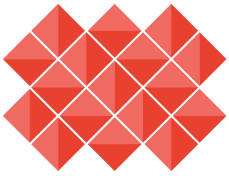

MAPBIOMAS
[FUEGO]

MapBiomas Argentina Fuego

ATBD

Algorithm Theoretical Basis Document

Collection 1

Version 1.0

General coordination (MapBiomas Argentina)

Ana Eljall

Technical coordination (MapBiomas Fuego)

Iván Barberá

Camilo Bagnato

Team

Lican Martínez
Ramón Peña Agrest
Gonzalo Dieguez Gaviola
Augusto Van Der Ploeg
Cecilia Evequoz
Jimena Albornoz
Daniela Arpigiani
Leónidas Lizárraga
Andrés Leszczuk
Camila Mateo
Juan Argañaraz
Jaime Moyano
Luna Schteingart
Mariana Lipori
Roxana Gimenez
Fernando Salvaré
Juan Gowda
José Daniel Lencinas
Clara Rey
Romina Ledesma
Carla Sauval
Santiago Verón
Thomas Kitzberger
Pablo Baldassini

September 2026

Contents

1 Introduction

1.1 Scope and purpose of this document

This document describes the theoretical basis, the methodological framework and the operational considerations behind Collection 1 of MapBiomas Argentina Fuego: an annual, 30 m burned-area product covering the whole of Argentina for the period 1999–2025.

This ATBD is conceptual, and it is deliberately not a complete specification of the algorithm. Every stage described below carries many decisions, but writing all of them down here would produce a document that nobody would read. What follows is the reasoning: what each stage measures, why it measures it that way, and what it hands to the next stage. Readers who want to reproduce the algorithm elsewhere, or who need the exact value of a parameter, should go to the per-step technical notes in the project repository, at `collection-01/docs` (`https://github.com/mapbiomas/argentina-fire`). There is one note per workflow step, numbered to match it, and each one points onward to the production code, the configuration files and the analysis notebooks that justify its choices. We cite those documents here as "see `docs/NN`".

Validation and the statistical characterisation of the product are outside the scope of this document. They are separate activities that consume the finished map rather than being stages of it, and they are documented on their own.

1.2 Objectives and improvements in Collection 1

In Collection 0 (the pilot) we assembled a working group capable of producing burned-area maps for MapBiomas Argentina, and developed a mapping algorithm of reasonable accuracy with a relatively low demand of human effort for training-data collection. We produced fire maps for a fraction of the country, in NW Patagonia. Collection 1 aimed at extending the mapping to the entire national territory, still covering the 1999-2025 period, and improving the mapping algorithm in the aspects that were known to be problematic. In particular, we used more complex models in the spectral analysis, changed the time-series metrics in the temporal analysis, used a fire-year-based spatial analysis to avoid splitting fires that cross calendar-years, and applied a model to separate fire from non-fire polygons (replacing an empirical decision tree).

1.3 Input datasets

Collection 1 was developed on complete Landsat time series plus two ancillary MapBiomas Argentina layers, processed almost entirely in Google Earth Engine (GEE; Gorelick et al. 2017).

1.3.1 Landsat imagery

All available Landsat Collection 2 Level-2 (surface reflectance) scenes at 30 m intersecting the territory and the period were used (U.S. Geological Survey 2021b). During the mapped period the active sensors were Landsat 5 TM (until 2013), Landsat 7 ETM+ (from mid-1999), Landsat 8 OLI (from 2013) and Landsat 9 OLI-2 (from 2021). Six reflective bands are read: blue, green, red, near infrared (NIR) and the two shortwave infrared bands (SWIR1, SWIR2). A set of spectral indices is derived from them (Section 3.6).

Scenes acquired on the same date are collapsed into a single date-mosaic before anything else happens, so that the unit the algorithm reasons about is a date, not a scene footprint.

1.3.2 Quality filters

Landsat observations are masked with the QA_PIXEL band delivered with the Collection 2 surface reflectance products, excluding cloud, dilated cloud, cloud shadow, snow and water.

No temporal interpolation and no spatial gap-filling are applied. A missing observation stays missing. This is a deliberate constraint rather than an omission: the temporal stage (Section 4) reads real acquisition dates and counts real observations. The direct consequence is that observation density varies in space and in time, strongly so before 2000, when only Landsat 5 and a newly launched Landsat 7 were flying. The algorithm carries that density forward as an explicit quality channel.

1.3.3 Spectral harmonisation between sensors

Landsat 5 TM and Landsat 7 ETM+ reflectances were transformed into the Landsat 8/9 OLI domain band by band, using the coefficients of Roy et al. (2016). OLI and OLI-2 observations were left untouched. This reduces artificial discontinuities at sensor transitions, which in a change detector would otherwise be read as change.

1.3.4 MapBiomas Argentina land cover

The annual land use and land cover (LULC) classification of MapBiomas Argentina (MapBiomas Argentina 2025) is used for two distinct purposes, and both read the previous year ($y - 1$) relative to the year being mapped:

1. it defines the fire class (`veg_fire`) of every pixel, which selects which burn-probability model judges that pixel, and which surfaces are excluded from mapping altogether as non-burnable (Section 3.4);
2. it supplies the object-level vegetation composition metrics used by the object classifier (Section 6.2).

We use $y - 1$ because the land cover of a burned pixel changes because it burned, so using the focal year's classification would let the fire we are trying to detect decide how it is detected. The same rule is applied everywhere the classification is consumed.

1.3.5 MapBiomas annual mosaic

The second ancillary layer is the MapBiomas Argentina annual mosaic: the reflectance summary that the land-cover initiative itself classifies. Forty of its bands are used: the six optical bands plus NDVI, NDWI, NPV and NDFI (Souza Jr., Roberts, and Cochrane 2005), each summarised as annual median, dry-season median, wet-season median and standard deviation. Again, the mosaic of the previous year is attached to every observation.

This layer is used to allow for the spectral information of a focal date to be interpreted in the context of what the vegetation characteristics were in the previous year. A burn scar is not a spectral value alone: we assign burn probability based on its spectral data but considering what it was in the previous year, from the land cover class and from this mosaic. The mosaic informs within-class variability in vegetation structure and moisture.

1.4 Computing platforms

Most of the pipeline runs in GEE, which stores the entire Landsat archive and the MapBiomas products and provides the compute to evaluate a model over them. Two stages run locally: the fitting of the burn-probability models, in R (R Core Team 2025) with `glmnet` (Friedman, Hastie, and Tibshirani 2010), and the entire object stage, that is vectorisation, object metrics and the object classifier, also in R, on tabular and raster data downloaded from GEE.

1.5 The processing grid

Every image-based step runs over the MapBiomas Argentina working grid, the national 1:250,000 chart series, locally called *cartas*. Each sheet covers roughly 14,000 km^2, and 248 of them intersect the buffered country. Tiling by chart sheet rather than by Landsat WRS-2 path/row is what lets one step read another step's tile and assemble the results on a single lattice with no resampling anywhere in the chain. All exports pin the projection and the affine transform explicitly, never a nominal scale, because in geographic coordinates a nominal 30 m scale is a different grid.

2 The burned-area mapping algorithm

The algorithm detects burned area at the finest temporal and spatial resolution Landsat allows, one 30 m pixel on one observation date, by exploiting the spectral, temporal and spatial signature of fire, in that order. Each stage consumes what the previous one produced and adds a kind of evidence the previous one could not see (Fig. 1).

