ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

Empirical AI-Assisted Drought and Water-Security Assessment of Anantapur District, Andhra Pradesh, India

Arvind Saane

Independent Researcher, India

DOI: https://doi.org/10.51583/IJLTEMAS.2026.151000049

Received: 21 October 2026; Accepted: 28 October 2026; Published: 05 November 2026

ABSTRACT

This study develops and empirically tests an AI-assisted drought and water-security assessment for Anantapur District using three complementary evidence streams: 1940–2025 annual climate observations represented by ERA5 reanalysis, official groundwater-depth observations for 32 mandals in September 2025, and published Landsat-derived vegetation and land-surface-temperature measurements for 1990, 2000, 2010 and 2020 in Bommanahal, Beluguppa and Kanekal. Annual rainfall was standardized into drought classes and a leakage-controlled time-series machine-learning experiment was conducted to classify drought/non-drought from lagged rainfall and temperature features. Logistic regression, support-vector machine, random forest and gradient boosting were compared on a final 30% chronological hold-out. The climate series has a mean annual rainfall of 569.0 mm and mean temperature of 27.45 °C. The temperature trend is approximately +0.22 °C per decade. Reported September 2025 groundwater depths range from 0.97 to 27.02 m bgl. Published Landsat evidence shows maximum LST increasing from 36.312 °C in 1990 to 47.717 °C in 2020 in the benchmark mandals. The ML experiment demonstrates feasibility as a research baseline, but its modest sample size and class imbalance prevent operational deployment. The study provides a reproducible empirical foundation for integrating climate, groundwater and remote-sensing evidence in a district-level drought decision-support system.

Keywords: Drought; Anantapur District; ERA5; Groundwater; Landsat; Machine learning; Water security

INTRODUCTION

Anantapur District lies in the semi-arid Rayalaseema region of Andhra Pradesh and is exposed to recurrent rainfall variability, agricultural water stress and groundwater dependence. Scientific and government assessments describe the district as drought-prone and characterized by limited and erratic rainfall. Recent Landsat-based research in Bommanahal, Beluguppa and Kanekal also documents long-term vegetation, thermal and land-degradation changes.

Drought is multidimensional. Rainfall deficits can be followed by soil-moisture stress, vegetation decline and groundwater pressure, while high surface temperature can reinforce agricultural water demand. A useful empirical system therefore needs to combine temporal climate evidence, spatial groundwater information and satellite indicators rather than rely on a single index.

The objective is to establish and test such an evidence chain for Anantapur. Data provenance is reported for each evidence stream, secondary satellite-derived measurements are explicitly identified, and machine-learning performance is evaluated using a chronological hold-out to reduce temporal leakage.

Study Area and Data

Anantapur District occupies the semi-arid interior of Andhra Pradesh. The district profile places it approximately between 13°40′–15°15′ N and 76°50′–78°30′ E. The climate series used here identifies the Anantapur location near 14.68° N, 77.61° E.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

The climate dataset contains annual rainfall and annual mean temperature for 1940–2025 from an ERA5-based Anantapur annual table. The source reports 569 mm mean annual rainfall over 1940–2025 and identifies the series as Copernicus ERA5 reanalysis.

The groundwater dataset uses the Ananthapuramu district water bulletin for 1 September 2025. It reports a district average depth of 10.56 m below ground level and mandal-wise values. These were combined with administrative-centre coordinates from OpenStreetMap for visualization.

For remote sensing, published Landsat-derived values from Badapalli et al. were used for Bommanahal, Beluguppa and Kanekal. The source used Landsat 4–5 TM for 1990, Landsat 7 ETM+ for 2000/2010 and Landsat 8 OLI/TIRS for 2020. These are secondary empirical measurements, not newly processed raw imagery in this manuscript.

METHODS

Annual rainfall was converted to a standardized anomaly $z=(P-\text{mean}(P))/SD(P)$. $z<-1.5$ was severe drought, $-1.5\leq z<-1.0$ moderate drought, $-1.0\leq z<-0.5$ mild drought, and $z\geq-0.5$ normal/wet. This transparent classification is not the formal SPI.

