

# Package: sheepdog (via r-universe)

October 24, 2026

**Title** Sample, Validate and Calibrate 'BirdNET' Detections

**Version** 3.2.1

**Description** Tools to check whether automated detections from 'BirdNET' custom classifiers in passive acoustic monitoring are real. The package builds clip sets for manual validation (top-score, random-occasion or random-recording designs), summarises the reviewed clips, calibrates confidence scores into probabilities of true detection with hierarchical logistic models, and turns calibrated detections into nightly detection histories and effort-standardised community matrices.

**License** MIT + file LICENSE

**URL** <https://github.com/diogoprov/sheepdog>,  
<https://provetelab.org/sheepdog/>

**BugReports** <https://github.com/diogoprov/sheepdog/issues>

**Depends** R (>= 4.2)

**Imports** dplyr (>= 1.1.1), pROC, purrr (>= 1.0.0), readr, stats,  
stringr, tibble, tidyr, tools, withr, zip

**Suggests** ape, av, brms, crywolf, ggplot2, gllvm, knitr, lme4,  
posterior, priorsense, rmarkdown, testthat (>= 3.0.0), vegan

**Additional\_repositories** <https://diogoprov.r-universe.dev>,  
<https://stan-dev.r-universe.dev>

**VignetteBuilder** knitr

**Config/testthat/edition** 3

**Encoding** UTF-8

**Roxygen** list(markdown = TRUE)

**Config/roxygen2/version** 8.1.1

**Config/pak/sysreqs** libicu-dev libx11-dev

**Repository** <https://diogoprov.r-universe.dev>

**Date/Publication** 2026-10-24 04:26:22 UTC

**RemoteUrl** <https://github.com/diogoprov/sheepdog>

**RemoteRef** main

**RemoteSha** 89a6cf3b5aea63dfb23174660bc5fd4ce24da921

## Contents

|                                        |    |
|----------------------------------------|----|
| assign_reviewers . . . . .             | 3  |
| build_community_matrices . . . . .     | 3  |
| build_training_set . . . . .           | 4  |
| calibrate . . . . .                    | 5  |
| calibration_diagnostics . . . . .      | 8  |
| calibration_effort . . . . .           | 8  |
| calibration_folds . . . . .            | 9  |
| calibration_models . . . . .           | 10 |
| calibration_prior . . . . .            | 11 |
| calibration_sensitivity . . . . .      | 12 |
| cell_status . . . . .                  | 12 |
| classify_review . . . . .              | 13 |
| community_matrices . . . . .           | 14 |
| community_matrices_draws . . . . .     | 15 |
| compare_calibration_models . . . . .   | 16 |
| consensus_verdicts . . . . .           | 17 |
| cut_clips . . . . .                    | 17 |
| detection_history . . . . .            | 18 |
| exclude_training_clips . . . . .       | 19 |
| extract_validation_clips . . . . .     | 19 |
| fit_calibration . . . . .              | 21 |
| history_draws . . . . .                | 22 |
| nightly_presence . . . . .             | 23 |
| pack_clips . . . . .                   | 24 |
| plot_calibration . . . . .             | 24 |
| plot_calibration_draws . . . . .       | 25 |
| plot_precision . . . . .               | 25 |
| post_validation . . . . .              | 26 |
| precision_by_bin . . . . .             | 27 |
| prior_predictive_calibration . . . . . | 28 |
| read_selection_tables . . . . .        | 28 |
| read_validation_tables . . . . .       | 29 |
| read_validations . . . . .             | 29 |
| retraining_clips . . . . .             | 30 |
| reviewer_agreement . . . . .           | 31 |
| score_detections . . . . .             | 31 |
| select_calibration_model . . . . .     | 32 |
| select_clips . . . . .                 | 33 |
| split_by_reviewer . . . . .            | 36 |
| suggest_core_species . . . . .         | 37 |
| suggest_thresholds . . . . .           | 37 |

|                                          |           |
|------------------------------------------|-----------|
| <code>assign_reviewers</code>            | 3         |
| <code>summarise_history</code> . . . . . | 38        |
| <b>Index</b>                             | <b>39</b> |

---

|                               |                                              |
|-------------------------------|----------------------------------------------|
| <code>assign_reviewers</code> | <i>Assign the clips to several reviewers</i> |
|-------------------------------|----------------------------------------------|

---

**Description**

Splits the clips of a manifest among reviewers, balanced within each species, and gives a fraction of them to a second reviewer, so that the agreement between reviewers can be measured with `reviewer_agreement()`. The overlap is drawn within each species, so every species gets doubly reviewed clips.

**Usage**

```
assign_reviewers(manifest, reviewers, overlap = 0.1, seed = 2026)
```

**Arguments**

- `manifest`           Clip manifest, as returned by `select_clips()` or `cut_clips()`.
- `reviewers`          Names of the reviewers (at least two).
- `overlap`            Fraction of the clips of each species reviewed by two reviewers.
- `seed`                Seed of the draws. The global random number generator is left as it was.

**Details**

Cut the clips with `cut_clips()` first, then copy each reviewer's share to their own folder with `split_by_reviewer()`.

**Value**

The manifest with `reviewer` (the first reviewer of each clip) and `second_reviewer` (NA for clips reviewed once).

---

|                                       |                                     |
|---------------------------------------|-------------------------------------|
| <code>build_community_matrices</code> | <i>Write the community matrices</i> |
|---------------------------------------|-------------------------------------|

---

**Description**

Reads the nightly detection history written by `calibrate()` and writes the matrices of `community_matrices()` to `output_dir`: `frequency_matrix_site_month_k<k>.csv`, `count_matrix_site_month_k<k>.csv` and `effort_site_month.csv` (which also lists the dropped cells). This is the function form of the v1.2.0 script `gerar_matrizes_comunidade.R`.

**Usage**

```
build_community_matrices(
  history_file,
  output_dir,
  core_species,
  k_min = 2,
  min_nights_month = 5
)
```

**Arguments**

|                  |                                                                                                                           |
|------------------|---------------------------------------------------------------------------------------------------------------------------|
| history_file     | Path to nightly_detection_history.csv (or a v1.2.0 historico_deteccao_noite.csv).                                         |
| output_dir       | Output folder; created if needed.                                                                                         |
| core_species     | Species codes (the species column) that enter the matrix.                                                                 |
| k_min            | Repetition rule: 1, 2 or 3 reliable detections on the night.                                                              |
| min_nights_month | Site x month cells with fewer recorded nights are dropped from the matrices (unstable frequencies) and flagged in effort. |

**Value**

The list returned by `community_matrices()`, invisibly.

---

|                    |                                                        |
|--------------------|--------------------------------------------------------|
| build_training_set | <i>Build the training folders for BirdNET-Analyzer</i> |
|--------------------|--------------------------------------------------------|

---

**Description**

Copies the reviewed clips into the folder layout that the training script of BirdNET-Analyzer expects: one folder per class, named after the label (species\_code, so the retrained classifier keeps the labels of the current one), with the negatives of a class in a folder with the same name prefixed by -. Clips of the sites in test\_sites go to a separate test folder, to be passed as --test\_data, so that the new classifier is evaluated on sites it was not trained on.

**Usage**

```
build_training_set(
  validation,
  output_dir,
  test_sites = NULL,
  negatives = TRUE,
  min_files = 100
)
```

## Arguments

|            |                                                                                                                                                                                                                                                                                       |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| validation | Reviewed clips, as returned by <code>classify_review()</code> , <code>post_validation()</code> ( <code>\$validation</code> ) or <code>read_validation_tables()</code> . It needs site, species_code, clip_file, verdict and reviewed_path (the path where the review left each clip). |
| output_dir | Output folder; train/ and, with test_sites, test/ are created in it.                                                                                                                                                                                                                  |
| test_sites | Sites held out for testing. NULL puts every clip in train/.                                                                                                                                                                                                                           |
| negatives  | Copy the negative clips into ~<species_code> folders.                                                                                                                                                                                                                                 |
| min_files  | Minimum number of positive training files per class below which a warning is raised.                                                                                                                                                                                                  |

## Details

The layout follows the documentation of BirdNET-Analyzer (version 2.4.0, `birdnet_analyzer.train`), which also recommends at least 100 files per species; classes below `min_files` in the training set raise a warning. Whole-recording clips of "random\_recordings" (longer than 3 s, confidence = NA) are left out: their label says only that the species was heard somewhere in the recording.

The table of copied clips is also written to `training_set.csv` in `output_dir`; keep it with the classifier, to leave these clips out of later validation rounds with `exclude_training_clips()`.

## Value

A tibble with one row per copied clip: site, species\_code, verdict, set ("train" or "test"), class\_folder, clip\_file, source\_audio, clip\_start, clip\_end, from and to, invisibly.

