INPUT PARAMETERS SELECTION FOR SOIL MOISTURE RETRIEVAL USING AN ARTIFICIAL NEURAL NETWORK

Soo-See Chai[1], Bert Veenendaal[1], Geoff West[1], Jeffrey P. Walker[2]

[1]Department of Spatial Sciences, Curtin University of Technology, Western Australia
[2] Department of Civil and Env. Eng., University of Melbourne, Melbourne
[1] soosee.chai@postgrad.curtin.edu.au, [1](b.veenendaal, g.west)@curtin.edu.au,
[2] j.walker@unimelb.edu.au

ABSTRACT

Factors other than soil moisture which influence the intensity of microwave emission from the soil include surface temperature, surface roughness, vegetation cover and soil texture which make this a non-linear and ill-posed problem. Artificial Neural Networks (ANNs) have been demonstrated to be good solutions to this type of problem. Since an ANN is a data driven model, proper input selection is a crucial step in its implementation as the presence of redundant or unnecessary inputs can severely impair the ability of the network to learn the target patterns. In this paper, the input parameters are chosen in combination with the brightness temperatures and are based on the use of incremental contributions of the variables towards soil moisture retrieval. Field experiment data obtained during the National Airborne Field Experiment 2005 (NAFE'05) are used. The retrieval accuracy with the input parameters selected is compared with the use of only brightness temperature as input and the use of brightness temperature in conjunction with a range of available parameters. Note that this research does not aim at selecting the best features for all ANN soil moisture retrieval problems using passive microwave. The paper shows that, depending on the problem and the nature of the data, some of the data available are redundant as the input of ANN for soil moisture retrieval. Importantly the results show that with the appropriate choice of inputs, the soil moisture retrieval accuracy of ANN can be significantly improved.

INTRODUCTION

Soil moisture controls several processes at or near the land surface. These processes include: the partition of rainfall into infiltration and run-off, the partition of available energy into latent and sensible heat, the drainage of water to ground water and/or surface water, and the growth of vegetation. All these processes are strong and nonlinear functions of soil moisture (Teuling *et al.* 2006). To forecast and predict these processes, accurate soil moisture observations are crucial.

Information on soil moisture can be obtained by either point measurements or by remote sensing techniques. Point-based measurements of soil moisture supply accurate information but they are costly and time consuming to obtain with limited spatial coverage. They provide in-situ measurements of soil moisture that can be repeated at the same locations at various times, but site specific calibration is required. In addition,

Chai, S.S., Veenendaal,B., West G. and J.P. Walker (2009). Input Parameter Selection for Soil Moisture Retrieval Using an Artificial Neural Network. In: Ostendorf, B., Baldock, P., Bruce, D., Burdett, M. and P. Corcoran (eds.), Proceedings of the Surveying & Spatial Sciences Institute Biennial International Conference, Adelaide 2009, Surveying & Spatial Sciences Institute, pp. 1181-1193. ISBN: 978-0-9581366-8-6.

practical and economic reasons restrict sampling to be spatially sparse with significant time between samples. For these reasons point-based sampling cannot capture the subtle variations in topography, soil texture and evapotranspiration at the required spatial and temporal scales.

Remote sensing, on the other hand, has the potential to provide a better spatial and temporal coverage of soil moisture when compared to point-based measurements, but can only sense the top few centimetres for areas with moderate to low vegetation cover. Platforms supporting remote sensing instruments can be either ground-based, aircraft or satellite based (Wu 1996). Among these, space platforms offer the optimum solution for the repeated mapping of large areas over long time periods. However, soil moisture measured using these techniques may suffer from inadequate spatial and/or temporal resolution.