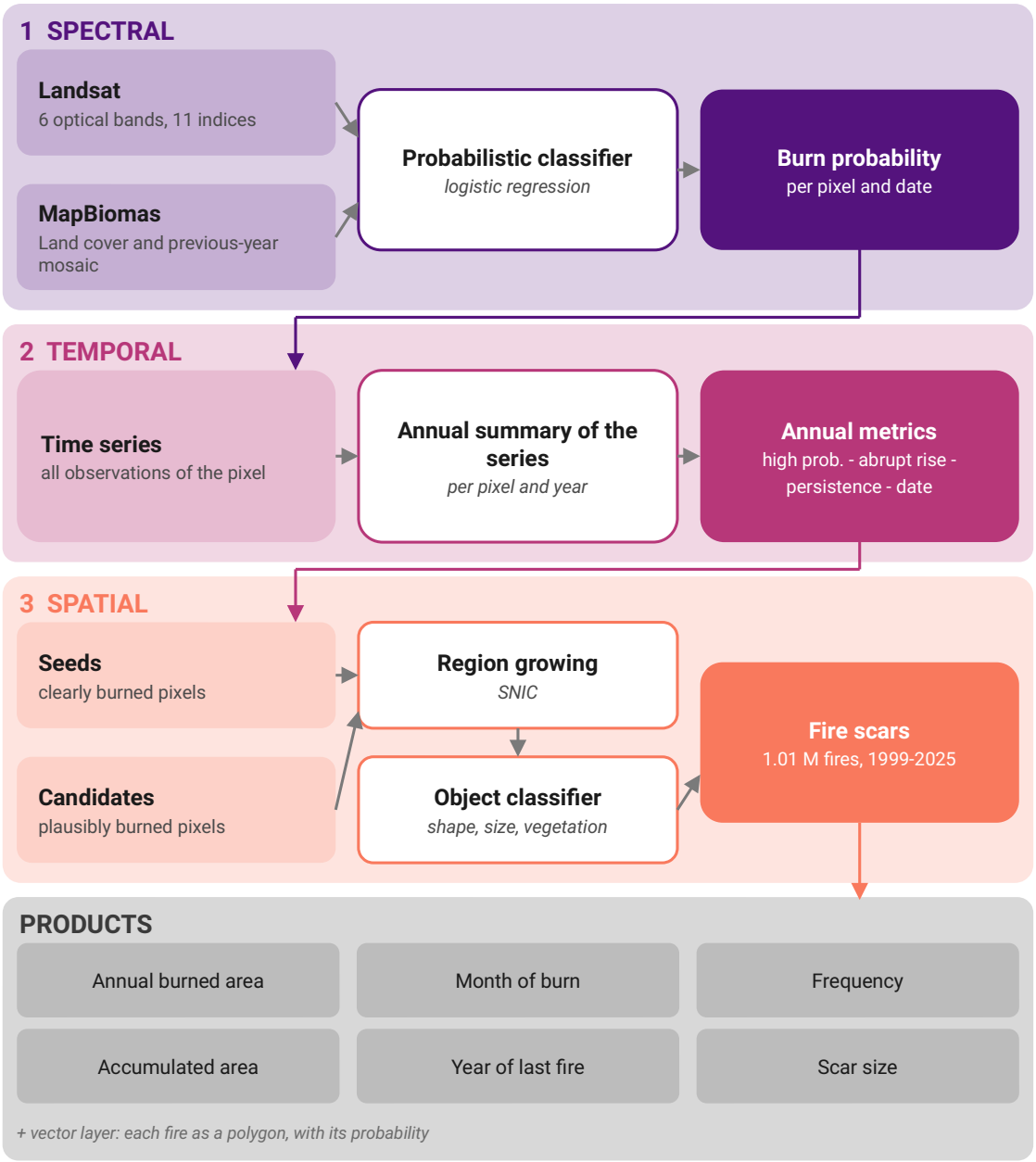

Figure 1: The Collection 1 algorithm. Spectral: Landsat bands and indices, plus MapBiomas land cover and the previous-year mosaic, feed a logistic-regression classifier that scores each pixel and date with a burn probability. Temporal: the per-pixel time series of that probability is summarised into annual metrics, capturing how high it rises, how abruptly, how long it persists, and when. Spatial: the metrics separate clearly burned seeds from plausibly burned candidates, which SNIC region-growing connects into segments; an object classifier filters those on shape, size and vegetation to yield the 1.01 M fire scars mapped over 1999–2025. The six published subproducts and the fire-object polygon layer are both built from that same set of scars.

Spectral: is this observation burned? A probabilistic classifier scores a single pixel on a single date, from its own reflectance and indices plus previous-year context. The output is a burn probability per observation.

Temporal: was this pixel burned this year, and when? High burn probability on one date is not necessarily fire. Cloud shadow and dark surfaces can raise it, and it can easily happen when the burn probability model is not trained on that rare conditions. In addition, low-productivity areas may look like burned, but they are steadily that way, not presenting change. Hence, what helps separate these cases is the time-series of burn probability. It rises, and it stays high. The second stage quantifies, per pixel and per calendar year, a maintained increase in burn probability, together with the date on which it happened.

Spatial: is this pixel part of a fire, and is this change-polygon fire? Fires are patches, not independent pixels. The third stage brings that context in twice: a region-growing algorithm converts per-pixel evidence into scar objects, and a classifier then judges each object on properties no pixel has, such as shape, size and vegetation composition.

A unifying idea in our approach is that quantitative information is preserved for as long as possible, so that every later step can weigh the earlier evidence instead of inheriting a hard classification. The temporal stage analyses a burn probability, not a hard burned/unburned class. The region-growing step judges pixels on numeric metrics, which is the only reason the seed/candidate distinction can be expressed at all. The object classifier consumes continuous metrics and returns a probability of fire, which is used to filter but is also published, so a user can apply a stricter or a looser threshold than ours.

3 Spectral stage: burn probability per observation

3.1 Foundations

The unit of judgement is the observation: a pixel-date. The model asks, for a single pixel on a single date, how likely it is that this pixel is burned, using spectral and land cover information only. Every usable Landsat observation is evaluated, so a fire is detected as a change in the series rather than as a signature in a summary image. The approach follows Long et al. (2019), where all Landsat observations are scored.

This is the opposite end of the design space from MapBiomas Fire Brazil (Alencar et al. 2022), which classifies an annual mosaic, the same way land cover is mapped everywhere in the network. Classifying a mosaic is cheaper, but it loses the temporal information, which requires more data and more flexible and expensive models to separate burned from unburned land.

3.2 Training data

Training points were collected interactively by domain experts, based mostly on fire events and areas similar to fire spectrally, to train the model on the false positive cases. For each selected fire the collector delimits a pre-fire and a post-fire period on the Landsat record, and then places burned and unburned points inside the scar and its surroundings, aiming to cover as much of the spectral variability of each class as possible and inspecting index time series wherever the classification is doubtful.

Collection 1 extends this protocol from 30 fires in one region to 198 fires across the five MapBiomas regions (13 in Atlantic Forest, 30 in Cuyo, 37 in Pampas, 47 in Patagonia, 71 in Chaco), holding 147,222 points. The full Landsat time series is then extracted at every point, which turns those points into 6,177,098 observations. Labels are assigned by window: observations of a burned point after the fire are labelled burned; observations of a burned point before the fire, and all observations of unburned points are labelled unburned.

A few of the sampled events are not "fires" but drought-affected steppe, volcanic-ash deposits or cropland harvest areas. These are events that look like fire in some indices and were the largest source of commission error, both in the pilot and in Collection 1. They enter the training set as unburned-only, and dropping them would remove the hardest negatives the model ever sees.

3.3 The cleaning gate

Observations were cleaned before the models were fit. We visually inspected the NBR and NBR2 time-series of all points (*spaghetti plots*) per event and further delineated pre- and post-fire periods to keep clear signal only. Of the 6,177,098 extracted observations, 5,923,062 pass the gate, of which 541,789 (9.1 %) are labelled burned.

3.4 Vegetation classes for fire

A separate model is fitted per vegetation type, because the effect of fire on the spectral properties varies among them. The unit of stratification is the `veg_fire` class, obtained by crossing the MapBiomas region with the MapBiomas LULC class of the previous year and remapping the result through a single lookup table. The table yields 23 fittable classes plus two non-fittable ones: non-burnable (water, rock, ice, urban, bare) and non-observed. Only the 23 are ever predicted, while the others are masked.

The class names carry both the vegetation type and the region it belongs to: `forest_pat`, `grassland_pampa`, `shrubland-open_chaco`, `agriculture_chaco`, and so on. The grouping is not purely ecological, because data availability was a hard constraint. A strictly within-region remap was considered first, but several classes are better off crossing regions: partly for spectral and land-use similarity, partly because data collection was organised by region and that did not guarantee that every class would be represented in every region. So a class may span regions, as `agriculture_cuyo-pat` covers both Cuyo and Patagonia. The remap is an open path for improvement: with more collected fires, groups currently merged for want of data could be split.

3.5 Model structure

As the burn probability model at the observation model had to compute predictions over the whole Landsat archive over the study area and period, it had to be deployed inside GEE (otherwise, it would require exporting all images, in a similar strategy as MapBiomas Brazil does with annual mosaics; Alencar et al. 2022). In addition, we needed a probabilistic classifier, so burn probability and not a hard burned/unburned class could be analysed as a function of time. Hence, we used a logistic regression. The fitted model is a set of coefficients plus a simple equation: trivial to store, to version and to ship as a CSV, and cheap to evaluate over billions of pixel-dates. The richer classifiers GEE offers natively, random forest and boosted trees, fail at both ends here. It is doubtful they could be fitted on this many training observations, and a fitted one cannot be saved as an asset, which is precisely what deploying over the whole Landsat archive requires. External machine-learning services would lift both limits at a large step up in complexity and at a monetary cost to the initiative.

The probability that the observation of pixel c at time t corresponds to burned is

$$\begin{aligned} \Pr(y_{c,t} = 1) &= \frac{1}{1 + \exp(-\eta_{c,t})}, \\ \eta_{c,t} &= \alpha_{k[c]} + \sum_{j=1}^{J} \beta_{j,k[c]}\, x_{j,c,t}, \end{aligned} \tag{1}$$

where α and β are estimated parameters, $k \in \{1, \ldots, 23\}$ is the `veg_fire` class, and the notation $k[c]$ indicates the class of pixel c in the year previous to the one time t belongs to. $J = 51$ is the number of regression terms in the deployed model.

All classes share the same predictor set, chosen globally rather than per class: only the coefficients differ between vegetation types. That keeps deployment cheap in the same way the model itself does. The prediction pipeline builds one band set once and reuses it for all 23 classes, instead of assembling a different design per vegetation type and paying for it at every pixel-date.

3.6 Predictors

Twenty-one spectral features are computed for every Landsat observation. Six of them are the optical bands; the other fifteen are indices: NBR (Key and Benson 2006), NBR2 (U.S. Geological Survey 2021a), MIRBI (Trigg and Flasse 2001), NDVI (Tucker 1979), the three tasselled-cap components (Baig et al. 2014), NDMI (Gao 1996), NDSI (Hall, Riggs, and Salomonson 1995), SAVI (Huete 1988), NDWI (McFeeters 1996), AFRI (Karnieli et al. 2001), kNDVI (Camps-Valls et al. 2021), EVI2 (Jiang et al. 2008) and NIRv (Badgley, Field, and Berry 2017). Forty previous-year mosaic bands are available alongside them.

