



Vocational opportunities for agricultural migrants in northern India: Insights from grassroots

SUSHMITA SAINI^{1*}, RAJARSHI ROY BURMAN^{1,2}, RABINDRA NATH PADARIA¹,
GIRIJESH SINGH MAHRA¹, SITARAM BISHNOI¹, KAUSTAV ADITYA³, NITHYASHREE M L¹,
SONALI MALLICK⁴ and SMRUTI RANJAN PADHAN¹

ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India

Received: 17 February 2024; Accepted: 28 March 2024

ABSTRACT

The livelihood of Indian farmers is primarily shaped by agriculture and its allied components. The degree of intricacy and contributions of each agricultural activity, although, vary with different social-ecological systems. Migration is considered an approach to livelihood adaptation; however, the migrants face challenges in new social-ecological settings due to their limited access to resources. This research aimed to develop a predictive model for sustainable agricultural employment options in Rajasthan and Uttar Pradesh (UP). In this study, states were selected purposively, while respondent farmers were chosen using a stratified multistage sampling design. A total of 480 resident and migrant farmers were selected to collect data. The machine learning algorithm, based on classification and regression tree (CART) analysis, was applied that could help to identify factors and prospects for migrant farmers in the agricultural sector. Results indicated that milk yield, operational land holding, and *rabi* crop yield were significant predictors. Further, milk yield, *rabi* crop yield, and *kharif* crop yield were observed to be the essential factors contributing to profitable ventures. The recommendations provided by Krishi Vigyan Kendra's professionals led to the identification of key income-generating activities, such as livestock management, vegetable production, and organic farming. These prospects are tailored to the location-specific context where migrant farmers reside. Overall, this research shed light on viable employment opportunities, ultimately contributing to the well-being of migrant workers in northern India, in addition to the policy interventions focused on capacity building and providing an enabling environment to the migrants.

Keywords: Classification and regression tree, Income-generating activities, Profitable ventures

Agriculture serves as a major source of livelihood in the world. In the year 2022, ~53 million individuals were internally migrated within their own countries across 25 nations facing food crises (GMDAC 2024). Additionally, around 20 million refugees and asylum seekers found shelter in 55 out of the 58 countries and territories experiencing food crises (GMDAC 2024). In India, the migration rate was 28.9% during 2020–21, with 26.5% and 34.9% in rural areas and urban areas, respectively (NSO 2022). Migrants, often originate from impoverished areas and seek employment in the agricultural sector as a means of livelihood and for economic stability (Saini *et al.* 2023a). However, limited access to reliable employment opportunities hampers the socio-economic development of the farmers as well as the farming community (Parnacli 2019). Agricultural

employment has been a substantial aspect of the labour market dynamics, rural development, poverty alleviation and social welfare (Kaika and Racelis 2021). The availability and accessibility of agricultural occupations in northern India are influenced by seasonal variations in crop demand (Keshri and Bhagat 2012), regional disparities, land fragmentation, and the use of traditional farming practices (Sarkar *et al.* 2019). Additionally, limited access to credit, skills and social networks further compounds the challenges faced by migrants in seeking employment (Saini *et al.* 2023b). Studies have highlighted the importance of understanding the socio-economic characteristics of migrants, their occupational mobility patterns, livestock holding and the barriers they encounter when accessing agricultural employment (Saini *et al.* 2024).

According to the Migration of India (2020) report rural males and rural females migrated in search of employment was 11.8% and 0.2%, respectively; urban males and urban females migrated in search of employment was 29.9% and 1.5%, during 2020–21 (NSO 2022) Migrarian (migration and agriculture) livelihoods now form a crucial part of India's economy and it has become one of the key

¹ICAR-Indian Agricultural Research Institute, New Delhi;

²Indian Council of Agricultural Research, New Delhi; ³ICAR-Indian Agricultural Statistics Research Institute, New Delhi; ⁴ICAR-Central Soil Salinity Research Institute, Regional Research Station, West Bengal. *Corresponding author email: sushmitasaini0117@gmail.com

approaches of livelihood adaptation across rural India. Moreover, the spatial distribution of agricultural jobs across different regions in northern India plays a crucial role in determining migrants' opportunities (Saini *et al.* 2024). The problem at hand revolves around the limited understanding of agricultural employment opportunities for migrants in northern India (Singh and Basu 2020).

