



# CODE MANUAL – USER GUIDE

María Bugallo Porto<sup>1</sup>

<sup>1</sup>Operations Research Center, Miguel Hernández University of Elche, Spain

## About the project

In this repository you will find the complete implementation of the model proposed in the paper entitled “*Zero-inflated negative binomial mixed models for predicting number of wildfires*”. In addition, we have included the necessary code to provide all graphical representations and tabular results.

## Installation and requirements

This project was performed on a desktop computer with a 16 GB RAM computer processing unit Intel®Core™ i7 and under the 64-bit Microsoft Windows 10 operating system. Moreover, it is pure implemented in R. Specifically, we have decided to use the version 4.1.3 (2022-03-10), entitled “One Push-Up”. As no other external software is used, no additional installation is required.

For data manipulation and preprocessing tasks, the following R packages are needed:

1. [readxl](#) (version 1.3.1).
2. [meza](#) (version 3.2-8).
3. [tidyverse](#) (version 1.3.2) and, in particular, [dplyr](#) (version 1.0.9).

The core structure have strong dependencies with the following R packages:

1. [RcmdrMisc](#) (version 2.7-2).
2. [MASS](#) (version 7.3-55).
3. [glmmTMB](#) (version 1.1.3), which is linked to the [TMB](#) (version 1.8.0) (Template Model Builder) package with the aim of providing generalised linear mixed model fits.

As optional dependencies to support the graphical representation of the results through the use of maps and colour scales, the following R packages are also required:

1. [maptools](#) (version 1.1-3).
2. [RColorBrewer](#) (version 1.1-2).

In order to facilitate the manipulation of the code, we recommend the GUI, “Graphical User Interface”, [R Studio](#). Once all the libraries are installed, simply download the project sources and run the scripts.

## Getting started

Some instructions are given to get a copy of the project ready to use on your local machine, understanding its purpose and all the results achieved. Contact us to report any problem.

To install the development version, run the following statements in the console (R or R Studio).

```
> # install.packages("devtools")
> devtools::install_github("mbugallo/aZINB11Fires")
```

## Repository structure

The data files and input-output information has been stored in the following files and formats:

- [AEMET\\_year\\_month\\_province.csv](#) and [unrate\\_year\\_month\\_province.csv](#): files containing information related to the auxiliary variables, aggregated by year, month and province.

Specifically, meteorological data (01/2000–12/2015 period), provided by the Spain Meteorological Agency (AEMET), and data related to unemployment rates (01/2002–12/2015 period), provided by the Spanish Statistical Office (INE), were used.

- [fires\\_year\\_month\\_province.xlsx](#): file containing all forest fires recorded in Spain at the provincial level for all months of the period 01/1999–12/2015. That is, the target variable.
- [raw\\_data.csv](#) and [prepared\\_data.csv](#) store the aggregated information of the target and auxiliary variables. The first file is prior to the preprocessing tasks and the second one is obtained as a result of them. Both files are stored in the folder **MODEL\_DATA**.
- [pred2015.txt](#) and [pred\\_summer2015.txt](#): provincial forecasts for 2015 and, in particular, for the months of June, July, August and September.
- [RSPE2015.txt](#): relative squared prediction error (RSPE) by month for 2015.
- [RRMSE\\_var\\_Fires500.csv](#): root mean squared error (RMSE), relative root mean squared error (RRMSE) and variance for 2015 forecasts, all estimated by bootstrap with 500 resamples.
- **SPAIN\_MAP**: folder that contains several files necessary for mapping the Spanish results.

The source code is structured in the following scripts:

- [Data\\_and\\_preprocessing.R](#): Script for data preprocessing tasks.
- [ZINBmodel.R](#): Main script to perform the selection, fitting and diagnosis of the model. In particular, a zero-inflated negative binomial mixed model is fitted to the data in order to describe both the number of fires and its non-occurrence, as well as providing good predictive tools. It includes plots and creates external files containing predictions and error measures.
- [Bootstrap\\_resampling500.R](#): This script calculates confidence intervals of model parameters and error measures of fire predictions at the provincial level using bootstrap resampling.
- [Spanish\\_maps.R](#): Essential script for the graphical interpretation of the results. It provides maps for 2015 Spanish predicted values and error measures.

## License

This project is licensed under the terms of the license GNU Affero GPL v3.0.

### INFORMATION

#### Researchers

María Bugallo Porto	–	Miguel Hernández University of Elche
María Dolores Esteban Lefler	–	Miguel Hernández University of Elche
Manuel Francisco Marey Pérez	–	University of Santiago de Compostela
Domingo Morales González	–	Miguel Hernández University of Elche

### DOWNLOAD

Github repository    →    [Click here](#)