Land surface models have been created to simulate the real-world conditions in order to predict the spatial and temporal variations of soil moisture. Some of the widely used land surface models include the Soil and Water Assessment Tool (SWAT) (Arnold *et al.* 1998), Integral Equation Model (IEM) (Fung *et al.* 1992) and the Soil-Plant-Air-Water (SPAW) model (Saxton and Willey 2006). These models require the estimation of a number of physical parameters which can be difficult to determine. To overcome this difficulty, researchers in soil moisture problems have been utilizing data driven forecasting approaches. Artificial Neural Networks (ANNs) are one of the most widely adopted data driven models (Elshorbagy and Parasuraman 2008). Various attempts at using this statistical method to estimate soil moisture have been reported. For example, recent research by Angiuli *et al.* (2008) studied the potential of ANN in the context of the European Space Agency (ESA)'s Soil Moisture and Ocean Salinity (SMOS) mission, focusing on land data. The goal of the SMOS mission is to monitor soil moisture with a Root Mean Square Error (RMSE) of 4% v/v with vegetation water content of less than 5 kg/m^2 (Kerr *et al.* 2001). A standard backpropagation algorithm with brightness temperatures at different incidence angles was used as the input by Angiuli *et al.* (2008). Their retrieval showed that the ANN can obtain a best result of RMSE of 5% v/v. To the knowledge of the authors when writing this paper, there is no study on the use of ancillary data on ANN soil moisture retrieval using passive microwave data, but ancillary data has been used as one of the input features for the ANN models. Atluri *et al.*(1999) for example, use surface temperature as an ancillary factor with the brightness temperature as input for the ANN model for soil moisture retrieval. Shou-Fang *et al.*(2002), on the other hand, argued that ancillary factors such as vegetation biomass, surface temperature, and surface roughness, are not required when using an ANN.

This paper examines the use of ancillary data together with brightness temperature as input data for soil moisture retrieval using an ANN. The incremental contribution of a variable (factor) will be used to decide whether a variable (factor) is to be excluded from the prediction model.

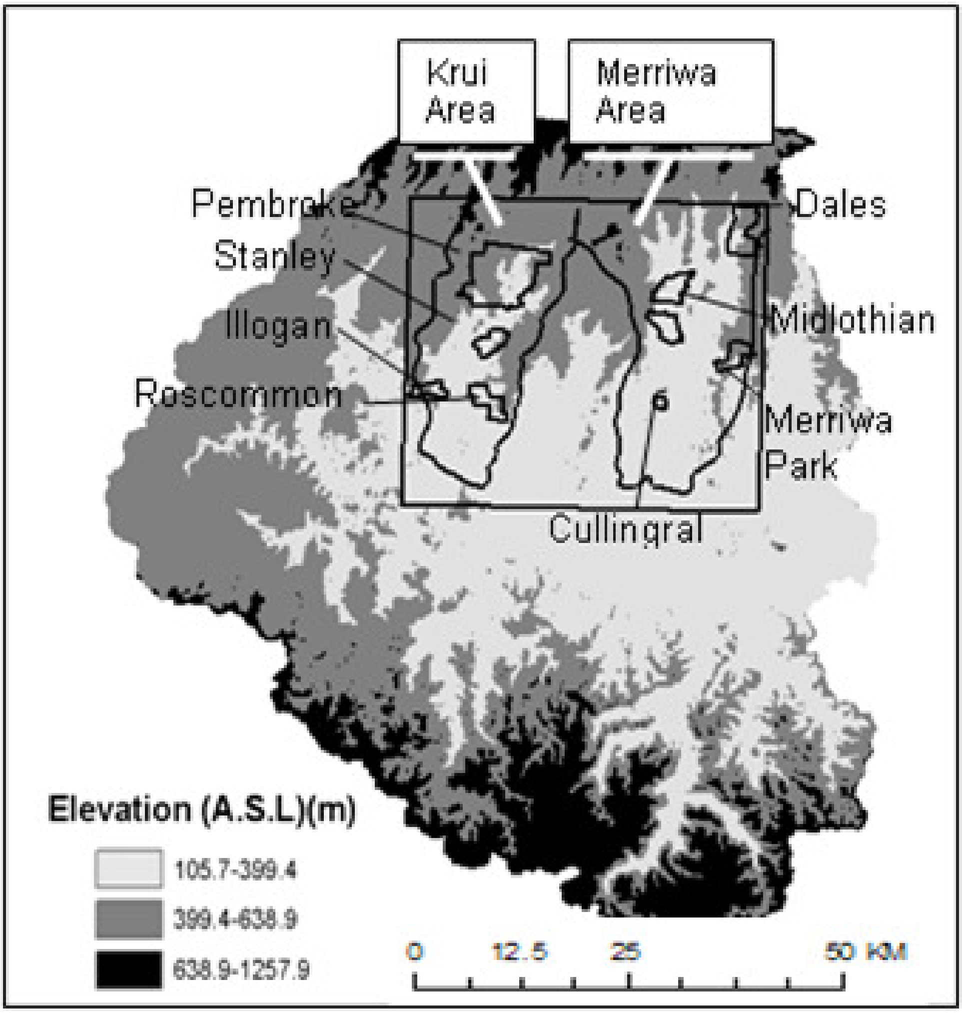
NAFE'05 DATA

The data used in this study have been collected during the month-long NAFE field campaign held in November 2005. The campaign included extensive airborne passive microwave observations together with spatially distributed and in-situ ground monitoring of soil moisture and related variables. For the purpose of this analysis, only pertinent details of the available data will be presented. Full detail description of the data can be obtained from Panciera *et al.*(2008).

