LEARNING VIIRS NIGHT RADIANCE FROM MODIS WITHOUT FIRE PRODUCTS: NEAR-SIMULTANEOUS PAIRS, A SHARED LATENT SCENE, AND A PRE-STATED EVALUATION

TECHNICAL REPORT

Joshua Julian Damanik
joshuajdmk@gmail.com

November 15, 2026

ABSTRACT

MODIS and VIIRS fire records differ in pixel size and detection limit, so continuing one with the other would jump at the 2012 sensor switch. Product-level harmonisations act on thresholded detections. We instead learn the raw signal: a small convolutional network (438,678 parameters) maps Aqua-MODIS 4 and 11 μm night radiance to S-NPP-VIIRS 375 m radiance. It is supervised by near-simultaneous pairs (within 30 minutes) and decodes a latent scene, background plus fire excess, rendered by fixed sensor views; no fire product is input or target. On 2,645 selection tiles the variants trade off: the Group-share loss leads on block average precision (0.62–0.63 against 0.46–0.49) but flags 2.67 times as many cells as VIIRS; the Gamma-BW variants have the right amount and flag count and rank blocks 0.017–0.028 AUC above the deployable Even spread, but not by the pre-stated margin above the oracle-background spread, which uses the measured VIIRS background. On one held-out episode (4 night scene pairs) the selected model fails 3 of 4 pre-stated lines; selection scores need not carry over, and four pairs give no rate. Recovery of fires below the detection threshold is not tested.

Keywords active fire · MODIS · VIIRS · harmonisation · image-to-image translation

1 Introduction

Active-fire detections from MODIS (1 km pixels) and VIIRS (375 m pixels) form the backbone of global fire monitoring, but the two sensors do not count alike. VIIRS reports many fires that MODIS does not see [10, 17, 28], so a record that continues the MODIS series with VIIRS data would jump at the 2012 sensor transition. Published harmonisations correct this at the level of the products, for detections and fire radiative power (FRP) [17, 25, 33].

We start from a different place. A fire detection is a threshold applied to the average radiance over a pixel footprint. A fire that fills part of a 1 km MODIS pixel can stay below the MODIS threshold and exceed the threshold of the 375 m VIIRS pixel that contains it, so the sensors disagree on counts by construction, and a detection product contains nothing below its threshold. This is our motivation, not a tested advantage: the evaluation starts from pixels where MODIS already shows a fire-level excess (Section 10). We therefore treat MODIS-to-VIIRS harmonisation as image-to-image prediction [15]: given the raw 4 and 11 μm radiance that MODIS measured, predict the radiance that VIIRS measured over the same ground. The target of such a model exists only where both sensors saw the same scene at nearly the same time. Aqua and Suomi-NPP fly the same afternoon orbit slot, and their overpasses fall within 30 minutes often enough to give a large set of real pairs (Section 4). The model reads the pairs through one idea: both sensors are views of one latent scene, a background plus a fire excess, rendered by fixed sensor views (Section 5). At use time only the MODIS side is needed, which is what would allow VIIRS-like radiance for the years before 2012.

The contributions are: (1) a raw-signal formulation in which a small network decodes a scene from either sensor and four reconstruction terms (a fifth, energy, in the Gamma runs) train it, with no fire product as input, target or loss term (Sections 3 and 5); (2) a near-simultaneous Aqua–S-NPP pairing, 1,017 night pairs on five continents giving 32,914

tiles (Section 4); (3) a pre-stated evaluation protocol with non-learning references, a fire-off control, pass lines fixed before scoring and one closed test (Section 7); and (4) a measured trade-off on the selection set and a negative result on held-out data (Sections 8 and 9): no loss family is best on all measures, and on the held-out episode the selected model fails 3 of its 4 pre-stated lines. The fire-off control shows that block-level ranking scores survive removing the fire output of the Group-share loss, while cell-level and amount measures fall in every family.

We call the model a *learned harmonization model*: a mapping from MODIS radiance to VIIRS-equivalent radiance, not a calibration of the sensors. To our knowledge, and within the documented search, ours is the first learned model that maps raw Aqua-MODIS 4/11 μm radiance to S-NPP-VIIRS 375 m radiance using near-simultaneous pairs, with no fire product as input or target (search protocol: Appendix A.7).

2 Background and related work

Sensors and geometry. MODIS on Terra and Aqua has observed the Earth since 2000 and 2002 and produces the long active-fire record distributed by FIRMS [8, 12]. VIIRS on Suomi-NPP (S-NPP) was launched in October 2011; its first fire detections date from January 2012, and a 375 m fire product followed the 750 m product designed to continue the MODIS record [21, 24]. Both detect fire with a contextual test on a mid-infrared band near 4 μm, where a small hot target dominates a cool background, and a window band near 11 μm [11, 24]. A MODIS pixel is 1 km at nadir and grows to about 9.7 km^2 at the swath edge; a VIIRS imagery pixel is 375 m and also grows toward the swath edge [24, 25]. The nominal detection limit differs by about a factor of seven (3.7 MW for MODIS, 0.5 MW for VIIRS, night, nadir) [25], and VIIRS detects several times more fire pixels than MODIS [10, 17, 28].

Product-level harmonisation. Published MODIS–VIIRS comparisons and corrections work on detections and FRP. A harmonised FRP record for 2012–2024 rescales MODIS FRP with one quadratic in view zenith angle fitted on pairs within five minutes and takes the VIIRS retrieval where one exists [33]; gridded FRP adjustments serve historical MODIS years over Africa [17]; observation operators for both sensors use overlapping granules one to two hours apart [25]. These methods model the view-angle dependence of their product explicitly, and they are strong on the quantity they target. They cannot recover what the product does not contain: nothing below the detection limit, and no background.

Coincident-overpass matching and learned translation. Matching two polar orbiters at near-simultaneous overpasses is the standard tool of radiometer intercalibration [3, 4, 26]. Learned cross-sensor mappings with concurrent VIIRS data exist for geostationary input: two journal papers train an autoencoder or U-Nets from GOES to the rasterised VIIRS fire product, with pairs within 2.5 and 5 minutes [1, 30], and a conference paper names VIIRS channels as targets (abstract only) [5]. Our scene decomposition follows coupled-unmixing fusion, where one latent scene is observed through known or estimated sensor models [31, 32].

Where this work sits. The pairing, the decomposition and the network components are established; the combination is the contribution. Our model works on radiance, not on a product, from polar orbiter to polar orbiter, where near-simultaneous overlaps are rare, recur on an 8-day cycle and must be found (Section 4). We did not run a head-to-head comparison with product-level methods: their outputs (FRP at detected pixels) and ours (radiance at 375 m cells) share no quantity that both can be scored on without further assumptions (Section 10). Table 9 in the appendix compares the closest published work.

3 Problem formulation and design principles

Task. The study works on tiles of 96×96 cells of 375 m (36 km). A tile is a night scene seen by one Aqua-MODIS overpass and one S-NPP-VIIRS overpass less than 30 minutes apart. For sensor $\sigma \in \{\mathrm{M}, \mathrm{V}\}$ and band $b \in \{4, 11\}$ μm the tile holds a radiance field $L_b^\sigma > 0$ on the cell grid Ω. A MODIS field is piecewise constant on its native pixels, each cell carrying the value of the native pixel that contains its centre (Section 4). A validity mask m marks cells that are clear land in both sensors; nothing outside m is used in the input or the loss. At use time only MODIS is available. The task is to predict the VIIRS field $\hat{L}_b^{\mathrm{V}}$ of the same ground. Claims are limited to the scale of 3×3 MODIS pixels (about 3 km); placement inside a pixel is measured separately and is only partly learnable (Section 9).

Design principles. Six decisions shape the work. Each is stated with its rationale, where it is realised and where it is tested. The rationales are design choices; Section 10 reports how far the evidence supports each.

P1. Harmonise the raw signal, not the products. A detection is a thresholded footprint average, so it carries each sensor's pixel size and threshold and contains nothing below it. Predicting the radiance before binarisation avoids

both and lets counts, FRP or flags be derived afterwards (MODIS FRP is linear in the 4 μm excess radiance over background [12, 29]). Realised: radiance in, radiance out, and a rule that no fire product enters as input, target or loss term; a FIRMS archive is used only to choose which nights and places to tile (Section 4). Not tested: no head-to-head against product-level methods was run, and the evaluation starts from pixels where MODIS already shows fire-level excess, so recovery of sub-threshold fires is not measured (Section 10).

P2. Supervise with near-simultaneous pairs. Only a pair of scenes taken close in time gives a measured target without labels. Aqua and S-NPP fly the same afternoon slot, so such overlaps exist and can be found from orbit geometry. The 30-minute limit is an intuition-based trade-off between the size of the training set and the stability of a fire between two looks; it was not derived or tuned. Realised in Section 4; tested for sensitivity in Section 9 (window experiment).

P3. One latent scene, rendered by fixed sensor views. Harmonisation is defined as a consistent representation: the same network decodes a scene (a background plus a fire excess) from either sensor, and fixed operators render that scene as MODIS or as VIIRS would measure it. The direction that matters is MODIS to VIIRS, because VIIRS exists only since 2012. An earlier formulation that pulled the two sensors' latent codes together directly did not give per-cell agreement in any of 11 runs and was replaced by this decoded scene. Realised in Section 5.

P4. A multi-scale convolutional network that sees the viewing geometry. MODIS pixel area grows from 1 km^2 to about 9.7 km^2 from nadir to swath edge, so the same fire is sampled differently across the swath. A convolutional network with several scales is given the view zenith angle and the pixel size of every pixel and learns the mapping from real pairs, which cover the whole swath. Product-level methods also model view angle [17, 33]; the claim is only that the network is exposed to the real geometry. Whether it uses it was not tested (Section 10).

