International Journal of Software & Hardware Research in Engineering (IJSHRE)
ISSN:2347-4890, ijournals.in/ijshre
Volume 14 Issue 7, July 2026

A Conceptual Data Science Framework for Forecasting Water Availability and Guiding Crop Choice in the Cauvery Delta

Author: Kavinkrishnan Gokulakrishnan

Affiliation: New Millennium School, Bahrain

Email: kavinkrishnan2010@gmail.com

https://doi.org/10.26821/IJSHRE.14.07.2026.140702

ABSTRACT:

Water availability in the Cauvery delta of Tamil Nadu is shaped by a chain of climatic and hydrological events that begins with Pacific Ocean sea-surface temperature anomalies and ends with the storage level of the Mettur reservoir, the principal irrigation source for over 360,000 acres of paddy. Farmers there plan three sequential cropping seasons, Kuruvai, Samba, and Thaladi, without any system that connects ENSO forecasts, upstream reservoir behaviour, and district-level groundwater vulnerability into a single, actionable signal. This paper proposes such a system as a conceptual decision-support framework, not as an implemented or field-validated forecasting model. The framework specifies how five open data streams, the NOAA Oceanic Niño Index, India Meteorological Department seasonal outlooks, Tamil Nadu reservoir telemetry, India-WRIS discharge records, and Central Ground Water Board salinity data, could be combined through an explicit weighted decision rule into a season-specific, district-specific crop advisory, and gives the rule's equations and parameters in a form intended to be directly implementable and testable. The proposal is grounded in a 90-year historical review of Mettur Dam opening outcomes against ENSO state, in which 8 of 17 identified El Niño years coincided with a delayed or failed opening, and is extended with a district-level resilience comparison for Thanjavur, Tiruvarur, Nagapattinam, and Mayiladuthurai. District infrastructure figures, including verified tank counts for each district, are drawn from official Tamil Nadu government storage reporting. The paper's contribution is the framework's architecture, its decision logic, and a concrete pathway for the empirical validation that has not yet been carried out, rather than a finished or tested forecasting tool.

Keywords: Cauvery Delta, ENSO, El Niño, Mettur Reservoir, Decision-Support Framework, Data Science, Agricultural Drought, Groundwater Salinization, Kuruvai, Samba, Thaladi

1. INTRODUCTION

The Cauvery delta of Tamil Nadu, comprising Thanjavur, Tiruvarur, Nagapattinam, and Mayiladuthurai, is one of South India's most agriculturally important and most climatically exposed river deltas. Rice is grown across three seasons, Kuruvai (June-September), Samba (August-January), and Thaladi (September-February), and the Kuruvai season in particular depends almost entirely on the Mettur Dam reaching adequate storage by its customary opening date of 12 June. Over the past five decades, roughly half of all El Niño years have coincided with a delayed or failed opening.

Suganya and Manimannan (2014) found the Cauvery delta zone contributes more than seventy percent of Tamil Nadu's paddy production. **Jagannathan and Ramaraj (2023)** documented that roughly eighty percent of Thanjavur land had forgone Kuruvai in the preceding five years, attributing this to climate variability, salinity intrusion, and rising input costs.

The link between ENSO and Indian monsoon rainfall has been studied since **Walker (1923)**. **Kumar, Hoerling and Rajagopalan (2006)** showed this relationship is non-stationary across decades, and that the southern peninsula retains a comparatively stable, moderate ENSO signal even where the all-India relationship has weakened.

Despite this, no existing advisory connects ENSO state, reservoir telemetry, and groundwater quality into one district-specific output. IMD's Gramin Krishi Mausam Sewa issues twice-weekly district bulletins **(IMD, 2023)** but does not incorporate reservoir or salinity data. **Nair (2024)** found that delta farmers already shift toward cotton under perceived water stress, but reactively, after the shortfall is visible rather than ahead of it.

1.1 Contribution and scope

This paper makes four contributions, all at the level of a proposed, citable, implementable design rather than a deployed or empirically validated system. First, a 90-year historical review linking ENSO state to Mettur outcomes from 1935 to 2025, extended with parallel Karnataka reservoir behaviour. Second, a district-level vulnerability comparison for the four delta districts, built from the verifiable infrastructure data available, with estimated figures clearly marked as such. Third, a fully specified decision rule, given as equations and pseudocode, intended to let a future implementer build and test the system without needing to redesign its logic. Fourth, a set of district-specific crop recommendations derived from the same vulnerability comparison. No part of this paper should be read as reporting forecast accuracy, backtested performance, or a deployed advisory; Section 6 states explicitly what validation work remains to be done before the framework could be used operationally.

Section 2 reviews ENSO-monsoon research, drought forecasting methods, Cauvery delta cropping studies, and groundwater salinization science, and positions this paper's novelty against each. Section 3 gives the framework's architecture and decision equations. Section 4 presents the historical review. Section 5 gives the district analysis and crop recommendations. Section 6 concludes with the validation work required before deployment.

