Only You Can Prevent Forest Fires

Background info from Cortez, Morais 2007

Forest fires cause major economical and ecological damage and endanger humans and animals. Millions of hectares (1 *ha* = 10,000 m^2) are destroyed in the world every year. Forest preservation relies tremendously on containment and prevention of wildfires. Some causes are natural (e.g. lightning); others are preventable (e.g. human negligence). In either case, fast detection is necessary. Once a fire has been detected and identified as having potential for excessive destruction, resources can be deployed appropriately. Fire management decisions include level of readiness, prioritizing target areas, etc. Poor decisions can result in loss of life and greater damage. In addition, if the current meteorological conditions are associated with a high probability of a wildfire, the public can be warned and given safety measures to protect themselves and their property.

Traditional fire detection methods such as human surveillance are expensive (both time and money) and can be subjective. Recently, more automatic detection devices are being used. Satellites are expensive and can have time delays; their prediction accuracy also depends on their resolution. Infrared scanners have very high maintenance and equipment costs. Meteorological conditions on the other hand can be collected from weather stations in real-time (i.e. little or no delay) and are relatively inexpensive. Factors such as temperature, air humidity, rainfall, and wind speed are known to be related to fire occurrence and spread.

Most meteorological information is combined into a constantly changing index, the *Fire Weather Index* (originally developed in Canada); if the index reaches a specific level, the chance of a fire is high. The *Fire Weather Index* has six components, four of which we will use in our analysis:

- *Fine Fuel Moisture Code*: moisture content surface litter; influences ignition and fire spread
- *Duff Moisture Code*: moisture content of shallow/deep organic layers
- *Drought Code*: a similar measure to the DMC, but doesn't take into account humidity
- *Initial Spread Index*: score correlated to fire velocity spread

For all four components, higher values suggest more severe burning conditions.

Portugal, in particular, is very affected by forest fires. From 1980 to 2005, over 2.7 million *ha* of forest were destroyed. In 2003, 4.6% of Portugal was destroyed (21 human deaths); in 2005, 3.1% destroyed (18 human deaths).

Your research group has been given forest fire data from the Montesinho natural park in the northeast region of Portugal (Tras-os-Montes). Montesinho is very high in floral and fauna diversity; its preservation is crucial. Your data were collected from January 2000 to December 2003. The Fire Weather Index components as well as the time, date, and spatial location of the fires were recorded by the inspector at Montesinho; the meteorological information was collected within a 30 minute time period of the fire occurrence by a station run by the Braganca Polytechnic Institute in the center of the Montesinho park. The area burned by the fire is in hectares; any small fire that burned less than 0.1 *ha* has been capped at 0.1 *ha*.

Your research group has been asked by Portugal's national forestry and emergency management services to analyze the data and develop a model to predict the area burned by wildfires.

You have been given the following variables:

area: total burned area (in *ha*)

X: x-axis coordinate from 1(west) to 9 (east)
Y: y-axis coordinate from 1(south) to 9 (north)
month: month of the year (Jan, Feb, etc)
day: day of the week (Mon, Tue, etc)
FFMC: Fine Fuel Moisture Code
DMC: Duff Moisture Code
DC: Drought Code
ISI: Initial Spread Index
temp: outside temperature (C^o)
RH: outside relative humidity (in $\%$)
wind: outside wind speed (in km/h)
rain: outside rainfall within the previous 30 minutes (in mm/m^2)