

Evaluation of three semi-empirical soil moisture estimation models with RADARSAT-2 imagery

by

Benjamin Deschamps

B.A. (Honours), Carleton University, 2008

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Geography

Carleton University
Ottawa, Ontario, Canada

© 2010, Benjamin Deschamps



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-71578-9
Our file *Notre référence*
ISBN: 978-0-494-71578-9

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

Performance of the Oh et al. (1992), Oh (2004) and Dubois et al. (1995a) soil moisture estimation models was evaluated with three RADARSAT-2 scenes over the Casselman, Ontario area. Root mean square error (RMSE), mean bias error (MBE), correlation coefficients and the number of valid pixels were used to assess the influence of boxcar speckle filter sizes and pre-extraction averaging filtering on the estimates. Highest agreement with the field data was obtained with speckle filter and pre-extraction filter sizes from 3 x 3 to 7 x 7. Bias of the estimates was found to be highly negatively correlated with incidence angle ($r^2 = 0.40, 0.15, 0.51$). Correlations between the estimated and observed soil moisture contents were not significant for most dates and the correlations with soil properties data were also weak. The effects of tillage method on the relative error of estimates were significant with lower relative errors in low-roughness tillage methods.

Acknowledgements

I would like to offer sincere thanks to my thesis co-supervisors, Dr. Scott Michell and Dr. Andrew Davidson for their constructive input and continued support since my arrival at Carleton University. Thanks also to Dr. Heather McNairn for sharing her expertise and arranging access to Agriculture and Agri-Food Canada data and facilities. Additional thanks to Dr. Amine Merzouki, Catherine Champagne and Anna Pacheco for their assistance throughout the process. At the Department of Geography and Environmental Studies, many thanks to Natalie Pressburger. I would also like to offer heartfelt thanks to my parents and sister for their continued encouragement and interest in my studies.

Finally, the completion of this thesis would not have been possible without the support and encouragement of my wife, Andrée-Anne, whose discipline, enthusiasm, wit and candour I admire, and for having inspired to me to pursue this path.

Contents

Abstract	ii
Acknowledgements	iii
List of Tables	vi
List of Figures	vii
List of Symbols	viii
List of Acronyms	ix
1 Introduction	1
2 Objectives and hypotheses	2
3 Theory	4
3.1 Radar remote sensing	4
3.2 Parameters influencing soil moisture estimation	5
3.2.1 Sensor parameters	5
3.2.2 Target parameters	5
3.3 Soil moisture retrieval models	8
3.3.1 Empirical models	9
3.3.2 Semi-empirical models	10
3.3.3 Physical models	16
4 Study area	16
5 Data and field methodology	19
6 Methods	23
6.1 Pre-processing	23
6.1.1 Ground control point collection	23
6.1.2 Filtering, resampling and orthorectification	25
6.1.3 Incidence angle map	26
6.1.4 Backscatter calibration	26
6.2 Application of the soil moisture models	28
6.3 Extraction of estimates at sample sites and pre-extraction averaging .	29
6.4 Determination of optimal speckle filter size, pre-extraction averaging and field sampling depth	29
6.5 Production of soil moisture maps at “optimal” filter combination . . .	32
6.6 Analysis of trends in relative error of estimates	33
6.6.1 Tillage methods	33
6.6.2 Soil properties	33

7	Results	34
7.1	Effects of speckle filter size, pre-extraction averaging and measurement depth	34
7.1.1	Number of sampling sites with non-zero values	34
7.1.2	Model error (RMSE)	36
7.1.3	Model bias (MBE)	39
7.1.4	Correlation between model estimates and field measurements .	41
7.1.5	Summary of model statistics and optimal filter sizes	43
7.2	Soil moisture maps at “optimal” filter sizes	44
7.3	Trends in relative error of estimates	49
7.3.1	Influence of tillage method	49
7.3.2	Influence of selected soil properties	54
7.4	Relation between model bias and incidence angle	54
7.5	Influence of actual soil moisture conditions	56
8	Discussion	57
8.1	Model implementation considerations	57
8.2	Optimal speckle filter size, pre-extraction averaging and field sampling depth	61
8.3	Relation between model bias and incidence angle	63
8.4	Trends in relative error of estimates	64
8.5	Limitations	66
9	Conclusions	68
	References	70
	A Script for generating model estimates	76
	B Script for extracting values, averaging and statistics	81
	C Script for pre-extraction averaging of entire map	85

