

Package: FlashAustria (via r-universe)

October 7, 2026

Version 0.4-3

Date 2026-10-05

Title Data and Model for Reanalyzing Flash Counts in Austria

Description Lightning discharges counted on a 32 x 32 sq-km grid over Austria, Europe, are provided for the years 2010-2018 and the months May-August. A subset of atmospheric parameters from the fifth ECMWF re-analysis (ERA5) is interpolated by applying the nearest neighbor method to the grid of the counted lightning discharges.

Depends R (>= 3.5.0)

Imports bamlss, gamlss.dist, gamlss.tr, sf

License GPL-2 | GPL-3

URL <https://zeileis.codeberg.page/FlashAustria/>

BugReports <https://codeberg.org/zeileis/FlashAustria/issues>

Config/pak/sysreqs libabsl-dev cmake libgdal-dev gdal-bin libgeos-dev libssl-dev libproj-dev libsqlite3-dev libudunits2-dev

Repository <https://zeileis.r-universe.dev>

Date/Publication 2026-10-06 13:30:17 UTC

RemoteUrl <https://codeberg.org/zeileis/FlashAustria>

RemoteRef HEAD

RemoteSha 98e7d562415e73b44d140bd00e4dc0fa51f7a471

Contents

FlashAustria	2
FlashAustriaModel	3
ThunderstormAustria	5

Index	7
--------------	---

Description

Lightning discharges counted on 32 x 32 sq-km grid over Austria, Europe, for the years 2010–2018 and the month May–August, detected by the ground-based Austrian Lightning Detection and Information System (ALDIS). A selection of atmospheric parameters from the fifth ECMWF re-analysis (ERA5) interpolated nearest neighbour method to the grid of the counted lightning discharges. The temporal resolution of the data is hourly.

Usage

```
data("FlashAustria", package = "FlashAustria")
```

Format

Three objects FlashAustriaTrain, FlashAustriaEval and FlashAustriaCase are added to the environment. FlashAustriaTrain and FlashAustriaEval are data.frames. FlashAustriaCase is an "sf" object, containing a geometry column providing the 32 x 32 sq-km grid as polygons.

counts ALDIS counts of cloud-to-ground lightning discharges.

hcc High cloud cover. The proportion of a grid box covered by cloud occurring in the high levels of the troposphere (approx. 6 km and above).

str Surface net thermal radiation. Thermal radiation refers to radiation emitted by the atmosphere, clouds and the surface of the Earth. str is the difference between downward and upward thermal radiation at the surface of the Earth. The units are joules per square metre as the radiation is accumulated over a period of one hour.

tciw Total column cloud ice water. The amount of ice contained within clouds in a column extending from the surface of the Earth to the top of the atmosphere.

q_prof_PC1 This variable has been derived from the vertical profile of specific humidity (labelled q at the ECMWF) by principal component analysis. This is the first principal component.

sqr_t_cape The square root of convective available potential energy. This is an indication of the (in)stability of the atmosphere.

Details

Only a random subset of the data given positive counts is provided for training and evaluation: FlashAustriaTrain and FlashAustriaEval contain 12000 and 6000 observations, respectively. The following prediction cases are provided: 2019-09-15 06 UTC, 2001-09-15 17 UTC, 2001-09-15 23 UTC and 2001-09-16 13 UTC.

References

To cite the flash count data (observations), please use:

Schulz W, Cummins K, Diendorfer G, Dorninger M (2005). Cloud-to-Ground Lightning in Austria: A 10-Year Study Using Data from a Lightning Location System. *Journal of Geophysical Research: Atmospheres*, **110**(D9), D09101. doi:[10.1029/2004JD005332](https://doi.org/10.1029/2004JD005332)

Schulz W, Diendorfer G, Pedebay S, Poelman DR (2016). The European Lightning Location System EUCLID – Part 1: Performance Analysis and Validation. *Natural Hazards and Earth System Sciences*, **16**(2), 595–605. doi:[10.5194/nhess165952016](https://doi.org/10.5194/nhess165952016)

To cite the reanalyses data (covariates), please use:

Copernicus Climate Change Service (C3S) (2017). ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. Copernicus Climate Change Service Climate Data Store (CDS), date of access: June 2019. <https://cds.climate.copernicus.eu/#!/home>

