Article

Model Prediction of the Soil Moisture Regime and Soil Nutrient Regime Based on DEM-Derived Topo-Hydrologic Variables for Mapping Ecosites

Zhengyong Zhao ^{1,2} , Qi Yang ¹, Xiaogang Ding ^{3,*} and Zisheng Xing ⁴

¹ Guangxi Key Laboratory of Forest Ecology and Conservation, College of Forestry, Guangxi University, Nanning 530004, China; u1v54@unb.ca (Z.Z.); qi66yang@yahoo.ca (Q.Y.)

² Faculty of Forestry, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

³ Guangdong Academy of Forestry, Guangzhou 510520, China

⁴ Brandon Research and Development Centre, Portage, MB R1N 3V6, Canada; zisheng.xing@canada.ca

* Correspondence: 27267152@sinogaf.cn

Abstract: Ecosites are required for stand-level forest management and can be determined within a two-dimensional edatopic grid with soil nutrient regimes (SNRs) and soil moisture regimes (SMRs) as coordinates. A new modeling method is introduced in this study to map high-resolution SNR and SMR and then to design ecosites in Nova Scotia, Canada. Using coarse-resolution soil maps and nine topo-hydrologic variables derived from high-resolution digital elevation model (DEM) data as model inputs, 511 artificial neural network (ANN) models were developed by a 10-fold cross-validation with 1507 field samples to estimate 10 m resolution SNR and SMR maps. The results showed that the optimal models for mapping SNR and SMR engaged eight and seven topo-hydrologic variables, together with three coarse-resolution soil maps, as model inputs, respectively; 82% of model-estimated SNRs were identical to field assessments, while this value was 61% for SMRs, and the produced ecosite maps had 67–68% correctness. According to the error matrix, the predicted SNR and SMR maps greatly alleviated poor prediction in the areas of extreme nutrient or moisture conditions (e.g., very poor or very rich, wet, or very dry). Thus, the new method for modeling high-resolution SNR and SMR could be used to produce ecosite maps in sites where accessibility is hard.

Keywords: soil moisture regime; ecosite; edatopic grid; artificial neural network

Citation: Zhao, Z.; Yang, Q.; Ding, X.; Xing, Z. Model Prediction of the Soil Moisture Regime and Soil Nutrient Regime Based on DEM-Derived Topo-Hydrologic Variables for Mapping Ecosites. *Land* **2021**, *10*, 449. <https://doi.org/10.3390/land10050449>

Academic Editors: Assem Abu Hatab, Nasem Baderldin and Zhen Liu

Received: 1 March 2021

Accepted: 21 April 2021

Published: 23 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ecosites, as the stand-level forest unit within the ecological land classification system, define a suite of site conditions that represent the general productivity of forest lands and describe an ecological setting for grouping forest and soil classes [1]. Ecosite maps are required for forest stand-level management and planning, such as forest best management practices and timber supply analysis [2]. The principal means of identifying ecosites in the field is based on a number of easily observable topographic, soil, and vegetation indicators [3]. However, obtaining ecosite maps with adequate resolution and accuracy through interpolating field-identified ecosite points requires a large number of field surveys due to the high spatial variation of ecosites across landscapes. Ecosites can be determined within a two-dimensional edatopic grid using soil nutrient regimes (SNRs) and soil moisture regimes (SMRs) as coordinates [4]. This makes it possible to produce high-resolution ecosite maps without the requirement of a large number of field soil surveys.

SNRs and SMRs are indices that reflect the combined interaction of climatic, topographic, geological, and biophysical factors on forest yield and growth [5]. The SNRs and SMRs are more tangible and operationally useful for mapping ecosites than plant-based methods [6]. In addition to field assessments using easily observable site features, indicator vegetation, and easily identified soil properties [7], many studies focused on the model

predictions of SNRs and SMRs. The models are often based on interpolation schemes [8] and statistics-based schemes [9,10] using varying model predictors from field-based plant indicators [7], model-based clay content [9], model-based soil drainage [11], remote sensing data [12], and map-based soil texture [10], with varying class number, map resolution, and model accuracies. However, there is a lack of model studies that estimated SNRs and SMRs with high resolution (i.e., ≤ 10 m) and high accuracy using easily accessible model predictors.