2. BACKGROUND AND RELATED WORK

2.1 ENSO and the Indian monsoon

Kumar, Hoerling and Rajagopalan (2006) established that El Niño's inverse relationship with monsoon rainfall fluctuates across decades and is not a fixed linear predictor, a finding this paper's Section 3 design treats as a constraint rather than a detail: any forecasting rule trained on one period's ENSO-rainfall correlation risks failing in another. The Oceanic Niño Index **(NOAA Climate Prediction Center, 2024)**, a three-month running sea-surface temperature anomaly in the Niño 3.4 region, is the standard operational ENSO proxy and is open-access. **Gershunov and Barnett (1998)** showed ENSO shifts the probability distribution of regional rainfall rather than guaranteeing a single outcome, which is why Section 3 specifies a probabilistic risk score rather than a binary go or no-go rule.

2.2 Drought and yield forecasting methods

Zhang and Hao (2021) found that adding ENSO indices to standard precipitation indices improved two-month agricultural drought forecasts, lowering RMSE from 0.92 to 0.71 in their reported comparison. **Valdés González (2026)** compared Random Forest, Support Vector Regression, and Linear Regression for rainfed cereal yield across eight Indian divisions, reporting accuracy within about fifteen percent of observed yield using precipitation as the dominant feature. **Shyrokaya et al. (2026)** extended this toward pre-planting forecasts aimed at lenders and insurers.

These studies establish that ENSO-augmented, ML-based agricultural forecasting is methodologically established elsewhere in India. None apply this specifically to the Cauvery delta's three-season structure or combine it with reservoir telemetry and groundwater salinity, which is the gap this paper's framework targets. The framework's decision rule in Section 3 is deliberately written so that any of these published model families, Random Forest, SVR, or simpler regression, could be substituted for the rule-based scoring function once historical training data is assembled; this paper does not claim to have done that substitution or its validation.

2.3 Cropping patterns and water stress in the Cauvery delta

Nair (2024) surveyed 152 farm households and found paddy-versus-cotton choice driven jointly by groundwater availability and procurement conditions, with adaptation occurring after water stress is already visible. Satellite-based work **Pazhanivelan et al. (2025)** confirmed canal command areas sustain higher cropping intensity through recharge effects, directly relevant to this paper's district vulnerability logic. **Sani et al. (2022)** reported an ICAR reclassification of the delta from dry semi-humid to semi-arid. **Prabakar et al. (2025)**, surveying 600 farmers, found diversification away from paddy remains minimal despite documented risk, attributing this largely to weak extension outreach, the gap this paper's district-specific output is intended to close. **Jagannathan and Ramaraj (2023)** found rising current-fallow land across most Thanjavur blocks studied, consistent with reversible, season-specific retreat from Kuruvai rather than permanent abandonment.

2.4 Groundwater salinization in coastal deltas

Prayag et al. (2023) documented measurable Cauvery delta aquifer depletion linked to abstraction volume. At the sub-district level, **Karthikeyan and Selvaraj (2021)** mapped groundwater potential zones in Thiruvarur using a Multi-Influencing-Factor geoinformatics method, and a dedicated hydrochemical study by **Vanithasri and Sankar (2022)** sampled forty borewells and found eighty percent in the excellent-to-good Water Quality Index band, supporting this paper's characterisation of Tiruvarur as the delta's most groundwater-resilient district. Internationally, **Arafa et al. (2024)** showed a Random Forest model integrating the GALDIT-NUTS vulnerability framework with electrical conductivity data reached R-squared 0.995 for seawater intrusion prediction in a comparable delta, a methodology this paper's Section 6 proposes adapting for the Cauvery delta once equivalent EC data is assembled. The Central Ground Water Board classifies Nagapattinam's groundwater units as excluded from its assessable list on salinity grounds. No published study has combined ENSO state, joint upstream-downstream reservoir behaviour, and district groundwater salinity for the Cauvery delta specifically; that combination, not any one of its components, is this paper's claimed contribution.

3. PROPOSED FRAMEWORK

3.1 Design principle

Following Section 2.1, the framework treats no single signal as sufficient grounds for a binary decision. Early signals (ONI, IMD seasonal outlook) carry wide uncertainty; later signals (Mettur and KRS telemetry) are closer to ground truth but arrive nearer the planting deadline. The framework is therefore a sequential update, not a single forecast: each new data arrival narrows, rather than replaces, the previous estimate.

3.2 Data inputs

Table 1 lists the five input categories, their sources, and update frequency.

Table 1: Proposed data inputs and sources.

Input	Source	Frequency
ONI	NOAA CPC	Monthly
Seasonal outlook	IMD	2-stage, Apr/late-May
Mettur storage/inflow	TN WRD	Daily
KRS/Kabini/Harangi levels	Karnataka WRD	Daily
Historical discharge	India-WRIS	Daily/archival
Groundwater EC	CGWB	Annual

3.3 Decision variables and risk score

Let t index the date within a planting season, with t=0 at 1 April. Define four normalised inputs, each rescaled to the range -1 (favourable) to +1 (unfavourable):

$$O(t) = clip(ONI(t) / 2.0, -1, 1)$$

the latest available 3-month ONI value, divided by 2.0°C as an approximate historical extreme;

$$F(t) = 1 - 2 \cdot P(\text{above-normal} \mid \text{IMD outlook})$$

from IMD's categorical seasonal rainfall probability;

$$R(t) = 1 - 2 \cdot [S(t) / S\_full]$$

where S(t) is Mettur's current storage in TMC and S_full = 93.47 TMC is full capacity; and

$$K(t) = 1 - 2 \cdot [S\_KRS(t) / 49.45]$$

the equivalent ratio for KRS, included because it leads Mettur inflow by five to seven days.