P5. Small, general, varied. The network has 438,678 parameters and larger networks were not better on a memorisation test (Section 9); it receives no location, date or time-gap input. Training pairs come from five continents. Only night scenes are used: at night the 4 μm radiance is thermal emission only, and an exploratory check made while planning measured a daytime background bias of $+7.2\%$ against $+0.2\%$ at night in the 4 μm band (not independently reproduced). The cost is that day scenes are neither trained on nor claimed.

P6. Honest evaluation. Simple non-learning references, a fire-off control (the same model with its excess set to zero), pass lines written down before scoring and one closed test on a held-out episode (Section 7).

4 Data: near-simultaneous pairing and preprocessing

Sources. Radiance comes from NASA Level-1B products: Aqua MODIS calibrated radiance MYD021KM with geolocation MYD03 [19], and S-NPP VIIRS 375 m imagery radiance VNP02IMG with the 750 m geolocation VNP03MOD [27]. The 375 m cell centres are interpolated from the 750 m geolocation. The NASA FIRMS archive [20] is used only to decide which nights and places to tile, never as an input or a target.

Pairing rule. Candidate pairs are Aqua and S-NPP granules found from NASA's public granule metadata, without orbit prediction. The granule footprints must overlap by at least 1 km^2 and the granule mid-times must differ by at most 30 minutes. Each 36 km tile is then checked on its own: its time gap is the mean VIIRS scan-line time minus the mean MODIS scan-line time over the tile, tiles with a gap above 30 minutes are dropped, and a tile must be at least half covered by both swaths. Terra is not paired (almost none of its overlaps with VIIRS fall within 30 minutes in the tropics). The 30-minute window keeps 7.1% of the 904,669 Aqua–VIIRS overlaps of a 120-day sample, and the pattern recurs every 8 days with Aqua's orbital phase; because the two orbits drift against each other, the window shrinks over the years (Figures 4 and 5; one pair in Figure 1).

Funnel and splits. Pairs were selected for the presence of night fire, not by time gap, so every number in this paper describes fire-enriched night scenes and not a random sample of land. Table 1 gives the counts. Spatial blocks of 100 km are assigned whole to one side, and whole fire episodes (tiles within 100 km and 5 days) stay on one side. A selection set (SELECT) chooses checkpoints and settings, a report set (REPORT) was reserved for a final score, and the Kalimantan 2015 episode is the one-time held-out test (Section 7). Training tiles are at least 100 km from every selection and report tile, and South America, the year 2023, one Sumatra block and Central Kalimantan in September–October 2015 are excluded from all three sets.

The 30-minute limit. The limit is a compromise between the size of the training set and the stability of a fire between the two looks. Published MODIS–VIIRS comparisons and learned GOES-to-VIIRS mappings match scenes within five minutes or less [1, 10, 17, 30, 33]; the only direct measurement of fire change over about 30 minutes we found, in the Amazon, reports a 38% rise in fire count and recommends a window of ± 15 minutes around the observations [7]. Within our window a post-hoc check on 15,279 MODIS fire pixels (665 scene pairs) found no measurable decline of MODIS–VIIRS agreement with the gap: rank agreement 0.44 / 0.38 / 0.38 for gaps of 0–5 / 5–15 / 15–30 minutes, and

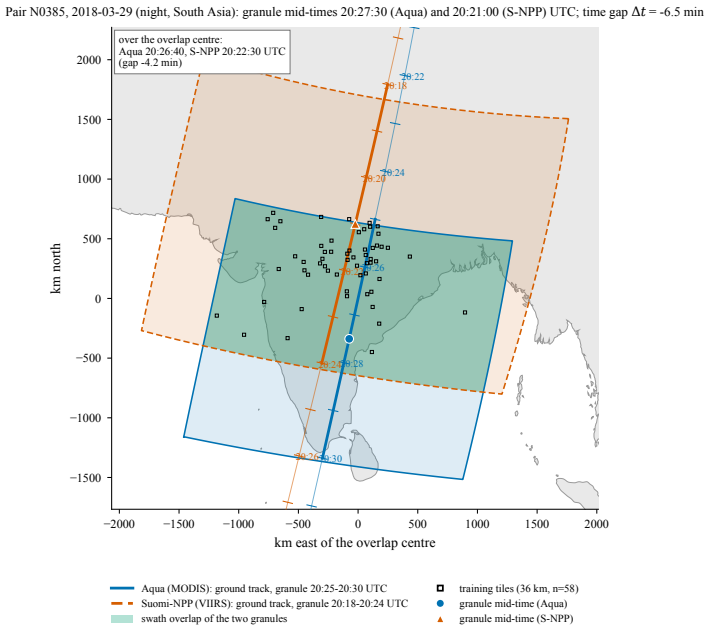


Figure 1: Two satellites, one ground. Ground tracks of Aqua and S-NPP (minute ticks) and the two granule footprints of one selected pair (29 March 2018, India); VIIRS passes the overlap 4.2 minutes before MODIS, and the tiles inside the overlap form training data. Orbits are computed from public elements and lie within 0.7 km (Aqua) and 0.2 km (S-NPP) of the positions recorded in the granules. One pair, training tiles only.

Table 1: From candidate pairs to the sets used. Counts are tiles of 96×96 cells unless stated. Mean absolute time gap of the kept tiles: 14.2 minutes. REPORT was not scored in this paper.

Stage	Count	Note
Candidate pairs (night-fire search)	7,462	Aqua × S-NPP, $\leq$ 30 min
Pairs selected	1,017	five continents; 971 produced tiles
Tiles built / kept	41,880 / 32,914	after mask and coverage filters
TRAIN	25,482	20,711 in the night training pack
SELECT	3,399	2,645 scored; 806 blocks
REPORT	3,945	after retiring 88 tiles (Statements)
Held-out episode (Kalimantan 2015)	1,499	697 night tiles, 4 scene pairs

an even-spread amount error that changes by +0.06 log units per 10 minutes (95% range -0.03 to $+0.13$) on a mean of 1.16. A rise of up to about 0.4 log units over 30 minutes cannot be excluded, and nothing beyond 30 minutes was tiled. The target has a noise floor of its own: two MODIS looks at the same fire less than a second apart already differ by a standard deviation of about 26.6% in fire radiative power [9], so no model can be expected to reproduce a VIIRS image taken up to 30 minutes later exactly. We do not claim that fires persist over 30 minutes; the effect on the trained model is tested in Section 9.

Preprocessing. Each tile cell is a bit-exact copy of one native pixel whose footprint contains the cell centre; valid covering pixels are preferred, then the nearest. VIIRS pixels deleted by the bow-tie correction are skipped, nothing is interpolated or blended, and radiance, not brightness temperature, is stored. The common mask keeps cells that are clear land in both sensors, using the products' cloud, water, glint and bow-tie classes but not their fire classes; VIIRS cells at the 367 K fill value that the fire product injects for saturated pixels are removed. The loss and the network see the same mask. Evaluation-only arrays (fire class, power, confidence) never reach the model.

5 Model

The model reads the image of *one* sensor per forward pass. It never receives a MODIS and a VIIRS image together; the other sensor enters only through a loss, in training. We describe one pass, then training as four such passes with shared weights, then use.

One pass. A pass takes the tile of a single sensor σ and returns a scene on the 375 m cell grid (Figure 2). Whatever the sensor, the input is a tensor of 12 channels on the same 96×96 grid of 375 m cells. For MODIS, each native 1 km pixel appears as the block of cells it covers, all holding that pixel's value; for VIIRS the cells are native pixels. The channels are computed from that sensor's radiance and its geometry only: the scaled log radiance of both bands relative to a 300 K Planck reference, the log excess ρ_b over the local background (both bands), the mask, the MODIS saturation and band-21 flags (zero for VIIRS), the view zenith angle, the native pixel size along and across the scan, the cosine of the solar zenith angle and the 4–11 μm contrast. Cells outside the mask are zeroed before any arithmetic, and every convolution and pooling is mask-aware [18]. A vector of seven tile-level conditions (sensor one-hot, mean geometry, valid fraction) modulates the Embedding through FiLM [22]. Time gap, latitude, longitude, date and land cover are not inputs.

The input goes through a *sensor-specific Embedding* (two masked 3×3 convolutions, 32 channels, FiLM; called "MODIS Embedding" or "VIIRS Embedding"), a shared three-level masked U-Net, the *Encoder* (32, 64 and 96 channels at full, half and quarter resolution, two skip connections, a 1×1 convolution to a 16-channel latent) that carries no sensor tag, and a shared *Decoder* (two convolution blocks, one with dilation 2, and a 1×1 head initialised to zero) that outputs four channels per cell (Figure 3). The Embeddings differ between sensors; the Encoder and Decoder weights are shared. The network has 438,678 parameters: 2 gains, $2 \times 13{,}248$ in the Embeddings, $379{,}536$ in the encoder, $32{,}644$ in the decoder.

Scene decomposition. Let A_b be the local masked median (window 7×7 taps, spacing 2 cells) of the input radiance, converted to the VIIRS band by Planck's law for MODIS. With head outputs $(d_4, d_{11}, o_4, o_{11})$ the scene is a background B_b and a fire excess E_b:

$$\ln B_b = \ln A_b + \delta_{\max} \tanh d_b, \tag{1}$$

$$s_b = \min\{\mathrm{softplus}(o_b) + g_b\, \rho_b^+,\ s_{\max}\}, \qquad E_b = B_b\,(e^{s_b} - 1). \tag{2}$$

The background may deviate from the local median by at most $\delta_{\max}$ in log units ($\delta_{\max} \in \{0.25, 0.5\}$). The excess is non-negative by construction when the learned gain g_b on the positive input excess ρ_b^+ is non-negative; g_b starts at 1 and is free to move. In the example checkpoint of Figure 2 the 4 μm gain ended at -0.048, so on the one example tile 17 of 9,057 valid cells have a slightly negative E_4 (minimum -0.017). The model sets $s_{\max} = 8$ and starts the excess near zero (head bias -4).