---

calibrate

*Calibrate confidence scores and build detection histories*


---

## Description

Runs the whole calibration stage: reads the reviewed clips, compares the candidate models by cross-validation, refits the best one to all the clips, optionally draws the effort curve, scores every detection of the complete selection tables and builds the nightly detection history. Writes to `output_dir`:

- `model_comparison.csv` (`compare_calibration_models()`);
- `calibration_folds.csv`, when `fold_group` is supplied: validation row number, site, species, grouping columns and assigned fold;
- `calibration_model.rds` and `calibration_model.txt` (the refitted best model and its summary);
- `calibration_curves.png` (`plot_calibration()`), when 'ggplot2' is installed;
- `effort_curve.csv` (`calibration_effort()`), when `effort_curve` is TRUE;
- `calibrated_detections_<site>.csv`, one per site (`score_detections()`);
- `nightly_detection_history.csv` (`detection_history()`);
- `detection_summary.csv` (`summarise_history()`).

**Usage**

```
calibrate(
  validation_files,
  prediction_tables,
  output_dir,
  site_map = NULL,
  min_confidence = 0.1,
  p_min = 0.95,
  min_detections_night = 1,
  dawn_hour = 12,
  k_folds = 5,
  select = c("one_se_logloss", "best_auc"),
  uncorrelated = FALSE,
  effort_curve = FALSE,
  effort_n = c(5, 10, 15, 20, 25, 30),
  effort_reps = 50,
  engine = c("lme4", "brms"),
  prior = NULL,
  brms_args = list(),
  rule = c("mean", "probability"),
  prob_min = 0.9,
  seed = 2026,
  fold_group = NULL
)
```

**Arguments**

|                                   |                                                                                                                                         |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| <code>validation_files</code>     | Paths to the <code>validation_by_clip.csv</code> files of the sites (v1.2.0 <code>validacao_por_clipe.csv</code> files are also read).  |
| <code>prediction_tables</code>    | Named character vector of the complete selection tables, named by site.                                                                 |
| <code>output_dir</code>           | Output folder; created if needed. The calibrated detections of a large project can take hundreds of MB.                                 |
| <code>site_map</code>             | Optional named character vector to fix diverging site names, as <code>c(old = "new")</code> , e.g. <code>c(BEP2 = "BEP")</code> .       |
| <code>min_confidence</code>       | Detections below this confidence are not scored.                                                                                        |
| <code>p_min</code>                | Threshold on the calibrated probability (not on the confidence score).                                                                  |
| <code>min_detections_night</code> | Repetition rule: minimum number of reliable detections for a night with detection.                                                      |
| <code>dawn_hour</code>            | Recordings that start before this hour belong to the night before.                                                                      |
| <code>k_folds</code>              | Number of folds.                                                                                                                        |
| <code>select</code>               | Rule to choose the model; see <a href="#">select_calibration_model()</a> .                                                              |
| <code>uncorrelated</code>         | Include the m4 variant without the species intercept-slope correlation among the candidates; see <a href="#">calibration_models()</a> . |

|                       |                                                                                                                                                                                                                                                                                                                                                                                                            |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| effort_curve          | Draw the effort curve (slow).                                                                                                                                                                                                                                                                                                                                                                              |
| effort_n, effort_reps | Training sizes and repetitions of the effort curve.                                                                                                                                                                                                                                                                                                                                                        |
| engine                | Engine of the final fit: "lme4" or "brms" (see <a href="#">fit_calibration()</a> ). The comparison of the candidate models always uses maximum likelihood (fast); with "brms", the selected model is then refitted in a Bayesian framework, its convergence diagnostics are written to calibration_diagnostics.csv, and the scored detections carry p_tp_lower and pr_reliable.                            |
| prior                 | engine = "brms": the prior, as a brmsprior object or as a function of the selected formula and the validations that returns one, e.g. calibration_prior (whose classes then match the selected model). Required with "brms".                                                                                                                                                                               |
| brms_args             | engine = "brms": list of further arguments for brms::brm() (e.g. list(chains = 4, cores = 4, backend = "cmdstanr")).                                                                                                                                                                                                                                                                                       |
| rule                  | When a detection is reliable: "mean", its calibrated probability (the posterior mean with engine = "brms") is $\geq p_{\min}$ ; "probability", the posterior probability that $P(TP) \geq p_{\min}$ (the pr_reliable of a model fitted with engine = "brms") is $\geq prob_{\min}$ . The second rule propagates the uncertainty of the calibration and is more conservative where few clips were reviewed. |
| prob_min              | rule = "probability": threshold on pr_reliable. pr_reliable is computed in <a href="#">score_detections()</a> for its p_min; use the same p_min in both.                                                                                                                                                                                                                                                   |
| seed                  | Seed of the fold assignment. The global random number generator is left as it was.                                                                                                                                                                                                                                                                                                                         |
| fold_group            | Optional character vector of grouping columns. Do not include species, verdicts or confidence scores: these would split recordings or nights by species or by outcome. Group IDs must be nonmissing and nonempty. NULL keeps the original clip folds.                                                                                                                                                      |

## Details

The candidate family follows the observed species and sites: one species uses the site-effect family of [calibration\\_models\(\)](#), and a single site omits non-estimable site random effects.

The model is chosen by [select\\_calibration\\_model\(\)](#) with the rule select: by default, the simplest model within one standard error of the lowest log-loss, leaving out models singular in most folds ("one\_se\_logloss"). Use select = "best\_auc" to reproduce the 2.x selection by highest mean AUC. When the refit is singular, [fit\\_calibration\(\)](#) warns; read calibration\_model.txt and the comparison before using the probabilities.

Converting BirdNET confidence scores into probabilities is recommended by the BirdNET authors (Wood & Kahl 2024, [doi:10.1007/s10336024021445](#)); calibration by logistic regression on logit(score) is applied, for instance, by Owens et al. (2026, [doi:10.64898/2026.06.01.729442](#)). This function extends that approach with a hierarchical multi-species structure. This is the function form of the v1.2.0 script calibracao.R.

## Value

A list with comparison, fit, effort (or NULL), scored and history, invisibly.

---

calibration\_diagnostics

*Convergence diagnostics of a Bayesian calibration model*


---

### Description

Reports the largest R-hat, the smallest bulk and tail effective sample sizes and the number of divergent transitions of a model fitted with engine = "brms", and whether they pass the usual limits.

### Usage

```
calibration_diagnostics(fit, rhat_max = 1.01, ess_min = 400)
```

### Arguments

|          |                                                                        |
|----------|------------------------------------------------------------------------|
| fit      | Model returned by <code>fit_calibration()</code> with engine = "brms". |
| rhat_max | Largest acceptable R-hat.                                              |
| ess_min  | Smallest acceptable bulk and tail effective sample size.               |

### Value

A one-row tibble with max\_rhat, min\_ess\_bulk, min\_ess\_tail, n\_divergent and ok.

---

calibration\_effort

*Effort curve of the calibration*


---

### Description

How many reviewed clips per site x species are needed to calibrate well? For each training size, draws that many clips per cell (keeping at least five for testing), fits the model and computes the AUC on the remaining clips. Repetitions with fewer than 50 test clips, or whose fit fails, are dropped.

### Usage

```
calibration_effort(
  validations,
  formula,
  n_train = c(5, 10, 15, 20, 25, 30),
  n_reps = 50,
  seed = 2026
)
```

**Arguments**

|             |                                                                                    |
|-------------|------------------------------------------------------------------------------------|
| validations | Table returned by <code>read_validations()</code> .                                |
| formula     | Model formula, usually the best one of <code>compare_calibration_models()</code> . |
| n_train     | Training sizes per site x species.                                                 |
| n_reps      | Repetitions per training size.                                                     |
| seed        | Seed of the draws. The global random number generator is left as it was.           |

**Details**

Cells with fewer than `n_train + 5` clips contribute all but five of their clips, so the training size actually used can be below `n_train`; it is reported in `n_train_mean`, with a warning. The test set shrinks as `n_train` grows (`n_test_mean`), which widens `auc_sd`.

**Value**

A tibble with `n_per_cell` (the requested training size), `auc_mean`, `auc_sd`, `n_train_mean` (mean clips per cell actually used for training), `n_test_mean` (mean size of the test set) and `n_failed` (repetitions dropped).

---

|                   |                                                      |
|-------------------|------------------------------------------------------|
| calibration_folds | <i>Assign cross-validation folds for calibration</i> |
|-------------------|------------------------------------------------------|

---

**Description**

With `fold_group = NULL`, assigns clips within each site x species, reproducing the original fold assignment. For grouped folds, give one or more column names, e.g. "source\_audio" for recordings or "night" for nights. Site is always part of the grouping key. All species and clips sharing that key go into the same fold; groups are shuffled and distributed among folds within site, balancing group counts rather than clip counts. Each site must have at least `k_folds` groups.

**Usage**

```
calibration_folds(validations, k_folds = 5, seed = 2026, fold_group = NULL)
```

**Arguments**

|             |                                                                                                                                                                                                                                                                    |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| validations | Calibration data including site, species and any columns named in <code>fold_group</code> . Use <code>read_validations(keep = )</code> to retain metadata from validation files.                                                                                   |
| k_folds     | Integer number of folds, at least two.                                                                                                                                                                                                                             |
| seed        | Seed of the assignment; the caller's random state is restored.                                                                                                                                                                                                     |
| fold_group  | Optional character vector of grouping columns. Do not include species, verdicts or confidence scores: these would split recordings or nights by species or by outcome. Group IDs must be nonmissing and nonempty. <code>NULL</code> keeps the original clip folds. |

Details

Grouped folds evaluate new recordings or nights at known sites, not transfer to new sites. Unlike clip folds, they cannot guarantee that each species appears in each training and test set. A test fold with only one verdict has no AUC but still contributes probability scores.

