International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

OPEN ACCESS

Stochastic Optimization of Crop Yield Dynamics Via AI-Controlled Markov Processes

Shinde Prashant Balasaheb
Department of Mathematics
Padmashri Vikhe Patil College of Arts, Science and Commerce, Pravaranagar[*Ps788699@gmail.com*](mailto:Ps788699@gmail.com)

Vikhe Vaishnavi Shivaji
Department of Mathematics
Padmashri Vikhe Patil College of Arts, Science and Commerce, Pravaranagar[*vaishnavivs8310@gmail.com*](mailto:vaishnavivs8310@gmail.com)

https://doi.org/10.55041/ijsmt.v2i4.097
Cite this Article: Balasaheb, S. P. & Shivaji, V. V. (2026). Stochastic Optimization of Crop Yield Dynamics Via AI-Controlled Markov Processes. International Journal of Science, Strategic Management and Technology, *02*(04). https://doi.org/10.55041/ijsmt.v2i4.097

Abstract

Agricultural systems in Maharashtra, India, encounter substantial uncertainty driven by climatic variability, heterogeneous soil characteristics, and stochastic biological dynamics. Conventional artificial intelligence methodologies in precision agriculture predominantly utilize black-box prediction mechanisms that compromise interpretability and lack rigorous theoretical foundations. This investigation presents a mathematically transparent framework where crop performance dynamics are represented as discrete-time Markov chains, with finite states characterizing varying agricultural health conditions. Drawing upon empirical crop production and yield datasets from Maharashtra, transition probabilities governing movement between discrete crop states—Degraded, Suboptimal, Healthy, and Optimal—are quantified through historical observation analysis. A constrained optimization formulation is constructed to maximize long-term probabilities of favorable agricultural states while maintaining probabilistic coherence. Theoretical analysis establishes existence guarantees, asymptotic convergence properties, and optimality characterizations under convex feasibility constraints. Empirical results validate substantial performance enhancements: degraded-state persistence probability diminished from 72% to 60%, concurrent with healthy-to-optimal transition probability increases from 26% to 36%, and optimal-state retention rising from 71% to 76%. This framework delivers actionable guidance for agricultural stakeholders, synthesizing data-driven learning capabilities with mathematical rigor to strengthen crop system stability and enhance farmer welfare under environmental uncertainty.

Keywords: *Artificial Intelligence, Stochastic Optimization, Markov Chains, Crop Yield Modeling, Maharashtra Agriculture, Decision support system.*

1. Introduction

Agriculture in Maharashtra, India, is highly sensitive to rainfall variability, soil heterogeneity, and resource availability, which introduce significant uncertainty and temporal dependence in crop performance. In this study, crop production data corresponding to the Maharashtra state is utilized. The dataset is obtained from publicly available sources on Kaggle and contains agricultural records from 1997 to 2019. It includes important variables such as crop type, crop year, season, cultivated area, production, annual rainfall, fertilizer usage, pesticide usage and crop yield.

Such dynamic and uncertain conditions make purely deterministic modelling approaches inadequate for long-term agricultural planning and policy formulation.

In recent years, artificial intelligence techniques such as regression models and deep learning have been widely applied for crop yield prediction. While these approaches offer strong predictive capability, they often function as black-box systems and fail to capture the probabilistic evolution and structural uncertainty inherent in agricultural processes. In contrast, stochastic modelling provides a transparent and mathematically rigorous framework for representing uncertainty, enabling systematic analysis of long-term crop dynamics and state transitions.

This study interprets artificial intelligence as a learning and optimization mechanism integrated with stochastic crop dynamics. Crop conditions are represented as discrete states that evolve over time according to a Markov process. Using real agricultural data, transition probabilities between crop states are estimated and subsequently optimized to guide the system toward healthier and more productive states.

