PAPER CHECK · EARTHQUAKES

A 92.70% score for finding labelled magnetic anomalies is no test of an earthquake link

Shiji Wen and three colleagues report a deep network that finds geomagnetic anomalies in Swarm-A satellite data with an F1 score of 92.70%, then apply it to “three representative earthquakes in 2024”, in an abstract that hedges with care [1 Abstract]. We could read its Highlights, Abstract, Keywords and §1, not §§2–6, so those earthquakes stay unnamed here; but the authors published their dataset under CC BY, which made this check possible, and we read all of it [2]. It holds passes from 14 dates across 21 days of August 2023, its “stable” training set spans a Kp 6+ storm, and its labels mark short magnetic bursts with no detectable tie to coming earthquakes on our illustrative settings, by our analysis. So the score says how well a detector matches those labels, not whether anomalies are seismic, and how it was scored, point adjustment above all, decides whether it beats a one-line threshold; we can’t read that.

By Fact-Checked Science · Published November 2, 2026 · Updated November 6, 2026

OUR ANALYSIS marks our own reasoning and calculations: the dating of the authors’ dataset against ESA’s, NASA’s, GFZ’s and OMNI’s public records; every count, score and baseline on its rows; the comparison with USGS’s catalogue; the base rate and the size estimate; and anything we infer about how the data were built. The verdicts are our assessment. Everything else comes from the numbered sources. For the paper [1], read only in its public view, Highlights, Abstract and §1 are those parts of it; §§2–6 are unread. Elsewhere p. gives a page, § a section, and Abstract or Methods the part of a source we rely on.

An open, exactly dated dataset, a score whose meaning depends on scoring we can’t read, and no detectable earthquake tie in the labels, on our illustrative settings

OUR ASSESSMENT

150 of 151

consecutive half-orbits of Swarm A, August 1–5, 2023, make up the training set the paper calls stable and randomly selected; they span a storm that reached Kp 6+, by our dating of the authors’ file [2][1 §1]

57% or 95%

what one threshold on one magnetic channel scores against the test labels, with no learning, at its best threshold, which uses the labels: 57.3% plainly, 94.9% with point adjustment (our analysis) [2]. The authors report 92.70%, and we can’t read how they scored it [1 Abstract]

5.4% and 4.8%

of labelled and of unlabelled test points lie in the zone of a shallow M ≥ 5.5 earthquake due within 90 days, on illustrative settings; the 95% limits on the difference, −1.5 to +3.0 points, span zero (our analysis) [2][36]

THE PAPER’S CLAIMS

MISLEADING

“A total of 150 samples with stable data were randomly selected to form the training set”

The paper’s §1 [1 §1]

A reader takes from this a random sample of quiet data, and that isn’t what the authors’ file holds: 150 of 151 consecutive half-orbits over five days, spanning a storm that reached Kp 6+, and more disturbed than the test set by Kp, Dst, residual size and burst rate (our analysis) [2]. “Stable” has two readings. If §2, unread, defines it as instrumentally nominal, the word can mostly stand: the vector flag is 0 on every training row, though 1.8% of rows carry a non-nominal scalar flag, against 0.9% in the test set [2][31]. If it means magnetically quiet, it doesn’t fit. “Randomly” stands against the file on either reading. See section 4.

UNVERIFIED

“achieving an F1 score of 92.70% for geomagnetic anomaly detection”

The paper’s abstract [1 Abstract]

The result would be in §4, with its scoring if stated. We can’t read that, and we found no code to re-run. How it was scored matters, point adjustment above all. With point adjustment, one threshold on one channel reaches 94.9% at its best threshold and 66–95% with no labels used, for anomaly ratios of 0.5–10%; random numbers reach about 74% at best. Scored plainly, the simple thresholds we tried top out at 57–64%, so 92.70% would be a large gain. We don’t know which applies (our analysis) [2]. Either way, it scores agreement with labels. See section 6.

UNVERIFIED

“Localized geomagnetic anomaly clustering is observed in selected earthquake cases.”

The paper’s Highlights [1 Highlights]; its abstract speaks of “three representative earthquakes in 2024” [1 Abstract]

The cases, their time windows and any controls would be in §5, which we can’t read, so we neither name the earthquakes nor guess them. “Selected” is the authors’ own word. A few selected cases couldn’t show a link by themselves: on illustrative settings, a given place lies in the zone of an M ≥ 5.5 earthquake due within 90 days on about 8% of days across 50°S–50°N, and far more often in the most active regions (our analysis) [36]. See section 7.

MISREAD

“Their study confirmed the correlation between the changes in these physical parameters and the earthquakes”

The paper’s §1, on Balasis and Mandea’s CHAMP study, which it says analysed “pre-seismic signals” [1 §1]

The study’s abstract says “some specific features are observed after the two earthquakes”, favours “an external source origin” for the 30 s pulsation, says a feature more than an hour later “could be associated with each earthquake”, and ends on “Further investigations are required” to answer whether the signals can be associated with earthquakes [4 Abstract]. Two conference abstracts by the same authors agree [5][6]. So: signals after, not before, and a question left open. The study’s full text is closed, so this rests on those abstracts. See section 8.

OVERSTATED

“the geomagnetic field is regarded as one of the most promising parameters for obtaining precursor information about impending earthquakes”

The paper’s §1, citing Hayakawa (2001) [1 §1]

Hayakawa calls electromagnetic phenomena as a whole “a very promising candidate” [8 §1], and the review’s “most exciting finding” is in radio propagation, not the magnetic field [8 Abstract]. That is one proponent’s review from 25 years ago, narrowed to the field. The emphasis has a basis: the review of record calls the ULF magnetic anomalies before Loma Prieta and Guam “The most convincing EM precursors”. But it goes on to say they “were likely to have been caused by solar-terrestrial interaction or sensor-system malfunction” [9 p. 335], and §1 leaves that out. See section 8.

1 · WHAT WE COULD READ

The paper’s opening, not its results, and, thanks to the authors, all of its data

Shiji Wen and Lingjia Gu of Jilin University, with Xiaofeng Li and Tao Jiang of the Northeast Institute of Geography and Agroecology [2], wrote this paper for *Advances in Space Research*, COSPAR’s peer-reviewed journal. It is an article in press, dated October 2026, and Jilin University funded it [1][37]. S0 News linked it on November 2, 2026 [3].

What we could read: the public view

The paper is closed, and its publisher reserves all rights [37]. We read its Highlights, Abstract, Keywords and §1, the introduction, which the site’s owner pasted from ScienceDirect’s public view [1]. We could not read §§2–6: the data, the method, the results, the earthquake analysis and the conclusions. We found no preprint or repository copy of the text, and no code.

The abstract says the method was applied to “three representative earthquakes in 2024” [1 Abstract]. That work is in §5, so this page neither names those earthquakes nor guesses them, and it rates only sentences we could read. Where a question needs the full text, the page says so, and section 11 lists them. If you have the article, please tell us through the correction form, and we will revise the page.

The authors published their dataset

SwarmA-MagSet is on Zenodo under CC BY 4.0, posted by the four authors on December 9, 2025, as “the dataset established in the submitted manuscript” [2]. That is what made this check possible, and the credit for it is theirs. The site’s owner downloaded the zip from Zenodo, and we read all 393,250 rows as data.

OUR ANALYSIS

What describing a file can and can’t do

We wrote our rules for describing the file before analysing it, and marked every later addition (section 10). This describes a public dataset. We don’t know how the authors built or labelled it, what selection came first, or how the paper used it beyond what §1 says. So nothing below rates what they did in §§2–6.

2 · WHAT THE PAPER SAYS

A deep network that flags anomalies point by point, then a look at three earthquakes

The paper’s question is practical: can a deep network find geomagnetic anomalies in Swarm-A satellite data, point by point, better than other time-series methods, and what does it find around three earthquakes [1 Abstract][1 §1]? Its abstract is careful about the second part.

What the abstract and §1 report

A dataset. SwarmA-MagSet has “150 training samples and 100 test samples”, each of 1,573 rows with “9 features” [1 Abstract]. The file has 15 columns, and which nine the paper uses is unread [2].

A method. The Period-aware Structural Temporal Feature Fusion Network, PaSTFF Net, reconstructs the data and picks out anomalies from “the residual data between reconstructed data and input data” [1 §1].