List of Tables

1	Ranges of validity for OH92, OH04 and DU95 models	10
2	Summary of imagery characteristics	20
3	Descriptive statistics for observed soil moisture	23
4	Registration error statistics	24
5	Summary of model statistics and “optimal” filter combinations	45
6	Linear regression of incidence angle and model bias at the field scale.	55

List of Figures

1	Map of study sites	17
2	Daily mean temperature and precipitation for study area	19
3	Example of field sampling design	22
4	Box-and-whisker plot of soil moisture at two depths for the three dates	24
5	Orthorectified imagery from May 5, 16 and 23 2008.	27
6	Number of sites with non-zero pixels for a range of filter sizes	35
7	Root mean squared error (RMSE) for a range of filter sizes	37
8	Mean bias error (MBE) for a range of filter sizes	40
9	Correlation coefficient for a range of filter sizes	42
10	OH92 maps of estimated volumetric soil moisture	46
11	OH04 maps of estimated volumetric soil moisture	47
12	DU95 maps of estimated volumetric soil moisture	48
13	Mean relative errors for the three models and various tillage methods.	53
14	Relation between incidence angle and model bias at the field scale. . .	55
15	Relation between relative error of estimates and measured volumetric soil moisture	56
16	Comparison of models used to convert from dielectric permittivity to volumetric soil moisture.	59

List of Symbols

Symbol	Definition
f	Central radar frequency
k	Wavenumber ($k = 2\pi f/c$ where c is the speed of light)
l	Correlation length
m_v	Volumetric soil moisture
p	Co-polarized ratio
q	Cross-polarized ratio
R	General symbol for roughness parameters
s	RMS height (roughness parameter)
S	Scattering matrix (elements denoted as S_{hh} , S_{hv} , etc.)
Γ	Fresnel reflectivity (Γ_0 is Fresnel reflectivity at nadir)
δ_p	Penetration depth
ε or ε_r	Relative dielectric permittivity (real and imaginary parts denoted by ε'_r and ε''_r respectively)
θ or θ_i	Incidence angle, radians or degrees depending on context
λ	Wavelength
σ^0	Sigma nought, radar backscatter. Subscript indicates either contribution (e.g. σ_s^0 for soil), units (e.g. σ_{dB}^0), polarization (e.g. σ_{hh}^0 or time (e.g. σ_{t1}^0))
τ	Canopy attenuation factor

List of Acronyms

Acronym	Meaning
AAFC	Agriculture and Agri-Food Canada
ANOVA	Analysis of variance
DEM	Digital elevation model
DOY	Day of year
DU95	Dubois et al. (1995a) model
MANOVA	Multiple analysis of variance
MBE	Mean bias error
NDVI	Normalized difference vegetation index
OH92	Oh et al. (1992) model
OH04	Oh (2004) model
RGB	Red - Green - Blue composite
RMSE	Root mean squared error
RS-2	Radarsat-2 imagery or satellite / sensor
SAR	Synthetic aperture radar
SLC	Single Look Complex (RS-2 mode)
UTM	Universal Transverse Mercator (coordinate system)

1 Introduction

Accurate and precise spatial estimates of soil moisture are required for many hydrological applications. In agriculture, soil water content is a limiting factor for plant growth, and estimates are useful in irrigation and soil management (Verhoest et al., 2008). Soil moisture is also a key variable in hydrological, drought and crop yield models where it acts as the link between surface and sub-surface hydrological processes, controlling infiltration and runoff rates (Engman et al., 1996). Traditionally, such models have used average soil moisture values for large areas. However, the use of spatially distributed (per pixel) soil moisture estimates could improve the accuracy of these models considerably (Moran et al., 2004).

