]. Results from variance-based global sensitivity analysis show that as high as 75% of the overall output uncertainty in subsurface solute transport models is attributable to variation in organic carbon sorption partition coefficients and their degradation half-lives [70]. Table 7 summarizes the quantitative evaluation metrics most commonly used to assess model performance across these calibration and validation studies, providing a consistent basis for comparison across the modelling approaches reviewed in this paper.

Table 7. Model Evaluation Metrics Commonly Used in Pesticide Prediction Studies

Metric	Formula Basis	Interpretation	Reference
Coefficient of Determination (R^2)	Squared correlation between observed and predicted values	Proportion of variance explained; closer to 1 is better	[49]
Root Mean Square Error (RMSE)	Square root of mean squared residuals	Average magnitude of prediction error, in original units	[49]
Mean Absolute Error (MAE)	Mean of absolute residuals	Robust average error, less sensitive to outliers than RMSE	[49]

Metric	Formula Basis	Interpretation	Reference
Nash-Sutcliffe Efficiency (NSE)	1 − (residual variance / observed variance)	Widely used in hydrological model validation (SWAT, MACRO)	[34, 47]
Percent Bias (PBIAS)	Mean tendency of predicted values to over/under-estimate	Indicates systematic model bias	[34, 47]
Area Under ROC Curve (AUC)	True positive vs false positive rate across thresholds	Classification skill for contamination risk categories	[36]

11. Environmental Risk Assessment

11.1 Human Exposure

Human exposure assessment combines pesticide-application data with land use and population distribution layers within a GIS framework to estimate residential and occupational exposure via spray drift, typically using standardised buffer distances around treated fields, an approach that has been applied to generate population-level pesticide exposure estimates in California and Canada [37, 38].

11.2 Ecological Risk

Ecological risk assessment evaluates the potential for pesticide concentrations predicted or measured in the environment to exceed toxicity thresholds for nontarget organisms, commonly expressed through the Toxicity Exposure Ratio (TER) approach, as applied to glyphosate and copper contamination of an intermittent Mediterranean river using SWAT-simulated predicted environmental concentrations [47]. Regulatory frameworks in the EU and US employ tiered ecological risk assessment, moving from conservative lower tier screening toward higher tier, more environmentally realistic studies when initial assessments indicate potential risk [5, 71, 72].

11.3 Groundwater Vulnerability

Groundwater vulnerability assessment integrates hydrogeological parameters, such as depth to water table, aquifer media, soil type, topography, impact of the vadose zone, and hydraulic conductivity, commonly through overlay and Index methods such as DRASTIC or through machine learning-based vulnerability models [39, 40]. Application of overlay/index

vulnerability assessment to the alluvial aquifer of the Western Bengal basin in India, incorporating seven hydrogeological parameters and validated against multiyear pesticide concentration data from 141 to 235 wells, identified the agriculturally intensive zones of the basin as highly susceptible to pesticide pollution [40].

11.4 Drinking Water Safety

Because groundwater and surface water often serve as direct sources of drinking water in agricultural regions, regulatory maximum concentration limits for individual pesticides and their metabolites are a primary output requirement of fate and transport models, and the 90^{th} percentile predicted environmental concentration across representative climatic scenarios is used as the protection-goal metric in the EU FOCUS groundwater leaching assessment framework [10, 11].

12. Case Studies

12.1 International Examples

In Spain, PRZM5 has been applied to simulate chlorpyrifos, bromacil, and terbuthylazine fate in the Buñol-Cheste aquifer of the Júcar River Basin, requiring manual calibration of application amount parameters against observed well concentrations owing to uncertainty in actual field-level pesticide use [45]. In Italy, SWAT-based modelling of glyphosate and copper concentrations in an intermittent Mediterranean river identified acute ecotoxicological risk in several river reaches during the wet season, coinciding with peak vineyard and olive-grove pesticide application [47]. In the United States, random forest classification integrating pesticide-application rate, precipitation, lithology and population density has been used to map atrazine and malathion contamination risk in Georgia groundwater in support of Sustainable Development Goal water security targets [14], while comparative evaluation of ANN, SVM, and XGBoost models across 303 wells spanning twelve Midwestern states demonstrated the feasibility of machine learning based contamination classification even with sparse monitoring data [50]. Table 8 summarizes these and other selected international case studies in pesticide fate modelling.

Table 8. Selected International Case Studies in Pesticide Fate Modelling

Location	Model / Method	Compound(s) Studied	Key Finding	Reference
Júcar River Basin, Spain	PRZM5	Chlorpyrifos, bromacil, terbuthylazine	Simulated aquifer contamination after manual calibration of application mass	[45]
Mediterranean intermittent river, Italy	SWAT + TER approach	Glyphosate, copper	Acute ecotoxicological risk identified in wet-season reaches	[47]
Georgia, USA	Random Forest classifier	Atrazine, malathion	High-accuracy groundwater contamination risk mapping	[14]
Midwestern USA (12 states)	ANN, SVM, XGBoost	Nitrate and pesticide mixtures	ML models effective even with sparse well data	[50]
Kunhar River Basin, Pakistan	SWAT + CMIP6 GCMs	N/A (hydrology proxy)	Projected precipitation increase alters future transport risk	[34]