Fixed sensor views. Two operators without learned parameters render a scene as a sensor would measure it:

$$\hat{L}_b^{\mathrm{V}}(c) = B_b(c) + E_b(c), \tag{3}$$

$$\hat{L}_b^{\mathrm{M}}(p) = \Pi_b\Big(\textstyle\sum_c k_p(c) B_b(c)\Big) + f_b \textstyle\sum_c k_p(c) E_b(c). \tag{4}$$

Here k_p is the normalised footprint weight of MODIS pixel p, Π_b converts a background between the VIIRS and the MODIS band of the same blackbody, and f_b is the fixed ratio of the two bands' Planck radiances at a fire temperature of 800 K. The fitted footprint kernel is practically the box of the own pixel.

Training: four passes, shared weights. One training step runs four single-sensor passes. The Encoder and Decoder weights are shared by all passes; each sensor has its own Embedding. The scene of a pass is rendered by a fixed view and compared with a measured image, in training only:

1. *cross-sensor loss*, MODIS image → scene → VIIRS view against the measured VIIRS radiance (weight $w_{mv} = 1$). This is the product: the VIIRS image is used only in the loss, never as input;
2. *cross-sensor loss*, VIIRS image → scene → MODIS view against the measured MODIS radiance ($w_{vm} = 0.5$), a physical check that a scene decoded from VIIRS explains MODIS;
3. *same-sensor loss*, MODIS image with hidden regions → scene → MODIS view ($w_{mm} = 0.1$);
4. *same-sensor loss*, VIIRS image with hidden regions → scene → VIIRS view ($w_{vv} = 0.1$). Windows of 9×9 cells are hidden from the input in a tile with probability 0.5: three windows centred (with a jitter of up to three cells) on the brightest fire-level cells and one random window; a hidden MODIS cell hides its whole native pixel. The same-sensor losses weigh hidden cells 1.0 and visible cells 0.25, the loss on a saturated MODIS pixel is one-sided

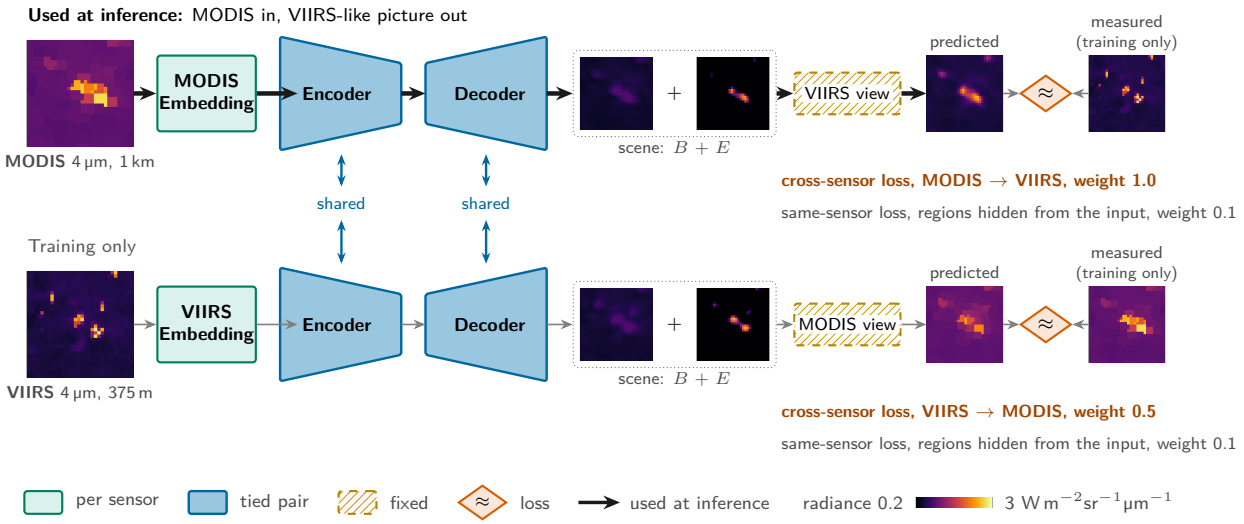


Figure 2: Two sensors, one set of Encoder and Decoder weights. Each sensor has its own Embedding; the Encoder and Decoder are the same network in both rows (weights shared). The Decoder gives a scene on the 375 m grid, a background B plus a fire excess E, and a fixed sensor view renders it as a sensor would see it. At inference only the top row is used: a MODIS image in (1 km pixels copied onto the 375 m grid), a VIIRS-like picture out. The comparison with the measured image happens in training only: a cross-sensor loss (MODIS to VIIRS, weight 1; VIIRS to MODIS, weight 0.5) and, for each sensor, a same-sensor loss on a pass with regions hidden from the input (weight 0.1). The energy term of the Gamma runs (weight 0.1) is not drawn. Every pass encodes one sensor's image only. Pictures: a 12 km crop of a tile from the training side (29 March 2018), 4 μm radiance on one scale, outputs of the Gamma-BW100 model; the images illustrate the mechanism, not skill.

(no penalty when the prediction is above the reading), and the three scales of the MODIS-to-VIIRS term have equal weights.

The Gamma runs add a fifth, *energy* term (weight 0.1): MODIS image → scene → MODIS view against MODIS. These five terms are the whole objective. There is no latent-consistency term and no term that compares the scenes decoded from the two sensors; optional block, allocation and hurdle terms exist in the code and are off in every reported run, and weight decay of the optimiser is the only other regulariser. All terms are evaluated inside the mask only.

Use. At inference only pass 1 runs: a MODIS image goes through the MODIS Embedding and the shared Encoder and Decoder, and the VIIRS view of the scene is the prediction. No VIIRS data and no fire product are needed.

Loss families and variants. The reconstruction loss ℓ comes in three forms. The *Group-share* loss is a Huber loss (threshold 0.1) on $\ln L$ with fixed shares $(0.4, 0.2, 0.4)$ for background, near-fire and fire-level cells, where the groups are defined from the *target* radiance. *Gamma* is the Gamma deviance $y/\mu - \ln(y/\mu) - 1$ on linear radiance, whose minimiser is the conditional mean [2, 13]; the MODIS-to-VIIRS term is evaluated at three scales (cell, MODIS pixel, 3×3 block). *Gamma-BW* multiplies each MODIS-to-VIIRS term by a smooth weight $1 + G\,\mathrm{clip}(\rho_4^{\mathrm{M}}, 0, 3)$ computed from the MODIS input only, with $G = 100$ (Gamma-BW100) or $G = 1000$ (Gamma-BW1000). A suffix B0.5 marks a background bound $\delta_{\max} = 0.5$ in place of 0.25. A Poisson deviance and a background-normalised squared error were tried in an earlier sweep (Section 9).

6 Training protocol

Training follows a staged, gated workflow written down before the runs reported here; a failed gate is treated as a bug to fix, not a setting to tune around. *Stage 0* freezes the train, selection, report and held-out lists, the references scored once on SELECT, the measures and the pass lines (Section 7). *Stage 1* runs sanity gates before comparisons: the network must memorise 8 fixed tiles with hiding switched off (loss fall above $10\times$, error at VIIRS fire cells below 0.1, energy at MODIS-bright cells within 0.9–1.1 of the measured), the training and evaluation paths must give the same loss on the same batch, and two identical short runs must give the same loss. All four gates pass for the Gamma family; the Group-share loss fails the energy line even when it memorises, painting about $1.37\times$ the measured energy, and a configuration without background correction ($\delta_{\max} = 0$) fails the memorise gate (loss fall $7.4\times$ against $10\times$) and was not trained. *Stage 2* trains each loss family at fixed budget with two seeds, then the two candidate settings (Gamma-

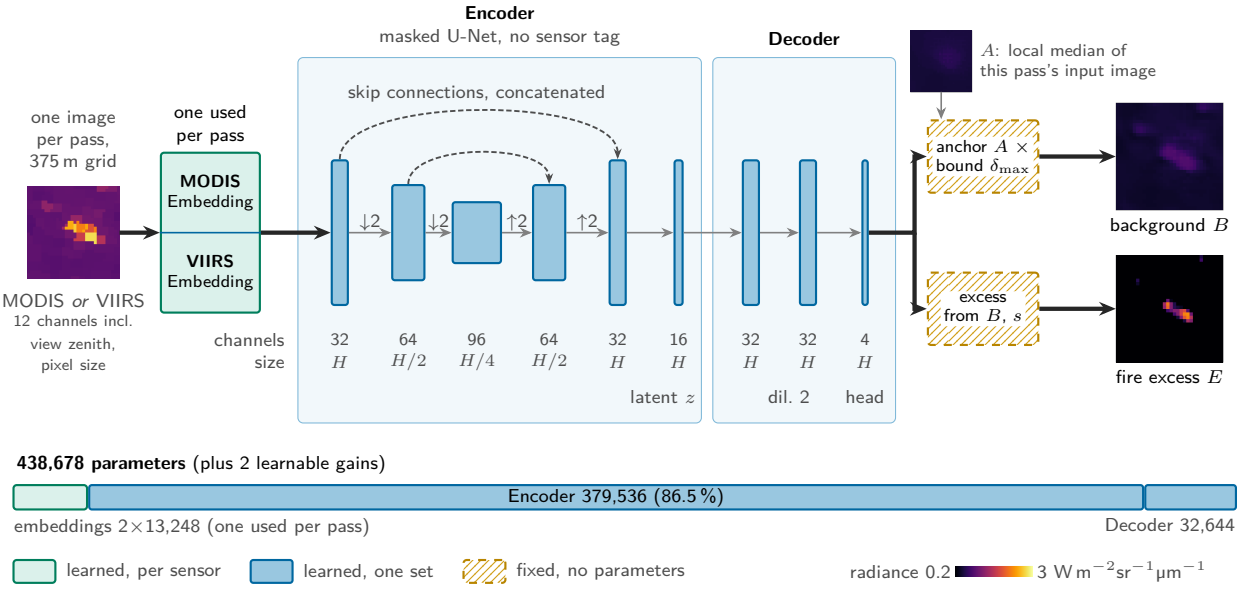


Figure 3: What is learned. One input image per pass enters the Embedding of its own sensor (the other is not used in that pass). The Encoder, a three-level masked U-Net (32, 64, 96, 64, 32 channels, two skip connections), maps it to a 16-channel latent; the Decoder (two blocks and a 1×1 head) gives four channels per cell. The fixed anchor A with a bounded learned correction gives B; E follows from B and the head. 438,678 parameters, 86.5% of them in the Encoder.