Due to the high local spatial variability of soil moisture, even within flat homogeneous land units, a large number of field samples are required to characterize its distribution adequately. This logistical constraint can be addressed using remotely sensed imagery, which allows for broader and repeatable coverage of any area at fine resolutions. Microwave (radar) remote sensing is of particular use for hydrological applications because the radar return is inherently sensitive to changes in surface soil moisture content.

Previous studies have identified site-specific empirical relations between radar backscatter and soil moisture, but these require re-calibration for extrapolation to new areas or in subsequent years (Verhoest et al., 2008). In many cases, accurate estimates could only be obtained at the field or watershed scale, primarily due to the occurrence of speckle in the imagery, and to problems related to the inadequate characterization of surface roughness or vegetation cover. Speckle is caused by constructive or destructive interference in interacting microwaves, and results in a “speckly” or “salt-and-pepper” type effect which must be corrected to obtain correct backscatter

values.

Semi-empirical models such as the Oh et al. (1992, 1994); Oh (2004) and Dubois et al. (1995a,b)¹ models have been developed using large databases of radar imagery with wide ranges of frequency, roughness characteristics and incidence angles. The performance of such models has been widely evaluated in the literature, with conflicting results. While estimates of soil moisture are relatively simple to produce using these models, little is known about the error and bias of estimates, the influence of processing on the estimates, or which factors (such as soil properties) affect the confidence of those estimates.

2 Objectives and hypotheses

The purpose of this research project is to evaluate the performance of the OH92, OH04 and DU95 semi-empirical soil-moisture models with RADARSAT-2 imagery in an Eastern Ontario agricultural area. More specifically, the objectives are to:

1. Evaluate the influence of speckle filter selection and window size, the influence of pre-extraction averaging and the influence of field sampling depth on model error, bias and correlation with the field data. It was hypothesized that:
 - (a) Root mean square error (RMSE) and mean bias error (MBE) would gradually decrease with larger speckle filter sizes (higher speckle suppression), then increase again when filtering is excessive;
 - (b) Error, bias and correlation would be better when comparing with the 3 cm field data than the 6 cm data (field soil moisture measurements were collected at two depths, see Section 5) due to the low radar penetration

¹These models will herein be referred to as the OH92, OH04 and DU95 models

depth; agreement may be better with the 6 cm data on drier dates (increased penetration depth), and better with the 3 cm data on wetter dates (decreased penetration);

- (c) The models would be positively correlated with the field data; ideal model performance would have a perfect correlation;
- (d) Pre-extraction averaging would further decrease RMSE and MBE, perhaps increase correlation and would increase the number of pixels with non-zero values. As with speckle filtering, these improvements should not be effective at excessively large averaging size.

2. Identify and quantify trends between the relative error of estimates and soil properties. It was hypothesized that the relative error of soil moisture estimates would be significantly related to:

- (a) Tillage method (i.e. conventional, minimal, no-till)
- (b) Soil composition (i.e. sand or clay fractions)
- (c) Soil order (i.e. gleysolic, etc.)
- (d) Amount of organic matter
- (e) Drainage characteristics (perfect vs. imperfect)

3. Examine other trends such as relations with incidence angle and “true” soil moisture conditions. It was hypothesized that:

- (a) Steeper incidence angles would lessen error and bias and increase correlation with field data by reducing the effects of roughness;
- (b) Lower “true” soil moisture conditions would increase the error of estimates due to increased penetration depth.

