



# Boussinesq Center for Hydrology

$$u_{tt} + \alpha u_{xx} + \beta(u^2)_{xx} + \gamma u_{xxxx} = 0$$



**Daily root-zone soil moisture (0–100 cm) ~~and evapotranspiration~~  
at **20 m (?)** with multimodal drivers and  
geospatial foundation model embeddings**

**Prajwal Khanal**

PhD candidate

## **Supervisors**

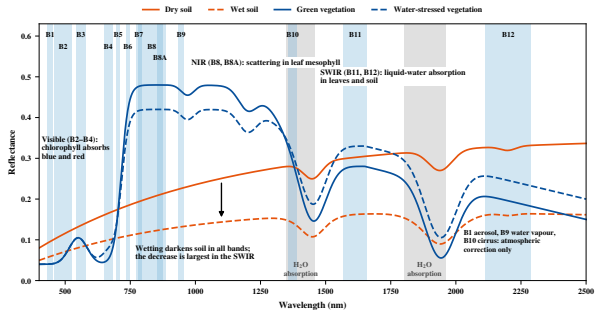
|                        |                                   |
|------------------------|-----------------------------------|
| prof.dr. Z. Su (Bob)   | Faculty ITC, University of Twente |
| dr. Y. Zeng (Yijian)   | Faculty ITC, University of Twente |
| dr. J. Beckers (Joost) | Planet Labs                       |

8 October 2026

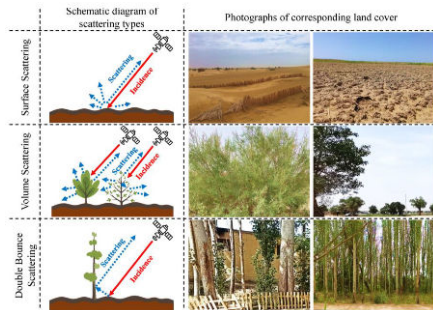
## Motivation Different ways of obtaining soil moisture

| Source                                           | How SM is obtained                                                                                                      | Depth                                                                 | Main limitation                                                                                                         |
|--------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| <b>Microwave</b><br>(SMAP, SMOS, ASCAT, S1)      | invert:<br>$\sigma^0$ or $T_B$ =<br>$f(\text{SM, roughness, vegetation})$                                               | top $\sim 5$ cm; informs part of the root zone (Feldman et al., 2023) | shallow; assumes vegetation and roughness; coarse (9–40 km) except S1                                                   |
| <b>Optical / thermal</b><br>(S2, Landsat, MODIS) | empirical relations with surface properties: reflectance and indices; LST-vegetation space (e.g. triangle method, TVDI) | skin / surface                                                        | indirect, site-specific calibration; clouds; canopy hides the soil                                                      |
| <b>Land-surface models</b>                       | $d\theta/dt = P - ET - Q - D$ , with parameters $p$                                                                     | full profile                                                          | forcing errors, soil and vegetation parameters, process representation (Koster et al., 2009; Seneviratne et al., 2010)  |
| <b>Reanalysis</b><br>(ERA5-Land)                 | LSM driven by reanalysis meteorology                                                                                    | full profile                                                          | model assumptions dominate deep layers with few observational constraints (Li et al., 2021; Miguez-Macho and Fan, 2021) |
| <b>In situ</b> (ISMN)                            | direct probe measurement                                                                                                | several depths                                                        | point scale; sparse, uneven global coverage                                                                             |
| <b>Data-driven</b><br>(ML / DL)                  | $\text{SM} = f_\theta(\text{inputs})$ , learned from in-situ labels; no ad-hoc process assumptions                      | any depth with labels                                                 | needs enough, representative sites; generalisation?; hard to interpret?                                                 |

# Motivation Multimodal remote sensing



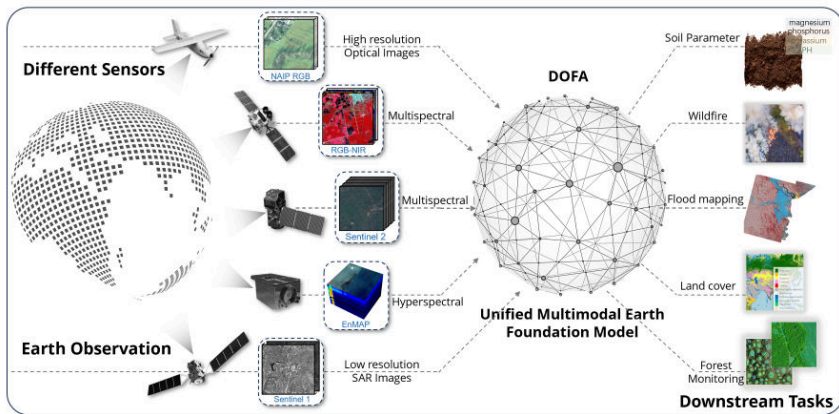
Optical: dry vs wet soil; green vs water-stressed vegetation



