Catena 154 (2017) 73–84

Contents lists available at ScienceDirect

Catena

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

Spatial soil nutrients prediction using three supervised learning methods for assessment of land potentials in complex terrain

Gwanyong Jeong [a,*,1], Hannes Oeverdieck [a], Soo Jin Park [b], Bernd Huwe [a], Mareike Ließ [a,2]

[a] Department of Geosciences/Soil Physics Division, University of Bayreuth, Universitaetsstrasse 30, 95447 Bayreuth, Germany
[b] Department of Geography, Seoul National University, Shilim-Dong, San 56-1, Kwanak-Gu, Seoul 151-742, South Korea

ARTICLE INFO

Article history:
Received 12 October 2015
Received in revised form 11 December 2016
Accepted 7 February 2017
Available online 6 March 2017

Keywords:
Mountain soil
Soil nutrients
Digital soil mapping
Land potential assessment

ABSTRACT

Mountain soils play an essential role in ecosystem management. Assessment of land potentials can provide detailed spatial information particularly concerning nutrient availability. Spatial distributions of topsoil carbon, nitrogen and available phosphorus in mountain regions were identified using supervised learning methods, and a functional landscape analysis was performed in order to determine the spatial soil fertility pattern for the Soyang Lake watershed in South Korea. Specific research aims were (1) to identify important predictors; (2) to develop digital soil maps; (3) to assess land potentials using digital soil maps.

Soil profiles and samples were collected by conditioned Latin Hypercube Sampling considering operational field constraints such as accessibility and no-go areas contaminated by landmines as well as budget limitations. Terrain parameters and different vegetation indices were derived for the covariates. We compared a generalized additive model (GAM) to random forest (RF) and support vector regression (SVR). For the predictor selection, we used the recursive feature elimination (RFE). A land potential assessment for soil nutrients was conducted using trimmed k-mean cluster analysis.

Results suggested that vegetation indices have powerful abilities to predict soil nutrients. Using selected predictors via RFE improved prediction results. RF showed the best performance. Cluster analysis identified four land potential classes: fertile, medium and low fertile with an additional class dominated by high phosphorus and low carbon and nitrogen contents due to human impact. This study provides an effective approach to map land potentials for mountain ecosystem management.

1. Introduction

Recent research has shown an increased interest in mountain soils and their essential role in ecological functions (Ballabio, 2009; Roman et al., 2010; Wilcke et al., 2013). Mountain soils are regarded as a key factor for conservation and sustainable management in mountain regions. They have effects on downstream water quality and quantity, and nutrient dynamics as an input source (Wohl, 2010). Mountain soils and particularly their organic layers contain large stocks of soil nutrients released by mineralization of the organic matter (Wilcke et al., 2010). Especially, nitrogen and phosphorus deficiencies exert influences on limitation of plant growth in mountain areas. It is an important issue to understand nutrient supply in mountain areas which influences plant productivity, diversity and compositions (Benner et al., 2010). Last but not least, spatial information on soil nutrients is required to understand and manage mountain ecosystems.

Properties of mountain soils have a high local variability (Holtmeier, 2009). This is because mountain areas underlie high variability in climatic conditions, topography, parent material and vegetation (Funnell and Parish, 2005). Several processes such as creeping, erosion, solifluction, and landslide lead to nutrient losses in some parts and accumulations in others. Organic matter accumulation, nutrient cycling and disturbance are enhanced by vegetation and animals (Brady and Weil, 2010). Furthermore, parent rock material influences the spatial distribution of base status and nutrient contents in the soil (Binkley and Fisher, 2012). Complex interactions among these various environmental factors generate a considerable spatial heterogeneity of physical and chemical soil properties. Mountain areas also provide various challenges to any soil mapping approach as they are difficult to access and have scarce data availability. Construction of soil dataset in mountain areas frequently requires considerable amounts of cost, time, and labor due to these reasons.

Digital soil mapping (DSM) can be a useful tool to reduce this kind of efforts as well as obtain reasonable results (Ballabio, 2009; Ließ et al., 2012). DSM is the procedure of creating spatial soil information systems

* Corresponding author.
E-mail address: gwanyong.jeong@uni-bayreuth.de (G. Jeong).
[1] Current address: Department of Geography, Chonnam National University, 77, Yongbong-ro, Buk-gu, Gwangju, 61186, South Korea.
[2] Current address: Department of Soil Physics, Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Str.4 D-06120 Halle (Saale), Germany.