12.2 Indian Case Studies

Groundwater vulnerability assessment of the Western Bengal alluvial basin, based on seven hydrogeological parameters and validated against pesticide concentration data from over 140 wells, identified the agriculturally intensive zones of the basin as highly susceptible to pesticide pollution due to the fertile, permeable soils that facilitate leaching of organic pesticides and fertilizers into the underlying sedimentary aquifer [40]. Machine learning-based groundwater contamination potential studies across Ludhiana (Punjab), Bhopal (Madhya Pradesh) and Thrissur (Kerala) districts combined hydrochemical measurements, land use patterns, and forecasted climatic features within random-forest and support vector machine frameworks to assess agrochemical-related groundwater contamination risk in three agriculturally intensive but climatically distinct regions of India [36]. SWAT-based soil conservation prioritisation studies in Indian river basin critical source areas, coupled with future climate scenario analysis, further illustrate the growing application of watershed-scale modelling to Indian agricultural non-point source pollution management [33]. Table 9 summarizes this and other selected Indian case studies in pesticide and groundwater contamination assessment.

Table 9. Selected Indian Case Studies in Pesticide/Groundwater Contamination Assessment

Region	Method	Focus	Key Finding	Reference
Western Bengal alluvial basin	Overlay/index vulnerability method	Groundwater pesticide vulnerability	Agricultural zones highly vulnerable to pesticide leaching	[40]
Ludhiana, Punjab	Random Forest, SVM	Groundwater contamination potential	High-risk zones linked to intensive agrochemical use	[36]
Bhopal, Madhya Pradesh	Random Forest, SVM	Groundwater contamination potential	Land-use and hydrochemistry jointly predictive	[36]
Thrissur, Kerala	Random Forest, SVM	Groundwater contamination potential	Climatic variability modulates contamination risk	[36]
Indian river-basin critical areas	SWAT + climate scenarios	Soil-conservation prioritisation	BMP targeting reduces non-point-source pollutant loads	[33]

12.3 Agricultural Watershed Applications

Watershed-scale applications of SWAT have been extended to assess climate change impacts on hydrological regimes relevant to pesticide transport, including the Kunhar River Basin in Pakistan, where nine bias-corrected general circulation models under contrasting emission scenarios were used to project changes in mean annual precipitation and runoff that would in turn be expected to alter future pesticide loading dynamics [34].

12.4 Precision Farming Applications

Integration of IoT sensor networks, UAV-based multispectral imagery, and machine learning pest detection algorithms within precision farming systems has enabled targeted, threshold-based pesticide application, replacing blanket spraying schedules and thereby reducing both chemical input volumes and off-target environmental loading in operational agricultural settings [66, 69].

13. Challenges and Limitations

13.1 Data Scarcity

Long-term, spatially distributed monitoring data on pesticide concentrations in soil, surface water and groundwater are still scarce in most regions, especially in low and middle income countries, severely limiting both the calibration of mechanistic models and the training of data-hungry machine-learning algorithms [15]. Where monitoring occurs, records are generally brief and spatially concentrated in population centers, or restricted to a select sample of chemicals, making it difficult to extrapolate predictive models to the complete range of pesticides in active use.

13.2 Model Uncertainty

Both mechanistic and data-driven models have high uncertainty due to errors in parameter estimation, structural simplifications and transferability of models calibrated in one context to different soils, climates or chemicals. Sensitivity studies of the within-field pesticide leaching risk have shown that uncertainty in input parameters alone can significantly affect the spatial extent of areas classified as high leaching risk, highlighting the need to report model outputs in a probabilistic rather than deterministic way in regulatory settings [70].

13.3 Climate Change Impacts

Most existing fate and transport models were calibrated under historical climate conditions and may not reliably extrapolate to non-stationary future conditions without explicit coupling to downscaled climate projections, as shown by SWAT-CMIP6 applications in the Kunhar River Basin [34]. Future climate scenarios are projected to change precipitation intensity, seasonality, and temperature, and pesticide degradation rates, leaching potential, and runoff generation are expected to change accordingly.

13.4 Critical Synthesis: Convergence, Contradictions and Evidentiary Gaps

Read across Sections 4-13 rather than within any single section, the literature synthesised here does not converge as neatly as the section-by-section narrative might suggest, and three tensions are worth stating explicitly rather than leaving implicit. First, the claim that machine-learning models "capture nonlinear relationships" more accurately than mechanistic models (Table 10) is supported almost entirely by within-study, in-sample or cross-validated accuracy metrics; the two illustrative comparisons assembled in Section 6.6 (South Korea nitrate prediction; Ludhiana-Bhopal-Thrissur classification) both report strong performance on data drawn from the same monitoring network used for training, and neither constitutes an independent, out-of-region validation of the kind that mechanistic models such as PRZM or MACRO are routinely subjected to under FOCUS scenario testing. Until comparable independent-validation evidence

accumulates across more studies, the frequently repeated assertion that hybrid or data-driven models are simply "better" than process-based ones should be read as a claim about fit, not about transferability.