BW100 and Gamma-BW1000) with four seeds. The checkpoint with the lowest validation loss on a half-sample of SELECT (best-validation) is the pre-stated choice, and the final epoch is also scored; scores are computed on the full SELECT, so the selection effect covers half of the scored tiles. *Stage 3* applies a model pick rule fixed before the later seeds were scored, then scores one frozen candidate once on the held-out episode (Section 8.4). The order of these steps is in Appendix A.3.

The settings are in Table 7: AdamW with a peak rate of 0.002, weight decay 0.0001, batches of 32 tiles (a quarter with fire) and up to 46 epochs of 647 steps. Each run used one RTX 4090 for about 2.7 hours (46 epochs, including in-training evaluation); the whole study used about 60 GPU-hours.

The Group-share and plain Gamma runs were stopped by hand after 32 and 20 of the planned 46 epochs, so comparisons across families are partly confounded with training length.

7 Evaluation protocol

Sets. SELECT (2,645 scored tiles) gives the validation curve and chooses checkpoints and settings; every number in Sections 8 and 9 comes from it and is therefore optimistic by construction. REPORT was reserved for a final score; it is not used here (aggregate scores of some of its tiles were exposed, see the statements). The Kalimantan 2015 episode (697 night tiles) is the one-time held-out test.

References and control. The model must beat methods that use no learning. The *Even spread* divides the MODIS excess (converted to the VIIRS band) equally over the cells of each pixel; it is deployable. The *Oracle-background spread* does the same but spreads only the heat above the *measured* local VIIRS background; it uses VIIRS data, so it is an upper bound for the non-learning route and not a method one could run. A Planck copy and a raw copy of the MODIS brightness and a constant complete the set. The *fire-off* control is the same trained model with its excess set to zero, so only its background prediction remains.

Definitions. The log excess of a field is $\rho_b^\sigma = \ln L_b^\sigma - \ln \operatorname{med} L_b^\sigma$, with the median over a 7×7-tap window of spacing 2 cells. An *item* is a MODIS pixel that is complete, unsaturated, inside the mask and *fire-level* in the MODIS input ($\rho_4^{\mathrm{M}} \geq 0.5$); items are chosen from the MODIS input only. The excess of a cell is its VIIRS 4 μm radiance above the measured local VIIRS median, and predictions are always scored against that measured background, so that only the excess is scored. The *flag rule* is a simplified MOD14/VNP14 contextual test on brightness temperatures T_4, T_{11} (thresholds in Appendix A.5); it is a metric only, never a training target. A 3×3-pixel block is *positive* when the rule

Table 2: Lines used on SELECT (stage-0 lines; five-line form fixed after seeds 0 and 1 were scored). L1 uses the better of the two spreads and requires the fire part to add. On the held-out episode L2 and L5 merge into one amount line (3 km ratio of sums in [0.8, 1.25] and the amount delivered at pixel and 3 km scale), and L1 is computed over scene pairs.

Line	Requirement
L1 block ranking	region-balanced AUC difference to the better spread at least 0.05, 95% range above 0, positive in at least 75% of episodes, and the model above fire-off
L2 energy	pixel energy ratio in $[0.8, 1.25]$
L3 flags	flag-count ratio in $[0.5, 2]$ and block hit rate at least 0.5
L4 concentration	hottest-cell share at least 0.33
L5 amount, 3 km	the model delivers the amount (above)

fires on measured VIIRS radiance in it. A scene pair *qualifies* for the ranking score with at least 10 positive and 100 negative blocks, and for the amount score with at least 10 items.

- **Block ranking (AUC)**: the chance that a positive block gets a higher predicted block excess than a negative one [14], averaged over qualifying pairs (the tables); line L1 uses the *region-balanced* AUC (mean over episodes within each of nine regions, then over regions) with an episode-clustered 95% range from 4,000 bootstrap resamples. **Block average precision (AP)** scores the same ranking by precision–recall.
- **Amount**: the energy ratio $\sum \hat{e} / \sum e$ (1 is right) and the mean absolute log error of the excess per item, averaged per pair. A model *delivers* the amount at a scale if its error is smaller than that of every reference, the pair-clustered 95% range of the paired difference is above zero, and it wins in at least 75% of the qualifying pairs. **Amount skill** is one minus the ratio of the model's error to the Even spread's.
- **Concentration (hottest-cell share)**: the share of a pixel's predicted excess on its hottest cell, over fire-level pixels (measured VIIRS 0.46, Even spread 0.15). It is not accuracy: all energy on a wrong cell scores 1. The **hottest cell exactly right** is the rate at which the predicted hottest cell is the measured one, against chance.
- **Flags**: the flag-count ratio (predicted over measured), the block *hit rate* (share of measured flagged blocks that the prediction also flags), the false-alarm share (share of predicted flagged blocks without measured flag) and the critical success index $\mathrm{CSI} = \mathrm{found}/(\mathrm{found} + \mathrm{missed} + \mathrm{invented})$, computed exactly from the hit rate and the false-alarm share. The **fractions skill score** of the flag maps tolerates displacement [23].

Lines and their status. Table 2 gives the lines used on SELECT. They were written down before any model of the final workflow was scored, except that their five-line form and the model pick rule were fixed after the first two seeds of each setting had been scored and before the later seeds were (Appendix A.3). The four held-out lines were fixed before the held-out test. Average precision, the fractions skill score, CSI, amount skill and the exact hottest-cell rate were defined after the selection results were seen: they describe and are marked post hoc (†).

Pick rule and held-out test protocol. Among the two candidate settings (four seeds each, best-validation checkpoints) the one with more seeds meeting all of L2–L5 wins. The rule picked Gamma-BW100 by one seed (2 of 4 against 1 of 4), within seed noise. The held-out test scores one frozen candidate once: by the rule, seed 0 of the winning setting, so that the best seed is not chosen; on SELECT it meets L2–L4 but not L5. The Kalimantan 2015 episode (Central Kalimantan, September 2015) was excluded from every training, selection and report list by construction, with a spatial and date buffer; REPORT was not used for the one-time test and has no score here. The four held-out lines were recorded in an amendment approved before the run. A line is testable with at least 3 scene pairs, the minimum; the 3 km amount line rests on exactly that many. A variant of the scorer that drops cells next to masked holes was run alongside and gives the same verdict. Any failing testable line gives the verdict fail.

8 Results

8.1 Selection set: learned variants against the references

Table 3 compares the four loss families at their best-validation checkpoints with the non-learning references on SELECT; Table 8 (appendix) names the variants, and every run is listed in the appendix. Entries are the range over seeds. Four readings follow.

Block ranking: Gamma variants beat the Even spread, not the oracle by the margin. Table 3 prints the pair-mean AUC, but line L1 is judged on the region-balanced, episode-clustered difference to the better reference, the Oracle-background spread (Section 7). On that statistic the Group-share loss is ahead by 0.07 (95% range from 0.03,

Table 3: Learned variants against the references on SELECT (2,645 tiles, range over seeds, best-validation checkpoints). "off": the same model with its fire output set to zero. AUC and AP rank 3×3-pixel blocks; energy and flag ratios: 1 is right; error: mean absolute log error of the 3 km block excess over 33 scene pairs (lower is better); amount skill against the Even spread; hottest share: concentration over MODIS fire-level pixels (measured VIIRS 0.46); hottest exact: rate of the right hottest cell over the same pixels. † post hoc.

	n	Block AUC		Block AP†		Energy	Flags	Err.	Amount	Hottest	Hottest
		on	off	on	off	ratio	ratio	3 km	skill†	share	exact†
Group-share	2	0.927–0.928	0.920–0.922	0.62–0.63	0.59–0.61	1.38–1.51	2.65–2.70	0.53–0.59	-0.05 to 0.03	0.23–0.26	0.31–0.32
Gamma	2	0.861–0.871	0.775–0.820	0.46–0.47	0.33–0.38	0.88–0.99	0.98–1.02	0.46–0.46	0.20 to 0.21	0.24–0.27	0.29–0.29
Gamma-BW100	4	0.872–0.879	0.765–0.817	0.46–0.49	0.30–0.39	0.95–1.06	0.93–1.02	0.43–0.48	0.17 to 0.28	0.32–0.34	0.30–0.31
Gamma-BW1000	4	0.868–0.876	0.792–0.833	0.47–0.48	0.33–0.39	0.97–1.03	0.92–1.02	0.43–0.47	0.21 to 0.26	0.32–0.36	0.30–0.30
Even spread	–	0.851	–	0.47	–	–	1.25	0.64	0	0.15	0.10
Oracle-background spread	–	0.881	–	0.54	–	–	1.25	0.55	–	–	0.10
Chance	–	–	–	–	–	–	–	–	–	–	0.11

positive in at least 81% of episodes) and passes L1; in pair-mean terms the margin is 0.046–0.047 (AUC 0.927–0.928 against 0.881), just under 0.05, so the table alone does not reproduce the line. No Gamma-family run passes: the best is ahead of the oracle by 0.02 with a range that includes zero, and only 3 of 22 runs have a pair-mean AUC above it at all (0.849–0.886). Against the deployable Even spread (0.851) the Gamma-BW variants rank better by 0.017–0.028 pair-mean AUC (region-balanced 0.02–0.03, lower end of the range above zero in every seed), a small gain.