A score. On the test set it “outperforms other advanced time-series anomaly detection methods by 9.73% on average, achieving an F1 score of 92.70% for geomagnetic anomaly detection” [1 Abstract].

An application. It was then applied to “three representative earthquakes in 2024”, before and after each [1 Abstract][1 §1].

A purpose. §1 says a point-by-point model would “provide reliability for seismic precursor analysis”, and that the anomalies it finds are then used to investigate links between the geomagnetic field and earthquakes [1 §1].

The hedges, as written A fair reading starts here. §1 is firmer in places; section 8 looks at that.

> *“may exhibit abnormal patterns”*
>
> Of geophysical parameters before an earthquake [1 Abstract]

> *“investigated as a possible source of information on earthquake-related variations”*
>
> Of the geomagnetic field [1 Abstract]

> *“to further investigate geomagnetic variations prior to seismic events”*
>
> Of the three earthquakes [1 Abstract]

3 · A 21-DAY SPAN IN AUGUST 2023

The dataset spans 21 days of one satellite, about 109 hours of data, in one season and two local-time sectors

§1 says the data are “observations collected in 2023” [1 §1]. That is true, and narrower than it sounds. The file carries no dates, but each row carries a position and the Kp and Dst of its moment, and those fix the time to the second. Each sample is one half-orbit between 50°S and 50°N: 1,573 rows, one a second, about 26 minutes. The 250 samples come from 14 dates across the span, with nothing on August 13–19 [2].

POSITION OUR ANALYSIS

393,250 rows

All match ESA’s Swarm A product at the assigned second: worst difference 2×10^{-7} degrees, and 1 cm in radius [30]. In NASA’s orbit data the first row’s median miss is 0.005° for Swarm A, against 0.22° for C and 0.12° for B [35].

DST OUR ANALYSIS

$< 10^{-13}$ nT

The largest difference from OMNI’s hourly Dst in 2,000 checks, once each hourly value is placed at mid-hour and joined by straight lines, as the file does [33][34].

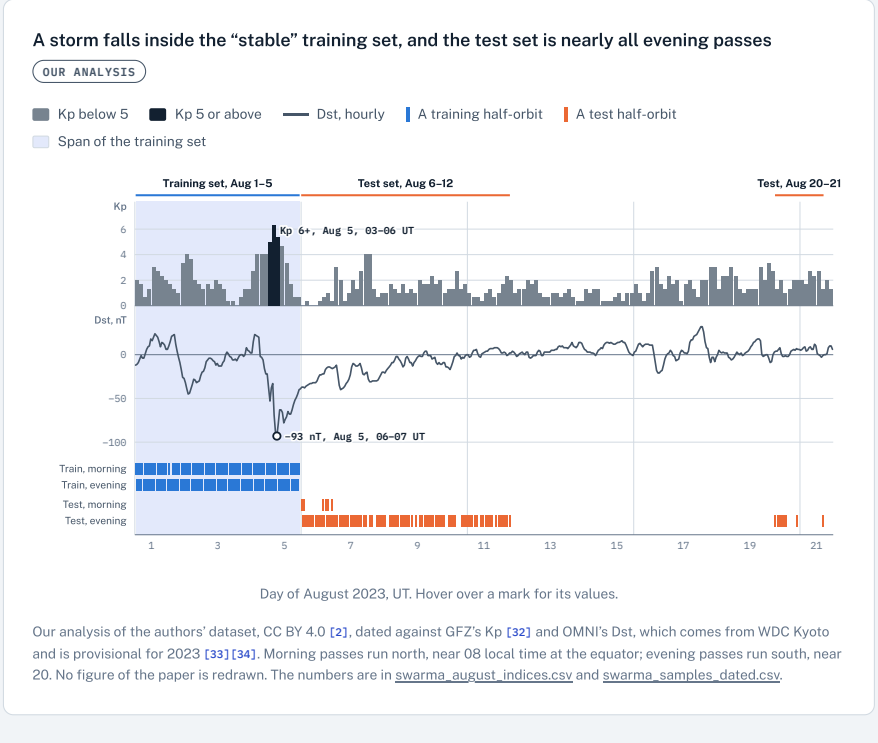
KP OUR ANALYSIS

0 of 2,000

Mismatches with GFZ’s Kp, with each three-hour value running from the middle of its block to the middle of the next, as the file does [32].

A storm falls inside the “stable” training set, and the test set is nearly all evening passes

OUR ANALYSIS



Day of August 2023, UT. Hover over a mark for its values.

Our analysis of the authors’ dataset, CC BY 4.0 [2], dated against GFZ’s Kp [32] and OMNI’s Dst, which comes from WDC Kyoto and is provisional for 2023 [33][34]. Morning passes run north, near 08 local time at the equator; evening passes run south, near 20. No figure of the paper is redrawn. The numbers are in swarma_august_indices.csv and swarma_samples_dated.csv.

Which passes each set holds

SET	FIRST ROWS, UT	HALF-ORBITS	LOCAL TIME AT THE EQUATOR
Training	August 1, 00:33, to August 5, 22:28	150 of 151 consecutive: 75 morning, 75 evening	08:08–08:35 and 20:09–20:35
Test	August 6, 00:02, to August 12, 06:11	92 of 192: 86 evening, 6 morning	19:35–20:08 and 08:03–08:08
Test	August 20, 05:59 to 21:42, and August 21, 16:34	8, all evening	18:44–18:52

Our analysis [2]. The last two test samples are out of time order in the file.

OUR ANALYSIS

Train and test: a clean split in time, and a shift

No pass is shared, no sample is duplicated, and the test set comes after the training set. That is good practice, since leakage between the two has inflated results across many fields [22 Abstract]. The split is adjacent, with the first test pass one orbit (1.57 h) after the last training pass, so the storm’s recovery straddles it.

The sets also differ. A per-pass AUC, where 0.5 means no separation and 1 full separation, tells them apart at 0.91 on mean radius, which drifts with time, 0.71 on highest Kp, and 0.75 on the rms of diff_res_BX. On August 6–12 the test set took 86 of the roughly 96 evening passes and 6 of the roughly 96 morning ones, so in effect it is evening passes (inference).

OUR ANALYSIS

The three earthquakes lie outside this file

The whole dataset comes from August 2023, and the paper’s earthquakes from 2024. Swarm’s orbit “drifted about 1 h in local time (LT), every 10 days” [15], so a network trained on these passes has seen only about 08 and 20 local time, in one season, and meets other hours and seasons in 2024. Whether it was retrained for them, and how it treats hours it hasn’t seen, would be in the closed sections, if stated. We draw no conclusion from that, except that it is a question for §5.

4 · “STABLE” AND “SIGNIFICANT”

By most measures the “stable” set is more disturbed than the one with “significant anomalies”

§1 describes the two sets this way: “A total of 150 samples with stable data were randomly selected to form the training set, while 100 samples with significant anomalies were labeled to establish the test set” [1 §1]. Here is the authors’ file beside those words (our analysis) [2].

The paper’s words beside the authors’ file

MEASURE	TRAINING SET, “STABLE DATA”	TEST SET, “SIGNIFICANT ANOMALIES”
Highest Kp	6.33 (6+)	4.00
Lowest Dst in the file	−92.6 nT	−39.4 nT
East residual, res_BY	−739 to +412 nT	−33 to +113 nT
Passes above 50 nT in |res_BY|	13 of 150	1 of 100
Passes failing a common quiet test, Kp ≤ 3 and |Dst| ≤ 20 nT	69 of 150	29 of 100
Rms of diff_res_BY per pass: median, largest	0.204, 5.39 nT/s	0.154, 0.82 nT/s
Points flagged by a burst rule: evening, morning passes	13.6%, 20.2%	10.7%, 15.5%
Largest jump in |diff_res_BY| per pass: median	1.69 nT/s	1.76 nT/s
Vector-instrument flag, Flags_B, not 0	No rows	4 rows
Scalar-instrument flag, Flags_F, not 0 or 1	4,327 rows, 1.8%	1,376 rows, 0.9%
Labels	No label file	92 passes labelled; 8 carry none: samples 3, 14, 16, 43, 58, 72, 84 and 87

res_ columns hold what is left after a field model is subtracted, and diff_res_ columns hold its change from one row to the next. Which model is unread; CHAOS-7 in the reference list suggests one (inference) [37][25]. The burst rule flags a point when the 31-row rms of diff_res_BY reaches 0.258 nT/s, the threshold that best matches the test labels. Samples count from 0.

