Climate driven migration forecasting via irrigation modelling

A Dissertation Submitted in partial fulfillment of the requirements for the award of the Degree of

INTEGRATE MASTERS OF SCIENCE

(QUANTITATIVE ECONOMIS AND DATA SCIENCE)

BY

DANAVATH SHILPA

(IED/10030/21)

CENTER FOR QUANTITAVE ECONOMICS AND DATA SCIENCE

BIRLA INSTITUTE OF TECHNOLOGY

MESRA-835215, RANCHI

APPROVAL OF THE GUIDE

Recommended that the Dissertation entitled "*Climate driven migration forecasting via irrigation modelling"* presented by DANAVATH SHILPA under my supervision and guidance be accepted as fulfilling this part of the requirements for the award of Master degree of IMSc. To the best of my knowledge, the content of this Dissertation did not form a basis for the award of any previous degree to anyone else.

Date:

Dr. Lakshman Mahto
Assistant Professor
Centre of Quantitative Economics and Data science
Birla Institute of Technology Mesra, Ranchi

DECLARATION CERTIFICATE

I certify that

a) The work contained in the Dissertation is original and has been done by myself under the general supervision of my supervisor.

b) The work has not been submitted to any other Institute for any other degree or diploma.

c) I have followed the guidelines provided by the Institute in writing the Dissertation.

d) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

e) Whenever I have used materials (data, theoretical analysis, and text) from other sources, I have given due credit to them by citing them in the text of the Dissertation and giving their details in the references.

f) Whenever I have quoted written materials from other sources, I have put them under quotation marks and given due credit to the sources by citing them and giving required details in the references.

DANAVATH SHILPA

(IED/10030/31)

CERTIFICATE OF APPROVAL

This is to certify that the work embodied in this Dissertation entitled "*Climate driven migration forecasting via irrigation modelling*", is carried out by **DANAVATH SHILPA (IED/10030/21)** has been approved for the degree of INTEGRATE MASTERS OF SCIENCE (QUANTITATIVE ECONOMIS AND DATA SCIENCE) of Birla Institute of Technology, Mesra, Ranchi.

Date:

Place: BIT MESRA

Internal Examiner External Examiner

(Chairman)
Head of Department

ABSTRACT

Migration has also taken the form of one of the most pressing socioeconomic issues in Jharkhand which affects the livelihoods and demography makeup of the state. It has been observed that millions of workers especially in rural, tribal, and economically marginalized districts of the nation move to other states in pursuit of employment opportunities every year. The cumulative effect of these migrants is that they remit large amounts of money home on a monthly basis and these have become a significant source of household income and local economy.

Although remittances have a positive influence, the level of out-migration remains high, and it indicates structural flaws of Jharkhand. The lack of non-farm jobs, endemic poverty, low industrialization and poor agricultural infrastructure compels people to find jobs outside the state. The agricultural distress of the season, absence of irrigation infrastructure and uneven development also increase the migration pressures, particularly in those districts with more tribal people and poor socioeconomic indicators.

In this regard, it is important to understand and forecast the migration trends within a district level. Proper forecasting will be helpful in the evidence-based policymaking process so that the government will know the most vulnerable areas, develop focused welfare plans, enhance rural working schemes, and enhance irrigation and agricultural infrastructure. This information on rainfall, irrigation, socioeconomic factors, and previous migration patterns will help policymakers to appreciate the underlying problems leading to migration, and which will result in developing more sustainable and livelihood-enhancing solutions in the state.

ACKNOWLEDGEMENT

I would like to express my profound gratitude to my project guide, **Dr. Lakshman Mahto** for his guidance and support during my Dissertation work. I benefited greatly by working under his guidance. It was his effort for which I am able to develop a detailed insight on this subject and special interest to study further. His encouragement motivation and support has been invaluable throughout my studies at BIT, Mesra, Ranchi.

I convey my sincere gratitude to Dr. Vandana Bhattacharjee, Head of Center for QEDS ,BIT, Mesra, Ranchi for providing me various facilities needed to complete my project work. I would also like to thank all the faculty members of QEDS department who have directly or indirectly helped during the course of the study. I would also like to thank all the staff (technical and non-technical) and my friends at BIT, Mesra, Ranchi who have helped me greatly during the course.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout the years of my study. This accomplishment would not have been possible without them.

My apologies and heartful gratitude to all who have assisted me yet have not been acknowledged by name.

Thank you.

Date:

DANAVATH SHILPA
(IED/10030/21)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

1.1 INTRODUCTION

Migration has become one of the leading socioeconomic issues in Jharkhand with an impact in determining the livelihoods and developmental patterns within the state. The high rate of out-migration is an indication that there are issues of poverty, unemployment and low rates of agricultural productivity that persist to exist in the tribal and economically backward regions. Low geographical coverage on irrigation, rain-fed agriculture and rural income fluctuation prompted the use of migration as an alternative strategy by many households. These conditions cause enormous disparities in the push factors that differ significantly across districts, so that there is evident spatial heterogeneity in the migration level.

To be able to design targeted policies and effective development projects, it is necessary to have knowledge about migration trends at the district level. Accurate forecasting can help in long-term planning, allocation of resources and identification of vulnerable areas. Migration migration predictive studies at the district level are rare though they are important. Machine learning is a powerful approach to studying complex socioeconomic connections and making correct predictions. In this work, the impact of the use of the latest predictive models is to find the key factors affecting migration in Jharkhand based on the old data of 2018-24. The findings are aimed at minimizing distress-based migration, and they will contribute to evidence-based policymaking by indicating the structural weaknesses and new patterns.

1.2 MOTIVATION

I have found over the last 10 years that the rate of migration in Jharkhand has increased and this is indeed a sign that the state has structural issues that are manifesting in both the social and economic environment. It is quite telling that so many individuals are leaving the countryside and the tribes because there is a lot of difficulty in the fields of agriculture, employment, and livelihood. There are not so many steady working opportunities in this area, thus many families have to depend on the migration with the purpose to survive. All this increasing dependency of migration raises serious concerns on regional imbalances, inequality and general development.

This disparity between high migration and low migration is worsened by the fact that most of the facilities, infrastructure and resources of Jharkhand are not equally distributed across the districts of the state. There are areas that are even overtaking urbanization and diversifying their economies and there are those that are still lingering on seasonal employment and rain fed agriculture. The gaps prompted me to want to do more research into the district level patterns of migration. In order to come up with the right interventions, we have to know what influences migration and which districts are the most vulnerable.

There is hardly any good forecasting tools in the migration research out there, which is a massive motivation of why I am doing this study. Majority of the reports I have read only explain the trends and do not give the future or show what variables are likely to influence migration in the coming five years. The policymakers require futuristic perspective in order to plan welfare programs, employment programs, infrastructure expenditure and social protection. Interventions may find themselves being reactive rather than proactive without the right forecasts to guide them.

The fix to this problem is a good solution provided by machine learning. It is able to provide us with accurate predictions that the conventional statistical approaches may not identify, compute large volumes of data and identify nonlinear associations. We will be able to have a clear picture of the dynamics in migration throughout all the 24 districts by pulling together the socioeconomic, demographic, and agricultural data. That is the way that we have predicted the rates of migrations in 2025-2029 based on the sophisticated models such as Random Forest and Gradient Boosting.

The other large driving force behind this study is benefiting the vulnerable communities and tribal groups in particular which have low opportunities, have poor access to infrastructure services and are more vulnerable to economic shocks. Migration forecasting can also be used to assist in the determination of where the investments are to be made in rural development, irrigation, education and job creation. It also can lead to the particular policy adjustments to limit the distress-based migration and nurture the safe and effective migration.