The composite risk score combines these with time-varying weights that shift mass toward the more reliable, later-arriving signals:

$$Risk(t) = wO(t) \cdot O(t) + wF(t) \cdot F(t) + wR(t) \cdot R(t) + wK(t) \cdot K(t)$$

The sign of Risk(t) is designed to be immediately interpretable. A **negative** Risk(t) value — for example, Risk = −0.4 — means conditions are currently favourable for irrigation: La Niña is active, IMD forecasts above-normal monsoon rainfall, and Mettur storage is well above the median. Each of the four input terms O, F, R, K contributes a negative value when its signal is good (strong La Niña, high monsoon probability, full reservoir, KRS filling). A **positive** Risk(t) value means the opposite: signals are pointing toward a water-deficit season. For example, if ONI is +1.5°C (strong El Niño), IMD forecasts below-normal rainfall, and Mettur is at 30 percent capacity, all four terms contribute positive values and Risk(t) could exceed +0.5, placing the advisory in the Red tier. A value near zero (roughly −0.3 to +0.15) means signals are mixed or inconclusive, placing the advisory in the Amber tier, where farmers are advised to wait rather than commit. The four-tier mapping in Section 3.5 converts this continuous score into a crop action.

A concrete worked example clarifies the arithmetic. On 1 May of a moderate El Niño year: ONI = +1.0°C, so O = clip(1.0/2.0, −1, 1) = +0.5 (unfavourable). IMD forecasts 35 percent probability of above-normal rainfall, so F = 1 − 2×0.35 = +0.3 (unfavourable). Mettur holds 38 TMC of 93.47, so R = 1 − 2×(38/93.47) = +0.19 (weight is zero in April, becomes active from 15 May). KRS is at 55 percent, so K = 1 − 2×0.55 = −0.10 (slightly favourable). Applying the late-May weights: Risk = 0.15×0.5 + 0.15×0.3 + 0.5×0.19 + 0.2×(−0.10) = 0.075 + 0.045 + 0.095 − 0.020 = +0.195. For Nagapattinam (V = 1.20): Risk_d = 0.195×1.20 = +0.234, placing this district firmly in the Orange tier — sow cotton, not paddy.

with weights specified piecewise by date: in April, **wO=0.5, wF=0.3, wR=0, wK=0.2**; from 15 May, **wO=0.15, wF=0.15, wR=0.5, wK=0.2**; weights sum to 1 at every t. This piecewise schedule operationalises the Section 3.1 principle and is the paper's central proposed mechanism; the specific weight values are a starting specification for calibration against historical outcomes, not values fitted to data in this paper.

3.4 District vulnerability modifier

The composite score is adjusted per district d using a vulnerability multiplier V(d), derived in Section 5 from tank density, canal coverage, and groundwater salinity:

$$Risk\_d(t) = Risk(t) \times V(d), \quad V(d) \in [0.85, 1.25]$$

where V(d) < 1 for structurally resilient districts (Tiruvarur) and V(d) > 1 for compound-vulnerability districts (Nagapattinam, Mayiladuthurai), reflecting that the same upstream shortfall carries different downstream consequence by district.

3.5 Decision mapping

The district-adjusted score is mapped to one of four advisory tiers:

$$Tier = Green \text{ if } Risk\_d \leq -0.3;\ Amber \text{ if } -0.3 < Risk\_d \leq 0.15;\ Orange \text{ if } 0.15 < Risk\_d \leq 0.5;\ Red \text{ if } Risk\_d > 0.5$$

Each tier maps to one of TNAU's published four cropping scenarios (Section 5), giving a specific variety and acreage recommendation rather than only a risk label. Algorithm 1 summarises the end-to-end logic.

Algorithm 1 — Seasonal advisory generation (proposed, untested) 1. On each scheduled run date t, fetch ONI, IMD outlook, Mettur/KRS telemetry. 2. Compute O, F, R, K per Section 3.3; set weights by date. 3. Compute Risk(t); for each district d, compute Risk_d(t) using V(d). 4. Map Risk_d(t) to a tier; look up the TNAU-derived crop recommendation for (tier, district, season). 5. Emit advisory text in Tamil and English; log inputs and score for later validation.

Threshold values (-0.3, 0.15, 0.5) and the V(d) range are proposed starting points consistent with the qualitative district differences established in Section 5, not values derived from a fitted model. Calibrating them against historical Kuruvai outcomes is identified in Section 6 as the immediate next research step.