SAR scattering types (Nurmemet et al., 2025)

**Soil and vegetation water change both optical reflectance and SAR backscatter**

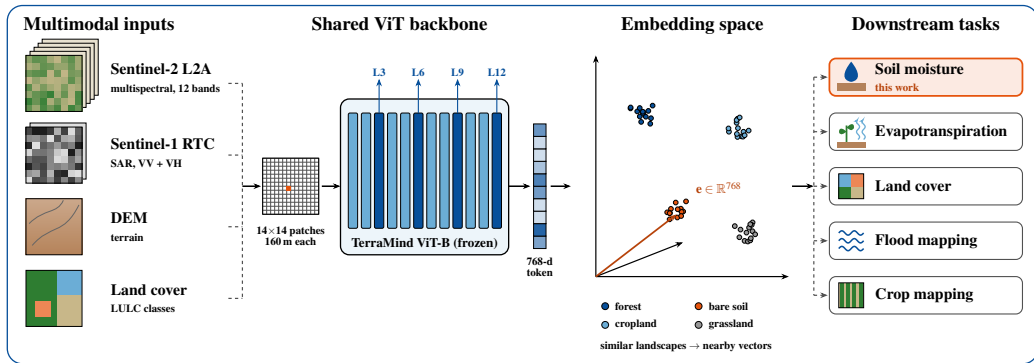
## Motivation Geospatial foundation models (GFMs)



*DOFA: one model for many sensors and many downstream tasks (Xiong et al., 2024, arXiv:2403.15356)*

**Once pre-trained on multimodal, unlabelled data, it works for many downstream tasks**

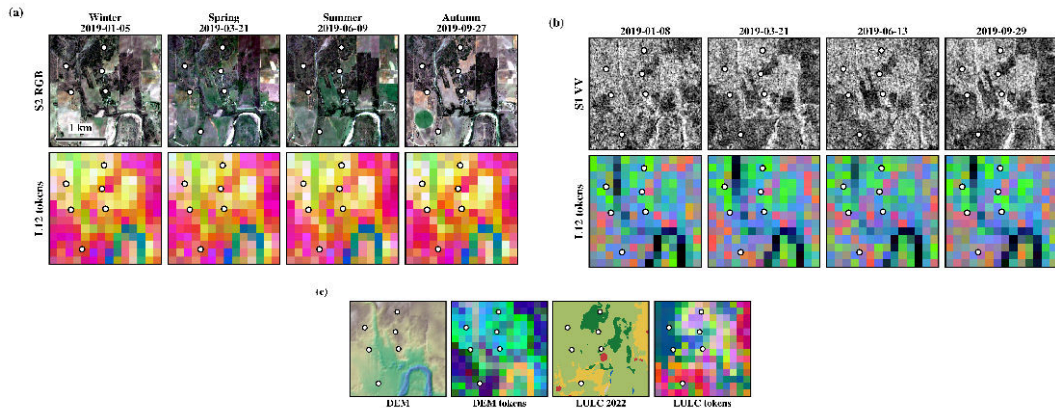
## Motivation TerraMind embeddings for soil moisture



*S2, S1, DEM and LULC → frozen TerraMind ViT-B → one 768-d embedding per 160m patch → downstream tasks*

**The same embedding can serve many tasks; in this case, soil moisture**

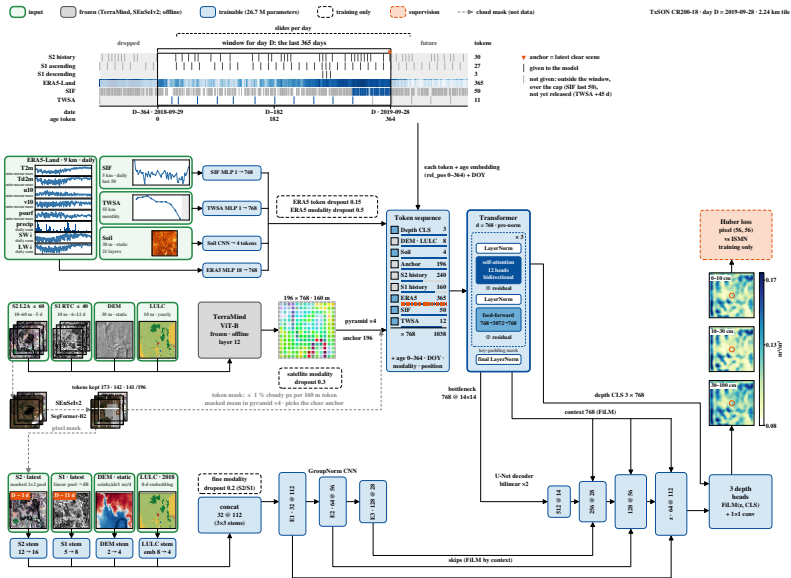
## Motivation Example: TerraMind embeddings at one site (TxSON, 2019)



