Remote Sensing of Environment 170 (2015) 153–165

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Comparison of passive microwave brightness temperature prediction sensitivities over snow-covered land in North America using machine learning algorithms and the Advanced Microwave Scanning Radiometer

Yuan Xue *, Barton A. Forman

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Article history:
Received 27 May 2015
Received in revised form 14 September 2015
Accepted 21 September 2015
Available online xxxx

Keywords:
Sensitivity analysis
Machine learning
Brightness temperature
Snow
Data assimilation

ABSTRACT

Recent studies showed that machine learning (ML) algorithms (*e.g.*, artificial neural network (ANN) and support vector machine (SVM)) reasonably reproduce passive microwave brightness temperature observations over snow-covered land as measured by the Advanced Microwave Scanning Radiometer (AMSR-E). However, these studies did not explore the sensitivities of the ML algorithms relative to ML inputs in order to determine the behavior and performance of each algorithm. In this current study, normalized sensitivity coefficients are computed to diagnose ML performance as a function of time and space. The results showed that when using the ANN, approximately 20% of locations across North America are relatively sensitive to snow water equivalent (SWE). However, more than 65% of locations in the SVM-based brightness temperature (Tb) estimates are sensitive relative to perturbations in SWE at all frequency and polarization combinations explored in this study. Further, the SVM-based results suggest the algorithm is sensitive in both shallow and deep SWE, SWE with and without overlying forest canopy, and during both the snow accumulation and snow ablation seasons. Therefore, these findings suggest that compared with the ANN, the SVM could potentially serve as a more efficient and effective measurement model operator within a Tb data assimilation framework for the purpose of improving SWE estimates across regional- and continental-scales.

1. Introduction and motivation

Snow is a critical component in the global energy and hydrologic cycle due to its control of mass and energy exchanges at the land surface (Robinson, Dewey, & Heim, 1993). However, direct quantification (*i.e.*, *in-situ* measurements) of the mass of snow (a.k.a., snow water equivalent (SWE)) is significantly complicated by spatial and temporal variability in snow processes. Therefore, space-borne passive microwave (PMW) SWE retrieval products have been employed to help fill observational gaps between ground-based sensors to better estimate SWE at the global scale based on the relationship between the measured electromagnetic response and the physical characteristics of SWE. Unfortunately, the highly nonlinear nature of the relationship is non-trivial to establish and numerous limitations exist that restrict the extensive application of PMW-based SWE estimates.

There are typically four ways to estimate SWE from space-borne sensors. One method is to merge relatively coarse, space-borne observations with *in-situ* measurements of finer resolution *via* spatial interpolation (Cao, Yang, & Zhu, 2008). However, this method is adversely impacted by sparse spatial coverage of *in-situ* observations, particularly in regions near the Arctic Circle (Takala et al., 2011), coupled with strong sub-grid scale snow variability in complex terrain (Foppa, Stoffel, & Meister, 2007). The second technique – space-borne PMW SWE retrieval – inverts (or retrieves) model states variables from the measured brightness temperature (Tb, defined as the physical temperature of an object times its emissivity) at specific frequencies by calibrating regression coefficients within the algorithm (Chang, Foster, & Hall, 1987; Goodison & Walker, 1994; Kelly, Chang, Tsang, & Foster, 2003; Chang, Foster, & Hall, 1996; Kelly, 2009). These satellite-based SWE retrieval models are often affected by errors arising from meteorological fields (*e.g.*, data aggregation, disaggregation, extrapolation and interpolation (Blöschl & Sivapalan, 1995)) used to force land surface models. They are also affected by significant uncertainties associated with snow stratigraphy (Derksen, Walker, & Goodison, 2005), snow grain size (Armstrong, Chang, Rango, & Josberger, 1993), depth hoar layer (Brucker, Royer, Picard, Langlois, & Fily, 2011; Hall, 1987; Hall, Chang, & Foster, 1986; Foster et al., 2005), ice crusts (Rees, Lemmetyinen, Derksen, Pulliainen, & English, 2010), lake fraction effects (Derksen et al., 2010), and snow morphology (Kelly et al., 2003), especially in densely-vegetated regions (Tedesco & Narvekar, 2010; Derksen et al., 2005) with relatively deep snow (Clifford, 2010).