Examples

```
data("FlashAustria", package = "FlashAustria")
head(FlashAustriaTrain)
nrow(FlashAustriaTrain)
nrow(FlashAustriaEval)

hist(FlashAustriaTrain$counts,
     breaks = c(seq(.5, 50.5), Inf),
     xlim = c(0, 50),
     col = "lightgray"
)

plot(log(counts) ~ log(sqrt_cape), FlashAustriaTrain)
plot(log(counts) ~ d2m, FlashAustriaTrain)
plot(log(counts) ~ q_prof_PC1, FlashAustriaTrain)
plot(log(counts) ~ cswc_prof_PC4, FlashAustriaTrain)
plot(log(counts) ~ t_prof_PC1, FlashAustriaTrain)
plot(log(counts) ~ v_prof_PC2, FlashAustriaTrain)
plot(log(counts) ~ sqrt_lsp, FlashAustriaTrain)

library("sf")
class(FlashAustriaCase)
head(FlashAustriaCase)
plot(st_geometry(FlashAustriaCase))
plot(FlashAustriaCase)
```

FlashAustriaModel

Fitted Distributional Regression Model Object (bamIss)

Description

A zero-truncated negative binomial `bamIss` model trained on the `FlashAustriaTrain` data using gradient boosting with subsequent MCMC sampling. The response variable are the lightning counts and the regression terms are P-splines based on ERA5 covariates.

Usage

```
data("FlashAustriaModel", package = "FlashAustria")
```

Format

An object of class `bamlss`.

References

Umlauf N, Klein N, Zeileis A (2018). BAMLSS: Bayesian Additive Models for Location, Scale and Shape (and Beyond). *Journal of Computational and Graphical Statistics*, **27**(3), 612–627. [doi:10.1080/10618600.2017.1407325](https://doi.org/10.1080/10618600.2017.1407325)

Examples

```
## Visualization of fitted model
if(require("bamlss") && require("gamlss.dist") && require("gamlss.tr")) {
  data("FlashAustria", package = "FlashAustria")
  data("FlashAustriaModel", package = "FlashAustria")
  ztSICHEL <- trun(0, family = "SICHEL", local = FALSE)
  plot(flash_model_ztSICHEL)
}

## Replication code:
## Not run:
## Required packages.
library("bamlss")
library("gamlss.dist")
library("gamlss.tr")

## Load the data.
data("FlashAustria", package = "FlashAustria")

## Generate zero truncated Sichel distribution.
ztSICHEL <- trun(0, family = "SICHEL", local = FALSE)

## Model formula, up to four parameters.
f <- list(
  counts ~ s(d2m, bs = "ps") + s(q_prof_PC1, bs = "ps") +
    s(cswc_prof_PC4, bs = "ps") + s(t_prof_PC1, bs = "ps") +
    s(v_prof_PC2, bs = "ps") + s(sqrt_cape, bs = "ps"),
  ~ s(sqrt_lsp, bs = "ps")
)

## Estimate models.
set.seed(123)
flash_model_ztnbinom <- bamlss(f,
  family = "ztnbinom", data = FlashAustriaTrain,
  optimizer = opt_boost, cores = 3, light = TRUE, binning = TRUE,
  maxit = 1000, thin = 3, burnin = 1000, n.iter = 2000)

set.seed(123)
```

```
flash_model_ztSICHEL <- bamlss(f,
  family = ztSICHEL, data = FlashAustriaTrain,
  optimizer = opt_boost, cores = 3, light = TRUE, binning = TRUE,
  maxit = 1000, thin = 3, burnin = 1000, n.iter = 2000)

## End(Not run)
```

ThunderstormAustria *Thunderstorm Occurrence in Austria 2010–2018*

Description

Thunderstorm occurrence on 32 x 32 sq-km grid over Austria, Europe, for the years 2010–2018 and the month May–August, derived by the ground-based Austrian Lightning Detection and Information System (ALDIS). A selection of atmospheric parameters from the fifth ECMWF re-analysis (ERA5) interpolated nearest neighbour method to the grid of the counted lightning discharges. The temporal resolution of the data is hourly.