The Airborne Data

The study area is situated in the northern part of the Goulburn River catchment, located in a semi-arid area of south-eastern Australia. The area monitored during NAFE'05 was a square area of approximately 40km×40km logistically divided into two sub-areas, the "Merriwa" area to the right and the "Krui" area to the left as shown in **Figure 1**.

Figure 1: Overview of the NAFE'05 study area showing the focus farms within the Krui and Merriwa areas



Flights were conducted between October 31st and November 25th with a small two-seater motor glider from the Airborne Research Australia national facility together with the Polarimetric L-band Multi-beam Radiometer (PLMR). The PLMR measures both H and V polarized brightness temperatures (Tb) using a single receiver with a polarization switch at incidence angles of +/-7º, +/-21º and +/-38.5º in either across track (push-broom) or along track configurations. For the purpose of this study, the Tb data was corrected to +/-38.5º.

Multi-resolution flights were undertaken with the PLMR instrument in "push-broom" configuration. The area mapped on each date was either the Krui or the Merriwa area. With four different flying altitudes, four different ground resolutions of: (i) 1km, (ii) 500m, (iii) 250m and (iv) 62.5m pixel sizes were obtained. During these flights, full coverage of the same ground area was guaranteed at each resolution and to avoid gaps in the data due to anomalous aircraft altitude or the terrain elevation effect on the swath, a full one pixel wide overlap was guaranteed for adjacent flight lines on all "push-brooms" flights.

The Ground Data

Spatial ground sampling was concentrated in the 40km×40km region and the eight focus farms, with the near-surface soil moisture data collected across the region and the farms at a range of spatial scales from 6.25m to 2km. Additionally, data were collected on surface roughness, skin and soil temperature, vegetation water content, land cover and the vegetation Normalized Difference Vegetation Index (NDVI). The timetable for ground sampling is shown in **Table 1**.

Table 1: Ground sampling timetable for NAFE'05

	Mon 31/10	Tues 1/11	Wed 2/11	Thurs 3/11	Fri 4/11
Regional flight					
Krui Area					
Merriwa Area					
	Mon 7/11	Tues 8/11	Wed 9/11	Thurs 10/11	Fri 11/11
Regional Flight					
Krui Area					
Merriwa Area					
	Mon 14/11	Tues 15/11	Wed 16/11	Thurs 17/11	Fri 18/11
Regional Flight					
Krui Area					
Merriwa Area					
	Mon 21/11	Tues 22/11	Wed 23/11	Thurs 24/11	Fri 25/11
Regional Flight					
Krui Area					
Merriwa Area					

Near-surface Soil Moisture Monitoring

The soil moisture within the top 5cm of the soil profile was monitored coincident with each aircraft flight either across the entire area or across the focus farms. Measurements of the top 5cm soil moisture content were taken using the Hydraprobe Data Acquisition System developed by the University of Melbourne that integrates a Global Positioning System and soil moisture sensor with a Geographic Information System (Panciera *et al.* 2006). During the regional area sampling, the soil moisture was sampled on a grid of approximately 2km. For other dates, the sampling was focused on two focus farms. The sampling in the focus farms was divided into a few resolutions. For very high resolution sampling, a 150m×150m area was selected at each farm and soil moisture was measured at 12.5m and 6.25m spacing. The area surrounding the very high resolution sampling areas was sampled at intermediate resolutions (125m to 250m spacing). The remaining extent of the focus farms was sampled at coarser resolution (500m and/or 1km spacing).

Vegetation Data

On each farm, the spatial variability of vegetation biomass and water content was characterized by collecting between four to sixteen 0.5m×0.5m quadrant samples across the high resolution soil moisture sampling area, supported by a minimum of five quadrant samples of the dominant vegetation types across the farm. This was undertaken once a week at fixed locations to monitor the temporal changes in vegetation biomass and water content. On all the other days, the vegetation water content samples were collected from two corners of the high resolution areas as a check on the temporal changes of the farm vegetation water content. The Normalized Difference Vegetation Index (NDVI) was measured for the high resolution areas of each focus farm using an Exotech Inc. Hand Held Radiometer 100BX at 50m spacing. This was done at least once during the campaign at each farm.