Value

Integer vector of fold numbers, in the input row order.

---

|                    |                                     |
|--------------------|-------------------------------------|
| calibration_models | <i>Candidate calibration models</i> |
|--------------------|-------------------------------------|

---

Description

The family of logistic calibration models compared by [compare\\_calibration\\_models\(\)](#), from a single logistic regression on the logit of the confidence score to a binomial GLMM with random slopes by species and a site x species intercept.

Usage

```
calibration_models(uncorrelated = FALSE, single_species = FALSE)
```

Arguments

- uncorrelated     Add m5\_species\_x\_site\_uncorrelated, the m4 structure without the correlation between the species intercepts and slopes. It avoids the singular fits of m4 when that correlation is estimated at +-1. FALSE keeps the candidates of 2.0.0.
- single\_species   Use a family for one species: global logistic regression, a site intercept, and a site intercept and slope. With uncorrelated = TRUE, also include independent site intercepts and slopes. FALSE preserves the multispecies family.

Details

The list goes from the simplest to the most complex model; that order is the complexity used by [select\\_calibration\\_model\(\)](#).

Value

A named list of formulas: m0\_global, m1\_species\_intercept, m2\_species\_slope, m3\_plus\_site, m4\_species\_x\_site and, with uncorrelated = TRUE, m5\_species\_x\_site\_uncorrelated. With single\_species = TRUE: m0\_global, m1\_site\_intercept, m2\_site\_slope and optionally m3\_site\_slope\_uncorrelated.

---

|                   |                                                       |
|-------------------|-------------------------------------------------------|
| calibration_prior | <i>Example prior for a Bayesian calibration model</i> |
|-------------------|-------------------------------------------------------|

---

## Description

Builds a weakly informative prior for a calibration formula, centred on the hypothesis that the BirdNET score is already calibrated: on the model scale ( $tp \sim \text{logit\_conf}$ ), intercept 0 and slope 1 mean  $P(TP) = \text{score}$ . The data then show how far the classifier departs from it. Only the classes present in the formula are returned:

## Usage

```
calibration_prior(formula, validations)
```

## Arguments

|             |                                                                  |
|-------------|------------------------------------------------------------------|
| formula     | Calibration formula (see <a href="#">calibration_models()</a> ). |
| validations | Table returned by <a href="#">read_validations()</a> .           |

## Details

- Intercept:  $\text{normal}(m, 1.5)$ , where  $m$  is the mean of `logit_conf` among the reviewed clips. 'brms' centres the predictors, so its intercept is the logit of  $P(TP)$  at the mean logit score; with a slope of 1, an intercept equal to that mean gives  $P(TP) = \text{score}$ ;
- $b$  (slope of `logit_conf`):  $\text{normal}(1, 1)$  truncated at 0, centred on a calibrated score, wide enough to allow slopes from near 0 to near 3, and assuming that a higher score never lowers  $P(TP)$ ;
- $sd$  (standard deviations of the random effects): `exponential(1)`;
- $cor$  (correlations between random effects): `lkj(2)`, which pulls the correlations away from  $\pm 1$  (the singular fits of `m4` with 'lme4').

This is an example, not a default: [fit\\_calibration\(\)](#) with `engine = "brms"` requires the prior to be given explicitly. Check what it implies on the probability scale with [prior\\_predictive\\_calibration\(\)](#) before fitting, and its influence on the posterior with [calibration\\_sensitivity\(\)](#). On weakly informative priors and prior predictive checks see McElreath (2020, [doi:10.1201/9780429029608](#)) and Gelman et al. (2026, *Bayesian Workflow*, Chapman & Hall).

## Value

A `brmsprior` object.

---

|                         |                                                                         |
|-------------------------|-------------------------------------------------------------------------|
| calibration_sensitivity | <i>Prior and likelihood sensitivity of a Bayesian calibration model</i> |
|-------------------------|-------------------------------------------------------------------------|

---

**Description**

Power-scaling sensitivity analysis with 'priorsense' (Kallioinen et al. 2024, [doi:10.1007/s11222-023103665](https://doi.org/10.1007/s11222-023103665)): how much the posterior changes when the prior or the likelihood is up- or down-weighted. A prior-data conflict or a strongly informative prior shows up as sensitivity to both.

**Usage**

```
calibration_sensitivity(fit, variable = NULL, ...)
```

**Arguments**

- fit                    Model returned by `fit_calibration()` with engine = "brms".
- variable             Parameters to analyse. NULL takes the population-level effects and the standard deviations and correlations of the random effects (b\_, sd\_, cor\_), leaving out the individual random effects.
- ...                   Further arguments for `priorsense::powerscale_sensitivity()`.

**Value**

The table of `priorsense::powerscale_sensitivity()`.

---

|             |                                                                      |
|-------------|----------------------------------------------------------------------|
| cell_status | <i>Status of each site x species cell for the community matrices</i> |
|-------------|----------------------------------------------------------------------|

---

**Description**

Classifies each site x species cell of the calibration by whether its detections can enter a community matrix, separating the absence of a species from a calibration that does not work:

**Usage**

```
cell_status(fit, validations, p_min = 0.95, min_n = 15)
```

**Arguments**

- fit                    Model returned by `fit_calibration()`.
- validations           Table returned by `read_validations()`, the clips the model was fitted to.
- p\_min                  Calibrated probability that makes a detection reliable, as in `detection_history()`.
- min\_n                  Minimum number of reviewed clips in a cell.

## Details

- "reliable": the calibrated curve reaches `p_min` within the range of reviewed scores of the cell, and the cell has at least `min_n` reviewed clips;
- "probably\_absent": clips were reviewed but none was a true positive. The species can enter the matrix with zeros at that site;
- "unreliable": true positives were reviewed, but the curve never reaches `p_min` within the reviewed scores, or the cell has fewer than `min_n` reviewed clips, or no clip at all was reviewed. A zero in such a cell could be a false negative.

`suggest_core_species()` keeps the species without unreliable cells. The final choice of core species remains with the analyst.

## Value

A tibble with one row per site x species: `site`, `species`, `n_reviewed`, `n_tp`, `max_confidence` (highest reviewed score), `threshold` (lowest grid score at which the calibrated probability reaches `p_min`; NA if never), `status` and `reason`. The grid contains both observed endpoints and steps of 0.01 within the reviewed range.

---

|                              |                                    |
|------------------------------|------------------------------------|
| <code>classify_review</code> | <i>Classify the reviewed clips</i> |
|------------------------------|------------------------------------|

---

## Description

Reads the result of the manual review done in the BirdNET GUI and joins it to the clip manifest. The GUI organises the review by moving each clip into subfolders such as `<species>/Positive` and `<species>/Negative`, so the verdict of each clip comes from the path where it is found. Clips still in the species folder are not reviewed.

## Usage

```
classify_review(
  manifest,
  review_dir,
  positive_pattern = "positiv",
  negative_pattern = "negativ"
)
```

## Arguments

|                                                               |                                                                                                                                                              |
|---------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>manifest</code>                                         | Clip manifest: a data frame or the path to <code>clip_manifest.csv</code> . Manifests written by v1.2.0 ( <code>manifesto_clipes.csv</code> ) are also read. |
| <code>review_dir</code>                                       | Folder with the reviewed clips (the <code>clips</code> folder of the validation package, after the review).                                                  |
| <code>positive_pattern</code> , <code>negative_pattern</code> | Regular expressions (case-insensitive) that identify the positive and negative folders in the path of each clip.                                             |

**Value**

The manifest with two more columns: `verdict` ("positive", "negative", "not\_reviewed" or "clip\_missing") and `reviewed_path`, the path where the clip was found. This is the table written to `validation_by_clip.csv`.

---

|                    |                                                      |
|--------------------|------------------------------------------------------|
| community_matrices | <i>Community matrices from the detection history</i> |
|--------------------|------------------------------------------------------|

---

**Description**

Turns the calibrated nightly detection history into site x month community matrices for multivariate ordinations (PCoA, NMDS, RDA). Because those analyses do not model false positives, the entry rule is strict:

- a species counts as detected on a night only when it had at least `k_min` reliable detections that night, i.e. detections with a calibrated probability  $\geq p_{min}$  of `calibrate()` (column `detection_k < k_min` of the history);
- only the core species enter the matrix: species whose calibration is reliable at every site. Other species would create artefactual beta diversity from differences in classifier recall among sites;
- each cell is the frequency of nights with detection, standardised by effort: nights with detection / nights recorded at that site x month.