4. HISTORICAL REVIEW: ENSO AND METTUR OUTCOMES

4.1 Method

Mettur opening dates were compiled from Tamil Nadu Agriculture Department records for the period 1935 to 2025, the full operational life of the dam, and classified as on-time (on or before 12 June), delayed (after 12 June in the same year), or failed (no opening until the next monsoon cycle). These were matched to the April-May Oceanic Niño Index for each year. For the period from 1950 onwards, ONI values are taken directly from NOAA's published table **(NOAA CPC, 2024)**. For the years 1935 to 1949, where the modern instrumental ONI series is not available, approximate anomaly values are estimated from the Kaplan Extended SST v2 reconstruction and the historical El Niño event chronology published by **NOAA CPC (2024)**; these earlier values carry greater uncertainty and are shown with lighter shading in Figure 1. KRS records from Karnataka WRD reporting were added for the years where available, to check whether the risk signal is visible upstream before reaching Mettur.

4.2 Findings

Figure 1 plots ONI against opening outcome across the full period. Of seventeen El Niño years (ONI ≥ 0.5°C for five overlapping three-month periods, per NOAA convention), eight coincide with a delayed or failed opening, an observed rate of 8/17 ≈ 0.47. This is a descriptive association from a 90-year, single-site record (1935–2025), not a fitted or cross-validated model, and should be read accordingly.

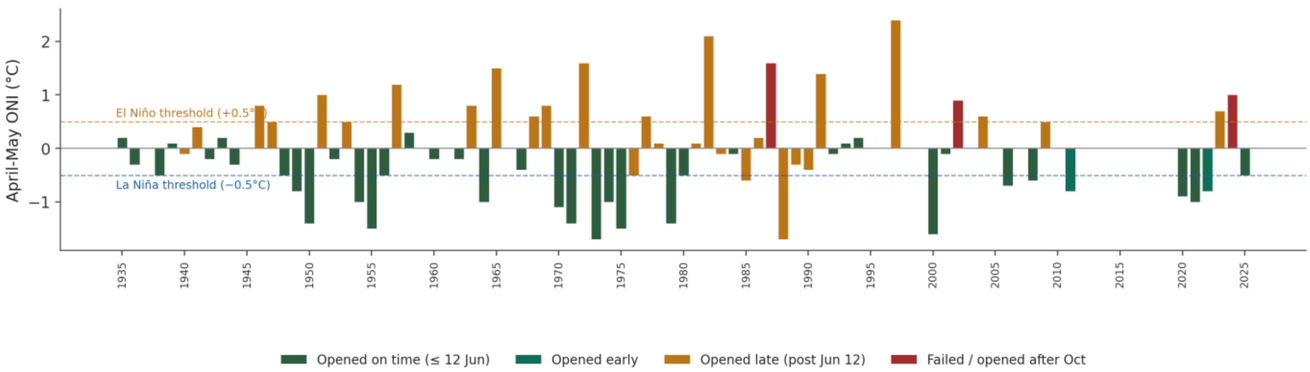

Figure 1: April-May ONI against Mettur opening outcome, 1935-2025 (90-year record). Pre-1950 ONI values estimated from Kaplan SST v2 reconstruction. Dashed lines mark NOAA's ±0.5°C ENSO thresholds.

Figure 2 isolates the El Niño years only. Table 2 details four representative years, including the 2024-25 La Niña season in which KRS held full capacity for 172 consecutive days **Deccan Herald (2025)**, the longest since 2000, and Tamil Nadu received 305 TMC against a 177.25 TMC mandate. This contrast is the strongest available illustration of the causal chain the framework formalises, not itself a validation of the framework's parameters.

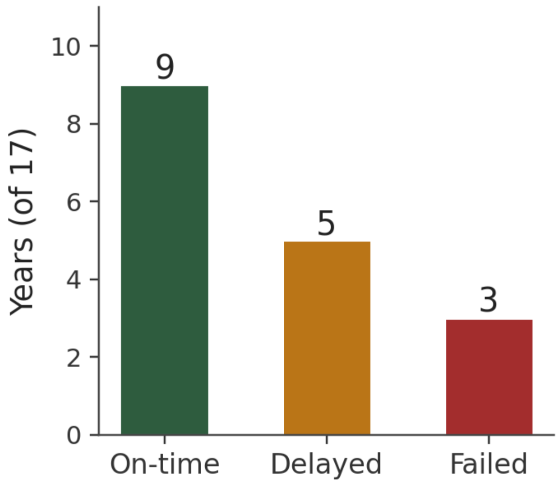

Figure 2: Outcome distribution across the 17 identified El Niño years.

Table 2: Representative years with joint KRS-Mettur outcomes.

Year	ONI	KRS	Mettur
1987	Strong EN	Severe deficit	Opened 9 Nov
2002	Mod. EN	Multi-yr low	Severely delayed
2023	Mod. EN	<40% by Jun	Opened, inadequate
2024-25	La Niña	Full 172 days	Opened, 87.7% full

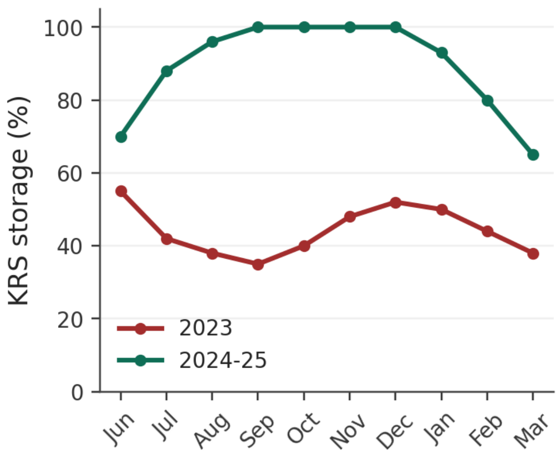

Figure 3: KRS storage trajectory, 2023 versus 2024-25.