It has been known for a long time that soil properties at a landscape or regional level can be greatly affected by bedrock and geology because soil particles and nutrient elements are mainly sourced from bedrock and parent material deposits [13]. Hydrological processes associated with local topography can heavily modify soil properties at a local level. This is because the weathering products of rock and parent material, such as nutrient elements, sediment, and soil particles, are constantly transferred with water movements along topographical gradients, which in turn affect soil properties [14]. Some researchers [15,16] reported that existing coarse-resolution soil maps can present the general information of bedrock, parent materials, and geological formations that are related to average soil properties at a large scale, and topo-hydrological parameters derived from high-resolution digital elevation model (DEM) data can capture the characteristics of hydrological processes that tune soil properties at a local level. It is perhaps possible to model SNRs and SMRs with existing coarse-resolution soil maps and high-resolution DEM data. However, the relationships between SNRs and SMRs and the predictors are complicated in nature and sometimes unknown. The existing empirical models and interpolation methods for mapping SNRs and SMRs may hinder the real relationships due to the use of hypotheses that predetermined the relationships [17]. In recent years, artificial intelligence models, especially artificial neural networks (ANNs), have proven effective and excellent in mapping the complicated relationships between multisource inputs without hypotheses. ANN models have been increasingly used to map soil properties, such as soil texture [18], soil drainage [19], and soil electrical conductivity [20]. Thus, the combined scheme, i.e., building an ANN model with predictors derived from the easily accessible DEM data, ought to be able to estimate high-resolution SNRs and SMRs.

The main objective of this research was to introduce new models to estimate SNRs and SMRs for mapping forest ecosites with high resolution. Specific objectives were to (1) model SNRs and SMRs using coarse-resolution soil maps and high-resolution DEM-derived variables with ANN models and (2) map ecosites using the produced SNR and SMR maps and test their accuracies.

2. Materials and Methods

2.1. Study Area

The province of Nova Scotia, Canada, is about 5.5×10^6 ha and 78% covered by forests. It is dominated by a continental climate with minor influence from the Atlantic Ocean. Various soil types and forest types within the province were formed with different combinations of topography, parent material, and climate. Hundreds of short rivers, streams, and lakes dotted over the countryside enhance the complexity of landscapes. The most common mixed-forest types in the region are made up of the major conifers, such as pines (*Pinus* spp.), balsam fir (*Abies balsamea*), and spruce (*Picea* spp.) and common hardwood species, such as aspen (*Populus* spp.), birch (*Betula* spp.), and maple (*Acer* spp.). The moisture and nutrient conditions of common soils, such as Regosols, Podzols, Gleysols, Brunisols, and Luvisols, in the area vary with topographic positions and soil texture [1].

2.2. Sample Plots

Field samples for this study came from 1507 fixed-area forest ecosystem classification (FEC) plots that have been investigated by the Nova Scotia Department of Natural Resources since 2000 [21]. The plots were established in order to cover the full range of SNRs and SMRs expected within the province and to produce an accurate FEC system. The com-

binations of field-identified soil properties, including soil depth, soil texture, forest floor humus form, and A horizon type, are used to assign SNR and SMR class on the basis of the developed FEC guides. SMRs included seven classes: wet, moist/wet, moist, fresh/moist, fresh, dry, and very dry. SNRs included five classes: very rich, rich, medium, poor, and very poor. Assigning ecosite type to each plot was based on field-identified vegetation type and field-assessed SNR and SMR classes within the rectangle-based edatopic grid. As a function of the tree growth data of the full FEC plots, the whole forest area of Nova Scotia was divided into the Maritime Boreal region and Acadian region, with 10 ecosites [11].

2.3. Predictors for Modeling

The candidates of model predictors in this study were composed of coarse-resolution soil data and high-resolution DEM-derived variables. Coarse geology, landscape, and topography variables were obtained from the biophysical land classification map that was compiled by the Nova Scotia Department of Natural Resources [22]. The map was delineated from the aerial photographs with a 1:50,000 scale. The unit of the map was defined as an area of forest land having a reoccurring pattern of landforms, soil, and vegetation. The size of the polygon varied with topography, but usually averaged 400–1600 ha. Coarse geology presented major rock types associated with the system soil's origin, including sedimentary (six types), metamorphic (five types), and igneous (six types). The landscape delineated conspicuous topographic features, including glacial (six types), organic (four types), marine and glacial marine (four types), alluvial (five types), and glacial fluvial (six types) features. Topography reported the topographic pattern, namely, smooth, hummocky, knob and knoll, drumlins, dissections, and ridged. The DEM data were sourced from Service, Nova Scotia, Canada. Before the elevation points were interpolated into a grid map with 10 m resolution, original points were compiled from SPOT 5 images. Using developed forest hydrology tools and the spatial analyst extension tools of ArcGIS [23], nine topo-hydrological variables were derived from the DEM data [18], namely, flow direction (FD), potential solar radiation (PSR), flow length (FL), aspect, depth to water (DTW), the sediment delivery ratio (SDR), slope, the topographic position index (TPI), and the soil terrain factor (STF).

2.4. Artificial Neural Network Models

The ANNs were developed to estimate SNRs and SMRs because they can realize any nonlinear mapping between model input and output data without hypotheses [24]. A three-layer structure was used for the ANNs in this research (Figure 1): an input layer containing predictor variables, an output layer consisting of estimated SNRs and SMRs, and a hidden layer, whose nodes linked the input layer and output layer and reflected the complexity of the ANNs. The ANNs were trained using the backpropagation technology with the Levenberg–Marquardt algorithm [25].

