

AP Water Intelligence

Data sources and the groundwater model

A technical note for review · figures from the weekly refresh of 19 October 2026



Depth to water in September, each mandal against its own usual September, as the platform's 3D view draws it: blue shallower than usual, slate about usual, amber deeper, red deeper than any September on record.

670

mandals mapped, each compared with its own history

21

public and official data sources, combined every week

0.91 m

average nowcast error; 13% better than the last reading

Mon

automated refresh; publishes only if about 430 tests pass

INSIDE

01

What it is

the question, three rules, the pipeline

02

Data sources

21 sources, what each gives and does

03

The model

method, tests, release rule

04

This week

four maps of 670 mandals

05

Agriculture

the crop, the season, the check

06

Open questions

limits, and where review helps

07

Closing the gap

six inputs, and what each moves

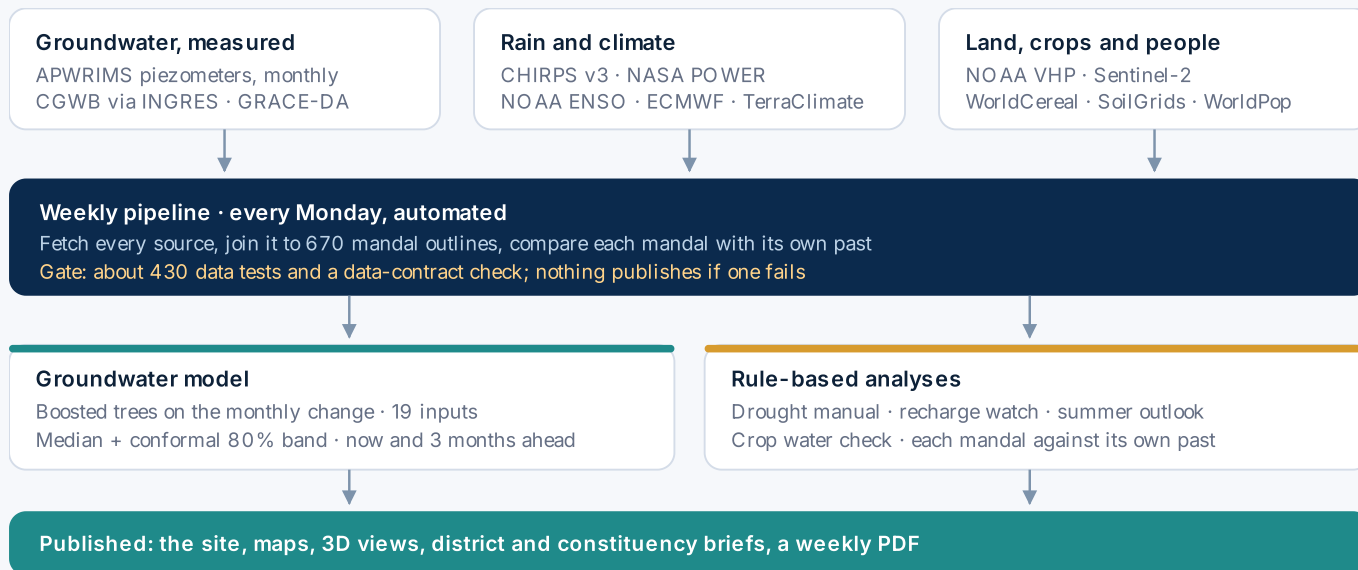
Every mandal's water, checked weekly

A working prototype for the Chief Minister's office and district officers. Every Monday it reports where groundwater, rain, soil moisture and crops stand in each of Andhra Pradesh's **670 mandals**, and what is likely next.

The question is local: **is this mandal's water table where it usually is for this month, is it heading somewhere it has not been, and do rain, soil and crops agree?** Three rules run through the whole platform.

