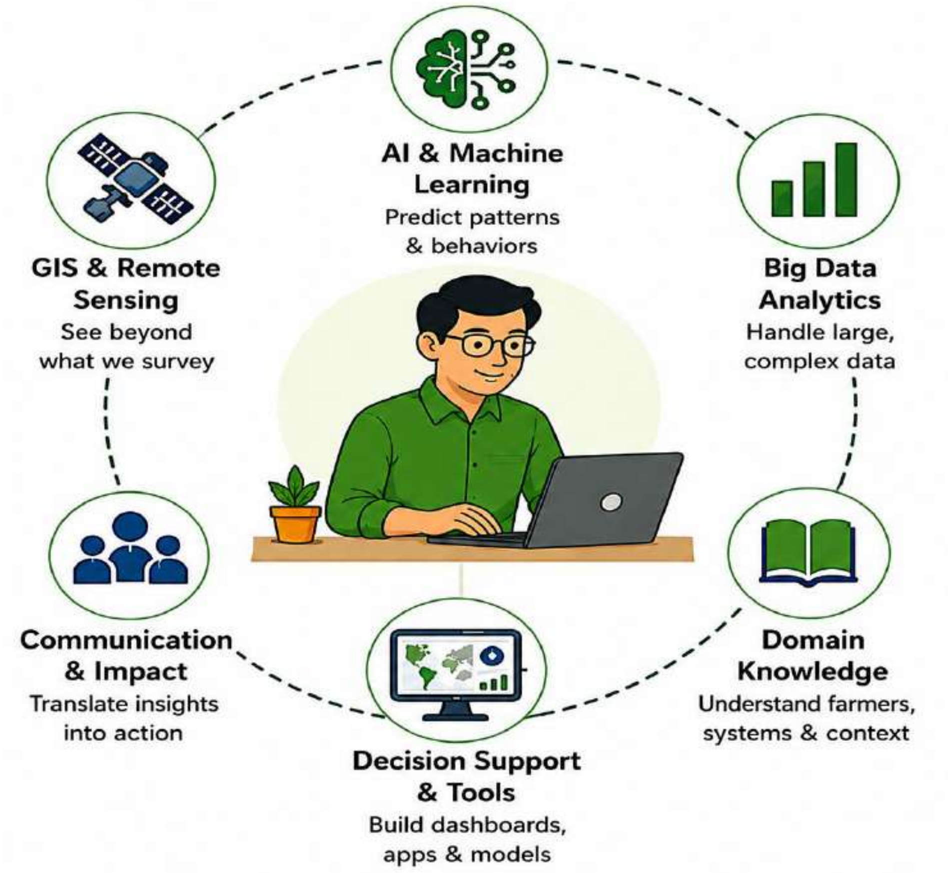
aesa Sharing, learning and networking for innovation Blog: 289 (August 2026) Agricultural Extension in South Asia

BEYOND REGRESSION: PREDICTION MODELLING WITH SPATIAL AND NON-SPATIAL DATA

In this blog, Paul Mansingh explores how integrating spatial and non-spatial data with predictive modelling can shift extension research from explaining past events to predicting future outcomes, enabling targeted, efficient, and evidence-based interventions.

CONTEXT

Over the years, agricultural extension researchers have used statistical methods such as correlation, regression analysis, path analysis, factor analysis, and structural equation modelling (SEM) to analyse farmers' behaviour. These approaches have enabled us to reach some important conclusions: What factors affect technology uptake? Why do some farmers not embrace innovations? What socio-economic and psychological factors are related to improved farm decision-making? These questions have continued to pose a problem. Today's farm, however, is a different place than 20 or 30 years ago.



THE EXTENSION RESEARCHER OF TOMORROW

A new breed of data-smart professionals

Precipitation patterns are changing due to climate change. Labour shortage drives changes in cropping systems. Farmers' access to information is changing due to digital technologies. Markets are becoming increasingly unpredictable. Digital agriculture and precision farming are receiving considerable attention and investment from governments and the agricultural sector. Meanwhile, an abundance of geospatial, remote sensing, and climate data is now openly accessible.

In a new reality, it's no longer sufficient to explain why something occurs.