* Corresponding author at: 0147B Engineering Laboratory Building, Department of Civil and Environmental Engineering, University of MD, College Park, MD 20742, United States
E-mail address: yxue12@umd.edu (Y. Xue).

http://dx.doi.org/10.1016/j.rse.2015.09.009

In an effort to overcome many of the limitations highlighted above, the third alternative involves merging measurements of remote sensing observations with estimates from physically-based models (Reichle, 2008; Forman, Reichle, & Rodell, 2012; Reichle, De Lannoy, Forman, Draper, & Liu, 2014) using data assimilation (DA). The goal of DA (with particular relevance to SWE discussed here) is to yield a merged estimate that is superior to either the observations or the model alone (Mclaughlin, 2002). Typically, a radiative transfer model (RTM) is used (Liang, Xu, Tsang, Andreadis, & Josberger, 2008) as a model operator to invert PMW Tb measurements into model (SWE) space. However, the practical utilization of a RTM is complicated by wet, moderately deep snowpacks (greater than the 100 mm SWE), snow located closer than 200 km to open water (Dong, Walker, Houser, & Sun, 2007), the existence of ice layers on or within the snowpack (Durand, Kim, & Margulis, 2011), and significant sub-grid scale variability (*e.g.*, mixed land cover within remotely sensed pixels (Andreadis, Liang, Tsang, Lettenmaier, & Josberger, 2008)).

The fourth method is to employ a machine learning (ML) technique (instead of a RTM) to estimate SWE, which has been conducted in a few studies (Chang & Tsang, 1992; Tsang, Chen, Oh, Marks, & Chang, 1992; Davis, Chen, Tsang, Hawang, & Chang, 1993; Tedesco, Pulliainen, Takala, Hallikainen, & Pampaloni, 2004; Cao et al., 2008). These studies focused on directly training an artificial neural network (ANN) using *in-situ* SWE observations. Reasonable performance was restricted to *in-situ* observation locations with less applicability to regions between these locations (Tedesco et al., 2004).

Recent research conducted by Forman, Reichle, and Derksen (2013); Forman and Reichle (2014) investigated the possibility of estimating Tbs (rather than SWE) by utilizing ML algorithms with an artificial neural network (ANN) or a support vector machine (SVM). In a DA context, the belief is the direct assimilation of Tb (rather than SWE) is preferable as it will avoid inconsistencies in the use of ancillary data between the assimilation system and the pre-processed geographical retrievals (Eyre, Kelly, McNally, Andersson, & Persson, 1993). It is further hypothesized that ML will serve as a more reliable model operator (relative to snow emission RTM) since current land surface models lack the fidelity at regional and continental scales to meet the needs of a snow emission model (*e.g.*, snow grain size, depth hoar development, internal ice layering) (Durand & Margulis, 2007).

It was concluded in Forman et al. (2013) and Forman and Reichle (2014) that both the ANN and SVM could eventually be used as measurement operators to estimate Tbs within a DA framework for the purpose of SWE estimation at regional and continental scales. However, a number of fundamental questions must first be addressed. For example, do the ANN and SVM reproduce Tb for the right (*i.e.*, physically-based) reasons? Further, what are the most significant input variables to the ML models? Are the accurate Tb estimates over snow-covered land associated with the snow-related variables (*e.g.*, SWE)? If so, under which conditions (*e.g.*, with or without overlying vegetation) will the ML models be sensitive to SWE? Or is the sensitivity of the ML model output due to non-snow-related state variables (*e.g.*, soil temperature and air temperature)? Therefore, the goal of this current study is to explore the ANN- and SVM-derived Tb sensitivities using a unified framework in an effort to answer the questions formulated above.