Despite the significant contribution of migrants to the agricultural sector, there is a lack of comprehensive information on the availability, accessibility, and distribution of employment opportunities (Singh and Basu 2020). This knowledge gap hampers effective policymaking and targeted interventions to address the employment challenges faced by migrants. Therefore, this research aims to utilize machine learning algorithms to scientifically investigate and delineate agricultural employment opportunities for migrants in northern India. By leveraging the power of data-driven analysis, this study seeks to provide valuable insights into exploring agricultural employment opportunities, identify factors influencing income generation activities and contribute to evidence-based recommendations for policymakers and stakeholders. Subsequently, the study contributes to the growing body of knowledge on labour market dynamics, specifically focusing on the agricultural sector.

MATERIALS AND METHODS

As reported by the Ministry of Labour and Employment, Govt. of India, a significant proportion of migrant workers, approximately 1.14 crore, returned to their home states during the COVID-19 lockdowns. The primary states to which these migrants returned were mainly Uttar Pradesh (UP), Bihar, West Bengal, Rajasthan, and Odisha. Specifically, 26% of the migrant workers returned to UP, while 11% returned to Rajasthan. In light of these indicators, both Uttar Pradesh and Rajasthan were intentionally selected to investigate vocational opportunities for agricultural migrants in their home states. A stratified multistage sampling design was used for data collection. A respondent was considered an agricultural migrant if, within the last 5 years, his usual place of residence (UPR) was different from the location of the survey for a minimum of 6 months. Additionally, he must have received more than ₹4000/month as the value of produce from agricultural activities; otherwise, they are classified as a resident farmer (NSO 2022). To gather data, a structured interview schedule was employed, and a total of 480 migrant and resident farmers each were interviewed face-to-face during 2022–23.

The primary objective was to investigate the income disparities between migrant and resident farmers by examining the factors influencing agricultural income. In pursuit of this objective, we adopted the push and pull factors theory of migration (Lee 1966) as the theoretical framework to contextualize migration in the agricultural sector. This theory offers valuable insights into the drivers as well as motivations behind migration decisions and is widely recognized as a foundational framework for understanding

migration patterns and their implications across various domains. By integrating this theoretical perspective into our study, we aimed to provide a comprehensive analysis of the complex interplay between migration dynamics and agricultural income.

The study employed machine learning algorithms, namely classification and regression tree (CART), and bigram text mining (BTM) to inquire into agricultural employment prospects for migrants in Northern India. The data analyses were carried out using the *rpart* and *tidyverse* packages in *R* software version 4.2.2. CART is a type of decision tree algorithm used for both classification and regression tasks. The basic structure of a CART model involves recursively partitioning the feature space into distinct regions, each associated with a predictive model (Loh 2014). For this study, the target variable was the monthly household farm income generated by resident farmers. Furthermore, the resulting decision tree are visually interpretable to evaluate the important variables in determining agricultural occupation opportunities for migrants.

Classification tree: It aims to predict the class label 'y' based on a set of input features *X*. For each node 'm' in the tree, let *R_m* represent the region of feature space associated with node 'm'. Let *pm(k|X)* denote the predicted probability of class 'k' for the observations in region *R_m*. The splitting criterion typically involves measures such as Gini impurity or entropy. The algorithm aims to recursively partition the feature space by selecting the split 's' that maximizes the reduction in impurity. The predicted probability *pm(k|X)* for a given observation *X* within region *R_m* is often estimated as the proportion of training observations in region *R_m* belonging to class 'k'.

Regression tree: It aims to predict a continuous target variable 'y' based on a set of input features *X*. For each node 'm' in the tree, let *R_m* represent the region of feature space associated with node 'm'. Let *cm* denote the predicted value for observations in region *R_m*. The splitting criterion typically involves measures such as mean squared error (MSE). The algorithm aims to recursively partition the feature space by selecting the split 's' that minimizes the prediction error. The predicted value *cm* for a given observation *X* within region *R_m* is often estimated as the average (or median) of the target variable for training observations in region *R_m*.

$$\text{Coefficient of determination (R}^2\text{)} = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\text{Mean square error} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

where *x*, Input features.