Four fire-sensitive indices are worth writing out, since they are the ones the pilot used and the ones that still carry most of the fire signal:

$$\mathrm{NBR} = \frac{\mathrm{NIR} - \mathrm{SWIR2}}{\mathrm{NIR} + \mathrm{SWIR2}}, \quad \mathrm{NBR2} = \frac{\mathrm{SWIR1} - \mathrm{SWIR2}}{\mathrm{SWIR1} + \mathrm{SWIR2}},$$
$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}}, \qquad \mathrm{NDMI} = \frac{\mathrm{NIR} - \mathrm{SWIR1}}{\mathrm{NIR} + \mathrm{SWIR1}}.$$

The candidate design was much larger than the deployed one. It began as several hundred terms: all focal features, all mosaic summaries and a large set of interactions. It was reduced in two steps, first to 129 terms and then, by ranking those terms globally and cutting the ranking, to the deployed set. Predictive skill is flat from the full fit down to about fifty terms and falls below that, so the smallest set on the flat part was deployed. The reduction is not cosmetic: the number of terms is the dominant cost of the prediction stage, since each term becomes one image band that must be built and multiplied at every pixel-date. The deployed set costs about 25 % less compute per tile than the 129-term one.

The deployed model carries an intercept plus 51 terms, in five blocks:

- 10 focal main effects: GREEN, RED, NIR, SWIR1, SWIR2, NBR, NBR2, NDVI, NDMI, NDSI, all measured on the observation being scored;
- 14 previous-year mosaic main effects: medians, dry/wet composites and standard deviations of the optical bands and of NDVI, NDWI, NPV and NDFI;
- 10 focal $\times$ focal interactions, for example NBR $\times$ NBR2 and SWIR1 $\times$ NDVI;
- 4 same-band interactions: a focal band with its own previous-year median (GREEN, NIR, SWIR2, NDVI), which is the model's explicit change term, the product of today's value with this pixel's baseline;
- 13 cross interactions between previous-year mosaic summaries and focal fire indices or focal bands.

Fitting a separate model per `veg_fire` class handles the between-type differences; the interactions with the previous-year baseline handle the within-type gradient in productivity, structure and seasonality.

3.7 Fitting, regularisation and cross-validation

The design is strongly collinear by construction, since indices share bands and interactions share factors, so the models are fitted by elastic net (Zou and Hastie 2005) with `glmnet` (Friedman, Hastie, and Tibshirani 2010), maximising the penalised Bernoulli likelihood. The mixing parameter α is tuned over $\{0.25, 0.5, 0.75\}$; pure ridge and pure lasso were dropped after never winning. λ is selected at the cross-validated minimum.

Cross-validation holds out whole fires. Observations from a single fire are strongly correlated in space and in time; splitting them across folds would report a skill the model does not have on a new fire. Folds are therefore grouped on a region-unique fire identifier, with stratified packing so that folds carry comparable numbers of positives, and K is adaptive: $K = \min(10, \text{fires with positives})$. Twenty-one of the 23 classes reach $K = 10$; two fit at 7 and 6. Out-of-fold predictions are stored per observation and are what the per-class diagnostics are computed on: calibration, omission and commission, by-fire breakdowns.

3.8 Deployment in GEE

The fitted models reach GEE as one small CSV of coefficients per class. Turning 23 such tables into a prediction over a whole country raises one design question, and the answer shapes the rest of the pipeline.

The obvious implementation builds 23 parallel branches of the same computation graph: evaluate each class's model on its own masked subset, then mosaic the results. Instead, the class selection is pushed into the data. Every term becomes one image band whose per-pixel value is that pixel's class's coefficient for that term, written by remapping the `veg_fire` image through the 23 coefficient values. A pixel of class 7 then carries class 7's coefficient for every term, and its neighbour of class 12 carries class 12's, in the same band. Prediction becomes a single arithmetic expression, identical everywhere, and choosing the model costs nothing at evaluation time. This is also why the predictor set has to be shared across classes in the first place.

The per-observation probability is never written to an asset. Over a padded year a carta tile is covered by roughly 150 dates, so storing the probability per observation would mean about 150 layers per tile-year where the annual summary that follows has 16 bands, an order of magnitude more data than a product that is already on the order of a terabyte. Nothing downstream requires a single observation's probability: what carries the evidence is the shape of the series. Since the model is a coefficient set plus a dot product, recomputing it inside the graph is far cheaper than writing it out and reading it back. The spectral and temporal stages are therefore two concepts but one computational step.

4 Temporal stage: burn-probability time-series metrics

4.1 Foundations

For every focal year and every carta, this stage reads the burn probability of every Landsat observation of every pixel and reduces that per-pixel series to a 16-band annual summary: how high the probability got, how persistently it stayed there, how sharply it jumped, when, and on how many observations all was computed. It never decides that a pixel burned. It summarises the evidence that the spatial stage classifies.

This is the fire team's annual mosaic, except that it is a prediction, not an aggregation. Every MapBiomas initiative reduces the raw Landsat archive to one annual layer per tile and builds everything downstream on it; for land cover that layer is the annual mosaic, a summary of reflectance. For the fire team it is a set of bands from the date of minimum NBR in a year. This layer is the equivalent of the MapBiomas Argentina Fire team, one level up, with a difference that decides a great deal: the model has already been applied, so what gets annualised are burn probability time-series metrics.

That provides more information density, since the vegetation-specific model, the previous-year context and the cloud masking are all resolved once, and no later stage ever touches a Landsat scene again. However, it also costs a dependency the plain mosaic does not have: this layer inherits the model, its training data and the land-cover collection that defines `veg_fire`, so improving any of the three means recomputing it, which is the most expensive step of the whole algorithm.

Burn probability has to be smoothed before it is read as detection. A single high observation is as likely to be shadow, ash or a wet scene as fire. As different vegetation types recover at very different speeds and image density varies by pixel and by year, the stage carries two temporal windows for aggregation, $K = 3$ and $K = 2$, and exports both, leaving the choice between them to the next stage, per pixel.

The burn evidence is three properties: magnitude, persistence and change. All three are measured in numbers of valid observations and never in elapsed time, because the Landsat series is irregular and its density is itself informative.

4.2 The burn probability jump

Let $\{p_t\}$ be the series of burn probabilities of a pixel, ordered by date, and let d_t be the day number of observation t counted from 1 January of the focal year, so that padding observations from the previous year are negative, those from the next year exceed 365, and every difference is an exact whole-day count across the year boundary.

For a window of K observations, we define

$$\begin{aligned} \mathrm{minfore}_K[t] &= \min\left(p_t, \ldots, p_{t+K-1}\right), \\ \mathrm{maxback}_K[t] &= \max\left(p_{t-K}, \ldots, p_{t-1}\right), \\ \mathrm{delta}_K[t] &= \mathrm{minfore}_K[t] - \mathrm{maxback}_K[t]. \end{aligned} \tag{2}$$

The asymmetry between the two operators is the core of the method. Taking the minimum forward demands that the high signal be sustained: a single isolated spike does not qualify, because one low value among the next K observations collapses the floor. Taking the maximum backward demands that the pixel was cleanly low before: one high previous value is enough to cancel the jump, which is what stops a noisy low observation inside an already-burned scar from manufacturing a second detection of the same fire. delta_K is therefore the most conservative reading of a step that the series admits.

Alongside the jump, the stage records the time structure around it (Fig. 2): $\mathrm{jumpgap}$, the number of days between the last low observation and the first high one; $\mathrm{prevwidth}$ and $\mathrm{postwidth}$, the spans in days of the back and fore windows; and $\mathrm{date\_post}$, the day of year of the post-jump observation. $\mathrm{jumpgap}$ is strong fire evidence when short. A long one may be slow change, such as drought, dieback or harvest, rather than fire, or a fire seen through a thin image series. The two widths report how much calendar time the windows actually cover, which under sparse imagery can be most of a season.