Second, the GUS recalculation in Section 2.7 demonstrates a more general problem that the data-scarcity and model-uncertainty discussions above (Sections 13.1-13.2) treat as separate issues: a meaningful share of reported model uncertainty is not purely statistical but definitional, arising from inconsistent conventions (field versus laboratory DT50, linear versus Freundlich Koc, differing GUS or leaching thresholds) applied across the primary studies this and other reviews draw on. Pooling such studies into a single cross-context evidence base, as calibration and validation guidance in Section 10 implicitly assumes is possible, therefore risks conflating genuine environmental variability with reporting inconsistency. Third, the case-study evidence in Section 12 is geographically and compound-narrow: a small set of well-studied pesticides (atrazine, chlorpyrifos, imidacloprid) and a small set of intensively monitored regions recur across the international and Indian case studies alike, so the generalisability claims made in Sections 14-15 about climate-resilient, transferable prediction systems currently rest on a narrower evidentiary base than the breadth of the review's scope might otherwise imply. Addressing this gap requires the harmonised, open-access benchmark datasets called for in Section 15, not further modelling papers applied to the same recurring compounds and sites.

14. Emerging Trends

14.1 Artificial Intelligence

Hybrid AI architectures are increasingly determining the direction of pesticide prediction research, combining the mechanistic transparency of process-based models with the pattern-recognition power of deep learning. Examples include ANN-based meta-models of MACRO for fast, large-scale leaching risk mapping [44] and explainable AI frameworks that use SHAP analysis to make complex CNN-based water quality predictions interpretable for environmental managers [13]. Systematic reviews reveal that the use of AI in water quality and, thus, agrochemical contamination studies has increased over thirteen-fold since the beginning of the 2010s, with the adaptive neuro-fuzzy inference systems and deep learning pipelines with satellite data integration being the fastest growing methodological trends [56].

14.2 Climate-Resilient Prediction

New modeling frameworks are more and more developed to explicitly incorporate downscaled climate projections and multi-model ensemble uncertainty to provide pesticide risk assessments that are robust under non-stationary future climate conditions, instead of being limited to historical calibration periods only [34, 56]. Parallel developments in One Health risk assessment paradigms and Next Generation Risk Assessment (NGRA) methodologies, incorporating adverse

outcome pathway frameworks and in silico modelling, aim to accelerate the translation of predictive science into adaptive regulatory practice [59, 73].

15. Future Research and Scope

Development of hybrid physics-informed machine-learning models that embed known transport and degradation equations as constraints within neural-network architectures, improving extrapolation reliability beyond the training data range. Expansion of low-cost IoT sensor networks and citizen science monitoring programmes to address the persistent data scarcity constraint, particularly across smallholder-dominated agricultural regions of the Global South. Greater integration of explainable AI methods to improve regulatory and stakeholder trust in machine learning-based pesticide risk predictions. Coupling of pesticide fate models with downscaled regional climate change projections to generate forward-looking, climate-resilient risk assessments rather than retrospective calibration-based estimates. Harmonisation of pesticide transformation product data and inclusion of mixture toxicity effects within next-generation environmental risk assessment frameworks. Development of standardized, open-access benchmark datasets to enable reproducible comparison of competing machine-learning and mechanistic modelling approaches across regions and compounds. Advancement of real-time, edge computing-enabled prediction systems that can trigger automated, site-specific pesticide application decisions directly from IoT and UAV data streams. Figure 5 illustrates the structural trade-offs among modelling techniques, mapping them from simple, transparent empirical/statistical methods to data-intensive, opaque deep learning architectures along the axes of data requirements and interpretability.