**Usage**

```
community_matrices(
  history,
  core_species,
  k_min = 2,
  min_nights_month = 5,
  drop_empty = FALSE
)
```

**Arguments**

|                               |                                                                                                                                                         |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>history</code>          | Nightly detection history, as returned by <code>detection_history()</code> (v1.2.0 <code>historico_deteccao_noite.csv</code> tables are also accepted). |
| <code>core_species</code>     | Species codes (the species column) that enter the matrix.                                                                                               |
| <code>k_min</code>            | Repetition rule: 1, 2 or 3 reliable detections on the night.                                                                                            |
| <code>min_nights_month</code> | Site x month cells with fewer recorded nights are dropped from the matrices (unstable frequencies) and flagged in effort.                               |
| <code>drop_empty</code>       | Drop the site x month rows in which no core species was detected. Distances such as Bray-Curtis are undefined for them.                                 |

**Details**

Effort (nights recorded) is approximated by the number of distinct nights in the history of the site, for any species. With dense detections, every night actually recorded appears in the history.

**Value**

A list with three tibbles, frequency and counts (one row per site x month, with site, month, n\_nights\_recorded and one column per core species) and effort (site, month, n\_nights\_recorded, included), plus trials: a numeric matrix with the dimensions of the species columns of counts holding the nights recorded, the Ntrials of a binomial model of the counts (e.g. `gllvm::gllvm()`, which needs a matrix when effort varies among samples). The row names of trials are <site>\_<month>.

---

community\_matrices\_draws

*Community matrices over posterior draws*

---

**Description**

Builds the community matrices of `community_matrices()` for each draw of `history_draws()` and summarises the frequency of each site x month x species cell by its posterior mean and an interval, so that the uncertainty of the calibration reaches the matrices. The per-draw frequencies are also returned, e.g. to repeat an ordination on each draw.

**Usage**

```
community_matrices_draws(
  history,
  core_species,
  k_min = 2,
  min_nights_month = 5,
  probs = c(0.05, 0.95)
)
```

**Arguments**

|                  |                                                                                                                           |
|------------------|---------------------------------------------------------------------------------------------------------------------------|
| history          | Table returned by <code>history_draws()</code> .                                                                          |
| core_species     | Species codes (the species column) that enter the matrix.                                                                 |
| k_min            | Repetition rule: 1, 2 or 3 reliable detections on the night.                                                              |
| min_nights_month | Site x month cells with fewer recorded nights are dropped from the matrices (unstable frequencies) and flagged in effort. |
| probs            | Probabilities of the interval.                                                                                            |

**Value**

A list with summary (site, month, species, n\_nights\_recorded, frequency\_mean, frequency\_lower, frequency\_upper) and draws (draw, site, month, species, frequency).

---

compare\_calibration\_models

*Compare calibration models by cross-validation*


---

## Description

Fits each candidate model with k-fold cross-validation and scores the predicted probabilities on the held-out folds. Folds are stratified within site x species by default. With `fold_group`, recordings or nights are kept intact across species; see `calibration_folds()`. Models with random effects need the 'lme4' package.

## Usage

```
compare_calibration_models(
  validations,
  models = NULL,
  k_folds = 5,
  seed = 2026,
  fold_group = NULL
)
```

## Arguments

|                          |                                                                                                                                                                                                                                                                                                                   |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>validations</code> | Table returned by <code>read_validations()</code> .                                                                                                                                                                                                                                                               |
| <code>models</code>      | Named list of model formulas. <code>NULL</code> chooses the family from the observed species and sites: site-effect models for one species, the original family for multiple species, and no site random effects when only one site is observed. Supply an explicit list to reproduce a particular candidate set. |
| <code>k_folds</code>     | Number of folds.                                                                                                                                                                                                                                                                                                  |
| <code>seed</code>        | Seed of the fold assignment. The global random number generator is left as it was.                                                                                                                                                                                                                                |
| <code>fold_group</code>  | Optional character vector of grouping columns. Do not include species, verdicts or confidence scores: these would split recordings or nights by species or by outcome. Group IDs must be nonmissing and nonempty. <code>NULL</code> keeps the original clip folds.                                                |

## Details

The AUC measures discrimination (whether true detections get higher probabilities than false ones). The log-loss and the Brier score are proper scoring rules: they reward probabilities close to the observed outcomes, which is what matters when the probabilities are compared with `p_min` (Gneiting & Raftery 2007, [doi:10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437)). The calibration intercept and slope come from a logistic regression of the verdicts on the logit of the predicted probabilities in each test fold: a well calibrated model has intercept 0 and slope 1; a slope below 1 means probabilities too extreme (overfitting).

**Value**

A tibble with model, formula, auc\_mean, auc\_sd, n\_failed (folds whose fit failed, left out of the means), logloss\_mean, logloss\_sd, brier\_mean, brier\_sd, cal\_intercept and cal\_slope (means over the folds), prop\_singular (share of the fold fits that were singular) and complexity (position of the model in models), ordered from the highest mean AUC. Failed folds also raise a warning. A test fold with only one verdict has missing AUC but contributes log-loss and Brier; it is not a failed fit.

---

|                    |                                                |
|--------------------|------------------------------------------------|
| consensus_verdicts | <i>Consensus verdicts of several reviewers</i> |
|--------------------|------------------------------------------------|

---

**Description**

Combines the verdicts of several reviewers into one validation table: each clip keeps the verdict its reviewers agree on; clips with conflicting verdicts get verdict = "disagreement" and are listed for a consensus review. Disagreements are left out of the precision, thresholds and calibration, which use only positive and negative verdicts.

**Usage**

```
consensus_verdicts(reviews)
```

**Arguments**

|         |                                                                                                                                                         |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| reviews | One table of reviewed clips (as returned by <code>classify_review()</code> ) with a reviewer column, or a named list of such tables, named by reviewer. |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------|

**Value**

A tibble with one row per clip: the columns of the reviewed tables (from the first reviewer of each clip), verdict, n\_reviewers and reviewers (names separated by ;).

---

|           |                               |
|-----------|-------------------------------|
| cut_clips | <i>Cut the selected clips</i> |
|-----------|-------------------------------|

---

**Description**

Cuts each clip of the manifest from its source recording into <output\_dir>/clips/<species\_folder>/<clip\_file>, one folder per species, ready for the review tab of the BirdNET GUI. Clips that already exist are not cut again, so an interrupted run resumes where it stopped. Missing or unreadable recordings do not stop the run; they are flagged in status.

**Usage**

```
cut_clips(manifest, output_dir, simulate = FALSE)
```

**Arguments**

|            |                                                                                                                                                          |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| manifest   | Clip manifest, as returned by <a href="#">select_clips()</a> .                                                                                           |
| output_dir | Output folder; created if needed.                                                                                                                        |
| simulate   | If TRUE, no audio is cut (the empty species folders are still created), but status tells what would be done. Useful to check a selection before cutting. |

**Details**

Cutting uses the 'ffmpeg' library bundled with the 'av' package.

**Value**

The manifest with the column status after begin\_path: one of "ok", "already\_exists", "audio\_not\_found", "simulated" or "error: <message>".

---

|                   |                                  |
|-------------------|----------------------------------|
| detection_history | <i>Nightly detection history</i> |
|-------------------|----------------------------------|

---

**Description**

Aggregates the scored detections by site x species x night. A night has a detection when at least min\_detections\_night detections have a calibrated probability  $\geq p_{\min}$ ; the columns detection\_k1 to detection\_k3 give the same rule with  $k = 1, 2$  and  $3$ , as a sensitivity analysis.

**Usage**

```
detection_history(
  scored,
  p_min = 0.95,
  min_detections_night = 1,
  rule = c("mean", "probability"),
  prob_min = 0.9
)
```

**Arguments**

|                      |                                                                                                                                                                                                                                                                                                                                                                                                            |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| scored               | Table returned by <a href="#">score_detections()</a> .                                                                                                                                                                                                                                                                                                                                                     |
| p_min                | Threshold on the calibrated probability (not on the confidence score).                                                                                                                                                                                                                                                                                                                                     |
| min_detections_night | Repetition rule: minimum number of reliable detections for a night with detection.                                                                                                                                                                                                                                                                                                                         |
| rule                 | When a detection is reliable: "mean", its calibrated probability (the posterior mean with engine = "brms") is $\geq p_{\min}$ ; "probability", the posterior probability that $P(TP) \geq p_{\min}$ (the pr_reliable of a model fitted with engine = "brms") is $\geq prob_{\min}$ . The second rule propagates the uncertainty of the calibration and is more conservative where few clips were reviewed. |

prob\_min            rule = "probability": threshold on pr\_reliable. pr\_reliable is computed in [score\\_detections\(\)](#) for its p\_min; use the same p\_min in both.

### Value

A tibble with site, species, night, n\_detections, n\_reliable, p\_tp\_max, detection\_k1, detection\_k2, detection\_k3 and detection.

---

exclude\_training\_clips

*Leave the training clips out of a manifest*

---

### Description

Drops from a clip manifest the clips that overlap, for the same species and the same recording, a clip used to train the classifier (the training\_set.csv written by [build\\_training\\_set\(\)](#)). Use it before reviewing the detections of a retrained classifier: scores of the clips it was trained on are biased upwards.