The candidate predictors composed of model inputs. Each candidate combination of model inputs corresponded to an ANN model. The coarse-resolution soil data were presented in each combination, whereas DEM-derived topo-hydrologic variables were gradually added from single to multiple in various combinations. For example, there were nine combinations when taking one from nine DEM-derived variables at a time ($C_9^1 = 9$) and 84 combinations when taking six from nine variables at a time ($C_9^6 = 48$). Thus, there were a total 511 combinations ($C_9^1 + C_9^2 + C_9^3 + C_9^4 + C_9^5 + C_9^6 + C_9^7 + C_9^8 + C_9^9 = 511$), which corresponded to 511 ANNs.

Figure 1. Schematic diagram showing structure and flow of the artificial neural network (ANN) model for predicting soil moisture regime (SMR) or soil nutrient regime (SNR). DEM: digital elevation model, TPI: topographic position index, PSR: potential solar radiation, STF: soil terrain factor, FEC: forest ecosystem classification.

2.5. Model Calibration, Validation, and Assessment

A method of 10-fold cross-validation [16] was used to build the ANNs and evaluate model performance with a total of 1507 field-assessed SNRs and SMRs. In the 10-fold cross-validation mode, after the entire dataset (1507 data) was randomly divided into 10 equal subsets, nine subsets as calibration data were used to calibrate an ANN model, and the remaining subset was used to validate the model; this process was repeated 10 times until all subsets were used as validation data. Following the requirement of developing ANNs with an early stopping technology, nine subsets, used as the calibration data, were randomly subdivided into a testing dataset (15% of the calibration dataset) and a training dataset (85% of the calibration dataset) within each model building process (each “fold”), which was repeated 100 times to obtain the optimal prediction model for this “fold”.

Three accuracy parameters from 10-fold cross-validation were used to screen the optimal ANN models, namely, coefficient of determination (R^2), root-mean-squared error (RMSE), and overall accuracy (OA). The processes of building ANNs, assessing model accuracies, and screening ANNs were programmed using MATLAB software.

2.6. Mapping Forest Ecosites

Before mapping forest ecosites, the entire province was divided into the Acadian and Maritime Boreal regions according to Table 1 in the study by Keys [1]. For each map unit (10 m × 10 m grid; equal to the size of an FEC plot) within each region, information on the model-estimated SNRs and SMRs was then used to project the map unit onto the rectangle-based edatopic grid [11] and assign an ecosite type that best matched its position on the edatopic grid. The produced ecosite maps were assessed with field-identified ecosites from FEC plots.

Table 1. Model accuracies evaluated using 1507 plots with 10-fold cross-validation and selection of DEM-derived variables, together with three coarse-resolution soil maps, for the optimal ANN models to estimate SNR and SMR.

	No. Variable	RMSE	R^2 *	OA (%)	The Best Combined DEM-Derived Variables
SMR	1	1.34	0.53	41	TPI
	2	1.19	0.63	45	TPI, SDR
	3	1.07	0.68	45	TPI, SDR, slope
	4	1.09	0.67	48	TPI, SDR, slope, FD
	5	1.00	0.73	52	TPI, SDR, slope, FD, STF
	6	0.98	0.75	56	TPI, SDR, slope, FL, aspect, PSR
	7	0.95	0.78	61	TPI, SDR, slope, FL, aspect, PSR, VSP
	8	0.97	0.75	56	TPI, SDR, slope, FL, aspect, PSR, VSP, FD
	9	0.98	0.75	55	TPI, SDR, slope, FL, aspect, PSR, VSP, FD, STF
SNR	1	0.73	0.52	55	TPI
	2	0.65	0.64	66	TPI, aspect
	3	0.58	0.74	72	TPI, FL, slope
	4	0.57	0.76	74	TPI, aspect, slope, FL
	5	0.55	0.78	76	TPI, aspect, slope, FL, SDR
	6	0.55	0.78	78	TPI, aspect, slope, FL, SDR, FD
	7	0.54	0.78	78	TPI, aspect, slope, FL, SDR, FD, VSP,
	8	0.51	0.81	82	TPI, aspect, slope, FL, SDR, FD, VSP, PSR
	9	0.53	0.80	79	TPI, aspect, slope, FL, SDR, FD, VSP, PSR, STF

* All of coefficient of determination (R^2) were significant at $p < 0.01$ based on an F-test. SMR: soil moisture regime, SNR: soil nutrient regime, DEM: digital elevation model, RMSE: root mean squared error, OA: overall accuracy. FD: flow direction, PSR: potential solar radiation, FL: flow length, DTW: depth to water, SDR: sediment delivery ratio, TPI: topographic position index, STF: soil terrain factor.