Temperature Data

The Goulburn catchment is instrumented with soil temperature monitoring infrastructure. Each of the soil moisture sites had a Stevens Water Hydraprobe which measures the soil temperature at 2.5cm. For the purpose of this study, the daily average soil temperature from 8.00 a.m. to 1.00 p.m. is used that corresponds to when the PLMR passed over the study area.

Roughness Data

The surface roughness was estimated once during the campaign at a minimum of four locations on each focus farm to capture the different roughness characteristics according to land cover type. Two 1m long roughness profiles were recorded for each measurement location, i.e. one north-south and one east-west oriented. The average Root Mean Square (RMS) roughness value for each of the focus sites is be used in this paper.

Soil Textural Properties

Volumetric samples of the top 5cm of soil were collected across the study area for the soil textural analysis. A total of 20 samples were collected at each focus farm, to characterize the different soil type and wetness conditions across the farms. Laboratory soil textural analysis was performed on the subset of the soil samples for fractions of sand, clay and silt.

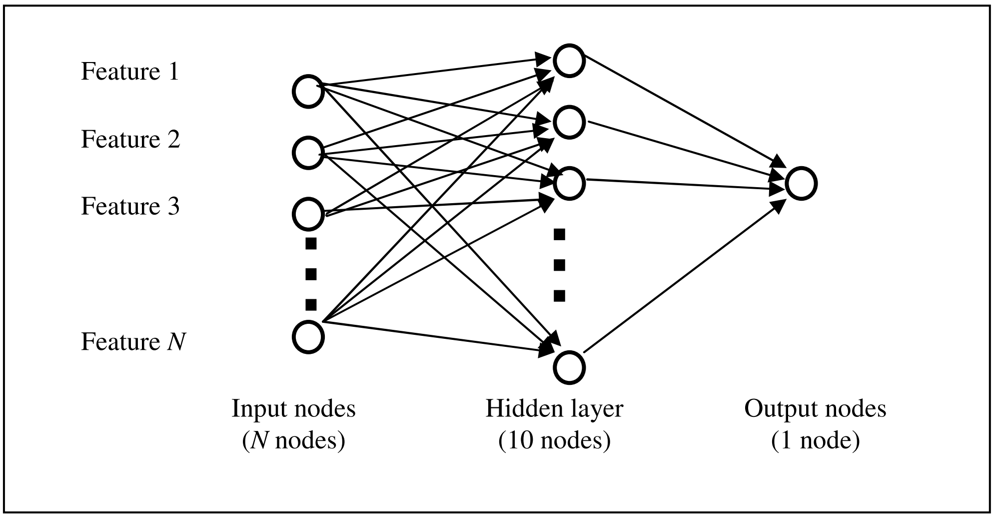
ARTIFICIAL NEURAL NETWORKS: THE PREDICTION MODEL

ANNs are mathematical models based on the biological neurons in human brains. They consist of interconnected artificial neurons which process information using a connectionism approach (Junlei *et al.* 2008). ANN has been applied successfully in many prediction cases as a non-linear statistical data modeling tool. Examples of such applications include: drought monitoring (Mishra and Desai 2006), flood forecasting (Campolo *et al.* 2003), and the classification and predicting of soil moisture (Atluri *et al.* 1999).

The relationship between brightness temperature and soil moisture is non-linear and complex. For this reason, the ANN is chosen for its suitability in modeling complex relationship between inputs and outputs (Lakhankar 2006). A three-layer feedforward NN network model was used for this research with input layers comprising the selected input variables, a hidden layer with 10 hidden neurons, and an output layer of predicted soil moisture values (denoted as $\{N, 10, 1\}$ where N is the number of inputs). The information from the inputs will move forward, without cycles, to the outputs (**Figure 2**). The sigmoid transfer function and the linear transfer function were used in the hidden and output layers, respectively.

The Quasi-Newton Backpropagation algorithm was employed to train the network During the forward phase, the inputs are propagated through the network layer to layer. Next, an error is computed by comparing the difference between the desired output and the predicted output. This error is propagated backward and the weights and biases are adjusted to minimize this error.