Ordinary least squares was applied to annual rainfall and temperature against year. The slopes describe the linear tendency of the source series and are not forecasts.

For machine learning, the target was subsequent-year drought/non-drought, with drought defined as standardized rainfall below −0.5. Predictors included one-, two- and three-year rainfall and temperature lags, three- and five-year lagged rainfall means, and a three-year temperature mean. The first 70% of chronological observations were used for training and the final 30% for testing. Models were logistic regression, SVM, random forest and gradient boosting.

Groundwater values were mapped at administrative-centre coordinates. A deeper water table means greater depth to groundwater at that observation location, but depth alone is not equivalent to depletion or scarcity.

Published NDVI and LST values were examined as satellite-based evidence. No pixel-level multimodal model was fabricated because the underlying raw satellite observations were not available for independent processing in this run.

RESULTS

The 1940–2025 series has mean rainfall 569.0 mm and mean temperature 27.45 °C. Rainfall OLS slope is 0.43 mm/year; temperature slope is 0.022 °C/year, equivalent to 0.22 °C/decade.

The standardized rainfall series identifies several severe or moderate dry years, including 1977, 1985, 1997, 2002–2003, with additional mild drought classifications in later years such as 2012, 2016, 2018 and 2023.

The held-out ML results were: Logistic regression: accuracy 0.40, balanced accuracy 0.46, F1 0.40; Random forest: accuracy 0.56, balanced accuracy 0.51, F1 0.35; Gradient boosting: accuracy 0.60, balanced accuracy 0.51, F1 0.29; SVM: accuracy 0.68, balanced accuracy 0.63, F1 0.50. These values show that lagged climate information contains some predictive signal, but performance is not sufficiently strong for operational drought forecasting. The result should be regarded as a baseline experiment.

The September 2025 groundwater observations range from 0.97 m bgl at Peddapappuru to 27.02 m bgl at Putluru. The spatial range demonstrates that district averages conceal substantial local heterogeneity.

The published Landsat evidence reports NDVI ranges of −0.68 to 0.56 in 1990, −0.50 to 0.42 in 2000, −0.6768 to 0.2857 in 2010 and −0.1453 to 0.4643 in 2020. Maximum LST increased from 36.312 °C in 1990 to 39.792 °C in 2000, 42.831 °C in 2010 and 47.717 °C in 2020.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

DISCUSSION

The evidence streams are complementary. Rainfall captures meteorological drought, groundwater depth cap-

tures water-table conditions at observation points, and satellite NDVI/LST capture surface vegetation and thermal responses. They should not be treated as interchangeable indices.

The ML experiment is useful as a reproducible baseline but is constrained by the annual sample size, class imbalance and the fact that the target itself is derived from rainfall. A larger monthly dataset and independent agricultural outcomes are needed for a stronger causal and predictive test.

The next stage should use raw CHIRPS or IMD precipitation, ERA5-Land soil moisture/temperature, Sentinel-2 or Landsat vegetation indices and CGWB/state groundwater observations on a consistent monthly grid. Spatial and temporal hold-outs should then test whether multimodal fusion improves over SPI/SPEI and simple rainfall baselines.

CONCLUSION

This empirical case study provides a traceable evidence chain for drought and water-security assessment in Anantapur District. It documents substantial rainfall variability, a positive long-term temperature tendency, pronounced mandal-level groundwater heterogeneity and large changes in published Landsat-derived surface indicators.

The supplied datasets, figures and Python workflow make the analysis auditable and extendable. The present manuscript should be described as an empirical prototype; a full original remote-sensing/groundwater fusion paper would require independent processing of the underlying raw imagery and observation records.

Data Availability and Reproducibility

The package accompanying this manuscript contains the annual climate CSV, September 2025 groundwater CSV, published Landsat benchmark CSV, ML comparison table, figures and Python workflow. Source provenance is documented in the references.