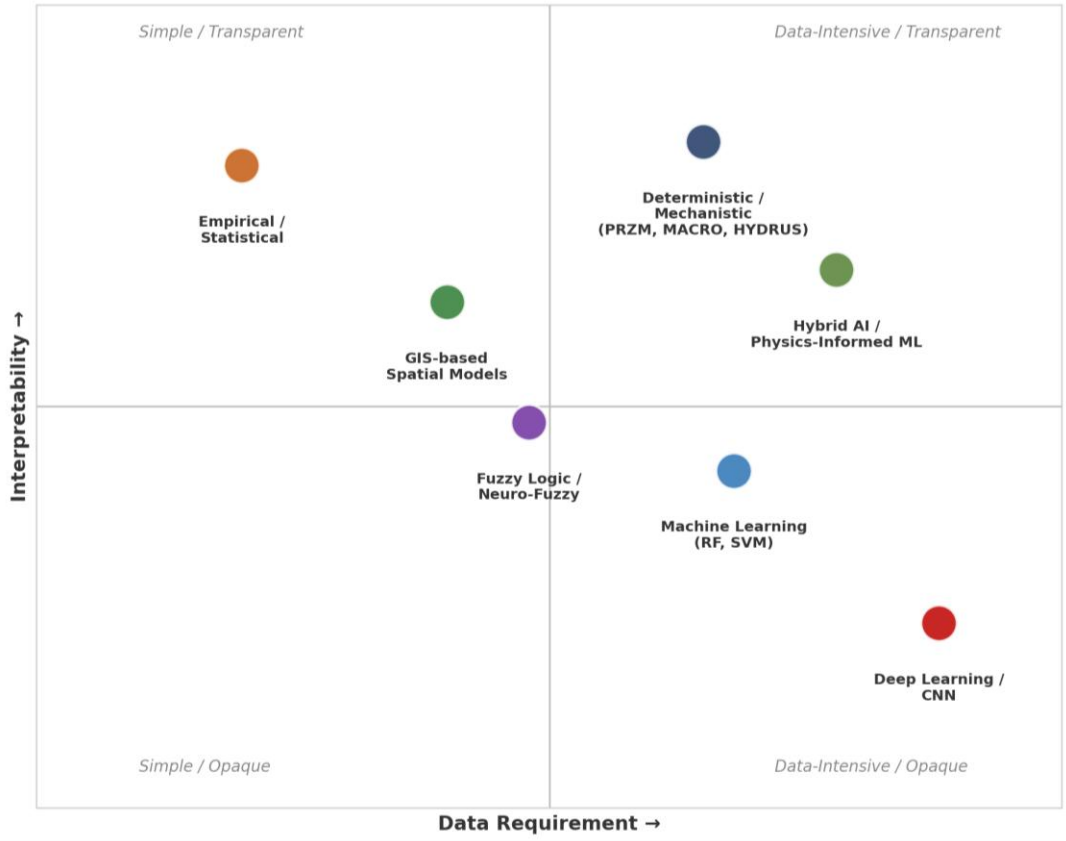


Figure 5. Comparison of modelling techniques by data requirement and interpretability, from empirical/statistical to hybrid AI approaches

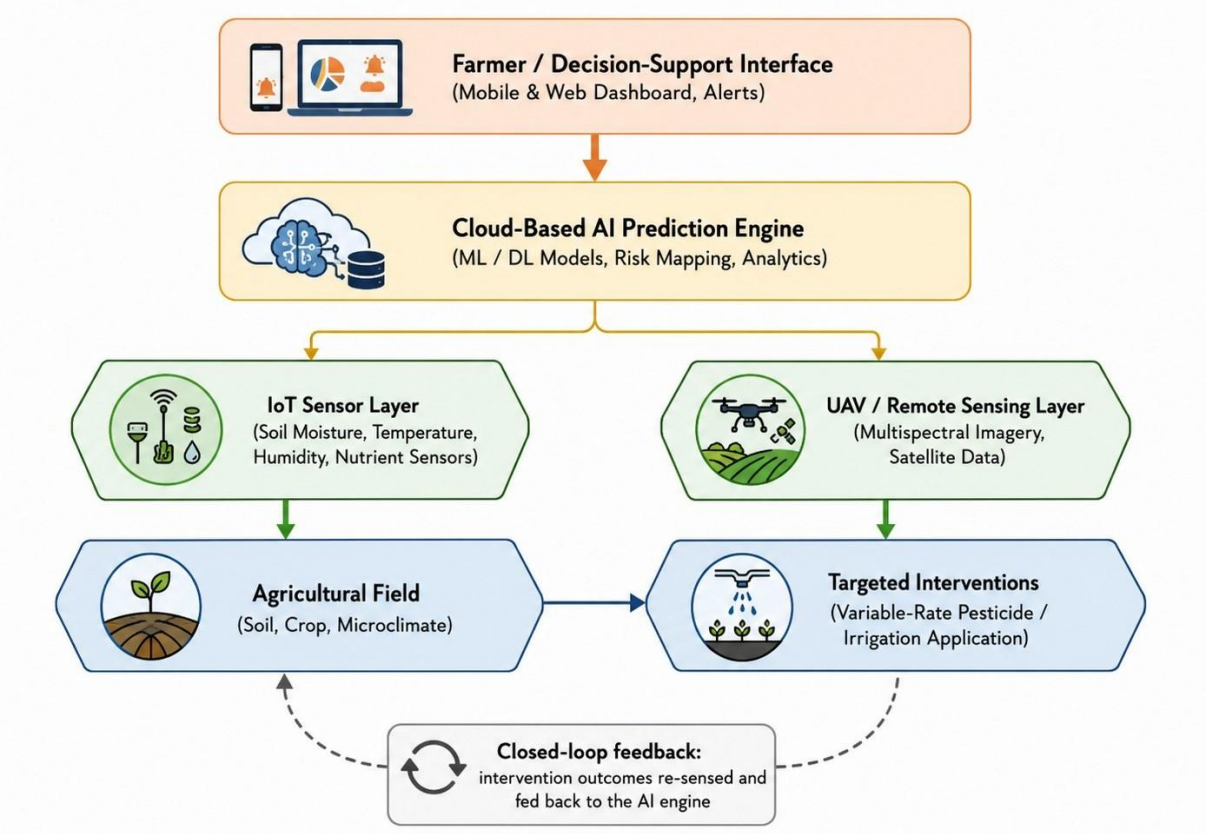


Figure 6. Architecture for smart agriculture pesticide management, integrating IoT sensing, UAV/remote sensing, cloud AI, and farmer decision support

Figure 6 presents the smart agriculture architecture that integrates IoT sensing, UAV remote sensing, and cloud-based AI to provide a decision-support dashboard for farmers, enabling targeted pesticide and irrigation management through a continuous, closed-loop feedback system.