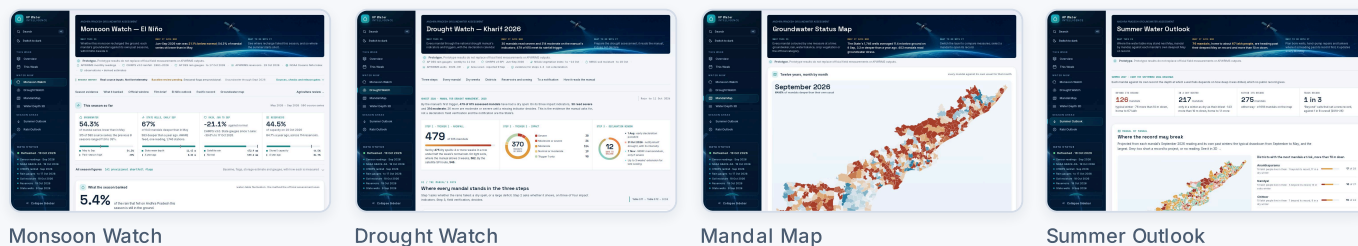


HOW IT FITS TOGETHER



Measured readings pass straight through to the pages. Only the model's estimates and the rule-based calls are derived, and each says so. The site's screens: a weekly overview, Monsoon Watch, Drought Watch, a mandal map, the 3D depth view, Summer and Rabi outlooks, a crop water check, and briefs by district and assembly constituency; a one-page digest is printed each week.

THE SITE ITSELF



One measured series, twenty supporting ones

Only the first row is what the model predicts; the rest are inputs, cross-checks or context, each **joined to the same 670 mandal outlines** before use.



Water the measured target and official context

SOURCE	WHAT IT GIVES	RESOLUTION · CADENCE	USED FOR
APWRIMS AP Water Resources Information & Management System	Depth to water below ground per mandal, from state piezometers, 2014 to date	mandal · monthly	The measured series the model learns from and every page reports
APWRIMS context feeds AP DES gauges · NRSC soil moisture · reservoirs	Mandal rain gauges, modelled soil moisture, reservoir storage and canal releases	mandal · daily to weekly	Drought triggers, the crop check, reservoir context
CGWB assessment via INGRES hosted by IIT Hyderabad	2024 Dynamic Ground Water Resource Assessment: stage of extraction and category	unit · annual	The official safe → over-exploited label
CGWB station levels via India Data Portal	Central network well readings, 2013 to 2023	station · seasonal	Checks the forecast on wells it never saw
Jal Jeevan Mission IMIS Dept. of Drinking Water and Sanitation	Drinking-water schemes by source; sources tested	block · weekly	Who drinks from wells past their record
NASA GRACE-DA GRACE gravimetry assimilated in a land model	Groundwater storage percentile	coarse grid · weekly	Regional storage context only
AP Data Lake AI Living Labs, access-controlled	The State's official groundwater snapshot and mandal outlines	mandal · on request	Official boundaries; well-count check



Rain and climate model inputs and climate context

SOURCE	WHAT IT GIVES	RESOLUTION · CADENCE	USED FOR
CHIRPS v3.0 UC Santa Barbara Climate Hazards Center	Satellite-and-gauge rainfall, 1981 to date	0.05° · monthly	The model's rainfall input , anomalies, SPI
IMD gridded rainfall India Meteorological Department	Observed rainfall (Pai et al. 2014)	0.25° · yearly release	Cross-check of CHIRPS; not the input
NASA POWER MERRA-2 reanalysis	Rainfall	≈55 km · monthly	Fallback where CHIRPS is missing
TerraClimate University of Idaho	Actual evapotranspiration, water balance	≈4 km · monthly	Rain-minus-ET context (not recharge)
NOAA CPC Oceanic Niño Index, ENSO outlook	El Niño state and official outlook	index · monthly	El Niño context and composites
NOAA ERSST v5 · GHCN-CAMS with NOAA GlobalTemp	Sea-surface and air temperature	2° / 0.5° · monthly	Pacific panel, temperature record
ECMWF open forecast via Open-Meteo	Reference ET and rain, 7 to 10 days ahead	0.25° · daily	The weekly crop water check