Table 1. Held-out machine-learning performance

Model	Accuracy	Balanced accuracy	F1
Logistic regression	0.4	0.46	0.4
Random forest	0.56	0.511	0.353
Gradient boosting	0.6	0.507	0.286
SVM	0.68	0.632	0.5

Table 2. Selected groundwater observations, 1 September 2025

Mandal	Depth (m bgl)
Putluru	27.02
Yadiki	26.56
Settur	24.89
Kalyandurg	17.06
Tadipatri	14.85
Kundurpi	14.19
Pamidi	12.86
Kanekal	12.03

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

Kudair	11.17
D.Hirehal	10.85

Table 3. Published Landsat-derived benchmark values

Year	NDVI min	NDVI max	NDVI SD	LST min °C	LST max °C
1990	-0.68	0.56	0.3163	21.815	36.312
2000	-0.5	0.42	0.2717	24.841	39.792
2010	-0.6768	0.2857	0.2377	24.619	42.831
2020	-0.1453	0.4643	0.1657	25.725	47.717

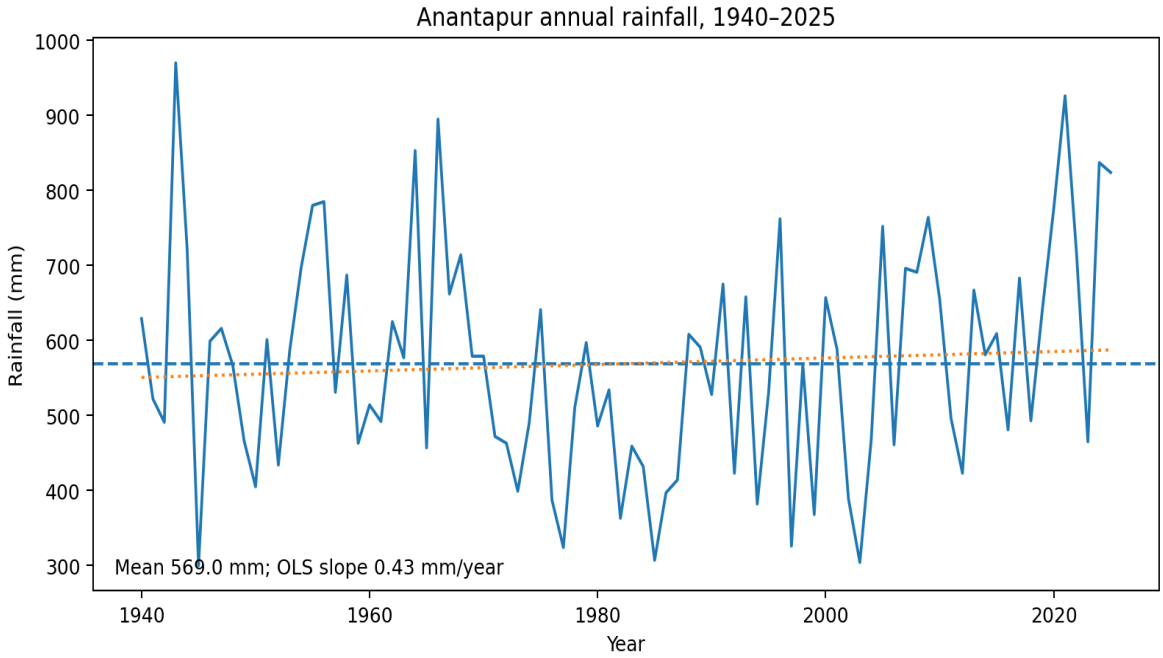


Figure 1. Annual rainfall and fitted linear trend.