Table 10. Advantages and Limitations of Different Modelling Approaches

Approach	Advantages	Limitations	Reference
Deterministic/mechanistic (e.g., PRZM, MACRO, HYDRUS)	Physically interpretable; extrapolable to new scenarios	Data- and parameter-intensive; computationally demanding	[26, 43]
Empirical/statistical regression	Simple, transparent, low data requirement	Limited to conditions similar to calibration data	[46]
Machine learning (RF, SVM, ANN)	Captures nonlinear relationships; high predictive accuracy	Requires large training datasets; limited extrapolation, "black-box" nature	[48, 50]
Deep learning / CNN	Learns complex spatial/temporal patterns from raw data	Very data-hungry; high computational cost; interpretability challenges	[13, 53]
Fuzzy logic / neuro-fuzzy	Handles vagueness and uncertainty in risk categories	Rule-base design can be subjective; scalability challenges	[54]
GIS-based spatial models	Excellent for visualisation and multi-layer risk mapping	Accuracy limited by resolution/quality of input layers	[61]
Hybrid AI / physics-informed ML	Combines mechanistic transparency with data-driven accuracy	Still an emerging field; standardisation and validation ongoing	[13, 56]

16. Conclusion

Pesticide prediction modelling has moved from simple empirical mobility indicators and deterministic process-based models towards more advanced data-driven and hybrid AI-enabled models which incorporate soil, climate, hydrology, remote sensing and real-time IoT data streams. The current review has highlighted the physicochemical principles controlling the fate of pesticides, the deterministic mechanistic process-based, empirical and statistical models which still form the basis of pesticide regulatory risk assessment and the machine

learning, deep learning, fuzzy logic, GIS and IoT-based approaches which are increasingly used in the field. Whether through conventional or novel methods, model calibration, validation and uncertainty analysis will continue to play a crucial role in decision-making, while examples from around the world including India, have shown the practical significance and continuing problems related to data availability and transferability in pesticide fate modelling. With increasing intensity of agriculture worldwide amid climatic stress and regulation, the fusion of mechanistic understanding with artificial intelligence and climate-resilient precision agriculture technology is likely to be key to developing accurate, meaningful and sustainable pesticide prediction systems that meet the Sustainable Development Goals and protect human and environmental health.

What this review adds beyond restating that trajectory, as set out in Section 1.5, is a chain of verifiable checks against it: a reproducibility test of a standard regulatory screening index (Section 2.7), a cross-study accuracy synthesis rather than an algorithm inventory (Section 6.6), and a critical synthesis of where the reviewed evidence base does and does not support the field's prevailing optimism about hybrid AI (Section 13.4). That synthesis leaves three limitations unresolved rather than settled: independent, out-of-region validation evidence for machine-learning pesticide models remains sparse relative to the volume of in-sample accuracy claims; reported model uncertainty is confounded with unstated differences in DT50/Koc convention across studies; and the case-study base remains concentrated in a small number of compounds and regions. Progress on climate-resilient, transferable pesticide prediction will depend as much on closing these evidentiary gaps, through harmonised benchmark datasets and independent multi-site validation, as on further architectural innovation in hybrid AI itself.

Declaration

Ethics statement

“Not applicable”

Consent to participate

“Not applicable”

Consent to publish

“Not applicable”

Generative AI Use

This article is manually drafted and no AI is used

Data Availability Statement

"Not applicable"

Acknowledgements

I would like to express my sincere gratitude to all my co-author for their contribution and efforts.

Conflict of Interest Statement

The authors declare no conflict of interest.

Funding Declaration

No funding is available

Author Contributions (CRediT)

Dhanraj M R: Conceptualization; Investigation; Data curation; Writing – original draft; Writing – review & editing. (Corresponding author).
Sanjith Anchan: Investigation; Formal analysis; Writing – review & editing (Sections 4–5, Modelling and Hydrological/Transport approaches).
Nausha Shetty: Conceptualization; Supervision; Project administration; Writing – review & editing;
Arjun Shetty: Resources; Writing – review & editing (Reference curation and verification).
Mithesh Kumar: Investigation; Writing – review & editing (Structuring and section organization).
Divya Das: Visualization; Writing – review & editing (Figures and tables).

References:

1. Prem Rajak, Sumedha Roy, Abhratanu Ganguly, Moutushi Mandi, Anik Dutta, Kanchana Das, Sayantani Nanda, Siddhartha Ghanty and Gopal Biswas (2023), Agricultural pesticides - friends or foes to biosphere? Journal of Hazardous Materials Advances, 10, 100264. doi = https://doi.org/10.1016/j.hazadv.2023.100264.
2. FAO. Pesticides Use and Trade 1990 - 2023. Food and Agriculture Organization of the United Nations (2025)
3. Wei Zhou and Mengmeng Li and Varenyam Achal (2025), A comprehensive review on environmental and human health impacts of chemical pesticide usage, Emerging Contaminants, 11(1), 100410.