Land, crops and people crop stress, cropland and exposure

SOURCE	WHAT IT GIVES	RESOLUTION · CADENCE	USED FOR
NOAA STAR VHP Vegetation Health Product	Vegetation Condition Index	4 km · weekly	Crop stress; the crop check's track record
Sentinel-2 L2A ESA Copernicus via Element 84, AWS	Reflectance and scene classification	10 m (read at 80–160 m) · 5 days	Crop check's field-scale score; water in tanks
JRC Global Surface Water EC Joint Research Centre, 1984–2021	Where water has come and gone	30 m · static	Tank beds, apart from lakes and aquaculture
ESA WorldCereal 2021	Irrigated versus rainfed cropland	10 m · one season	Isolating rainfed fields
ESA WorldCover 2021	Land cover and cropland	10 m · one year	Cropland mask
ISRIC SoilGrids	Soil texture and water capacity	250 m · static	Soil water holding in the crop check
WorldPop 2020 UN-adjusted	Population	1 km · one year	People living in mandals at risk

Rules rather than data: the national *Manual for Drought Management 2020* (drought triggers), FAO-56 (crop water use), and CGWB's convention of comparing pre-monsoon (May) depths year to year.

This month's depth, with an 80% band

Gradient-boosted trees estimate this month's depth and the depth three months ahead. Retrained before every month, it errs by **0.91 m** on average, 13% better than carrying the last reading forward, and its 80% band held **81.1%** of readings.

$$\hat{d}_{m,t} = d_{m,t-1} + w_b \cdot \Delta \hat{d}_{m,t}$$

It predicts the month-on-month **change** and adds it to the last reading: trees cannot extrapolate depth, but its change is stable. w_b pulls toward the last reading per depth band (0.30 to 0.80).

HOW IT IS BUILT

01

19 inputs

Depth lags of 1 to 12 months, CHIRPS rain over 1, 3 and 12 months, season, location, aquifer type and specific yield (CGWB)

02

Learner

scikit-learn HistGradientBoosting: the median plus the 10th and 90th percentiles under quantile loss

03

Blend

Pulled toward the last reading by a weight set per depth band on a held-out year (2023)

04

Band

10th to 90th percentile range widened per aquifer and season by split-conformal calibration

05

No leakage

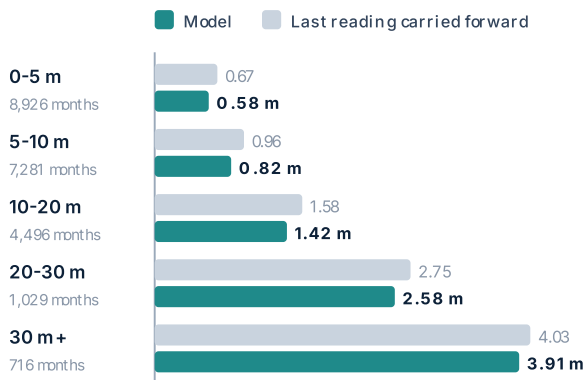
Each mandal's latest reading is held out before its own estimate; inputs use only what was known then

HOW IT IS TESTED

TEST	MONTHS	MODEL	BASELINE	GAIN
Nowcast, held out (trained to Dec 2023), Jan 2024 to Sep 2026	22,448	1.02 m	1.15 m last	11%
Nowcast, retrained every month	16,328	0.91 m	1.05 m last	13%
Forecast 3 months ahead, rolling origins	20,958	1.90 m	2.22 m no change	14%
80% band coverage	22,448	81.1%	80% nominal	±1.6 m
Rise or fall on CGWB's wells, never trained on (2018–23)	1,199	75% right	77% State wells	89 / 38%
IMD rain in place of CHIRPS, retrained every month (2024–25)	16,351	1.04 m	1.06 m CHIRPS	2%, a year late
Ungauged mandal (whole mandal held out)	82,332	4.9 m	state average	not released

Error grows with depth; the model wins in every band

Average absolute error on held-out months



Release rule. A horizon is published only if it beats both no-change and same-month-last-year by at least 5% under rolling-origin validation, with no leakage and consistent results by terrain. Three months passes; one, six and twelve do not yet, so they are not shown.