Amount: lower error in every run, weak evidence. All 22 Gamma-family runs have a lower 3 km amount error than the Oracle-background spread (0.41–0.52 for the 20 brightness-weighted runs against 0.55). These are settings and seeds on the same 33 scene pairs, not independent confirmations: the paired 95% range excludes zero in 15 of 22 runs, L5 is met in 4, a pass needs a win in 25 of the 33 pairs and the weakest passing run has 25 (one pair decides), and one Australian block supplies 72% of the SELECT fire pixels. The Gamma-BW variants have the right energy ratio and flag count; the Group-share loss has an energy ratio of 1.38–1.51, flags 2.65–2.70 times as many cells as VIIRS does and has no amount skill.

Detection scores favour the Group-share loss. Block AP is 0.62–0.63 for the Group-share loss against 0.46–0.49 for the Gamma-BW variants (even spread 0.47, Oracle-background spread 0.54), and its CSI is higher (0.63–0.67 against 0.51–0.63), partly because it flags too much. The Gamma-BW variants do not find fire better; they flag the right number of cells, which it does not.

The two Gamma-BW weights are not ranked. The pick rule chose Gamma-BW100 by one seed (Section 7); over four seeds the ranges of weights 100 and 1000 overlap on every measure in Table 3.

8.2 What the fire-off control shows

Setting the fire output to zero removes the model's excess and leaves its background. For the Group-share loss the block-level ranking scores survive: AUC falls from 0.927–0.928 to 0.920–0.922 and AP from 0.62–0.63 to 0.59–0.61. Cell-level and amount measures do not: the fractions skill score at 375 m falls from 0.36–0.37 to 0.10–0.14, the exact hottest-cell rate from 0.31–0.32 to 0.25, and amount skill to -2.62 to -2.30. For the Gamma-BW variants the fire output matters on every measure, including block ranking (AUC 0.872–0.879 to 0.765–0.817; AP 0.46–0.49 to 0.30–0.39). A block-ranking score can thus be high without the fire output; it does not show that the Group-share fire output is uninformative, since the cell-level measures fall.

8.3 Placement and sharpness

At 375 m the learned families place flagged cells no better than each other (fractions skill score 0.36–0.37 Group-share, 0.35–0.39 Gamma-BW, 0.27 even spread). From about 1 km up the Gamma-BW variants are better than the Group-share loss (3 km: 0.60–0.66 against 0.54–0.56) and better than the references at every scale, narrowly at 10 km. Concentration separates the weighted variants: the hottest-cell share is 0.32–0.34 (Gamma-BW100) and 0.32–0.36 (Gamma-BW1000) against 0.15 for the even spread and 0.46 measured (Table 3); the final-epoch checkpoints of the bound-0.25 settings are more concentrated still (0.41–0.45). The exact hottest cell is found at 0.29–0.32 of MODIS fire-level pixels (pair mean) by every learned family against 0.10 for the best reference and a chance rate of 0.11; on the population of pixels flagged in VIIRS the same rate is 0.32–0.35. This separates learned from non-learned models, not the learned ones; concentration is not accuracy.

Table 4: Held-out test (Kalimantan 2015, night, 4 scene pairs; 3 at 3 km). The verdict is the pre-stated one.

Line	Value on the held-out episode	Verdict
1 block ranking	AUC 0.741 against 0.852 for the Oracle-background spread (shortfall 0.110, range 0.060–0.156), below the even spread (0.783); above fire-off (0.675) by 0.066 (0.018–0.107)	fail
2 amount	3 km ratio of sums 1.00 (pixel 0.91); delivered at pixel scale (error 0.94 against 1.18 for the best reference) but not at 3 km (0.65 against 0.60 Oracle-background spread, 0.62 Planck copy, 0.78 even spread)	fail
3 hottest-cell share	0.359 against the line 0.33 (measured 0.489, even spread 0.141; 337 pixels)	pass
4 flags	flag ratio 0.70 (inside 0.5–2), block hit rate 0.39 (line 0.5; 10,784 measured flagged cells)	fail

8.4 The held-out episode

The frozen candidate (Gamma-BW100, seed 0, best-validation, epoch 17) was scored once on the Kalimantan 2015 night pack (697 tiles, 4 scene pairs; Table 4). The verdict is **fail**: 3 of 4 pre-stated lines fail and one passes. On this episode the model keeps its total energy (ratio of sums 1.00 at 3 km) and its sharpness (hottest-cell share 0.359 against 0.141 for the Even spread), and the fire output still adds to block ranking over fire-off. It does not rank blocks better than the references: its AUC is below that of the Even spread and the Oracle-background spread, and at 3 km its amount error is no better than the oracle spread or the Planck copy (it is better than the Even spread and the raw copy). The flag rule finds 0.39 of the measured fire blocks while flagging 0.70 of their number. The scorer variant that drops cells next to masked holes gives the same verdict (AUC 0.738, hit rate 0.39). The day pack of the same episode (802 tiles) has no fire pixels, so no line is testable there.

All scores fell relative to the selection set (block AUC, selection set to held-out episode: Even spread 0.851 to 0.783, Oracle-background spread 0.881 to 0.852, this candidate 0.872 to 0.741), the model's by more. Selection optimism and a shift in region and fire type (a peat-fire episode in a region excluded from training) cannot be separated with one episode. Four scene pairs, three at 3 km, support a clear statement about this episode and a warning about Table 3; they do not support a failure rate for other regions or years. The set-up of the test is in Section 7.

9 Ablations and sensitivity

Table 5 lists the design choices that were tested, each with its setup and a one-line reading. All are on SELECT (or, for the loss sweep, a small earlier set); only the first four use the large multi-continent data. With two to four seeds, only differences larger than the seed spread are readable. Settings were compared with the same scorer; the window experiment and the short-budget runs are short and not converged.

What the table supports. Only three of its tests move a measure beyond the seed spread: the loss family (ranking against over-painting), the brightness weight (concentration) and the memorise gate (background correction). The pairing-window test does not show that a tighter window helps; it also does not show that 30 minutes is optimal, and a window narrower than five minutes with more data was not tried.

10 Discussion and limitations

What the evidence says about each principle. *P1, raw signal.* The Gamma-BW variants have the right amount and flag count on the selection set (Table 3). The model is evaluated where MODIS already shows a fire-level excess; the motivating claim that the raw signal recovers fires below a detection threshold, which a product cannot contain, is not tested here, because every item-level measure starts from MODIS fire-level pixels. Whether the raw-signal route is better than product-level harmonisation is not testable here: no head-to-head exists in the literature or in our results, and the product-level methods report FRP at detected pixels while we score radiance at 375 m cells. *P2, pairing.* The 30-minute limit shows no decline of MODIS–VIIRS agreement within the window (Section 4) and, at equal data, a tighter window gave no detectable benefit (Table 5); both are weak tests (two seeds, short budget, nothing beyond 30 minutes). *P3, shared scene.* The decoded scene trains, and its MODIS-to-VIIRS pass is the only one needed at use. *P4, geometry.* The model is exposed to the real swath geometry through real pairs, but no experiment shows that

Table 5: Ablations and sensitivity tests. n: seeds. † post hoc. Short: 1,750 steps, not converged.

Change	Setup	Reading
Loss family	small data, 9-pair validation, 4 Group-share and 20 new-loss runs (U-Net and three-convolution network pooled), 30 epochs	Group-share ranks blocks better (AUC 0.785–0.800 against 0.671–0.718) but paints 2.07–2.42 times the energy (new losses 0.91–1.25); Gamma, Poisson and squared error are not distinguishable. Tuning-contaminated.
Brightness weight 0 / 100 / 1000	large data, best-validation, $n = 2/4/4$	Hottest-cell share 0.24–0.27 / 0.32–0.34 / 0.32–0.36: beyond the seed spread. Block ranking and amount error do not move materially.
Background bound $\delta_{\max}$ 0.25 / 0.5	same, $n = 4$ / 2	AUC 0.872–0.879 against 0.877–0.882 (weight 100); lower hottest-cell share with the wider bound (0.28–0.30); L5 met in 0 of 2 seeds against 2 of 4. The wider bound trades a slightly higher AUC for a lower hottest-cell share and no L5 pass; the three Gamma runs whose AUC exceeds the oracle's are all B0.5 runs.
Best-validation / final epoch	same	Final epochs of the bound-0.25 settings are more concentrated (0.41–0.45) and rank worse; the choice trades sharpness for ranking.
Small list / multi-continent data†	Gamma, $n = 2$ each	AUC 0.861–0.866 against 0.861–0.871, hottest-cell share 0.19–0.23 against 0.24–0.27. Differences are of seed size; training length also differs.
No background correction	short, memorise gate	Fails the gate (loss fall 7.4× against 10×); not trained. The bounded correction is needed.
Energy term; hidden-region terms†	short, Gamma, $n = 2$	Removing either stays within or near the seed spread; the brightness weight is the only ingredient that moves the lines.
Network width	memorise 8 tiles, 1 seed	Up to 6,841,878 parameters memorise no better than 438,678 (loss fall 5.2–5.6× for Gamma).
4 μm-only output	small data, Group-share, $n = 2$	No gain; a block score 0.012 lower than two bands, about twice their seed spread.
Pairing window†	Gamma-BW100, short, $n = 2$, equal data (449 tiles, 97 with fire) at $\leq$ 5 / 15 / 30 min	3 km amount gain over the even spread 0.072 / 0.072 / 0.084, block ranking gain 0.008 / 0.006 / 0.017: a tighter window gives no detectable benefit at equal data. All 2,340 tiles together rank best (0.032). Only large effects are visible with two seeds.
Scan angle†	Gamma-BW1000 seed 0, pixel scale	Error flat (0.92–1.04) up to 40° view angle, 1.40 in the outermost bin; whether the model uses its geometry inputs was not tested.

it uses the geometry inputs or handles the swath better than a view-angle-corrected product method. *P5, small and general.* A larger network did not memorise better. *P6, honest evaluation.* The fire-off control and the references changed the conclusions: a block-ranking score can be high without the fire output, and the strongest reference is an oracle-background spread. The held-out test failed three of four lines; without it the selection-set scores would have been read as skill on unseen data.