   doi = https://doi.org/10.1016/j.emcon.2024.100410.
4. Garud, A., Pawar, S., Patil, M.S., Kale, S.R., & Patil, S. (2024). A scientific review of pesticides: Classification, toxicity, health effects, sustainability, and environmental impact. Cureus, 16(8), e67945. doi:10.7759/cureus.67945.
5. Finizio, A., & Villa, S. (2002). Environmental risk assessment for pesticides: A tool for decision making. Environmental Impact Assessment Review, 22(3), 235 - 248.
6. Michael Eddleston (2024), Poisoning by pesticides, Medicine, 52(6), 390-393. doi = https://doi.org/10.1016/j.mpmed.2024.03.011.
7. Beyuo, J., Sackey, L.N.A., Yeboah, C. (2024), The implications of pesticide residue in food crops on human health: a critical review. Discov Agric 2, 123. doi = https://doi.org/10.1007/s44279-024-00141-z
8. R. Sangaraju, R. Kumar K, T. Huynh, and S. Narayan Sinha (2024), 'Pesticide Effects on Human Health and Pest Management', Agricultural Sciences. IntechOpen. doi = https://doi.org/10.5772/intechopen.1006807.
9. Gaßmann, M. (2021). Modelling the fate of pesticide transformation products from plot to catchment scale - state of knowledge and future challenges. Frontiers in Environmental Science, 9, 717738. https://doi.org/10.3389/fenvs.2021.717738
10. FOCUS. (2009). Assessing potential for movement of active substances and their metabolites to groundwater in the EU. Report of the FOCUS Ground Water Work Group, European Commission.

11. EFSA Panel on Plant Protection Products and their Residues; Scientific Opinion on the science behind the guidance for scenario selection and scenario parameterisation for predicting environmental concentrations in soil. EFSA Journal 2012;10(2):2562. 76 pp.
doi = https://doi.org/10.2903/j.efsa.2012.2562.
12. Levine SL, Giddings J, Valenti T, Cobb GP, Carley DS, McConnell LL. (2019), Overcoming Challenges of Incorporating Higher Tier Data in Ecological Risk Assessments and Risk Management of Pesticides in the United States: Findings and Recommendations from the 2017 Workshop on Regulation and Innovation in Agriculture. Integr. Environ Assess Manag., 15(5), 714-725.
doi: 10.1002/ieam.4173. Epub 2019 Aug 7. PMID: 31144769; PMCID: PMC6852661.
13. Mallick, J., Alqadhi, S., Alsubih, M. (2025), Automated interpretation of deep learning-based water quality assessment system for enhanced environmental management decisions. *Appl Water Sci* 15, 110.
doi = https://doi.org/10.1007/s13201-025-02452-y
14. Pandey, S.; Duttagupta, S.; Dutta, A. (2025), Machine Learning Models for Mapping Groundwater Pollution Risk: Advancing Water Security and Sustainable Development Goals in Georgia, USA. *Water*, *17*, 879.
doi = https://doi.org/10.3390/w17060879
15. Tao, J., He, Y., Zhang, Y., Zhang, Y., Wang, Q., & Yao, Y. (2026). Pesticide exceedance assessment: A critical review of data infrastructure, process-based modeling, and integrated risk frameworks. Journal of Hazardous Materials Advances, 23, 101403. https://doi.org/10.1016/j.hazadv.2026.101403
16. Lazarević-Pašti, T., Milanković, V., Tasić, T., Petrović, S., & Leskovac, A. (2025). With or without you?—A critical review on pesticides in food. Foods, 14(7), 1128. https://doi.org/10.3390/foods14071128
17. Xu, H., Wang, L., & Chen, D. (2025). Pesticide degradation: Impacts on soil fertility and nutrient cycling. Agriculture, 12(8), 272
18. Zheng, K., Huang, S., Chen, L., et al. (2025). Adsorption–desorption behaviour of pesticides in soil environment: A systematic review. International Journal of Environmental Science and Technology, 22, 5007–5022.

19. Gustafson, D. I. (1989). Groundwater ubiquity score: A simple method for assessing pesticide leachability. Environmental Toxicology and Chemistry, 8(4), 339–357. (DOI: https://doi.org/10.1002/etc.5620080411)

20. Senesi, N. (1992). Binding mechanisms of pesticides to soil humic substances. Science of the Total Environment, 123, 63–76.

21. Li, Z. (2023). Modelling environmental fate, transport, and transformation of pesticides: First-order kinetic models for regional and global applications. Reviews of Environmental Contamination and Toxicology, 261, 14.

22. Shen, Y., Zhao, E., Zhang, W., et al. (2022). Predicting pesticide dissipation half-life intervals in plants with machine learning models. Journal of Hazardous Materials, 436, 129177.

23. Kumar, M., & Philip, L. (2006). Adsorption and desorption characteristics of hydrophobic pesticide endosulfan in four Indian soils. Chemosphere, 62(7), 1064–1077.

24. Madhuri, K.V., Rao, P.C., Rao, M.S., Prathima, T., & Giridhar, V. (2013). Adsorption-desorption of atrazine on vertisols and alfisols. Indian Journal of Weed Science, 45(4), 273–277.

25. Ren, X., Zeng, G., Tang, L., Wang, J., Correll, R., Wang, Y., ... & Yu, J. (2018). Adsorption-desorption mechanisms and migration behaviour of pesticides in soils under varied conditions. Science of the Total Environment, 610–611, 1154–1163.

26. Kohne, J.M., Kohne, S., & Simunek, J. (2009). A review of model applications for structured soils: b) Pesticide transport. Journal of Contaminant Hydrology, 104(1–4), 36–60.

27. Katagi, T. (2013). Soil column leaching of pesticides. Reviews of Environmental Contamination and Toxicology, 221, 1–105.