OUR ANALYSIS

“Stable”: two readings

Instrumental. If §2 defines “stable” as nominal instrument data, the word can mostly stand: the vector flag, Flags_B, is 0, nominal in ESA’s handbook, on every training row, though 1.8% of training rows carry a non-nominal scalar flag, against 0.9% of test rows [31]. We can’t read §2, so we can’t rule this reading out, and it is why we rate the claim Misleading and not False.

Geomagnetic. If it means magnetically quiet, it doesn’t fit. The set takes in a storm that reached Kp 6+ and Dst −93 nT on August 5, and nearly half its passes fail a common quiet test. In fairness, the typical largest jump per pass is about equal in the two sets. A reader of a paper on earthquake signals would likely take “stable” this second way, which is what makes the sentence misleading.

The largest training residuals, up to 739 nT in the East component, lie near 57–60° geomagnetic latitude south of Australia on August 2 and 5, because a 50°S limit reaches that far there. Field-aligned currents are expected at those latitudes (inference).

OUR ANALYSIS

“Randomly selected” and “significant”

“Randomly.” The training set is 150 of 151 consecutive half-orbits, with one skipped early on August 2. Whatever the word refers to, this isn’t a random draw from 2023, or from anything longer than those five days. That holds whichever way “stable” is read, so the verdict doesn’t rest on that word alone.

“Significant anomalies.” Labelled, yes, mostly: 92 of the 100 passes. By these measures their disturbances are no larger than those of the training set, which is labelled nowhere.

Why it matters. A reconstruction method judges new data by how far they stray from what it learned [1 §1]. What it learns as normal here includes a storm, so what counts as anomalous is relative to that (our reading; §§3–4 may say more). A set that spans a storm may be deliberate, so that storms aren’t scored as anomalies, and §§2–3 may say. “Stable” would still be the wrong word for it.

5 · WHAT THE LABELS MARK

The labels mark short bursts in the vector channels, peaking either side of the magnetic equator

14,074 of the 157,300 test points, 8.95%, carry a label, in 500 runs of 1 to 330 rows; the median run is 15 rows, about 15 seconds [2]. Who marked them, and by what rule, is unread. The labels follow one thing closely and little else (our analysis).

What the labels track

DOES THE LABEL TRACK…	AUC (95% LIMITS)	READING
31-row rms of diff_res_BY	0.958 (0.951–0.966)	Yes
Kp	0.564 (0.511–0.605)	No; slightly toward low Kp
Dst	0.537 (0.482–0.582)	No
|Dst|	0.529 (0.479–0.571)	No
|Latitude|	0.562 (0.511–0.607)	Slightly, toward low latitude
|Geomagnetic latitude|	0.553 (0.506–0.595)	Slightly; see the chart
Local time, hours from midnight	0.553 (0.503–0.595)	Slightly
Flags_F	0.511 (0.496–0.528)	No
Flags_B	0.500 (0.500–0.500)	No

AUC: the chance that a labelled point scores higher than an unlabelled one, shown as the larger of AUC and 1 − AUC, so 0.5 means no relation. Limits come from resampling the 100 passes [2].

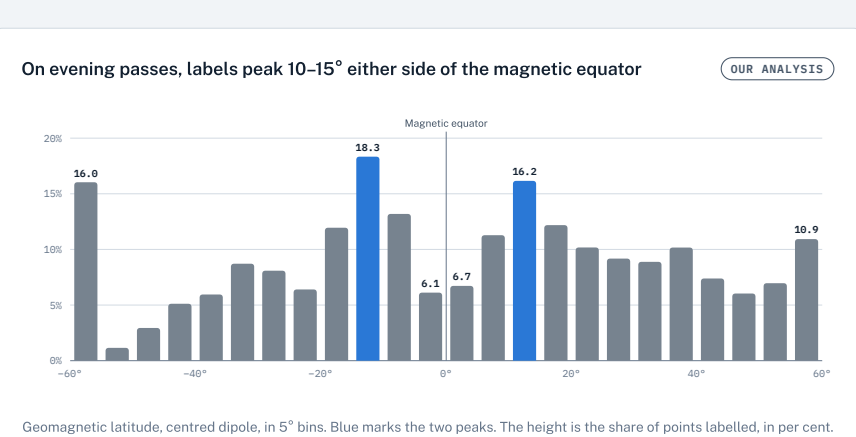
Inside labelled runs and outside them

MEDIAN CHANGE PER ROW	INSIDE	OUTSIDE
East component, diff_res_BY	**0.270 nT/s**	**0.040**
Centre component, diff_res_BZ	0.088	0.026
North component, diff_res_BX	0.070	0.025
Total field, diff_res_F	0.0131	0.0119

The direction of the field jumps inside the runs, and its strength barely changes. Absolute values; our analysis of [2].

On evening passes, labels peak 10–15° either side of the magnetic equator

OUR ANALYSIS



Geomagnetic latitude, centred dipole, in 5° bins. Blue marks the two peaks. The height is the share of points labelled, in per cent.

Our analysis of the authors’ dataset, CC BY 4.0 [2]: the 94 evening passes of the test set. The outermost bins hold about 1,400 points each, against about 7,400 near the equator. No figure of the paper is redrawn.

OUR ANALYSIS

What this suggests, marked as inference

Currents. Changes across the field with little change in its strength are what field-aligned currents produce. De Santis et al. write the same of their East-component anomalies, “field aligned current processes that do not practically affect the total intensity” [13 Methods]. For this file, that is our inference.

The magnetic equator. On evening passes, 17.25% of points are labelled at 10–15° from it, against 6.42% within 5° of it, a ratio of 2.69 (95% limits 1.51–5.59). Ionospheric currents and plasma are organised around the magnetic equator [27], and earthquakes aren’t. That suggests many labelled bursts are ionospheric. It doesn’t prove it.

Side of the orbit. Evening passes are 9.41% labelled and morning passes 1.76%, yet the burst rule flags more on morning passes, 15.5% against 10.7%. So the labels aren’t simply bursts: on the six morning passes, most bursts go unmarked.

OUR ANALYSIS

Plasma bubbles explain a part, not the whole

Where ESA’s Ionospheric Bubble Index flags a point, 1,716 points in the test set, 58.9% are labelled, against 9.1% where it reads 0: a 6.5-fold excess [30][26]. So some labelled anomalies are what ESA’s own product calls plasma bubbles, an ionospheric phenomenon. But those points are only **7.2% of the labels**, and only 23 of the 500 runs touch one. Most labels aren’t flagged as bubbles, another 1,099 labelled points lie where the index makes no call, and this doesn’t explain them.

Nor do instrument flags or simple rules

Rows that ESA’s handbook marks as doubtful in the scalar channel, Flags_F 5, 17 and 21, are nearly all labelled, 86 of 95, but they make up 0.6% of the labels [31]. Rule families fixed in advance reach an F1 of 59.8% at best, with an overlap of 42.6%, short of our bar of 80%. So the labels stay unexplained by these rules. They could be marked by hand or come from a rule we didn’t try; §§2–4 may say.

6 · WHAT 92.70% CAN MEAN

What the score means depends on how it was scored, point adjustment above all, and we can’t read that

F1 balances two questions about a detector: of the points it flags, how many carry a label, and of the labelled points, how many does it flag? So it scores agreement with the labels, and can’t say what the labels are. On this test set, flagging every point scores 16.4%, the score to beat (our analysis) [2].

Two ways to score the same flags

Plain F1 counts each point on its own. Point adjustment credits a whole labelled run if any one point in it is flagged. Kim and colleagues warn that it “has a great possibility of overestimating the detection performance; that is, even a random anomaly score can easily turn into a state-of-the-art TAD method”, TAD being time-series anomaly detection [19 Abstract].

Most such methods use the step, Kim and colleagues say [19 Abstract]. The public code of TimesNet, a model the authors cite, applies it in its anomaly test as “(4) detection adjustment”, after setting its threshold at a percentile of the scores over train and test together [21][37]. That doesn’t show that the authors used it.