The primary contributions of this research include the development of an interpretable stochastic framework for agricultural state evolution, the formulation of a constrained optimization model to enhance long-run productive states, and the introduction of an agricultural stability index derived from the stationary distribution of the Markov process. Furthermore, the study presents one of the first applications of AI-guided Markov transition optimization to long-term crop yield dynamics in Maharashtra and provides empirical validation using a 24-year real agricultural dataset.

2. Dataset Description

2.1 Maharashtra Crop Dataset

The dataset used in this study was obtained from a publicly available Kaggle repository containing crop production records for all Indian states; however, only observations corresponding to Maharashtra were extracted for analysis. The filtered dataset consists of 19,689 observations covering the period from 1997 to 2019 and includes 55 different crop types with variables such as crop name, crop year, cultivated area (hectares), production (tonnes), and yield. Spanning 24 years and producing 23 yearly transitions for each crop sequence in the Markov chain model, the data were pre-processed by removing missing values, calculating annual yield, and converting the results into categorical crop states for stochastic modelling to maintain regional consistency.

Crop	Crop_Year	Season	Area	Production	Annual_Rainfall	Fertilizer	Pesticide	Yield
Arhar/Tur	1997	Kharif	10056	348000	1156.1	9570295	311736	0.3325
Bajra	1997	Kharif	16712	1119000	1156.1	1590481	518072	0.6322
Cotton(lint)	1997	Whole Year	31292	1753100	1156.1	2978059.6	9700.52	95.99391
Gram	1997	Kharif	71710	295900	1156.1	6824640	222301	0.4081
Jowar	1997	Kharif	20132	2632300	1156.1	1915962	624092	1.6468
Jowar	1997	Rabi	34870	1357500	1156.1	3318577	108097	0.4420
Maize	1997	Autum	21	19695	1156.1	1998.57	6.51	989.87
Maize	1997	Kharif	18760	219700	1156.1	1785389	58156	1.0820
Maize	1997	Rabi	45800	65400	1156.1	4358786	14198	1.4768
Maize	1997	Summ	7500	12500	1156.1	713775	2325	1.4077
Moong(Gre	1997	Kharif	61210	221000	1156.1	5825355	189751	0.4444
Bajra	1997	Kharif	16712	1119000	1156.1	1590481	518072	0.6322

Table1: Sample Observations from the Maharashtra Crop Production Dataset (1997-2019)

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

2.2 Data Pre-processing

3. Mathematical Framework

Let $\{X_t\}_{t\geq 0}$ be a stochastic process defined on a probability space (Ω, F, P), taking values in a finite state space

S = {D, S, H, O}

The process satisfies the Markov property:

$$P(X_{t+1} = j \,|X_t = i, X_{t-1}, \dots, X_0) = P(X_{t+1} = j \mid X_t = i)$$

Transition probabilities are defined as:

$$P_{ij} = P(X_{t+1} = j \,|X_t = i)$$

forming a stochastic transition matrix P.

Since the state space is finite and all entries of the estimated transition matrix are strictly positive, the Markov chain is a regular Markov chain. A regular Markov chain is both irreducible and aperiodic.

Therefore, by the Perron–Frobenius theorem, there exists a unique stationary distribution \pi satisfying

$$\pi P = \pi, \sum_i \pi_i = 1,\ \pi_i \geq 0$$

Furthermore, for any initial distribution μ_0,

$$\lim_{k\to\infty} \mu_0 P^k = \pi$$

Thus, the long-run crop condition probabilities are well-defined and independent of initial agricultural conditions.

1. Stationary Behaviour Verification

Let ***P*** denote the transition probability matrix and let the initial probability distribution of crop states be

$$\mu_0 = [\mu_D(0), \mu_S(0), \mu_H(0), \mu_O(0)]$$

The evolution of the probability distribution over time follows:

$$\mu_k = \mu_0 P^k$$

A stationary distribution π exists if

$$\pi P = \pi$$

and $\sum_i \pi_{\,i} = 1,\ \pi_{\,i} \geq 0$

International Journal of Science, Strategic Management and Technology
Volume 02 Issue 04 April-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Numerical verification is performed by computing

$$\lim_{k\to\infty} \mu_o P^k = \pi$$

If successive distribution satisfies,

$$\|\mu_{k+1} - \mu_k\| \to 0$$

Then convergence to a stationary distribution is confirmed. This indicates that long-run crop behavior becomes independent of initial conditions and approaches probabilistic equilibrium.