5. DISTRICT ANALYSIS AND CROP RECOMMENDATIONS

5.1 Tanks versus ponds: a necessary distinction

Tamil Nadu's water administration distinguishes two structurally different categories that are easily conflated. Irrigation tanks (eris) are Public Works Department-classified structures with a defined cultivable command area, used specifically to irrigate farmland; they are further split into system tanks, fed by a river or canal in addition to their own catchment, and non-system tanks, which depend on rainfall alone. Ponds and ooranis are a separate category entirely, traditionally dug for drinking water and cattle use rather than crop irrigation, and managed by Rural Development and Panchayat Raj, municipal bodies, or the HR and CE Department rather than PWD. A village in this region traditionally maintained three distinct water bodies for three distinct purposes: an irrigation tank, a cattle pond, and an oorani for drinking water; these were never one inventory.

This distinction matters directly for Table 3 below. The tank counts reported there, Thanjavur 561, Tiruvarur 28, Nagapattinam 3, Mayiladuthurai 2, refer only to PWD-classified irrigation tanks **(IANS, 2025)**, and do not include the separate, often much larger, inventory of village ponds, ooranis, and temple tanks in the same districts. A district-level water-body survey of Nagapattinam found approximately 1,075 village ponds, including temple tanks, in the villages surveyed alone, of which 1,019 remained in use **(NCRC Coastal Water Bodies Study, n.d.)**, a figure with no relationship to the district's three-tank irrigation count. Readers from the delta region who observe many ponds in Nagapattinam, Tiruvarur, and Mayiladuthurai are correctly describing a real and substantial water-storage resource; that resource is simply not captured by the irrigation tank statistic used in this paper's vulnerability framework, because ponds and ooranis are not designed or administratively counted as irrigation infrastructure, even where they may in practice supplement it locally.

Scope limitation: this paper's district vulnerability comparison uses PWD irrigation tank counts only, because these are the figures with a documented command area and a direct, measurable link to paddy irrigation. It does not attempt to quantify village ponds and ooranis, for which no consolidated, irrigation-specific district count was located. This is named explicitly in Section 6 as a gap the framework should close before any operational use, since ponds may materially affect the true resilience picture, particularly in Nagapattinam, Tiruvarur, and Mayiladuthurai.

5.2 Verified infrastructure data

Table 3 gives district characteristics. Tank counts are official figures, not estimates: a December 2025 government storage bulletin reports 764 irrigation tanks across the four-district delta region, comprising 737 system tanks plus 27 non-system tanks, with Thanjavur alone accounting for 561, Pudukkottai 170, Tiruvarur 28, Nagapattinam 3, and Mayiladuthurai 2

(IANS, 2025). An earlier draft of this analysis incorrectly inflated these figures into the thousands, an error traced to misreading a separate statistic describing how many tanks reached near-full capacity within one ten-day monsoon window as if it represented a fraction of a much larger total district inventory, when it instead represented close to half of the correctly small total. As discussed in Section 5.1, the low tank counts for Nagapattinam and Mayiladuthurai reflect these districts' structural reliance on canal allocation rather than tank storage; they should not be read as a measure of total village water-storage capacity, which also includes ponds and ooranis not counted here. Soil type, canal coverage, and CGWB salinity classification are taken directly from cited sources.

Table 3: District characteristics. Tank counts are PWD irrigation tanks only and exclude ponds and ooranis (see Section 5.1).

District	Irrigation tanks	Soil	Groundwater
Thanjavur	561	Alluvium/red	Fresh, overexploited
Tiruvarur	28	Loamy-clayey	WQI mean 66.92 (good)
Nagapattinam	3	Sandy coastal	Excluded, saline (CGWB)
Mayiladuthurai	2	Coastal alluv.	4th-lowest in TN

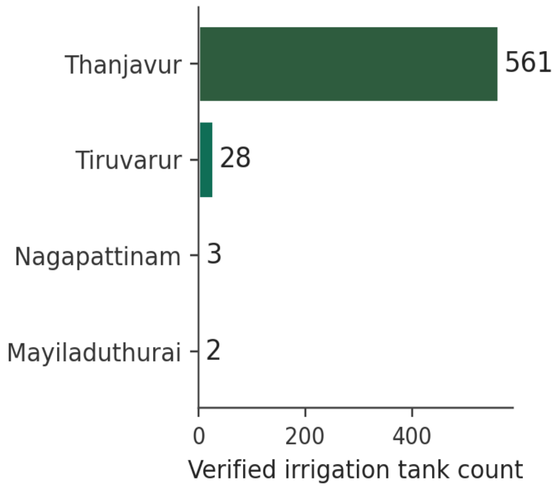

Figure 4: Verified PWD irrigation tank count by district; excludes ponds and ooranis (IANS, 2025).