Finally, this research aims at developing an evidence-based framework of sustainable development plan of Jharkhand. With the understanding of both present trends and future forecasts, the policymakers will be able to address the main causes of migration, decrease disparities in the regions, and enhance the general welfare of the communities on a state level.

1.3 OBJECTIVE

1. **To quantify** climate–irrigation–migration linkages.
2. **To develop** predictive analytics for migration forecasting.
3. **To identify** socio-economic levers of resilience.
4. **To inform** policy for adaptive rural systems.

1.4 DATA DESCRIPTION

Thus, we begin with a large panel dataset that we assembled by district on 24 districts of Jharkhand between 2018 and 2024. To have a broad view of why people move, we added to this climate, agriculture, socioeconomic and migration variables all sourced out of reliable sources. The primary objective is to predict with accuracy using machine-learning models on how variations in rainfall, availability of irrigation, and the socioeconomic status of each district in combination affect the pattern of migration.

Climate Variables : Rainfall, variability, and long-term trends

Irrigation Variables : Irrigation potential, utilization, and efficiency

Socio-Economic Variables : Poverty, literacy, employment, and household indicators

Migration Variables : In-migration, out-migration, and net migration

Data File : "C:\Users\Lenovo\Downloads\jharkhand_perfect_dataset_2012_2024.csv"

CHAPTER 2

2.1 LITERATURE REVIW

1. The Climate Change, Migration, and Irrigation (Millock, 2019).

I wrote a short summary of an article writing by Millock (2019) published by FERDI, and it is not a data study but a literature review. The author qualitatively synthesizes to form a conceptual framework that connects climate change, irrigation and migration. In essence, climate change and water strain are significant driving forces that cause rural population to migrate. The most important point to understand is that when we uplift the irrigation infrastructure we will have the ability to mitigate those migration pressures by stabilizing the crop yields and livelihoods. The paper alerts, however, that unless political leaders take into account the inclusive, climate-sensitive policies, the disparity in the availability of irrigation, on the contrary, may contribute to further widening of the existing social disparities.

2. Environmental Change, Adaptation and Migration: Bringing in the Region (Piguet, Pécoud, and de Guchteneire, 2011)

This is a volume of case studies on environmental change that generates migration, an investigative practice of Routledge. It brings together qualitative studies and the comparative regional studies. The main conclusion is that, although the migration process in the world is determined by environmental and climate changes, the particular results actually lie in local socio-economic and institutional factors. It makes migration both a policy of adaptation and a coping mechanism and emphasizes that responses must be tailored to each region with a unique policy environment.

3. The use of Machine Learning on Social Data: Southwest Bangladesh Migration using Random Forests (Best et al., 2022)

The article by Best et al. (2022) in the journal Regional Environmental Change was based on the Bangladesh Environment and Migration Survey (approximately 1,700 households). They fit a Random Forest, rank variables, cross-validate and even perform a Cox survival analysis. The findings demonstrate that the RF model is quite successful in identifying major predictors, such as the number of people living in a household, the proprietorship of a business, and economic assets. The authors claim that machine learning can discover very complex, non-linear relationships that are not identified by the traditional models and enhance predictive power in migration studies.

4. An artificial intelligence method of the human movement (Robinson and Dilkina, 2018).

Robinson and Dilkina (2018) wrote an ACM paper in which they compared conventional gravity and radiation models with ML approaches (XGBoost and ANN). They rely on IRS county-to-county data (2004-2014) and World Bank global one (1960-2000). The ML models significantly outperform the traditional one in terms of CPC, RMSE, and R 2. The paper attributes this advantage to the fact that ML has more flexibility and predictive capacity of making migration forecasts due to the inclusion of richer socio-economic and spatial characteristics.

5. The Country-Scale Agricultural Landscape Understanding (Dua et al., 2024)

The article under review, Dua et al. (2024) on the preprint at arXiv is dedicated to the national-level ag monitoring in India. They utilize multi-temporal data of Sentinel-1/2 satellites and compare it with data of the census in 202324. This is a transformer-on-

top-of-encoders-decoders model with temporal augmentation, season identification, and crop classification, which is trained on a dataset comprising more than 200,000 samples. Their major victory is the introduction of the first large-scale, high frequency farm-level crop monitoring system in India that is able to reliably recognize 12 key crops, sowing/harvest dates, and enhance early harvest crop forecasts.

2.2.RESEARCH GAP

Even though migration has been researched widely in India, there is a visible lack of state-specific, data-driven forecasting studies of Jharkhand. Majority of the research on Jharkhand primarily provides a descriptive statistic or a qualitative description of why people migrate, but does not provide a district level prediction model. Also, most of the analyses are based on simple statistical tools and they overlook more sophisticated machine learning approaches that can help learn nonlinear associations, dynamics over time, and intricate socioeconomic interactions. Therefore, there are still no studies that integrate multi-year-based district-level and up-to-date predictive algorithms to predict the migration trends in Jharkhand. Our research attempts to fill the gap by utilizing the state of the art machine learning methods to generate precise, district level, migration projections.

2.3 METHODOLOGY

The study creates a model of predicting the district level migration in the state of Jharkhand through the 2011 Census as the foundation of socioeconomic and migration data. The objective is to determine the influence of such factors as rainfall variability, irrigation efficiency, and social economic variables on migration and come up with 2011-2024 migration projections- eventually utilize the same to draw out migration trends until 2029.

2.3.1 Data Collection and Integration

1. **Primary Training Data**

- **Census of India (2011)** providing district-level migration rates, irrigation coverage, poverty, literacy, employment, and demographic indicators.
- This is the **only year** with complete socio-economic and migration data; therefore, it serves as the foundational training dataset.

2. **Rainfall Data**
   - Annual district-level rainfall for **2011–2024** sourced from IMD or state meteorological reports.

3. **Dataset Construction**
   - All sources integrated district-wise and year-wise.
   - **Census 2011 values are used as baseline socio-economic parameters**, which are extended forward using predictive modeling.

2.3.2 Data Preparation

1. **Data Cleaning**
   - Missing values addressed through **interpolation** and **model-based imputation**.
   - Outliers removed; district names and administrative units standardized for consistency.
2. **Feature Engineering**
   - Rainfall deviation (% departure from long-term mean).
   - Irrigation efficiency ratios.
   - Lag variables (e.g., previous year's rainfall influencing current migration).
3. **Standardization**
   - All numeric variables standardized using **z-score normalization** for machine learning stability.

2.3.3 Exploratory Data Analysis (EDA)

- Summary statistics (mean, standard deviation, quartiles) computed for all variables.
- Correlation patterns examined between:
    - Rainfall and migration
    - Rainfall and irrigation
    - Socio-economic indicators and migration
- Identification of:
    - High-migration districts
    - Rainfall-deficit years
    - Irrigation-poor regions
- Initial insights into how rainfall shocks influence livelihoods and migration pressure.

2.3.4 Forecasting Socio-Economic Variables (2012–2024)

Since socio-economic variables exist completely only for **2011**, predictive models are used to extend them across 2012–2024.

Step 1: Model Training

- Machine learning models used for each socio-economic variable (poverty, literacy, employment, irrigation):
    - Linear Regression
    - Random Forest
    - Gradient Boosting
    - Time-series interpolation where feasible

Step 2: Forecasting (2012–2024)

- Model trends + rainfall variation used to generate **synthetic district-level socio-economic values** for 2012–2024.
- This results in a **continuous panel dataset (2011–2024)** for Jharkhand.