### Usage

```
exclude_training_clips(manifest, training_set)
```

### Arguments

manifest            Clip manifest, as returned by [select\\_clips\(\)](#) or read from clip\_manifest.csv.  
 training\_set        Table returned by [build\\_training\\_set\(\)](#), or the path to its training\_set.csv.

### Value

The manifest without the overlapping clips.

---

extract\_validation\_clips

*Build the validation clip set*

---

### Description

Runs the whole first stage: reads the selection tables, selects the clips, cuts them, writes clip\_manifest.csv and zips one package per species. This is the function form of the v1.2.0 script extrair\_clipes\_validacao.R.

**Usage**

```
extract_validation_clips(
  data_dir,
  output_dir,
  ...,
  table_pattern = "^BirdNET_SelectionTable\\.txt$",
  make_zip = TRUE,
  simulate = FALSE,
  classifier = NULL
)
```

**Arguments**

|                            |                                                                                                                                                                                                                           |
|----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>data_dir</code>      | Root folder, searched recursively.                                                                                                                                                                                        |
| <code>output_dir</code>    | Output folder; created if needed.                                                                                                                                                                                         |
| <code>...</code>           | Sampling design, passed to <a href="#">select_clips()</a> (e.g. <code>sampling_mode</code> , <code>n_top</code> , <code>threshold</code> , <code>target_total</code> , <code>remap_from</code> , <code>remap_to</code> ). |
| <code>table_pattern</code> | Regular expression for the table file names.                                                                                                                                                                              |
| <code>make_zip</code>      | Zip one package per species (skipped when <code>simulate</code> is <code>TRUE</code> ).                                                                                                                                   |
| <code>simulate</code>      | If <code>TRUE</code> , no audio is cut (the empty species folders are still created), but <code>status</code> tells what would be done. Useful to check a selection before cutting.                                       |
| <code>classifier</code>    | Label of the classifier that produced the selection tables (e.g. "FrogNet v2, BirdNET-Analyzer 2.4.0"), recorded in <code>extraction_info.csv</code> . <code>NULL</code> records NA.                                      |

**Details**

Output structure:

```
<output_dir>/clips/<species_folder>/*.wav
<output_dir>/packages/<species_folder>.zip
<output_dir>/clip_manifest.csv
<output_dir>/extraction_info.csv
```

`extraction_info.csv` records the provenance of the run (one key, value row per item): the date, the versions of `sheepdog` and `R`, the folder of the selection tables, the classifier label and every argument passed to [select\\_clips\(\)](#). Keep it with the manifest: it tells later rounds which classifier produced the clips (for instance, to leave out of a validation round the clips used to retrain that classifier).

**Value**

The manifest written to `clip_manifest.csv`, invisibly.

## Examples

```
## Not run:
# Check the selection first, then cut
extract_validation_clips("N:/Arara", "N:/Arara/validation",
                        remap_from = "N:/", remap_to = "/Volumes/Arara/",
                        simulate = TRUE)

## End(Not run)
```

---

|                 |                                  |
|-----------------|----------------------------------|
| fit_calibration | <i>Fit the calibration model</i> |
|-----------------|----------------------------------|

---

## Description

Fits a calibration model to all the reviewed clips. The species and site levels of validations are stored with the model, so that [score\\_detections\(\)](#) can score new detections.

## Usage

```
fit_calibration(
  validations,
  formula,
  engine = c("lme4", "brms"),
  prior = NULL,
  ...
)
```

## Arguments

|             |                                                                                                            |
|-------------|------------------------------------------------------------------------------------------------------------|
| validations | Table returned by <a href="#">read_validations()</a> .                                                     |
| formula     | Model formula, usually the best one of <a href="#">compare_calibration_models()</a> .                      |
| engine      | "lme4" or "brms".                                                                                          |
| prior       | engine = "brms": the prior (a brmsprior object), e.g. from <a href="#">calibration_prior()</a> . Required. |
| ...         | engine = "brms": further arguments for <code>brms::brm()</code> (e.g. chains, iter, cores, backend, seed). |

## Details

With engine = "lme4" (default), the model is fitted by maximum likelihood ('lme4' for random effects, `glm()` otherwise). A model with random effects whose fit is singular (a variance estimated at zero or a correlation at +-1) raises a warning: the fixed effects and the predictions can still be used, but the variance components on the boundary should not be interpreted, and a simpler model may predict as well. Check them with `lme4::VarCorr()`. The other messages of `lme4::glmer()` are passed on, prefixed by `lme4:`, and every message of the fit is kept in the attribute `sheepdog_messages` of the model (convergence warnings of 'lme4' are warnings and are passed on as such).

With `engine = "brms"`, the model is fitted in a Bayesian framework with 'brms' (Bernoulli family, logit link). The prior must be given explicitly; `calibration_prior()` builds an example centred on an already calibrated score. After the fit, `calibration_diagnostics()` runs and a warning is raised when R-hat, the effective sample sizes or the divergent transitions fail the usual limits. Check the prior with `prior_predictive_calibration()`, the fit with `brms::pp_check()` and the sensitivity to the prior with `calibration_sensitivity()`.

## Value

A `glm`, `glmerMod` or `brmsfit` model, with the attribute `sheepdog_levels` (the species and site levels), used by `score_detections()` and `plot_calibration()`.

---

|               |                                                         |
|---------------|---------------------------------------------------------|
| history_draws | <i>Nightly detection histories over posterior draws</i> |
|---------------|---------------------------------------------------------|

---

## Description

Propagates the uncertainty of a Bayesian calibration model to the detection histories: for each of `n_draws` draws of the posterior, a detection is reliable when its probability in that draw is  $\geq p_{\min}$ , and the nightly history is rebuilt. The histories can then be turned into community matrices with uncertainty by `community_matrices_draws()`.

## Usage

```
history_draws(
  fit,
  scored,
  p_min = 0.95,
  n_draws = 200,
  min_detections_night = 1,
  min_confidence = 0.1,
  grid_step = 0.01,
  seed = 2026
)
```

## Arguments

|                                   |                                                                                      |
|-----------------------------------|--------------------------------------------------------------------------------------|
| <code>fit</code>                  | Model returned by <code>fit_calibration()</code> with <code>engine = "brms"</code> . |
| <code>scored</code>               | Table returned by <code>score_detections()</code> .                                  |
| <code>p_min</code>                | Threshold on the calibrated probability (not on the confidence score).               |
| <code>n_draws</code>              | Number of posterior draws.                                                           |
| <code>min_detections_night</code> | Repetition rule: minimum number of reliable detections for a night with detection.   |
| <code>min_confidence</code>       | Detections below this confidence are not scored.                                     |

|           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| grid_step | Models fitted with engine = "brms": step of the grid of scores on which the posterior is computed.<br>With a model fitted with engine = "brms", the posterior of P(TP) is computed once on a grid of scores (step grid_step) for each site x species and assigned to the detections by their score rounded to the grid, so that the cost does not depend on the number of detections; the rounding error is small next to the uncertainty of the calibration. Each detection then gets the posterior mean (p_tp), a lower quantile (p_tp_lower) and the posterior probability that P(TP) >= p_min (pr_reliable). <a href="#">detection_history()</a> can use either rule. |
| seed      | Seed of the choice of draws. The global random number generator is left as it was.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

### Details

As in [score\\_detections\(\)](#), the probabilities are computed on a grid of scores (step grid\_step) for each site x species and assigned to the detections by their rounded score, which keeps the cost independent of the number of detections.

### Value

A tibble with draw, site, species, night, n\_detections, n\_reliable, detection\_k1 to detection\_k3 and detection.

---

|                  |                                    |
|------------------|------------------------------------|
| nightly_presence | <i>Confirmed presence by night</i> |
|------------------|------------------------------------|

---

### Description

Summarises the verdicts by site, species and night. A night is "confirmed" when at least one clip is positive, "not\_detected\_in\_reviewed" when clips were reviewed but none was positive, and "not\_reviewed" otherwise.

### Usage

```
nightly_presence(validation)
```

### Arguments

validation      Table returned by [classify\\_review\(\)](#).

### Value

A tibble with site, species\_folder, night, n\_clips, n\_reviewed, n\_positive, presence and max\_confidence\_positive.

---

|            |                                                |
|------------|------------------------------------------------|
| pack_clips | <i>Pack the clips into one zip per species</i> |
|------------|------------------------------------------------|

---

### Description

Zips each species folder of <output\_dir>/clips into <output\_dir>/packages/<species\_folder>.zip: light validation packages to send to the reviewers.

### Usage

```
pack_clips(manifest, output_dir)
```

### Arguments

|            |                                                             |
|------------|-------------------------------------------------------------|
| manifest   | Clip manifest, as returned by <a href="#">cut_clips()</a> . |
| output_dir | The output_dir given to <a href="#">cut_clips()</a> .       |

### Value

The paths of the zip files, invisibly.

---

|                  |                                    |
|------------------|------------------------------------|
| plot_calibration | <i>Plot the calibration curves</i> |
|------------------|------------------------------------|

---

### Description

Calibrated probability of a true detection against the confidence score, one panel per species and one line per site, with the observed precision by confidence bin as points. Needs the 'ggplot2' package.