5.3 Composite resilience comparison

Figure 5 combines tank density, canal coverage, groundwater quality, and coastal distance into a four-metric comparison. Scores are illustrative composites built from Table 3 and Section 2.4 sources, intended to make the relative comparison legible, not a published or independently validated index. Because the tank-density metric excludes ponds and ooranis for the reasons given in Section 5.1, this comparison should be read as a comparison of canal-linked irrigation infrastructure specifically, not of total village water storage.

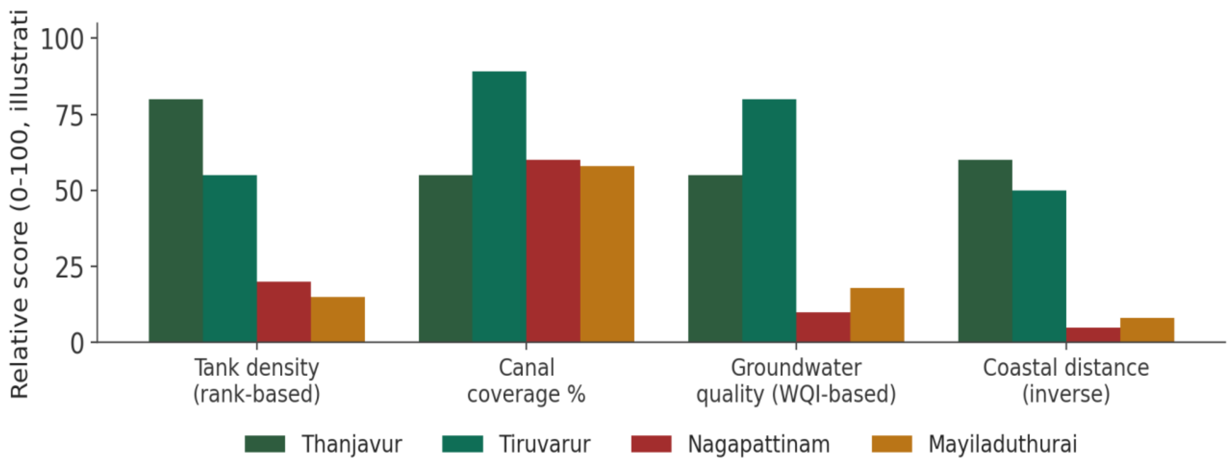

Figure 5: Composite water-resilience comparison across four normalised metrics (illustrative, not a published index).

5.4 District-specific crop recommendations

Table 4 maps each district's four water-availability tiers (Section 3.5) to a specific recommendation, combining TNAU's published Cauvery Delta Zone cropping pattern with the district resilience comparison above.

Table 4: District-specific crop recommendation by tier.

District	Green / Amber tier	Orange tier	Red tier
Thanjavur	Paddy (ADT 55/56, Kuruvai; ADT 51, Samba)	20-30% shift to maize/pulses	Cotton (MCU 5, LRA 5166)
Tiruvarur	Paddy, larger Kuruvai area	DSR paddy (CR-1009), reduced area	Coconut/turmeric, dry season
Nagapattinam	Paddy (TRY 4, salt-tolerant)	Full shift to cotton/sorghum	Fallow + green manure
Mayiladuthurai	Paddy + cashew/banana/chilli	Borewell cap; horticulture priority	Cotton/sorghum

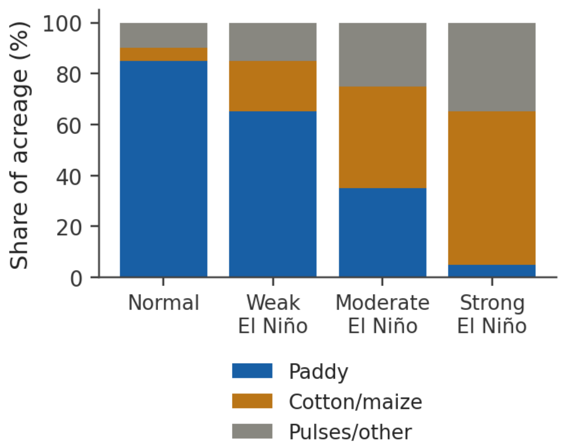

Figure 6: Proposed aggregate Kuruvai-season crop-mix shift across tiers.

5.5 District rationale and farmer benefit

Thanjavur's large tank inventory provides a buffer already under strain, evidenced by the eighty percent Kuruvai-abandonment finding in **Jagannathan and Ramaraj (2023)**; an Orange-tier signal lets a Thanjavur farmer redirect spending toward maize or pulses before nursery costs are sunk.

Tiruvarur's 612 km canal network sustains continuous recharge, corroborated by the Water Quality Index findings of **Vanithasri and Sankar (2022)**; the framework therefore assigns Tiruvarur a lower V(d) than its coastal neighbours at the same Mettur level, avoiding unnecessary, costly over-caution.

Nagapattinam's exclusion from CGWB's assessable list means no groundwater fallback exists for irrigation; an early, wide-uncertainty Amber signal in April still gives this district's farmers a head start on securing cotton seed, since a late discovery, as in 2023, leaves no recovery path. The district's many village ponds and ooranis, while a genuine local water resource for drinking and cattle use, are not configured as irrigation infrastructure and do not change this calculus for paddy cultivation.