OUR ANALYSIS

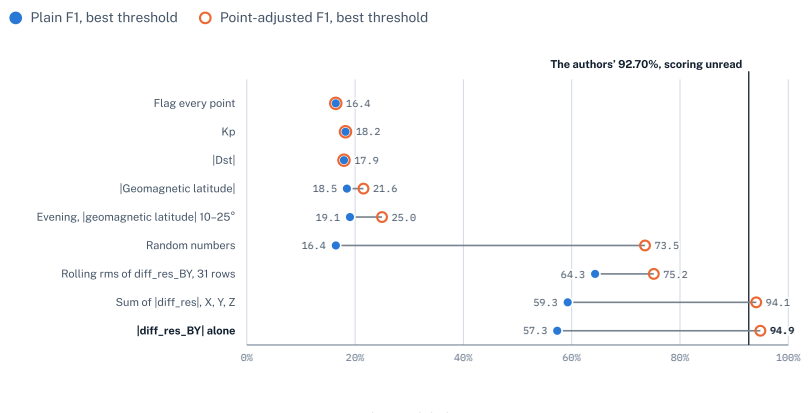
What we did

We scored simple detectors with no learning, each using one number per row, against the 14,074 labelled points. The chart shows each at its best threshold, which uses the labels, so these are upper bounds. The table adds a threshold chosen on half the passes and scored on the other half, and TimesNet’s recipe, which uses no labels. Of the five ratios we fixed in advance for that recipe, 5% scored best, so it isn’t a neutral choice; the tables give the others.

We stated before scoring that this says how hard the labelled task is, not how good the network is, and nothing about earthquakes either way (rule A6). The authors’ 92.70% is the one number here we didn’t compute.

Point adjustment lifts random numbers from 16% to about 74%, and one channel from 57% to 95%

OUR ANALYSIS



F1 against the test labels, in per cent.

Our analysis of the authors’ dataset, CC BY 4.0 [2]. Random numbers: 200 seeds, each at its own best threshold, with a mean of 73.5% and a range of 70.8–75.8%. The exact best for |diff_res_BY| is 94.85%, with 96.0% precision and 93.7% recall. The line is the authors’ figure [1 Abstract]; the chart doesn’t say which way to compare it, because its scoring is unread.

OUR ANALYSIS

If the 92.70% is point-adjusted

Then a one-channel threshold with no learning can reach it, but only in a band. Under TimesNet’s recipe, a ratio of about 2.3–7% gives |diff_res_BY| 92.7–94.9%. At 1% and 0.5%, ratios in TimesNet’s public scripts, it gives 82.7% and 66.1%; at the file’s own labelled share, 8.95%, 90.3%. Random numbers reach about 72% at 5% and about 60% at 1%. So on this reading the score wouldn’t by itself show that the network beats a one-line rule, and whether it does depends on how its threshold was set, which is unread.

OUR ANALYSIS

If it is plain

Then the simple thresholds we tried top out at 57–64%, and 92.70% would be a large gain at reproducing the marks, far beyond one-line rules. That would be a real technical result, and the abstract would be right to report it. We can’t tell which of these two main readings holds, and other forms exist, such as a window-level or per-sample F1. §4 is unread and no code is public. We draw nothing from the 9.73% margin either, since we can’t read what the rivals scored or how.

On either reading, the number measures agreement with the labels, so it is no test of an earthquake link.

What to ask of §4, and why it is worth asking

Was the F1 point-adjusted? How was the threshold set? Did any baseline with no learning appear? §1 names only rivals built on RNN, CNN and attention [1 §1]. These questions are general ones in machine learning. Wu and Keogh argue, in a paper titled for it, that benchmarks solved by a line of code create an illusion of progress in anomaly detection [20]. Mignan and Broccardo, on 77 papers applying neural networks to earthquake prediction, found that “far simpler (and traditional) models seem to offer similar predictive powers, if not better ones” [23 Abstract].

7 · ARE LABELLED POINTS NEAR EARTHQUAKES?

Labelled points are no detectably nearer coming earthquakes than unlabelled ones, which doesn’t contradict the abstract

The abstract reports “geomagnetic anomaly detection” [1 Abstract], so a weak tie between these labels and earthquakes contradicts nothing it says of the score. §1 goes further, saying such a model would “provide reliability for seismic precursor analysis” [1 §1]; a score on these labels can’t show that. What the comparison shows is that the F1 on these labels isn’t a test of “pre-seismic”. Our settings are illustrative, not the paper’s, which we can’t read: shallow USGS earthquakes, under 50 km, within zones of Dobrovolsky radius $10^{0.43M}$ km, about 230 km at M 5.5 and 1,020 km at M 7 [36][28].

Test points against earthquakes still to come

M ≥ 5.5, DUE WITHIN	10 DAYS	30 DAYS	60 DAYS	90 DAYS
Passes touching a zone, training set	10.0%	17.3%	41.3%	48.7%
Passes touching a zone, test set	12.0%	20.0%	41.0%	46.0%
Labelled points inside a zone	0.64%	1.61%	4.79%	5.43%
Unlabelled points inside a zone	0.85%	1.91%	3.86%	4.80%
Labelled minus unlabelled, points (95% limits, by sample)	−0.21 (−0.89, 0.55)	−0.30 (−1.39, 0.95)	0.93 (−1.14, 3.20)	0.63 (−1.45, 2.95)
The same places on all days of 2014–2025: labelled places	1.32%	3.72%	6.90%	9.75%
The same places on all days of 2014–2025: unlabelled places	1.09%	3.08%	5.75%	8.16%

Our analysis of [2] and [36]; limits come from resampling the 100 test passes. The point rows use every fifth row and the same-places rows every 50th; all rows give 5.46% against 4.81% at 90 days and 1.51% against 1.92% at 30. Higher magnitude limits look similar: at 90 days, 3.64% against 2.89% for M ≥ 6, and 0.18% against 0.12% for M ≥ 7.

OUR ANALYSIS

What the table shows

No detectable excess. About 95% of labelled points have no qualifying earthquake due within 90 days, and at every lead time the difference from unlabelled points is within the noise. The places under labelled runs are a little more active on all days too, 9.75% against 8.16% at 90 days. Sliding the catalogue to 300 other dates, which keeps the geography, the real gap of +0.63 points is smaller than the typical one, a mean of +1.67 with a spread of 2.35, and 63% of dates give a gap at least as large.

The test’s reach. It is coarse. With that spread, and a 95th percentile of +5.3 points, a tie short of about doubling the 4.8% wouldn’t show in 100 passes.

Wider settings. Allowing any depth widens the gap, to 12.83% against 7.20% at 90 days and 11.55% against 5.93% at 60, and so does doubling the radius, to 22.85% against 16.19%. Their limits by pass exclude zero, but they stay within the spread at other dates, which match or exceed them on 17%, 9% and 34% of dates. A review reached these after trying about 50 settings. Earthquakes just past give 9.09% against 6.28%, matched on 32% of dates. All are illustrative, and were added after the first results (rule A13).

No seismic split. “Stable” passes and those with “significant anomalies” touch these zones about equally often, so the split between train and test isn’t a split between quiet and pre-seismic ground.

No cluster. The events most often largest for a pass are spread out: Morocco, M 6.8, September 8, 2023 (8 training passes, 4 test passes); Chile, M 6.6, October 31 (3, 5); Macquarie, M 6.3, October 11 (6, 5). So the set wasn’t built around particular earthquakes, as far as this shows (inference).

OUR ANALYSIS

What it doesn’t show

It says nothing about the three earthquakes of 2024, or what §5 found around them. That analysis uses other data, and we can’t read it.

It also isn’t a test of whether anomalies ever precede earthquakes: so short a span can’t do that. It shows only that this score doesn’t measure it.

An F1 can speak to earthquakes when its labels do. SafeNet, a deep network on Swarm data, labels windows around real earthquakes and reports an F1 of 0.846 [16]. Even there, the negatives are artificial events at least 10° and 15 days from any earthquake, so they differ in place as well as in signal.