In response to a new set of questions, researchers, policymakers, and extension professionals are on the lookout for answers:

- Which farmers are most likely to adopt a new technology?
- Which villages are most vulnerable to climate shocks?
- Where should extension officers concentrate their efforts?
- Which interventions are likely to generate the highest impact?
- Can we predict adoption behaviour before implementing expensive programmes?

These questions cannot be answered adequately using conventional statistical approaches alone. They require a new research paradigm—one that combines prediction models, artificial intelligence, geospatial analytics, and the integration of spatial and non-spatial data. The future of agricultural extension research lies not only in understanding the past but also in predicting the future.

WHY TRADITIONAL STATISTICAL ANALYSIS IS NO LONGER ENOUGH

Conventional statistical approaches are insufficient to answer these questions adequately. Agricultural extension has benefited greatly from traditional statistical methods. A correlation analysis is used to find associations between variables. Regression analysis is a means of measuring relationships and estimating the effects of independent variables. Structural Equation Modelling (SEM) is used to test theories about the relationships among latent constructs such as attitude, motivation, self-efficacy, and innovation proneness. These methods still play an essential role in the development of theories. They are, to a large extent, used for explanation, however.

Let's imagine classical agricultural extension research that is exploring the uptake of climate-smart agriculture. Upon surveying 500 farmers, the regression analysis might identify four factors as significant influences on adoption: education, extension contact, risk orientation, and farm size. The paper includes regression coefficients, p-values, the R^2 and any consideration of policy implications.
Although the data is statistically valid, there is one question that has not been answered: Who will be the farmers who need to be targeted in the future?

This is because traditional regression cannot easily identify the individual farmers/villages that are least likely to adopt an innovation or estimate the probability of adoption for each household. Extension professionals need decision support tools – not just coefficients with significance. That's where predictive analytics can have an impact.

AGRICULTURE IS SPATIAL—WHY ISN'T MOST EXTENSION RESEARCH?

One of the most significant limitations of conventional agricultural extension research is its tendency to ignore geography. Most extension studies rely exclusively on questionnaire-based survey data.

Researchers collect information such as:

- Age
- Education
- Family size
- Farm size

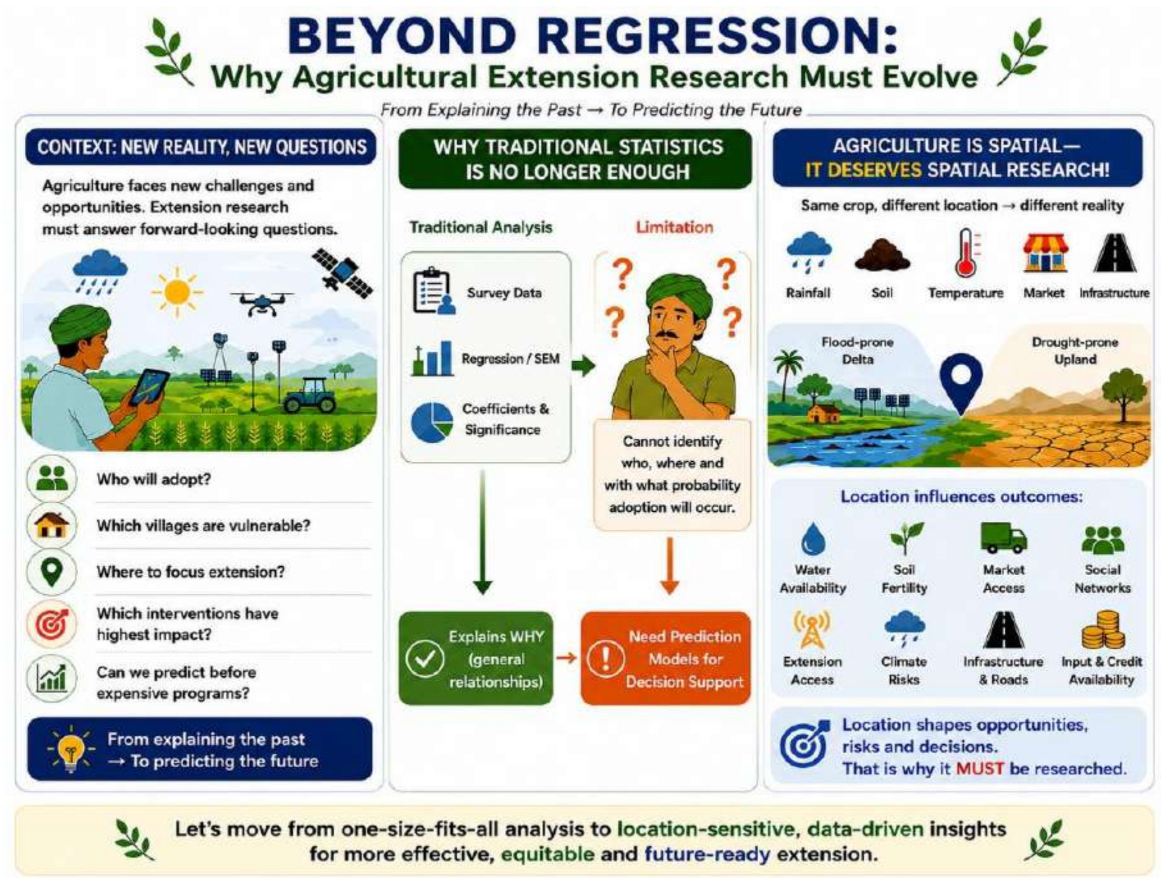
- Income
- Extension contact
- Social participation
- Risk orientation
- Innovativeness
- Information-seeking behaviour