The trade-off. The Group-share loss ranks blocks and scores detection best but over-paints (2.67 times as many flags as VIIRS, no amount skill); the Gamma-BW variants hold the right amount and flag count, are sharper and place better from about 1 km, and rank 0.017–0.028 AUC above the Even spread but not by the pre-stated margin above the Oracle-background spread. The held-out episode adds that the Gamma-BW100 candidate keeps energy and sharpness there, but ranks and flags worse than on the selection set (AUC 0.741 against 0.872, hit rate 0.39 against 0.56); the references fell too, so selection optimism and a shift in region and fire type cannot be separated.

Limitations.

- **Night only.** Day scenes are neither trained on nor claimed; most fire activity may fall in the day overpasses, which we did not quantify.
- **Few seeds, short study.** Two or four seeds per setting and a study of about two days; differences of the size of the seed spread are not results. One scene pair decides the 3 km amount line on SELECT.
- **One held-out episode, four scene pairs.** The held-out result is a case study of one region and season. It supports a statement about that episode and no rate across regions.
- **Selection and post-hoc measures.** Settings and checkpoints were chosen on SELECT, so its scores are optimistic; average precision, fractions skill, CSI, amount skill and the exact hottest-cell rate were defined after seeing results.
- **Same-source data.** All tiles come from the same two archives with the same tiler and masks; splits are spatial and by episode, but there is no independent sensor pair or reprocessing. One Australian block supplies 72% of the SELECT fire pixels, and pairs were selected for night fire.
- **VIIRS as truth.** The target is a measurement made up to 30 minutes from MODIS with a different view geometry, with cell centres interpolated from 750 m geolocation.
- **Sub-threshold signal not tested.** Every item-level measure starts from MODIS fire-level pixels, so fires that MODIS does not detect are outside the evaluation. The target also has a noise floor: MODIS looks at the same fire less than a second apart differ by about 26.6% in FRP [9], and no ceiling for a 0–30 minute gap was measured.
- **No head-to-head with product-level methods.** A fair comparison needs a radiance-to-FRP conversion we have not validated and a common support; it is future work.

11 Conclusion

A small network trained on near-simultaneous Aqua–S-NPP pairs can map MODIS night radiance to VIIRS-like radiance without any fire product. On the selection set its Gamma-BW variants have the right amount and flag count and sharper placement than the references, and rank blocks slightly better than the deployable Even spread; none ranks them better than the oracle-background spread by the pre-stated margin, and the Group-share loss that does so overpaints. On the one held-out episode, with four scene pairs, the selected model keeps its total energy and sharpness but fails three of four pre-stated lines; selection-set scores need not carry over, and four pairs give no rate. Future work: recovery of fires below the detection threshold (a stratified re-score of the stored checkpoints), narrower and wider pairing windows with more data, more held-out regions and seeds, day scenes, and a head-to-head with product-level methods.

Statements

Data and code availability. Code, run records, score files and the sources of this paper are in private repositories; processed tiles, checkpoints and scores are in a private object store. Inputs are NASA open data and can be re-downloaded. All runs are identified by checkpoint hash in the repository, and every number in the text is generated from stored result files by scripts in the repository; whether the repositories become public is the author's decision.

Licences and acknowledgements. MODIS (MYD021KM, MYD03) and VIIRS (VNP02IMG, VNP03MOD) Level-1B data are NASA EOSDIS open data (LAADS DAAC, LP DAAC); the FIRMS archive is used under NASA ESDIS data-use guidance. Software: PyTorch (BSD-3-Clause), NumPy and SciPy (BSD), pandas (BSD), pyproj, netCDF4 and pyhdf (MIT), matplotlib (PSF-style); orbit tracks in Figure 1 use SGP4 propagation of two-line elements obtained through the IAU CPS SatChecker service (Space-Track data; licence not stated on the pages read); typesetting with Tectonic (MIT) and the arxiv-style template by George Kour (MIT). We acknowledge the use of data and imagery from NASA LANCE FIRMS. GPU compute was rented from RunPod.

Author contributions and use of SI tools. The author conceived the study, led the design ideation, made all design decisions and directed the work. Super Intelligence (SI, also known as AI) assistants (Claude, Anthropic) were used under the author's direction and supervision for writing code, implementing training, data analysis, figure generation, drafting and internal checks. The network is trained from scratch; no pre-trained model is used. The manuscript has not been peer reviewed, and the author takes full responsibility for its content.

Protocol deviations. A cross-check list of 589 tiles built for a second look turned out, on an independent review, to contain 88 tiles of the report set, at most 45 of them in the scored pack; aggregate scores of earlier runs on it were seen and had influenced an overnight preference for one setting (the queue of extra seeds was built with that winner in mind; both settings still received four seeds). The 88 tiles were retired, no number from that list is used anywhere, and the setting was decided again by the pick rule on selection scores only. The scorer now refuses report and retired tiles.

References

[1] Mukul Badhan, Kasra Shamsaei, Hamed Ebrahimian, George Bebis, Neil P. Lareau, and Eric Rowell. Deep Learning Approach to Improve Spatial Resolution of GOES-17 Wildfire Boundaries Using VIIRS Satellite Data. *Remote Sensing*, 16(4):715, 2024. doi:10.3390/rs16040715.

[2] Arindam Banerjee, Srujana Merugu, Inderjit S. Dhillon, and Joydeep Ghosh. Clustering with Bregman Divergences. *Journal of Machine Learning Research*, 6(58):1705–1749, 2005. URL `https://www.jmlr.org/papers/v6/banerjee05b.html`.

[3] Changyong Cao, Michael Weinreb, and Hui Xu. Predicting Simultaneous Nadir Overpasses among Polar-Orbiting Meteorological Satellites for the Intersatellite Calibration of Radiometers. *Journal of Atmospheric and Oceanic Technology*, 21(4):537–542, 2004. doi:10.1175/1520-0426(2004)021<0537:PSNOAP>2.0.CO;2.

[4] Gyanesh Chander, Tim J. Hewison, Nigel Fox, Xiangqian Wu, Xiaoxiong Xiong, and William J. Blackwell. Overview of Intercalibration of Satellite Instruments. *IEEE Transactions on Geoscience and Remote Sensing*, 51 (3):1056–1080, 2013. doi:10.1109/tgrs.2012.2228654.

[5] Haonan Chen and Changyong Cao. Super-Resolution of GOES ABI Data Based on VIIRS Measurements and Auxiliary Environmental Information. In *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, pages 3889–3892. IEEE, 2023. doi:10.1109/igarss52108.2023.10283330.

[6] Zuoqi Chen, Bailang Yu, Chengshu Yang, Yuyu Zhou, Shenjun Yao, Xingjian Qian, Congxiao Wang, Bin Wu, and Jianping Wu. An extended time series (20002018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth System Science Data*, 13(3):889–906, 2021. doi:10.5194/essd-13-889-2021.

[7] Ivan A. Csiszar and Wilfrid Schroeder. Short-term observations of the temporal development of active fires from consecutive same-day ETM+ and ASTER imagery in the Amazon: Implications for active fire product validation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(4):248–253, 2008. doi:10.1109/JSTARS.2008.2011377.

[8] D.K. Davies, S. Ilavajhala, Min Minnie Wong, and C.O. Justice. Fire Information for Resource Management System: Archiving and Distributing MODIS Active Fire Data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):72–79, 2009. doi:10.1109/tgrs.2008.2002076. URL https://doi.org/10.1109/tgrs.2008.2002076.

[9] Patrick H. Freeborn, Martin J. Wooster, David P. Roy, and Mark A. Cochrane. Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite-based active fire characterization and biomass burning estimation. *Geophysical Research Letters*, 41(6):1988–1994, 2014. doi:10.1002/2013gl059086. URL https://doi.org/10.1002/2013gl059086.

[10] Yuyun Fu, Rui Li, Xuewen Wang, Yves Bergeron, Osvaldo Valeria, Raphaël D. Chavardès, Yipu Wang, and Jiheng Hu. Fire Detection and Fire Radiative Power in Forests and Low-Biomass Lands in Northeast Asia: MODIS versus VIIRS Fire Products. *Remote Sensing*, 12(18):2870, 2020. doi:10.3390/rs12182870.

[11] Louis Giglio, Jacques Descloitres, Christopher O. Justice, and Yoram J. Kaufman. An Enhanced Contextual Fire Detection Algorithm for MODIS. *Remote Sensing of Environment*, 87(2-3):273–282, 2003. doi:10.1016/s0034-4257(03)00184-6. URL https://doi.org/10.1016/s0034-4257(03)00184-6.