3 Theory

3.1 Radar remote sensing

Active microwave remote sensing sensors operate by emitting pulses of energy in a side-looking configuration. The energy interacts with the surface and is reflected back towards the sensor. This reflected energy is known as backscatter, and consists of the desired signal plus speckle, a form of noise due to constructive and destructive interference in the reflected signal. The result is a backscatter image showing the intensity of the returned signal.

This study will make use of RADARSAT-2 (RS-2) imagery. RS-2 is an active sensor and as such does not require solar illumination. Because it operates at longer wavelengths than optical sensors, RS-2 can image through clouds, rain, fog and haze, and is not highly susceptible to atmospheric effects. RS-2 operates in the portion of the electromagnetic spectrum designated as C-band; it has a wavelength of 0.055 m and operates at a frequency of 5.405 GHz. The sensor is fully polarimetric (captures both amplitude and relative phase of returned signal in all polarizations: HH, VV, HV, VH, where the first letter represents the transmitted horizontal (H) or vertical (V) polarization and the second the received polarization).

Both sensor and target parameters influence backscatter. Understanding these factors is important because estimates of soil moisture are directly related to them. The most significant sensor parameters influencing the radar backscatter are the incidence angle, wavelength and polarization of the sensor, while important target parameters include surface roughness, the dielectric properties of the surface, and vegetation cover.

3.2 Parameters influencing soil moisture estimation

3.2.1 Sensor parameters

Incidence angle and sensor wavelength affect penetration depth and perceived surface roughness. Longer wavelengths and steeper incidence angles (the angle between the radar beam and the vertical where the beam meets the surface) usually lead to increased penetration depth, and incidence angles of 10 to 20 degrees are recommended for soil moisture mapping (Ulaby et al., 1982). Backscatter varies according to the polarization configuration, primarily due to the depolarizing effect of vegetation. Soil moisture is strongly correlated with backscatter in the HH and VV polarization configurations (Sokol et al., 2004).

3.2.2 Target parameters

The adequate characterization of surface roughness is the largest obstacle to the accurate retrieval of soil moisture information from radar backscatter, as the sensitivity of backscatter to roughness is greater than that of soil moisture (Álvarez-Mozos et al., 2009). The high spatial variability of surface roughness makes it difficult to measure accurately across large areas (Verhoest et al., 2008). Roughness is also linked to soil type, tillage practises and rainfall events. Though the effects of surface roughness are eliminated when considering relative changes in soil moisture estimates, they need to be considered when attempting to estimate absolute values in empirical, semi-empirical and physical models. Roughness is most frequently characterized by three parameters: RMS height, correlation length and an autocorrelation function. These parameters are derived from soil profiles obtained with pin or laser profilometers. A pin profilometer consists of a horizontal bar (usually 1 meter in length) holding long small diameter, closely-spaced metal pins that move vertically to conform to the

shape of the ground below and “replicate” the surface height along a linear transect. Statistical measures are then derived from a digitized photograph of the profilometer reading.

The RMS height (s) is analogous to the standard deviation of soil heights about the mean height and is given by:

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2} \quad (1)$$

where n is the number of height measurements along a transect, z_i is a height measurement and \bar{z} is the mean height. According to Bryant et al. (2007), accurate characterization of RMS height requires a minimum of 20 transects of 3 m in length. In the case of laser profilometers, thousands of points are recorded. With pin profilometers, the number of points corresponds to the number of pins on the instrument.

A discrete autocorrelation function, showing the relation between height at points x and x' distant from x can be derived from a one-dimensional linear transect. For lags of $x' = (j - 1)\Delta x$ where j is an integer ≥ 1 representing the discrete step and Δx the discrete interval:

$$\rho(x') = \frac{\sum_{i=1}^{N+1-j} z_i z_{j+i+1}}{\sum_{i=1}^N z_j^2} \quad (2)$$

where N is the number of discrete steps along the transect. The correlation length (l) corresponds to the horizontal distance such that $\rho(x') = 1/e$; height measurements taken at separations greater than the correlation length are considered statistically independent (Álvarez-Mozos et al., 2005)

Finally, ks and kl parameters are used as wavelength-dependent representations of soil roughness where k represents the wavenumber (a function of the radar frequency,

with $k = 2\pi f/c$ where c is the speed of light). Zribi and Dechambre (2003) introduced a new parameter named Z-index, which combines the effects of both RMS height and correlation length. The Z-index is given as: $Z = s^2/l$.