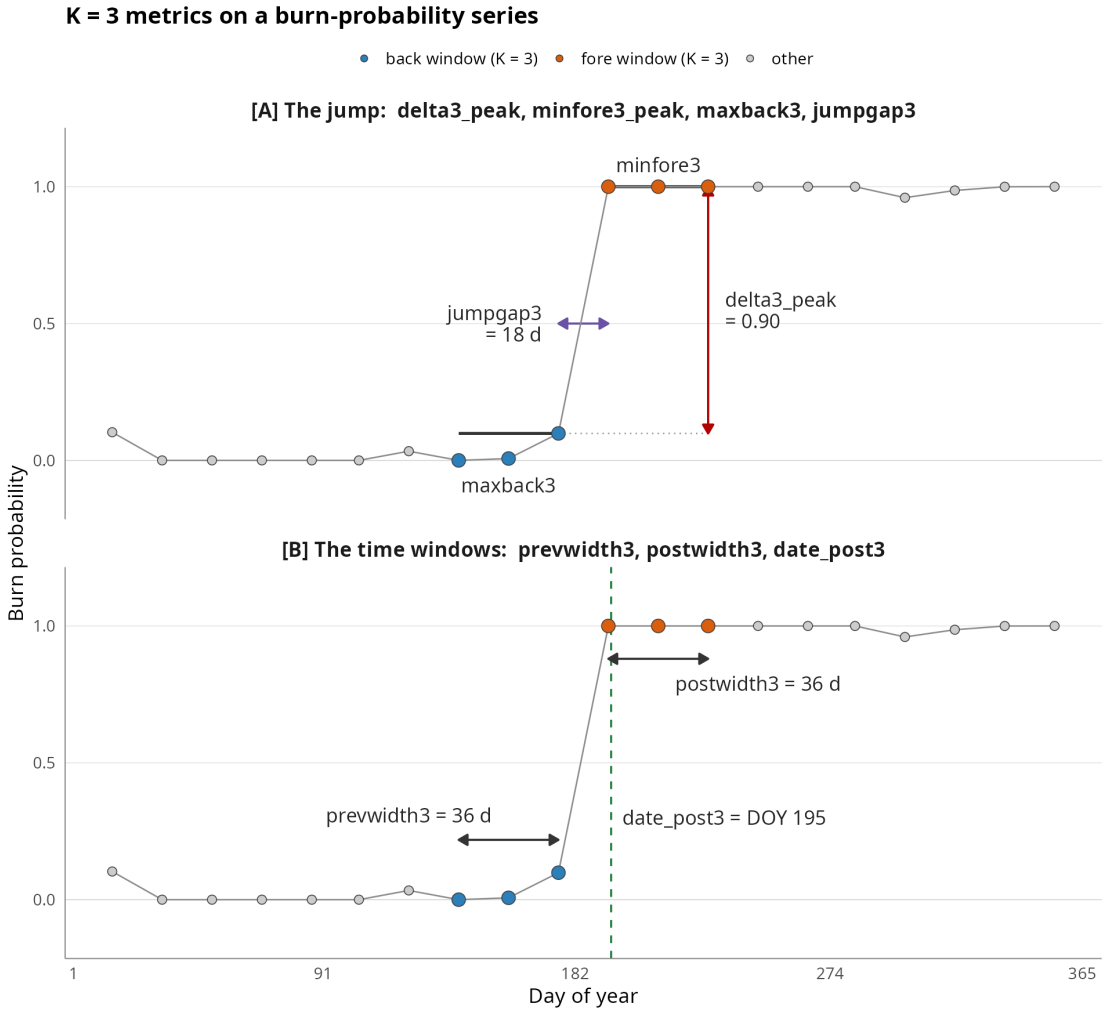

Figure 2: The $K = 3$ time-series metrics on a synthetic burn-probability series, which is the same over (A) and (B). (A) The jump. The black segment over the three back observations marks $\mathrm{maxback}_3$ (their maximum); the black segment over the three fore observations marks $\mathrm{minfore}_3$ (their minimum); the red arrow is delta_3, the magnitude of the sustained jump, and the purple arrow is $\mathrm{jumpgap}_3$. (B) The time the windows span: $\mathrm{prevwidth}_3$ and $\mathrm{postwidth}_3$ in days, and the green dashed line at $\mathrm{date\_post}_3$, the anchor from which the burn date is estimated. The $K = 2$ family is identical in construction but uses windows of two observations, which makes it more permissive and more sensitive, since fewer consecutive observations are needed to confirm the jump, and therefore useful where image density is low.

4.3 The peak, and what is stored

Each delta_K is collapsed independently to the observation t^* at which it is maximal within the focal year, and its whole bundle of companion values is read at that observation. The two families are collapsed separately on purpose: the optimal transition observation differs by window, and tying both to one peak would mis-anchor the other.

Three further bands are not tied to any peak: the maximum raw probability of the year, and the maxima of $\mathrm{minfore}_2$ and $\mathrm{minfore}_3$ over the whole series, which measure sustained magnitude regardless of whether it arrived as a jump. Finally, n, the count of valid focal observations, is exported as the sole density and quality metric; it also carries values marking non-burnable and non-observed pixels, and it is the one band that is never masked.

Table 1 lists the 16 bands. All are encoded as signed 16-bit integers, probabilities scaled by 10^4 and days as days, which roughly halves the size of the asset. Missing data is never given a probability: a pixel with no valid observations comes back with $n = 0$ and everything else masked. A pixel whose series is too short for a given window has that window's bands masked, which is a quality flag rather than an error.

Table 1: The 16 exported bands of the temporal stage. Each jump metric exists in two families, $K = 3$ (conservative) and $K = 2$ (permissive), collapsed at their own t^*.

Concept	Band(s)	What it measures
Change	`delta3_peak`, `delta2_peak`	the sustained jump, the central fire indicator
Persistence	`minfore3_peak`, `minfore2_peak`	the probability floor just after the jump
Sharpness	`jumpgap3`, `jumpgap2`	days between the last low and the first high obs.
Window span	`prevwidth3`, `prevwidth2`	days spanned by the back window
	`postwidth3`, `postwidth2`	days spanned by the fore window
Date	`date_post3`, `date_post2`	day of year of the post-jump observation
Magnitude	`pmax1`	maximum raw probability of the year
	`pmax2`, `pmax3`	maximum sustained probability of the year
Quality	`n`	count of valid focal observations

Note that $\mathrm{maxback}_K$ is not stored: it is exactly $\mathrm{minfore}_K - \mathrm{delta}_K$ and is recovered whenever it is needed. Note also that `date_post` is not the fire date. It is the day of the post-jump observation, and the burn date used downstream is the mid-point between the last low and the first high observation, $\mathrm{date\_post}_2 - \mathrm{jumpgap}_2/2$.

4.4 Crossing the year boundary

Window metrics starve at the extremes of a series, since $\mathrm{maxback}_3$ needs three observations behind it and $\mathrm{minfore}_3$ two ahead, and many parts of Argentina the start and the end of the calendar year are in the middle of the burning season. The focal year therefore borrows three observations from the previous year and two from the next: asymmetric on purpose, because symmetric padding of two would under-supply the back window at the first focal observation.

The neighbouring observations cannot simply be taken from the adjacent scenes, since the literal nearest scene may be cloud-masked at this pixel. Burn probability is instead computed for a margin of a few months on each side and the three latest previous and two earliest next valid observations are kept per pixel; the margin is widened for 1999 and 2000, where the archive is thin. Padded observations use the focal year's previous-year context, which is strictly the wrong year for them but is the better approximation over three-plus-two observations.

Crucially, the search for t^* is always restricted to focal observations. Padding is only ever context, so a burn can never be detected twice in two adjacent years, which means that there are no cross-year duplicates to remove later.

4.5 Cost

This is the most expensive artifact in the collection, in compute and in storage: a complete year is roughly 55 GiB over the 248 cartas, and the 27 years are on the order of 1.4 TiB. Per-account task parallelism in GEE is limited, so a whole-country year cannot be produced by one person in reasonable time; production ran as a distributed, multi-account export, one contributor per year.

The high computational cost of this step requires a careful planning on the timing of improvements to be made upstream.

5 Spatial stage I: segmentation into scar objects

5.1 Foundations

From the time series metrics we can define pixels that are clearly burned and those that are plausible burned. In both cases, but mostly in the second ones, the spatial context is highly informative: plausibly burned pixels that are isolated are probably noise, while those connected to clearly burned pixels are probably fire. We exploit this information by computing a region growing algorithm from *seeds*, the clearly burned pixels to *candidates*, the plausibly burned ones. Candidate islands with no seeds inside are dropped (Fig. 3).

This strategy exploits a favourable trade-off between the two error types. Seeds have high omission and low commission error; candidates, the opposite. A pixel whose own evidence is weak is accepted when it belongs to a patch whose evidence is strong, and rejected when it stands alone. Region growing has been used this way in burned-area mapping for decades, exactly because fire is spatially contagious (Adams and Bischof 1994; Haralick and Shapiro 1985; Fraser, Li, and Cihlar 2000; Giglio et al. 2009; Hawbaker et al. 2020; Long et al. 2019).

The stage errs deliberately towards recall: precision is recovered at the object classifier of Section 6.4, which can see the patch's shape, size and seed density, things no pixel can report.