### Usage

```
plot_calibration(
  fit,
  validations,
  min_confidence = 0.1,
  p_min = 0.95,
  model_name = ""
)
```

### Arguments

|                |                                                                        |
|----------------|------------------------------------------------------------------------|
| fit            | Model returned by <a href="#">fit_calibration()</a> .                  |
| validations    | Table returned by <a href="#">read_validations()</a> .                 |
| min_confidence | Detections below this confidence are not scored.                       |
| p_min          | Threshold on the calibrated probability (not on the confidence score). |
| model_name     | Name shown in the subtitle.                                            |

**Value**

A ggplot object.

---

plot\_calibration\_draws

*Calibration curves of posterior (or prior) draws*

---

**Description**

Draws the population-level calibration curve,  $P(TP)$  against the score, for `n_draws` draws of a Bayesian calibration model, without the species and site effects. With a model from [prior\\_predictive\\_calibration\(\)](#) it shows what the prior implies on the probability scale (a prior predictive check); with a fitted model, the uncertainty of the curve. Needs the 'ggplot2' package.

**Usage**

```
plot_calibration_draws(fit, n_draws = 100, p_min = 0.95, seed = 2026)
```

**Arguments**

|                      |                                                                                                                               |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------|
| <code>fit</code>     | Model returned by <a href="#">fit_calibration()</a> with engine = "brms" or by <a href="#">prior_predictive_calibration()</a> |
| <code>n_draws</code> | Number of curves.                                                                                                             |
| <code>p_min</code>   | Threshold drawn as a dashed line.                                                                                             |
| <code>seed</code>    | Seed of the choice of draws. The global random number generator is left as it was.                                            |

**Value**

A ggplot object.

---

plot\_precision

*Plot the precision curves*

---

**Description**

One panel per species: precision by confidence bin, with the target precision as a dashed line and bins with low support shown faded. With more than one site, each site is a line of its own colour. Needs the 'ggplot2' package.

**Usage**

```
plot_precision(
  precision,
  bins = seq(0.1, 1, by = 0.1),
  target_precision = 0.9,
  min_n = 15
)
```

**Arguments**

|                  |                                                                                                                      |
|------------------|----------------------------------------------------------------------------------------------------------------------|
| precision        | Table returned by <code>precision_by_bin()</code> .                                                                  |
| bins             | Breaks of the confidence bins (left-closed).                                                                         |
| target_precision | Target precision.                                                                                                    |
| min_n            | Minimum of reviewed clips for a bin to be taken seriously; bins with fewer are flagged in <code>low_support</code> . |

**Value**

A ggplot object.

---

|                 |                                        |
|-----------------|----------------------------------------|
| post_validation | <i>Summarise the manual validation</i> |
|-----------------|----------------------------------------|

---

**Description**

Runs the whole post-validation step: classifies the reviewed clips and writes, to `output_dir`:

- `validation_by_clip.csv`: manifest + verdict of each clip (`classify_review()`);
- `precision_by_bin.csv`: precision by species x confidence bin (`precision_by_bin()`);
- `suggested_thresholds.csv`: lowest confidence reaching the target precision, by species (`suggest_thresholds()`);
- `nightly_presence.csv`: confirmed presence by site x species x night (`nightly_presence()`);
- `retraining_clips.csv`: positive clips for retraining (`retraining_clips()`);
- `precision_by_bin.png`: precision curves (`plot_precision()`), when 'ggplot2' is installed.

**Usage**

```
post_validation(
  review_dir,
  manifest_path,
  output_dir,
  positive_pattern = "positiv",
  negative_pattern = "negativ",
  bins = seq(0.1, 1, by = 0.1),
  target_precision = 0.9,
  min_n = 15
)
```

**Arguments**

|                                    |                                                                                                                   |
|------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| review_dir                         | Folder with the reviewed clips (the clips folder of the validation package, after the review).                    |
| manifest_path                      | Path to the manifest of the same run as the reviewed clips (clip_manifest.csv, or a v1.2.0 manifesto_clipes.csv). |
| output_dir                         | Output folder; created if needed.                                                                                 |
| positive_pattern, negative_pattern | Regular expressions (case-insensitive) that identify the positive and negative folders in the path of each clip.  |
| bins                               | Breaks of the confidence bins (left-closed).                                                                      |
| target_precision                   | Target precision.                                                                                                 |
| min_n                              | Minimum of reviewed clips for a bin to be taken seriously; bins with fewer are flagged in low_support.            |

**Details**

This is the function form of the v1.2.0 script pos\_validacao.R.

**Value**

A list with the tables validation, precision, thresholds, presence and retraining, invisibly.

---

|                  |                                    |
|------------------|------------------------------------|
| precision_by_bin | <i>Precision by confidence bin</i> |
|------------------|------------------------------------|

---

**Description**

Precision (share of positive verdicts) of the reviewed clips, by site, species and confidence bin. Clips without confidence (whole recordings of "random\_recordings") are left out.

**Usage**

```
precision_by_bin(validation, bins = seq(0.1, 1, by = 0.1), min_n = 15)
```

**Arguments**

|            |                                                                                                        |
|------------|--------------------------------------------------------------------------------------------------------|
| validation | Table returned by <a href="#">classify_review()</a> .                                                  |
| bins       | Breaks of the confidence bins (left-closed).                                                           |
| min_n      | Minimum of reviewed clips for a bin to be taken seriously; bins with fewer are flagged in low_support. |

**Value**

A tibble with site, species\_folder, bin, n\_reviewed, n\_positive, precision and low\_support.

---

prior\_predictive\_calibration

*Prior predictive calibration curves*

---

### Description

Samples the calibration model from the prior only (no data enter the likelihood), so that the curves it implies on the probability scale can be checked before the model is fitted. Plot them with [plot\\_calibration\(\)](#).

### Usage

```
prior_predictive_calibration(formula, validations, prior, ...)
```

### Arguments

|             |                                                                                    |
|-------------|------------------------------------------------------------------------------------|
| formula     | Calibration formula (see <a href="#">calibration_models()</a> ).                   |
| validations | Table returned by <a href="#">read_validations()</a> .                             |
| prior       | Prior, e.g. from <a href="#">calibration_prior()</a> .                             |
| ...         | Further arguments for <code>brms::brm()</code> (e.g. chains, iter, backend, seed). |

### Value

A `brmsfit` sampled from the prior, usable in [plot\\_calibration\(\)](#).

---

read\_selection\_tables *Read BirdNET selection tables*

---

### Description

Finds every BirdNET selection table (Raven format, tab-separated) under `data_dir`, reads them and stacks them in one table. Each table holds the predictions of one site, and the site is the name of the folder that contains it. The `output_` prefix of BirdNET output folders is removed, so `output_UEM/` becomes `site UEM`.

### Usage

```
read_selection_tables(
  data_dir,
  table_pattern = "^BirdNET_SelectionTable\\.txt$"
)
```

### Arguments

|               |                                              |
|---------------|----------------------------------------------|
| data_dir      | Root folder, searched recursively.           |
| table_pattern | Regular expression for the table file names. |

**Details**

Column names are standardised to lower case without units or punctuation, e.g. Begin Time (s) becomes begin\_time and File Offset (s) becomes file\_offset.

**Value**

A tibble with one row per prediction, the column site first, followed by the standardised table columns. The columns begin\_time, end\_time, species\_code, confidence, begin\_path and file\_offset are required.

---

read\_validation\_tables

*Read and stack the validation tables*


---

**Description**

Stacks the validation\_by\_clip.csv tables of several runs or sites (output of [post\\_validation\(\)](#)) and keeps every column and every verdict, unlike [read\\_validations\(\)](#), which keeps only what the calibration needs. Tables written by v1.2.0 (validacao\_por\_clipe.csv) are also read.

**Usage**

```
read_validation_tables(files)
```

**Arguments**

|       |                                                                                      |
|-------|--------------------------------------------------------------------------------------|
| files | Paths to the validation_by_clip.csv files. Missing files are skipped with a warning. |
|-------|--------------------------------------------------------------------------------------|

**Value**

A tibble with the columns of the validation tables; night and recording\_date are character.

---

read\_validations

*Read the validation tables for calibration*


---

**Description**

Stacks the validation\_by\_clip.csv tables of several sites (output of [post\\_validation\(\)](#)) and keeps the clips with a positive or negative verdict and a confidence score. Tables written by v1.2.0 (validacao\_por\_clipe.csv) are also read. Whole-recording clips of "random\_recordings" have no confidence and are left out.

**Usage**

```
read_validations(files, site_map = NULL, keep = NULL)
```

**Arguments**

|          |                                                                                                                                                                                                                  |
|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| files    | Paths to the validation_by_clip.csv files. Missing files are skipped with a warning.                                                                                                                             |
| site_map | Optional named character vector to fix diverging site names, as <code>c(old = "new")</code> , e.g. <code>c(BEP2 = "BEP")</code> .                                                                                |
| keep     | Optional character vector of input columns to retain, e.g. <code>c("source_audio", "night")</code> for grouped cross-validation. Missing columns raise an error rather than silently falling back to clip folds. |

**Value**

A tibble with site and species (factors; species is the species folder code), confidence, logit\_conf (logit of the confidence, bounded at  $1e-4$  and  $1 - 1e-4$ ) and tp (1 = true positive, 0 = false positive), plus the columns requested in keep.