Figure 2: Architecture of the ANN used for prediction of soil moisture value.



INCREMENTAL CONTRIBUTIONS OF VARIABLES

When considering what inputs to use for an ANN, the effect of adding/removing an input can be used as an indication. The incremental contribution of an input can be explained by the reduction of the explained variance of the dependent variable (output) when we exclude an explanatory variable (input) (Kaashoek and Dijk 2002). A natural candidate for quantification of the network performance is the square of the correlation, R^2. The network performance with only one input deleted can be measured in a similar way.

For the purpose of this study, the ANN architecture was first optimized using all the available features as input. The optimization produces a correlation of R_y^2. When the contribution of an input of a feature is set to zero, the same network after this first optimization without this particular input produces a correlation of R_i^2. The incremental contribution of this particular input is then defined as:

$$R_y^2 - R_i^2 \qquad (1)$$

If the value of Equation (1) is low for some input n compared to all other inputs, then this input is a candidate for exclusion from the network. In the research by Kaashoek and Dijk (2002), a feature is to be considered for exclusion if the value of Equation (1) is less than one tenth of the feature with the highest contribution.

EXPERIMENTS AND RESULTS

For this study, experiments were carried out to determine the selection of input factors from the available inputs for the study area: H- and V-polarized brightness temperatures (TbH, and TbV), surface temperature (Ts), vegetation water content (VWC), Normalized Difference Vegetation Index (NDVI), average RMS roughness value (Rs), and percentage of silt (%Si), sand (%Sa) and clay (%Cl) of soil textural properties.

Two sets of experiments were carried out and discussed here. These two experiments will compare the effects of constant values of the ancillary data. The first experiment was conducted using the data of the Pembroke focus farm and the second experiment was conducted using data from a combination of three focus farms, Roscommon, Stanley and Cullingral. Pembroke was chosen as it is the largest among the eight focus farms which will provide adequate data for training of the ANN. Experiment 2 requires focus farms of different characteristics in term of topography and land cover. For these reasons, Roscommon, Stanley and Cullingral with the characteristics stated in **Table 2** were selected. In addition to this, Pembroke has a high average vegetation biomass of 1.5 kg/m^2 comparing to Roscommon 0.6 kg/m^2, Stanley 0.5 kg/m^2 and Cullingral 0.5 kg/m^2. The average spatial variability of soil moisture for Pembroke was 4.5% v/v, Roscommon 3.3% v/v, Stanley 5.8% v/v and Cullingral 11% v/v. Roscommon was considered as the "control" site with minimum soil moisture heterogeneity. The spatial data resolution of 500m was used. The data at this resolution is the coarsest with adequate number of data available on the study area. As ground-based ancillary data will be used, most of these data will be either the daily average or a constant for a focus farm, data at finer resolutions (i.e. 250m, 125m and 62.5m) will have highly detailed ground information which could not be represented using these ancillary data. The characteristics of each of these farms are summarized in **Table 2**.

Table 2: Characteristics of the data for each of the focus farms used in the experiments