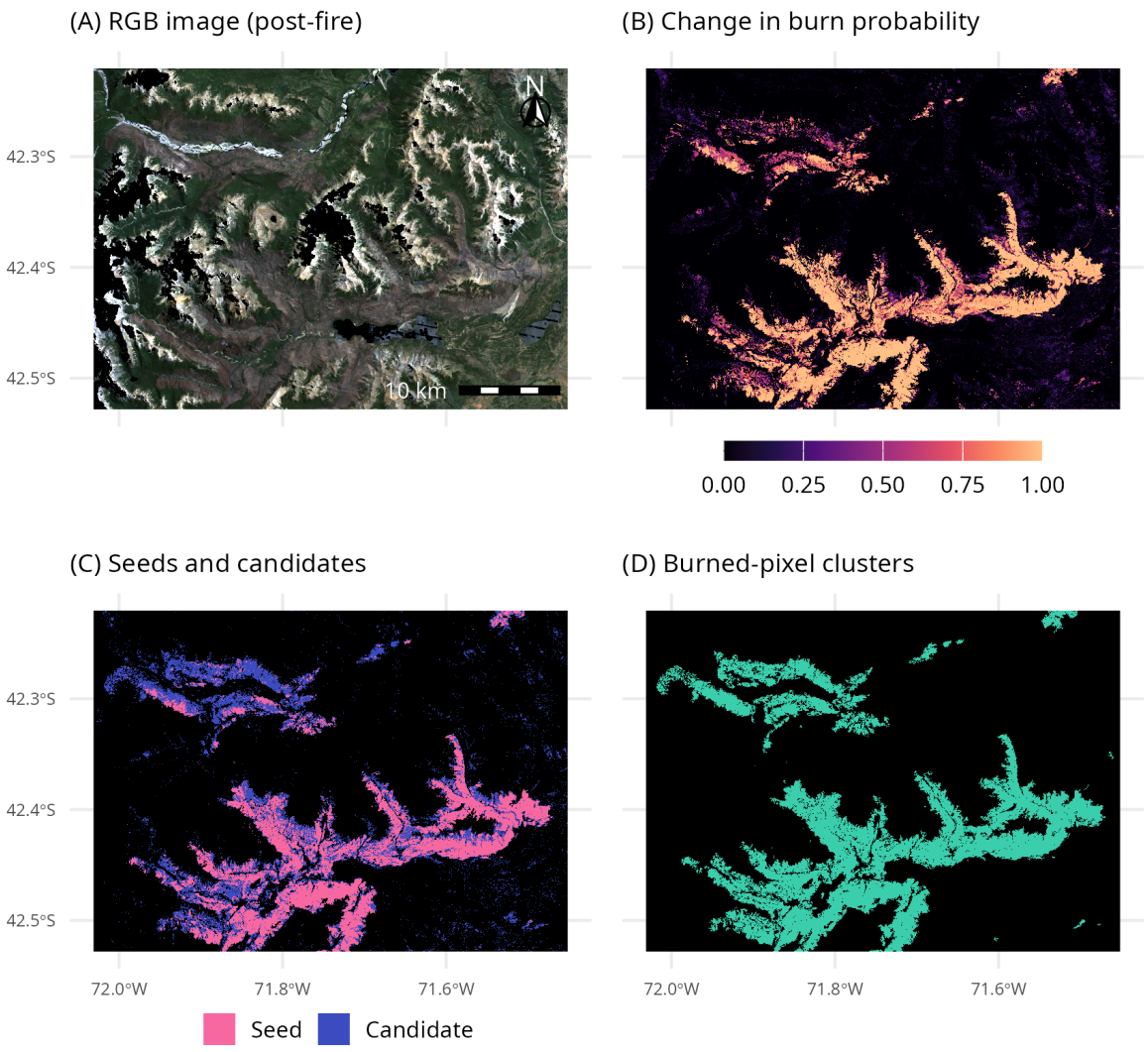

Figure 3: The region-growing step, illustrated on two fires that burned in 2015 in Chubut province (Río Turbio to the north, Cholila to the south). (A) RGB composite of the following summer, for reference only; black areas are masked by the quality filters. (B) The per-pixel evidence of burning as `delta2_peak`. (C) Seeds and candidates, obtained by applying a strict and a permissive cut to that evidence. (D) The output of the region growing: the candidate pixels connected to at least one seed.

5.2 The fire year

A fire is an object in space and in time, and this stage buys the time axis with a calendar rather than with an algorithm. Segmenting calendar years independently splits every scar that straddles 31 December and duplicates it across two years. We initially tried to pad the previous years and apply a temporal firebreak inside the segmentation, but it did not work, because the per-pixel dates are too noisy to serve as a barrier. The problem disappears almost entirely if the year boundary is placed where almost nothing is burning.

The *fire year* therefore runs from 1 May of year Y_1 to 30 April of year Y_1+1 and is named after its starting year. May is the country-wide trough of fire activity, in MODIS and VIIRS active fires and in our own burn dates, so no region's season is cut in half: the summer burners of the south and centre and the winter–spring season of the north both fall inside a single May-to-April year. One national boundary is a compromise; Patagonia would ideally break in June–July. Defining a separate fire year per region was tried and abandoned, because processing the whole country in one pass brought far more benefit than the optimum boundary would have.

Since the annual metrics cover calendar years 1999–2025, the fire years mapped are 1998–2025, with two partial edges: fire year 1998 holds only the January–April 1999 tail, and fire year 2025 only the May–December 2025 head. Both are flagged as partial in the asset metadata, and anything that averages across fire years has to know.

Although the spatial analysis stage works in fire years, the products are delivered in calendar years, in line with what the MapBiomas Fire network delivers (Alencar et al. 2022).

5.3 Seeds, candidates and dates

A fire year spans two calendar-year metric images, and each is thresholded before they are combined.

Candidates are pixels whose delta_2 exceeds a permissive cut, always the $K = 2$ family, because it gives the broadest footprint. Seeds are pixels whose delta_K exceeds a strict cut, with K chosen per pixel from its vegetation class and its observation count n: a pixel whose series is dense enough for its class is judged on the conservative $K = 3$ metric, and a pixel in a thin part of the archive on the permissive $K = 2$ one. A temporal-gap gate also applies, rejecting seeds whose $\mathrm{jumpgap}$ exceeds an n-adaptive ceiling: that is, rejecting changes too slow to be fire, which is precisely the drought signature that forced hand-drawn masks in the pilot. The deployed cuts are global, set by visual exploration of the metric images rather than fitted; per-class cuts calibrated on data were tried and dropped after failing out of sample. Non-vegetated classes receive an unreachable cut so that they can never fire.

Each pixel's burn date is taken from the $K = 2$ family as $\mathrm{date\_post}_2 - \mathrm{jumpgap}_2/2$, converted to an absolute day count since a fixed epoch, so that every later date comparison is safe across year boundaries.

The two calendar images are then filtered by date and combined. Only pixels whose burn date falls inside $[1\ \mathrm{May}\ Y_1,\ 1\ \mathrm{May}\ Y_1{+}1)$ are kept: the Y_1 image contributes May–December, the Y_1+1 image January–April, and a detection dated before May of Y_1 belongs to the previous fire year and is dropped. Where both images claim a pixel, they are combined by rank, with seed beating candidate and candidate beating nothing, and the date follows the image that won. This is what turns two calendar layers into one non-calendar fire year.

5.4 Slow post-fire dieback in Andean Patagonia

When low severity fires burn Andean Patagonian forest, they die slowly: part of a real scar crosses the change thresholds only in the following fire year, when the canopy finally collapses. For forest and shrubland pixels west of a longitude cut, a pixel that is seed or candidate in the next calendar image with a mid-winter to spring date is added to the focal year as a candidate, flagged with its own code. It is never added as a seed, because dieback must be able to extend a detected scar but

never to create one. Since padding survives region growing only where it connects to a genuine focal seed, it cannot manufacture a fire. These pixels carry a next-year date, so they are excluded from every date statistic downstream and inherit their parent object's date instead.

5.5 Supervised region growing

The region growing itself uses SNIC (Simple Non-Iterative Clustering, Achanta and Süsstrunk 2017), as implemented in GEE. SNIC is an unsupervised segmentation method, and its most common use is to partition a wall-to-wall image from uniformly placed seeds and spectral features. Here it is used in a far more guided way: we supply the seeds and we supply the footprint through which they may grow, so what would ordinarily be a segmentation becomes a classification. In this *supervised SNIC*, SNIC is the mechanism and the seed and candidate definition is the model.

Seed clumps below a small size are demoted to candidates before growing: they no longer seed, but they remain in the footprint so that a genuine cluster can still grow through them.

The algorithm runs on the whole country in one pass per fire year, not region by region: SNIC's memory footprint depends on its internal tile size and neighbourhood buffer, not on the extent of the export, so a national run costs what a regional one costs, and a large neighbourhood heals the internal seams. Only the mask of the result is kept; the cluster identifiers are discarded, because the objects are relabelled globally in the next stage.

The output is one image per fire year with a single band distinguishing candidate, seed and dieback padding, masked everywhere else. Burned pixels are 0.3–1 % of the country, which is what makes the whole national layer small enough to download and process the next stage locally.

6 Spatial stage II: fire objects and the object classifier

6.1 From pixels to objects

This stage materialises the fire patch so that it can be judged as a patch. A patch has properties no pixel has, such as shape, size and vegetation composition, variables useful to separate a real scar from a field of spectrally similar noise.

Objects are the connected components of the burned pixels, labelled one fire year at a time, whole-country and untiled, so that an object is global within its year and a scar is never split by a processing seam. Labelling uses a union-find structure (Tarjan 1975), which needs only one integer per burned cell; the country's 30 m lattice holds 9.16 billion cells while a heavy fire year holds around 116 million burned ones, three orders of magnitude fewer, and every stage of this part of the pipeline is sized by the latter.

Connectivity is wider than eight-neighbour adjacency, and selectively so. A real scar frequently breaks into fragments one or two pixels apart, and gluing them recovers one fire instead of many. The rule is equivalent to dilating the burned mask by one pixel, labelling, and dropping the halo, but it is implemented as a wider union window, which avoids ever materialising the dilation. Two burned cells are joined if their Chebyshev distance is at most 3 when both have enlarged context, at most 2 when only one does, and at most 1, plain eight-connectivity, when neither does. A cell is denied enlarged context when it is agriculture, pasture or temperate grassland, where bridging distinct burned fields would inflate commission error, or when it is a dieback pixel, which may only extend a scar. Everything else keeps it, that is forests, shrublands and the arid grasslands, because in sparse fuels bridging recovers a single real scar.

Each labelled object is then vectorised into one, possibly multi-part, polygon. Disconnected fragments joined by the bridging rule dissolve into a single multipolygon, so the bridge is preserved natively in the geometry.

6.2 Object metrics

The metric set is deliberately narrow, only what the classifier uses, and is computed directly over the burned cells rather than over the polygons, which is both faster and exact for area.

- Seed fraction: the share of the object's pixels that are seeds. This is the single most discriminating metric for fire against noise: real scars are densely seeded.
- Neighbourhood density: the mean, over the object's pixels, of the burned fraction within square windows of one, two and three pixels' radius. Sparse, threadlike objects score low; solid scars score high.
- Size: area in hectares, and pixel count.
- Dates: the median, minimum and maximum burn date of the object's pixels, and the calendar year that most of its pixels fall in.
- Vegetation composition: the fraction of the object's pixels in each `veg_fire` class.
- Shape: perimeter, convexity (area over convex-hull area), bounding-box fill and elongation, circularity $4\pi A/P^2$ and shape index $P/(2\sqrt{\pi A})$.