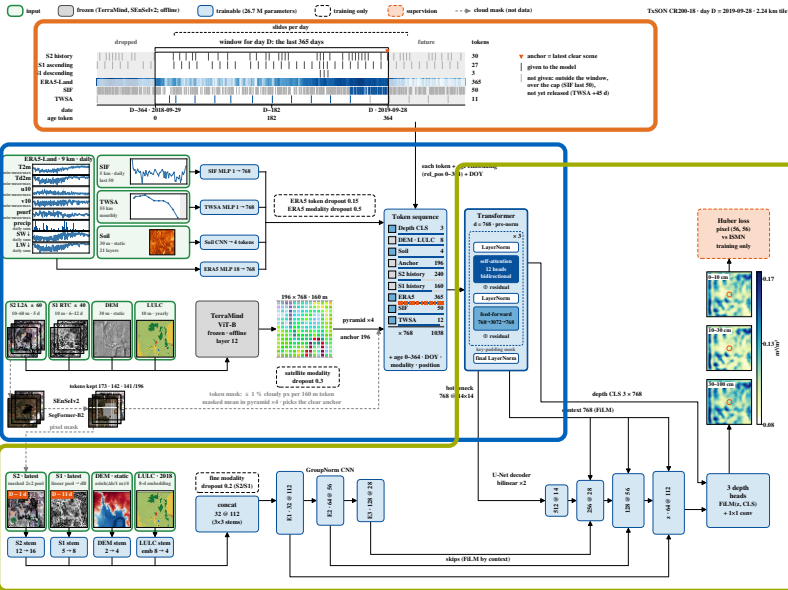
(a) S2 in four seasons, (b) S1 VV, (c) DEM and LULC, each with its TerraMind L12 tokens

Each 160 m patch  $\rightarrow$  one 768-d vector; colours show its first 3 principal components