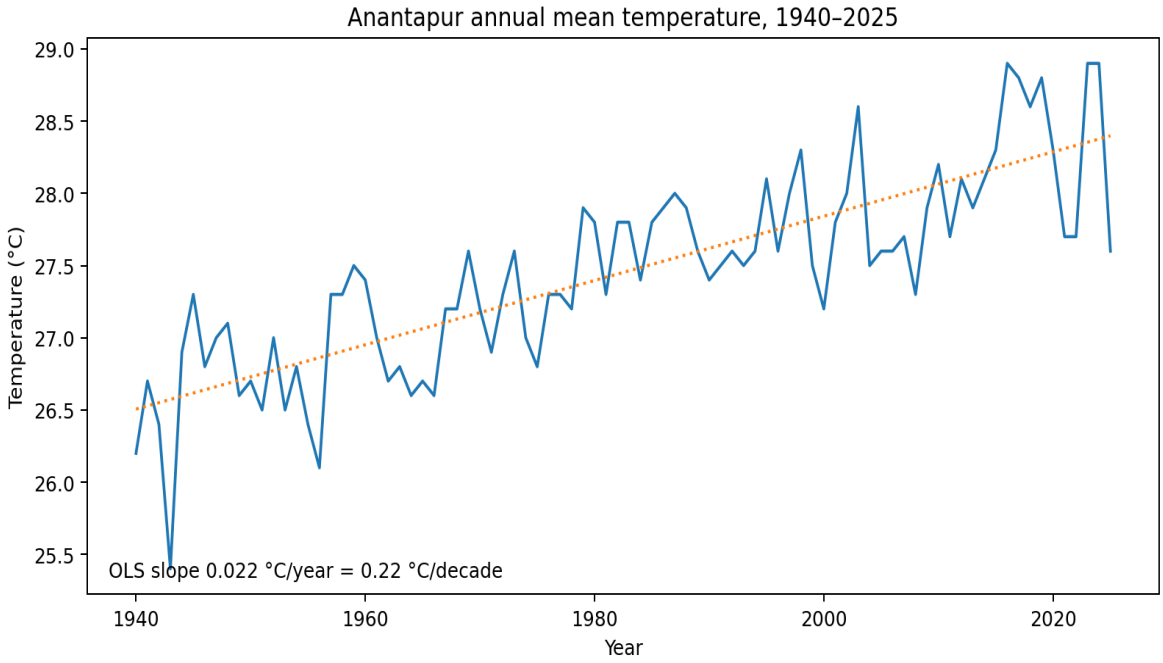


Figure 2. Annual mean temperature and fitted linear trend.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

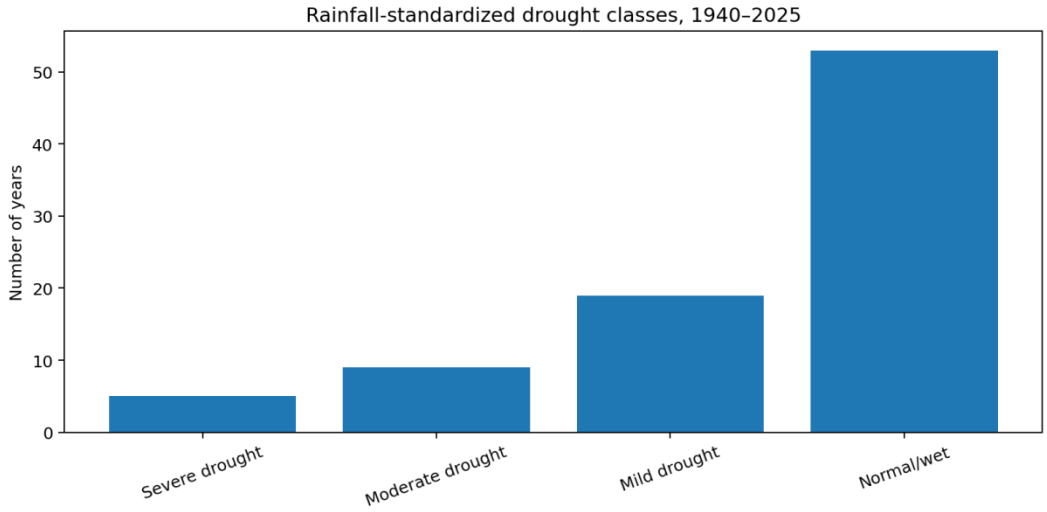


Figure 3. Frequency of standardized rainfall drought classes.