---

|                  |                                      |
|------------------|--------------------------------------|
| retraining_clips | <i>Positive clips for retraining</i> |
|------------------|--------------------------------------|

---

**Description**

Lists the clips confirmed as positive, ready to retrain the classifier.

**Usage**

```
retraining_clips(validation)
```

**Arguments**

|            |                                                    |
|------------|----------------------------------------------------|
| validation | Table returned by <code>classify_review()</code> . |
|------------|----------------------------------------------------|

**Value**

A tibble with site, species\_code, species\_folder, clip\_file, reviewed\_path, confidence, night, recording\_hour and source\_audio.

---

|                    |                                    |
|--------------------|------------------------------------|
| reviewer_agreement | <i>Agreement between reviewers</i> |
|--------------------|------------------------------------|

---

### Description

Measures, for each species, the agreement between reviewers on the clips reviewed more than once: the proportion of agreement and kappa, which discounts the agreement expected by chance. When a species has exactly two reviewers, this is Cohen's kappa, with the marginal frequencies of each reviewer (Cohen 1960, [doi:10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)); with more reviewers, or different pairs of reviewers on different clips, Fleiss' kappa (Fleiss 1971, [doi:10.1037/h0031619](https://doi.org/10.1037/h0031619)). Only positive and negative verdicts count.

### Usage

```
reviewer_agreement(reviews)
```

### Arguments

|         |                                                                                                                                                         |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| reviews | One table of reviewed clips (as returned by <code>classify_review()</code> ) with a reviewer column, or a named list of such tables, named by reviewer. |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------|

### Value

A tibble with `species_folder`, `n_clips` (clips with at least two verdicts), `n_reviewers`, `agreement`, `kappa` and `method`.

---

|                  |                                                         |
|------------------|---------------------------------------------------------|
| score_detections | <i>Score every detection with the calibration model</i> |
|------------------|---------------------------------------------------------|

---

### Description

Reads the complete BirdNET selection tables (not only the reviewed clips) and predicts, for each detection, the probability that it is a true positive. Only species present in the calibration are scored. The night of each detection is its recording date, or the day before for recordings that start before dawn\_hour.

### Usage

```
score_detections(
  tables,
  fit,
  min_confidence = 0.1,
  dawn_hour = 12,
  p_min = 0.95,
  lower_prob = 0.05,
  grid_step = 0.01
)
```

**Arguments**

|                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| tables            | Named character vector of selection-table paths, named by site (the site names of the calibration).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| fit               | Model returned by <code>fit_calibration()</code> .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| min_confidence    | Detections below this confidence are not scored.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| dawn_hour         | Recordings that start before this hour belong to the night before.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| p_min, lower_prob | Models fitted with engine = "brms": the threshold of <code>pr_reliable</code> and the probability of the lower quantile <code>p_tp_lower</code> .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| grid_step         | Models fitted with engine = "brms": step of the grid of scores on which the posterior is computed.<br><br>With a model fitted with engine = "brms", the posterior of P(TP) is computed once on a grid of scores (step <code>grid_step</code> ) for each site x species and assigned to the detections by their score rounded to the grid, so that the cost does not depend on the number of detections; the rounding error is small next to the uncertainty of the calibration. Each detection then gets the posterior mean ( <code>p_tp</code> ), a lower quantile ( <code>p_tp_lower</code> ) and the posterior probability that $P(TP) \geq p\_min$ ( <code>pr_reliable</code> ). <code>detection_history()</code> can use either rule. |

**Value**

A tibble with site, species, date, night, hour, confidence, `p_tp`, `begin_path` and `file_offset`; with a model fitted with engine = "brms", also `p_tp_lower` and `pr_reliable`.

---

select\_calibration\_model

*Select the calibration model*

---

**Description**

Picks a model from the table of `compare_calibration_models()`.

**Usage**

```
select_calibration_model(
  comparison,
  rule = c("one_se_logloss", "best_auc"),
  k_folds = 5
)
```

**Arguments**

|            |                                                               |
|------------|---------------------------------------------------------------|
| comparison | Table returned by <code>compare_calibration_models()</code> . |
| rule       | Selection rule.                                               |
| k_folds    | Number of folds of the comparison (for the standard error).   |

### Details

- "best\_auc": the model with the highest mean AUC (the rule of `calibrate()` in 2.x).
- "one\_se\_logloss": among the models whose mean log-loss is within one standard error ( $\text{logloss\_sd} / \sqrt{k\_folds}$ ) of the lowest, the simplest (lowest complexity); models singular in most folds are left out, unless all are, which raises a warning. The log-loss rewards correct probabilities, not only their ranking, and the one-standard-error rule prefers the simpler model when the data cannot tell the candidates apart. This is the default of `calibrate()` since 3.0.0.

### Value

The name of the selected model.

---

|              |                                     |
|--------------|-------------------------------------|
| select_clips | <i>Select the clips to validate</i> |
|--------------|-------------------------------------|

---

### Description

Applies the sampling design to the predictions read by `read_selection_tables()` and returns the clip manifest: one row per clip to cut, with its site, species, night, source audio, start and end. No audio is read or written; cut the clips with `cut_clips()`.

### Usage

```
select_clips(
  detections,
  sampling_mode = c("top", "random_detections", "random_recordings"),
  occasion_unit = c("file", "minute"),
  threshold = NULL,
  n_top = 2,
  strata = "night_hour",
  recording_fraction = 0.01,
  n_recordings_per_site = NULL,
  recordings = NULL,
  audio_pattern = "\\.(wav|flac|mp3)$",
  recording_duration = NULL,
  min_confidence = 0.1,
  sites = NULL,
  date_start = NULL,
  date_end = NULL,
  species = NULL,
  hour_start = NULL,
  hour_end = NULL,
  alternate_days = FALSE,
  target_total = NULL,
  balance_by_hour = TRUE,
```

```

max_per_species = NULL,
remap_from = NULL,
remap_to = NULL,
seed = 2026
)

```

## Arguments

|                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| detections         | Predictions, as returned by <code>read_selection_tables()</code> .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| sampling_mode      | "top", "random_detections" or "random_recordings". See <i>Sampling modes</i> .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| occasion_unit      | The occasion of the occupancy model: "file" (the whole recording) or "minute" (minute of the file, from file_offset).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| threshold          | Confidence threshold of the occupancy analysis. It defines the candidate occasions of "random_detections" and the n_detections_occasion column. NULL leaves that column NA.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| n_top              | "top": predictions kept per site x species x night x hour. "random_detections": occasions drawn per stratum (strata).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| strata             | <p>"random_detections": the strata of the occasion draw (ignored by the other modes). Either one of the presets, all within site x species:</p> <ul style="list-style-type: none"> <li>• "night_hour": night x hour (the rule of 2.0.0);</li> <li>• "night": night;</li> <li>• "hour": hour of the day, pooling the nights (covers the diel cycle without fixing the night);</li> <li>• "week": ISO week of the night (YYYY-Www);</li> <li>• "month": month of the night (YYYY-MM), which spreads the draw over the season;</li> <li>• "site_species": no further stratification;</li> </ul> <p>or a character vector of column names of detections, e.g. c("site", "species_code", "habitat", "night"), for strata defined by the user. Columns can be added to detections before the call (a habitat, a recorder, the moon phase, the rain of the night); month and week can also be used and are derived from the night when absent. site and species_code are always added to the strata, which keeps inclusion_prob exact when target_total or max_per_species draw again within site x species. Columns that depend on the score (confidence, rank) are refused, since the observation-confirmation design needs strata independent of the score, and every stratum column must be constant within an occasion.</p> <p>With one recording per hour and occasion_unit = "file", "night_hour" holds a single occasion per stratum and draws nothing (every inclusion_prob is 1); a coarser stratum makes the draw random. With strata without the hour and balance_by_hour = TRUE under target_total or max_per_species, inclusion_prob is conditional on the first draw (a warning says so).</p> |
| recording_fraction | <p>"random_recordings": fraction of the recordings (or minutes) drawn at each site, rounded, with at least one per site. The default (1 %) suits large projects: with fewer than 150 recordings per site it draws a single recording; raise it or use n_recordings_per_site in small projects.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| n_recordings_per_site | "random_recordings": fixed number of recordings (or minutes) per site, used instead of recording_fraction.                                                                                                                                                                                                                                                                                                                                      |
| recordings            | "random_recordings": the list of all recordings, as a data frame or the path to a CSV file with the columns site and source_audio (the v1.2.0 columns local and caminho_audio are also read). NULL lists the files matching audio_pattern in the audio folders that appear in begin_path; folders without any prediction are then left out.                                                                                                     |
| audio_pattern         | Regular expression for the audio files listed when recordings is NULL.                                                                                                                                                                                                                                                                                                                                                                          |
| recording_duration    | Length of the recordings (s), for occasion_unit = "minute" in "random_recordings" and for the end of whole-recording clips. NULL reads each file with the 'av' package when needed (minute occasions) and leaves clip_end as NA otherwise.                                                                                                                                                                                                      |
| min_confidence        | Minimum confidence of a candidate prediction in "top". NULL applies no filter. The random modes use threshold.                                                                                                                                                                                                                                                                                                                                  |
| sites                 | Restrict to these sites. NULL keeps all.                                                                                                                                                                                                                                                                                                                                                                                                        |
| date_start, date_end  | Date window of the recordings ("YYYYMMDD"). NULL leaves that side open.                                                                                                                                                                                                                                                                                                                                                                         |
| species               | Restrict to these short species codes (the last part of species_code, e.g. "boa-ran"). NULL keeps all.                                                                                                                                                                                                                                                                                                                                          |
| hour_start, hour_end  | Hour window (0-23). It can cross midnight: hour_start = 17, hour_end = 2 keeps 17h-23h and 0h-2h, and the grouping becomes by night (small hours count as the night before). NULL applies no filter.                                                                                                                                                                                                                                            |
| alternate_days        | Keep alternate days (or nights), anchored on the first date of each site.                                                                                                                                                                                                                                                                                                                                                                       |
| target_total          | Total clips per site, split evenly among species by water filling: species with fewer clips than the quota enter whole and the surplus goes to the others. In "random_detections" it counts occasions. Ignored in "random_recordings". NULL sets no target. Because the target is split evenly, rare species have higher inclusion probabilities than common ones: estimates pooled across species must weight the clips by 1 / inclusion_prob. |
| balance_by_hour       | Within each species, draw round-robin across the hours of the day (covers the diel cycle) when applying target_total or max_per_species.                                                                                                                                                                                                                                                                                                        |
| max_per_species       | Simpler alternative to target_total: a fixed limit per site x species. Ignored when target_total is set.                                                                                                                                                                                                                                                                                                                                        |
| remap_from, remap_to  | Replace the prefix remap_from of the begin_path paths by remap_to, e.g. "N:/" and "/Volumes/Arara/". Backslashes are turned into forward slashes first.                                                                                                                                                                                                                                                                                         |
| seed                  | Seed of every random draw. The global random number generator is left as it was.                                                                                                                                                                                                                                                                                                                                                                |