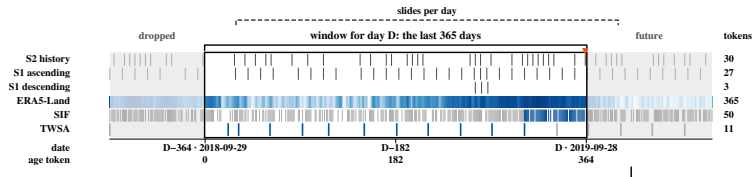
# Method Model architecture



# Method Model architecture

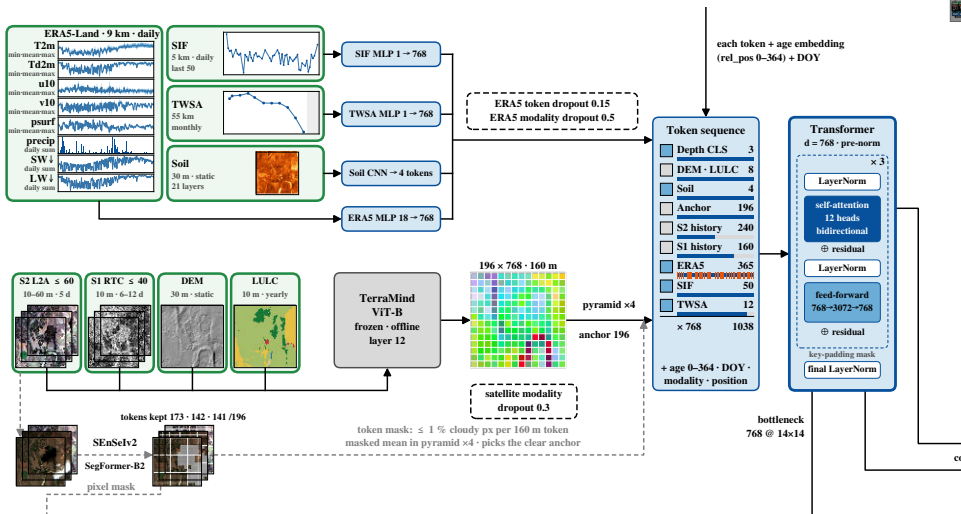


# Method Inputs: what the model sees for day D

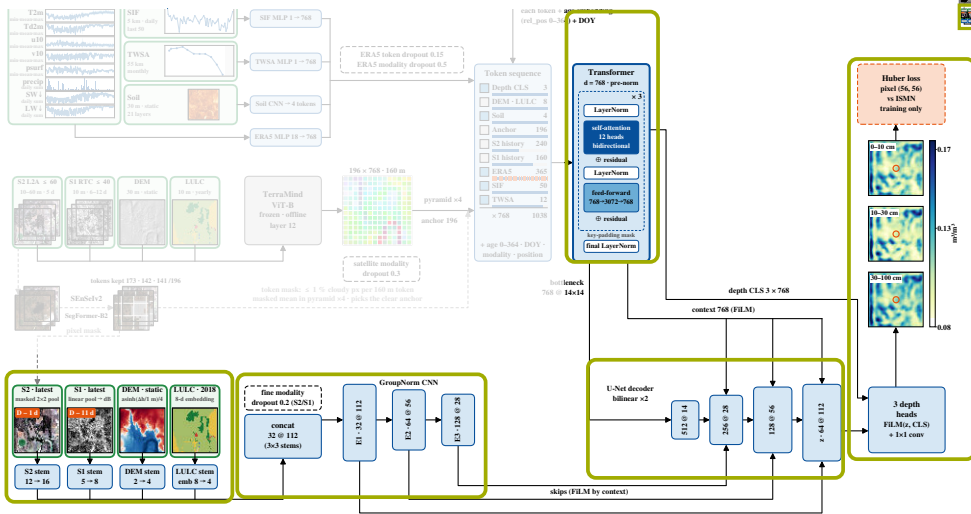


- ▼ anchor = latest clear scene
- | given to the model
- | not given: outside the window, over the cap (SIF last 50), not yet released (TWSA +45 d)

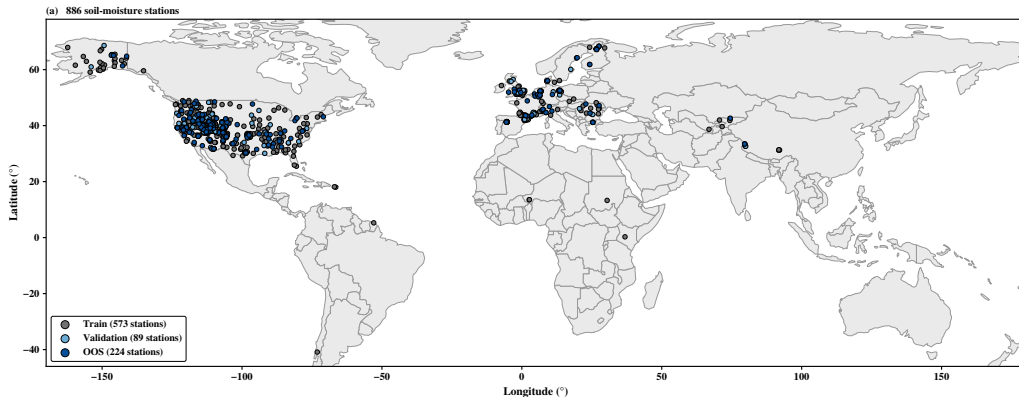
# Method Coarse path: tokens $\rightarrow$ transformer



# Method    Fine path: 20 m CNN → U-Net decoder → depth heads



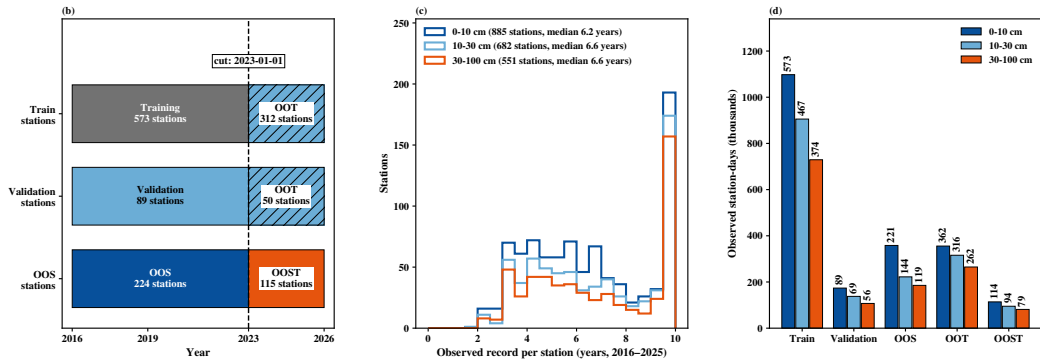
## Data In-situ soil moisture: global coverage



*886 soil-moisture stations (ISMN and flux sites): train 573, validation 89, out-of-sample (OOS) 224*