Tried and rejected, on evidence. The El Niño index made it worse. Rain over official outlines: no gain (1.797 vs 1.800 m). NASA near-real-time rain read twice as wet as CHIRPS and raised error 0.62 → 0.96 m. A month whose satellite rain is unpublished is estimated without it and says so (0.89 → 1.04 m). The State's seven carried-forward months of 2021 are treated as missing.

Six readings of the same 670 mandals

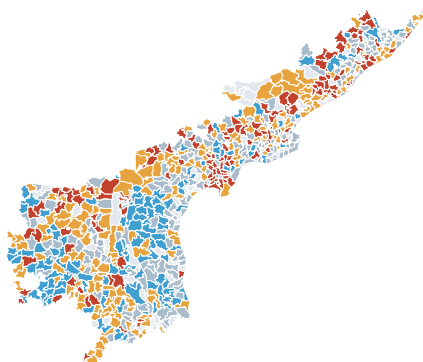
Each map and figure is drawn from this week's published data. Each compares a mandal with its own history, and each method carries a backtest or a rule from an official manual.

WATER DEPTH 3D

September against its usual September

309 mandals deeper than usual; 128 deepest on record

Against the median of the same mandal's earlier Septembers; within 1 m is usual.



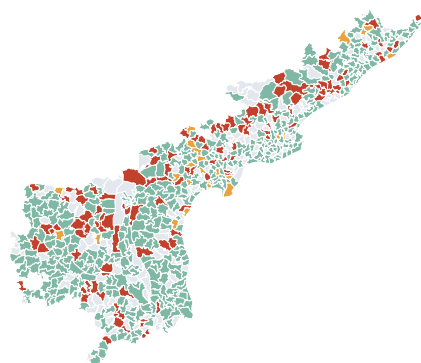
■ Shallower than usual **128** ■ Within 1 m **171**
■ Deeper than usual **181** ■ Deepest on record **128**

MONSOON WATCH

Did this monsoon recharge the ground?

54.3% of 580 mandal series lower than in May

May-to-now against the same mandal's median over ten seasons; flagged at 1 m short and twice its own spread. Rain 21.1% below normal.



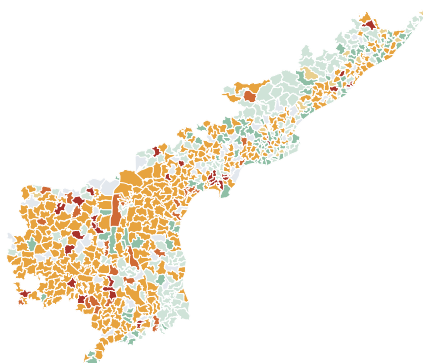
■ Normal **403** ■ Short **23** ■ Severe shortfall **107**

DROUGHT WATCH

What the national drought manual says

30 · 314 severe · moderate, of 615 assessed

Rain deficit and dry spells under the Manual for Drought Management 2020, then impact on vegetation, soil, groundwater and reservoirs. It reads the manual; it does not declare. **479 meet Trigger 1 on department gauges, 546 on CHIRPS SPI.**