These variables are undoubtedly important. However, they overlook one of the most influential determinants of agricultural behaviour: Location.

Agriculture is inherently spatial. A farmer cultivating paddy in a flood-prone delta faces entirely different production risks than another farmer growing the same crop in a drought-prone upland area.

Farmers separated by only a few kilometres may differ substantially in:
- rainfall received,
- groundwater availability,
- soil fertility,
- elevation,
- market accessibility,
- road connectivity,
- irrigation infrastructure, and
- climate vulnerability.

Ignoring these spatial differences inevitably limits our understanding of technology adoption and rural development.



THE POWER OF COMBINING SPATIAL AND NON-SPATIAL DATA

Research in modern agriculture has made an unprecedented number of spatial datasets available. In most parts of the globe, satellite data, climate data, Digital Elevation Models (DEMs), land-use information, road maps, irrigation facilities, and weather data are freely available. With the addition of the household survey data, researchers can now explore farmer behaviour within the environmental context.

The questionnaire collects:

- age,
- education,
- digital literacy,
- climate knowledge,
- extension contact,
- risk orientation,
- training received.

Now imagine that every farmer is also GPS-tagged during the survey.
Using GIS, each farmer's location can be linked with:

- annual rainfall,
- soil organic carbon,
- vegetation indices (NDVI),
- elevation,
- groundwater availability,
- flood frequency,
- drought occurrence,
- distance to markets,
- distance to extension offices,
- road accessibility.

Suddenly, the dataset becomes dramatically richer.

You don't just try to analyse who the farmer is; you are also aware of where the farmer is coming from, what the constraints are and how the geography influences the behaviour. This holistic perspective offers insights that are not captured by traditional surveys.

HOW IS AI TRANSFORMING AGRICULTURAL EXTENSION RESEARCH?

Another key development in the history of agricultural extension is the advent of machine learning. Unlike classic regression models, where researchers must define relationships in advance, machine learning algorithms learn patterns from data.

Algorithms such as:

- Random Forest,
- XGBoost,
- Artificial Neural Networks,
- Support Vector Machines,
- Gradient Boosting,
- Bayesian Networks,

can examine hundreds of interacting variables simultaneously. More notably, they can be used to model complex nonlinear relationships. Technology use might increase with higher education, but only among farmers with adequate irrigation water.