The questions a reader should put to this paper

QUESTION	WHAT CAN BE SAID NOW
A control: comparable windows with no earthquake?	Unread for §5. The train and test sets aren’t one, as the table shows. Zhu et al. show what one can look like: four pseudo-epicentres and two low-seismicity periods [15].
Are solar and magnetospheric effects removed?	The readable text doesn’t mention solar activity, storms, Kp or Dst, and §1 describes the field as “generated by internal sources”, which leaves out the external currents a satellite also sees [1 §1][27]. The file holds Kp and Dst columns, with no filter on them [2]. Periodic sources in such data include the 90-minute orbit, the drift in local time [15] and the daily variation [27]; whether “period-aware” features learn them is unread. Being periodic or not says nothing of a signal’s origin. Six of the 53 references concern modelling geomagnetic storms [37], and Kp and Dst are columns the network may take as inputs, so §§2–5 may say more; we can’t tell (inference).
False alarms and misses, with skill against a reference?	Only F1 is readable. Skill needs a reference: a related study of Swarm anomalies reports accuracy of 72–95% beside hit rates of about 1–19% [14], and for rare events accuracy above 50% doesn’t beat chance (our analysis).
Radius and lead windows set in advance? Declustering?	Unread. Dobrovolsky et al. in the reference list suggest a radius (inference) [37].
Anomalies verified as seismic, or only labelled?	Labelled, by an unread process. They show no detectable tie to coming earthquakes on our illustrative settings, some are ESA-flagged bubbles, and they peak either side of the magnetic equator.
Would it pass the rules we apply to any earthquake-timing claim?	Freeze the rules first, compare with chance, and hold back part of the record or wait for new events: see our check of a forecasting paper, and our check of magnetic surveys in Uzbekistan for a long record with no link. §5 would have to show them.

8 · THE RECORD

The introduction leans on the hopeful side of a contested record

§1 is firmer than the abstract. “Numerous studies on earthquakes demonstrate that multiple geophysical parameters exhibit anomalous behavior prior to an earthquake”, it says [1 §1], where the abstract says “may exhibit” [1 Abstract]. None of the critical sources below is among the paper’s 53 references [37]. That is a fact about the list, and we don’t say why.

The CHAMP study: what §1 says, and what the study says

§1 says Balasis and colleagues used CHAMP satellite data “to analyze pre-seismic signals” for two Sumatra earthquakes, and that “Their study confirmed the correlation” [1 §1]. The study’s title is a question. Its abstract reports that “some specific features are observed after the two earthquakes, with periods of about 16 and 30 s”, says “Our results favor an external source origin for the 30 s pulsation”, adds that a feature more than an hour later “could be associated with each earthquake, respectively”, the nearest it comes to §1’s claim, though still after the earthquakes, and ends: “Further investigations are required in order to answer the question of whether these signals can be associated with earthquakes” [4 Abstract].

The authors’ conference abstracts of 2007 and 2008 say the same; the second has “We continue to investigate in order to answer the question” [5][6]. A 2010 paper with Mandea as co-author sums up the work: “only a statistical study based on large earthquakes recorded during the CHAMP magnetic mission could bring an answer” [7 §1]. The study’s full text is closed, so our verdict rests on these abstracts and that summary. The full text would have to differ from all of them to change it.

“One of the most promising”: what Hayakawa says

§1 cites Hayakawa’s 2001 review for the geomagnetic field being among the most promising parameters [1 §1]. The review’s own words are that “Electromagnetic phenomena are recently considered to be a very promising candidate for the short-term earthquake prediction”, all of them, from radio to the ground [8 §1]. Hayakawa’s “most exciting finding” concerns seismo-ionospheric effects on VLF/LF radio propagation [8 Abstract]. For ULF magnetic signals the review leans on the Spitak, Loma Prieta and Guam cases [8 §5].

The later commission’s report on earthquake forecasting calls the Loma Prieta and Guam anomalies “The most convincing EM precursors”, then says reanalysis found they “were likely to have been caused by solar-terrestrial interaction or sensor-system malfunction” [9 p. 335]. So §1’s emphasis has a basis in its time; it leaves out what came after.

The later record, in its reviewers’ words None of these is cited in the paper.

> *“The search for diagnostic precursors has not yet produced a successful short-term prediction scheme”*
>
> International Commission on Earthquake Forecasting, 2011 [9 p. 319]

> *“a causal relationship between preearthquake magnetic changes and impending earthquakes has never been demonstrated”*
>
> Masci and Thomas (2015) [10 Abstract]

> *“In all considered events, the “anomalous” geomagnetic field behavior can be explained by global geomagnetic activity, and it is apparently not associated with seismic activity”*
>
> Pilipenko and Shiokawa (2024), who don’t rule out a usable signal [11 Abstract]

The other side, and the middle

The positive case is real research. De Santis and colleagues, for example, report from 4.7 years of Swarm data and 1,312 earthquakes of M 5.5 or more that anomalies cluster before earthquakes more than in random simulations [13]. Zhu and colleagues test one earthquake against controls [15]. This page doesn’t assess those papers. Most of the studies §1 lists look at one earthquake, after the fact, and nothing in their abstracts or records contradicts §1’s one-line summaries [1 §1]. We read Zhu et al. in full, Fan et al. by abstract and the others by title; those others aren’t in our source list.

Reviews stay divided: “the statistical significance of the results are highly debated, and more in general a common consensus is still missing” [12 Abstract]. A study of seven years of Swarm electron density and 527 earthquakes of M ≥ 6 relates most disturbances to ionospheric activity and some to seismic activity, finds that seismic part very weak and hidden, and urges more caution [18]. Proponents themselves call the possibility “controversial in the scientific community” [14 Abstract]. For more, see our check of a paper on crustal signals.

OUR ANALYSIS

PARTLY TRUE

Older methods ignore time, links

The abstract says “existing research primarily relies on traditional mathematical methods, which fail to account for variable relationships and temporal characteristics” [1 Abstract]. “Primarily” is fair, and some of it holds: De Santis et al. work on one magnetic component [13]. But “fail to account” is too sweeping. §1 itself lists several studies that use more than one parameter [1 §1], and deep networks on Swarm data already exist, SafeNet and Akhoondzadeh and colleagues’ LSTM among them [16][17 Abstract]. We rate this in the text only, not in the panel.

OUR ANALYSIS

The base rate, for illustration

How often does a place and a moment sit “before an earthquake” by chance? These settings are ours, not the paper’s: USGS 2014–2025, shallow, Dobrovolsky zones [36][28].

M ≥ 5.5, DUE WITHIN	10 DAYS	30 DA
All of 50°S–50°N	1.1%	3.
Japan	6.1%	16
Peru–Chile	2.9%	8.
California	1.8%	5.
Interior Australia, a quiet comparison	0.1%	0.
A whole pass, 50°S–50°N, touches a zone	12.4%	29

On the same settings, 42 of the 500 labelled runs touch such a zone within 90 days, and 13 within 30, in a file with no detectable tie. Coincidences are easy to find, which is why controls matter, and why selected cases can’t show a link alone. Removing aftershocks by Gardner and Knopoff’s windows [29] lowers these a little overall: the 56.0% becomes 51.5%, and the 7.9% becomes 7.1%; Japan falls more, by 17–28%.

OUR ANALYSIS

How big could a signal be at the satellite?

The best-instrumented test is Parkfield 2004, M 6.0: changes of 0.2–0.5 nT at the moment of the earthquake, at five sites, and “No preseismic magnetic field changes exceeding background noise levels” [24 Abstract]. Take 0.5 nT at about 10 km. At Swarm A’s height of about 490 km it falls to about 0.0002 nT from a line of dipoles ($1/r^2$, so $0.5 \times (10/490)^2$), and about 0.000004 nT from a single dipole ($1/r^3$).

Swarm’s error is “less than 0.3 nT, with a typical value of 0.1 nT” [13 Methods], and CHAOS-7 fits its East component to a weighted rms of 1.91 nT [25]. So nothing crustal from an earthquake of this size reaches the satellite directly. Scaling that dipole by seismic moment, about 32-fold per unit of magnitude, gives about 0.0001 nT at M 7, 0.004 nT at M 8 and 0.13 nT at M 9, so it stays far below the error short of the very largest. Any seismic signal seen by a satellite therefore needs a relay through the atmosphere and ionosphere. Proposals exist, but one review finds them “still unable to explain the large variety of the phenomenology” [12 Abstract].

9 · QUAKE WARNINGS FROM ORBIT?