Date	TbH(K) (max/min)	TbV(K) (max/min)	SM(v/v) (max/min)	VWC (kg/m^2)	NDVI	Rs	%Cl	%Si	%Sa	No. of Data
Focus Farm : Pembroke (Area : 6400 ha, Topography : Hilly/Gently rolling, Land cover: Wheat and Barley)										
8/11	258.7/243.4	277.4/260.3	0.63/0.28	0.54	0.71	0.84	62.5	29.5	8	88
15/11	269.7/256.4	283.2/272.1	0.35/0.14	2.03	0.71	0.84	62.5	29.5	8	88
17/11	270.3/258.1	282.0/271.1	0.36/0.14	0.91	0.71	0.84	62.5	29.5	8	88
22/11	273.7/263.7	284.5/276.1	0.22/0.06	2.41	0.71	0.84	62.5	29.5	8	88
Focus Farm : Roscommon (Area : 940 ha, Topography : Flat/Gently rolling, Land cover: Grassland)										
1/11	237.0/205.1	262.9/235.0	0.38/0.19	0.77	0.60	0.62	6.5	8.5	85	11
8/11	244.2/222.9	262.7/253.5	0.26/0.11	0.83	0.60	0.62	6.5	8.5	85	11
15/11	271.0/260.2	285.1/278.4	0.10/0.04	0.48	0.60	0.62	6.5	8.5	85	11
22/11	275.4/269.2	286.3/282.3	0.05/0.01	0.44	0.60	0.62	6.5	8.5	85	11
Focus Farm : Stanley (Area : 720 ha, Topography : Hilly, Land cover: Grassland)										
3/11	255.1/243.7	275.1/265.4	0.46/0.23	0.37	0.73	1.07	39.2	39.4	21.4	17
10/11	255.6/241.7	272.4/260.7	0.47/0.21	0.07	0.73	1.07	39.2	39.4	21.4	17
17/11	268.7/260.4	281.3/274.4	0.25/0.05	0.31	0.73	1.07	39.2	39.4	21.4	17
24/11	248.4/237.3	267.8/255.2	0.37/0.16	0.29	0.73	1.07	39.2	39.4	21.4	17
Focus Farm : Cullingral (Area : 220 ha, Topography : Flat, Land cover:Wheat and Barley)										
4/11	255.5/249.9	273.2/265.7	0.41/0.11	0.87	0.60	0.65	0	6	94	8
9/11	249.0/240.5	270.0/258.5	0.64/0.14	0.48	0.60	0.65	0	6	94	8
18/11	275.1/264.8	285.9/277.4	0.24/0.006	0.42	0.60	0.65	0	6	94	8
25/11	271.8/262.1	281.9/273.6	0.23/0.09	0.36	0.60	0.65	0	6	94	8

Experiment 1: Pembroke

Data obtained for the Pembroke focus farm on 8th Nov and 15th Nov 2005 were used for training of the ANN. Of the 176 samples available on these two dates, 3% (5 samples) of the data were randomly selected for validation and 3% (5 samples) for testing. Training was stopped when either the ANN met the maximum number of validation failures (5 times) or reached the maximum number of epochs (200). The ANN which produced the lowest RMSE on the testing samples was then taken as the trained ANN. At this stage, the weights and biases were held constant. The contribution of each of the input was evaluated by setting the weights of the chosen input to be zero and the trained ANN run to examine the correlation R^2. The incremental contribution of each of the inputs is shown in **Table 3**.

Table 3: Incremental contribution of each of the inputs for Experiment 1.

Input Excluded	R^2	Incremental Contribution	RMSE (% v/v)
None	0.5031	-	4.09
I1 (TbH)	0.4893	0.0138	3.94
I2 (TbV)	0.4482	0.0549	4.06
I3 (Ts)	0.5041	-0.0010	3.58
I4 (VWC)	0.5311	-0.0280	4.15
I5 (NDVI)	0.5040	-0.0009	4.06
I6 (Roughness)	0.5040	-0.0009	4.06
I7 (%Clay)	0.5031	0	4.09
I8 (%Silt)	0.5031	0	4.09
I9 (%Sand)	0.5031	0	4.09

From Table 4, the incremental contribution for TbH and TbV are higher compare to the other parameters (i.e. more than one tenth of the contributions of all nine inputs). The incremental contributions of I3(Ts), I4(VWC), I5(NDVI), I6(Roughness), I7 (%Clay), I8(%Silt), and I9(%Sand) are very small and hence these inputs are candidates for exclusion. After these inputs are excluded from the ANN, the ANN was re-trained to obtain the best parameters and performance. Then the process of checking the incremental contribution of each of the remaining variables is again undertaken with the results shown in **Table 4**.

Table 4: Incremental contribution of TbH and TbV after exclusion of other inputs.

Input Excluded	R^2	Incremental Contribution	RMSE (%v/v)
None	0.5307	-	3.92
I1 (TbH)	0.2233	0.3074	4.03
I2 (TbV)	0.3576	0.1731	6.32

Table 4 shows that the incremental contributions for each of the two inputs are of similar magnitude and the lowest is greater that one tenth of the largest. Hence no further exclusion is needed. This results in a network of {2, 10, 1} i.e. two inputs, 10 hidden nodes and one output node. To verify that these inputs alone produce either superior or almost the same accuracy with inclusion of any other ancillary data, the lowest RMSE value for each of a number of combinations of ancillary factors with the brightness temperature is shown in **Table 5**.

