From Forest Nursery Notes, Winter 2010

61. © Reforestation planning using Bayesian networks. Galan, C. O., Matias, J. M., Rivas, T., and Bastante, F. G. Environmental Modelling and Software 24:1285-1292. 2009.

Environmental Modelling & Software 24 (2009) 1285–1292

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Reforestation planning using Bayesian networks

C. Ordóñez Galán [a,*], J.M. Matías [b], T. Rivas [a], F.G. Bastante [a]

[a] Departamento de Ingeniería de los Recursos Naturales y Medio Ambiente, Escuela de Minas, Universidad de Vigo, 36310 Vigo, Spain
[b] Departamento de Estadística, Escuela de Minas, Universidad de Vigo, 36310 Vigo, Spain

ARTICLE INFO

Article history:
Received 3 June 2008
Received in revised form
18 May 2009
Accepted 19 May 2009
Available online 21 June 2009

Keywords:
Reforestation
Environmental variables
Bayesian networks
Machine learning
GIS

ABSTRACT

The aim of this research was to construct a reforestation model for woodland located in the basin of the river Liébana (NW Spain). This is essentially a pattern recognition problem: the class labels are types of woodland, and the variables for each point are environmental coordinates (referring to altitude, slope, rainfall, lithology, etc.). The model trained using data for existing wooded areas will serve as a guideline for the reforestation of deforested areas. Nonetheless, with a view to tackling reforestation from a more informed perspective, of interest is an interpretable model of relationships existing not just between woodland type and environmental variables but also between and among the environmental variables themselves. For this reason we used Bayesian networks, as a tool that is capable of constructing a causal model of the relationships existing between all the variables represented in the model. The prediction results obtained were compared with those for classical linear techniques, neural networks and support vector machines.

1. Introduction

Reforestation projects are generally entrusted to experts who use their own knowledge and personal experience of an area to plan and implement a reforestation project. In recent years, however, mathematical models have become increasingly popular as decision making tools. Based on current woodland distribution, these models evaluate the feasibility of different types of woodland in different zones in the studied area. The models take into account a series of physical and biological variables that have a bearing on the distribution of woodlands, such as the lithology, altitude, slope and temperature; these are handled as statistical variables in a pattern recognition problem in which each class corresponds to an existing type of woodland in the studied area.

Choosing the statistical model is a matter of some importance, particularly in terms of the starting hypotheses implied by each. For example, a linear relationship between the dependent variable and influencing variables is often assumed, as in the case of methods based on linear regression models, descriptions of which are relatively abundant in the literature (Felicísimo et al., 2002; Narumalani et al., 1997; Van de Rijt et al., 1996). However, given that relationships in nature are not always linear, alternative methods may be more suitable. Non-linear approaches to vegetation modelling, nonetheless, are much less frequently encountered in the literature (Muñoz and Felicísimo, 2004; Ordóñez et al., 2005).

With a view to tackling reforestation from a more informed perspective, of interest is an interpretable model of the relationships existing between the variables—whether between woodland type and environmental variables or between and among environmental variables. This information is generally not provided by the typical pattern recognition techniques.

Bayesian networks (BNs) have begun to be used in recent years in order to tackle problems in different fields, for example, in medicine (Andreassen et al., 1991; Tucker et al., 2005), the environment (Marcot et al., 2006; Martin de Santa Olalla et al., 2007; Morteza et al., 2009), remote sensing (Ouyang et al., 2006), industry (Vehtari and Lampinen, 1999) and geographical information systems (Stassopulou et al., 1998), given their capacity for combining information from different sources, relating variables, interpretating and inferring. Recently, Farmani et al. (2009) used BNs combined with an evolutionary multiobjective optimization (EMO) technique for optimum management of a groundwater contamination problem under uncertainty. The optimization algorithm was used to generate the state variable values, which are fed into the Bayesian belief network.

The usefulness of the BNs compared to mechanistic and empirical models has been analysed by Castelletti and Soncini-Sessa (2007) for a water resource management problem. These authors came to the conclusion that BNs are useful when knowledge about the system to be modelled is poor structured and mainly empirical in nature and can easily be integrated with other

* Corresponding author at: E.T.S.I MINAS, Universidad de Vigo, Campus Universitario, 36310 Vigo, Spain. Tel.: +34 98 681 4052; fax: +34 98 681 2201.
E-mail address: cgalan@uvigo.es (C. Ordóñez Galán).

1364-8152/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2009.05.009