**Dense in North America and Europe; sparse in the tropics and the Southern Hemisphere**

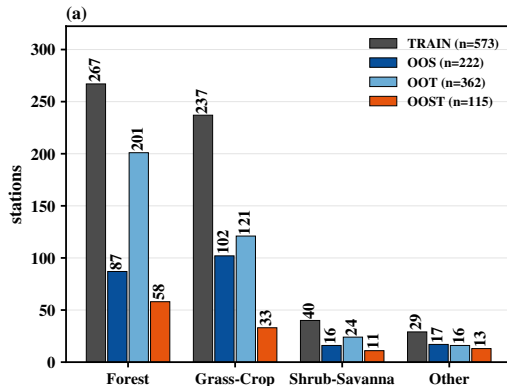
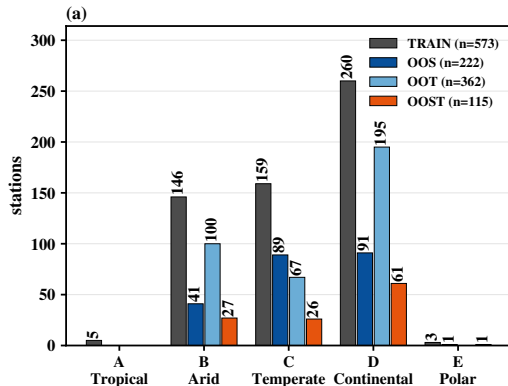
## Data Depth coverage, record length and splits



(b) train / validation / OOS stations, time cut 2023-01-01; (c) observed record per station; (d) observed station-days

**0–10 cm: 885 stations (median 6.2 yr); 10–30 cm: 682 (6.6 yr); 30–100 cm: 551 (6.6 yr)**

## Data Climate and land-cover coverage



*Stations per Köppen climate zone (left) and land-cover class (right), by split*

**Mostly continental, temperate and arid; forest and grass-crop; almost no tropics or polar**

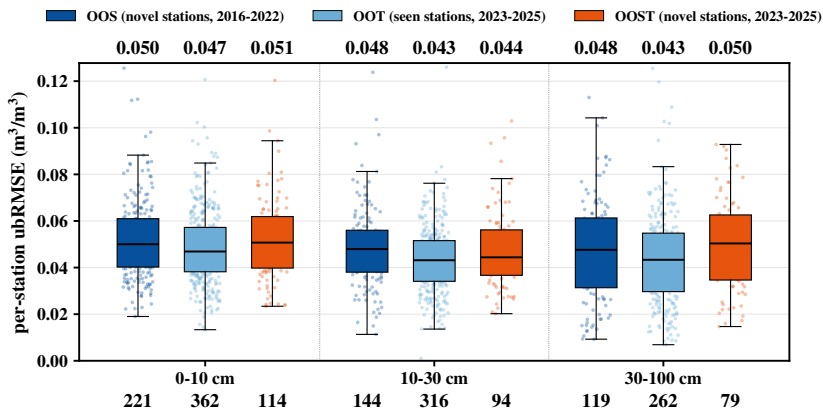
---

|                                   |                                                                                |
|-----------------------------------|--------------------------------------------------------------------------------|
| <b>Stations</b>                   | 990 in inventory; <b>886</b> with soil moisture (train 573 / val 89 / OOS 224) |
| <b>Depths</b>                     | 0–10 cm: 885   10–30 cm: 682   30–100 cm: 551 stations                         |
| <b>Period</b>                     | 2016–2025; out-of-time test from 2023-01-01                                    |
| <b>Soil-moisture observations</b> | <b>5.14 M</b> observed daily values (2.10 M / 1.68 M / 1.37 M per depth)       |
| <b>Sentinel-2 L2A</b>             | 272,281 scenes (994 sites); cloud < 75 % per scene, cloudy tokens masked       |
| <b>Sentinel-1 RTC</b>             | 496,835 scenes: 251,310 ascending + 245,525 descending                         |
| <b>ERA5-Land</b>                  | 18 variables, 2.46 M station-days (990 stations)                               |
| <b>Static / other</b>             | DEM, LULC, soil (OpenLandMap), SIF, TWSA                                       |