3. Results

3.1. SMR Model

Prediction accuracies of the 511 ANN models for SMR, assessed with 10-fold cross-validation, are shown in Table 1. The table shows that with the increase in DEM-derived variables, the performance of models was significantly improved in general. However, the improvement in model performance was limited when the number of used DEM-derived variables exceeded a certain limitation. In the table, the performance of ANN models significantly improved when DEM-derived variables were gradually added from one to seven, and the values of OA were increased from 41% to 45%, 45%, 48%, 52%, 56%, and 61%, respectively, but there was a minor decrease in OA value when DEM-derived variables were further added (eighth and ninth). The improvements in RMSE and R^2 showed the same trend. This result was not a surprise because the order of adding DEM-derived variables was the same as the order of importance in relation to SMR accumulation. Thus, the best model for estimated SMR classes in Table 1 was the ANN model that engaged seven DEM-derived variables, together with three coarse-resolution soil maps, as model inputs.

The results that involved the best combinations of DEM-derived variables for the SMR in Table 1 also illustrated that TPI was the most important DEM-derived variable for estimating SMRs, which, together with three coarse-resolution soil maps, could directly account for 53% of the total variation, following by SDR and slope, which accounted for 10% and 5%, respectively. This result was not a surprise since the spatial distribution of SMR was heavily affected by local topography that decided surface water movement [21]. This result agrees with a similar study [26].

An error matrix of model accuracy was calculated using field-assessed SMRs from FEC plots and model-estimated SMRs from the optimal ANN engaging seven DEM-derived variables in Table 1 (Table 2). Accordingly, 61% of model-estimated SMRs were identical with field assessments of 90% within ± 1 class.

Table 2. Error matrices of the soil nutrient and soil moisture regime.

		SMR Model-Estimation							Producer's Accuracy * (%)		
		VD	D	F	FM	M	MW	W	Total	OA	OA ± 1 Class
Acadian-region field assessment	VD	4	3	4	0	0	0	0	11	36	64
	D	0	40	38	11	0	1	0	90	45	87
	F	2	21	523	83	24	9	3	665	79	94
	FM	1	3	80	163	26	18	0	291	56	92
	M	0	2	13	79	67	27	14	202	33	86
	MW	0	0	4	23	35	57	10	129	44	79
	W	0	1	3	6	16	34	59	119	50	78
	Total	7	70	665	365	168	146	86	1507	61	90

		SNR Model-Estimation					Producer's Accuracy (%)		
		VP	P	M	R	VR	Total	OA	OA ± 1 Class
Mar-Bor-region field assessment	VP	31	36	3	1	0	71	44	95
	P	9	321	65	3	0	398	81	99
	M	5	58	739	12	2	816	91	99
	R	0	7	62	136	1	206	66	67
	VR	0	1	1	7	7	16	44	88
	Total	45	423	870	159	10	1507	82	99

* OA was equaled to the percentage ratio between the sum of correctly estimated plots (the dark grey cells) and the sum of plots in the same line; OA ± 1 class was equaled to the percentage ratio between the sum of estimated plots within ± 1 class (the dark grey and light grey cells) and the sum of plots in the same line. SMR: soil moisture regime, SNR: soil nutrient regime, W: wet, MW: moist/wet, M: moist, FM: fresh/moist, F: fresh, D: dry, VD: very dry, VR: very rich, R: rich, M: medium, P: poor, VP: very poor.

3.2. SNR Model

The performance of the ANN models in estimating SNRs is also reported in Table 1, showing similar results to the SMR models. Within a certain limitation, model accuracies of SNR also obviously improved with an increase in the number of used DEM-generated variables. The optimal model for estimating SNR was the ANN model that used eight DEM-generated variables, together with three coarse-resolution soil maps, as model inputs, with an OA of 90%, R^2 of 0.81, and RMSE of 0.51.

The results of screening DEM-generated variables for SNR in Table 1 showed that, as when modeling SMR, the TPI was also the most important DEM-generated variable for estimating SNR, which, together with three coarse-resolution soil maps, could account for 51% of total variation, followed by aspect, which accounted for 12%. These results were generally expected because soil nutrients are usually reassigned along the hillside at a local level.

Using the optimal ANN in Table 1 with eight DEM-derived variables to produce the SNR map, the error matrix showed that 90% of model-estimated SNRs were identical with field assessments of 99% within ± 1 class (Table 2).

3.3. Mapped Forest Ecosite

After each map unit (10 m $\times$ 10 m grid) was projected onto the rectangle-based edatopic gird using its model-estimated SNR and SMR classes, the accuracies of produced ecosite maps for Maritime Boreal region and Acadian region are presented in Figure 2.

Figure 2. Comparison of prediction accuracy of Acadian region and Maritime Boreal region ecosites using the model-estimated SNRs and SMRs with that of field assessment-based.