The dielectric properties of soils are closely related to their moisture content and are the second most important factor controlling backscatter. The broad term *dielectric properties* describes the ability of a wave to propagate through a material. Mathematically, it is represented by the value of the complex relative dielectric permittivity (ϵ_r) given as $\epsilon_r = \epsilon'_r + \epsilon''_r$ where ϵ'_r is the dielectric constant and ϵ''_r is a measure of wave attenuation in the medium. Values for the dielectric constant typically range from 0 to 80, with water having a value approaching 80. Thus, the value of the dielectric constant is highly influenced by water content resulting in higher values in soils with higher moisture contents. Volumetric soil moisture (m_v) is related to the dielectric constant by polynomial models such as the widely used Topp et al. (1980) model, which holds for a wide range of soil types:

$$m_v = -5.3 + 2.92(\epsilon'_r) + 0.055(\epsilon'_r)^2 + 0.0043(\epsilon'_r)^3 \quad (3)$$

Soil moisture probes operating on the principles of time domain reflectometry measure permittivity directly with an accuracy of $\pm 1\%$, to which is added an error of 3-4% in the conversion from the dielectric constant to volumetric soil moisture (Topp et al., 1980). This relation is accurate for volumetric soil moisture contents of less than 40-55%, independently of the wavelength of incoming energy. Beyond this range, the model can under-predict soil moisture by 10% or more and becomes somewhat frequency dependent (Curtis, 2001; Topp et al., 1980).

Penetration depth (δ_p) of the radar signal is also affected by the dielectric properties of soils, and can be approximated by:

$$\delta_p \approx \frac{1}{k} \cdot \frac{\sqrt{\epsilon_r''}}{\epsilon_r''} \quad (4)$$

where k is the wavenumber. For increasing soil moisture, the real portion of the complex permittivity increases faster than the imaginary portion and is always greater (Ulaby et al., 1982). C-band penetration in very dry sandy soils can approach depths of 0.52 m (Schaber and Breed, 1999) while penetration in most soils suitable for cultivation is limited to between 0.01 and 0.05 m (Bruckler et al., 1988), signifying that the radar return is only representative of near-surface conditions (Dobson and Ulaby, 1986).

As will be discussed in more detail in Section 3.3.1, the presence of vegetation cover influences backscatter, and this backscatter is not linearly related to soil moisture due to complex multi-bounce interactions between the canopy and underlying soil. In empirical and semi-empirical models, the influence of vegetation is removed using scattering models that take into account the distribution of water in the plant canopy and the geometry of the canopy or using empirical relations when the canopy is composed of a single species (Bindlish and Barros, 2001).

3.3 Soil moisture retrieval models

Three broad categories of soil moisture models exist for the retrieval of soil moisture estimates from radar backscatter: empirical, semi-empirical and physical models (Verhoest et al., 2008). This section presents an overview of the three semi-empirical models evaluated in this study. Empirical models are also presented briefly as the precursors of semi-empirical models, and physical models are mentioned as an alternative to them.

