

Imputation Techniques to complement the Irish Field-based National Forest Inventory

Daniel McInerney¹, Maarten Nieuwenhuis.
School of Biology and Environmental Science,
University College Dublin,
Belfield, Dublin 4.
Ireland.

¹Tel: ++ 353 (0)1 716 7787, e-mail: daniel.mcinerney@ucd.ie

Abstract

Ireland's forest cover currently represents 9.6% of the total land area. It increased rapidly over the past two decades as a direct result of an active afforestation policy largely funded by the European Union and coordinated by the Irish Forest Service. As a result of these recent developments, Ireland's forests have not been monitored on a regular basis as is usual in other parts of the world. However, in late 2004, the first national field-based forest inventory was initiated by the Irish Forest Service with a view to it being repeated on a five year cycle. It is based on a systematic sample of all forests with both timber and non-timber forest variables being measured and assessed at each plot. These data are being used by the Irish Forest Service to generate provincial and national estimates, but the current methodology limits its use for the production of information at finer scales.

In order to provide more detailed, localised estimates of forest parameters, a range of imputation techniques are being tested and optimised for Irish conditions. The approach is being implemented in two study areas in Ireland, which are representative of the Irish forest estate, both in terms of species and age composition. A hierarchical classification approach is being followed, that will generate statistics and thematic maps of forest types, quantitative timber parameters (volume per hectare and basal area per hectare) and a number of non-timber parameters used as indicators of sustainable forest management. The imputation techniques being tested include: k-Nearest Neighbour, most similar neighbour and classification and regression trees. Cross-validation and feature selection algorithms are being used to identify the optimal number of neighbours and combination of explanatory variables respectively. Different weighting techniques are also being researched to identify the optimal weights for the explanatory variables.

An overview of the data processing system is also presented, which combines PostgreSQL/PostGIS, GRASS-GIS, 'R' and PHP/Mapscript.