For the 10 ecosites from the Acadian region, there was a total of 67% correctness with a range from 23% to 76%. The ecosites that offset from the middle of the edatopic grid, such as e.g., ecosite 1 and ecosite 10, were less well represented than those ecosites that appeared in the center, such as ecosite 5 and ecosite 6 (Figure 3). For 10 ecosites from Maritime Boreal region, there was a total of 68% correctness with a range from 25 to 94%. Most of ecosites that were placed along the edges of the edatopic grid, such as ecosites 9 and ecosite 10, were less well represented than those ecosites in the center, such as 5 and ecosite 6 (Figure 3).

Figure 3. Comparison of average model-estimated SNR and SMR values of field-assessed ecosites (model-estimated centers) and the responded Yang-reported centers [11], and the transfers from Yang-reported centers to model-estimated centers within the edatopic grid for 10 Acadian-region ecosites (a) and 10 Maritime-Boreal-region ecosite (b). W: wet, MW: moist/wet, M: moist, FM: fresh/moist, F: fresh, D: dry, VD: very dry, VR: very rich, R: rich, M: medium, P: poor, VP: very poor.

4. Discussion

4.1. The Performance of SMR and SNR Models and Ecosite Maps

The accuracy of the produced SMR map, i.e., 61% correctness (with 90% within ± 1 class), was higher than the reported 48% correctness (with 83% within ± 1 class), where SMRs, using the same database as in this study, were estimated from modeling soil

drainage classes [11], as well as higher than the reported 55% correctness (with 94% within ± 1 class), where SMRs with six classes were estimated from plant indicators [7], but a little lower than the reported 65% correctness, where SMRs with four classes were generated from a rule-based GIS model [10]. Compared with Yang's matrix [11], although many very dry plots in the estimated SMR map were still classified as dry or fresh plots and several wet plots were classified as moist or moist/wet plots in this study, there was an increase in OA value from 0% to 37% for very dry plots and from 23% to 50% for wet plots, respectively. Compared with the biggest problem in Yang's report [11], whereby estimated SMRs were clumped into the middle of SMR classes (Figure 3), which resulted in too many estimated mesic plots on the SMR map, the result in this study showed that the situation was greatly alleviated. This result was also proven by the fact that total kappa index was increased from 0.24 to 0.45. After mining all potentials of DEM-derived data by modeling in this study, introducing plant information, as prompted by Wang's research [7], should be a good choice to improve model predictions in the next step.

The accuracy of the produced SNR map, 82% correctness (with 99% within ± 1 class), was evidently higher than the reported 58% correctness, where SNRs, using the same database as in this study, were estimated from modeling clay content [9], as well as higher than that obtained by Wang [7], who reported 59% correctness when using plant-based indicators. As with SMRs, an increase in OA value from 14% to 44% for very poor plots and from 19% to 44% for very rich plots, as well as an increase in total kappa index from 0.47 to 0.76, also supported that the produced SNR map in this study greatly alleviated the trend of clumping into the middle of SNR classes that was reported in Yang's research [9] (Figure 3). Considering that most ecosite classes exist in more than one SNR class [11]—for instance, the SNR for the Maritime Boreal region ecosite 1 ranges from poor to very poor—the obtained 82% correctness of model-estimated SNRs with 99% within ± 1 class should be good enough to map ecosites in this study.

The accuracy of the produced ecosite map (67–68% correctness) was higher than the reported 59–61% correctness, using the same database as in this study, reported by Yang [11], and close to the 66–70% correctness recorded by MacMillan [27] for estimated 25 m resolution ecosite maps in British Columbia, Canada, using a mixed method of manual and automated procedures that could realize ecology–landform relationships with rule-based models.

4.2. Effects of Coarse-Resolution Soil Maps and High-Resolution DEM-Derived Maps

Coarse-resolution geology, landscape, and topography maps were used in each ANN model input in order to represent their general conditions in a region. The boundaries of the coarse-resolution soil maps were still visible in the produced SNR, SMR, and ecosite maps (Figure 4g–j), which reflected the effects of the maps on SNR, SMR, and ecosite distribution. The estimated SNRs (Figure 4g) within an area of shale-based soil (Figure 4b), for instance, were obviously richer than those of granite- and quartzite-based soil (Figure 4b). The estimated SMRs (Figure 4h) within an area of granite-based soil (Figure 4b) were generally drier than those of other parent material-based soil. Thus, the area of granite-based soil (Figure 4b), due to its poor nutrient level and dry moisture, signified that there was more SNRs in the Acadian region ecosite 1 (Figure 4i,j), where jack pine/black spruce is the main forest type [1]. All of these results proved that coarse-resolution maps affected the distributions of SNR, SMR, and ecosite on a large scale.