Area is measured on the ellipsoid, which means a Landsat pixel is not 0.09 ha: at these latitudes one Landsat pixel covers about 831 m^2 at 22°S and 517 m^2 at 55°S. A size class is therefore a range of pixel counts, and the same 15-pixel object changes size class between the north of the country and Patagonia.

6.3 Object labels

The classifier is fitted on about five thousand fire / non-fire labels collected by hand. Collectors worked in the GEE Code Editor on the seed-and-candidate layer itself, not on uploaded polygons, because that layer already shows, per pixel, candidate against seed on exactly the clusters the region growing kept, which makes shape and seed density, the most informative signal, fully visible without any vector data. Each collaborator exports a single asset; drawing-layer names carry the class and the fire year. Labels are then matched to objects, not to pixels, and the object inherits the label. This process yielded a fitting set of 5,255 objects: 2,788 fire and 2,467 non-fire.

The labelled sample is not a random sample of objects, and this single fact conditions every number the model produces. Labels were collected where fires were known to be, so per-year prevalence in the fitting set runs from 0.00 to 1.00, seven fire years have no labels at all, and small objects are strongly under-sampled. The model's ranking and its uncertainty are the product, not a calibrated probability; the deployed decision thresholds are a lower bound; and the cross-validation design has to be chosen with this sampling in mind. Fixing it is a collection-level task: it needs a randomly sampled set of small-object labels.

6.4 The classifier

The model is a probit Bayesian Additive Regression (and Classification) Tree (BART; Chipman, George, and McCulloch 2010; Albert and Chib 1993), fitted locally in R with `stochtree` (Herren et al. 2026). The choice follows from what five thousand labels allow. There is no honest way to tune a boosted ensemble's hyperparameters on a set this small, whereas BART's defaults are calibrated regularisation priors rather than placeholders. More importantly, the posterior yields a per-object interval on the probability: both an uncertainty statement that can be published and the targeting signal for a second round of label collection.

The model is applied to all 28 fire years and 1,689,419 objects.

6.5 Predictors

Twenty predictors are used. Fifteen are size, shape, density, seed and date-structure metrics; the other five are aggregated vegetation fractions: agriculture, flooded grassland, pasture, temperate grassland and woody. The aggregation replaced the 23 raw class fractions, which were 58 % of the design matrix and consumed the model's split budget without adding skill; membership of each group is derived from the remap table by name, so a remap change follows through automatically. The five groups are deliberately not a partition and do not sum to one.

No predictor may name the year or act as a proxy for it. The proportion of burned objects in the fitting set varies among years as an artifact of sampling, so the fire year, the calendar year and the mean observation count per object were all dropped: the first two leak the labels directly, and the third is a proxy for the era, since observation density rises as sensors come online. They remain product attributes, but none of them is a predictor. Season is legitimate information and does enter the model, as $\sin$ and $\cos$ of the day of year, which keeps it circular across the turn of the year without carrying any absolute time coordinate.

6.6 The decision threshold

A probability cut of 0.5 was far from optimal, and the right cut rises with object size. Thresholds were swept on out-of-fold probabilities and selected by Youden's J (Youden 1950), which is the only criterion here that does not move with prevalence. Sweeping in sample is never done, since the cut would then be chosen against answers the model has already seen.

Table 2: Deployed decision thresholds by size band, selected on out-of-fold probabilities by Youden's J. The $<$1 ha row is shown for completeness but is not leaned on: there, a hard size cut rather than a threshold is the right tool.

Size band	labels	prevalence	cut	J at the cut	J at 0.5
$<$ 1 ha	114	0.254	0.250	0.826	0.643
1–50 ha	3217	0.468	0.202	0.600	0.492
50–300 ha	1192	0.576	0.436	0.645	0.631
$\geq$ 300 ha	732	0.773	0.690	0.706	0.662

The bootstrap intervals on the three deployed cuts are close to disjoint, so the differences are signal rather than noise: the model is far more confident on large objects, and a single threshold would be simultaneously too high for small objects and too low for large ones. The gain lands where the error was: sensitivity in the 1–50 ha band rises from 0.596 to 0.837.

Two qualifications. First, the threshold governs object counts, not the area headline: on a heavy year the band cuts call some 64,000 objects fire against 50,500 at a flat 0.5, for an increase of well under 2 % in burned area, because area is dominated by large objects the model is confident about. Second, the cuts are a lower bound. J is prevalence-invariant as a measure, but the cut it selects is optimal for the prevalence of the set it was selected on, and that set is not the population (Section 6.3).

Where a human label exists for an object, it overrides the model's call. The published call is therefore the collected label where there is one and the model's call otherwise, and the three columns are kept separate in the product so that the provenance of every call is visible.

6.7 Performance

Cross-validation blocks space into 0.5° cells assigned to five folds. This removes the adjacency leak, since objects from one drawn polygon are neighbours, while keeping every region in every fold, which is the deployment condition. Leave-one-region-out was run for comparison and is the harshest possible

test, but it is not the condition the model is used in: every region does have labels in production, and held-out prevalence swings from 0.10 in Patagonia to 0.83 in the Pampas, so a fold's model would be trained on a class mix it is never scored on. The fold design decides the answer, and the choice has to be argued rather than defaulted.

Table 3: Spatially blocked cross-validated performance, pooled over folds, at a flat 0.5 cut and at the deployed per-band cuts. AUC is 0.891 pooled; the observation-weighted within-year AUC is 0.845 over the eighteen years carrying both classes, and that is the figure to watch for leakage.

	at 0.5	at the deployed cuts
Accuracy	0.790	0.814
Sensitivity	0.717	0.845
Specificity	0.872	0.780

Performance is not uniform across sizes. Above 300 ha the model is nearly clean (AUC 0.923); the weak band is 1–50 ha (AUC 0.871), which is also where 83 % of the objects are, and where at a flat 0.5 cut the model would miss two fires in five. That band is 61 % of the labels but 83 % of the population, the same under-sampling that the threshold caveat and the posterior widths both report.

6.8 Additional filters and minimum mapped size

Two commission patterns survived the object classifier and were only visible once the model was applied out of sample, over the whole country and all fire years: agricultural fields mapped as fire across the country, and objects of `grassland_pampa` mapped as fire in the agricultural Pampa. Harvest, tillage and stubble burning look like a burn scar to a single-date spectral model, and the `veg_fire` remap deliberately keeps agriculture burnable, so those pixels were never excluded from the candidate set. Both patterns are compact and densely seeded, which is exactly the case the object classifier cannot resolve, so they are removed by two explicit exclusion rules applied to the objects themselves rather than by masking the rasters. Excluding whole objects keeps the vector layer and every raster identical by construction.

Rule A removes objects that are almost entirely Pampa grassland (a `grassland_pampa` fraction above 0.70), burned between 1 July and 15 November, smaller than 150 ha, and lying in the agricultural Pampa, delimited by a hand-drawn polygon. The three confinements are what make the rule safe. The season matters because Pampa grassland does burn outside that window and that fire is real; the area and size limits matter because the same vegetation class covers the marshes of the Delta del Paraná, which burn inside the window: an unconfined version of the rule deleted two thirds of the Delta's burned area, including the 121 kha Islas del Paraná fire of 2020. Rule B removes objects whose agriculture fraction exceeds 0.40, counting the three annual-cropping `veg_fire` classes but not perennials and orchards. It applies everywhere and in every season.

The two rules are of similar size and cannot both fire on the same object at these thresholds. On fire year 2020 rule A removes 3.2 % of the burned area and rule B 3.8 %, together 7.0 %. Over the 28 fire years they take the mapped total from 69.12 Mha to 63.33 Mha, a reduction of 8.4 %.

The last filter is a minimum mapped size of one hectare, applied as a hard rule. The argument for it is cost and benefit rather than model uncertainty: it discards 3.4 % of the objects for 0.044 % of the area. Posterior uncertainty does fall with object size, from a mean interval width of 0.41 below half a hectare to 0.13 above a thousand hectares, but the model is unsure everywhere, with a global mean width of 0.32 and 31 % of all objects straddling their own decision cut, so small objects are not distinctively harder to classify. The one-hectare cut also satisfies, more strictly, the solitary-pixel filter of the network's common post-processing. Objects below it are scored like everything else but are not published.

7 Products

Our algorithm ends with a set of fire objects per fire-year; afterwards, all that remains is postprocessing to make the MapBiomas Fire network raster products. The fire-year objects are re-partitioned into calendar years, with the year and the month assigned per pixel from that pixel's own burn date rather than per object, and the published layers are derived from those pixels.

7.1 The six subproducts

The published products are the ones the MapBiomas Fuego network defines. They are not Argentina's design and they are not justified here: every participating country publishes the same set, in the same shape and with the same pixel encodings, so that the layers are comparable across the network and readable by the same platform. Argentina delivers all six: annual, monthly, accumulated, frequency, year of last fire and scar size. Some countries deliver only the first four.

The six become twelve assets: four have a coverage variant crossing the burned pixels with the published MapBiomas Argentina land cover of the same year, and the scar-size product comes with the scar identifier and the scar area it derives from. Every asset is a single multiband image with one band per year, not one image per year: 27 bands (1999–2025) for the annual, monthly and year-of-last-fire products, 53 for the frequency and accumulated ones, one per two-sided window (Table 4).

Table 4: The published raster subproducts. M is the month of burn (1–12, masked elsewhere) and L the land-cover class of the same calendar year. Asset names follow the network convention `mapbiomas_argentina_fire_collection1_<subproduct>_v<N>`.