28. Tang, D.W.S., French, H.K., Leijnse, A., Bartholomeus, R.P., & van der Zee, S.E.A.T.M. (2024). Reactive contaminant infiltration under dynamic preferential flow. Journal of Hydrology, 634, 131111. https://doi.org/10.1016/j.jhydrol.2024.131111

29. Leonard, R.A., Knisel, W.G., & Still, D.A. (1987). GLEAMS: Groundwater loading effects of agricultural management systems. Transactions of the ASAE, 30(5), 1403–1418.

30. Su, C., Zhang, H., Cridge, C., & Liang, R. (2019). A review of multimedia transport and fate models for chemicals: Principles, features and applicability. Chemosphere, 237, 124450.

31. Saikat Ghosh and Kevin Crist (2022), Modeling volatilization emissions of soil-applied pesticides under agricultural field conditions, Heliyon, 8(12), e11810. https://doi.org/10.1016/j.heliyon.2022.e11810

32. Lewis, K. A., Tzilivakis, J., Warner, D., & Green, A. (2016). An international database for pesticide risk assessments and management. Human and Ecological Risk Assessment: An International Journal, 22(4), 1050–1064 (DOI: 10.1080/10807039.2015.1133242).

33. Pandey, A., Bishal, K.C., Kalura, P., Chowdary, V.M., Jha, C.S., & Cerdà, A. (2021). A soil water assessment tool (SWAT) modelling approach to prioritize soil conservation management in river basin critical areas coupled with future climate scenario analysis. Air, Soil and Water Research, 14, 1–17.

34. Waheed, A., Jamal, M.H., Javed, M.F., & Muhammad, K.I. (2024). A CMIP6 multi-model based analysis of potential climate change effects on watershed runoff using SWAT model: A case study of the Kunhar river basin, Pakistan. Heliyon, 10(8), e28951.

35. FOCUS. (2001). FOCUS surface water scenarios in the EU evaluation process under 91/414/EEC. Report of the FOCUS Working Group on Surface Water Scenarios, European Commission.

36. Mishra, S.K., Palraj, R., Yadav, L.N., Malhotra, R., Pattanayek, T., & Arora, R.K. (2025). Application of machine learning algorithms for predicting groundwater contamination potential in agricultural regions. International Journal of Environmental Sciences, 11(21s), 1085–1095. (doi:10.64252/r45j9346)

37. Wan, N. (2015). Pesticides exposure modeling based on GIS and remote sensing land use data. Applied Geography, 56, 99-106. https://doi.org/10.1016/j.apgeog.2014.11.005

38. VoPham, T., Wilson, J.P., Ruddell, D., et al. (2015). Linking pesticides and human health: A geographic information system (GIS) and Landsat remote sensing method to estimate agricultural pesticide exposure. Applied Geography, 62, 171-181. https://doi.org/10.1016/j.apgeog.2015.04.009

39. Zhang, Q., Li, Y., & Sun, T. (2025). An improved groundwater vulnerability evaluation model based on random forest algorithm and spatio-temporal change prediction method. Journal of Hydro-environment Research (in press).

40. Chowdhury, A., Sikder, S., & Roy, S. (2020). Groundwater vulnerability to pesticide pollution assessment in the alluvial aquifer of Western Bengal basin, India using overlay and index method. Chemosphere, 246, 125826.

41. Talib, N., Abdul Maulud, K., Mohd, F., Mohd Taib, A., & Abdullah, M. (2026). Assessing the Impact of Topography on Soil Erosion in Agricultural Areas Using UAV and GIS-Based Analysis. International Journal of Geoinformatics, 22(1), 47–55. (https://doi.org/10.52939/ijg.v22i1.4719)

42. Zhu, Y., Fox, R.H., & Toth, J.D. (2025). Simulation of pesticide transport in 70-m-thick soil profiles in response to large water applications. Journal of Contaminant Hydrology (in press).

43. Šimůnek, J., Sejna, M., Saito, H., Sakai, M., & van Genuchten, M. Th. (2008). The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably saturated media

44. Stenemo, F., Lindahl, A.M., Gardenas, A., & Jarvis, N. (2007). Meta-modelling of the pesticide fate model MACRO for groundwater exposure assessments using artificial neural networks. Journal of Contaminant Hydrology, 93(1–4), 270–290.

45. Pérez-Indoval, R., Rodrigo-Ilarri, J., Cassiraga, E., & Rodrigo-Clavero, M.E. (2021). Numerical modelling of groundwater pollution by chlorpyrifos, bromacil and terbuthylazine: Application to the Buñol-Cheste aquifer (Spain). International Journal of Environmental Research and Public Health, 18(7), 3511.

46. Ahmadianfar, I., Kisi, O., & Elshafie, A. (2023). Recent advances in surface water quality prediction using artificial intelligence models. Water Resources Management, 37(15), 5905–5936.

47. Centanni, M., Ricci, G.F., De Girolamo, A.M., & Gentile, F. (2024). Modelling pesticides and ecotoxicological risk assessment in an intermittent river using SWAT. Scientific Reports, 14(1), 6389.

48. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

49. Lee, J.M., Ko, K.-S., & Yoo, K. (2023). A machine learning-based approach to predict groundwater nitrate susceptibility using field measurements and hydrogeological variables in the Nonsan Stream Watershed, South Korea. Applied Water Science, 13, 242.