In BTM, the focus is on extracting pairs of consecutive words from a given text corpus and analyzing their frequencies for various natural language processing tasks. Pointwise mutual information (PMI) score, which measures the association strength between two words occurring

Table 1 Operationalization of variables in the study

Variables	Description of measurement	Mean (*SD)	
		Migrants	Residents
Age (No. of years)	Number of chronological years since the birth	39.08 (5.82)	48.5 (7.7)
Farmer type	Landless farmers	3.5%	0.6%
	Marginal farmers	60.8%	42.7%
	Small farmers	31.9%	45.0%
	Semi-medium farmers	31.9%	11.7%
Farming experience (No. of years)	Practicing farming as an income-generating activity for a household is measured in years	13.77 (3.94)	24.4 (4.4)
Operational landholding (ha)	The area is actively utilized and managed by the farmer for agricultural purposes	0.7 (0.4)	1.2 (0.5)
<i>Kharif</i> yield (q/ha)	The production per hectare of a crop obtained during the <i>kharif</i> season	18.2 (11.9)	34.7 (12.4)
<i>Rabi</i> yield (q/ha)	The production per hectare of a crop obtained during the <i>rabi</i> season	25.3 (7.5)	35.6 (6.1)
Milk yield (litres/ animal)	The production of milk per animal was obtained during a year	900 (500)	1200 (800)

*SD, Standard deviation in the given variable; %, Percent of farmers belong to the given class

together was used. The BTM method was utilized to analyse the diverse and profitable ventures of resident farmers and the suggestions provided by Krishi Vigyan Kendra (KVK) professionals to enhance vocational opportunities for migrants in their location-specific context. Given a bigram w_i and w_j , and their frequencies $f(w_i)$ and $f(w_j)$, and the frequency of their co-occurrence $f(w_i, w_j)$ within a corpus, the PMI score is calculated as:

$$PMI = \log_2 \left(\frac{f(w_i) \times f(w_j)}{f(w_i, w_j)} \right) \times N$$

where $f(w_i, w_j)$, Frequency of the bigram w_i and w_j occurring together in the corpus; $f(w_i)$ and $f(w_j)$, Frequencies of the individual words w_i and w_j , respectively; N , Total number of tokens (words) in the corpus.

RESULTS AND DISCUSSION

The study found significant differences in socio-economic variables between migrant and resident farmers in northern India (Table 1).

Classifying factors contributing to high income generation: The classification tree of the machine learning algorithm CART was employed to analyse and classify the factors contributing to the high-income generation of resident farmers. The target variable, monthly farm income, was classified into binary output, where the decision “no” was taken by the algorithm if resident farmers earned less than fifteen thousand, and “yes” otherwise (NSO 2022). The

“yes” decision signifies the high income earned by resident farmers, leading to better and diversified opportunities in agriculture. Conversely, the “no” decision indicates less than 15,000 income earned by the resident farmers, resulting in a more probability to earn. The dataset of 480 respondents was divided into two halves for training and testing the model to classify the factors. The training dataset ($n = 240$) was used to first train the classification tree, which was then further tested to predict the model for determining significant factors. The significant factors that contributed to building the classification tree were milk yield (m^3 /animal per year), operational landholding of residents (ha), and *rabi* crop yield (q/ha) as visualised in Fig. 1 and the probability of outcome with predication is illustrated in Table 2.

Significant contributing factors employed by the classification model for predicting high income are given

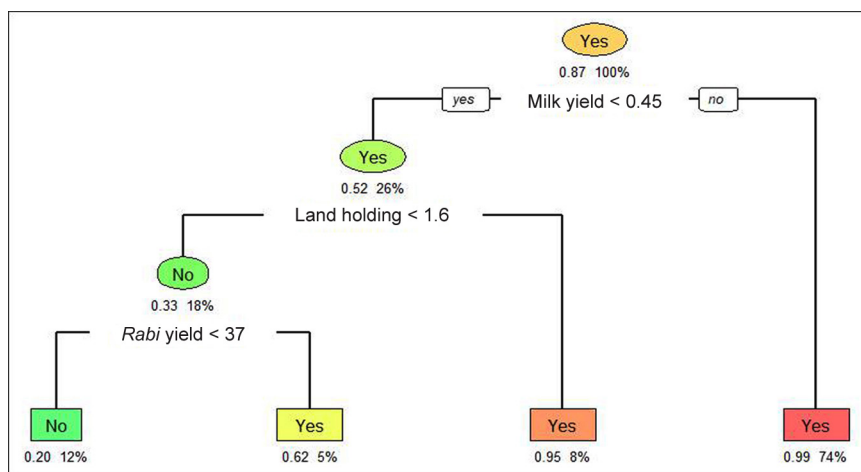


Fig. 1 Classification tree predicting significant variables for high monthly household income.