http://dx.doi.org/10.1016/j.catena.2017.02.006

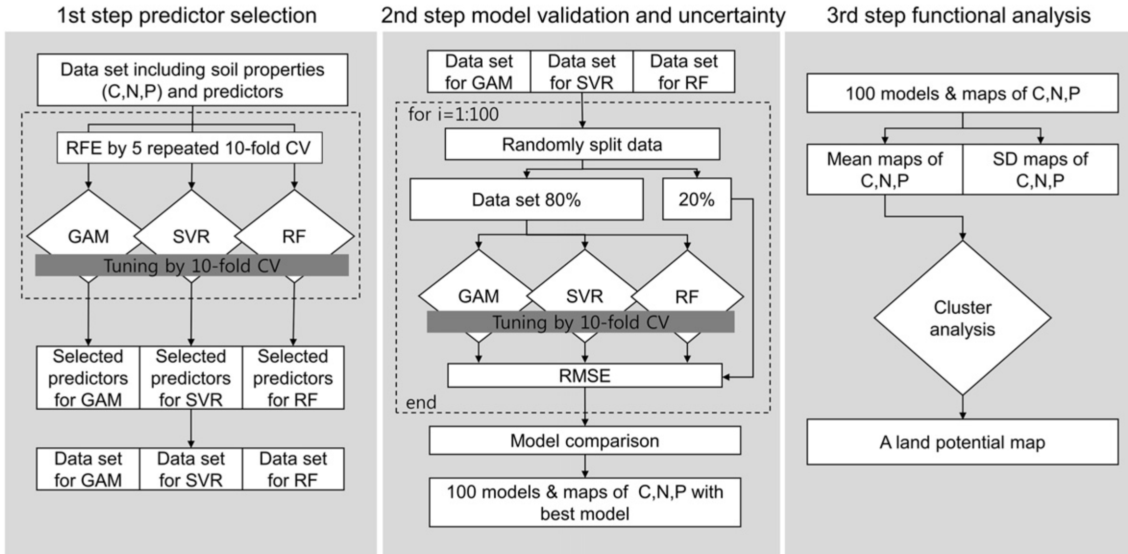

Fig. 1. Flowchart of the proposed procedure. C = carbon, N = nitrogen, P = available phosphorus, RFE = recursive feature elimination, GAM = generalized additive model, SVR = support vector regression, RF = random forest, CV = cross validation, RMSE = root mean squared error, SD = standard deviation.

based on numerical relationships between soil observation and related environmental predictors to understand spatial patterns of soil types and soil properties (Lagacherie et al., 2007). DSM encompasses many attempts to investigate the quantitative relationships based on the Jenny state factor equation (McBratney et al., 2003). DSM has been successfully used for predicting various soil types and properties with environmental predictors derived from geographical information system and remote sensing techniques and machine learning methods such as random forest (Breiman, 2001), artificial neural networks (Bishop, 1995), support vector regression (Smola and Schölkopf, 2004) at the local, regional, national, continental, and global scale (Dobos and Hengl, 2009; Grunwald, 2009; McBratney et al., 2003; Minasny and McBratney, 2015).

Soil fertility refers to the soil's ability to provide nutrients in available forms and appropriate amounts for plant growth and reproduction (Osman, 2013). Soil fertility plays a key role in land potential or capacity (Osman, 2013; Stockdale et al., 2013). In mountain areas, comprehensive information on the soil fertility is needed because the nutrient availability determines plant and hence forest growth and it is important for sustainable mountain ecosystem management (Jenny, 1980). Some research on land potentials performed spatial predictions based on digital soil mapping but most studies focused on agricultural areas (Al-Shamiri and Ziadat, 2012; Harms et al., 2015; Sun et al., 2012). Moreover, there were a few studies on spatial predictions for soil fertility in mountain areas. In this study, we tried to develop soil nutrient maps in order to investigate land potentials in the Soyang watershed

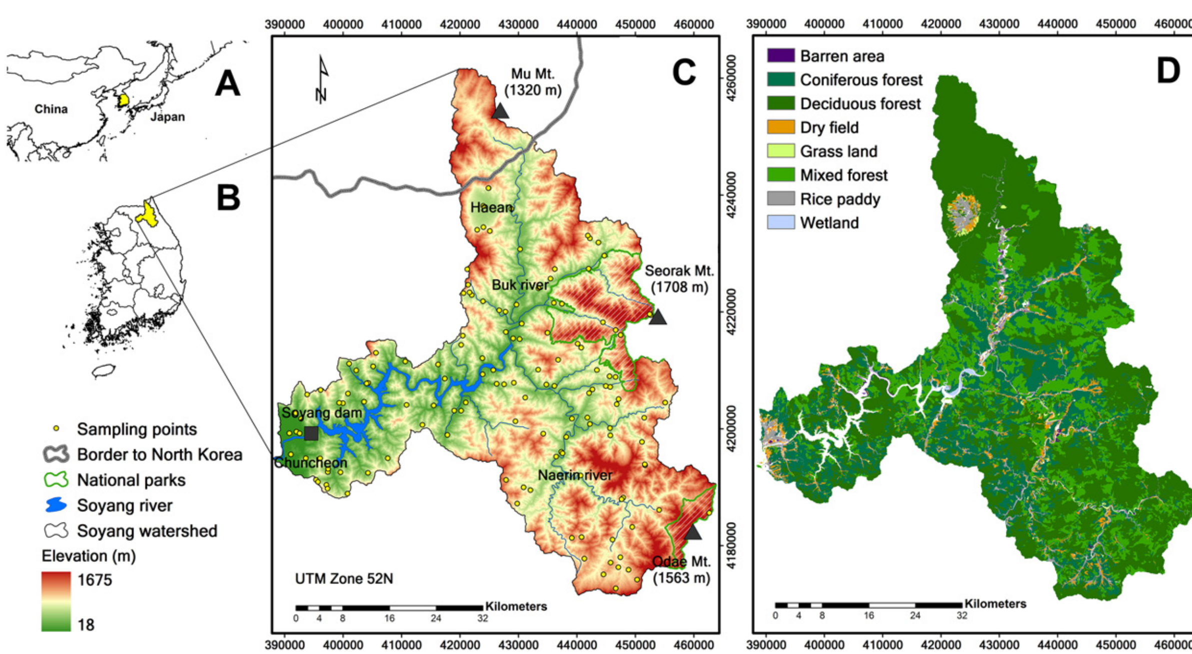

Fig. 2. Research area. (A) The map of the Korean Peninsula. (B) The Soyang watershed is located in the north-eastern part of South Korea. (C) The map shows the spatial distributions of sampling points. The Soyang watershed is near the border to North Korea and includes two national parks (Seorak and Odae). The Soyang river originates from Mu, Seorak and Obdae Mountain. (D) The landuse map of Soyang watershed from Korean Environmental Geographic Information Service (EGIS) (http://egis.me.go.kr).