The high-resolution DEM-derived predictors were used in each ANN model input in order to express how soil nutrient and moisture were redistributed by hydrological processes in local topography. On one hand, the TPI (Figure 4d), aspect (Figure 4e), and SDR (Figure 4f) derived from DEM data (Figure 4c) showed the relative topographic position, direction of the steepest slope, and efficiency of transporting sediment, respectively, in detail, proving that DEM-derived topo-hydrological predictors can capture the characteristics of hydrological processes that tuned soil properties at a local level. On the other hand, when compared to existing coarse-resolution soil maps (Figure 4b), the estimated SNR, SMR, and

ecosite maps (Figure 4g–j) contained more detailed information, and reflected the influence of the high-resolution topo-hydrological maps. For example, the estimated SNR map (Figure 4g) and SMR map (Figure 4h) were somewhat similar to the TPI map (Figure 4d). This means that the TPI played an important role in the distribution of SNR and SMR, which was in keeping with the analysis of model input screen discussed above. As also shown in Figure 4d, the TPI was a ridge in areas far away from the waterway, typically signifying a poor SNR class (Figure 4g) and dry SMR class (Figure 4h); alternatively, where the TPI was a valley, this usually signified a wet SMR class. These variations are indistinguishable in the coarse-resolution maps, which only referred to one value for each polygon. All of these results proved that high-resolution topo-hydrological predictors affected the distributions of SNR, SMR, and ecosite in detail.

Figure 4. Location of a sample area of Nova Scotia (a), and maps of coarse geology (b), digital elevation model (DEM), (c), DEM-derived topographic position index (TPI), (d), aspect (e), and sediment delivery ratio (SDR), (f), and model-predicted soil moisture regime (SMR), (g), soil nutrient regime (SNR), (h), and ecosite (i) and ecosites eliminated the polygons with less than 1 ha area (j).

4.3. Limitations and Future Improvements

Analyzing plot data indicated that the major reason for the lower accuracies in areas of extreme conditions, such as ecosite 1, ecosite 3, ecosite 7, and ecosite 10 from the Acadian region and ecosite 9 and ecosite 10 from the Maritime Boreal region, might be the problems in model estimations of SNR and SMR (Figure 3). These observations indicated that it was certain difficult to capture the extreme nutrient/moisture conditions, such as very rich/very poor and very dry/wet, using only the established models in this study, which was identical to the results of estimation error analysis discussed above (Table 2). Considering a total of 80–84% correctness for the Acadian region and Maritime Boreal region ecosites according to field assessment of SNRs and SMRs (Figure 1) and the remainder (16–20%) according to the field-identified vegetation type, it will be possible to improve the accuracy of model-estimated ecosites in areas with extreme conditions by introducing vegetation information in the future. In addition to remote sensing image-based high-resolution vegetation information, ecozone, ecoregion, and ecodistrict maps from the current ecological land classification system of the Province of Nova Scotia, which capture macroclimate and macro-topographic differences in vegetation [28], may also be a good candidate.

Although the best combinations of the nine DEM-generated terrain attributes, together with three coarse-resolution soil maps, could account for 81% of SNR total variation and 78% of SMR total variation, introducing more topo-hydrological indices [29] and more advanced neural networks, e.g., deep learning neural networks with multiple hidden layers and improved training algorithms [30] (if a large number of field data are available), would improve model performance. However, it should also be noted that a combined building scheme, using k-fold cross-validation with an early stopping technology, may avoid overfitting, which had poor model generalization outside of the training zone, to a certain extent, as well as perform well in areas with a similar environment to the training zone; however, the model predictions would not be justified in areas with a significantly different environment. Thus, the best SNR and SMR models of Nova Scotia should be retrained before being applied to areas with a significantly different environment. Furthermore, an ANN model is a kind of “black box” where there is little or no possibility of understanding its internal behavior [31]; thus, generating equations based on ANN [32] to delineate how various predictors affect SNR and SMR should be done in the future.

5. Conclusions

A new modeling method was introduced to map SNRs and SMRs with high resolution, and then produce ecosite maps in Nova Scotia, Canada. Using the biophysical land classification maps and high-resolution DEM-derived variables as model inputs, 511 ANNs were developed and evaluated on the basis of 1507 field data with 10-fold cross-validation. The results showed that the best model for mapping SNRs and SMRs engaged seven and eight DEM-generated variables, together with three coarse-resolution soil maps, as model inputs, respectively, whereby 61% of model-estimated SMRs were identical with field assessments, while this value was 82% for SNRs. Model prediction accuracy for the Acadian region ecosite map and Maritime Boreal region ecosite map revealed 67% and 68% correctness, respectively. The accuracies were evidently higher than those obtained in previous studies. Based on the role of field-identified vegetation type in identifying SMR, SNR, and ecosite type in plots and edatopic grids, it will be possible to improve the accuracy of model-estimated ecosites in areas with extreme conditions by introducing forest-cover information in the future. Meanwhile, introducing more topo-hydrological indices and applying deep learning neural networks would improve model performance; however, retraining the best ANNs for Nova Scotia will be required before their application to areas with a significantly different environment.