Subproduct	Band name	What a pixel holds
`annual_burned`	`burned_area_<year>`	1 if burned that year
`annual_burned_coverage`	`burned_coverage_<year>`	L where burned
`monthly_burned`	`burned_monthly_<year>`	M, the month of burn
`monthly_burned_coverage`	`burned_coverage_<year>`	$M \times 100 + L$
`frequency_burned`	`fire_frequency_<y1>_<y2>`	number of years burned in the window
`frequency_burned_coverage`	`fire_frequency_<y1>_<y2>`	frequency $\times 100 + L$
`accumulated_burned`	`fire_accumulated_<y1>_<y2>`	1 if burned at least once in the window
`accumulated_burned_coverage`	`fire_accumulated_<y1>_<y2>`	L where burned at least once
`year_last_fire`	`classification_<year+1>`	calendar year of the most recent fire
`annual_burned_id`	`scar_id_<year>`	identifier of the scar the pixel belongs to
`annual_burned_area_ha`	`scar_area_ha_<year>`	area in hectares of that scar
`annual_burned_scar_size_range`	`scar_area_ha_<year>`	size class of that scar, 1–8

A scar is a connected group of burned pixels within one calendar year, labelled with plain eight-connectivity: the network's definition, and deliberately not the wider connectivity the fire objects of Section 6 use. The eight size classes are the platform's published legend: $<$10, 10–250, 250–500, 500–5,000, 5,000–10,000, 10,000–50,000, 50,000–100,000 and $\geq$100,000 ha. Argentina populates all eight. The scars also exist as vectors, one FeatureCollection per calendar year.

Two reading rules follow from how the pixels are built. The month of burn is measured, not inferred: it comes from the pixel's own burn date rather than from the date of a minimum-NBR composite, which is how most of the network derives it. And a fire that straddles 31 December is two scars, one in each calendar year. That is the price of assigning the date per pixel, and what makes the annual, monthly and scar-size layers agree pixel for pixel.

7.2 The fire-object vector database

Alongside the six products, Argentina publishes the fire objects themselves as a single vector layer: `burned_area_polygons_v2`, 1,012,645 objects covering 63.33 Mha, every mapped fire of all 28

fire years in one FeatureCollection. It is the output of Section 6 under the published selection, with nothing recomputed and no geometry touched.

Each polygon carries ten properties: `oid`, the stable object identifier that joins back to the object database and its twenty metrics; `fire_year`, the non-calendar mapping year; `calendar_year`; `area_ha`; the three burn dates `date_med`, `date_min` and `date_max`, written as `YYYY-MM-DD`; `p_mean` and `p_width`, the object classifier's posterior mean fire probability and the width of its credible interval; and `seed_mean`. Every feature is stamped with `system:time_start` from `date_med`, so the collection answers `filterDate()`; `system:time_end` is deliberately not set, so that one fire falls in exactly one time bucket and filtered areas stay summable.

Users of this layer must consider what follows:

- It is a fire-year layer, and the rasters are calendar-year layers. `calendar_year` is the majority year of the object's pixels, while every raster above assigns year and month per pixel. A fire straddling 31 December is split across two years in the rasters, but lands whole in one year here. Neither is wrong, but cross-tabulating the two without knowing this produces apparent "missing" area.
- Fire year 1998 exists here and in no published raster. The calendar series starts in 1999, so the November–December 1998 tail, 3,845 polygons and about 76 kha, is only in this layer.
- Area must be summed per object, not per row. One very large fire year 2000 object exceeded GEE's vertex limit on ingest and is stored as four features that each repeat the whole object's attributes. Summing `area_ha` row by row overstates the layer by 5.1 Mha.

7.3 Useful assets for expert users

Our algorithm provides intermediate products, like the time series metrics and the SNIC layers, which experienced users may request. They are available as GEE assets, and can be made public on request.

7.4 Where the products live

The first route to the products is the MapBiomas Argentina platform, and the website around it:

`https://plataforma.argentina.mapbiomas.org/`

`https://argentina.mapbiomas.org/`

The platform's fire section is the interactive view of exactly the layers described above: they can be browsed by year, month and frequency, crossed with territories and land cover, and downloaded as statistics. That is the entry point for the general public and the one we point users to.

The second route is GEE. The six subproducts and their coverage variants are copied, by an explicit subproduct list, into the MapBiomas network's public asset collection:

`projects/mapbiomas-public/assets/argentina/fire/collection1/`

These copies are readable by anyone and are what the platform itself ingests, so they are the layers to use in any GEE analysis. Each one is named `mapbiomas_argentina_fire_collection1_<subproduct>_v1`: the version token belongs to the published copy and is set by the network, so it does not follow our internal product version. The download page,

`https://argentina.mapbiomas.org/descargas/`

serves the same products as cloud-optimised GeoTIFFs, one file per band, for users who work outside GEE.

The fire-object vector database is not one of the network's six, so it is not swept into that public copy. It is Argentina's own layer, and we share it as a GEE asset with read access open to anyone, rather than through the network's public collection:

projects/mapbiomas-argentina/assets/FIRE/COLLECTION-1/FINAL_PRODUCTS/
burned_area_polygons_v2

7.5 Using the products in Google Earth Engine

The platform is the quickest route to look at the products; this section is for users who want to work with them in Google Earth Engine. The snapshot below loads the annual and monthly layers for one year, the frequency layer for the whole series, and the fire-object polygons, with the visualisations the platform uses:

https://code.earthengine.google.com/178cf543c607f361419aff569ec09d44

The script it opens is the following, which can also be pasted into a blank Code Editor tab:

```
var ROOT = 'projects/mapbiomas-public/assets/argentina/fire/collection1
    /';
var NAME = 'mapbiomas_argentina_fire_collection1_';
var V    = '_v1';   // version token of the published copy
var year = 2020;

// Annual burned area: one band per year, 1 = burned.
var annual = ee.Image(ROOT + NAME + 'annual_burned' + V)
               .select('burned_area_' + year).selfMask();
Map.addLayer(annual, {palette: ['#ff0000']}, 'burned ' + year);

// Month of burn: 1-12.
var monthly = ee.Image(ROOT + NAME + 'monthly_burned' + V)
                .select('burned_monthly_' + year).selfMask();
Map.addLayer(monthly, {min: 1, max: 12, palette: ['#2c7bb6','#ffffbf','
    #d7191c']},
             'month of burn ' + year);

// Fire frequency over the whole series: number of years burned.
var freq = ee.Image(ROOT + NAME + 'frequency_burned' + V)
             .select('fire_frequency_1999_2025');
Map.addLayer(freq, {min: 1, max: 10}, 'fire frequency');

// The fire-object vector database: one feature per mapped fire.
    Argentina's own
// layer, shared by us, so it is not under the network's public
    collection.
var POLYGONS = 'projects/mapbiomas-argentina/assets/FIRE/COLLECTION-1/'
             + 'FINAL_PRODUCTS/burned_area_polygons_v2';
var fires = ee.FeatureCollection(POLYGONS)
              .filterDate(year + '-01-01', (year + 1) + '-01-01');
var style = {color: '0000ff', fillColor: '0000ff66', width: 1};
Map.addLayer(fires.style(style), {}, 'fire objects ' + year);

Map.setCenter(-64, -34, 5);
```

Note the two patterns every consumer needs: a subproduct is one image and a year is one band of it, and the polygon layer is filtered by date rather than by a year property, because each fire carries a single timestamp taken from its median burn date.

8 Reproducibility

The code and the documentation needed to reproduce this workflow are hosted on GitHub. The repository is organised as one directory per collection; Collection 1 lives in `collection-01/`, with the

numbered pipeline scripts in `workflow/`, the configuration tables that every stage reads in `config/`, the fitted coefficients in `models/`, and the analysis notebooks in `notebooks/`. The per-step technical notes in `collection-01/docs` are the intended entry point for anyone reproducing or adapting the method: each one states what its stage measures, lists its inputs and outputs, records the decisions and their reasons, and warns about the failure modes that were actually encountered. The GEE JavaScript used for interactive threshold tuning and for label collection lives in a separate Code Editor repository, referenced from those notes.

9 Limitations and known challenges

Productive, actively managed landscapes are the hardest case. In agricultural regions the region-growing step produces poor clusters: harvest patches enter the candidate footprint and are occasionally merged into a genuine scar. No spatial strategy can repair this, because the ambiguity is already present in the per-observation probability: a harvested field and a burned field look alike to a single-date spectral model. Improving fire mapping in croplands and sown pastures requires better spectral models, which in turn requires collecting training data specifically on those confusions.

The object classifier was conceived from the pilot's experience, where many false positives were easily separable from fire by shape and seed density: dendritic, sparse objects associated with drought and volcanic ash. It does that job well, and it also removes clearing-type disturbances such as roads and urbanisation, which the pilot confounded with fire. But in regions where the false positives are compact and densely seeded, it helps much less. It is a filter on patch geometry and density, and it cannot recover information the spectral stage did not capture.

Understorey fires in the most productive systems remain largely undetected, as they were in the pilot: a fire that burns beneath an intact canopy leaves little signal in a 30 m top-of-canopy reflectance series.

10 Prioritized lines of improvement

These are listed in what we judge to be decreasing order of effect on mapping accuracy. They do not consider incorporating sensors other than Landsat, since for the first collections the priority is to refine the mapping without compromising the temporal extent of the product. In fact, Collection 2 aims at covering 1986–2026, the same span as the MapBiomas Argentina land-cover products.