Table 2 Predictions of classification tree for high monthly farm household income

Probability of outcome	Conditions for outcome	n = 480 (%)	Prediction
0.99	When milk yield $\geq$ 4500 litres per animal in a year.	74	High-income
0.95	When milk yield $<$ 4500 litres per animal in a year and land holding $\geq$ 1.6 (ha).	8	High-income
0.62	When milk yield $<$ 4500 litres per animal in a year and land holding $<$ 1.6 (ha) and <i>Rabi</i> yield $<$ 37 (q/ha).	5	High-income
0.20	When milk yield $<$ 4500 litres per animal in a year and land holding $<$ 1.6 (ha) and <i>Rabi</i> yield $\geq$ 37 (q/ha).	12	Less-income

in Table 2 provides the conditions of profit maximization for migrants. These findings suggest that optimizing milk yield within the range of 4000 to 4500 litre/animal annually can significantly enhance the annual income for migrant farmers engaged in agricultural activities. Fine-tuning milk production to this threshold may serve as a strategic approach to strengthening income stability as well as fostering sustainable livelihoods within migrant communities. These opportunities will help in generating 1.5 to 2 lakh income/annum to agricultural migrants. The results indicate consistent metrics values during the training and testing phases. The accuracy showed an improvement of 4.1%, while the reliability increased by 17.8%. Therefore, this model can aid in enhancing the farm profit for migrant farmers. The model performance of the classification tree, is presented in terms of accuracy, reliability (Kappa), sensitivity, and specificity (Table 3).

The realm of sustainable agricultural employment opportunities for migrants recognizes the importance of promoting inclusive growth, reducing poverty, and ensuring the well-being of both migrants and the agricultural sector. The income disparity between migrant and resident farmers serves as a primary economic impetus compelling farmers to migrate aligns with push and pull theory (Lee 1966). Kundu and Das (2019) highlight the rural-to-urban migration trends driven by the dearth of employment opportunities and suboptimal agricultural incomes in India. The significance of factors such as milk yield, operational land holding, and yields from both *kharif* and *rabi* crops in generating income for migrant farmers emphasizes the multifaceted nature of sustainable agricultural employment. The role of livestock and milk yield is crucial as part of coping mechanisms against natural adversities, such as drought and biotic and abiotic stresses, thereby mitigating the need for temporary migration (Sarkar *et al.* 2019). Jokisch (2002) also found

that farmers with small operational landholdings migrate without significant agricultural investment, instead shifting towards housing and land, thus transforming the region into a peri-urban landscape. Thus, changing human-environment landscape in peri-urban region is becoming one of the frontier areas with new changes influenced by the migrants, and posing challenges to the sustainability of agricultural production (Marshall and Dolley 2019). As among the migrants the younger were predominant, it is imperative to learn that age, as a factor, plays pivotal role in taking decision whether to migrate from rural to urban region. Therefore, the developmental agencies and policy makers need to learn how the decision process of younger migrants are influenced by the gender and their choices of livelihoods contextualized with multipole stress (Resurreccion *et al.* 2019). The size of operational landholdings directly determines the yields of *rabi* and *kharif* crops (Azumah *et al.* 2022). The transitional nature of agricultural families, who strategically navigate livelihood vulnerabilities and migration decisions as adaptive strategies, resonate with the findings of Singh and Basu (2020). These dynamic stresses the complex interplay of socio-economic factors influencing migration patterns and livelihood choices among agricultural communities.

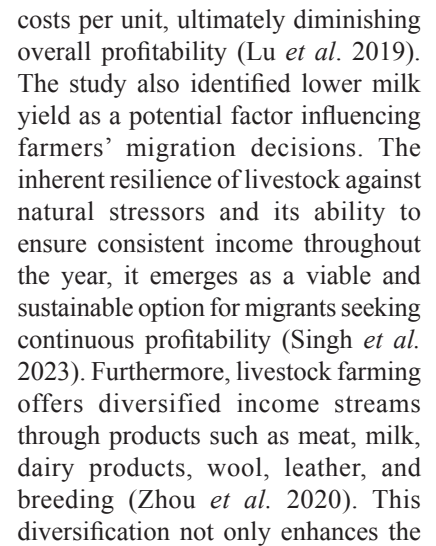
Predicting the significant factors of a profitable venture using a regression tree: The significant factors contributing to monthly household farm income for resident farmers were analysed using regression tree of the CART method for predicating vocational opportunities for migrant farmers. The training dataset (n = 240) was utilized to initially train the regression tree, which was subsequently tested to predict the model for determining significant factors. The significant factors that contributed to building the regression tree were milk yield (m^3/animal per year), *rabi* crop yield, and *kharif* crop yield (q/ha) as depicted in Fig. 2.