[12] Louis Giglio, Wilfrid Schroeder, and Christopher O. Justice. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, 178:31–41, 2016. doi:10.1016/j.rse.2016.02.054. URL https://doi.org/10.1016/j.rse.2016.02.054.

[13] Tilmann Gneiting. Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, 106 (494):746–762, 2011. doi:10.1198/jasa.2011.r10138. URL https://doi.org/10.1198/jasa.2011.r10138.

[14] J A Hanley and B J McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982. doi:10.1148/radiology.143.1.7063747. URL https://doi.org/10.1148/radiology.143.1.7063747.

[15] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, 2017. doi:10.1109/CVPR.2017.632.

[16] Fangjun Li, Xiaoyang Zhang, Shobha Kondragunta, and Ivan Csiszar. Comparison of Fire Radiative Power Estimates From VIIRS and MODIS Observations. *Journal of Geophysical Research: Atmospheres*, 123(9): 4545–4563, 2018. doi:10.1029/2017jd027823. URL https://doi.org/10.1029/2017jd027823.

[17] Fangjun Li, Xiaoyang Zhang, and Shobha Kondragunta. Biomass Burning in Africa: An Investigation of Fire Radiative Power Missed by MODIS Using the 375 m VIIRS Active Fire Product. *Remote Sensing*, 12(10):1561, 2020. doi:10.3390/rs12101561. URL https://doi.org/10.3390/rs12101561.

[18] Guilin Liu, Fitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image Inpainting for Irregular Holes Using Partial Convolutions. In *Computer Vision – ECCV 2018*, Lecture Notes in Computer Science, pages 89–105. Springer International Publishing, 2018. doi:10.1007/978-3-030-01252-6_6. URL https://doi.org/10.1007/978-3-030-01252-6_6.

[19] MODIS Science Team. MYD021KM MODIS/Aqua Calibrated Radiances 5-Min L1B Swath 1km, 2017. URL https://www.earthdata.nasa.gov/data/catalog/laads-myd021km-6.1.

[20] NASA Earthdata. Fire Information for Resource Management System (FIRMS). https://www.earthdata.nasa.gov/data/tools/firms, 2026. URL https://www.earthdata.nasa.gov/data/tools/firms.

[21] NASA VIIRS Land Science Investigator Processing System (SIPS). Visible Infrared Imaging Radiometer Suite (VIIRS) 375 m & 750 m Active Fire Products: Product User's Guide, Version 1.4. https://lpdaac.usgs.gov/documents/427/VNP14_User_Guide_V1.pdf, 2018. URL https://lpdaac.usgs.gov/documents/427/VNP14_User_Guide_V1.pdf.

[22] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. FiLM: Visual Reasoning with a General Conditioning Layer. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 2018. doi:10.1609/aaai.v32i1.11671. URL https://doi.org/10.1609/aaai.v32i1.11671.

[23] Nigel M. Roberts and Humphrey W. Lean. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, 136(1):78–97, 2008. doi:10.1175/2007mwr2123.1. URL https://doi.org/10.1175/2007mwr2123.1.

[24] Wilfrid Schroeder, Patricia Oliva, Louis Giglio, and Ivan A. Csiszar. The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143:85–96, 2014. doi:10.1016/j.rse.2013.12.008. URL https://doi.org/10.1016/j.rse.2013.12.008.

[25] Mikhail Sofiev and Rostislav Kouznetsov. Observation operator and detection limits for MODIS and VIIRS Fire Radiative Power products. *Geoscientific Model Development*, 19(20):10123–10142, 2026. doi:10.5194/gmd-19-10123-2026. URL https://doi.org/10.5194/gmd-19-10123-2026.

[26] Sirish Uprety, Changyong Cao, Xiaoxiong Xiong, Slawomir Blonski, Aisheng Wu, and Xi Shao. Radiometric Intercomparison between Suomi-NPP VIIRS and Aqua MODIS Reflective Solar Bands Using Simultaneous Nadir Overpass in the Low Latitudes. *Journal of Atmospheric and Oceanic Technology*, 30(12):2720–2736, 2013. doi:10.1175/JTECH-D-13-00071.1.

[27] VIIRS Calibration Support Team (VCST). VIIRS/NPP Imagery Resolution Bands 6-Min L1B 6-Min Swath- 375m, 2021. URL https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/VNP02IMG/.

[28] Benjamin Leonardo Alves White. Quantitative differences in detection of fire pixels using Aqua/MODIS and NPP-Suomi/VIIRS satellites. *Journal of Hyperspectral Remote Sensing*, 12(6):379–390, 2023. doi:10.29150/jhrs.v12.6.p379-390. URL https://doi.org/10.29150/jhrs.v12.6.p379-390.

[29] Martin J. Wooster, B. Zhukov, and D. Oertel. Fire radiative energy for quantitative study of biomass burning: derivation from the BIRD experimental satellite and comparison to MODIS fire products. *Remote Sensing of Environment*, 86(1):83–107, 2003. doi:10.1016/s0034-4257(03)00070-1. URL https://doi.org/10.1016/s0034-4257(03)00070-1.

[30] Ryota Yagi, Mukul Badhan, Majid Bavandpour, Kasra Shamsaei, Dani Or, George Bebis, Neil P. Lareau, Qunying Huang, and Hamed Ebrahimian. Near-Real-Time Wildfire Segmentation and Brightness Temperature Estimation from Geostationary Satellite Imagery Using Deep Learning. *Fire*, 9(10):445, 2026. doi:10.3390/fire9100445.

[31] Naoto Yokoya, Takehisa Yairi, and Akira Iwasaki. Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):528–537, 2012. doi:10.1109/TGRS.2011.2161320.

[32] Ke Zheng, Lianru Gao, Wenzhi Liao, Danfeng Hong, Bing Zhang, Ximin Cui, and Jocelyn Chanussot. Coupled Convolutional Neural Network With Adaptive Response Function Learning for Unsupervised Hyperspectral Super Resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3):2487–2502, 2021. doi:10.1109/tgrs.2020.3006534.

[33] Yang Zheng, Ke Ding, Lian Xue, Zilin Wang, Guanjie Jiao, Yifan Zhu, Jinying Zhang, and Qianyu Ren. Fire Radiative Power Correction and Spatiotemporal Fusion Based on MYD14 and VNP14IMG. *Remote Sensing*, 18(10):1650, 2026. doi:10.3390/rs18101650.

A Appendix

A.1 Per-run results

Table 6 lists every SELECT score file used (block AUC with the fire output on and off, pixel energy ratio, flag-count ratio, hit rate, hottest-cell share, 3 km error, and whether L5 is met). Table 7 gives the full hyper-parameters.

A.2 Closest published work

A.3 Protocol order

A.4 Further pairing figures

A.5 Flag rule

The flag rule is a simplified MOD14/VNP14 contextual test on brightness temperatures T_4 and T_{11} (dynamic potential-fire thresholds of Collection 6; the contextual constants are the repository's). A cell is flagged when it is valid, passes the potential-fire gate (T_4 above a threshold equal to the background mean plus 5 K, clamped to 300–330 K at night,

Table 6: Per-run SELECT results.

Setting	Seed	AUC	AUC off	Energy	Flags	Hit	Hottest share	Err 3 km	Del.
Group-share (best-val)	0	0.927	0.922	1.38	2.70	0.90	0.26	0.53	–
Group-share (best-val)	1	0.928	0.920	1.51	2.65	0.88	0.23	0.59	–
Gamma (best-val)	0	0.871	0.820	0.99	0.98	0.58	0.27	0.46	–
Gamma (best-val)	1	0.861	0.775	0.88	1.02	0.54	0.24	0.46	–
Gamma-BW100 (best-val)	0	0.872	0.790	0.95	0.93	0.56	0.33	0.48	–
Gamma-BW100 (best-val)	1	0.873	0.817	1.06	1.02	0.62	0.34	0.43	Y
Gamma-BW100 (best-val)	2	0.879	0.765	1.00	0.99	0.60	0.32	0.44	–
Gamma-BW100 (best-val)	3	0.873	0.804	0.99	0.95	0.61	0.34	0.43	Y
Gamma-BW1000 (best-val)	0	0.876	0.792	0.97	1.00	0.63	0.35	0.44	–
Gamma-BW1000 (best-val)	1	0.872	0.833	0.97	0.98	0.58	0.32	0.45	–
Gamma-BW1000 (best-val)	2	0.871	0.793	1.03	1.02	0.62	0.36	0.43	Y
Gamma-BW1000 (best-val)	3	0.868	0.812	1.00	0.92	0.53	0.32	0.47	–
Gamma-BW100 (final)	0	0.869	0.817	1.06	0.95	0.63	0.43	0.43	–
Gamma-BW100 (final)	1	0.849	0.815	0.78	0.77	0.54	0.45	0.52	–
Gamma-BW1000 (final)	0	0.863	0.826	0.94	0.94	0.63	0.41	0.41	Y
Gamma-BW1000 (final)	1	0.859	0.803	1.09	0.96	0.61	0.42	0.45	–
Gamma-BW100-B0.5 (best-val)	0	0.877	0.807	0.99	0.92	0.59	0.30	0.44	–
Gamma-BW100-B0.5 (best-val)	1	0.882	0.827	1.01	1.02	0.60	0.28	0.43	–
Gamma-BW1000-B0.5 (best-val)	0	0.886	0.842	1.04	1.07	0.62	0.30	0.44	–
Gamma-BW1000-B0.5 (best-val)	1	0.873	0.840	0.99	0.98	0.58	0.28	0.42	–
Gamma-BW100-B0.5 (final)	0	0.861	0.800	1.01	0.95	0.62	0.40	0.44	–
Gamma-BW100-B0.5 (final)	1	0.864	0.800	0.92	0.87	0.59	0.39	0.46	–
Gamma-BW1000-B0.5 (final)	0	0.881	0.836	1.20	1.04	0.68	0.41	0.42	–
Gamma-BW1000-B0.5 (final)	1	0.870	0.771	0.85	0.82	0.58	0.39	0.47	–
Gamma (small data)	0	0.861	0.772	1.25	1.22	0.55	0.23	0.49	–
Gamma (small data)	1	0.866	0.749	1.06	1.03	0.41	0.19	0.54	–
Group-share (small data)	0	0.920	0.906	1.46	3.61	0.92	0.23	0.56	–
Group-share (small data)	1	0.921	0.926	1.41	2.80	0.90	0.26	0.52	–

Table 7: Hyper-parameters of the large-data runs (run record of one Gamma-BW1000 run; other settings differ only in loss family, weight and bound).