**Value**

A tibble, the clip manifest without status, with the columns site, species\_code, species\_folder, recording\_date, night, recording\_hour, rank, confidence, clip\_file, clip\_start, clip\_end, duration, source\_audio, begin\_path, sampling\_mode, occasion\_id, n\_detections\_occasion and inclusion\_prob. See vignette("crywolf-integration") for their definitions.

**Sampling modes**

- "top" keeps the n\_top predictions with the highest confidence per site x species x night x hour. It serves classifier retraining and the score-to-probability calibration of `calibrate()`. Do not use it to estimate false-positive rates: top-score clips hold fewer false positives than a random sample.
- "random\_detections" draws, at random, n\_top occasions per stratum (by default site x species x night x hour; see strata) among the occasions with at least one prediction of the species with confidence  $\geq$  threshold, and keeps every prediction  $\geq$  threshold of each drawn occasion. This is the observation-confirmation design of Chambert et al. (2015, doi:10.1890/141507.1).
- "random\_recordings" draws recordings (or minutes) among all those recorded, with or without predictions, in proportion to the effort of each site, for annotation by ear. This is the multiple-method design of Miller et al. (2011, doi:10.1890/101396.1). The clip is the whole recording (or minute), copied once per species to annotate, and confidence is NA.

In both random modes, use the same threshold as the occupancy analysis.

---

split\_by\_reviewer

---

*Copy each reviewer's clips to their own folder*


---

**Description**

Copies the cut clips of a manifest with reviewers (`assign_reviewers()`) to <output\_dir>/<reviewer>/<species\_folder>. one folder tree per reviewer, ready for the review tab of the BirdNET GUI. Clips with a second reviewer are copied to both folders. After the review, classify each reviewer's folder with `classify_review()`.

**Usage**

```
split_by_reviewer(manifest, clips_dir, output_dir)
```

**Arguments**

|            |                                                                                  |
|------------|----------------------------------------------------------------------------------|
| manifest   | Manifest returned by <code>assign_reviewers()</code> .                           |
| clips_dir  | Folder of the cut clips (the clips folder written by <code>cut_clips()</code> ). |
| output_dir | Output folder.                                                                   |

**Value**

A tibble with reviewer, species\_folder, clip\_file and to, invisibly.

---

`suggest_core_species`     *Suggest the core species of the community matrices*

---

### Description

Keeps the species with no unreliable site x species cell in the table of `cell_status()`: reliable at every site, or reliable where present and probably absent elsewhere.

### Usage

```
suggest_core_species(status)
```

### Arguments

`status`                      Table returned by `cell_status()`.

### Value

A character vector of species codes (the species column of the history).

---

`suggest_thresholds`        *Suggest a confidence threshold per species*

---

### Description

For each site and species, the lowest confidence  $c$  such that the precision among the reviewed clips with confidence  $\geq c$  reaches `target_precision`, with at least `min_n` clips. Candidate cuts are the observed confidences rounded to two decimals.

### Usage

```
suggest_thresholds(validation, target_precision = 0.9, min_n = 15)
```

### Arguments

`validation`                Table returned by `classify_review()`.

`target_precision`  
                              Target precision.

`min_n`                      Minimum of reviewed clips for a bin to be taken seriously; bins with fewer are flagged in `low_support`.

### Value

A tibble with `site`, `species_folder`, `threshold`, `n_support` and `precision_at_threshold` (NA when no cut reaches the target). Without any reviewed clip with a confidence, the tibble has no rows.

---

|                   |                                                            |
|-------------------|------------------------------------------------------------|
| summarise_history | <i>Summarise the detection history by site and species</i> |
|-------------------|------------------------------------------------------------|

---

**Description**

Summarise the detection history by site and species

**Usage**

```
summarise_history(history)
```

**Arguments**

history      Table returned by [detection\\_history\(\)](#).

**Value**

A tibble with site, species, nights (nights with any scored detection) and nights\_k1 to nights\_k3 (nights with detection under each repetition rule).

# Index

assign\_reviewers, 3  
assign\_reviewers(), 36

build\_community\_matrices, 3  
build\_training\_set, 4  
build\_training\_set(), 19

calibrate, 5  
calibrate(), 3, 14, 33, 36  
calibration\_diagnostics, 8  
calibration\_diagnostics(), 22  
calibration\_effort, 8  
calibration\_effort(), 5  
calibration\_folds, 9  
calibration\_folds(), 16  
calibration\_models, 10  
calibration\_models(), 6, 7, 11, 28  
calibration\_prior, 11  
calibration\_prior(), 21, 22, 28  
calibration\_sensitivity, 12  
calibration\_sensitivity(), 11, 22  
cell\_status, 12  
cell\_status(), 37  
classify\_review, 13  
classify\_review(), 5, 17, 23, 26, 27, 30, 31, 36, 37  
community\_matrices, 14  
community\_matrices(), 3, 4, 15  
community\_matrices\_draws, 15  
community\_matrices\_draws(), 22  
compare\_calibration\_models, 16  
compare\_calibration\_models(), 5, 9, 10, 21, 32  
consensus\_verdicts, 17  
cut\_clips, 17  
cut\_clips(), 3, 24, 33, 36

detection\_history, 18  
detection\_history(), 5, 12, 14, 23, 32, 38

exclude\_training\_clips, 19

exclude\_training\_clips(), 5  
extract\_validation\_clips, 19

fit\_calibration, 21  
fit\_calibration(), 7, 8, 11, 12, 22, 24, 25, 32

history\_draws, 22  
history\_draws(), 15

nightly\_presence, 23  
nightly\_presence(), 26

pack\_clips, 24  
plot\_calibration, 24  
plot\_calibration(), 5, 22, 28  
plot\_calibration\_draws, 25  
plot\_precision, 25  
plot\_precision(), 26  
post\_validation, 26  
post\_validation(), 5, 29  
precision\_by\_bin, 27  
precision\_by\_bin(), 26  
prior\_predictive\_calibration, 28  
prior\_predictive\_calibration(), 11, 22, 25

read\_selection\_tables, 28  
read\_selection\_tables(), 33, 34  
read\_validation\_tables, 29  
read\_validation\_tables(), 5  
read\_validations, 29  
read\_validations(), 9, 11, 12, 16, 21, 24, 28, 29  
retraining\_clips, 30  
retraining\_clips(), 26  
reviewer\_agreement, 31  
reviewer\_agreement(), 3

score\_detections, 31  
score\_detections(), 5, 7, 18, 19, 21–23

`select_calibration_model`, 32  
`select_calibration_model()`, 6, 7, 10  
`select_clips`, 33  
`select_clips()`, 3, 18–20  
`split_by_reviewer`, 36  
`split_by_reviewer()`, 3  
`suggest_core_species`, 37  
`suggest_core_species()`, 13  
`suggest_thresholds`, 37  
`suggest_thresholds()`, 26  
`summarise_history`, 38  
`summarise_history()`, 5