Training data that represents the false positives. As in the pilot, the largest available gain is data on the processes that mimic fire: harvest, drought, flooding, logging, insect defoliation, ash deposition. They have to be collected explicitly as negatives, region by region. The Collection 1 training set already contains some of these, and they demonstrably helped; there are nowhere near enough of them. Collecting more data in productive areas is a main priority.

A randomly sampled set of object labels. The second-largest gain costs the least effort. A few thousand labels drawn at random from the object population, rather than from where fires were known, would convert the object classifier's output from a ranking into a calibrated probability, put the size-band thresholds on a defensible footing, and let the population fire rate be estimated rather than assumed. Objects with wide posterior intervals are the natural targets for a second round.

Thresholds for seeds and candidates. The deployed cuts were set by visual exploration and are global (Section 5), and they are the largest single lever in the segmentation: everything the object stage sees is what they let through. Investing the time to define them per `veg_fire` class, on data rather than by eye, would probably improve the quality of the fire objects more than anything else at this stage.

Better use of the temporal metrics in the segmentation. The sixteen exported bands were fixed before the segmentation stage existed, and some of them are never read. The interesting question is not only which to drop but which the segmentation should have been using: the seed and candidate cuts currently rest on the jump metrics and the observation count, leaving the sharpness and the window-span bands, the ones that best separate abrupt change from gradual change, almost unused. This is the cheapest available improvement in commission error, because it requires no re-export.

Splitting the vegetation remap. Several `veg_fire` classes are currently merged across regions for want of data rather than on ecological grounds (Section 3.4). With more collected fires, those groups can be split and fitted separately, which directly improves the per-class fit where it is weakest.

The spectral model itself. The logistic regression can be extended to a categorical regression predicting the probability of several classes at once, such as unburned, burned, cloud, water, ash, clearing and drought, instead of a binary one; or simple auxiliary models can be multiplied into it, for instance a brightness-based probability of cloud, snow or ash. Both remain compatible with the constraint that the model has to run inside GEE over the entire archive. This is also where richer classifiers would pay off most, if a deployment path for them existed.

Planning and versioning a change to the spectral model. Any change to the spectral model has to be paid for in the temporal stage, which is the single largest commitment a collection can make: about a terabyte of assets and some fifteen days of distributed export for the whole country and series. That cost is what makes provenance worth fixing first. Every exported annual-metrics asset records its year and its tile but not which coefficient set produced it, and because the coefficients enter through a per-pixel remap, a model improvement affecting one vegetation class changes nothing in the tiles where that class is absent. A partial re-export is therefore technically sound, and the only thing that stops it being auditable is the missing stamp. Recording the deployed model version on every asset costs nothing, and it is what would let an improvement be rolled out over the part of the country it actually changes instead of forcing a full re-run.

References

Achanta, Radhakrishna and Sabine Süsstrunk (July 2017). "Superpixels and Polygons Using Simple Non-Iterative Clustering". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4895–4904. DOI: 10.1109/CVPR.2017.520.

Adams, Rolf and Leanne Bischof (1994). "Seeded Region Growing". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.6, pp. 641–647. DOI: 10.1109/34.295913.

Albert, James H. and Siddhartha Chib (1993). "Bayesian Analysis of Binary and Polychotomous Response Data". In: *Journal of the American Statistical Association* 88.422, pp. 669–679. DOI: 10.1080/01621459.1993.10476321.

Alencar, Ane A. C. et al. (2022). "Long-Term Landsat-Based Monthly Burned Area Dataset for the Brazilian Biomes Using Deep Learning". In: *Remote Sensing* 14.11, p. 2510. DOI: 10.3390/rs14112510.

Badgley, Grayson, Christopher B. Field, and Joseph A. Berry (2017). "Canopy near-infrared reflectance and terrestrial photosynthesis". In: *Science Advances* 3.3, e1602244. DOI: 10.1126/sciadv.1602244.

Baig, Muhammad Hasan Ali et al. (2014). "Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance". In: *Remote Sensing Letters* 5.5, pp. 423–431. DOI: 10.1080/2150704X.2014.915434.

Camps-Valls, Gustau et al. (2021). "A unified vegetation index for quantifying the terrestrial biosphere". In: *Science Advances* 7.9, eabc7447. DOI: 10.1126/sciadv.abc7447.

Chipman, Hugh A., Edward I. George, and Robert E. McCulloch (2010). "BART: Bayesian additive regression trees". In: *The Annals of Applied Statistics* 4.1, pp. 266–298. DOI: 10.1214/09-AOAS285.

Fraser, Robert H., Zhanqing Li, and Josef Cihlar (2000). "Hotspot and NDVI Differencing Synergy (HANDS): A New Technique for Burned Area Mapping over Boreal Forest". In: *Remote Sensing of Environment* 74.3, pp. 362–376. DOI: 10.1016/S0034-4257(00)00139-7.

Friedman, Jerome, Trevor Hastie, and Rob Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1, pp. 1–22. DOI: 10.18637/jss.v033.i01.

Gao, Bo-Cai (1996). "NDWI — A normalized difference water index for remote sensing of vegetation liquid water from space". In: *Remote Sensing of Environment* 58.3, pp. 257–266. DOI: 10.1016/S0034-4257(96)00067-3.

Giglio, Louis et al. (2009). "An Active-Fire Based Burned Area Mapping Algorithm for the MODIS Sensor". In: *Remote Sensing of Environment* 113.2, pp. 408–420. DOI: 10.1016/j.rse.2008.10.006.

Gorelick, Noel et al. (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". In: *Remote Sensing of Environment* 202. Publisher: Elsevier, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Hall, Dorothy K., George A. Riggs, and Vincent V. Salomonson (1995). "Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data". In: *Remote Sensing of Environment* 54.2, pp. 127–140. DOI: 10.1016/0034-4257(95)00137-P.

Haralick, Robert M. and Linda G. Shapiro (1985). "Image Segmentation Techniques". In: *Computer Vision, Graphics, and Image Processing* 29.1, pp. 100–132. DOI: 10.1016/S0734-189X(85)90153-7.

Hawbaker, Todd J. et al. (2020). "The Landsat Burned Area Product". In: *Remote Sensing of Environment* 235, p. 111486. DOI: 10.1016/j.rse.2019.111486.

Herren, Drew et al. (2026). *stochtree: Stochastic Tree Ensembles (XBART and BART) for Supervised Learning and Causal Inference*. R package version 0.4.5. DOI: 10.32614/CRAN.package.stochtree. URL: https://CRAN.R-project.org/package=stochtree.

Huete, A. R. (1988). "A soil-adjusted vegetation index (SAVI)". In: *Remote Sensing of Environment* 25.3, pp. 295–309. DOI: 10.1016/0034-4257(88)90106-X.

Jiang, Zhangyan et al. (2008). "Development of a two-band enhanced vegetation index without a blue band". In: *Remote Sensing of Environment* 112.10, pp. 3833–3845. DOI: 10.1016/j.rse.2008.06.006.

Karnieli, Arnon et al. (2001). "AFRI — aerosol free vegetation index". In: *Remote Sensing of Environment* 77.1, pp. 10–21. DOI: 10.1016/S0034-4257(01)00190-0.

Key, C. H. and N. C. Benson (2006). "Landscape Assessment (LA): Sampling and Analysis Methods". In: *FIREMON: Fire Effects Monitoring and Inventory System*. General Technical Report RMRS-GTR-164-CD. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Long, Tengfei et al. (2019). "30 m resolution global annual burned area mapping based on Landsat Images and Google Earth Engine". In: *Remote Sensing* 11.5, p. 489. DOI: 10.3390/rs11050489.

MapBiomas Argentina (2025). *MapBiomas Argentina*. Annual land use and land cover classification. URL: https://argentina.mapbiomas.org/.

McFeeters, S. K. (1996). "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features". In: *International Journal of Remote Sensing* 17.7, pp. 1425–1432. DOI: 10.1080/01431169608948714.

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Roy, David P. et al. (2016). "Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity". In: *Remote Sensing of Environment* 185, pp. 57–70. DOI: 10.1016/j.rse.2015.12.024.

Souza Jr., Carlos M., Dar A. Roberts, and Mark A. Cochrane (2005). "Combining spectral and spatial information to map canopy damage from selective logging and forest fires". In: *Remote Sensing of Environment* 98.2-3, pp. 329–343. DOI: 10.1016/j.rse.2005.07.013.

Tarjan, Robert Endre (1975). "Efficiency of a Good But Not Linear Set Union Algorithm". In: *Journal of the ACM* 22.2, pp. 215–225. DOI: 10.1145/321879.321884.

Trigg, S. and S. Flasse (2001). "An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah". In: *International Journal of Remote Sensing* 22.13, pp. 2641–2647. DOI: 10.1080/01431160110053185.

Tucker, C. J. (1979). "Red and photographic infrared linear combinations for monitoring vegetation". In: *Remote Sensing of Environment* 8, pp. 127–150. DOI: 10.1016/0034-4257(79)90013-0.

U.S. Geological Survey (2021a). *Landsat Burned Area Product Guide*. Tech. rep. Version 1.1. United States: Department of the Interior.

— (2021b). *Landsat Collection 2 Level-2 Science Products*. Accessed via Google Earth Engine. URL: https://www.usgs.gov/landsat-missions/landsat-collection-2.

Youden, W. J. (1950). "Index for rating diagnostic tests". In: *Cancer* 3.1, pp. 32–35. DOI: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

Zou, Hui and Trevor Hastie (2005). "Regularization and variable selection via the elastic net". In: *Journal of the Royal Statistical Society: Series B* 67.2, pp. 301–320. DOI: 10.1111/j.1467-9868.2005.00503.x.