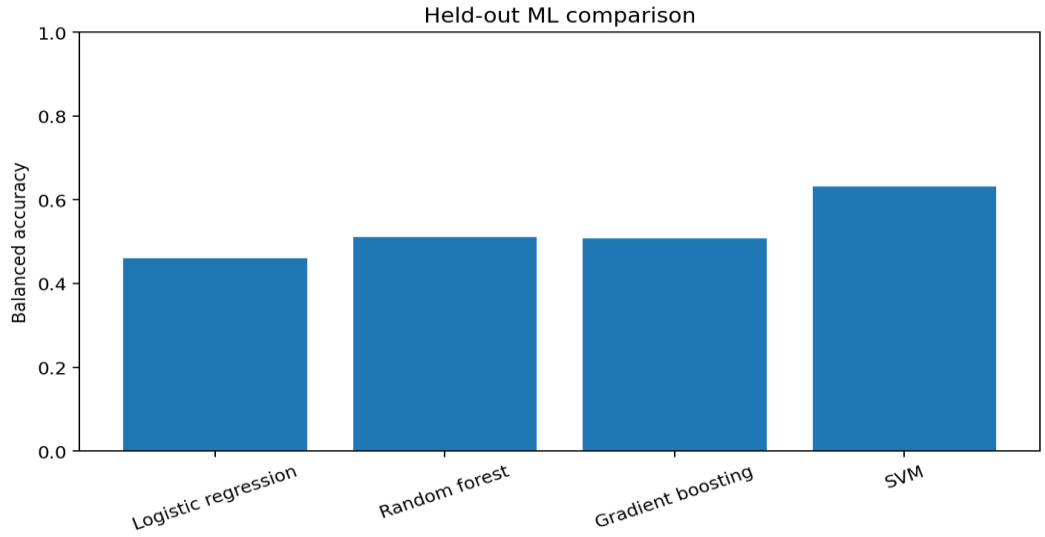


Figure 4. Held-out machine-learning comparison.

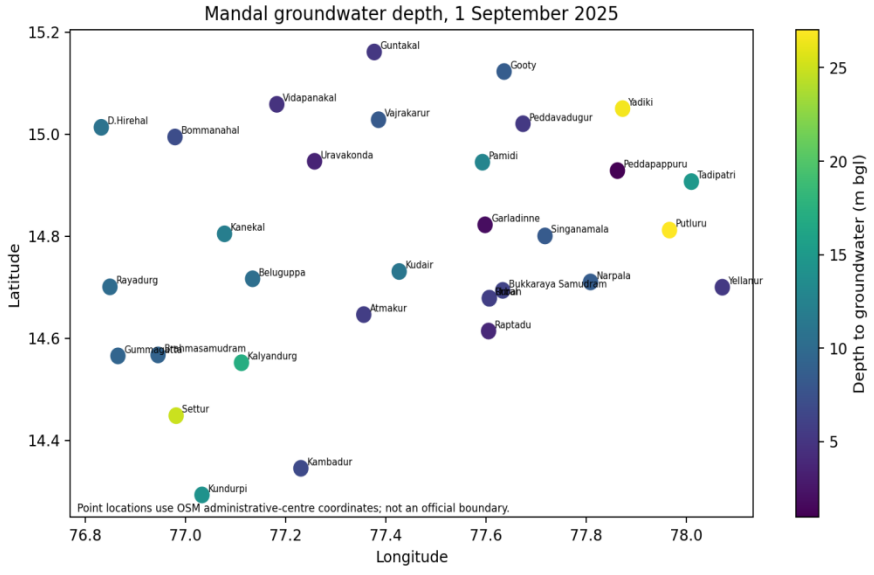


Figure 5. Mandal groundwater-depth visualization.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