---

**About 5.1 million daily soil-moisture observations at 886 stations, 0–100 cm**

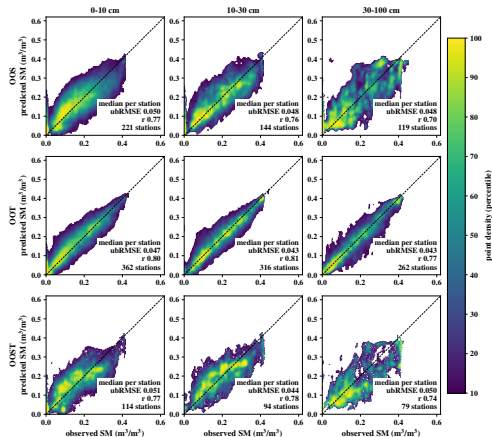
## Results Per-station ubRMSE by depth



*One point per station; medians above, station counts below*

**Median per-station ubRMSE  $0.043\text{--}0.051 \text{ m}^3/\text{m}^3$  at every depth and split**

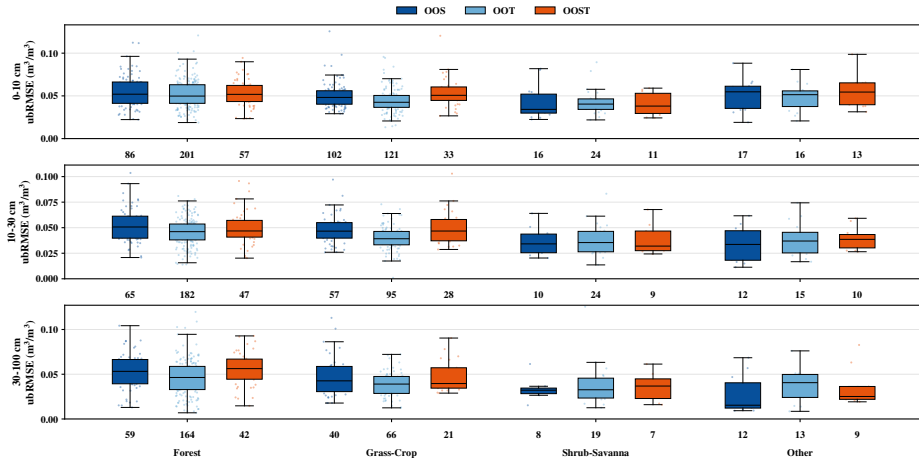
## Results Predicted vs observed, by split and depth



Daily samples coloured by point density (sparsest 10 % not drawn); metrics = median over stations; rows OOS, OOT, OOST

**New years at known sites fit far better than new sites**

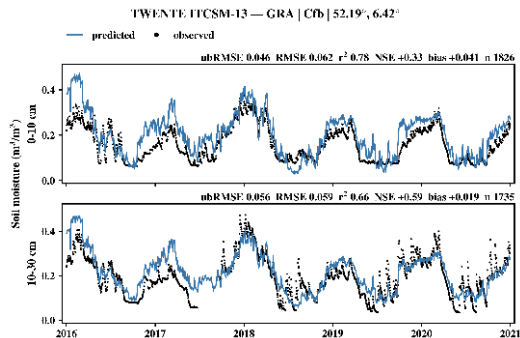
## Results Per-station ubRMSE by land cover



Rows: depths; groups: land-cover class; colours: OOS, OOT, OOST

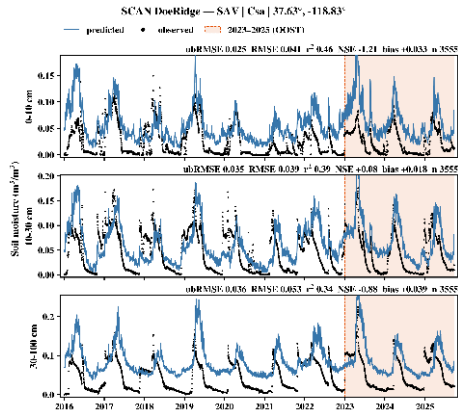
**Similar error across land cover; forest slightly highest, shrub-savanna lowest**

## Results Sites where the time series match



TWENTE ITCSM-13, Netherlands

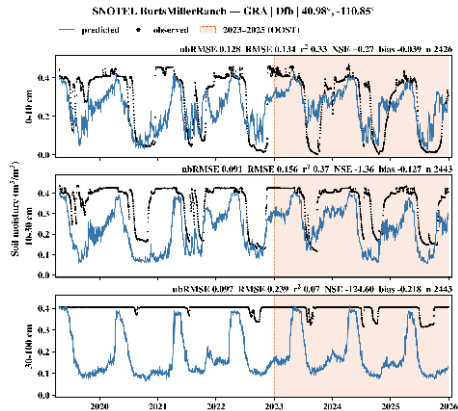
*OOS stations; blue predicted, black observed; Twente has two depths, Doe Ridge three*



SCAN Doe Ridge, California

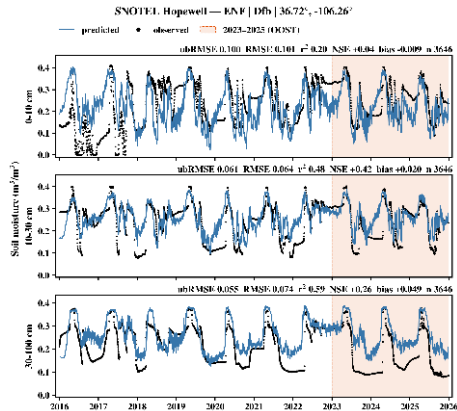
**Dynamics right at every depth ( $r$  0.57–0.88, ubRMSE 0.024–0.056); level off by up to 0.04**