2. Machine learning and model formulation

Arthur Samuel (1959) first defined ML as a field of study that gives computers the ability to learn without being explicitly programmed. An alternative definition is the process of identifying a set of categories (sub-populations) where a new observation belongs on the basis of a training set of data containing observations whose category membership is known (Hastie, Tibshirani, Friedman, & Franklin, 2005). Based on properly constructed systems with proper parameterizations, ML algorithms are capable of learning about the regularities present in the training data such that constructing and generalizing rules can be extended to the unknown data during the training phase.

There is a plethora of ML algorithms to choose from depending on what type of question needs to be addressed. An ANN and a SVM framework are selected in this study (with particular relevance to SWE) because (1) they are data-driven models (He, Wen, Liu, & Du, 2014) used in cases where the underlying physical relationships between the electromagnetic response and SWE characteristics are not fully understood and (2) they can be used to reproduce nonlinear processes *via* iterations without prior knowledge about the relationship between the parameters (Suykens, Vandewalle, & De Moor, 2001) (*e.g.*, snow grain size and SWE).

However, some differences between these two types of ML techniques are also evident. For example, the existence of local minima (Smola & Schölkopf, 2004) could prevent an ANN from finding the unique global minimum solution to a constrained optimization problem, which is not the case for a SVM, which possesses a more simple geometric interpretation (Burges, 1998) characterized by convex optimization problems and thereby a unique global optima will always be found. Additionally, if the size of the training examples is not large enough, the SVM is expected to perform well based on a properly-selected mechanism of model parameters since the number of support vectors in the decision (feature) space is far less than the number of training points (Tsang, Kwok, Cheung, & Cristianini, 2005) whereas an ANN is always in need of a relatively large number of training points.

Both ANN- and SVM-based techniques in this study utilize the same model inputs derived from the NASA Catchment land surface model (Catchment; Koster, Suarez, Ducharne, Stieglitz, & Kumar, 2000) and output Tbs at three different frequencies (10.65 GHz, 18.7 GHz, and 36.5 GHz) at both horizontal and vertical polarization (see Table 1). Uncertainty and errors in Catchment-derived model output, including SWE, were discussed in detail in Reichle et al. (2011). In general, the Catchment model is unbiased (Reichle et al., 2011) and the brightness temperatures produced from the machine learning algorithms are also unbiased (Forman et al., 2013; Forman & Reichle, 2014). Therefore, it is hypothesized that the first statistical moment related to the mode of estimated SWE in the Catchment model is reasonably characterized.

Each ML technique is trained with the same nine-year (2002–2011) training dataset of Tb observations from the Advanced Microwave Scanning Radiometer — Earth Observing System (AMSR-E) where forest and atmospheric effects were not removed prior to ANN or SVM training in this study. All Catchment-based inputs (*i.e.*, the 11 model inputs listed in Table 1), AMSR-E training data, ANN-based output, and SVM-based output (*e.g.*, six different Tbs listed in Table 1) are generated on the

Table 1
Model (ANN and SVM) inputs and outputs (reproduced from Forman et al., 2013).

Inputs	Symbol	Unit
Top layer snow density	ρ_{sn1}	kg/m^3
Middle layer snow density	ρ_{sn2}	kg/m^3
Bottom layer snow density	ρ_{sn3}	kg/m^3
Snow liquid water content[a]	SLWC	kg/m^3
Snow water equivalent[a]	SWE	m
Near-surface air temperature	T_{air}	K
Near-surface soil temperature	T_{p1}	K
Skin temperature	T_{skin}	K
Top layer snow temperature	T_{sn1}	K
Bottom layer snow temperature	T_{sn3}	K
Temperature gradient index	TGI	–
Outputs	**Symbol**	**Unit**
Brightness temperature at 10.65 GHz, H-polarization	10 H	K
Brightness temperature at 10.65 GHz, V-polarization	10 V	K
Brightness temperature at 18.7 GHz, V-polarization	18 H	K
Brightness temperature at 18.7 GHz, H-polarization	18 V	K
Brightness temperature at 36.5 GHz, V-polarization	36 H	K
Brightness temperature at 36.5 GHz, H-polarization	36 V	K

[a] Column-integrated quantity.