Table 5 shows that using just the brightness temperatures: TbH and TbV, produces the best accuracy for both dates. The ANN model was used to predict the soil moisture for the two different dates (17th Nov. and 22nd Nov.) after training on two previous dates (8th Nov. and 15th Nov.). Although good accuracy was obtained on the training data, the accuracy were not within the acceptable SMOS mission requirement of 4% v/v for the 17th and 22nd of November dates. These results are similar to the results obtained by Angiuli *et al.*(2008) whose ANN model was unable to predict soil moisture values which were out of range of the training data.

Table 5: RMSE testing when using different combinations of inputs

Combination	RMSE (% v/v)	
	17th Nov	22nd Nov
TbH+TbV	4.93	8.85
TbH+TbV+Ts	12.41	30.58
TbH+TbV+Ts+VWC	6.43	12.52
TbH+TbV+Ts+VWC+NDVI	10.34	10.13
All Nine Inputs	9.21	12.50

Experiment 2: Combination of Roscommon, Stanley and Cullingral Farms

Experiment 1 considered a single farm whereby the values of NDVI, RMS roughness value, %Clay, %Silt and %Sand were constant throughout the whole farm. To investigate if this is a result of relatively uniform parameters across this farm, a combination of farms was used. The ANN was trained using data for the first three dates of each of three farms (Roscommon: 1st, 8th & 15th Nov., Stanley: 3rd, 10th & 17th Nov., Cullingral: 4th, 9th & 18th Nov.). A total of 108 samples were obtained and of theses, six samples were randomly selected for each of the validation and testing samples. The same process of input analysis used in Experiment 1 was used. The correlation of the incremental contributions when all nine inputs were used is shown in **Table 6**.

Table 6: Incremental contribution of each of the variables in Experiment 2.

Input Excluded	R^2	Incremental Contribution	RMSE (%v/v)
None	0.5210	-	7.31
I1 (TbH)	0.5044	0.0166	7.37
I2 (TbV)	0.3904	0.1306	8.01
I3 (Ts)	0.0030	0.5180	11.68
I4 (VWC)	0.5472	-0.0262	7.21
I5 (NDVI)	0.3629	0.1581	8.21
I6 (Roughness)	0.5774	-0.0564	8.62
I7 (%Clay)	0.0251	0.4959	15.60

I8 (%Silt)	0.2423	0.2787	9.82
I9 (%Sand)	0.4851	0.0359	9.08

Table 6 shows that inputs I1(TbH), I4(VWC), I6(Roughness), and I9(%Sand) are candidates for exclusion. This results in a network of {5,10,1}. It can be seen that the constant values for %Clay and %Silt do contribute to the mapping of the function using the ANN and hence is not a parameter to be excluded at this stage. After the exclusion of the candidate inputs, the ANN was retrained and results shown in **Table 7**. The inputs I5(NDVI) and I7(%Clay) are now candidates to be excluded. This results in a network of {3,10,1}.

Table 7: Incremental contributions of each of the variables.

Input Excluded	R^2	Incremental Contribution	RMSE (%V/V)
None	0.5878	-	6.76
I2 (TbV)	0.0096	0.5782	17.41
I3 (Ts)	0.1075	0.4803	16.95
I5(NDVI)	0.5333	0.0545	8.42
I7(%Clay)	0.7655	-0.1777	7.25
I8(%Silt)	0.4255	0.1623	9.44

Table 8 shows the results after further training which results in input I8 (%Silt) being considered for exclusion. The network of size {2,10,1} is again retrained and the results shown in **Table 9**. As the contribution of each of the two inputs is almost the same, no further reduction is needed. A verification of this combination of inputs compared to other combinations is shown in **Table 10**.

Table 8: Incremental contributions of each of the variables.

Input Excluded	R^2	Incremental Contribution	RMSE (%v/v)
None	0.6685	-	6.21
I2 (TbV)	0.4133	0.2552	8.57
I3 (Ts)	0.1803	0.4882	17.47
I8(%Silt)	0.8268	-0.1583	10.86

Table 9: Incremental contributions of each of the variables.

Input Excluded	R^2	Incremental Contribution	RMSE (%v/v)
None	0.6039	-	6.56
I2 (TbV)	0.3411	0.2628	8.86
I3 (Ts)	0.1826	0.4213	18.08

Table 10: Accuracy for different combinations of input.