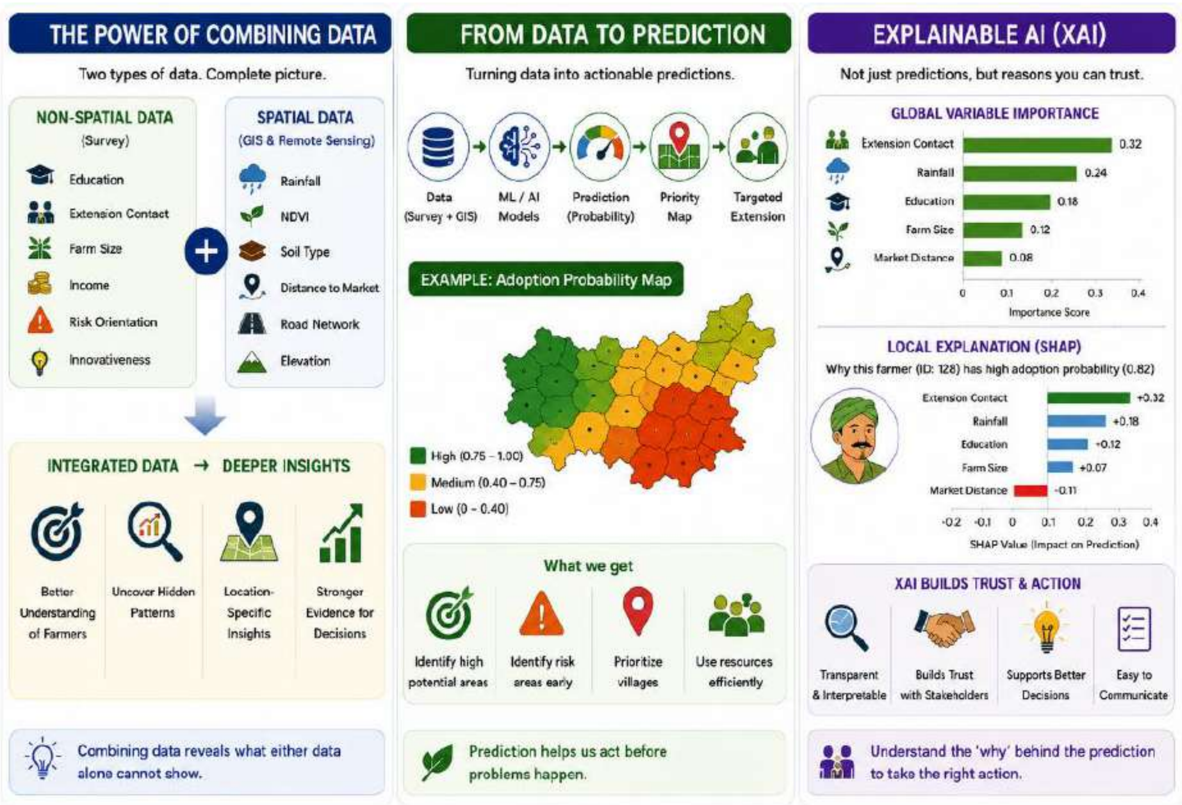
Likewise, adoption of digital literacy could be improved only as long as mobile internet connectivity is not lacking. Usually, these interactions are hard to identify with traditional regression analysis. Machine learning algorithms automatically identify these relationships. This can be especially useful in the context of agricultural systems, where many interacting social, climatic, institutional, and environmental factors are at play.

EXPLAINABLE AI: TAKING THE "BLACK BOX" OUT OF THE AI

One of the frequently raised objections to artificial intelligence is that it's akin to a "black box. Researchers may obtain very accurate forecasts, even if they do not know the reasons for the conditions that gave rise to them. This restriction has been largely overcome by Explainable Artificial Intelligence (XAI). Machine learning models can be interpreted using techniques such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model Explanations), Partial Dependence Plots, and Accumulated Local Effects.

For example, if a random forest predicted the accuracy of prediction as 92%, researchers can identify:

- which variables contributed most,
- whether rainfall increased or decreased adoption probability,
- how extension contact influenced predictions,
- and which combinations of variables produced the highest likelihood of adoption.



Explainable AI bridges the gap between predictive accuracy and scientific understanding.

SPATIAL INTELLIGENCE FOR BETTER EXTENSION PLANNING

Imagine that an agricultural department has funding to conduct climate-smart agriculture demonstrations in only 100 villages.
Which villages should be selected?