The predictive outcomes for receiving more monthly household income, using regression tree analysis are presented in Table 4. The condition with the highest coverage of respondents forecasted that achieving a sustained income for migrant farmers is facilitated by a milk yield exceeding 1600 litres per animal annually, along with *rabi* and *kharif* crop yields surpassing 22 and 16 q/ha, respectively. These observations were based on the milk productivity of milching animals and associated resources available at the farmer's level. These prospects will aid in creating 3.0 to 3.5 lakh returns per annum to agricultural migrants. The performance

Table 3 Model performance of classification tree

Statistics of class: Yes	Training model	Testing model
Accuracy	0.9083	0.9458
Kappa	0.6413	0.7559
Sensitivity	0.965	0.9713
Specificity	0.625	0.7742
P-value	0.000***	0.000***

***Significant at P -value $<$ 0.001.



overall financial security of migrants, but also contributes to the resilience of their livelihoods in the face of climatic uncertainties as well as market volatility (Mulwa and Visser 2020).

Exploring profitable ventures for migrant farmers through bigram text mining: In order to analyze location-specific employment opportunities of resident farmers from Rajasthan and Uttar Pradesh for continuous income generation among migrant farmers, a network analysis of bigrams was accompanied. Fig. 3 depicts a network graph of bigrams on suggestions of both residents and KVK professionals, representing the employment prospects for stakeholders. The network consists of nodes representing income generation activities and connections between these nodes based on the degree of centrality of job prospects. The colour of the links represents the strength of connections. The network analysis revealed three thematic clusters that exhibited the highest number of connections (Fig. 3). The first thematic cluster revolves around the production of vegetables, flowers, fruits, mushrooms, pulses, and oilseeds.

of the regression tree in predicting the model, demonstrates consistent fitting of observed and predicted values of outcome variables in both the training and testing datasets (0.45 and 1.97e-15, respectively). The decreased root mean square error (RMSE) in the testing model dataset indicates the robust performance of the model.

60

The second cluster is centered on integrated, organic, and natural farming practices in conjunction with sheep and goats. The third theme emerges from the management of agricultural residue and livestock. The network analysis provides valuable insights into the interconnections between different employment opportunities and helps visualize the relationships among them. These thematic clusters highlight additional potential agricultural avenues for profit generation among migrant farmers, and also offer valuable insights for stakeholders seeking to optimize agricultural productivity and sustainability. Moreover, the alignment of Lee's theory of push and pull factors with our findings emphasizes the significance of considering vocational opportunities in the UPRs for agricultural migrants.

Conclusions and policy implications

Migrations from rural to urban area has been one of the major concerns in the changing socio-ecological and climatic scenario. This study could lead to learn that the vocational opportunities in agricultural employment for migrants in northern India has a complex dynamic. These migrants face challenges associated with accessing and improving employment opportunities, and thus their overall livelihood strategies are impacted. This suggests that interventions aimed at expanding opportunities in local agricultural sectors can serve as effective measures to mitigate migration pressures and facilitate agricultural development within the communities. The options explored by resident farmers could be utilized for the skill development programmes in enhancing migrants' employability and productivity within the agricultural sector. Further, the inclusion of migrant farmers in economic growth can reduce income disparities and alleviate poverty in agriculture. The key insights of this study may help in boosting the productivity and creating appealing livelihood opportunities for migrant farmers in northern India. To address the challenges faced by migrant farmers, policymakers can think to plan and implement capacity building programmes in the livestock management and integrated farming systems; facilitating the access to credit and digital resources, and ensuring the social protection for migrant farmers as well as farm-women. Despite the valuable insights gained from this study, this was limited to northern India and subjective to the personal bias of resident as well as migrant farmers. Therefore, generalization of results is subjected to the similar conditions and opportunities of residents and migrants related to agricultural sector.

Ethical statements

This research study adheres to ethical standards, ensuring informed consent, participant anonymity, and integrity in data collection and reporting for agricultural migrants in northern India.

ACKNOWLEDGEMENTS

The authors acknowledge the ICAR-Indian Agricultural Research Institute and the Indian Council of Agricultural

Research, New Delhi, India, for providing the necessary facilities in the research. Without the respondents' participation and information, this study was not possible, so authors acknowledge their inputs greatly.