The paper’s results neither support nor rule out magnetic warnings of earthquakes, and its own data show no detectable tie on our illustrative settings

If magnetic anomalies reliably came before earthquakes, satellites such as Swarm might one day warn of great ones. The paper’s abstract stops short of that, saying only that parameters before an earthquake “may exhibit abnormal patterns” [1 Abstract]. Its §1 goes further (section 8): it calls early warnings “crucial”, says that numerous studies “demonstrate” such anomalies and calls satellite observations of the field “particularly valuable for detecting seismic precursors” [1 §1]. Its results can’t test that. The 92.70% measures agreement with labels (section 6), and in the authors’ own file the labelled points are no detectably nearer coming earthquakes than the rest, in a test too coarse to see a small tie (section 7) [2][36].

OUR ANALYSIS

What the labels look like instead

On evening passes the labels peak 10–15° either side of the magnetic equator, around which ionospheric currents and plasma are organised and earthquakes aren’t, and some are what ESA’s own product calls plasma bubbles (section 5) [27][30]. That suggests many of the labelled anomalies are ionospheric. It doesn’t prove it, and it says nothing about the three earthquakes of 2024, whose analysis is in §5, which we can’t read [1 Abstract][1 §1].

OUR ANALYSIS

A signal from the crust would need a relay

On our rough estimate, scaled from Parkfield’s, a magnetic change from the crust itself would be about 0.0001 nT at Swarm’s height for an M 7 earthquake and 0.13 nT for an M 9, against Swarm’s error of “less than 0.3 nT, with a typical value of 0.1 nT” (section 8) [24 Abstract][13 Methods]. So, short of the very largest earthquakes, a warning from orbit would need the atmosphere and ionosphere to carry a signal up, and one review finds the proposed processes “still unable to explain the large variety of the phenomenology” [12 Abstract]. The record on such signals is still contested: De Santis and colleagues report, from 4.7 years of Swarm data and 1,312 earthquakes, anomalies clustering before earthquakes more than in random simulations [13], while Masci and Thomas find that a causal link between magnetic changes and coming earthquakes “has never been demonstrated” [10 Abstract] (section 8).

10 · REPEAT OUR NUMBERS

The scripts and small outputs are here; the authors’ zip stays at Zenodo

One command re-runs everything of ours from the authors’ zip and the public records, using Python’s standard library only. On November 2, 2026, it ran from clean copies in under three minutes and reproduced each output below byte for byte. This page hosts no copy of the dataset and no ESA file.

PY **run_all.py** Runs every script below in order · 2 KB

PY **fetch_public_inputs.py** Fetches the public inputs and says whether each is the file we used; ESA’s files only if you accept ESA’s terms · 6 KB

CSV **swarma_samples_dated.csv** All 250 samples dated: set, direction, time, local time, and the file’s Kp and Dst beside GFZ’s and OMNI’s · 26 KB

CSV **swarma_label_runs.csv** The 500 labelled runs: sample, start, length, mid-run position and time, largest changes, and bubble-flagged points · 33 KB

JSON **swarma_baselines.json** Every feature’s AUC and F1, the rule families, the labels by date and latitude, and ESA’s bubble flag · 39 KB

TXT **swarma_rules.txt** The rules, written before the analysis, with each later addition marked and dated · 13 KB

Run it

Download SwarmA-MagSet.zip from Zenodo, then, in a folder holding these scripts:

```
python3 -I fetch_public_inputs.py inputs --accept-esa-terms
```

```
python3 -I run_all.py SwarmA-MagSet.zip inputs out
```

The first fetches about 84 MB; read ESA’s terms at the top of the script before passing that flag. Without it, the second skips the two steps that need ESA’s files. The second reads the zip without extracting it and writes to `out`. GFZ, OMNI and USGS revise what they serve, so a later download can differ slightly.

Set in advance, or added after

The rules came first. Additions A1–A7 came after the file’s structure and dates and the paper’s text, but before any score or test of the labels. A8 fixed a crash before any score was seen. A9 came after the first baselines showed how far apart plain and point-adjusted F1 are, and A10 and A11 came while building this page, when a re-count showed that our first figure of 94.8% was cut short, not rounded, and that our dating table had read the indices at the wrong offset. A12 and A13 came after a fresh review of the page: more anomaly ratios, and the test against other dates.

The data and their terms

SwarmA-MagSet: Wen, Gu, Li and Jiang, CC BY 4.0 [2]. Our files date it and summarise it per pass and run; we changed nothing in it. **Kp:** GFZ, CC BY 4.0 [32]. **OMNI:** CC0 1.0 [33]. Its Dst comes from WDC Kyoto, for scientific use and no commercial applications, with acknowledgement of the WDC and its data suppliers and the data DOI [34]. The site is non-commercial, and that condition carries over to the Dst in our CSV files. **USGS’s catalogue:** public domain [36]. **NASA’s SSCWeb:** public, with no terms stated on its front page; we host only distances from it [35].

ESA: Data provided by the European Space Agency (ESA). © ESA (2026). Swarm products are free, and their access and use fall under ESA’s Data Policy and Terms and Conditions, which limit passing the files on [30]. So we host no ESA file, and report only counts derived from them.

Acknowledgements. We thank the observatories that supply the data behind the Kp index and the Dst index (Kakioka, Honolulu, San Juan and Hermanus), the WDC for Geomagnetism, Kyoto, GFZ, NASA’s Space Physics Data Facility, ESA, and the authors of SwarmA-MagSet.

11 · LIMITS

What this check can’t tell us, and what would change it

The paper, mostly unread

§§2–6 could change any verdict about the paper here, and claims 1, 2 and 3 most [1].

The network, not re-run

We found no code. Every score of ours is for a simple baseline, never for PaSTFF Net.

A file, not a paper

We describe what the authors published. It may not be all they used, and a description isn’t a test of their method.

Illustrative settings

The earthquake tables use one catalogue, one radius rule, and magnitude and depth limits of our choosing, and 100 passes can’t see a small tie. The paper may use others [36][28].

A coarse magnetic frame

Geomagnetic latitude here comes from a centred dipole with one pole position for the whole record. It is enough for the humps in the chart, not for fine structure.

Sources read in part

Balasis and Mandea’s study by its abstract [4] and conference abstracts [5][6]. He et al. through an index, so paraphrased [18]. Several others by abstract or record, as the list says. The paper’s reference titles are as Crossref lists them [37].

What only §§2–6 can settle

QUESTION	WHERE IT WOULD BE
What “stable” and “randomly” mean, and which nine features are used	§2, unread; the file has 15 columns [2]
Who made the labels, and by what rule	§§2–4, unread; our rule families don’t reproduce them
Whether the F1 is point-adjusted, how the threshold was set, what the rivals scored, and what 9.73% is a percentage of	§4, unread
Whether any baseline with no learning was run	§4, unread; §1 names only deep rivals [1 §1]
Which three earthquakes, which time windows, which controls and catalogue, and whether the network was retrained	§5, unread; not named or guessed here
What the authors conclude, and any statement on data and competing interests	§6 and the closing matter, unread

What would change these verdicts

01

§2 defining “stable” in a stated, instrumental sense, and how the samples were drawn: claim 1 could soften or firm up. A statement that “stable” meant magnetically quiet would make it False against the file.

02

§4 showing a plain F1 of 92.70%, or a re-run of the network on the public file doing so: claim 2 could become True as a score, a large gain over simple thresholds. §4 showing it point-adjusted: claim 2 could become Misleading if a single channel does as well under the same recipe.

03

§5 showing the earthquakes, windows and controls: claim 3 could become True as an observation, and a test keeping real anomalies against shifted earthquake times, on many earthquakes, could begin to say whether clustering beats chance.

04

The full text of Balasis and Mandea’s study confirming a pre-seismic correlation, against its own abstract: claim 4 would fall.

OUR ANALYSIS

Bottom line: an open dataset and a careful abstract, a score we can’t yet interpret, and a number that is no test of an earthquake link.