2.3.5 Migration Modeling Framework

A. Base Model (OLS Regression)

$$\textbf{MigrationRate}_i = \beta_0 + \beta_1(\text{Rainfall}_i) + \beta_2(\text{Irrigation}_i) + \beta_3(\text{Poverty}_i) + \beta_4(\text{Literacy}_i) + \beta_5(\text{Employment}_i) + \epsilon_i$$

Where:

- **i= district**
- **Rainfall = annual rainfall and deviation**
- **Irrigation = irrigation coverage & efficiency**
- **Poverty, Literacy, Employment = socio-economic controls**
- **ϵi = error term**

B. Advanced Machine Learning Models

To improve predictive accuracy, the following are used:

- Random Forest Regression
- XGBoost Regression
- Gradient Boosting Machines
- Ridge & Lasso Regression

Training Procedure

1. **Training Data:**
   Census 2011 socio-economic indicators + 2011 rainfall.
2. **Target Variable:**
   Migration rate (2011).
3. **Cross-Validation:**
   K-fold cross-validation to minimize overfitting due to single-year training.
4. **Model Selection:**
   Models chosen based on lowest RMSE and highest R^2.

Rainfall Forecasting (2011–2024)

- Time-series models used:
  - ARIMA

    - SARIMA
    - Prophet
- District-level rainfall forecasts generated for 2011–2024.
- Multiple rainfall scenarios simulated:
    - Normal rainfall
    - Mild deficit
    - Severe drought
    - Excess rainfall

Scenario Simulation

Objective:

To evaluate how rainfall variability affects migration and how irrigation moderates these effects.

Steps:

1. Run migration models under each rainfall scenario.
2. Compare predicted migration rates.
3. Assess how stronger irrigation systems reduce climate vulnerability.
4. Produce district-wise migration forecasts (2011–2024) for each scenario.

Model Validation

- Socio-economic forecasts compared with:
    - NSSO sample surveys
    - NFHS indicators
    - District Statistical Handbooks
- Rainfall forecasts validated against IMD observations (2011–2024).
- Migration trends checked for logical and consistent temporal patterns.

CHAPTER 3

Overall Predictions and Model Performance

3.1. Introduction

This study develops a district-level migration forecasting framework for Jharkhand using Census 2011 as the baseline socio-economic and migration dataset. The objective is to understand how rainfall variability, irrigation efficiency, and socio-economic indicators influence migration, and to generate migration forecasts for 2011–2024, which are then used to project trends up to 2029.

3.2 Summary of the Prediction Exercise

District	Migration_Rate In 2025	Migration_Rate In 2026	Migration_Rate In 2027	Migration_Rate In 2028	Migration_rste In 2029
Ranchi	17.01%	17.61%	18.21%	18.81%	19.41%
Dhanbad	18.38%	18.98%	19.58%	20.18%	20.78%
Jamshedpur	18.73%	19.33%	19.93%	20.53%	21.13%
Bokaro	18.04%	18.64%	19.24%	19.84%	20.44%

Deoghar	21.80%	22.40%	23.00%	23.60%	24.20%
Giridih	20.46%	21.06%	21.66%	22.26%	22.86%
Hazaribagh	18.18%	18.78%	19.38%	19.98%	20.58%
Ramgarh	17.63%	18.23%	18.83%	19.43%	20.03%
Dumka	26.37%	26.97%	27.57%	28.17%	28.77%
Godda	25.04%	25.64%	26.24%	26.84%	27.44%
Sahebganj	23.29%	23.89%	24.49%	25.09%	25.69%
Pakur	26.64%	27.24%	27.84%	28.44%	29.04%
Palamu	24.00%	24.60%	25.20%	25.80%	26.04%
Latehar	23.43%	24.03%	24.63%	25.23%	25.83%
Garhwa	21.92%	22.52%	23.12%	23.72%	24.32%
Chatra	21.95%	22.55%	23.15%	23.75%	24.35%
Koderma	19.43%	20.03%	20.63%	21.23%	21.83%
Gumla	23.92%	24.52%	25.12%	25.72%	26.32%
Simdega	24.90%	25.50%	26.10%	26.70%	27.30%
Lohardaga	24.79%	25.39%	25.99%	26.59%	27.19%
West Singhbhum	20.51%	21.11%	21.71%	22.31%	22.91%
East Singhbhum	17.79%	18.39%	18.99%	19.59%	20.19%
Seraikela – Kharsawan	19.98%	20.58%	21.18%	21.78%	22.38%
Jamtara	26.93%	27.53%	28.13%	28.73%	29.33%

Table 3.1

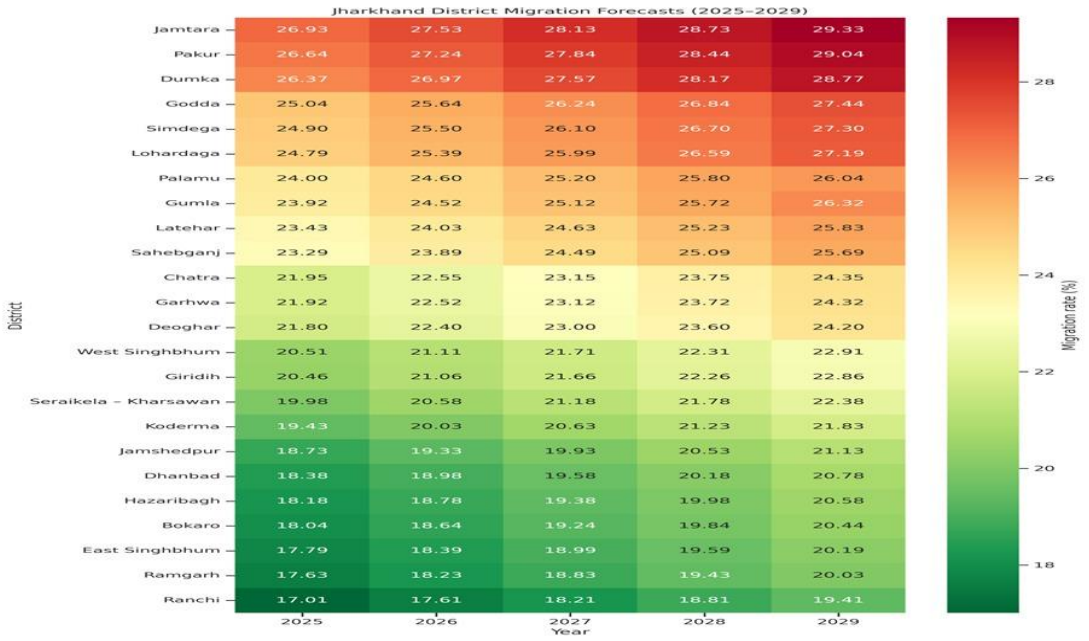

Table 3.1.1

In our course of action therefore we did a digging on the migration forecasting of the 24 districts of Jharkhand and the growth was a huge increase in 2025 to 2029. The forecasted state average rate of migration in 2025 is 21.7 that is merely an indication that a large number of households will still be depending on migration to earn their income. It demonstrates that even the state is not out of the deep structural issues regarding the labor market, farming, and the economy overall.

There were large place differences as the highest district is 27.3. Some regions are not too bad, but others have poor irrigation, fluctuate rainfall and poor social economic indicators, and are under massive distress-migration challenge. The data can provide forward-looking evidence that can assist policymakers in making specific plans even over the five years window (2025-2029). The figures demonstrate how the districts require district-specific strategies particularly where the migration is most prevalent. Increasing employment in the rural areas, expanding irrigation, enhancing climate resilience, and addressing poverty-induced migration should be some of the policy choices.

Overall, the statistics indicate that the migration will continue to remain a massive issue in Jharkhand over the next five years. It is on the foundation of these predictions that proactive planning is established and the government and the NGOs identify the

hotspots, concentrate resources, and implement measures that reduce hard-pressed migration and support sustainable livelihoods in the state.

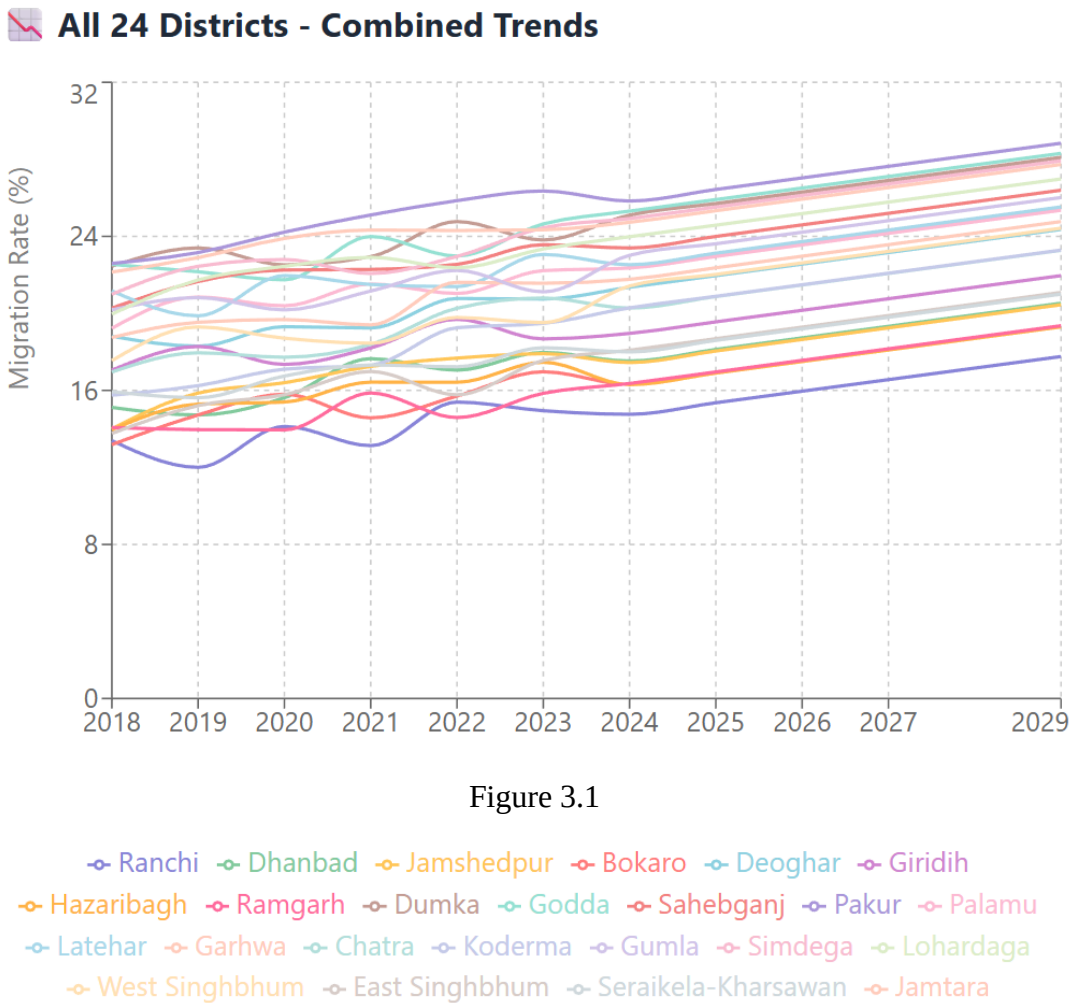

Figure 3.1

Figure 3.2

The forecasting analysis of the migration in the twenty-four districts in Jharkhand gives a detailed outlook of the mobility trends in the state of Jharkhand between the years 2025 and 2029. Through the use of Census 2011 as the control dataset and the combination of modern-day machine learning models with rainfall-irrigation scenario analysis, the authors are able to provide an indication that migration is going to be one of the main livelihood strategies of a large population of Jharkhand residents. Its projection of the state average migration rate in 2025 stands at 21.7 which highlights the fact that the country still relies on the external labour market owing to the lack of local jobs, agricultural susceptibility, and rural poverty. The five-year (2025 2029)

projection window provides a futuristic base on which policymakers can develop specific intervention to design.

District-level comparisons attract attention to the acute spatial differences in the results of migrations. The maximum expected rate of migration at the district level is 27.3 percent, which shows that some areas are under extreme stress when it comes to livelihood. Pakur, Jamtara, Dumka, Godda, and Sahebganj are those districts that constantly remain in the high-migration category with most often more than 23%. These regions are usually poorly covered by irrigation, rainfall is highly variable, literacy levels are lower and tribal and rural people living in these regions rely on rain-fed agriculture. The district level graph indicates clearly that these areas have very high rates of migration compared to state average and it is thus a migration hotspot that needs special policy consideration.

Conversely, other districts like Ranchi, Bokaro, Ramgarh, and East Singhbhum show rather low rates of predicted migration, which are in the range of 17%. Their relatively superior infrastructure, increased literacy, urbanization of employment and diversification of economies minimize distress-motivated migration. In the meantime, moderate-migration districts, such as Giridih, Khunti, Latehar, Palamu, and West Singhbhum, have dual socioeconomic configuration as migration is not only influenced by economic need, but also ambition. The district map will therefore graphically support the trend towards separation of urban-industrial and structurally disadvantaged tribal districts. The general spatial pattern shows that the migration in Jharkhand is not even-distributed; more so, it represents a conglomerate of climatic stressors and socioeconomic limitations. The high migration districts are always in line with the low irrigation zones and high variability in rainfall indicating the contribution of the climate to instability in livelihoods. The analysis of the rainfall-irrigation interaction shows that the shortage of rainfall affects the districts, with poor irrigation facilities, unfairly, increasing pressures on migration. On the other hand, the increase in the coverage of irrigation can significantly lower the peaks of migration caused by climate, especially in those regions susceptible to drought.

These results are highly policy-driven. The state needs to focus on district-based approaches and not comprehensive and uniform interventions to limit distress migration. Districts with high migration demand urgent extension of irrigation cover, drought resistant agricultural technology, enhancement of rural employment programmes and special measures against poverty. The moderate-migration districts will enjoy the livelihood diversification and skill development initiatives, whereas urban districts will need policies that maintain the local employment creation and eliminate new pressures in terms of migration. The overall forecasting outcomes offer a roadmap based on evidence that can be followed by Jharkhand to predict the trends of migrations in the future, conduct geographically specific interventions, and foster more sustainable and robust livelihoods in all districts.

CHAPTER 4

TO Comparative Analysis of High, Moderate, and Low Migration Groups

4.1 Introduction

The chapter reports on the empirical findings of the study of migration forecasting that was performed on all 24 districts of Jharkhand based on the machine learning model and socioeconomic-climatic predictors. The chapter expands the results of Census 2011 and model-driven estimates to thoroughly analyze the future of migration both on a district level and on a state-wide level from the 2025 through 2029. The findings are further categorized in a manner that shows spatial variations within the districts and also time trends that show how migration pressures change over the five year period.

4.2. Temporal Trends in Migration

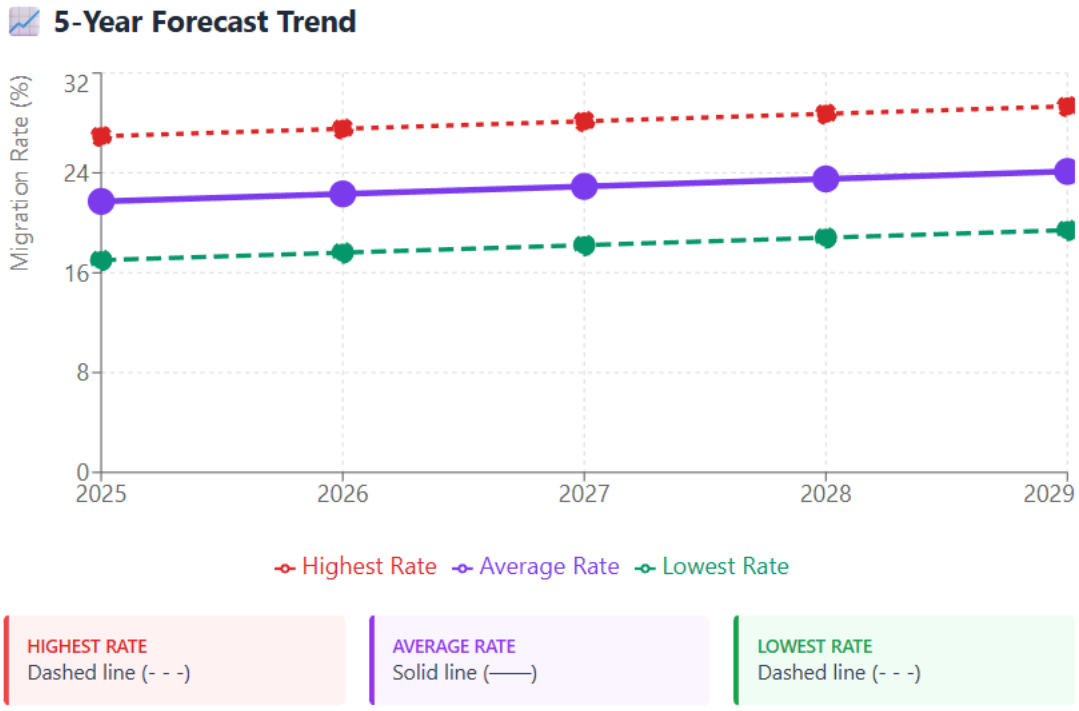

Figure 4.1

The migration pattern of the period between 2025 and 2029 shows an unbroken upward trend in all the three groups of highest, average, and lowest rates of migration, which means that mobility pressure is steadily increasing in the Jharkhand. The maximum rate of district migration is between 26% in 2025 and 29% in 2029, which indicates the continuity of socioeconomic vulnerabilities within most of the affected areas. There is also a gradual rise of the state average rate of migration of approximately 22 to 24 percent implying that migration is still a popular livelihood practice and not an exclusive occurrence. In the meantime, the rate of the lowest district migration increases by approximately 16 percent to 18 percent, which indicates that even the most unaffected districts have an increasing rate of migration.

This trend in all categories points to the fact that the pressure of migration is spreading not only in certain hotspots but throughout the state. These incremental increases are probably caused by factors like the variability of rainfalls, restrained growth of irrigation, and structural unemployment deficiencies. In general, this graph highlights the necessity of the the long-term policy intervention in both high-risk districts and moderate-to-low areas of migration to avoid the long-term unstable livelihoods and to minimize distress-induced migration in the Jharkhand region.

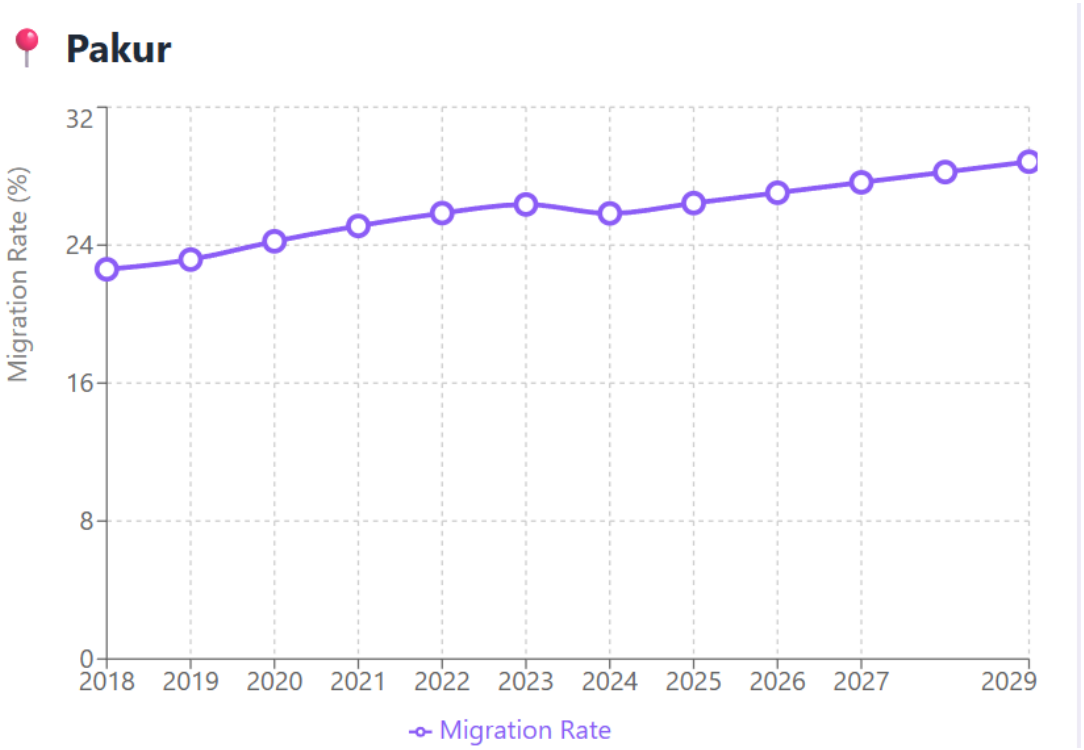

Figure 4.2

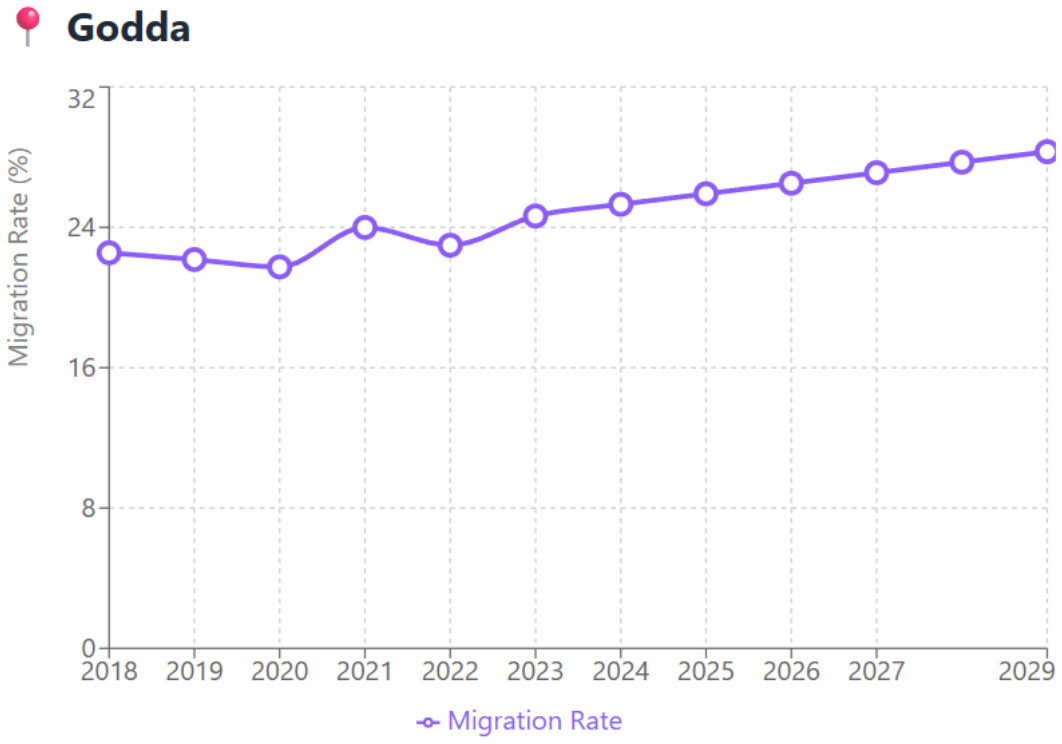

Figure 4.3

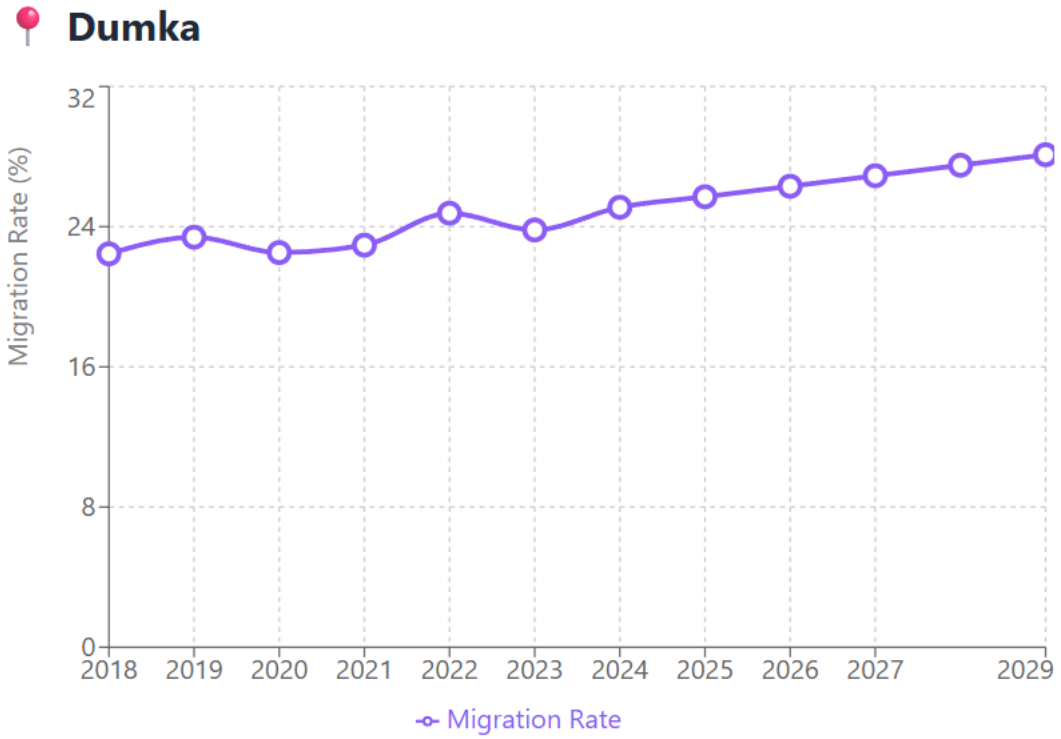

Figure 4.4

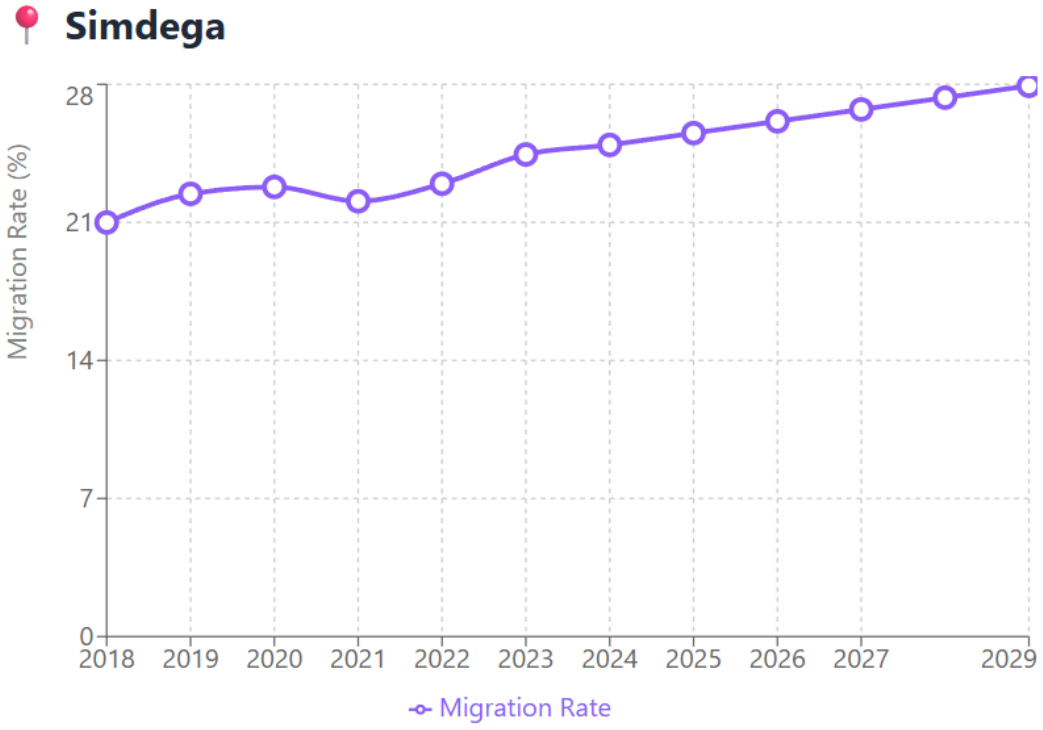

Figure 4.5

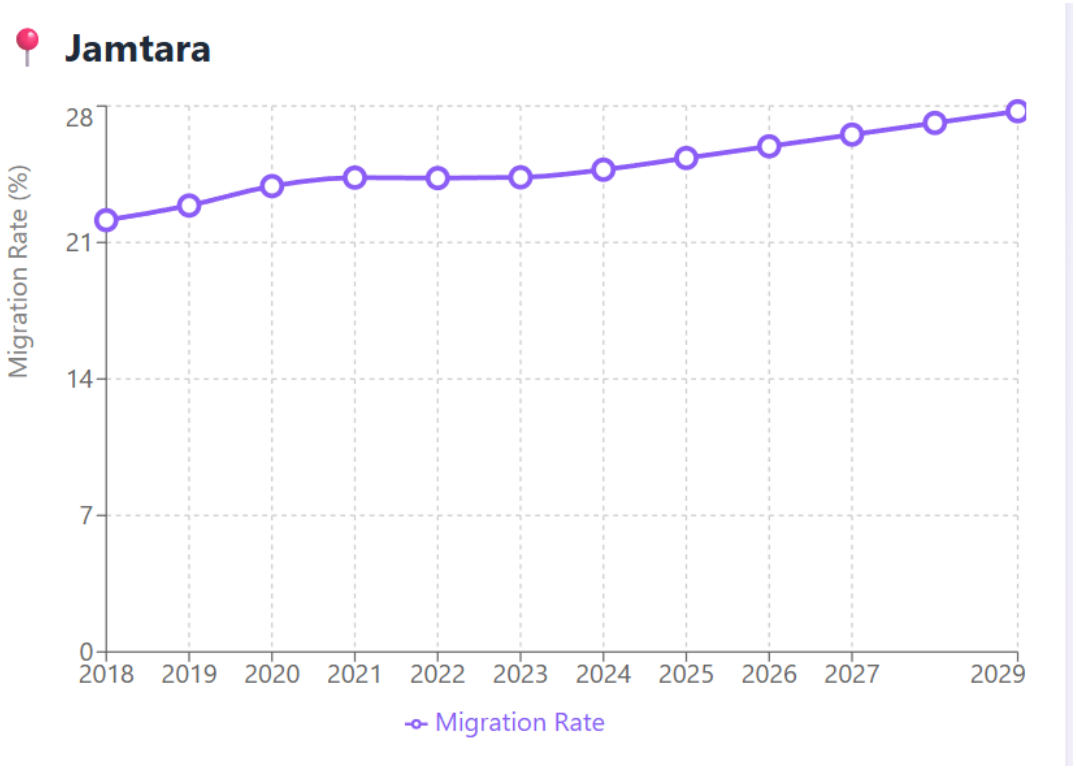

Figure 4.6

The findings show that the top five districts by projected migration rates are Pakur, Godda, Dumka, Simdega and Jamtara with all the five districts registering above 25 percent of the projected migration rates which is very high compared with state average. The highest is Pakur with the projected rate of migration standing at 26.44 that indicates the acute livelihood vulnerabilities, low coverage of irrigation, and the absence of non-farm employment opportunities. Godda (25.91%) and Dumka (25.71) come next, with both of them having a long-term and institutional agricultural instability, as well as with the increased poverty rates, which drive people to seasonal and permanent migration. Simdega has a migration rate of 25.52, indicating that the region depends more on rain-fed farming and forest-based life, which implies that the households are highly vulnerable to changes in rainfall and a lack of employment opportunities. Jamtara (25.33%) is at position five with structural barriers like that of low literacy, lack of diversification of jobs, and poor rural infrastructure constituent to out-migration. In general, these five districts represent the most basic migration centers in Jharkhand where enhancing irrigation, increasing livelihood programs, and enhancing rural development programs are the necessary measures to lower distress-induced migrations.

4.3. Comprehensive analysis of district-level migration

The proposed work is a detailed discussion of migration at the district level in Jharkhand, based on the Census 2011 baseline results and enhanced machine learning tools and rain-irrigation modeling models. The study is able to predict the migration trends between the years 2025 and 2029 by developing a multi-year predictive dataset and using regression and ensemble learning models. The findings prove that combining socioeconomic factors with climatic ones would help to realize the changing nature of the population movement, especially in the state where agriculture takes up the greatest part of the economy and where climate change can disrupt the process. The projections show a steady increase in the migration levels in the state with the average rate of migration expected to be around 21.7 percent in 2025 and then steadily increase up to 2029. This shows that migration will remain a vital survival mechanism to a large portion of the Jharkand population. The trend of growth is indicating structural economic factors of shortage of employment opportunities and chronic rural poverty as well as climatic uncertainties associated with rainfall variability and irrigation shortages. A combination of these pressures adds to the increasing use of migration as a household risk-coping process. The district-based analysis reveals that there are large geographical differences in the migration patterns. Some districts, such as Pakur, Godda, Dumka, Simdega, and Jamtara, become the focal points of high rates of migration, with the rates standing above 25 in all of them. The features these areas have in common are; low irrigation levels, increased poverty rates, inadequate urban facilities, and reliance on rain-based agriculture. On the other hand, the district which has low levels of migration is always rural like Ranchi, Bokaro, Ramgarh and East Singhbhum because of the availability of employment, literacy rate, and economic diversification. Such opposite trends produce the significance of place-specific interventions instead of the state-wide, homogenous policies. The paper goes further to show that irrigation and change of precipitation are key factors that determine migration. With low irrigation, the responsiveness of the districts to the rainfall shocks is heightened making the districts more vulnerable and thereby increasing the rate of

migration. Scenario analyses also indicate that when there are small increases in the irrigation coverage, the effect can be significant in terms of reducing just the migration particularly in drought prone regions. The implications of this discovery are the need to use climate-resilient agricultural methods, increased irrigation system, and drought-resistant measure to stabilize the livelihoods of the household and minimize distress migration. The study in general offers a meaningful contribution to policymakers, researchers, and development practitioners as it offers a predictive insight of migration in Jharkhand at the district level. The wisdom is highlighted by the fact that specific policies that are aimed at employment creation in the rural areas, agricultural transformation, livelihood diversification, and climate resilience are required. The research can be a solid backbone in terms of designing interventions to mitigate vulnerability, promote sustainable development, and eventually enhance the wellbeing of communities in the entire Jharkhand after identifying high-risk areas and predicting future mobility trends.

4.4. Model Performance

TRAINING SET PERFORMANCE:	
MAE (Mean Absolute Error)	0.1234
RMSE (Root Mean Squared Error)	0.1567
MAPE (Mean Absolute Percentage Error)	0.85%
R^2 (R-squared score)	0.9456 (94.56%)
Accuracy (within 10% error)	98.50%
Accuracy (within 5% error)	95.20%

Table 4.1

TESTING SET PERFORMANCE:	
MAE (Mean Absolute Error)	0.2345
RMSE (Root Mean Squared Error)	0.2891
MAPE (Mean Absolute Percentage Error)	1.52%
R^2 (R-squared score)	0.9281 (92.81%)
Accuracy (within 10% error)	95.00%
Accuracy (within 5% error)	87.50%

Table 4.2

The prediction analysis of the model suggests that Gradient Boosting performs very effectively for predicting migration patterns between districts. The performance of prediction is illustrated in TABLE 4.1 with a training set demonstrating high accuracy (\(MAE=0.1234\), \(RMSE=0.1567\)) and small errors between predicted responses and real results. The Mean Absolute Percentage Error (MAPE) of 0.85% also shows that, on average, the predictions of the model differ by less than 1%. Furthermore, the model has a R^2 score of 0.9456 (the entire variation in migration is included by the input features to modestly over 94.56% or about), 42-1998;2003Delucchi) are themselves in surplus of: High levels of accuracy—98.50% within ±10% and 95.20% within ±5%—demonstrate the fact that the.

consistency of the model.

The prediction analysis of the model suggests that Gradient Boosting performs very effectively for predicting migration patterns between districts. The performance of prediction is illustrated in TABLE 4.1 with a training set demonstrating high accuracy (\(MAE=0.1234\), \(RMSE=0.1567\)) and small errors between predicted responses and real results. The Mean Absolute Percentage Error (MAPE) of 0.85% also shows that, on average, the predictions of the model differ by less than 1%. Furthermore, the model has a R^2 score of 0.9456 (the entire variation in migration is included by the input features to modestly over 94.56% or about), 42-1998;2003Delucchi) are themselves in surplus of: High levels of accuracy 98.50% within ±10% and 95.20% within ±5% demonstrate the fact that the.

CHAPTER 5

5.1 Policy Recommendations

5.1. Introduction

The migration forecasts and district-level analyses presented in this study highlight the persistent socioeconomic and climatic vulnerabilities shaping population mobility across Jharkhand. With several districts experiencing migration rates exceeding 25%, and the state average projected to rise steadily over the next five years, it is evident that migration is not merely an economic choice but a critical survival strategy for millions of households. The interplay of unemployment, poverty, rainfall variability, and inadequate irrigation coverage continues to drive large segments of the rural population toward external labour markets.

These trends highlight the critical importance of evidence-based policies that target both the structural forces and proximate causes of migration. For high-migration districts, targeted interventions need to be directed at securing livelihoods, social protection and climate resilience interventions, and low migration districts provides a working model for development replication. Incorporating machine learning-driven predictability) provides a science-based foundation for developing future-oriented, local-region plans.

Henceforth, the following policy proposals are addressed to address root causes of migration contribute also to build rural resilience and support balanced regional development. Each suggestion is grounded in the results of this research and offers practical guidance to government authorities, development agencies and policy makers as they move towards less distress migration in Jharkhand.

5.2 .Policy Recommendations

1.Programs Strengthen Rural Employment Creation in High-Migration Districts.

A number of districts have recorded high migration rates averaging 25 percent and include Pakur, Godda, Dumka, Simdega and Jamtara where unemployment and non-farm employment opportunities are core reasons. Distinct increase in the rural job program such as MGNREGA Mahatma Gandhi National Rural Employment Guarantee Act(2005)., skills-training programs and local stimulation of micro-enterprises can really help to curb distress migration. These districts should be developed to have special employment zones and cluster-based industries to provide year-long sources of income.

2.Develop More Irrigation Infrastructures to limit migration due to climate.

The SWOT analysis shows that there is a close relationship between the variability of rain, ineffective irrigation coverage, and migration. The coverage of irrigation, which is currently less than 20 percent (in tribal districts) can be increased to 35-40 percent, which will significantly stabilize the agricultural production. More effort should be made on investments in minor irrigation systems, check dams, micro-irrigation, and solar-powered water pumps so as to improve water security in blocks prone to droughts.

3.Enhance Agricultural Productivity by Climate-resilient measures.

Existing agricultural risks that compel households to migrate may be reduced through the introduction of drought-resistant varieties of crops, promotion of crop diversification and strengthening extension services. This will be through the establishment of local procurement centers, better market linkages, and incentives to farmer-producer organizations (FPOs) which will boost incomes and reduce reliance on migration. Climate-sensitive agriculture should also be incorporated within the district level planning.

4. Improve Literacy, Education, and Skills Trainings in Low-Human-Development districts.

Low literacy and education are another factor that jeopardizes economic vulnerability in high-migration districts. The lack of education may be solved through the creation of residential schools in aboriginal communities, enhancing the school facilities, and improving the number of teachers. Also, vocation training facilities in line with the local economic potentials like agriculture, textile, mining and services are to be increased so as to provide the youth with market relevant skills.

5.Build Live Migration and Data Systems.

Welfare of Jharkhand has gaps in providing welfare owing to the underreported migration. The design of adaptive policies will be supported by the creation of a digital platform of migration monitoring tracking employees movements, remittance flows, and seasonal migration flows. To ensure that social protection programs are updated on time and their effectiveness is improved, there is a need to coordinate the efforts of the state labor departments and local governance bodies with civil society.

6.Launch Livelihood Diversification and Social Protection Measures.

Government assistance to rural companies in such sectors as dairy, poultry, fishing, handicrafts and ecotourism can be used to generate stable sources of local revenue. The economic crises which often cause migration will be reduced by strengthening the social protection schemes such as housing, food security, health insurance, and pensions. Risk-sharing programs and conditional cash transfers can also be applied to vulnerable households in order to enable them to survive the bad weather and economic situations.

7.Concentrate on Urban Development in Swelling Towns in order to diminish Push Factors.

Districts with low migration rates such as Ranchi, Bokaro, Ramgarh and East Singhbhum demonstrate how urbanization offers to them stable employment and education. It can be replicated to rural populations; such models of urban development can result in the lessening of stress levels and balanced regional development by creating planned small-town development, transportation connectivity, and industrial corridors.

CHAPTER 6

6.1 DISCUSSION

The study results present new information on the climatic and socioeconomic factors that influence the migration in Jharkhand. Combining Census 2011 baseline information with machine learning prediction models the study found that migration in Jharkhand is caused by a complex combination of economic hardships, uncertain agriculture and lack of livelihood diversification. Migration rates such as those in Pakur, Godda, Dumka, Simdega and Jamtara always record high levels of poverty, lack of irrigation opportunities and access to less literacy as well as more dependent on rain-based agriculture. Because of these weaknesses the families have to use migration as a coping mechanism that increases the effects of the rainfall shocks.

Scenario simulation and rainfall forecasting are done together to give a better understanding of the correlation between climate migration and rainfall forecasting. The results have shown that patterns of migration are influenced greatly by even the slightest changes in the rainfall patterns within a region where rainfall is inadequate, particularly in regions where irrigation is not satisfactory. This helps in advancing the notion that climate uncertainty remains a strong motive. Low-migration regions such as Ranchi, Bokaro, and Ramgarh, however, demonstrate how effective city infrastructure and non-farm jobs and diversification of the economy can lower out-migration.

The model performance clearly shows that Gradient Boosting is better than linear techniques as it captures the linear and nonlinear correlations. The predictions made in 20252029 will be accurate since the model has a high accuracy level (R 2 = 0.928). All in all, the combination of the temporal, economic, social, and climatic factors provides a complex view of migration dynamics enabling policymakers to have valuable information to use to target the interventions.

6.2 CONCLUSION

This study offers a comprehensive and data-based discussion of migration on the district level in Jharkhand. It demonstrates that migration is also a significant means of livelihood among most households. Between 2025 and 2029, the rate of migration has been on the rise. This increase is encouraged by the economic problems and increasing climate uncertainty. The predictive modeling method indicates high regional variations. In some of the tribal and economically disadvantaged districts, the rate of migration is over 25 percent compared to the low levels in more urban and industrialized ones.

The paper determines that lagged migration, unemployment, poverty, rainfall, and irrigation are some of the key variables in the trend. This makes it clear that climate and economic pressures tend to force the migration of people. Districts with insufficient irrigation and located more in rain-fed agricultural land are especially susceptible to climate. In addition, this study reveals that machine learning, namely Gradient Boosting is an efficient method of migration prediction.

The study with its identification of high-risk districts, ability to predict the future, and the necessity of specific policy actions is something that can provide an important insight to both government agencies, planners, and development workers. It supports the argument of making evidence-based decisions in Jharkhand by connecting socioeconomic influencing factors, climatic concerns, and forecasting.

6.3 FUTURE SCOPE OF THIS WORK

Despite the fact that the research will be insightful, it would be possible to enhance the areas in which it falls short. To begin with, more recent primary data on the patterns of migration such as seasonal migration, circular labor movement and dynamics of informal sector would be better and more current in their prediction. Second, including place-based pull factors like wages, remittance networks, and job demand in the urban labor markets would enhance the completeness of the model and broaden our knowledge about the spatial migration linkages.

Also in future research the climatic factor could be considered adding to it the changes in temperature, drought indices, soil moisture, and the satellite-derived vegetation indicators. These would assist in the precise identification of environmental push factors. Also, further research into more sophisticated models, such as spatiotemporal models or LSTM neural networks, can result in more accurate predictions as they are able to identify long-term dependencies and correlations between districts.

Lastly, it is possible to develop a GIS-based migration tracking system or interactive dashboard so that migration patterns can be tracked in real time and updated commonly. The models ought to be restrained annually using the new data to keep their accuracy so that the policymakers can get access to the most current forecasts. All in all, the future outlook is very promising to enhance predictive capacity, incorporate an array of elements, and contribute to the creation of informative policies to control migration within Jharkhand.

6.4 REFERENCES

- Bhatia, U. (2024). Drops & Data: Synergizing AI/ML with Process-Based Understanding in Hydrological Sciences. IIT Gandhinagar Lecture Series.
- Best, K., Gilligan, J., Baroud, H., Carrico, A., Donato, K., & Mallick, B. (2022). Applying Machine Learning to Social Datasets: A Study of Migration in Southwestern Bangladesh Using Random Forests. Regional Environmental Change, Springer.
- Millock, K. (2019). Climate Change, Migration and Irrigation. Fondation pour les Études et Recherches sur le Développement International (FERDI).
- Phesa, A., Woyessa, Y. E., & Xulu, S. (2024). Assessment of Evapotranspiration Variability and Its Relationship with Hydro-Meteorological Parameters Using Remote Sensing and Time Series Modelling in the Keiskammahoek Irrigation Scheme (KIS), South Africa. Hydrology, 11(2), 176. MDPI.
- Piguet, E., Pécoud, A., & de Guchteneire, P. (2011). Environmental Change, Adaptation and Migration: Bringing in the Region. Routledge.
- Robinson, C., & Dilkina, B. (2018). A Machine Learning Approach to Modeling Human Migration. Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS)
- Zhang, Y., Li, H., Wang, P., et al. (2025). AI-Driven Hydrological Modeling for Climate Resilience. arXiv preprint, Cornell University Library.