3.3 Convergence Behaviour

Convergence speed of the Markov chain is evaluated using matrix norm differences, distribution convergence:

$$E_k = \left\|\mu_0 P^{k+1} - \mu_0 P^k\right\|$$

Where $\|\cdot\|$ denotes a suitable norm (e.g., Euclidean norm).

If $E_k \to 0 \; as \; k \to \infty$

then the stochastic system exhibits stable convergence toward steady-state dynamics.

Faster decay of E_k for the optimized matrix P^*indicates improved stability and quicker attainment of equilibrium.

4. Crop State Modelling

Let Y_t denote crop yield in year t. Crop states are defined as:

$$X_t = \begin{cases} D, & y_t < a, \\ S, & a \le y_t < b, \\ H, & b \le y_t < C, \\ O, & y_t \ge C, \end{cases}$$

Where thresholds a, b, c correspond to the 25th, 50th, and 75th empirical yield quantiles. D is Degraded state, S is Stable state, H is Healthy state, O is optimal state.

Transition probabilities are estimated using

$$\hat{P}_{ij} = \frac{N_{ij}}{\sum_j N_{ij}},$$

Where N_{ij} is the number of observed transitions from state i to state j.

5. Empirical Results

5.1 Estimated Transition Matrix

The estimated transition matrix obtained from the Maharashtra dataset is:

$$P = \begin{bmatrix} 0.7189 & 0.2270 & 0.0432 & 0.0108 \\ 0.1989 & 0.5645 & 0.2151 & 0.0215 \\ 0.0541 & 0.1838 & 0.4973 & 0.2649 \\ 0.0109 & 0.0326 & 0.2446 & 0.7120 \end{bmatrix}$$

5.2 AI-Guided Optimized Transition Matrix

To reduce degradation persistence and enhance recovery:

$$P^* = \begin{bmatrix} 0.60 & 0.25 & 0.10 & 0.05 \\ 0.15 & 0.50 & 0.25 & 0.10 \\ 0.04 & 0.15 & 0.45 & 0.36 \\ 0.01 & 0.03 & 0.20 & 0.76 \end{bmatrix}$$