50. Bedi, S., Samal, A., Ray, C., & Snow, D. (2020). Comparative evaluation of machine learning models for groundwater quality assessment. Environmental Monitoring and Assessment, 192, 776.

51. Wagner, P., Klotz, D., & Fuchs, S. (2020). A machine learning-based algorithm for water network contamination source localization. Water, 12(4), 1030.

52. Torres, J., Ramos, D., & Sela, L. (2021). Machine learning and simulation-optimization coupling for water distribution network contamination source detection. Water, 13(4), 546.

53. Karimanzira, D., Weis, J., Wunsch, A., Ritzau, L., Liesch, T., & Ohmer, M. (2023). Application of machine learning and deep neural networks for spatial prediction of groundwater nitrate concentration to improve land use management practices. Frontiers in Water, 5, 1193142.

54. Zhang, T., Liu, H., & Wu, Q. (2025). Enhancing water quality early-warning systems through fuzzy logic and artificial neural networks: A comparative analysis of algorithms and applications. Water, 17(19), 2841.

55. Toukak, A.S., et al. (2025). Interpretable machine learning framework for predicting pesticide phytotoxicity in wastewater reuse: Integrating molecular, quantum, and experimental descriptors. Environmental Research, 289, 123389.

56. Al-Khafaji, M., Abdulameer, L., Al-Shammari, M.A., Al Maimuri, N.M.L., Dulaimi, A., & Al-Jumeily, D. (2025). Revolutionizing water quality monitoring with artificial intelligence: A systematic review. Journal of Studies in Science and Engineering, 5(1), 358–385. https://doi.org/10.53898/josse2025528

57. Frincu, R.M. (2025). Artificial intelligence in water quality monitoring: A review of water quality assessment applications. Water Quality Research Journal, 60(1), 164–176. https://doi.org/10.2166/wqrj.2024.049

58. Anandhi, G., & Iyapparaja, M. (2024). Systematic approaches to machine learning models for predicting pesticide toxicity. Heliyon, 10(7), e28752. https://doi.org/10.1016/j.heliyon.2024.e28752

59. D'Amore, T. (2026). New challenges in the monitoring, risk assessment, and management of pesticides and biocides in the "One Health era." Journal of Xenobiotics, 16(1), 35. https://doi.org/10.3390/jox16010035

60. Galimberti, F., Bopp, S., Carletti, A., et al. (2025). From parcels to people: development of a spatially explicit risk indicator to monitor residential pesticide exposure in agricultural areas. Scientific Reports, 15, 96928 (https://doi.org/10.1038/s41598-025-96928-1).

61. AGSRT GIS Blogs. (2024). GIS applications in precision agriculture, including soil and crop health monitoring. AGSRT Publications.

62. Ranjan, R., & Deshmukh, P. (2025). Remote sensing application in plant protection and its usage in smart agriculture to hasten decision making of the farmers. Journal of Plant Diseases and Protection, 132, 1–18.

63. EOS Data Analytics. (2026). GIS for agriculture: Solutions, applications, benefits. EOS Blog.

64. Tiwari, G. (2022). Monitoring and management of agricultural insect pest with GIS. Oregon State University Extension Blog.

65. Srbinovska, M., Gavrovski, C., Dimcev, V., Krkoleva, A., & Borozan, V. (2015). Environmental parameters monitoring in precision agriculture using wireless sensor networks. Journal of Cleaner Production, 88, 297–307.

66. Botta, A., Cavallone, P., Baglieri, L., et al. (2022). A review of robots, perception, and tasks in precision agriculture. Applied Sciences, 12(11), 5691.

67. Farooqui, N., Ahmad, R., & Dar, S. (2025). Integration of smart sensors and IoT in precision agriculture: Trends, challenges and future prospectives. Frontiers in Plant Science, 16, 1587869.

68. Dutta, M., Gupta, D., Tharewal, S., Goyal, D., Sandhu, J.K., Kaur, M., Alzubi, A.A., & Alanazi, J.M. (2025). Internet of Things-based smart precision farming in soilless agriculture:

Opportunities and challenges for global food security. IEEE Access, 13, 34238-34268. https://doi.org/10.1109/access.2025.3540317

69. Alharbi, N., & Al-Fuqaha, A. (2025). The IoT and AI in agriculture: The time is now—A systematic review of smart sensing technologies. Sensors, 25(12), 3583.

70. Lindahl, A.M., Söderström, M., & Jarvis, N. (2008). Influence of input uncertainty on prediction of within-field pesticide leaching risks. Journal of Contaminant Hydrology, 98(3–4), 106–114.

71. European Commission. (2003). Technical guidance document on risk assessment in support of Commission Directive 93/67/EEC, Commission Regulation (EC) No. 1488/94, and Directive 98/8/EC — Part II: Environmental risk assessment. Office for Official Publications of the European Communities.

72. Society of Environmental Toxicology and Chemistry (SETAC) Workshop Group. (2019). Overcoming challenges of incorporating higher tier data in ecological risk assessments and risk management of pesticides in the United States. Integrated Environmental Assessment and Management, 15(3), 342–351.

73. Buonsenso, F. (2025). Scientific and regulatory perspectives on chemical risk assessment of pesticides in the European Union. Journal of Xenobiotics, 15(5), 173. https://doi.org/10.3390/jox15050173

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ResponsetoEditorandscientificreports.docx