International Journal of Agriculture and Food Science 2025; 7(7): 133-140

ISSN Print: 2664-844X
ISSN Online: 2664-8458
NAAS Rating: 4.97
IJAFS 2025; 7(7): 133-140
www.agriculturaljournals.com
Received: 14-05-2025
Accepted: 16-06-2025

Soumyadeep Thakur
University Research Scholar, Bidhan Chandra Krishi Viswavidyalaya (Agriculture University), Mohanpur, Nadia, West Bengal, India

Debabrata Basu
Vice Chancellor, Uttar Banga Krishi Viswavidyalaya (Agriculture University), Pundibari, Cooch Behar, West Bengal, India

Corresponding Author:
Soumyadeep Thakur
University Research Scholar, Bidhan Chandra Krishi Viswavidyalaya (Agriculture University), Mohanpur, Nadia, West Bengal, India

Institutional and grassroots disconnect in jute marketing: A multidimensional analysis of information dissemination gap in west Bengal using feature engineering

Soumyadeep Thakur and Debabrata Basu

DOI: https://www.doi.org/10.33545/2664844X.2025.v7.i7b.507

Abstract
Jute, as a strategic fibre crop in the agrarian landscape of West Bengal, continues to face substantial impediments in marketing efficiency, primarily due to a persistent disjunction between the availability of market-related information and its actual assimilation at the grassroots level. The present study was conducted to systematically explore and quantify this information asymmetry by adopting a comparative framework between institutional supply and farmer-level accessibility of jute marketing information. A multi-stage stratified purposive sampling design was employed, encompassing 275 jute-growing farmers across 12 blocks in four jute-producing districts-Murshidabad, Nadia, Purba Burdwan, and Cooch Behar-and 30 officials representing agricultural marketing departments, cooperatives, and extension systems. A multidimensional interview schedule operationalizing 30 distinct categories of marketing information was administered, with established validity and reliability. All informational variables were normalized using z-score transformation, and normality tests (Shapiro-Wilk and Anderson-Darling) were conducted to guide appropriate inferential pathways. Advanced statistical methods-including Cohen's d, Cliff's Delta, bootstrap confidence intervals, empirical cumulative distribution divergence, overlap coefficient, Kullback-Leibler divergence, Wasserstein distance and lasso regression-were employed. Findings revealed a substantial perceptual and functional gap between the information availability and accessibility along with significant predictors of the both domains. The study recommends a targeted feature specific reorientation of extension strategies through digital inclusion, localized market intelligence, and decentralized policy dissemination, while integrating machine learning-based decision support tools to address context-specific informational deficits in jute marketing systems.

Keywords: Jute marketing information, information asymmetry, comparative analysis, statistical divergence metrics, machine learning

Introduction

Agricultural marketing plays a pivotal role in empowering farmers by bridging the gap between production and market realization, particularly in the case of traditional commercial crops such as jute. In India, and more specifically in West Bengal-the leading jute-producing state-the effective dissemination and utilization of marketing information is critical to ensuring remunerative returns for farmers (Chowdhury and Roy, 2017; Sulaiman and Holt, 2002) [1, 2]. Despite a robust institutional framework comprising government agencies, cooperatives, and marketing systems, discrepancies persist between the volume of structured jute marketing information disseminated and the actual level of information accessed and internalized by farmers (Sarkar *et al.* 2018) [3]. Prior studies have primarily focused on either the availability of agricultural information or the broader challenges in rural information systems but have seldom studied the perceptual and functional divide between the supply and demand sides of commodity-specific information ecosystems (Meera *et al.*, 2004) [4]. This oversight has led to a persistent research gap in understanding how institutional efforts in information dissemination translate into practical awareness, comprehension, and utility at the grassroots level (Aker, 2011) [5]. The problem becomes particularly acute in the context of jute, where volatile markets, weak price forecasting mechanisms, inadequate awareness of policy supports and procurement structures further exacerbate farmers' vulnerabilities (Rao, 2007; Chand, 2012) [6, 7].

Recognizing this gap, the present study adopted a comparative diagnostic approach to systematically investigate the divergence between the availability of jute marketing information-emanating from institutional sources-and the actual accessibility or uptake experienced by jute farmers across key producing districts of West Bengal.

Methodology

The present study was methodologically anchored in a comparative framework to investigate the disparity between the availability of jute marketing information from the supply side, primarily government agencies, and its actual accessibility to farmers in West Bengal (Glendenning *et al.*, 2010) [8]. For the purpose of constructing a representative and logically profound sampling framework, four jute-producing districts of West Bengal were identified based on their substantial contribution to jute cultivation: Murshidabad, Nadia, Purba Burdwan and Cooch Behar. Within these districts, key blocks known for their high jute production were purposively selected (Kumar, 2014) [9]. From Murshidabad, the blocks of Raghunathganj-I, Lalgola, and Bhagwangola-II were chosen; in Nadia, the blocks included Krishnanagar-I, Nabadwip, and Tehatta-II; from Cooch Behar, the blocks of Tufanganj-I, Dinhata-II, and Cooch Behar-I were selected; and in Purba Burdwan, Kalna-I, Purbasthali-I, and Katwa-II were selected. A multi-stage stratified purposive sampling design was employed, wherein three blocks from each of the four districts were selected based on jute production intensity, followed by the selection of two to three villages from each block (Babbie, 2020) [10]. From each village, seven to eight farmers actively engaged in jute cultivation and marketing were selected using purposive judgment based on criteria such as landholding size, exposure to market systems, and interaction with institutional mechanisms, culminating in a total sample of two hundred and seventy-five farmers (approximately sixty-nine farmers per district). Additionally, thirty government personnel across these districts were selected, including officials from agricultural marketing departments, Krishi Vigyan Kendras (KVKs), Assistant Director of Agriculture and cooperative societies involved in jute procurement, extension, and policy implementation, thereby ensuring a balanced representation from both the supply and demand sides of the jute information ecosystem (Chowdhury and Roy, 2017) [1].

To effectively capture the multifaceted nature of jute marketing, a comprehensive and multidimensional interview schedule was meticulously constructed. This instrument aimed to operationalize key constructs across thirty distinct information categories, each representing a critical facet of marketing intelligence (Meera *et al.*, 2004) [4]. These categories spanned a wide spectrum, including the accessibility of daily and weekly price data from official sources, the dynamics and causal factors behind price fluctuations, and the dissemination of Minimum Support Price (MSP) and procurement centre details. Furthermore, the schedule delved into the provision of real-time market insights at local, national, and international levels, alongside future price forecasts and demand-supply trends (Sarkar *et al.*, 2018) [3]. It also incorporated modules on buyer and exporter databases, marketing channel specifics, wholesale and retail price spreads, export procedures, governmental assistance, trade agreements, subsidies, storage schemes, and value-added product policies. The instrument extended to encompass digital market information, online trading platforms, training initiatives, intelligence reports, information-sharing systems, branding strategies, and packaging techniques (Aker, 2011) [5]. Reliability testing yielded a Cronbach's alpha coefficient of 0.81, demonstrating strong internal consistency (Tavakol and Dennick, 2011) [11]. The tool's validity was established through expert judgment, utilizing a panel of forty subject matter specialists from the field of agriculture (Haynes *et al.*, 1995) [12].

Given the considerable disparity in sample sizes, z-score normalization was utilized to standardize all informational variables (Field, 2018) [13]. This transformation facilitated the adjustment of variable distributions onto a common scale with a mean of zero and unit variance, preserving the internal structure of the data while enabling rigorous and unbiased comparative analyses. To ascertain the statistical assumptions underpinning the comparative framework, normality diagnostics were carried out using both the Shapiro-Wilk and Anderson-Darling tests for sensitivity to distributional symmetry and skewness (Razali and Wah, 2011) [14]. To enrich the interpretive robustness, a series of advanced statistical techniques was executed using python libraries. Effect size estimations were meticulously conducted utilizing Cohen's d and Cliff's Delta, serving to quantify both the magnitude and directional predominance of distributional disparities, independent of the variations in sample size (Lakens, 2013) [15]. To ascertain confidence intervals surrounding mean differences, bootstrap resampling techniques using feature engineering were deftly employed (Efron and Tibshirani, 1993) [16]. The exploration of divergence between empirical cumulative distributions was facilitated through the application of Kolmogorov-Smirnov distance and the Overlap Coefficient (OVL), while Kullback-Leibler divergence provided insights into asymmetric informational loss (Lv *et al.*, 2024) [17]. Furthermore, Wasserstein Distance was utilized to quantify the geometric separation between distributions, and the Gardner-Altman estimation method was employed to visually represent effect sizes along with their corresponding confidence intervals (Ho *et al.*, 2019) [18]. Conclusively, Lasso regression was implemented for feature selection purposes, effectively identifying the most salient predictors and integrating machine learning methodologies into the overarching analytical framework (Tibshirani, 1996) [19].

Results and Discussion

The results of this investigation are presented herein, followed by a comprehensive discussion of their significance.

Table 1: Normality Test Results for Information Availability and Information Accessibility

Serial No.	Variable	Shapiro-Wilk Test Value	p-value (Shapiro-Wilk)	Anderson-Darling Test Value	p-value (Anderson-Darling)	Normality Conclusion
1	Information Availability (n = 30)	W = 0.9114	0.0161	0.8903	0.712	Non-Normal
2	Information Accessibility (n = 275)	W = 0.9948	0.4821	0.5139	0.776	Normal

Table 1 summarized the normality assessment of two key constructs using the Shapiro-Wilk and Anderson-Darling tests. For information availability, the Shapiro-Wilk statistic (W = 0.9114, p = 0.0161) revealed a significant deviation from normality at the 5 percent level, while the Anderson-Darling test (value = 0.8903, p = 0.712) indicated a comparatively better distributional fit (Razali and Wah, 2011) [14]; however, the Shapiro-Wilk test was prioritized for its superior sensitivity with smaller samples, pointing to skewed information dissemination patterns. In contrast, information accessibility demonstrated normal distribution characteristics with Shapiro-Wilk (W = 0.9948, p = 0.4821) and Anderson-Darling (value = 0.5139, p = 0.776).

Table 2: Mann-Whitney U Test Results Comparing Information Accessibility and Information Availability

Particular	Sample Sizes	Mann-Whitney U Value	p-value
Mann-Whitney U Test	Information Accessibility (n)= 275, Information Availability (n) = 30	8250	0.0001

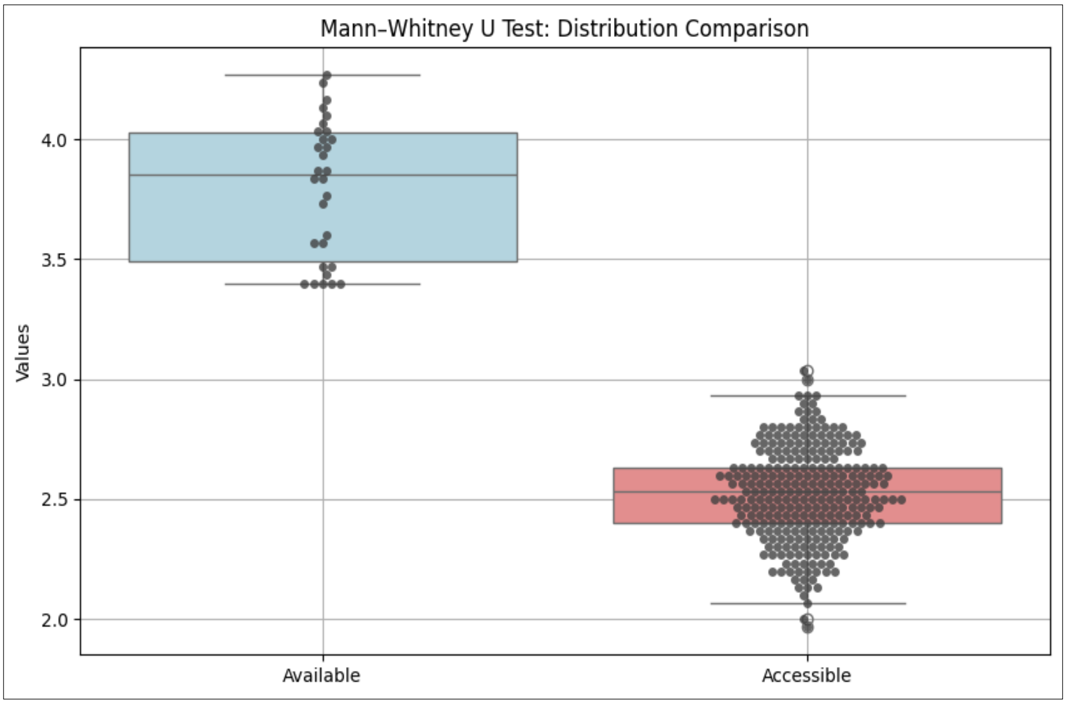

Fig 1: Mann-Whitney U Test for Comparative Analysis

The Mann-Whitney U test in table 2, employed to assess differences between the two independent constructs without assuming normality, yielded a U statistic of 8250.0000 with a p-value of less than 0.0001 at 5 percent level of significance. This result revealed a highly significant difference between the standardized distributions of information availability and accessibility (figure-1), confirming that the central tendencies of the two construct domains diverged meaningfully.

Table 3: Effect Size Estimates Comparing Information Accessibility and Information Availability

Statistical Measure	Value	Sample Sizes	Threshold Interpretation
Cohen's d	0.8473	Accessibility (n) = 275	Large effect size (≥ 0.8)
Cliff's Delta	0.947	Availability (n) = 30	Very strong ordinal dominance (≈ 1.0)

To determine the magnitude of disparity between the domains of information availability and accessibility in the context of jute marketing in West Bengal, effect size estimations were presented using Cohen's d and Cliff's Delta in table 3. Cohen's d value, computed at 0.8473, exceeded the conventional threshold of 0.8, indicating a large and practically meaningful separation between the standardized means of the two groups (Lakens, 2013) [15]. This suggested a marked divergence in the structural diffusion of jute marketing information versus its actual grassroots-level assimilation. Complementing this, Cliff's Delta value reached the maximum of 0.947, reflecting complete ordinal dominance of one group over the other. This highlighted an exceptional and unequivocal distributional dissimilarity, reinforcing the existence of a critical information asymmetry in the jute marketing information ecosystem.

Bootstrapping and Confidence Interval Estimation

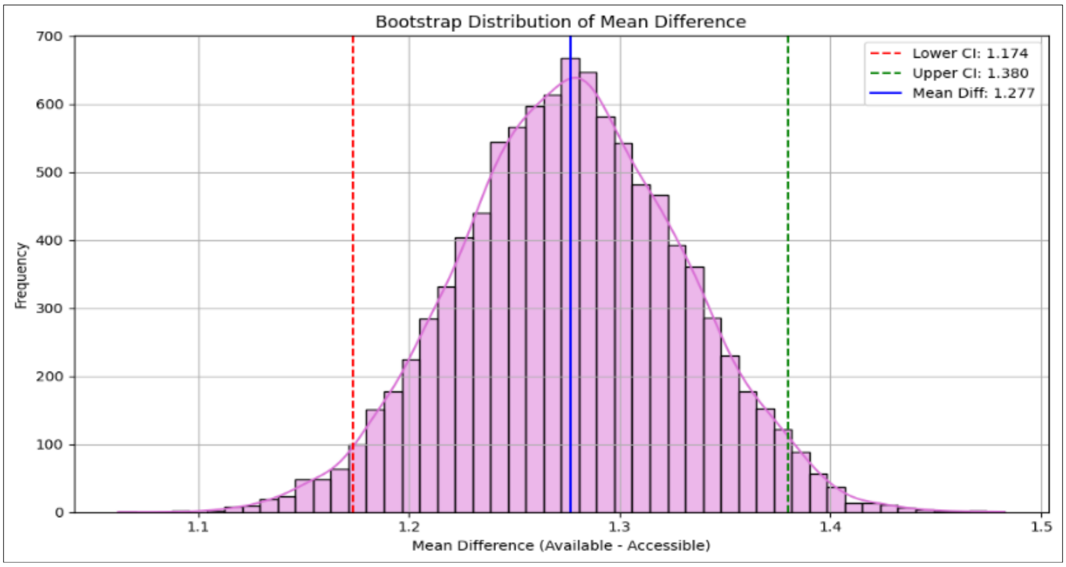

Fig 2: Bootstrapping and Confidence Interval Estimation

To mitigate the influence of sampling variability, a bootstrap resampling procedure was implemented (figure-2), to approximate the sampling distribution of the mean difference. The resulting 95 percent confidence interval for the mean difference between the standardized information available and accessible scores, (1.1735, 1.3802), thus reinforcing the statistical significance of the observed disparity. Consequently, the bootstrap analysis corroborated the prior inferential findings, providing supplementary statistical validation (Efron and Tibshirani, 1993) [16]. The procedure confirmed the persistent and substantial mean difference between perceived information availability and accessibility, even after accounting for resampling fluctuations, thereby bolstering the credence and consistency of the identified information asymmetry.

Empirical Cumulative Distributional Divergence Analysis

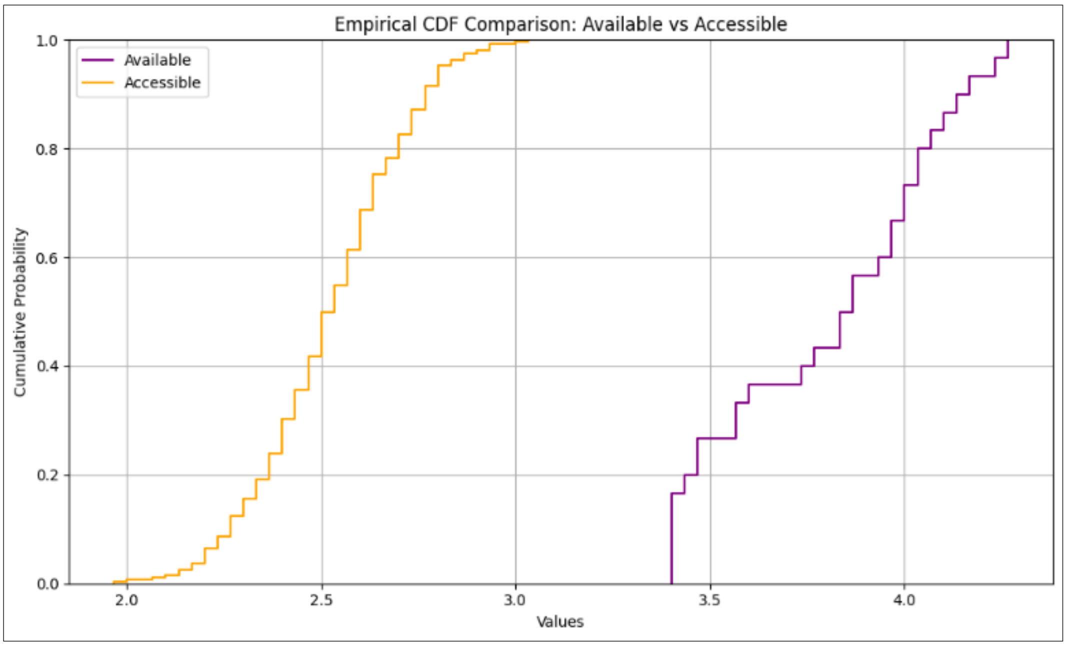

Fig 3: Empirical Cumulative Distributional Divergence Analysis

To delve deeper into the distributional differences between information availability and accessibility domain, the Kolmogorov-Smirnov (KS) test was conducted (figure-3), comparing the cumulative distribution functions of the standardized variables. The test yielded a KS Statistic of 0.968, with a p-value of 0.0000, revealing a highly significant divergence between the two distributions. The KS statistic indicated a complete absence of overlap between the empirical cumulative distributions, signifying a fundamental distinction beyond mere differences in central tendency and variance (Albano *et al*, 1995) [20]. This result corroborated the findings from the Mann-Whitney U Test, effect size analysis, and bootstrap confidence interval estimation, emphasizing the structural and experiential gap in information availability accessibility.

Domain Overlap Analysis

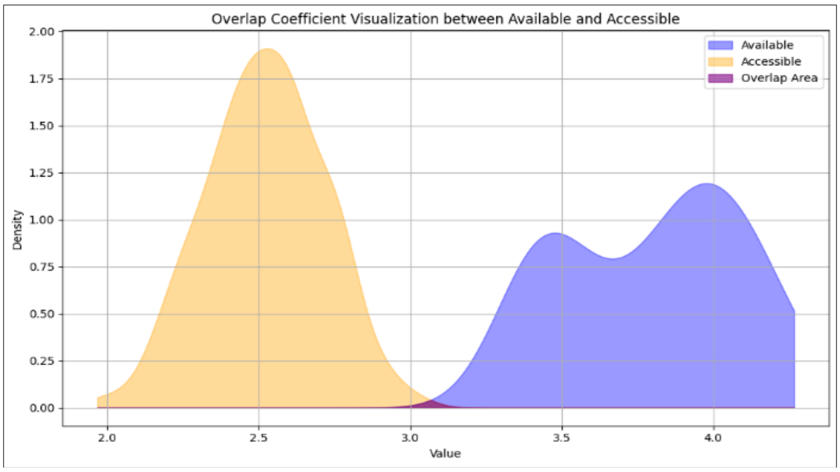

Fig 4: Domain Overlap Analysis

To quantify the dissimilarity between the information available and accessible probability distributions, the domain overlap analysis (figure-4) was computed. The analysis, measuring the overlapping area of the two distributions, registered at 0.0041, indicating a negligible overlap. This result corroborated the findings from the Kolmogorov-Smirnov Test and Cliff's Delta, reinforcing the stakeholders' perception of fundamental distinctiveness between the two constructs (Kim *et al*, 2016) [21]. The minimal overlap suggested a stark perceptual disconnect between information presence and accessibility, implying that high availability ratings did not correspond with similar accessibility ratings.

Information Disparity Quantification using Kullback-Leibler Divergence and Wasserstein Distance Analysis

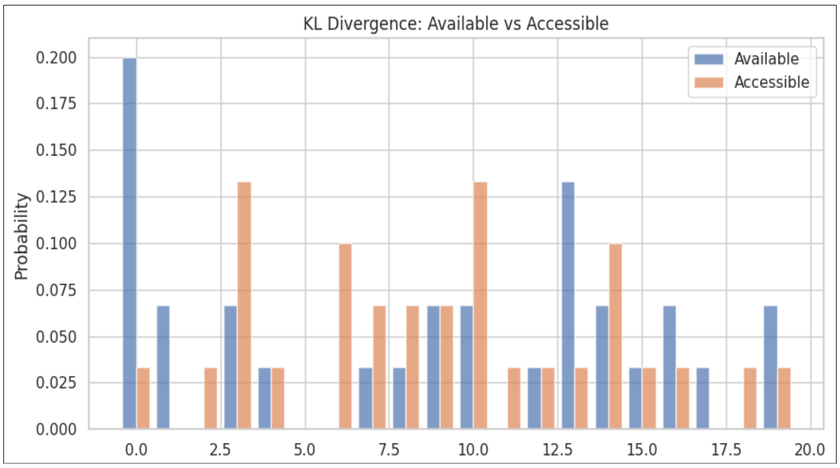

Fig 5: Kullback-Leibler Divergence Analysis

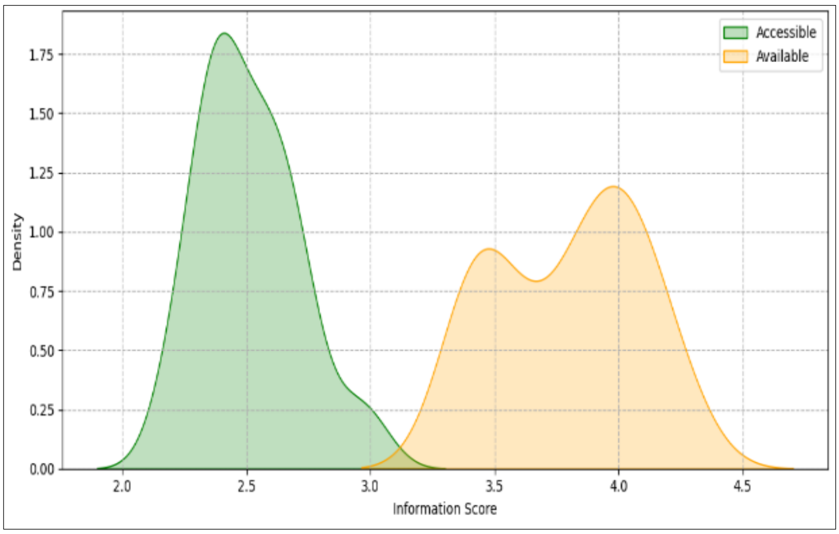

Fig 6: Wasserstein Distance Analysis

The findings from the Kullback-Leibler (KL) divergence and Wasserstein Distance analyses provided critical insights into the extent of informational disparity between institutional availability and farmer-level accessibility of jute marketing information. The KL divergence (figure-5), computed at 2.03, indicated a substantial asymmetry in the probability distributions of the two domains, reflecting a significant informational loss when the accessibility pattern was used to approximate the availability structure. This highlighted the perceptual and functional disconnect in how market-related data, though structurally disseminated by institutional sources, failed to align with the cognitive and practical absorption by the farming community. Complementarily, the Wasserstein Distance value (figure-6) of 0.2241 underscored a moderate yet notable geometric separation between the distributions, further validating the existence of a tangible gap in informational reach (Lv *et al*., 2024) [17]. These results collectively emphasized that while the formal information ecosystem possessed a well-articulated structure, its grassroots penetration remained uneven, thereby necessitating strategic interventions to realign the flow of critical jute marketing knowledge to end users.

Gardner- Altman Test

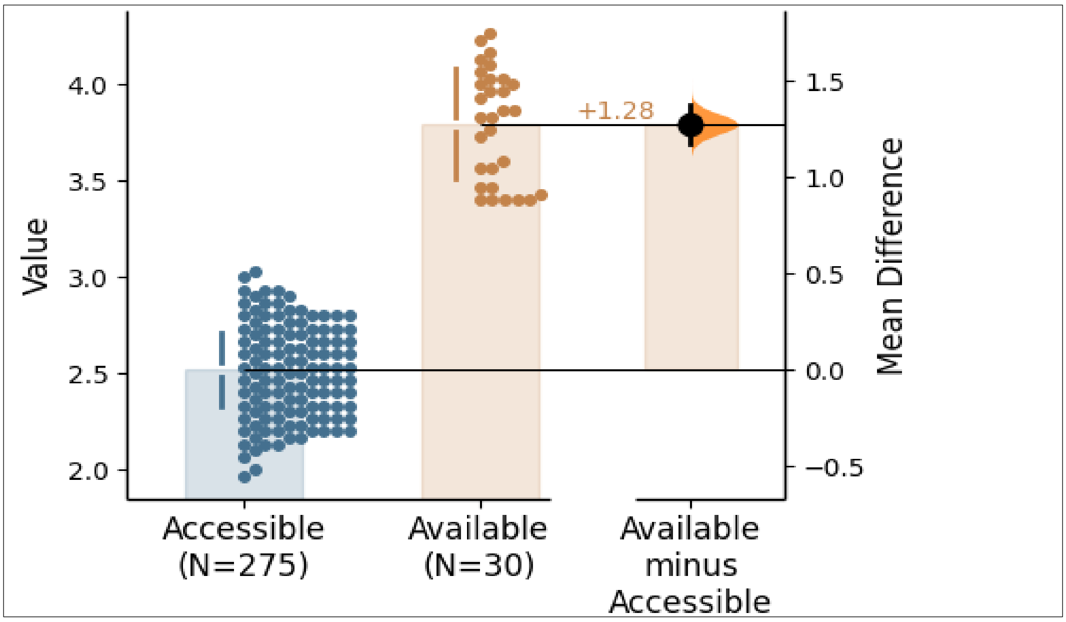

Fig 7: Bee Swarm Plot of Gardner- Altman Test

The analysis (figure-7) revealed a statistically significant disparity between the accessible and available information constructs, as evidenced by the Gardner-Altman test. It was observed that the available group exhibited a markedly elevated mean value, approximately 1.28 units higher, than the accessible group, a divergence that suggested a substantial effect of the availability variance. This test underscored a robust difference between the two groups, indicating that the availability context significantly influenced the measured variable (Reed *et al*., 2023) [22]. The dispersion of data points within each group, visualized via the bee swarm plot, further illustrated the distributional differences, with the available group demonstrating both a wider spread and a higher central tendency.

Table 4: Feature Importance Extraction of Information Accessibility with Lasso Regression (n=275)

Serial No.	Particular	Value	Rank
1	Availability of Daily/Weekly Prices, Official Sources	0.0326	1
2	Price Fluctuations Availability, Reasons for Fluctuations	0.0297	2
3	Wholesale/Retail Price Spreads	0.0281	3
4	Exporting Countries Information, Market Demands	0.0272	4
5	Digital Platform Training	0.0262	5
6	Branding, Certification, Labelling Awareness	0.0260	6
7	Financial Assistance for Processing/ Export	0.0256	7
8	Jute Marketing/Storage Schemes Awareness	0.0251	8
9	Export Procedures, Certification, Documentation Awareness	0.0251	9
10	Jute Marketing Subsidies/ Incentives	0.0249	10
11	Market Intelligence Reports	0.0248	11
12	Eco-Friendly Product Market Demand	0.0239	12
13	MSP Information, Procurement Center Information	0.0235	13
14	Real-time Local Market Info, Real-time National Market Info	0.0211	14
15	International Buyer/Exporter Database	0.0103	15

Table 5: Feature Importance Extraction of Information Availability with Lasso Regression (n=30)

Serial No.	Particular	Value	Rank
1	Consumer Trends, Branding Strategies	0.0889	1
2	Price Fluctuations Availability and Reasons for Fluctuations	0.0882	2
3	Exporting Information and Market Demands	0.0708	3
4	Export Procedures, Certification, Documentation Awareness	0.0612	4
5	Emerging Jute Market Reports	0.0474	5
6	Market Intelligence Reports	0.0415	6
7	International Buyer/Exporter Database	0.0368	7
8	Buyer/Exporter Database	0.0353	8
9	Import Tariffs, Trade Agreements, Trade Restrictions	0.0321	9
10	Digital Platform Training	0.0291	10
11	Value-Added Jute Product Policies Awareness	0.0285	11
12	Middlemen Details, Commission Charges, Marketing Margins	0.0213	12
13	Online Jute Trade Platform	0.0204	13
14	Government Export Assistance	0.0045	14
15	Farmer/Trader Information Sharing System	0.0029	15

Lasso regression-an advanced machine learning modelling was employed to identify the most influential predictors (Tibshirani, 1996) [19] within two critical domains: information accessibility and information availability presented in table 4 and 5. In the information accessibility domain, top five key features included the availability of daily or weekly price information from official sources, access to data on price fluctuations along with the reasons behind such changes, knowledge of wholesale and retail price spreads, availability of information related to exports and market demand, and exposure to training on digital platforms. These findings suggested that stakeholders, particularly farmers, faced considerable challenges in obtaining timely and actionable market data, especially in areas requiring digital engagement and export-oriented decision-making. Similarly, in the information availability domain, the analysis revealed top five primary predictors that contributed significantly to the observed information gaps. These were awareness of consumer trends and branding strategies, availability of information on price fluctuations and their causes, access to data on exporting countries and related market demands, knowledge of export procedures including certification and documentation, and access to emerging market reports for jute. The recurrence of price fluctuation information in both domains highlighted its critical role in informed decision-making.

Conclusion

The study illuminated a stark asymmetry between institutional-level availability and grassroots-level accessibility of jute marketing information in West Bengal, exposing systemic inefficiencies in information flow, assimilation, and responsiveness. Despite the structured dissemination of market intelligence by formal institutions, its fragmented absorption at the farmer level impeded both market efficiency and the adaptive capacity of smallholders in an increasingly volatile agri-market landscape. In alignment with the national imperatives of *doubling farmers' income* and the *vocal for local* movement, the findings called for a paradigm shift toward hyper-localized, digitally enabled, and context-specific transfer of information architectures that elevate farmer agency and embed inclusivity. The integration of machine learning-driven decision support tools, fortified village-level information nodes, and multilingual, real-time advisory platforms emerged as critical enablers for democratizing market access and fostering participatory, data-driven governance. Such transformative interventions, grounded in localized empowerment and technological precision, are essential to future-proof the jute marketing ecosystem and actualize the vision of a self-reliant, resilient, and income-secure agrarian economy.

Acknowledgement

The authors gratefully acknowledge the instrumental and academic support provided by the Research Methodology Laboratory, Bidhan Chandra Krishi Viswavidyalaya, throughout the course of this research. No conflicts of interest are associated with this work.

References

1. Chowdhury A, Roy A. Role of agricultural marketing in ensuring food security and rural development. Indian J Agric Econ. 2017;72(3):289-299.
2. Sulaiman RV, Holt G. Extension, poverty and vulnerability in India: Country study for the Neuchâtel Initiative. Working Paper. Neuchâtel Group; c2002.
3. Sarkar S, Maity A, Mondal T, Jose S, Sangma RC, Khan AM, *et al*. Participatory Agro-ecological and Socio-economic Analysis of Jute Cultivation in West Bengal. J Community Mobil Sustain Dev. 2018;13(1):43-51.
4. Meera SN, Jhamtani A, Rao DUM. Information and communication technology in agricultural development: A comparative analysis of three projects from India. Agric Res Ext Netw Pap No. 135. London: ODI; 2004.
5. Aker JC. Dial "A" for agriculture: A review of information and communication technologies for agricultural extension in developing countries. Agric Econ. 2011;42(6):631-47.
6. Rao CHH. Policy distortions in agriculture: Implications for growth, sustainability and welfare. Indian Econ Rev. 2007;42(1):1-21.
7. Chand R. Development policies and agricultural markets. Econ Polit Wkly. 2012;47(52):53-63.
8. Glendenning CJ, Babu SC, Asenso-Okyere K. Review of agricultural extension in India: Are farmers' information needs being met? IFPRI Discussion Paper 01048. Washington (DC): International Food Policy Research Institute; 2010.
9. Kumar P. Agricultural extension in India: Status, challenges and policy support. Agric Econ Res Rev. 2014;27(2):159-68.

10. Babbie ER. The practice of social research. 15th ed. Boston: Cengage Learning; c2020.
11. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011; 2:53-5.
12. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. Psychol Assess. 1995;7(3):238-47.
13. Field A. Discovering statistics using IBM SPSS statistics. 5th ed. London: Sage; c2018.
14. Razali NM, Wah YB. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Anal. 2011;2(1):21-33.
15. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. Front Psychol. 2013; 4:863.
16. Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton: Chapman & Hall/CRC; c1993.
17. Lv J, Yang H, Li P. Wasserstein distance rivals kullback-leibler divergence for knowledge distillation. Adv Neural Inf Process Syst [Internet]. 2024; 37:65445-65475.
18. Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A. Moving beyond p-values: Data analysis with estimation graphics. Nat Methods. 2019;16(7):565-6.
19. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc B. 1996;58(1):267-88.
20. Albano AM, Rapp PE, Passamante A. Kolmogorov-Smirnov test distinguishes attractors with similar dimensions. Phys Rev E. 1995;52(1):196.
21. Kim MC, Zhu Y, Chen C. How are they different? A quantitative domain comparison of information visualization and data visualization (2000-2014). Scientometrics. 2016; 107:123-165.
22. Reed D, Kazemi A, Gulati V, Fan S, Sit AJ, Toris CB, Moroi SE. Applying Estimation and Robust Methods to Compare the Intraocular Pressure Measured by Tonometers in the Eye Dynamics and Engineering Network (EDEN) Aqueous Humor Dynamics (AHD) Clinical Trial. Invest Ophthalmol Vis Sci. 2023;64(8):1377.