## Results Sites where the time series have some issues



SNOTEL Burts Miller Ranch, Utah

*OOS stations observed at all three depths, highest ubRMSE; shaded = 2023–2025 (OOST)*

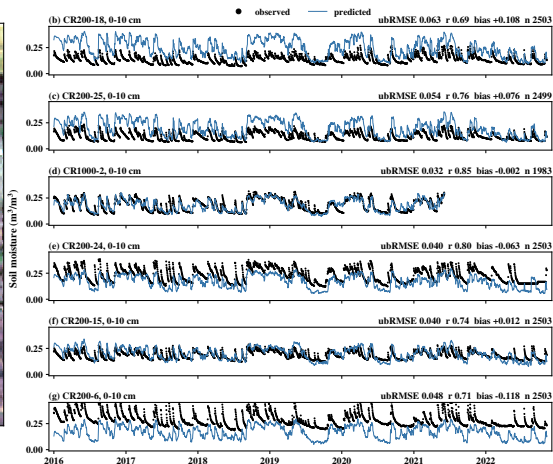


SNOTEL Hopewell, New Mexico

**Deep layer (30–100 cm) wrong in summer: too dry at Burts Miller, too wet at Hopewell**

## Results TxSON: six stations inside one 2.2 km tile

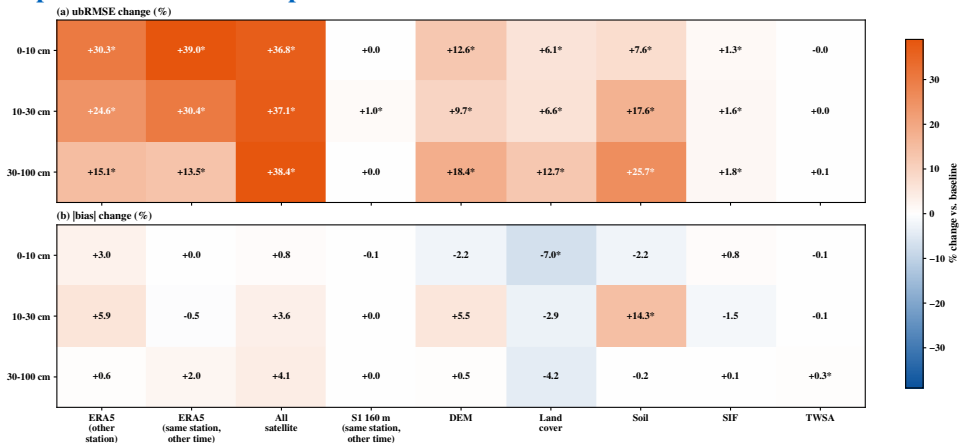
(a) Sentinel-2 true colour, 2018-07-14



(a) Sentinel-2 true colour with the six stations; (b–g) 0–10 cm, predicted vs observed

**Timing right at every station (r 0.69–0.85); offsets remain (bias –0.12 to +0.11)**

## Results Input ablation: which inputs does the model use?



One input shuffled at a time; 221 OOS stations; \* = significant

**ERA5 matters most at the surface (+30–39 % vs +14–15 % deep); satellite equally at all depths (+37–38 %); DEM, land cover and soil grow with depth**

## Conclusions    Answers to the research questions

### RQ1 Do frozen GFM embeddings add skill beyond a multimodal driver baseline?

Ans. **GFM embeddings carry a site-specific signal, comparable to climate.** Swapping satellite tokens between stations raises ubRMSE by 37–38 % at every depth (ERA5 from another station: +15–30 %); reordering S1 in time at the same station costs  $\approx 0$  %. Added skill still needs a drivers-only run.

### RQ2 How does the model generalise to unseen sites?

Ans. **New sites are slightly worse than seen stations.** Median per-station ubRMSE 0.048–0.050 at new sites vs 0.043–0.047 at seen stations.

### RQ3 How does skill vary within the same tile?

Ans. **Timing is captured, level may not be.** At TxSON,  $r$  is 0.69–0.85 and ubRMSE 0.032–0.063 at all six stations, while bias runs from  $-0.12$  to  $+0.11$  inside one 2.2 km tile.

