Prediction of MPB outbreaks using Markov process logistic regression

Past and future project milestones

Mario Pineda
MPK Analytics, Inc.
Suite 300, 8507 - 112 Street
Edmonton, AB T6G 2L7
Canada

Version: 2012-10-17 20:23:58

Abstract

This vignette provides an abbreviated log of milestones that have been reached and that still need to be reached.

About this document

This document was created using Sweave, a literate programming tool for R. To generate this documentation (i.e. weaving) may require a UNIX platform to allow for system calls. Weaving may take a long time. To speed up repeated weaving intermediate data and figures may be cached. For system requirements see end of document. This document (incl. the source) is availiable at `https://github.com/mariopineda`.

MPK Analytics, Inc.
Suite 300, 8507 - 112 Street
Edmonton, AB T6G 2L7
Canada

Ph: 866.406.0187
Fax: 780.800.6929
Email: info@mpkanalytics.com

[DONE] Set up of password protected backup (sirius) for the project and data.
[DONE] Look into running the model with the Revolution Analytics Enterprise R version using the lab HPC infrastructure. Explored options together with Mark Duller but model ran slower so this apporach was abandoned in favour of optimizing the model to run on my laptop.
[DONE] Develop Empirical Bayes model in C++. (See vignette *Preparation and analysis of weather station data and temperature interpolations*)
[DONE] Computational optimization of predictive model.
[DONE] Preliminary runs of MPLR model for single years and part of BC - proof of concept. (see vignette *Autologistic model outline*)
[DONE] Preliminary runs of MPLR model for all years and part of BC- proof of concept.
[DONE] Obtained raw red top data from Fangliang He with an email confirmation its Open Data status. Cleaned up, prepared and analyzed red top data. (See vignette *Red top data*)
[DONE] Clean up, preparation and analysis of weather station data that was obtained from Tomas. (See vignette *Preparation and analysis of weather station data and temperature interpolations*)
[DONE] Interpolations of temperature variables (annual degree days and minimum winter temperature). (See vignette *Preparation and analysis of weather station data and temperature interpolations*)
[DONE] Generated summer Climate Suitability Indices (CSI) data with Barry using the Powel and Logan recipies.
[DONE] Obtained new DEM data and BC outline data from GIS Analyst. Prepared and analyzed the the data. (See vignette *Preparation and analysis of the Digital Elevation Model and BC outline data*)
[DONE] Obtained pine cover data from Barry, GIS Analyst performed preprocessing and Mario did the final clean up and analysis. (See vignette *Preparation and analysis of the BC pine cover data*)

[IN PROGRESS] Clean up and organize the sirius repository. This version of the project will be handed over to Devin and will not be developed further by Mario.
[IN PROGRESS] Development of a series of R packages containing the various components of the project. Mario will continue working on the project within the context of these R packages.
[IN PROGRESS] Final model version running all years for all of BC, computationally optimized for best performance and debugged.
[IN PROGRESS] Perform model selection. We would like to use BIC. Mark had an argument for why BIC would be prefered over AIC in this case. Request a copy of Alex's write-up on the virtues of BIC.

[TODO] Once the sirius repository has been finalized it has to be moved from Mario's account to its final resting place.

[TODO] Integration of project components across the mplr R packages (smartly).

[TODO] Parallelize the predictive model.

[TODO] Preliminary runs of MPLR model for all years and all of BC - proof of concept.

[TODO] Run model using the CSI variables (i.e. by replacing the temperature variables).

[TODO] Run model using temperature predictions (see vignette *Temperature predictions*).

[TODO] Figure out what results we are interested in, i.e. what message should the manuscript have?

[TODO] Evaluate the accuracy of the temperature predictions using the empirical Bayes (EB). The EB probably needs to be calibrated using some satistical black magic.

[TODO] Visualize spatial data on maps using a proper projection model.

[TODO] Check on CSI data integrity. BioSim flaked out over parts of BC in the initial runs, need to check that the corrupted data has been removed and replaced by the correct version.

[TODO] Fix data inconsistencies at the edges of the DEM. This will have no effect on the results but is required for the model to run for all of BC.

[TODO] It appears the `nn1` and `nn2` variables include the focal location in the summation. Figure out how to deal with this.

[TODO] For the manuscript we should run the model with red top data and temperature data up to the present year.

[ON HOLD] Follow up with GIS Analyst on CSI data conversion.

Session information

- R version 2.15.1 (2012-06-22), `i386-apple-darwin9.8.0`
- Locale: `C/en_US.UTF-8/C/C/C/C`
- Base packages: base, datasets, grDevices, graphics, methods, stats, utils
- Loaded via a namespace (and not attached): tools 2.15.1

The R code in this chapter took 0 seconds to execute, the last time the souce file was modified was on 2012-10-17 20:24:13 and the last time the PDF file was modified was on 2012-10-17 20:23:58.

MPK Analytics — Transforming Data Into Insight
Suite 300, 8507 - 112 Street, Edmonton, AB T6G 2L7, Canada
Ph: 866.406.0187, Fax: 780.800.6929
Email: info@mpkanalytics.com