Author Contributions: Conceptualization, Z.Z.; methodology, Q.Y.; software, Z.X.; validation, X.D.; formal analysis, Q.Y.; investigation, Z.X.; resources, X.D.; data curation, X.D.; writing—original draft

preparation, Z.Z.; writing—review and editing, Z.X.; visualization, Q.Y.; supervision, Z.Z.; project administration, X.D.; funding acquisition, Q.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from the Guangxi Natural Science Foundation of China (Grant No. 2018GXNSFAA050135 and 2018GXNSFBA138035), and Guangdong Forestry Science and Technology Plan of China (Grant No. 2019-07). The authors are also grateful to the special founding from Guangxi Hundred-Talent Program to Zhengyong Zhao and Qi Yang.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Keys, K.; Neily, P.; Quigley, E.; Stewart, B. *Forest Ecosystem Classification for Nova Scotia Part III: Ecosites*; (2010); Report FOR 2011-2; Nova Scotia Department of Natural Resources, Renewable Resources Branch: Truro, SK, Canada, 2011; pp. 15–25.
2. Montigny, M.K.; MacLean, D.A. Using heterogeneity and representation of ecosite criteria to select forest reserves in an intensively managed industrial forest. *Biol. Conserv.* **2005**, *125*, 237–248. [[CrossRef](#)]
3. McLaughlan, M.S.; Wright, R.A.; Jiricka, R.D. *Field Guide to the Ecosites of Saskatchewan's Provincial Forests*; Saskatchewan Ministry of Environment: Prince Albert, Canada, 2010; pp. 34–46.
4. Sims, R.A.; Uhlig, P. The current status of forest site classification in Ontario. *For. Chron.* **1992**, *68*, 64–76. [[CrossRef](#)]
5. Taylor, K.C.; Arnup, R.W.; Merchant, B.G.; Parton, W.J.; Nieppola, J. *A Field Guide to Forest Ecosystems of Northeastern Ontario*, 2nd ed.; NEST Field Guide FG-001; Queen's Printer for Ontario: Toronto, ON, Canada, 2000; pp. 68–77.
6. New Brunswick Department of Natural Resources (NBDNR). *Our Landscape Heritage: The Story of Ecological Land Classification in New Brunswick*; New Brunswick Department of Natural Resources: Fredericton, NB, Canada, 2007; pp. 10–21.
7. Wang, G.G. Use of understory vegetation in classifying soil moisture and nutrient regimes. *For. Ecol. Manage.* **2000**, *129*, 93–100. [[CrossRef](#)]
8. Emadi, M.; Shahriari, A.R.; Sadegh-Zadeh, F.; Seh-Bardan, B.; Dindarlou, A. Geostatistics-based spatial distribution of soil moisture and temperature regime classes in Mazandaran province, northern Iran. *Arch. Agron. Soil Sci.* **2016**, *62*, 502–522. [[CrossRef](#)]
9. Zhao, Z.; Ashraf, M.I.; Keys, K.S.; Meng, F.-R. Prediction of soil nutrient regime based on a model of DEM-generated clay content for the province of Nova Scotia, Canada. *Can. J. Soil. Sci.* **2013**, *93*, 193–203. [[CrossRef](#)]
10. Akumu, C.E.; Baldwin, K.; Dennis, S. GIS-based modeling of forest soil moisture regime classes: Using Rinker Lake in northwestern Ontario, Canada as a case study. *Geoderma* **2019**, *351*, 25–35. [[CrossRef](#)]
11. Yang, Q.; Meng, F.-R.; Charles, P.-A.B.; Zhao, Z. Production of high-resolution forest-ecosite maps based on model predictions of soil moisture and nutrient regimes over a large forested area. *Sci. Rep.* **2017**, *7*, 10998. [[CrossRef](#)] [[PubMed](#)]
12. Bulut, B.; Yilmaz, M.T.; Afshar, M.H.; Sorman, A.U.; Yucel, I.; Cosh, M.H.; Şimşek, O. Evaluation of remotely-sensed and model-based soil moisture products according to different soil type, vegetation cover and climate regime using station-based observations over Turkey. *Remote Sens.* **2019**, *11*, 1875. [[CrossRef](#)]
13. Brady, N.C.; Weil, R.R. *The Nature and Properties of Soils*, 14th ed.; Pearson Education: Upper Saddle River, NJ, USA, 2008; pp. 134–153.
14. Arp, P.A. *Soils for Plant Growth: Field and Laboratory Manual*; University of New Brunswick: Fredericton, NB, Canada, 2005; pp. 21–56.
15. Zhao, Z.; Meng, F.-R.; Yang, Q.; Zhu, H. Using artificial neural networks to produce high-resolution soil property maps. In *Advanced Applications for Artificial Neural Networks*; El-Shahat, A., Ed.; IntechOpen: London, UK, 2018; pp. 51–75.
16. Zhao, Z.; Yang, Q.; Sun, D.; Ding, X.; Meng, F.-R. Extended model prediction of high-resolution soil organic matter over a large area using limited number of field samples. *Comput. Electron. Agr.* **2020**, *169*, 105172. [[CrossRef](#)]
17. Zhao, Z.; Yang, Q.; Benoy, G.; Chow, T.L.; Xing, Z.; Rees, H.W.; Meng, F.-R. Using artificial neural network models to produce soil organic carbon content distribution maps across landscapes. *Can. J. Soil Sci.* **2010**, *90*, 75–87. [[CrossRef](#)]
18. Ding, X.; Zhao, Z.; Yang, Q.; Chen, L.; Tian, Q.; Li, X.; Meng, F.R. Model prediction of depth-specific soil texture distributions with artificial neural network: A case study in Yunfu, a area of Uduits Zone, South China. *Comput. Electron. Agr.* **2020**, *169*, 105217. [[CrossRef](#)]
19. Beucher, A.; Møller, A.B.; Greve, M.H. Artificial neural networks and decision tree classification for predicting soil drainage classes in Denmark. *Geoderma* **2019**, *352*, 351–359. [[CrossRef](#)]
20. Hateffard, F.; Dolati, P.; Heidari, A.; Zolfaghari, A.A. Assessing the performance of decision tree and neural network models in mapping soil properties. *J. Mt. Sci.* **2019**, *16*, 1833–1847. [[CrossRef](#)]