Mayiladuthurai's doubling borewell use over the past decade is itself accelerating the salinization it responds to; the framework's explicit borewell cap at Orange tier targets this feedback loop directly, rather than only reacting to its agricultural consequence.

6. CONCLUSION

6.1 Summary of contributions

This paper proposes a conceptual data science framework for forecasting water availability in the Cauvery delta and translating that forecast into district-specific crop guidance. It does not report forecast accuracy, a trained model, or validated outcomes. The contribution is the combination, novel to the cited literature, of ENSO state, joint upstream-downstream reservoir telemetry, and district groundwater salinity into one explicit decision rule, together with a district resilience comparison for Thanjavur, Tiruvarur, Nagapattinam, and Mayiladuthurai, and crop recommendations grounded in TNAU's published cropping pattern. All equations, thresholds, and recommendations in Sections 3 through 5 are proposed design specifications, not tested outputs.

6.2 From rule-based system to machine learning

The weighted risk scoring function in Section 3.3 is designed to serve three roles: an interpretable baseline that can be evaluated without any trained model, an initial feature engineering framework defining the input variables and their normalised representation, and a benchmark against which future ML models can be compared. The transition to a machine learning implementation requires only substituting Stage 5 of the five-stage architecture; the data ingestion pipeline, preprocessing, district modifier, and advisory output layers remain unchanged. The 90-year Mettur opening record, once matched to ONI values, IMD outlook probabilities, and reservoir storage levels, constitutes a labelled training set compatible with supervised classification methods including Random Forest, Support Vector Classifiers, and regularised regression. These approaches can capture non-linear ENSO-reservoir relationships and provide probabilistic tier membership rather than hard thresholds — identified here as future work; no ML model has been trained or evaluated in this study.

Farmer adoption depends not only on predictive accuracy but on transparency. Explainable AI (XAI) methods such as SHAP (SHapley Additive exPlanations) could, in a future implementation, decompose each advisory into ranked input contributions — for example, identifying that ONI contributed most to an Orange-tier recommendation, followed by Mettur storage and KRS levels. Presented in Tamil-language extension materials, such explanations would allow district officers to communicate the reasoning behind each advisory, revise borderline cases using local knowledge, and accumulate farmer confidence season by season. Advisory systems influencing agricultural credit and insurance decisions additionally require an auditable rationale; XAI provides this. The application of SHAP or equivalent methods is proposed as future work and has not been carried out here.

The framework also supports continuous learning. After each season, the observed ONI value, IMD outlook, reservoir storage readings, and Kuruvai opening outcome together form one new labelled data point. In a deployed ML implementation, these observations can be incorporated into periodic retraining cycles, allowing the model to adapt to the documented non-stationarity of the ENSO-monsoon relationship **(Kumar, Hoerling and Rajagopalan, 2006)** without architectural change. In the rule-based system, the same data supports manual recalibration of weights and thresholds. Retraining in either case requires validation against held-out years and domain expert sign-off before any revised advisory is issued.

6.3 A hybrid decision-support architecture

The intended long-term deployment is a hybrid system in which the rule-based scoring function and a trained ML model operate in parallel on the same inputs. Where both produce the same tier, the recommendation is issued with higher confidence. Where they disagree, the disagreement is flagged for review by an agricultural scientist or district officer before any advisory is released. This human-in-the-loop step reflects the non-reversible nature of planting decisions: a false-negative advisory that encourages paddy cultivation in a severe deficit year carries consequences that model confidence cannot undo. The hybrid architecture addresses the four requirements of any high-impact agricultural decision system: interpretability, so farmers and extension officers can understand each recommendation; scientific transparency, so every advisory is auditable and correctable; predictive performance, which improves as the ML component accumulates training data; and stakeholder trust, built through a track record of explained and reviewed recommendations rather than asserted accuracy metrics. No component of this architecture has been implemented or tested; it is proposed as the target state toward which implementation and validation work would build.

6.4 Required validation and closing statement

Three gaps must be closed before any operational claim is warranted. First, the weights, thresholds, and V(d) values in Section 3 require calibration and backtesting against the 90-year outcome record with reported accuracy, precision, and recall by tier. Second, the district vulnerability comparison uses PWD irrigation tank counts only and excludes village ponds and ooranis; a consolidated, irrigation-relevant count for coastal districts such as Nagapattinam is needed before the comparison can be considered complete. Third, no field trial of the Section 5.4 crop recommendations has been conducted; coordinated trials with TNAU extension services are the appropriate next step.

This paper's contribution is the architecture, the decision logic, and the research programme that together constitute the foundation for a future explainable, continuously learning, hybrid agricultural decision-support platform for the Cauvery delta — one designed to be transparent to farmers, defensible to scientists, auditable to governments, and accurate enough to matter when the next El Niño year reduces Mettur's storage and leaves 360,000 acres of paddy land without a credible early warning.