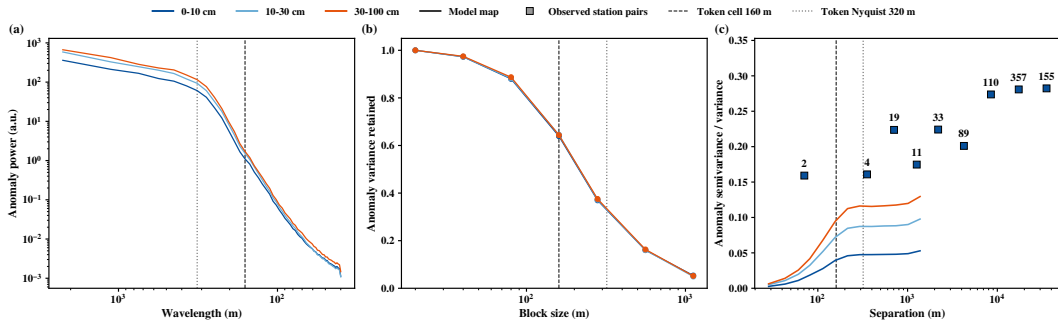
### RQ4 Which modalities matter, at which depth?

Ans. **Depth changes what the model relies on.** ERA5 dominates the surface; satellite matters equally at all depths; soil, DEM and land cover gain weight at depth.

## What I could not present

1. Effective resolution of the map
2. Where the station bias comes from
3. Voids in the density plot
4. Distribution of soil moisture

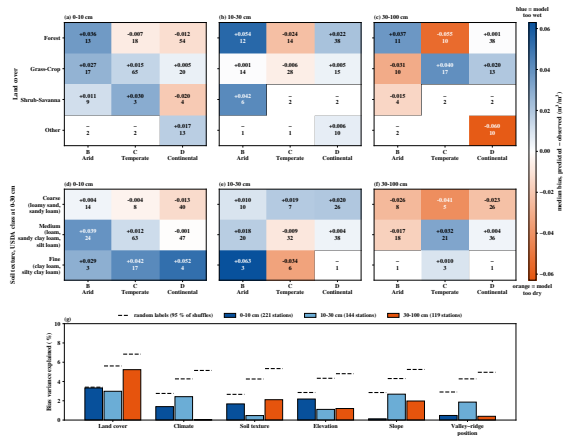
## Other analysis 1. Effective resolution of the map (TxSON)



30 OOS TxSON tiles. (a) Spectrum of map anomalies; (b) variance kept after block averaging; (c) map variogram vs observed station pairs (numbers = pairs)

**Map has structure down to 250 m**

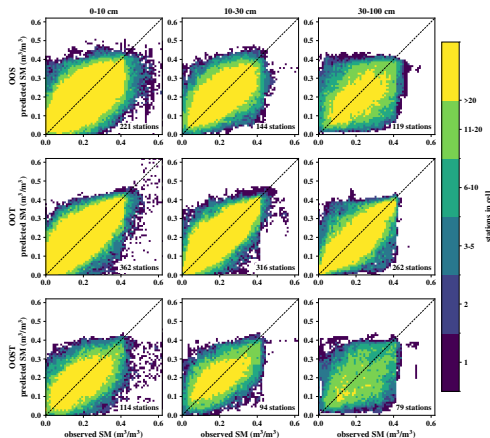
## Other analysis 2. Where the station bias comes from



Median OOS bias by land cover, soil texture and climate; (g) bias variance explained vs random labels

**Bias is not explained by land cover, climate, soil texture or terrain ( $\leq 5\%$ , none above random)**

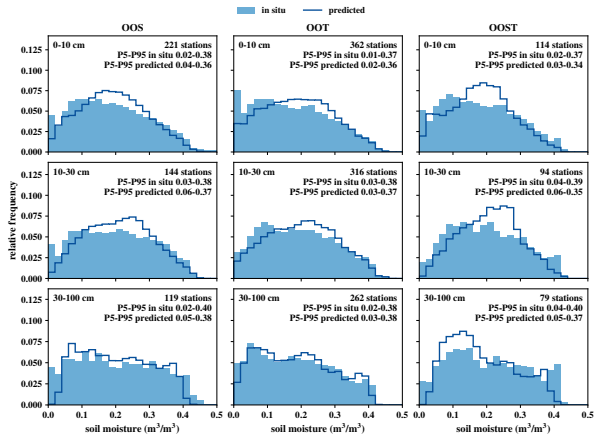
## Other analysis 3. Voids in the density plot



*Same grid as the density plot; colour = number of distinct stations with samples in each cell*

**Voids are not empty: many stations pass through briefly, a few long-record stations make the dense blobs**

## Other analysis 4. Distribution of soil moisture



All daily samples; filled = in situ, line = predicted; box: P5-P95 range of each

**Distribution matches; the model narrows the dry end**