Setting	Value
Optimiser steps per epoch	647
Epochs (cap)	46
Batch size (tiles)	32
Peak learning rate	0.002
Weight decay	0.0001
Warm-up steps	100
Learning-rate schedule	constant after warm-up
Fire tiles per batch	0.25
Early stopping	none (fixed 46 epochs)
Planck band constants	centre
Gradient-norm clip	1
Trunk precision	bf16
Tile size (cells)	96×96
Network width / latent channels	32 / 16
Background bound $\delta_{\max}$	0.25
Max excess log-ratio $s_{\max}$	8
Loss term weights mv / vm / mm / vv	1 / 0.5 / 0.1 / 0.1
Energy-consistency weight	0.1
Hiding window / tiles hidden	9×9 / p=0.5
Median background window / dilation	7×7 / 2
MODIS view kernel / T_{fire}	tri / 800 K
Minimum valid fraction per tile	0.1
Day/night	night
Hardware per run	one NVIDIA RTX 4090

and $T_4 - T_{11}$ above one clamped to 10–35 K), and either passes the absolute test ($T_4 \geq 320$ K at night) or passes the contextual tests: T_4 above the local background mean plus three standard deviations, $T_4 - T_{11}$ above the background mean plus 3.5 standard deviations and at least 6 K, and T_{11} above its background mean plus one standard deviation minus 4 K. The background is a 7×7-tap window (spacing 8 cells for MODIS and 2 for VIIRS) with at least 8 valid cells.

Table 8: Model variants and display names. B0.5 marks the background bound $\delta_{\max} = 0.5$ (default 0.25). Best-validation and final-epoch checkpoints share a name; run identifiers appear only here and in the appendix.

Name	Loss	BW	Bound	Run id (prefix)	Seeds
Group-share	group-share (Huber on ln radiance, fixed shares for background, near-fire and fire-level cells from the target)	0	0.25	`oldloss-nf2`	0,1
Gamma	Gamma deviance	0	0.25	`gamma-nf2`	0,1
Gamma-BW100	Gamma deviance	100	0.25	`gammabw100-nf2`	0,1,2,3
Gamma-BW1000	Gamma deviance	1000	0.25	`gammabw1000-nf2`	0,1,2,3
Gamma-BW100-B0.5	Gamma deviance	100	0.5	`gammabw100-d05-nf2`	0,1
Gamma-BW1000-B0.5	Gamma deviance	1000	0.5	`gammabw1000-d05-nf2`	0,1

Table 9: Closest published work, read from the full texts. "Product" means the active-fire products (detections with fire radiative power, FRP); none of the MODIS–VIIRS fire papers reads radiance. †: abstract only, full text not openly available. Numbers are the authors' own and are not comparable across rows: quantities, references and test sets differ.

Work	From → to	Pairs (time window)	Region, period	Method; held-out	Authors' headline result
Product-level MODIS–VIIRS harmonization (statistical)					
Zheng et al. [33]	MYD14 FRP → VIIRS-equivalent FRP (MW) per MODIS fire pixel; fused daily record	VIIRS fire pixels inside the MODIS footprint, $\leq$ 5 min; 7.8×10^4 validation pairs (10%)	global, 2012–2024; day/night not stated	one quadratic in view zenith (factor 1.19–1.52); random 90/10 split	pixel slope $0.74 \rightarrow 1.03$, RMSE $113 \rightarrow 93$ MW, R $0.79 \rightarrow 0.81$
Li et al. [17]	MYD14 grid FRP → VIIRS-level grid FRP (0.05°–5°)	grids with both sensors, $\pm$5 min, 44 days	Africa, 2017, day	quadratic in view zenith per grid size; none	MODIS continental FRP about 49% below VIIRS
Sofiev and Kouznetsov [25]	FRP → FRP each sensor would report (observation operators)	overlapping granules, no window (1–2 h apart)	global, 2000–2024, day and night	fitted detection curves vs. pixel area; none	2020 totals 128 and 177 TW become 88 and 106 TW
Li et al. [16]	comparison only (VIIRS 750 m product)	$\pm$5 min granules; 619 clusters within 20 min	global, 2016–2017	regression; n/a	VIIRS FRP about 20% larger in fire-prone 1° cells
Fu et al. [10]	comparison only	VIIRS pixels in the MODIS footprint, $\leq$ 5 min	Northeast Asia, 2012–2017, day	regression; n/a	65% / 83% of VIIRS fire pixels missed by MODIS
Learned geostationary → VIIRS					
Badhan et al. [1]	GOES-17 band 7 (2 km) → rasterised VIIRS fire product: mask and brightness temperature, 375 m	within 2.5 min; 5,869 windows	western USA, 2019–2021	autoencoder; random split of windows	IoU 0.14
Yagi et al. [30]	GOES bands 7, 14, 15 → rasterised VIIRS fire product: mask and brightness temperature	$\pm$5 min; 257 fires, 3,242 snapshots	USA, 2018–2024	two U-Nets, 3 seeds; split by fire	IoU 0.155 (GOES product 0.109); fire-region RMSE 22.7 K
Chen and Cao [5]†	GOES bands 2, 7, 14, 15 → VIIRS I1–I5, 375 m	"concurrent overlapping"; not stated	not stated	ESRGAN; not stated	none in the abstract
Learned older sensor → VIIRS-like record (night lights, not fire)					
Chen et al. [6]	DMSP night lights + vegetation index → VIIRS-like annual night lights	same calendar year (2013)	global, 2000–2012	autoencoder; held-out year (2012)	pixel R^2 0.87
This work	Aqua MODIS 4 and 11 μm radiance → VIIRS-like radiance, 375 m cells; no fire product	same ground, $\leq$ 30 min; 1,017 scene pairs	nine regions, night, 2012–2022	small network with fixed sensor views; 100 km blocks; fire-off control	see Section 8; selection-set numbers only

A.6 Earlier latent-pull trials

Before the decoded-scene formulation, 11 runs of an earlier, counts-based trial pulled the latent codes of the two sensors together; none gave per-cell agreement (the representation collapsed), and the trial was stopped. The records of those runs are not part of this paper's result files, so no collapse measure is reported here.

A.7 Literature search

The statement in Section 1 rests on a search made on 15 November 2026 by an SI agent under the author's direction. Works were located by web search and fetched as DOI records (Crossref, DataCite) with the deposited abstract; full

Table 10: Order of protocol steps relative to the scores they could influence.

	Step
1	Raw-signal formulation fixed and reviewed; staged workflow written down.
2	Stage-0 pass lines recorded before any model of the final workflow was trained.
3	Memorise criterion revised after the first gate result, original numbers kept.
4	Five-line decision table written after the first scores (seeds 0 and 1 of each setting).
5	Cross-check list dropped and the 88 report tiles retired; pick rule fixed before seeds 2 and 3 were scored.
6	Held-out lines approved and the candidate frozen; one held-out scoring.

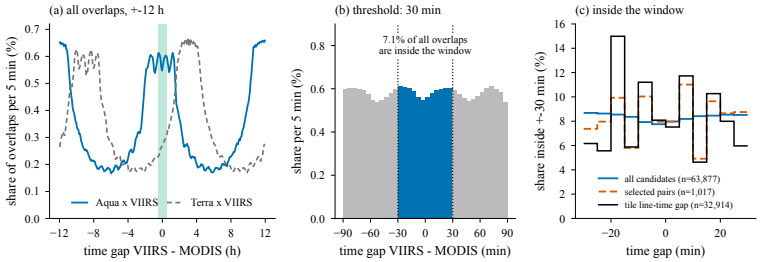


Figure 4: Time gaps of all Aqua–VIIRS overlaps, the 30-minute window, and the selected pairs and tiles inside it.

texts were read for 8 of the 11 closest works. Limits: Google Scholar was not queried directly, Scopus and Web of Science were not searched, non-English literature and patents were not searched. The search was not exhaustive.

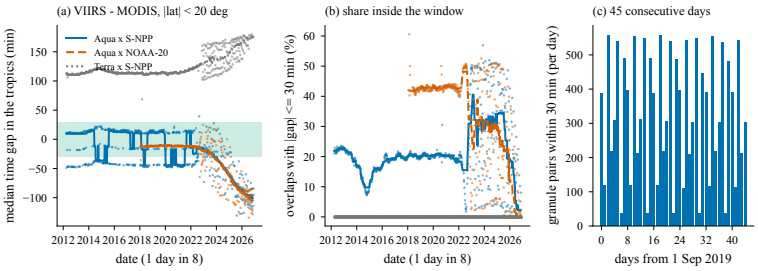


Figure 5: Time gap and share of overlaps inside the window by date: the window drifts with Aqua's orbit and near-simultaneous overlaps recur every 8 days.