Traditionally, villages may be chosen based on administrative convenience or subjective judgment. A predictive geospatial model offers a far better alternative. By integrating household surveys with GIS layers and machine learning, researchers can generate adoption probability maps.

Villages can then be classified as:

- High adoption potential
- Moderate adoption potential
- Low adoption potential
- Climate-vulnerable
- Resource-constrained
- Priority intervention zones

Extension resources can therefore be allocated objectively rather than intuitively. This approach significantly improves programme efficiency.

TOWARDS PRECISION AGRICULTURAL EXTENSION

Precision farming is already in use in agriculture. Advanced technology is making GPS-guided tractors, drones, satellite imagery, IoT sensors, automated irrigation and digital weather forecasting more widely used. Extension to agriculture should also undergo a similar change. This is a new paradigm that may be referred to as Precision Agricultural Extension.

Precision Extension means delivering:

- the right advisory,
- to the right farmer,
- at the right place,
- at the right time,
- through the right communication channel.

Achieving this vision requires integrating:

- Farmer surveys,
- GIS,
- Remote sensing,
- Climate databases,
- Mobile advisory systems,
- Artificial Intelligence,
- Machine Learning,
- Explainable AI,
- Spatial Statistics,
- Decision Support Systems.

Rather than implementing identical extension programmes across entire districts, interventions can be tailored to specific agro-ecological zones, farmer segments, and vulnerability profiles.

A VISION FOR FUTURE AGRICULTURAL EXTENSION RESEARCH

The next generation of agricultural extension studies should move beyond isolated statistical analyses. Instead of asking only:
"What factors influence adoption?"

Researchers should also investigate:
- Who will adopt?
- Where will adoption occur?
- When should interventions be implemented?
- Which villages require immediate attention?
- What will happen if rainfall patterns change?
- How will digital advisory services influence future adoption?
- Which farmers are at greatest risk of exclusion?

Answering these questions requires integrating disciplines that have traditionally operated independently.

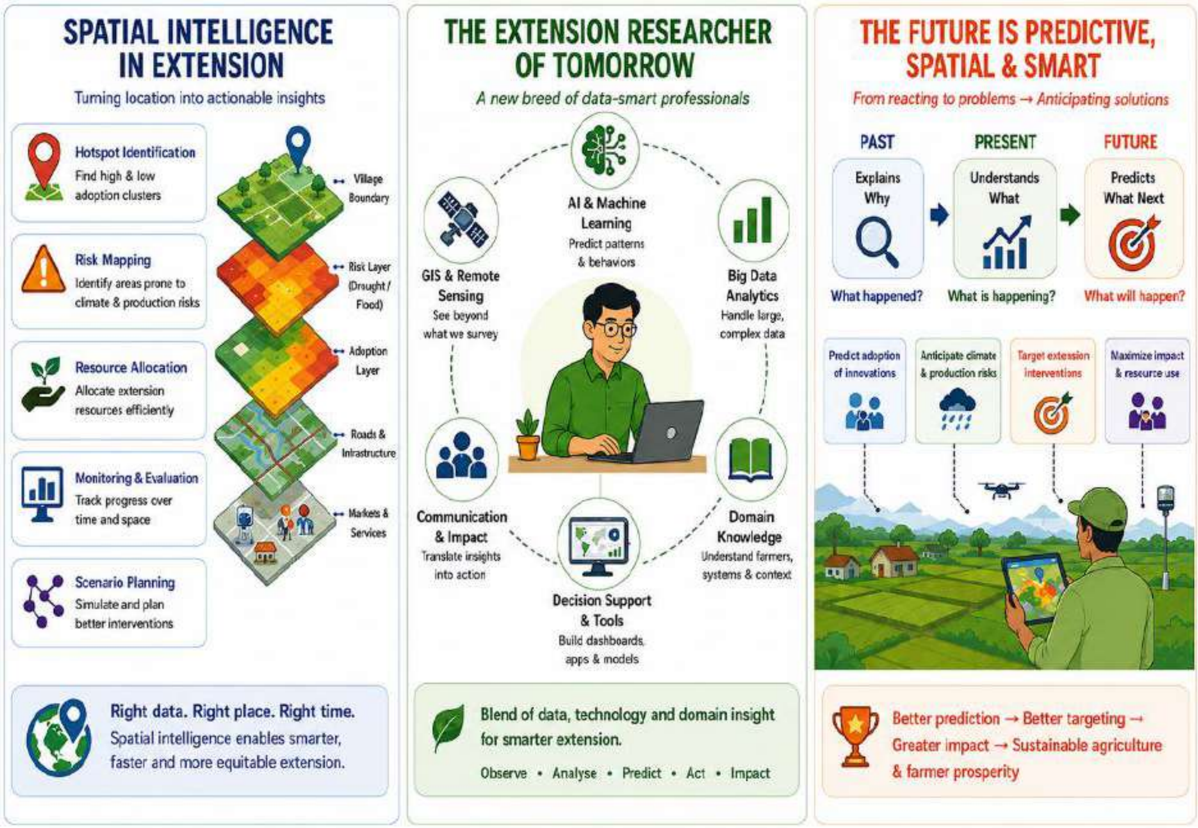
Agricultural Extension must increasingly collaborate with:
- Geographic Information Systems (GIS),
- Remote Sensing,
- Data Science,
- Artificial Intelligence,
- Climate Science,
- Spatial Econometrics,
- Computational Social Science,
- Precision Agriculture,
- Behavioural Analytics.