3.3.1 Empirical models

Empirical models establish regression equations between backscatter and volumetric soil moisture as collected in field sampling. For bare or barely vegetated soils, backscatter is a function of RMS height (roughness) and volumetric soil moisture: $\sigma_s^0 = f(R, m_v)$ where σ_s^0 is the fraction of backscatter attributed to soil returns, R is the soil roughness and m_v is soil moisture². Backscatter is linked to soil moisture by $m_v = a + b\sigma_s^0$ where a and b are regression coefficients that are sensor, land use and soil type dependent (Moran et al., 2004). Shoshany et al. (2000) report that the intercept coefficient a ranged from -12.33 and -14.89 dB in multiple studies, while the slope coefficient b was more variable, ranging from 0.2 to 0.4. The Baghdadi and Zribi (2006) formulation, which separates contributions from soil moisture ($f()$) and from roughness effects ($g()$, introduced through the s or ks term) is given as

$$\sigma_{dB}^0 = f(m_v, \theta, mn, \lambda)_{dB} + g(s, \theta, mn, \lambda) = am_v + be^{-ks} + c \quad (5)$$

where mn is the polarization configuration (HH, HV or VH), and a , b , and c are calibration coefficients that account for the influences of frequency, incidence angle and polarization. The notion of the exponential ks term will also be of importance in most semi-empirical models.

For vegetated surfaces, backscatter contributions from the vegetation canopy and from ground-canopy interactions are added (Engman et al., 1996): $\sigma^0 = \tau\sigma_s^0 + \sigma_{dv}^0 + \sigma_{int}^0$ where τ is a factor representing the attenuation of incoming energy by the vegetation canopy, σ_{dv}^0 is the directly backscattered energy from the vegetation canopy, and σ_{int}^0 is the contribution to backscatter from multi-bounce scattering between the ground

²Note that the volumetric soil moisture will be represented using the term m_v , and not by the traditionally used letter θ to avoid confusion with incidence angle which is represented by θ or θ_i .

Table 1: Ranges of validity for OH92, OH04 and DU95 models

Model	Source	m_v (%)	ks (cm)	θ_i (degrees)	ϵ_r
OH92	Oh et al. (1992)	9 - 31	0.1 - 6.0	10 - 70	-
OH04	Oh (2004)	4 - 29.1	0.13 - 6.98	10 - 70	-
DU95	Dubois et al. (1995a,b)	0 - 35	$< 2\pi/3$	30 - 60	0 - 20

and the vegetation canopy (Ulaby et al., 1996). Thus, removing the influence of the vegetation canopy yields a direct relation to the backscatter due to the soil layer (and thus to soil moisture). The influence of the vegetation canopy can be removed through the use of physical scattering models such as the water cloud model that simulates a vegetation canopy composed of a cloud of individual water droplets (Ulaby et al., 1978; Bindlish and Barros, 2001). In applications with bare soils or barely vegetated surfaces, the influences of vegetation can be ignored.

3.3.2 Semi-empirical models

Semi-empirical models establish relations between backscatter (or backscatter ratios) and relative dielectric permittivity. The most widely used semi-empirical models are the Oh et al. (1992), Oh (2004) and Dubois et al. (1995a,b) models (Baghdadi and Zribi, 2006). Based on the data they are derived from and on the model's formulation, ranges of validity can be defined for the three models (Table 1). All three models are valid through L, C and X bands.

The DU95 and OH92 models, which both rely on the polynomial relation described by Topp et al. (1980) could theoretically be extended to 40-45% volumetric soil moisture; as stated in Section 3.2.2. However, the behaviour in terms of error and uncertainty is not known for those soil moisture levels.

Oh (1992) model The OH92 model was developed using ground scatterometer, laser profilometer and relative dielectric permittivity measurements, and is valid for

incidence angles between 10 and 70 degrees, soil roughnesses of $0.1 < ks < 6.0$, $2.6 < kl < 19.7$ and $0.09 < m_v < 0.31$ (Oh et al., 1992). The model derives empirical relations between the co-polarized (p) and cross-polarized (q) ratios and ks :

$$\sqrt{p} = \sqrt{\frac{\sigma_{hh}^0}{\sigma_{vv}^0}} = 1 - \left[\left(\frac{2\theta}{\pi} \right)^{1/3\Gamma_0} \cdot e^{-ks} \right] \quad (6)$$

$$q = \frac{\sigma_{hv}^0}{\sigma_{vv}^0} = 0.23\sqrt{\Gamma_0} [1 - e^{-ks}] \quad (7)$$

where Γ_0 is the Fresnel reflectivity of the surface at nadir, given as:

$$\Gamma_0 = \left| \frac{1 - \sqrt{\varepsilon_r}}{1 + \sqrt{\varepsilon_r}} \right|^2 \quad (8)$$

and θ is the incidence angle, k is the wavenumber, s is the surface RMS height, and ε_r is the relative dielectric permittivity.