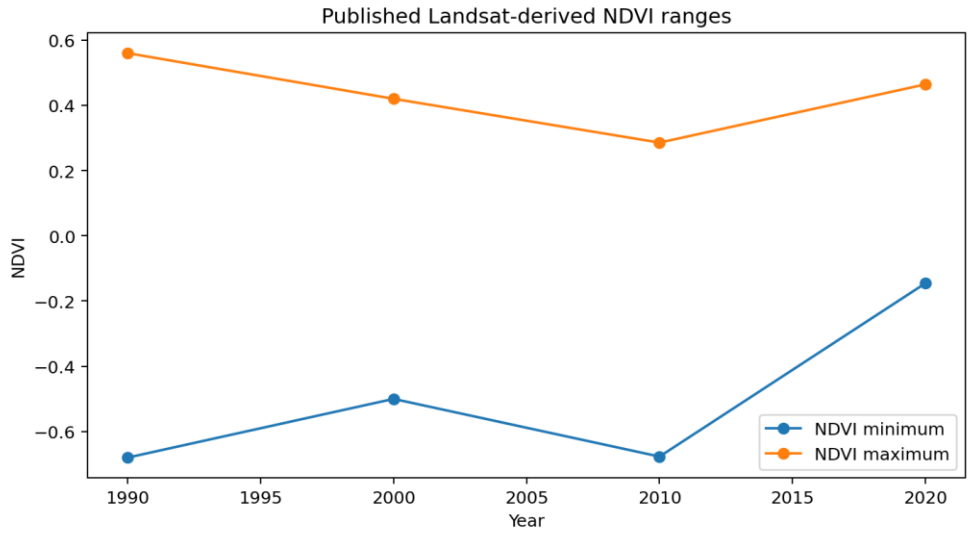


Figure 6. Published Landsat-derived NDVI ranges.

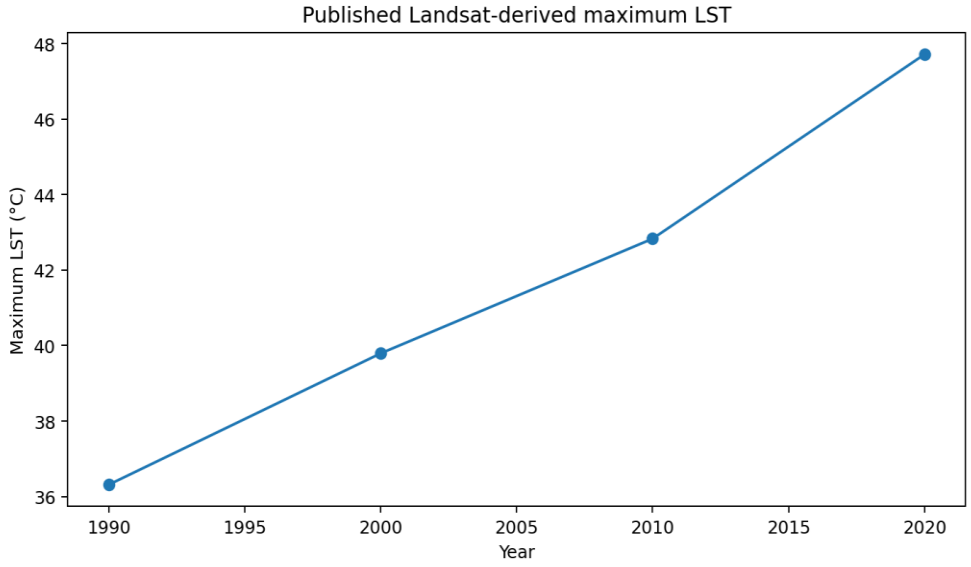


Figure 7. Published Landsat-derived maximum LST.

REFERENCES

1. Badapalli, P. K., Nakkala, A. B., Kottala, R. B., Pujari, P. S., & Gugulothu, S. (2026). Aeolian sand migration induced land degradation and desertification hotspots identification in the semi-arid rain shadow regions of Anantapur, India. Scientific Reports, 16, 1875. https://doi.org/10.1038/s41598-025-31610-0
2. Central Ground Water Board. (2025). Ground water level bulletin: Andhra Pradesh, August 2025. Government of India.
3. Climate Memory. (2026). Anantapur climate — normals and trends. ERA5-based climate series, 1940–2025.
4. European Centre for Medium-Range Weather Forecasts. (2026). ERA5-Land: Reanalysis dataset. Copernicus Climate Change Service.
5. Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324
6. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue X, October2026

7. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
8. OpenStreetMap Wiki. (2022). Sub districts in Anantapur district.
9. Reddy, B. S., Sudheer, K. V. S., Kumar, K. A., Reddy, A. M., & Kumari, C. R. (2020/2021). Assessment of variability in monthly, seasonal and annual rainfall received during 140 years at Ananthapuramu district of Andhra Pradesh. Indian Journal of Dryland Agricultural Research and Development, 35(1), 13–19. https://doi.org/10.5958/2231-6701.2020.00003.2