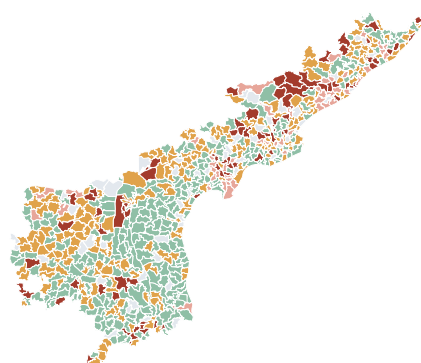
■ Severe **30** ■ Moderate to severe **26** ■ Moderate **314**
■ Normal **90** ■ No trigger **136**

SUMMER OUTLOOK

Where May 2027 may break the record

126 mandals past their deepest May in a typical winter

Latest reading plus the mandal's own past winter drawdowns, against its deepest May. Backtest: past "beyond" calls came true 1 in 3, against 1 in 8 overall.



■ Beyond record, 10 m+ **76** ■ Beyond, under 10 m **50**
■ Only in a dry winter **217** ■ Within its record **275**

DRINKING WATER · JAL JEEVAN MISSION

96% of rural drinking-water schemes draw on groundwater

1,24,219 sources sit in the 126 mandals heading past their deepest May; 11,337 tested above a chemical limit.

TANKS · SENTINEL-2 OVER JRC TANK BEDS

76% of their usual water, 15 September to 15 October

143 of 318 tank mandals emptier than usual; the typical tank bed is 26% wet, the lowest of 8 years.

What the season has done to the crop

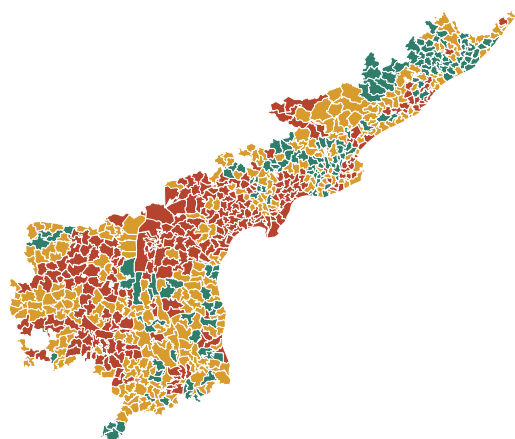
Groundwater matters here because of what grows above it. **437 of 670 mandals** are mostly rainfed, and across the State under a quarter of mapped cropland is irrigated — so for most of Andhra Pradesh the monsoon is the irrigation.

CROP CONDITION · NOAA VHP, WEIGHTED TO CROPLAND

How the crop has read these four weeks

244 mandals severe; 281 moderate

Vegetation Condition Index against each cell's own record, weighted by cropland share, averaged over the four weeks to 14 October 2026. Table 3.4 of the drought manual: 60–100 normal, 40–60 moderate, under 40 severe.



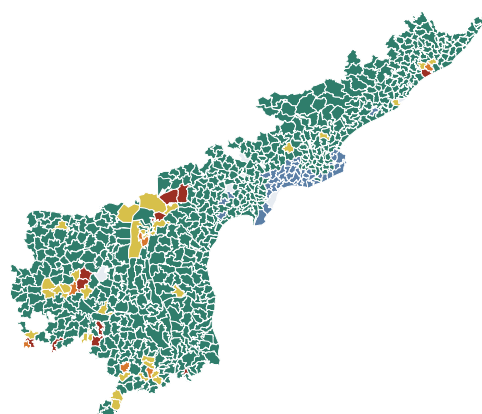
■ Normal 145 ■ Moderate 281 ■ Severe 244

OFFICIAL ASSESSMENT · CGWB AND THE STATE, VIA INGRES

What the annual balance says

577 mandals safe; 11 over-exploited

The 2025–26 Dynamic Ground Water Resources assessment, GEC-2015: a year's recharge against a year's draft. It is not a reading of this week, and the two maps are not in competition.



■ Safe 577 ■ Semi-critical 30 ■ Critical 6
■ Over-exploited 11 ■ Salinity 40

THE SEASON THE CROP IS STANDING IN

-77.7%

north-east monsoon so far:
27 mm against 120.7 mm
normal, 1 October to 17
October 2026

603

of 688 mandals deficient or
scanty on 3,532 department
gauges since June; -39.6%
overall

76%

of their usual water in 4,012
tank beds, read by Sentinel-
2; 143 mandals emptier than
usual, 8 fuller

23.9%

of mapped cropland detected
as irrigated statewide in one
rabi season; the median
mandal 26%

THE CROP WATER CHECK, AND HOW IT IS BEING SCORED

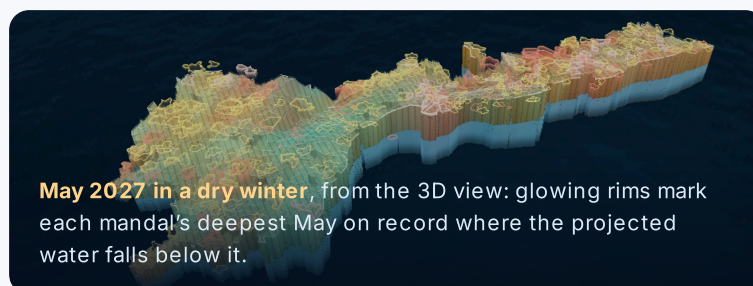
Each week the check compares a mandal's rain and reference evapotranspiration, from the ECMWF open forecast, with what its crop would need. **Every call is frozen the day it is made** and scored three weeks later against the vegetation that followed — Sentinel-2 NDVI over rainfed cropland at 160 m, and the vegetation index. 5 weeks are on the record, **none scored yet**; the first verdict falls due on 31 October 2026.

A negative result, kept on the record. Pooled across mandals the check does not separate vegetation at all: the places called short most often are the chronically dry ones. It is therefore scored within a mandal and season, against that same mandal when it was called comfortable.

What is missing, and it is the heart of it. Sown area, crop and stage, and what the canals actually delivered to a field are all held by the State and not released. Until they are, the check runs on FAO-56 reference coefficients and a stated assumption about the crop — not on the mandal's own sown record. That single release would turn a water platform into an agricultural one.

Where a hydrogeologist's view would help most

The model knows depth well and storage, aquifers and pumping poorly. These are the questions we would most value a review on.



KNOWN LIMITS

- **Aquifer properties are coarse.** Aquifer type is a district proxy; for 259 of 688 series the measured specific yield disagrees with the label.
- **Pumping is unobserved.** Extraction enters only through the official stage-of-extraction figure.
- **Mandal averages.** APWRIMS reports a mandal's mean across piezometers; a village's wells can differ.
- **Data status.** Research sample pending authorization; some series await an official location crosswalk.
- **Depth is not storage.** A metre lost in hard rock holds far less water than in the delta; volumes are under review.
- **670 outlines, 688 mandals.** The State's record lists 688; the public sub-district map predates the latest splits, so 14 newer town mandals share their former town's outline and 5 have none. All are still measured and estimated — only the map is short.
- **Ungauged mandals are hard.** Estimating from neighbours alone errs by 4.9 m, so it is not released.

QUESTIONS FOR REVIEW

- 1 Is there a defensible mandal-level aquifer and specific-yield layer (for example CGWB's NAQUIM mapping) to replace the district proxy?
- 2 Would a physically based recharge and abstraction component (water-table fluctuation, irrigated area from Sentinel-2) beat the statistical model at six to twelve months, where it does not yet pass?
- 3 Which public proxy for pumping is most credible in Andhra Pradesh: agricultural power use, irrigated area, or well census counts?
- 4 Is a 5% gain over persistence and seasonal baselines, under rolling-origin validation, the right bar for releasing a forecast to administrators?
- 5 Can GRACE-scale storage be usefully downscaled to mandals, or should it stay regional context?