Combination	RMSE (% v/v)		
	Roscommon (22/11)	Stanley (24/11)	Cullingral (25/11)
TbV+Ts	1.77	9.11	5.86
TbH+TbV	7.56	10.06	5.09
All Nine Inputs	2.06	6.90	7.17

Table 10 shows the accuracies determined for the best two inputs from Experiment 2, the best two inputs from Experiment 1, and all nine inputs when evaluated on data from different dates to those used for training. It can be seen that the TbH was determined to be a good input for a single farm and for multiple farms. Note that TbV gave the best result in combination with TbH in experiment 1 whereas Ts gave the best result in combination with TbV in experiment 2. For Experiment 2, the results show that using all nine inputs or just the best ones determined through the incremental analysis show variation in the results. However the results do show that analysis of the incremental contributions of variables helps to reduce the number of inputs needed resulting in less complex ANNs and the need for less inputs.

CONCLUSIONS

This paper has analyzed the use of brightness temperature and ancillary data for soil moisture retrieval. It is important to analyze the ANN for different combinations of inputs to determine those that improve the accuracy of soil moisture measurement, and those that, could, reduce the performance by essentially confusing the ANN. **Table 5** shows that, for one farm, the use of ancillary data reduces the accuracy and hence appears not to be beneficial. **Table 10** shows that more inconclusive results for the usefulness of ancillary data occur for a more representative training and testing set of data i.e. over more farms. Through the analysis of the incremental contribution of each input, the number of inputs needed can be reduced. Although this paper shows that incremental contribution of a variable can be used for parameter selection for the ANN model, the validity of the ancillary data should be further investigated.

ACKNOWLEDGEMENTS

The National Airborne Field Experiments have been made possible through recent infrastructure (LE0453434 and LE0560930) and research (DP0557543) funding from the Australia Research Council, and the collaboration of a large number of scientists from throughout Australia, United States and Europe. Special thank also to Malaysia Higher Ministry of Education for the scholarship of the PhD study.

REFERENCES

Angiuli, E., F. del Frate, et al. (2008). Application of Neural Networks to Soil Moisture Retrievals from L-Band Radiometric Data. IEEE International Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large Area Hydrologic Modeling And Assessment Part I: Model Development. *Journal of the American Water Resources Association* 34 (1): 73-89.

Atluri, V., H. Chih-Cheng, et al. (1999). An artificial neural network for classifying and predicting soil moisture and temperature using Levenberg-Marquardt algorithm. IEEE Southeastcon '99. Proceedings.

Campolo, M., A. Soldati, et al. (2003). "Artificial neural network approach to flood forecasting in the River Arno." Hydrological Sciences Journal 48(3): 381-398.

Elshorbagy, A. and K. Parasuraman (2008). "On the relevance of using artificial neural networks for estimating soil moisture content." Journal of Hydrology 362(1-2): 1-18.

Fung, A. K., Z. Li, et al. (1992). "Backscattering from a randomly rough dielectric surface." IEEE Transactions on Geoscience and Remote Sensing 30(2): 356-369.

Junlei, S., W. Dianhong, et al. (2008). Soil Moisture Prediction with Feature Selection Using a Neural Network. Proceedings of the 2008 Digital Image Computing: Techniques and Applications, IEEE Computer Society.

Kaashoek, J. F. and H. K. v. Dijk (2002). "Neural network pruning applied to real exchange rate analysis." Journal of Forecasting 21(8): 559-577.

Panciera, R., J. P. Walker, et al. (2008). "The NAFE'05/CoSMOS Data Set: Toward SMOS Soil Moisture Retrieval, Downscaling, and Assimilation." IEEE Transactions on Geoscience and Remote Sensing 46(3): 736-745.

Saxton, K. E. and P. H. Willey (2006). The SPAW model for agricultural field and pond hydrologic simulation. Watershed Models. V. P. Singh and D. K. Frevert. Florida USA, CRC Press: 401–435.

Shou-Fang, L., L. Yuei-An, et al. (2002). "Retrieval of crop biomass and soil moisture from measured 1.4 and 10.65 GHz brightness temperatures." IEEE Transactions on Geoscience and Remote Sensing 40(6): 1260-1268.

Teuling, A. J., R. Uijlenhoet, et al. (2006). "Estimating spatial mean root-zone soil moisture from point-scale observations." Hydrol. Earth Syst. Sci. 10(5): 755-767.

Wu, S. T. S. (1996). Microwave remote sensing of land surfaces soil moisture at Global Hydrology and Climate Center. Geoscience and Remote Sensing Symposium, 1996 (IGARSS '96).