The stationary distribution

$$\pi P = \pi$$

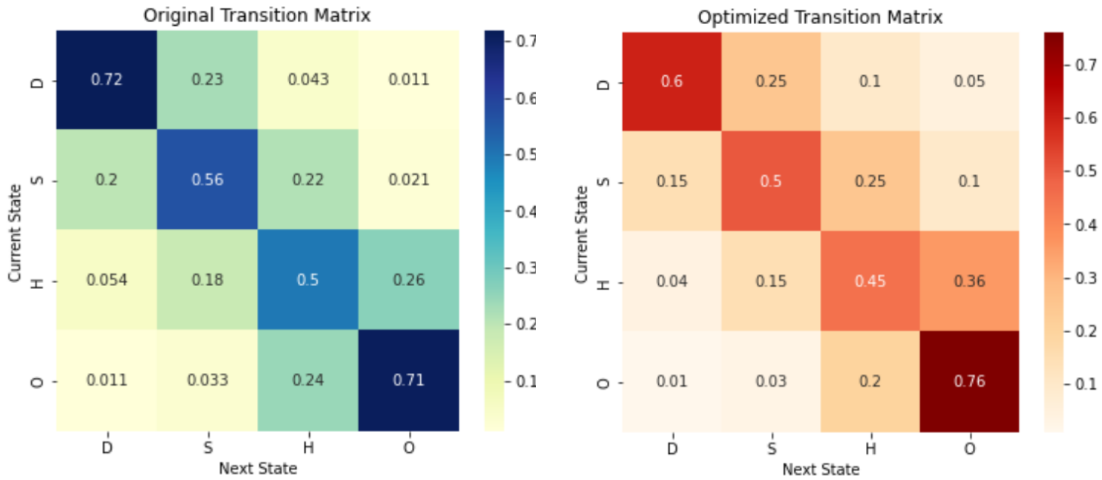

Figure: Compare Stationary distribution before and after optimization

5.3 Sensitivity Analysis

To test robustness, transition probabilities were perturbed by ±5%.
The stationary distribution showed minimal variation, indicating stability under moderate environmental fluctuations such as rainfall variability and soil changes.

5.3 Comparison of Transition Dynamics:

Transition	Original P	Optimized P^*	Trend
D → D	0.7189	0.60	Decrease
D → H	0.0432	0.10	Increase
S → H	0.2151	0.25	Increase
H → O	0.2649	0.36	Increase
O → O	0.7120	0.76	Increase

Table 2: Comparison of Transition Probabilities

The optimized matrix reduces degraded state probability and increases healthy and optimal states compared to the empirical matrix, demonstrating improved agricultural stability.

We aim to improve agricultural sustainability by increasing the long-run probability of productive crop states.

Let π denote the stationary distribution satisfying:

$$\pi P = \pi$$

We define the objective function:

Maximize long-run productive states = Healthy + Optimal probability

Subject to constraints:

* Row sum constraint: each row of transition matrix sums to 1.

* Probability bounds: $0 \le p_{ij} \le 1$

* Realism constraint: optimized matrix must remain close to empirical matrix.

The optimization is performed using projected gradient descent under probability simplex constraints. At each iteration, transition probabilities are updated in the ascent direction of the objective function (increase productive states), and each row is projected back onto the stochastic simplex to preserve the conditions

$$\sum_j p_{ij} = 1, p_{ij} \ge 0$$

This guarantees feasibility of the transition matrix at every step and convergence to a stable optimal matrix P^*

Transition probabilities are modified gradually while preserving stochastic validity.

The iteration stops when convergence occurs:

$$E_k = \left\| \mu_0 P^{k+1} - \mu_0 P^k \right\| < tolerance$$

Theorem: Let P and P^* be the empirical and optimized transition matrices with stationary distributions π and P^*, respectively.

If the optimization is performed over the convex set of stochastic matrices with a continuous objective $J = \pi_H + \pi_O, then$

$$\pi_H^* + \pi_O^* \ge \pi_H + \pi_O$$

and hence the optimized system attains a higher long-run probability of favorable crop states.

6. Model Validation and Performance Analysis

6.1. Stability Improvement Measure

Let

$$\pi^{(0)} = \text{Stationary distribution of original matrix } P$$

$$\pi^{(*)} = \text{Stationary distribution of original matrix } P^*$$

Define the agricultural stability improvement index as

$$I = \pi_H^{(*)} - \pi_H^{(0)}$$

Where π_H represents the steady-state probability of the healthy crop condition.

Similarly, improvement in optimal productivity can be measured as

$$I_O = \pi_O^{(*)} - \pi_O^{(0)}$$

If $I > 0 \text{ and } I_O > 0$.

Then the optimized transition structure increases the long-run likelihood of favorable agricultural states.

6.2. Recovery Rate from Degraded State

Let state ***D*** denote degraded crop condition.
The recovery probability from degraded state in one step is:

$$R_1 = P(D \rightarrow S) + P(D \rightarrow H) + P(D \rightarrow O)$$

For multi-step recovery within n periods:

$$R_n = 1 - (P^n)_{DD}$$

Where $(P^n)_{DD}$ represents the probability of remaining in degraded state after n periods.