Given the nature of Equations 6 and 7, the estimates of the co-polarized ratio p are always lower than 1, and increase with increasing ks , converging towards 1 for $ks > 3$. For $ks < 3$, the co-polarized ratio decreases with increased theta (shallower incidence angles) and/or higher soil moisture. Estimates of the cross-polarized ratio q are always much lower than 1 because of a higher sensitivity to the surface roughness term and lessened influence of soil moisture. The cross-polarized ratio increases with increasing ks , converging towards a value depending on soil moisture. As reflected by Equation 7, the cross-polarized ratio is insensitive to incidence angle.

When trying to derive soil moisture or surface roughness from this model using remotely sensed imagery, σ_{hh}^0 , σ_{hv}^0 , σ_{vv}^0 , θ and λ are known. We can rearrange Equation 6 and Equation 7 for e^{-ks} to eliminate that term, equating the two remaining terms to obtain the following nonlinear equation:

$$\left(\frac{2\theta}{\pi}\right)^{1/3\Gamma_0} \cdot \left[1 - \frac{q}{0.23\sqrt{\Gamma_0}}\right] + \sqrt{p} - 1 = 0 \quad (9)$$

where θ is expressed in radians. The above equation can be solved for Γ_0 iteratively (using the Newton-Raphson method, for example). Finally, the real portion of the relative dielectric permittivity can be calculated using Equation 8 under the assumption that the imaginary portion is negligible for natural soil materials (Oh et al., 1992) and volumetric soil moisture can be deduced using the polynomial relations given by Hallikainen et al. (1985) or Topp et al. (1980). A modified formulation of the q term that included the influence of incidence angle was proposed by Oh et al. (1994); additional modifications were made by Oh et al. (2002) to include correlation length, but this modification was undone in OH04 after Baghdadi et al. (2000) showed that correlation length was not an accurate representation of roughness conditions (these intermediary versions of the model are not considered here). The formulation of the OH04 model is sufficiently different both in its formulation and implementation that it was included separately in the analysis.

Oh (2004) model The OH04 model is expressed as a function of the co- and cross-polarized ratios, with the addition of the cross-polarized backscatter coefficient. It is given by:

$$\sigma_{vh}^0 = 0.11m_v^{0.7}(\cos \theta)^{2.2} [1 - \exp(-0.32(ks)^8)] \quad (10)$$

$$p = \frac{\sigma_{hh}^0}{\sigma_{hh}^0} = 1 - \left(\frac{\theta}{90^\circ}\right)^{0.35m_v^{-0.65}} \cdot \exp(-0.4(ks)^{1.4}) \quad (11)$$

$$q = \frac{\sigma_{vh}^0}{\sigma_{vv}^0} = 0.095(0.13 + \sin(1.5\theta)^{1.4} \{1 - \exp[-1.3(ks)^{0.9}]\}) \quad (12)$$

where p and q are the co- and cross-polarized ratios, σ_{ij}^0 represents the backscatter coefficient (emitted polarization i and received polarization j), m_v is the volumetric soil moisture, θ is the incidence angle obtained for each pixel from an incidence angle map, $k = 2\pi f/c$ is the wavenumber (f is the radar central frequency and c the speed of light), and s is the soil RMS height, a measure of surface roughness. The ks term gives a frequency-specific representation of the surface roughness. Prior to generating estimates for a pixel, the maximum allowable value of p given the local incidence angle is calculated using Equation (11) with $s = 0.055$ m and $m_v = 0.01\%$. Pixels with $p > p_{max}$ are excluded *a priori*.