The authors did something useful: they published their data, which made this check possible, and they hedged their abstract. The data let us date the set to a 21-day span of August 2023 and see what §1’s description leaves out: the “stable” training set spans a storm and isn’t a random draw. The labels mark short bursts in the vector channels, peaking either side of the magnetic equator, with some plasma bubbles among them and no detectable tie to coming earthquakes on our illustrative settings, in a test too coarse to see a small one. Whether 92.70% beats a one-line threshold depends on how it was scored and how its threshold was set, which we can’t read. Either way, it measures agreement with those labels, so it is no test of an earthquake link. What the network finds around three earthquakes of 2024 may be interesting, and it is unread. The introduction’s account of the record is more confident than the sources it cites.

Related checks

FALSE Olivine and earthquakes The Disaster Cycle, Fact-Checked

Sources

Bracketed numbers refer to this list. For the paper [1] we read only ScienceDirect’s public view, pasted by the site’s owner: Highlights, Abstract and §1 are those parts of it, quoted where a sentence is rated or in short phrases; §§2–6 are unread, and no figure of the paper is redrawn. For the dataset [2], samples count from 0, and res_ and diff_res_ are its column names. Elsewhere p. gives a page, § a section, and Abstract or Methods that part of the cited work; where only an abstract or a record was read, the entry says so; an abstract that reached us only through an index is paraphrased. UT is Universal Time and nT the nanotesla; Kp is given to two decimals or in thirds, so 6+ is 6.33; F1 and AUC are explained where they first matter. Links checked November 2, 2026.

1 Wen, S., Gu, L., Li, X., and Jiang, T. (2026). Detection and analysis of pre-seismic geomagnetic anomalies using the period-aware structural temporal feature fusion network. *Advances in Space Research*, peer reviewed, article in press, dated October 2026; no volume or pages on record. © 2026 COSPAR, published by Elsevier B.V., all rights reserved, not open access. **Public view only:** the site’s owner pasted its Highlights, Abstract, Keywords and §1 from ScienceDirect’s public view on November 2, 2026. §§2–6, which hold the data, the method, the results, the earthquake analysis and the conclusions, are paywalled, and we could not read them. No figure of the paper is redrawn.

2 Wen, S., Gu, L., Li, X., and Jiang, T. (2025). SwarmA-MagSet, version 1 [data set]. Zenodo, published December 9, 2025. CC BY 4.0. The zip holds train.csv, test.csv and label.csv, 36,271,184 bytes, SHA-256 beginning 2513b67e; the whole hash is in fetch_public_inputs.py. The site’s owner downloaded it from Zenodo on November 2, 2026, and we read every row as data. Our tables and charts date it and summarise it; we changed nothing in it and host no copy.

3 SpaceWeatherNews (S0) (November 2, 2026). Pole Shift Changes Reported, Solar Coronal Hole | S0 News Nov.2.2026. YouTube video, https://www.youtube.com/watch?v=bSJbUC5WIgU, published at 12:02 UT. It links the paper in its description; no transcript was available.

4 Balasis, G., and Mandea, M. (2007). Can electromagnetic disturbances related to the recent great earthquakes be detected by satellite magnetometers? *Tectonophysics* 431(1–4), 173–195. Closed. Abstract only, from GFZ’s public record; we could not read the full text.

5 Balasis, G., and Mandea, M. (2007). CHAMP satellite observations during recent destructive megathrust earthquakes. *Geophysical Research Abstracts* 9, 02314. Conference abstract, read in full.

6 Balasis, G., and Mandea, M. (2008). A study of CHAMP satellite magnetic data for electromagnetic signatures possibly related to earthquakes. *Geophysical Research Abstracts* 10, EGU2008-A-02264. Conference abstract, read in full.

7 Zubaidah, T., Korte, M., Mandea, M., Quesnel, Y., Kanata, B., and Arumdati, N. (2010). Geomagnetic field anomalies over the Lombok Island region: an attempt to understand the local tectonic changes. *International Journal of Earth Sciences* 99, 1123–1132. GFZ’s postprint; we rely on one sentence of its introduction, on [4].

8 Hayakawa, M. (2001). Electromagnetic phenomena associated with earthquakes: review. *IEEJ Transactions on Fundamentals and Materials* 121(10), 893–898. Free on J-STAGE; read in full.

9 Jordan, T. H., Chen, Y.-T., Gasparini, P., Madariaga, R., Main, I., Marzocchi, W., Papadopoulos, G., Sobolev, G., Yamaoka, K., and Zschau, J. (2011). Operational earthquake forecasting: state of knowledge and guidelines for utilization. Report of the International Commission on Earthquake Forecasting for Civil Protection (ICEF). *Annals of Geophysics* 54(4), 315–391. Open access; p. gives the journal’s page. Quotations as checked for our earlier page, which read it in full; reference numbers inside them are left out.

10 Masci, F., and Thomas, J. N. (2015). Are there new findings in the search for ULF magnetic precursors to earthquakes? *Journal of Geophysical Research: Space Physics* 120(12), 10289–10304. Abstract only.

11 Pilipenko, V., and Shiokawa, K. (2024). A closer cooperation between space and seismology communities: a way to avoid errors in hunting for earthquake precursors. *Russian Journal of Earth Sciences*, 1–22. Abstract only, as deposited with Crossref.

12 Picozza, P., Conti, L., and Sotgiu, A. (2021). Looking for earthquake precursors from space: a critical review. *Frontiers in Earth Science* 9, 676775. CC BY 4.0. Abstract only.

13 De Santis, A., and 20 others (2019). Precursory worldwide signatures of earthquake occurrences on Swarm satellite data. *Scientific Reports* 9, 20287. CC BY 4.0. Read in full; cited here as an example of the field’s positive evidence and for its account of Swarm’s error. This page doesn’t assess it.

14 Marchetti, D., De Santis, A., Campuzano, S. A., and 18 others (2022). Worldwide statistical correlation of eight years of Swarm satellite data with M5.5+ earthquakes: new hints about the preseismic phenomena from space. *Remote Sensing* 14(11), 2649. CC BY 4.0. Read in full.

15 Zhu, K., Fan, M., He, X., and 6 others (2021). Analysis of Swarm satellite magnetic field data before the 2016 Ecuador (Mw = 7.8) earthquake based on non-negative matrix factorization. *Frontiers in Earth Science* 9, 621976. CC BY 4.0. Read in full.

16 Xiong, P., Marchetti, D., De Santis, A., Zhang, X., and Shen, X. (2021). SafeNet: SwArm for earthquake perturbations identification using deep learning networks. *Remote Sensing* 13(24), 5033. CC BY 4.0. Read in full.

17 Akhoondzadeh, M., De Santis, A., Marchetti, D., and Wang, T. (2022). Developing a deep learning-based detector of magnetic, Ne, Te and TEC anomalies from Swarm satellites: the case of Mw 7.1 2021 Japan earthquake. *Remote Sensing* 14(7), 1582. CC BY 4.0. Abstract only.

18 He, Y., Zhao, X., Yang, D., Wu, Y., and Li, Q. (2022). A study to investigate the relationship between ionospheric disturbance and seismic activity based on Swarm satellite data. *Physics of the Earth and Planetary Interiors* 323, 106826. Abstract only, through an index, so paraphrased.

19 Kim, S., Choi, K., Choi, H.-S., Lee, B., and Yoon, S. (2022). Towards a rigorous evaluation of time-series anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence* 36(7), 7194–7201. Read in full as its preprint, arXiv:2109.05257.

20 Wu, R., and Keogh, E. (2021). Current time series anomaly detection benchmarks are flawed and are creating the illusion of progress. *IEEE Transactions on Knowledge and Data Engineering*. Read in full as its preprint, arXiv:2009.13807.

21 Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., and Long, M. (2023). TimesNet: temporal 2D-variation modeling for general time series analysis. *International Conference on Learning Representations*. With its public code, Time-Series-Library: exp/exp_anomaly_detection.py, run.py and the TimesNet scripts in scripts/anomaly_detection, which we read as data and did not run.

22 Kapoor, S., and Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* 4(9), 100804. Abstract only.

23 Mignan, A., and Broccardo, M. (2020). Neural network applications in earthquake prediction (1994–2019): meta-analytic and statistical insights on their limitations. *Seismological Research Letters* 91(4), 2330–2342. Abstract only, as deposited with Crossref.

24 Johnston, M. J. S., Sasai, Y., Egbert, G. D., and Mueller, R. J. (2006). Seismomagnetic effects from the long-awaited 28 September 2004 M 6.0 Parkfield earthquake. *Bulletin of the Seismological Society of America* 96(4B), S206–S220. Abstract only, as quoted on our earlier page.