REFERENCES

- Azumah S B, Mahama A, Yegbemey R N and Dapilah F. 2022. Climate perception, migration and productivity of maize farmers in Ghana. *Journal of Agricultural Studies* **10**(1): 82.
- GMDAC. 2024. Food security and human mobility. IOM Global Migration Data Analysis Centre, Berlin. Retrieved from: <https://www.migrationdataportal.org/food-security> (Accessed on 3rd Nov 2023).
- Jokisch B D. 2002. Migration and agricultural change: The case of smallholder agriculture in highland Ecuador. *Human ecology* **30**: 523–50.
- Kaika A and Racelis A. 2021. Civic agriculture in review: Then, now, and future directions. *Journal of Agriculture, Food Systems, and Community Development* **10**(2): 1–22.
- Keshri K and Bhagat R B. 2012. Temporary and seasonal migration: Regional pattern, characteristics and associated factors. *Economic and Political Weekly* **47**(4): 81–88.
- Kundu A and Das S. 2019. Push and (or) pull? Drivers of labour force participation in Indian agriculture. *Indian Journal of Labour Economics* **62**(3): 413–30.
- Lee E S. 1966. A theory of migration. *Demography* **3**: 47–57.
- Loh W Y. 2014. Fifty years of classification and regression trees. *International Statistical Review* **82**(3): 329–48.
- Lu H, Xie H and Yao G. 2019. Impact of land fragmentation on marginal productivity of agricultural labour and non-agricultural labour supply: A case study of Jiangsu, China. *Habitat International* **83**: 6572.
- Marshall F and Dolley J. 2019. Transformative innovation in peri-urban Asia. *Research Policy* **48**(4): 983–92.
- Mulwa C K and Visser M. 2020. Farm diversification as an adaptation strategy to climatic shocks and implications for food security in northern Namibia. *World Development* **129**: 104906.
- NSO. 2022. Migration in India 2020–2021 (Periodic Labour Force Survey, July 2020–June 2021), pp. 25. National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India.
- Parmaceli D M. 2019. Production efficiency in unstable agriculture: Methodological aspects. *International Agricultural Journal* **1**: 24–28.
- Resurrección B P, Goodrich C G, Song Y, Bastola A, Prakash A, Joshi D, Liebrand J and Shah S A. 2019. In *The Shadows of the Himalayan Mountains: Persistent Gender and Social Exclusion in Development*. Wester P, Mishra A, Mukherji A and Shrestha A (Eds). The Hindu Kush Himalaya Assessment-Springer. https://doi.org/10.1007/978-3-319-92288-1_14
- Saini S, Burman R R, Padaria R N, Mahra G S, Bishnoi S, Aditya K, Nithyashree M L, Mallick S, Mukherjee S and Padhan S R. 2023a. Mapping the research trends of migration behaviour in agricultural households: A bibliometric analysis. *Frontiers in Sustainable Food Systems* **7**: 1241716.
- Saini S, Jirli B and Padhan S R. 2023b. Analysis of factors promoting the usage of electronic national agriculture market in Rajasthan, India. *Current Science* **125**(6): 643–48.
- Saini S, Burman R R, Padaria R N, Mahra G S, Bishnoi S, Aditya K, Nithyashree M L, Padhan S R and Mallick S. 2024. Overcoming constraints through innovative agricultural ventures in the agro-

- ecological regions of migrant farmers. *Journal of Community Mobilization and Sustainable Development* **19**(special issue): 50–56.
- Sarkar A, Dasgupta A and Sensarma S R. 2019. Climate change and food security in India: Adaptation strategies and major challenges. *Sustainable Solutions for Food Security: Combating Climate Change by Adaptation* **2019**: 497–520.
- Singh C and Basu R. 2020. Moving in and out of vulnerability: Interrogating migration as an adaptation strategy along a rural–urban continuum in India. *The Geographical Journal* **186**(1): 87–102.
- Singh R, Maiti S, Garai S and Rachna. 2023. Sustainable intensification-reaching towards climate resilience livestock production system: A review. *Annals of Animal Science* **23**(4): 1037–047.
- Sloat L L, Davis S J, Gerber J S, Moore F C, Ray D K, West P C and Mueller N D. 2020. Climate adaptation by crop migration. *Nature Communications* **11**(1): 1243.
- Zhou H, Yan J, Lei K, Wu Y and Sun L. 2020. Labour migration and the decoupling of the crop-livestock system in a rural mountainous area: Evidence from Chongqing, China. *Land Use Policy* **99**: 105088.