Outbreaks

Mladen Cucak

5/4/2020

## Libraries

## [1] "All packages were successfully loaded."

##Background I guess the whole idea about this is related to some of my PhD projects that I could not finish. I had a pile of yearbooks with initial disease outbreaks and follow-up assessments from early 1950s to late 1990s. The thickness of these yearbooks was slimmer and slimmer towards end of the last century, hence, beside the historical and theoretical value they would not be of much use for my current project. I have given up on digitising them, because both myself and few of my students started hating me, because they are really messy.  
(Very relevant paper by Zwankhuizen & Zadoks)[<https://bsppjournals.onlinelibrary.wiley.com/doi/full/10.1046/j.1365-3059.2002.00738.x>], which I think I have mentioned to Neil.  
The main idea related to the current data set is that it is a bit more relevant to current agrosystem. It might be interesting to see how the initial outbreaks, their distribution, rate of increase, etc., is changes over years. This is a observational “study” and comes with all its caveats. These are reports by blight scouts, and it is hard to know if some scouts are more active than others, and similar variables. However, this data is collected by Louise Cooke, and I am sure she would be happy to answer any queries, if it comes to that.

## Data

The disease outbreak data consisted of the date and coordinates of 352 late blight outbreaks from across Northern Ireland over an 11- year period (2005-2014). The NI data were collected every year as part of the Agriculture and Horticulture Development Board (AHDB) Potatoes ‘Fight Against Blight’ campaign (<https://potatoes.ahdb.org.uk/>). The coordinates of blight outbreak locations were obtained using the ‘geocoding’ function from the ‘ggmap’ package (Kahle and Wickham 2013) and confirmed manually. The same data set was used in revision of the Hutton Critera, previously known as Smith Periods (some details [here](https://euroblight.net/fileadmin/euroblight/Workshops/AArhus/Proceedings/5._Siobhan_Dancy-p53-58.pdf)).

A look at data set.

Table continues below

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| date | yr | source | variety | varietyII | Location |
| 2003-05-21 | 2003 | crop | Lady Felicia | NA | Ards |
| 2003-05-31 | 2003 | crop | Home Guard | NA | BallyLagen,BClare |
| 2003-06-05 | 2003 | crop | Kerrs Pink | NA | Benagh, Kilkeel |
| 2003-06-06 | 2003 | crop | Kerrs Pink | NA | Glenkeen Colraine |
| 2003-06-06 | 2003 | crop | Home Guard | NA | Pollee B.Mena |
| 2003-06-07 | 2003 | crop | Kerrs Pink | NA | Dunamoy B.Clare |

Table continues below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| County | Comments | loc\_full | lon | lat |
| Down | NA | Ards, Down, UK | -5.664 | 54.66 |
| Antrim | NA | BallyLagen,BClare, Antrim, UK | -5.922 | 54.76 |
| Down | protected crop | Benagh, Kilkeel, Down, UK | -6.078 | 54.05 |
| Londonderry | NA | Glenkeen Colraine, Londonderry, UK | -7.305 | 54.99 |
| Antrim | NA | Pollee B.Mena, Antrim, UK | -6.121 | 54.55 |
| Antrim | NA | Dunamoy B.Clare, Antrim, UK | -6.041 | 54.78 |

|  |
| --- |
| jday |
| 141 |
| 151 |
| 156 |
| 157 |
| 157 |
| 158 |

Varieties present in the data. I know most of them, and all are traditional varieties with not partial resistance.