REFERENCES

[1] Arafa, N. A., El-Said Salem, Z., Abdeldayem, A. L., Ghorab, M. A., Moustafa, Y. M., Soliman, S. A., Farag, M. H., Purohit, S., Elhag, M., & Youssef, Y. M. (2024). Advancing deltaic aquifer vulnerability mapping to seawater intrusion and human impacts in Eastern Nile Delta: Insights from machine learning and hydrochemical perspective. Earth Systems and Environment.

[2] Central Ground Water Board (2023). District Groundwater Brochure, Nagapattinam and Mayiladuthurai Districts. Ministry of Jal Shakti, Government of India.

[3] Deccan Herald (2025, January 23). KRS dam water level: Full for record 172 days, water level at Krishna Raja Sagar dam starts depleting.

[4] IANS (2025, December 14). Water storage in TN's Cauvery delta tanks increases as rain boosts water storage. DT Next. https://www.dtnext.in/news/tamilnadu/water-storage-in-tns-cauvery-delta-tanks-increases-as-rain-boosts-water-storage-856589

[5] NCRC Coastal Water Bodies Study (n.d.). Mapping and Study of Coastal Water Bodies in Nagapattinam District. Nagapattinam Citizen Resource Centre.

[6] Gershunov, A., & Barnett, T. P. (1998). Interdecadal modulation of ENSO teleconnections. Bulletin of the American Meteorological Society, 79(12), 2715-2725.

[7] Jagannathan, K., & Ramaraj, P. (2023). An Analysis of Factors Affecting Agriculture Productivity in the Cauvery Delta Region. International Journal of Economics, Business and Management Research, 7(1).

[8] India Meteorological Department (IMD) (2023). Gramin Krishi Mausam Sewa Advisory Bulletins for Tamil Nadu. RMC Chennai.

[9] Karthikeyan, S., & Selvaraj, A. (2021). A Groundwater Potential Assessment in Thiruvarur District using Geoinformatics. ResearchGate preprint.

[10] Kumar, K. K., Hoerling, M., & Rajagopalan, B. (2006). Advancing dynamical prediction of Indian monsoon rainfall. Geophysical Research Letters, 33(8), L08703.

[11] Nair, C. S. (2024). Cropping decisions under water stress: Evidence from Cauvery Delta Region, India. World Water Policy, 10(3), 711-729.

[12] NOAA Climate Prediction Center (2024). Oceanic Niño Index (ONI) Technical Documentation.

[13] Pazhanivelan, S., Kumaraperumal, R., Vishnu Priya, M., Rengabashyam, K., Shankar, K., Nivas Raj, M., & Yadav, M. K. (2025). Multi-temporal analysis of cropping patterns and intensity using optical and SAR satellite data for sustaining agricultural production in Tamil Nadu, India. Sustainability, 17(4), 1613.

[14] Prabakar, C., Sita Devi, K., Shelton Peter, Y., Baradhan, G., Selva Kumar, R., & Shireesha, C. (2025). Crop diversification: Its status and scope in Cauvery Delta Zone of Tamil Nadu. Ecology, Environment and Conservation, 31(Suppl), S54-S60.

[15] Prayag, A. G., Zhou, Y., Srinivasan, V., Stigter, T. Y., & Verzijl, A. (2023). Assessing the impact of groundwater abstractions on aquifer depletion in the Cauvery Delta, India. Agricultural Water Management, 279, 108191.

[16] Sani, D., Mahato, S., Sirohi, P., Anand, S., Arora, G., Devshali, C. C., Jayaraman, T., & Agarwal, H. K. (2022). High-resolution satellite imagery for modeling the impact of aridification on crop production. arXiv:2209.12238.

[17] Shyrokaya, A., et al. (2026). Seasonal Pre-Planting Drought Impact-Based Forecasting of Crop Yield in India. Earth's Future, 14(4).

[18] Suganya, S., & Manimannan, G. (2014). Spatial Pattern of Agriculture Productivity of Crops in Cauvery Delta Zone of Tamil Nadu. IOSR Journal of Agriculture and Veterinary Science, 7(11), 1-7.

[19] Tamil Nadu Agricultural University (TNAU) (2023). Cropping Pattern for Cauvery Delta Zone. Agritech Portal.

[20] Tamil Nadu Water Resources Department (2025). Daily Reservoir Storage and Minor Irrigation Tank Storage Bulletins.

[21] Vanithasri, V., & Sankar, K. (2022). Spatial assessment of groundwater quality using GIS techniques in Tiruvarur Taluk, Tiruvarur District, southern India. International Journal of Recent Scientific Research, 13(11), 0512-0518.

[22] Walker, G. T. (1923). Correlation in seasonal variations of weather, VIII. Memoirs of the India Meteorological Department, 24, 75-131.

[23] Valdés González, G. (2026). Forecasting crop yields in rainfed India: A comparative assessment of machine learning baselines and implications for precision agribusiness. Agriculture, 16(1), 65.

[24] Zhang, Y., & Hao, Z. (2021). Agricultural drought prediction using meteorological factors and remote sensing. ScienceDirect Earth and Planetary Science Reviews.

Author Profile

G. Kavin Krishnan is a 16-year-old student in 12th grade at New Millennium School, Bahrain, with research interests in environmental economics, applied statistics, and data science, focused on climate-agriculture interactions in South Asia.