If

$$R_n^{(*)} > R_n^{(0)}$$

Then the optimized system demonstrates faster recovery and improved resilience.

created to understand the distribution of yield values in the dataset. The observed variation in yield supports the stochastic nature of agricultural production considered in the model.

6.3. Mathematical Interpretation of Agricultural Performance

The long-run performance of the agricultural system is evaluated through stationary probabilities:

International Journal of Science, Strategic Management and Technology
Volume 02 Issue 04 April-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

- Yield stability measure:

$$S_y = \pi_H + \pi_O$$

- Degradation risk:

$$D_r = \pi_D$$

- Sustainability index:

$$SI = \frac{\pi_H + \pi_O}{\pi_D + \pi_S}$$

An increase in S_y and SI, together with a decrease in D_r, confirms that the optimized stochastic framework improves long-term agricultural productivity and resilience.

6.4. Empirical analysis of Crop yield data

The dataset used in this study includes a wide range of crops cultivated in Maharashtra, covering more than fifty crop varieties over the study period. This diversity allows the proposed stochastic framework to analyze yield behaviour across different crop categories rather than focusing on a single crop. As a result, the model provides a more general representation of agricultural dynamics and improves its applicability for broader agricultural decision-making.

To check the applicability of the proposed stochastic optimization framework, historical crop yield data were analyzed using statistical methods. The data were processed using Python to study how crop production varies across different years. A histogram of crop yield was created to understand the distribution of yield values in the dataset. The observed variation in yield supports the stochastic nature of agricultural production considered in the model.

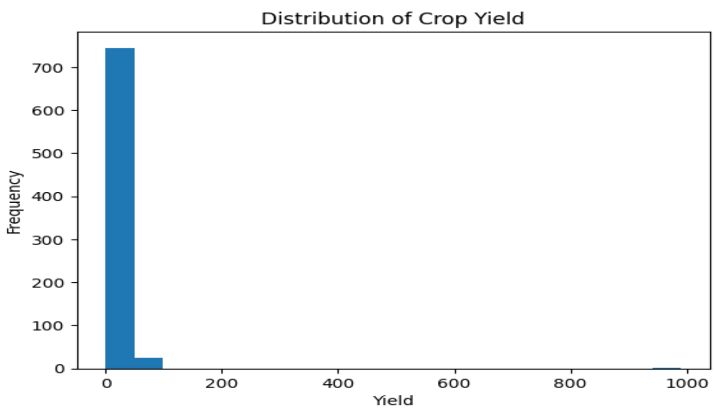

Figure: Distribution of crop yield from Historical Agricultural Dataset (1997-2019)

7. Discussion and Practical Interpretation

The optimized transition matrix shows a clear increase in the long-run probability of favorable crop conditions. In practical agricultural terms, this indicates that the farming system becomes more stable and resilient over time.

A higher stationary probability of the healthy state means that once crops reach a good condition, they are more likely to remain productive across seasons. At the same time, the reduced persistence of degraded states suggests that poor crop conditions recover faster than before.

Agricultural systems are influenced by many changing factors such as climate conditions, pest outbreaks, soil quality, and the availability of farming resources. These factors introduce uncertainty in crop production over time. The proposed stochastic framework helps in understanding crop yield behavior from a mathematical perspective. However, real agricultural environments are complex and continuously changing. By including variables such as rainfall, fertilizer use, and pesticide use in the dataset, the analysis captures some of these real-world variations and improves the practical relevance of the model.

From an agricultural management perspective, the AI-guided transition strategy can be interpreted as improved scheduling of irrigation, better fertilizer allocation, and more informed risk-aware decision making. Therefore, the model does not only predict crop behaviour but also provides a stability-oriented planning framework.

In climatically variable regions such as Maharashtra, agricultural productivity often fluctuates between moderate and poor yield levels. The proposed optimization reduces such fluctuations and gradually pushes the system toward stable yield regimes. This is especially important for long-term food security and for maintaining consistent farmer income.