## [1] "Lady Felicia" "Home Guard"   
## [3] "Kerrs Pink" "British Queen"   
## [5] "Desiree" "Saxon"   
## [7] NA "Up to Date"   
## [9] "Dunbar Standard" "D.Standard"   
## [11] "King Edward" "Dundrod"   
## [13] "Sante" "Nicola"   
## [15] "Red Duke of York" "Navan"   
## [17] "Rooster" "Pentland Squire"   
## [19] "Arran Victory" "Milagro"   
## [21] "Sharpes Express" "Arran Banner"   
## [23] "Caberet" "Dunluce"   
## [25] "Maris Piper" "Desiree & Nicola"   
## [27] "Lady Claire" "Golden Wonder"   
## [29] "Pentland Ivory" "Marfona"   
## [31] "Pentland Javelin" "Pink Fir Apple"   
## [33] "Avalanche" "UpToDate"   
## [35] "Santee" "Romano"   
## [37] "Piccolo Star" "various"   
## [39] "Charlotte" "Lady Rosetta"   
## [41] "Paramount" "R1"   
## [43] "Up To Date" "Cara"   
## [45] "Sarpo Mira" "Hermes"   
## [47] "Cabaret" "Duke of York"   
## [49] "Tomato Money Maker" "Tomato Gardeners Delight"  
## [51] "Up-to-Date" "AFLO1\_1"   
## [53] "VR808" "Golden Nugget"   
## [55] "Aphrodite" "Kifli"   
## [57] "Ramos" "Up-To-Date"

This is the data summary for different sources.

