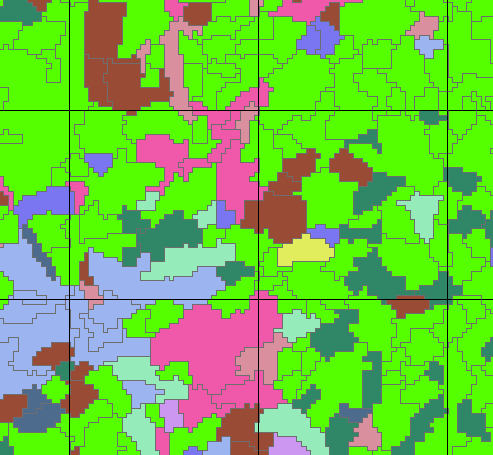
Hi coding club folks,

For next week’s workshop, I’m hoping some of you crack coders out there can give us a crash course in streamlining scripts using custom functions. Here’s the scenario:

Before Christmas, I was working on a script to assign ten kilometre squares (‘tenkms’) habitat classifications using data extracted from the CEH Land Cover Map 2000. The LCM2000 map shows discrete habitat parcels in 26 habitat types, and looks like this (grid is one kilometre squared; ‘onekm’):



I extracted the area of each habitat type per onekm (using ArcMap)\* and saved it in a file called ‘onekm\_LCM2000.csv’.

I reassigned the 26 habitat types into 10 broad classes, based on a similar assignment used in Scotland and LCM2000 guidelines. These were: Broad-leaf Woodland; Conifer Woodland; Arable; Improved Grassland; Semi-natural Grassland; Mountain/Heath/Bog; Open Water; Urban; Coast; Sea.

A key to the 26 habitat types (named Codes) and broad habitat classifications (named Class) are stored in a file called ‘LCM2000\_class.csv’

An additional classification, All Woodland, was created by adding together the area of Broad-leaf Woodland and Conifer Woodland in the script.

The purpose of the script is to identify when tenkms contain certain habitat classifications, over a given threshold (two-step threshold method explained below). The output should be a .csv where each tenkm is a row, and whether or not it contains a particular habitat classification is denoted by a ‘Y’ in columns, one for each habitat classification. E.g.:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| tenkm | IGRAS | NGRAS | MHEBO | OWATR | URB | CST | SEA |
| IC51 |  |  | Y |  |  |  |  |
| IC52 | Y |  |  |  |  |  |  |
| IC60 | Y |  | Y |  |  |  |  |
| IC61 | Y |  | Y |  |  |  |  |
| IC62 | Y |  |  |  |  |  |  |
| IC63 | Y |  |  |  |  | Y |  |

Tenkms are defined as contributing to a landscape type using two steps, developed for Scottish landscape types:

1. First, identify 1-km squares with 30% or more of the particular habitat.
2. Second, select a 10-km square as contributing to a particular landscape type if 70% or more of 1-km squares are defined as contributing to that landscape type (except for fragmented Montane, Coastal and Urban habitats, where the threshold is reduced to 30% or more).

... Therefore none of the landscape types is mutually exclusive.

Broad landscape types are being used because:

1. Bird Atlas data collection was designed to be reported at 10-km scale; and
2. Scotland tends not to have large areas of discrete habitat types and, rather, habitat mosaics (i.e. variations over the landscape) are more usual. This is even more so in NI.

As you will see in the script, the way I get my final .csv involves separate if-else statements for each habitat classification, and repeats over the two steps. That’s A LOT of if else statements! I suspect I could streamline this by making functions, so that is the challenge for the group!

You will certainly find lots of things that could be improved! Please don’t laugh! Any additional general guidance would be of interest along the way.

NB: you may notice when running the script that the end product using the thresholds above does not end in all tenkms receiving at least one habitat classification. This is a problem! I have played around with the thresholds, but it is a bit arbitrary. If anyone has ideas, please let me know!

\*Bonus points for anyone who can demonstrate how to do this using R