Usage

```
data("ThunderstormAustria", package = "FlashAustria")
```

Format

Two objects ThunderstormAustriaTrain and ThunderstormAustriaEval are added to the environment, both of which are `data.frames`.

tunderstorm factor indicating the occurrence of a tunderstorm.

d2m 2m dewpoint temperature.

q_prof_PC1 First principal component of the vertical profile of specific humidity.

cswc_prof_PC4 Fourth principal component of the vertical profile of specific snow water content.

t_prof_PC1 First principal component of the vertical profile of temperature.

v_prof_PC2 Second principal component of the vertical profile of v component of wind.

sqrt_cape Square root of convective available potential energy.

sqrt_lsp Square root of large scale precipitation.

Details

The World Meteorological Organization (WMO) defines a thunderstorm as follows: “One or more sudden electrical discharges, manifested by a flash of light (lightning) or rumbling sound (thunder).”

Following this definition the thunderstorm occurrence is derived from the ALDIS lightning detection data: If at least one lightning discharge was detected in a 32 x 32 sq-km grid, it indicates the occurrence of a thunderstorm. Only a random subset of the data given, thunderstorm occurrence is provided for training and evaluation: ThunderstormAustriaTrain and ThunderstormAustriaEval contain 40000 and 20000 observations, respectively.

References

To cite the flash count data (observations), please use:

Schulz W, Cummins K, Diendorfer G, Dorninger M (2005). Cloud-to-Ground Lightning in Austria: A 10-Year Study Using Data from a Lightning Location System. *Journal of Geophysical Research: Atmospheres*, **110**(D9), D09101. doi:10.1029/2004JD005332

Schulz W, Diendorfer G, Pedebay S, Poelman DR (2016). The European Lightning Location System EUCLID – Part 1: Performance Analysis and Validation. *Natural Hazards and Earth System Sciences*, **16**(2), 595–605. doi:10.5194/nhess165952016

To cite the reanalyses data (covariates), please use:

Copernicus Climate Change Service (C3S) (2017). ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. Copernicus Climate Change Service Climate Data Store (CDS), date of access: June 2019. <https://cds.climate.copernicus.eu/#!/home>

The WMO's definition of thunderstorm:

<https://cloudatlas.wmo.int/thunderstorm.html>

Examples

```
data("ThunderstormAustria", package = "FlashAustria")
head(ThunderstormAustriaTrain)
nrow(ThunderstormAustriaTrain)
nrow(ThunderstormAustriaEval)

prop.table(table(ThunderstormAustriaTrain$thunderstorm))

plot(thunderstorm ~ log(sqrt_cape), ThunderstormAustriaTrain, ylevels = 2:1,
     breaks = 0:17/4)
plot(thunderstorm ~ d2m, ThunderstormAustriaTrain, ylevels = 2:1)
plot(thunderstorm ~ q_prof_PC1, ThunderstormAustriaTrain, ylevels = 2:1)
plot(thunderstorm ~ cswc_prof_PC4, ThunderstormAustriaTrain, ylevels = 2:1,
     breaks = c(-Inf, -4, -3.3, -3.2, -3, -2, 0, 10, Inf) / 1e6)
plot(thunderstorm ~ t_prof_PC1, ThunderstormAustriaTrain, ylevels = 2:1)
plot(thunderstorm ~ v_prof_PC2, ThunderstormAustriaTrain, ylevels = 2:1)
plot(thunderstorm ~ sqrt_lsp, ThunderstormAustriaTrain, ylevels = 2:1,
     breaks = c(0, 0.0001, 0.002, 0.005, 0.008, 0.012, 0.015, 0.02, 0.05))
```

Index

* datasets

FlashAustria, [2](#)

FlashAustriaModel, [3](#)

ThunderstormAustria, [5](#)

bamlss, [3](#), [4](#)

flash_model_ztnbinom

(FlashAustriaModel), [3](#)

flash_model_ztSICHEL

(FlashAustriaModel), [3](#)

FlashAustria, [2](#)

FlashAustriaCase (FlashAustria), [2](#)

FlashAustriaEval (FlashAustria), [2](#)

FlashAustriaModel, [3](#)

FlashAustriaTrain, [3](#)

FlashAustriaTrain (FlashAustria), [2](#)

ThunderstormAustria, [5](#)

ThunderstormAustriaEval

(ThunderstormAustria), [5](#)

ThunderstormAustriaTrain

(ThunderstormAustria), [5](#)