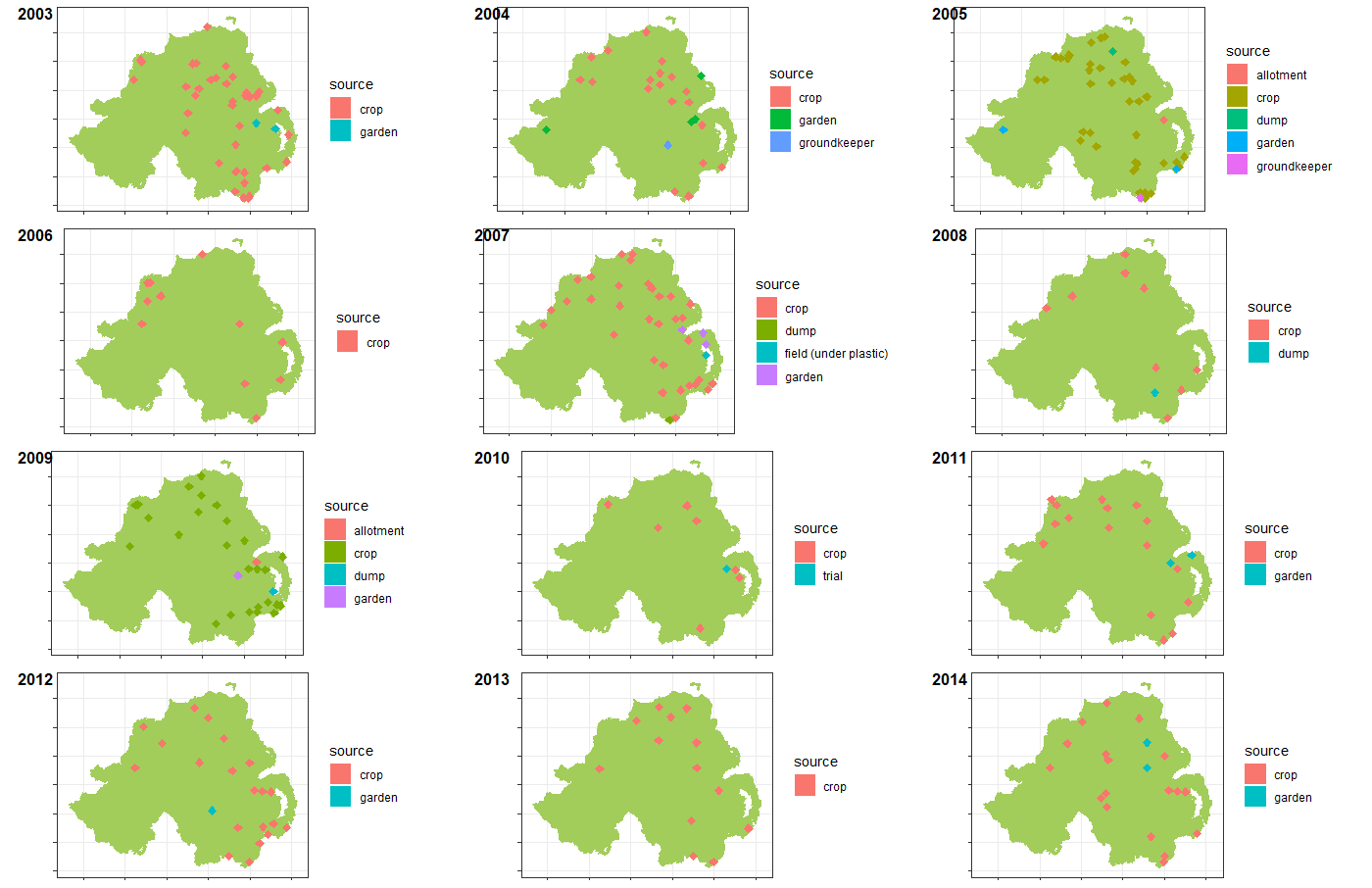
Table continues below

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| source | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
| allotment | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 |
| crop | 43 | 21 | 46 | 15 | 43 | 9 | 36 | 9 |
| dump | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| field (under plastic) | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| garden | 2 | 4 | 2 | 0 | 3 | 0 | 1 | 0 |
| groundkeeper | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| trial | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| 2011 | 2012 | 2013 | 2014 |
| 0 | 0 | 0 | 0 |
| 25 | 35 | 16 | 32 |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 2 |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |

The entire data set can be seen [here](https://github.com/mladencucak/Outbreaks/blob/master/dat/outbreaks_fin.csv)

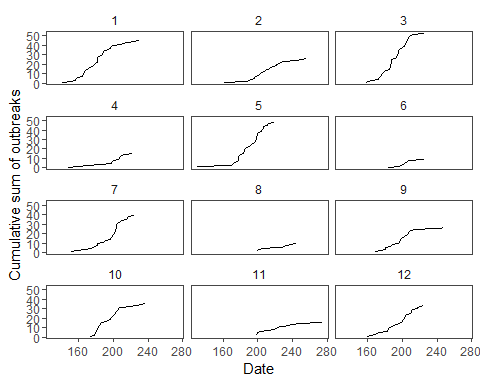
!!TASK plot outbreak maps for each year

 `

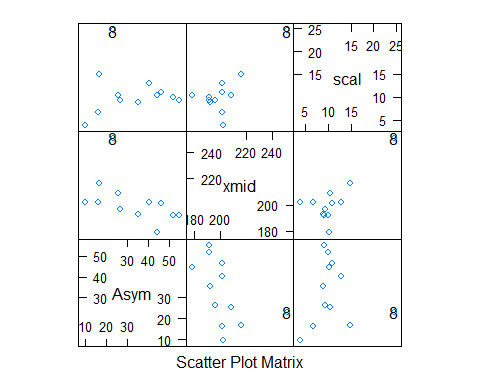
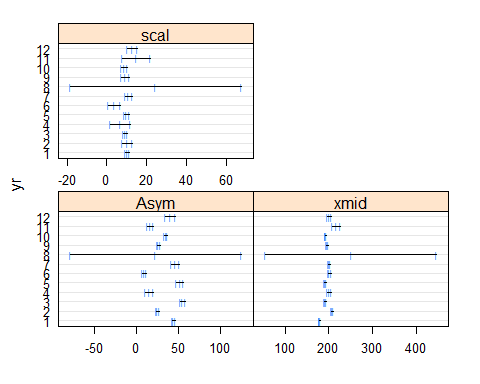
## Model fitting

Initially, a simple cumulative sum per year was calcualted with julian day and year as covariates.