SOURCES

Platform prucodes.github.io/ap-groundwater-fusion · [/methodology](#)
Groundwater apwrims.ap.gov.in · ingres.iith.ac.in · indiadataportal.com (CGWB wells) · nasagrace.unl.edu · ejalshakti.gov.in (drinking water)
Rain data.chc.ucsb.edu/products/CHIRPS/v3.0 · imd pune.gov.in (gridded) · power.larc.nasa.gov · climatologylab.org/terraclimate

Climate cpc.ncep.noaa.gov (ONI) · psl.noaa.gov (ERSST v5) · open-meteo.com (ECMWF open data)
Land star.nesdis.noaa.gov (VHP) · earth-search.aws.element84.com (Sentinel-2) · global-surface-water.appspot.com (JRC) · esa-worldcereal.org · esa-worldcover.org · soilgrids.org · worldpop.org
Methods Manual for Drought Management 2020 · FAO-56 (Allen et al., 1998) · scikit-learn

Six inputs, and the number each one moves

In an ordinary season the forecast is already at the practical ceiling for any method. Almost all of the headroom that remains sits in seasons that break the pattern — when the water table is set by what is pumped rather than by what fell, and **pumping is the one driver the model cannot see**.

RANKED BY THE ERROR EACH WOULD REDUCE

- 1 **Station-level DWLR history, with coordinates.** An ungauged mandal is interpolated between mandal means today and errs by 4.9 m, five times the nowcast. Well locations would let it run between wells inside aquifer structure, and let an estimate be made below mandal scale.
- 2 **The NAQUIM aquifer framework — boundaries, thickness, specific yield, transmissivity.** Aquifer type is a district proxy, contradicted by the measured specific yield in 259 of 688 series. Hard rock is the worst terrain in both tests: 1.40 m against 0.78 m on the coast for the nowcast, 5.7 m against 3.8 m when a mandal is ungauged. Real aquifer properties would also make depth-to-storage defensible.
- 3 **Agricultural feeder supply hours, and the pump-set census.** Extraction reaches the model only through the official stage-of-extraction figure, which is annual. Feeder supply hours against connected load would give a weekly, mandal-scale proxy for what is actually drawn — the one input aimed squarely at the regime where the forecast fails.
- 4 **DWLR telemetry, not only the history.** Readings are monthly, so even a perfect estimate is a month behind. Telemetry would make instrumented mandals measured and current rather than modelled, and turn a backtest into continuous validation.
- 5 **An authorized APWRIMS export, full history.** Training begins June 2015; a longer authorized record is accuracy at no modelling cost. It is also the gate on any official use, since the present series is a research sample pending authorization.
- 6 **Field verification on twenty to thirty mandals.** Every figure in this brief is a backtest. Measuring a small sample after an estimate has been published, then comparing, is the one step that turns validation into verification. There is no field ground truth today.

Which of these reaches the ceiling. In an ordinary season the forecast calls direction right 89%, ahead of the 87% that the State's own wells and CGWB's agree with each other — the practical ceiling for any method. All the headroom left is in the pattern-breaking season, where it falls to 38% against the measured wells' 49%. **The first and third — well locations and a pumping proxy — are aimed at that regime.**

WHAT THIS PROJECT CAN GIVE BACK

- **A systematic gap between the two well networks.** Over 16,097 readings paired within 20 km and the same month, CGWB's wells read 4.3 m shallower than the State's, correlating 0.41.
- **259 series whose aquifer label its own specific yield contradicts,** already listed mandal by mandal and ready to hand over.
- **Quality control on the State's own record.** The same model that estimates a level also flags a reading inconsistent with that well's history and its neighbours — which is how a drifting recorder or a transcription error is found across a network of this size.

What that would build. With the first three in place, this same weekly machine — 21 sources, 670 mandals, every Monday — would estimate a level where no well stands, inside real aquifers rather than a district proxy. The pipeline already runs; what it lacks is the department's measurement, not the method.