21. Keys, K.; Neily, P.; Quigley, E.; Stewart, B. *Field Manual for Forest Ecosystem Classification in Nova Scotia*; Manual FOR 2007-1; Nova Scotia Department of Natural Resources: Truro, NS, Canada, 2007; pp. 112–124.
22. Zhao, Z.; Ashraf, M.I.; Meng, F.-R. Model prediction of soil drainage classes over a large area using limited number of field samples: A case study in the Province of Nova Scotia, Canada. *Can. J. Soil Sci.* **2013**, *93*, 73–83. [[CrossRef](#)]
23. ESRI Inc. *The Help Document*; ESRI Inc.: Redlands, CA, USA, Copyright 1999–2013.
24. Li, Z. Supervised classification of multispectral remote sensing image using B-P Neural Network. *J. Infrared Milli. Wave* **1998**, *17*, 153–156.
25. Sigillito, V.G.; Hutton, L.V. Case study II: Radar signal processing. In *Neural Network PC Tools*; Eberhart, R.C., Dobbins, R.W., Eds.; Academic Press Professional, Inc.: San Diego, CA, USA, 1990; pp. 96–114.
26. Zhao, Z.; Chow, T.L.; Rees, H.W.; Yang, Q.; Xing, Z.; Meng, F.R. Predict soil texture distributions using an artificial neural network model. *Comput. Electron. Agr.* **2009**, *65*, 36–48. [[CrossRef](#)]
27. MacMillan, R.A.; Moon, D.E.; Coupe, R.A. Automated predictive ecological mapping in a forest region of B.C., Canada, 2001–2005. *Geoderma* **2007**, *140*, 353–373. [[CrossRef](#)]
28. Neily, P.; Quigley, E.; Benjamin, L.; Stewart, B.; Duke, T. *Ecological Land Classification for Nova Scotia: Volume 1—Mapping Nova Scotia's Terrestrial Ecosystems*; Report DNR 2003-2; Nova Scotia Department of Natural Resources: Truro, NS, Canada, 2003; pp. 1–15.
29. Li, Q.; Wei, X.; Yang, X.; Giles-Hansen, K.; Zhang, M.; Liu, W. Topography significantly influencing low flows in snow-dominated watersheds. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 1947–1956. [[CrossRef](#)]
30. Pyo, J.; Hong, S.M.; Kwon, Y.S.; Kim, M.S.; Cho, K.H. Estimation of heavy metals using deep neural network with visible and infrared spectroscopy of soil. *Sci. Total Environ.* **2020**, *741*. [[CrossRef](#)]
31. Lopez, C. Looking inside the ANN “black box”: Classifying individual neurons as outlier detectors. In *Proceedings of the International Joint Conference on Neural Networks*, Washington, DC, USA, 10–16 July 1999; pp. 1185–1188.
32. Zhao, Z.; MacLean, D.A.; Bourque, C.P.-A.; Swift, D.E.; Meng, F.-R. Generation of soil drainage equations from an artificial neural network-analysis approach. *Can. J. Soil Sci.* **2013**, *93*, 329–342. [[CrossRef](#)]