|  |  |  |
| --- | --- | --- |
| yr | jday | cums |
| 1 | 141 | 1 |
| 1 | 151 | 2 |
| 1 | 156 | 3 |
| 1 | 157 | 5 |
| 1 | 158 | 6 |
| 1 | 164 | 7 |
| 1 | 168 | 13 |
| 1 | 170 | 14 |
| 1 | 171 | 16 |
| 1 | 177 | 19 |
| 1 | 178 | 20 |
| 1 | 179 | 21 |
| 1 | 181 | 22 |
| 1 | 182 | 27 |
| 1 | 184 | 28 |
| 1 | 186 | 29 |
| 1 | 188 | 32 |
| 1 | 189 | 33 |
| 1 | 191 | 34 |
| 1 | 197 | 37 |

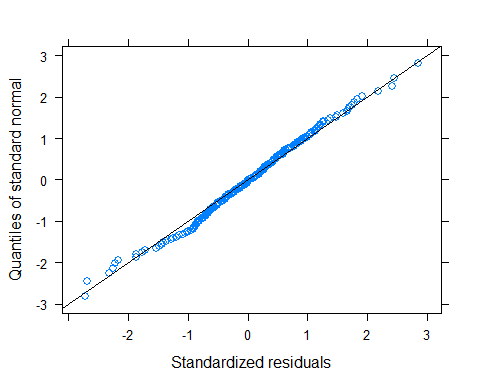
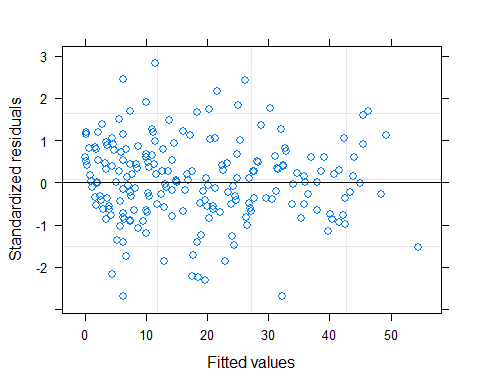


Get the starting values.



Logistic fit with julian day as predictor and year as random grouping variable.

## Nonlinear mixed-effects model fit by maximum likelihood  
## Model: cums ~ SSlogis(jday, Asym, xmid, scal)   
## Data: .   
## AIC BIC logLik  
## 912.2414 935.6378 -449.1207  
##   
## Random effects:  
## Formula: list(Asym ~ 1, xmid ~ 1)  
## Level: yr  
## Structure: General positive-definite, Log-Cholesky parametrization  
## StdDev Corr   
## Asym 15.257826 Asym   
## xmid 9.429709 -0.768  
## Residual 1.581969   
##   
## Fixed effects: list(Asym ~ 1, xmid ~ 1, scal ~ 1)   
## Value Std.Error DF t-value p-value  
## Asym 31.19484 4.458924 195 6.99605 0  
## xmid 199.64214 2.816207 195 70.89044 0  
## scal 9.96144 0.258903 195 38.47564 0  
## Correlation:   
## Asym xmid   
## xmid -0.728   
## scal 0.065 0.081  
##   
## Standardized Within-Group Residuals:  
## Min Q1 Med Q3 Max   
## -2.70462897 -0.56815659 0.02036488 0.60186916 2.83759410   
##   
## Number of Observations: 209  
## Number of Groups: 12



## Approximate 95% confidence intervals  
##   
## Random Effects:  
## Level: yr   
## lower est. upper  
## sd(Asym) 10.1930868 15.2578264 22.8391333  
## sd(xmid) 6.1632929 9.4297087 14.4272563  
## cor(Asym,xmid) -0.9243657 -0.7684395 -0.3925328  
##   
## Within-group standard error:  
## lower est. upper   
## 1.429334 1.581969 1.750904

Calculate the predictions.

Plot predicted values. 