25 Finlay, C. C., Kloss, C., Olsen, N., Hammer, M. D., Tøffner-Clausen, L., Grayver, A., and Kuvshinov, A. (2020). The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly. *Earth, Planets and Space* 72, 156. CC BY 4.0. Read for its fit to Swarm data.

26 Park, J., Noja, M., Stolle, C., and Lühr, H. (2013). The Ionospheric Bubble Index deduced from magnetic field and plasma observations onboard Swarm. *Earth, Planets and Space* 65, 1333–1344. Record only, for what the index is.

27 Yamazaki, Y., and Maute, A. (2017). Sq and EEJ: a review on the daily variation of the geomagnetic field caused by ionospheric dynamo currents. *Space Science Reviews* 206, 299–405. CC BY 4.0. Record only.

28 Dobrovolsky, I. P., Zubkov, S. I., and Miachkin, V. I. (1979). Estimation of the size of earthquake preparation zones. *Pure and Applied Geophysics* 117, 1025–1044. Record only; the radius formula as [13] gives it.

29 Gardner, J. K., and Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America* 64(5), 1363–1367. Record only, for its aftershock windows; the formula we used is in base_rate.py.

30 European Space Agency. Swarm Ionospheric Bubble Index, product IBIATMS_2F, Swarm A, August 1–12, 20 and 21, 2023, version 0502, fetched November 2, 2026 from ESA’s dissemination server, which needs no account. Data provided by the European Space Agency (ESA). © ESA (2026). Access and use fall under ESA’s Data Policy and Terms and Conditions.

31 European Space Agency. Swarm Product Data Handbook, entry SW_MAGx_LR_1B, for what the values of Flags_B and Flags_F mean. Read November 2, 2026.

32 Matzka, J., Bronkalla, O., Tornow, K., Elger, K., and Stolle, C. (2021). Geomagnetic Kp index, version 1.0. GFZ Data Services. CC BY 4.0. File Kp_ap_since_1932.txt, read November 2, 2026; GFZ. The index is described in Matzka, J., Stolle, C., Yamazaki, Y., Bronkalla, O., and Morschhauser, A. (2021), The geomagnetic Kp index and derived indices of geomagnetic activity, *Space Weather*.

33 Papitashvili, N. E., and King, J. H. (2020). OMNI Hourly Data Set [data set]. NASA Space Physics Data Facility. File omni2_2023.dat, read November 2, 2026. CC0 1.0.

34 World Data Center for Geomagnetism, Kyoto, Nose, M., Iyemori, T., Sugiura, M., Kamei, T., Matsuoka, A., Imajo, S., and Kotani, T. (2015). Geomagnetic Dst index [data set]. Provisional for 2023, read through [33], under the WDC’s data usage rules: scientific use, no commercial applications, with acknowledgement of the data suppliers and the WDC, and the data DOI.

35 NASA Goddard Space Flight Center. Satellite Situation Center Web (SSCWeb). One-minute positions of Swarm A, B and C, July 31 to September 1, 2023, read November 2, 2026.

36 U.S. Geological Survey. Earthquake catalog (ComCat), FDSN event service. Earthquakes of M 5 or more, 2014–2025, queried November 2, 2026. Public domain.

37 Crossref. Metadata record for doi:10.1016/j.asr.2026.10.049. Read November 2, 2026: the date, the funder, Elsevier’s copyright line, “All rights are reserved, including those for text and data mining, AI training, and similar technologies”, and the paper’s 53 references. We use their titles as Crossref lists them, and don’t read them as telling us how the paper uses them.

METHOD

How this was checked

On November 2, 2026, we tried the routes the site allows to the paper’s text: Crossref, OpenAlex, Semantic Scholar, Europe PMC, CORE, BASE, the journal’s feed, preprint servers and the authors’ pages. None holds it; where a site showed a bot check, we stopped there. The page rests on the paper’s public view, pasted by the site’s owner; on the 53 references Crossref lists; on the authors’ dataset, which the site’s owner downloaded from Zenodo and whose 393,250 rows we read as data; on 26 other papers, reviews and conference abstracts (eight read in full, three in part, 11 by abstract and four by record); and on public records from ESA, NASA, GFZ, WDC Kyoto through OMNI, USGS and Crossref: the 37 sources listed. Quotations were checked word for word against the saved sources, and those from the paper are limited to rated sentences and short phrases.

The verdicts and the items tagged “Our analysis” are our own assessment, built on the cited sources. They include the dating of the dataset, every statistic on its rows, the baselines, the earthquake comparisons, the base rate, the size estimate, and our reading of what the readable text and its reference list do and don’t say. Anything we infer about how the data were built is marked as inference. §§2–6 could change any verdict about the paper; section 11 says how. If you have the article, please tell us through the correction form.

Our scripts, hosted under Repeat our numbers, take the authors’ zip and the public records as inputs and write every number of ours on this page except the one-line size estimate, whose arithmetic is in the text. Run from clean copies on November 2, 2026, they reproduced their outputs byte for byte. Rules for describing the file were written before the analysis, and later additions are marked in them. We could not re-run the authors’ network, for which we found no code, or recompute anything in §§2–6. No figure of the paper is redrawn; the charts are our own, of the authors’ CC BY dataset and of public indices. A fresh review, which re-counted our numbers from the zip with its own code and found them sound, then changed what the page concludes from them: the point-adjusted reading now rests on a band of anomaly ratios and not on one, “21 days” became passes from 14 dates across a 21-day span, the title and “no tie” were softened to what the test can reach, a test against other dates was added, and §1’s tie to seismic analysis, the sentence in Balasis and Mandea’s abstract nearest to a link, the paper’s six storm references and the basis for “Overstated” were added in fairness.

On November 6, 2026, the page was revised to review the paper on its own. No transcript of the S0 News episode that linked it was available, and without one the site no longer comments on how an episode presented a paper. So the verdict on the link’s label was removed, and section 9, on the episode, was replaced by one on what the paper means for magnetic warnings of earthquakes, built only from what this page already showed; the paper’s claims are now numbered 1 to 5, and the title in search results now uses the paper’s own “pre-seismic”. The episode is cited only as having linked the paper [3]. The verdicts on the paper’s claims, and the checks behind them, are unchanged.

This page was researched and written with Claude, an AI assistant made by Anthropic, and then re-checked against its sources in a separate verification pass.

The verdicts on this page

MISLEADING	Contains something true but leaves a false impression, usually by leaving out key context.
OVERSTATED	Has some basis but claims much more than the evidence supports.
MISREAD	Cites a real source but gets wrong what it says.
UNVERIFIED	We couldn’t find the source or record it depends on.
PARTLY TRUE	Part of the claim holds up and part doesn’t.

All ten verdicts, and how we check

CORRECTIONS

Corrections and changes

If we get something wrong, we fix it and record the change below with the date. If a correction changes a verdict, the log says what the old verdict was.

Change log

Nov 6, 2026	Revised to review the paper on its own. No transcript of the S0 News episode that linked it was available, and without one the site no longer comments on how an episode presented a paper. The verdict on the link’s label is removed, the section on the episode is replaced by one on what the paper means for magnetic warnings of earthquakes, and the verdicts on the paper’s claims, now numbered 1 to 5, are unchanged. The title in search results now uses the paper’s own “pre-seismic”.
Nov 2, 2026	Published on factchecked.science after a full check against the sources. Built from the paper’s public view, pasted by the site’s owner, since §§2–6 are paywalled; its 53 references; the authors’ CC BY dataset, read row by row; the episode’s title and description, as no transcript was available; the literature on the point; and our own analysis of public indices and catalogues, whose scripts are hosted on the page. Six claims are rated on the site’s scale, one the episode’s and five the paper’s, plus one in the text. No figure of the paper is redrawn. Before publication a fresh review re-counted the numbers and changed what the page concludes from them: the point-adjusted reading now rests on a band of anomaly ratios, “21 days” became passes from 14 dates across a 21-day span, the title changed from “says nothing about earthquakes” to “is no test of an earthquake link”, “no tie” became “no detectable tie” with a test against other dates, claim 1 went from True to Context, and several points were added in fairness to the authors and to S0 News.