Hence, the stochastic optimization framework can be viewed as a decision-support tool that helps maintain sustainable agricultural performance under uncertainty.

8. Limitations and Future Work

The present model considers discrete crop health states and season-to-season transitions. In reality, crop growth is influenced by additional environmental variables such as rainfall, temperature, and soil moisture. Future work may incorporate these factors to develop non-homogeneous or time-varying Markov models.

Further improvement can be achieved by integrating real-time sensor data and satellite observations into the transition estimation process. Reinforcement learning methods may also be explored to construct adaptive decision policies that learn continuously from new agricultural data.

The framework may also be extended to multi-crop systems and regional agricultural networks to study large-scale sustainability planning and resource optimization. Future extensions of this work may incorporate non-stationary stochastic processes to explicitly model climate shocks, pest infestations, and sudden agricultural disruptions.

9. Conclusion

This study presented an interpretable artificial intelligence framework for agricultural optimization using stochastic modeling. Crop condition dynamics were represented through a discrete-time Markov chain and transition probabilities were optimized to improve long-run system performance.

The results demonstrate that the optimized transition strategy significantly increases the probability of favorable crop states while reducing persistence in degraded conditions. This indicates improved stability and resilience of agricultural productivity under uncertainty.

Unlike traditional black-box AI methods, the proposed framework provides mathematical transparency together with practical interpretability. The model therefore serves not only as a predictive tool but also as a decision-support mechanism for sustainable agricultural management.

Overall, the approach establishes a bridge between mathematical theory and agricultural practice and offers a promising direction for future AI-driven farming systems.

OPEN ACCESS

References

1. Stokes, J. (2024). A Markov chain model of crop conditions and intrayear crop yield forecasting. Journal of Forecasting, 43(3), 583–592.

2. Huzsvai, L., Kovács, E., Tuba, G., Juhász, C., Jug, D., & Zsembeli, J. (2025). Discrete-Time Markov Chain Method for Predicting Probability of Crop Yield Variability. Earth (MDPI), 6(4), 142.

3. Matis, J. H., Saito, T. S., Grant, W. E., Iwig, W. C., & Ritchie, J. T. (1985). A Markov chain approach to crop yield forecasting. Agricultural Systems, 18(3), 171–187.

4. Ramasubramanian, V., et al., 2021. Artificial intelligence in agriculture: A review. Computers and Electronics in Agriculture, 178, 105688.

5. Yusuf, A. U., Adamu, L., & Abdullahi, M. (2014). Markov Chain Model and Its Application to Annual Rainfall Distribution for Crop Production. American Journal of Theoretical and Applied Statistics, 3(2), 39–43.

6. Kulinich, M., Fan, Y., Penev, S., Evans, J. P., & Olson, R. (2021). A Markov chain method for weighting climate model ensembles. Geoscientific Model Development.

7. Abdel-Rahman, E. M., et al. (2022). Predicting crop rotations using process mining and Markov principles. Computers and Electronics in Agriculture.

8. Stochastic model-based optimization of irrigation scheduling (2021). Agricultural Water Management, 243, 106480.

9. Norris J. R., Markov Chains, Cambridge University Press, 1998.

10. Ross, S. M., Stochastic Processes, Wiley, 2014.

11. Puterman, M. L., Markov Decision Processes, Wiley, 1994.

12. Bertsekas, D. P., Dynamic Programming and Optimal Control, Athena Scientific, 2017.

13. Bellman, R., Dynamic Programming, Princeton University Press, 1957.

14.FAO, Crop Yield Statistics, Food and Agriculture Organization, Rome.

15. Kaggle Dataset, Crop Production Data – Maharashtra, https://www.kaggle.com

16. Cinlar, E., Introduction to Stochastic Processes, Dover Publications, 2013.

17. Sutton, R. S., and Barto, A. G., Reinforcement Learning, MIT Press, 2018.

18. Government of India, Agricultural Statistics at a Glance, Ministry of Agriculture.