Such interdisciplinary research will produce findings that are scientifically robust, operationally relevant, and directly useful for policymakers.

THE EXTENSION RESEARCHER OF TOMORROW

The agricultural extension researcher of the future will require skills that extend well beyond questionnaire design and conventional statistical analysis.

Tomorrow's researchers should be comfortable working with:
- GPS-enabled surveys,
- GIS software such as QGIS or ArcGIS,
- Google Earth Engine,
- R or Python,
- Machine Learning algorithms,
- Explainable AI techniques,
- Spatial statistics,
- Geospatial databases,
- Interactive dashboards,
- Decision-support systems.



Fortunately, these technologies are becoming increasingly accessible through open-source software and freely available datasets. For a better understanding, refer to the enclosed AI-generated infographic, which explains the power of predictive models and spatial and non-spatial data.

The challenge is no longer access to technology—it is our willingness to adopt new research paradigms.

CONCLUSION: BUILDING A SMARTER AGRICULTURAL EXTENSION SYSTEM

Fortunately, these technologies are becoming increasingly available as open-source software and freely available datasets. Access to technology is no longer the challenge; it's our openness to new research paradigms.

The very nature of agricultural extension has been, and always will be, to support farmers in their decision-making. Now, researchers have an unparalleled opportunity to assist extension systems in decision-making as well. Theory development and understanding of relationships among behaviours will remain important through traditional statistical analyses such as correlation, regression, and SEM. They no longer, however, should be the final goal of agricultural extension research.

Researchers can go beyond mere description of farmer behaviour to prediction using spatial and non-spatial data, predictive analytics, machine learning, and explainable AI. They know which interventions provide the most benefit, who will benefit most, and where public resources can be used more effectively.

With climate uncertainty, digital transformation and the abundance of data, the future of agricultural extension is no longer about understanding the past but about forecasting the future.

A new frontier in agricultural extension research is already here. The issue is not if, but when, to develop the skills, partnerships, and institutional structure necessary to enjoy the potential of these approaches.

Dr J. Paul Mansingh is Professor and Head of the Department of Agricultural Extension & Economics at the VIT School of Agricultural Innovations and Advanced Learning (VAIAL), Vellore Institute of Technology. He specialises in advanced statistical and econometric analysis, structural equation modelling, quantitative and qualitative research methods, AI-enabled predictive modelling, simulation modelling, and computational analytics for decision support and policy evaluation. He can be reached at paul.mansingh@vit.ac.in

AESA Secretariat: Centre for Research on Innovation and Science Policy (CRISP)
Road No 10, Banjara Hills, Hyderabad 500034, India

www.aesanetwork.org **Email: aesanetwork@gmail.com**