The model is inverted to retrieve an estimate of soil moisture and RMS height. First, Equation(10) is rearranged for ks :

$$ks = \left[-3.125 \ln \left(\frac{\sigma_{vh}^0}{0.11m_v^{0.7}(\cos\theta)^{2.2}} \right) \right]^{5/9} \quad (13)$$

The substitution of Equation (13) into Equation (11) gives a non-linear equation for m_v which can be solved using iterative numerical methods such as Newton's method, using an initial value of $m_v = 0.5$:

$$1 - \left(\frac{\theta}{90^\circ} \right)^{0.35m_v^{-0.65}} \cdot \exp \left(-0.4 \left\{ \left[-3.125 \ln \left(\frac{\sigma_{vh}^0}{0.11m_v^{0.7}(\cos\theta)^{2.2}} \right) \right]^{5/9} \right\}^{1.4} \right) - p = 0 \quad (14)$$

A first estimate of s can also be obtained directly from Equation (13) by substituting the first estimate of m_v . A second estimate of s can be derived from Equation (12):

$$ks = \left[\ln \left(1 - \frac{q}{0.095(0.13 \sin(1.5\theta))^{1.4}} \right) / -1.3 \right]^{10/9} \quad (15)$$

Using this second estimate of s , two more estimates of m_v can be derived from Equations (10) and (11):

$$m_v = \left[\frac{\sigma_{vh}^0}{0.11(\cos \theta)^{2.2} \cdot [1 - \exp(-0.32(ks)^{1.8})]} \right]^{10/7} \quad (16)$$

$$m_v = \left[\ln \left(\frac{1 - p}{\exp(-0.4(ks)^{1.4}/0.35 \ln \theta)} \right) \right]^{-20/13} \quad (17)$$

This process results in two estimates of s and three estimates of m_v . Estimates of either parameter that are smaller than zero will be excluded from the weighted mean that determines the final estimates by being attributed a weight of zero. All positive estimates are assigned a weight of 1 and combined according to the weighted means:

$$s = \frac{s_1 * w_1 + s_2 * 0.25w_2}{w_1 + 0.25w_2} \quad (18)$$

$$m_v = \frac{m_{v1} * w_3 + m_{v2} * w_4 + m_{v3} * w_5}{w_3 + w_4 + w_5} \quad (19)$$

The final step is a verification of whether the estimates generated are within the range of validity of the model; there is a possibility that they might fall outside the range of validity because of the averaging in the previous step.

Dubois (1995) model The DU95 model is simpler than the OH92 model, using only the terms of the co-polarized ratio. It links ks directly with backscatter in the HH and VV polarizations, based on empirical fit. The model is optimized for $ks < 2.5$, incidence angles of 30 degrees or greater and volumetric soil moisture content less

than or equal to 35%. Like the OH92 model, it is valid for bare soils, but presents the additional criterion of having a normalized difference vegetation index (NDVI) of less than 0.4. Given the strong relation that exists between NDVI and the cross-polarized ratio, it is possible to mask out areas with high amounts of vegetation using that ratio. All pixels with a cross-polarized ratio exceeding -11 dB are excluded *a priori*. Dubois et al. (1995a) report RMS errors of less than 4.2% in volumetric soil moisture. The model is expressed as:

$$\sigma_{hh}^0 = 10^{-2.75} \frac{\cos^{1.5} \theta}{\sin^5 \theta} \cdot 10^{0.028 \varepsilon \tan \theta} (ks \sin \theta)^{1.4} \lambda^{0.7} \quad (20)$$

$$\sigma_{vv}^0 = 10^{-2.35} \frac{\cos^3 \theta}{\sin^3 \theta} \cdot 10^{0.046 \varepsilon \tan \theta} (ks \sin \theta)^{1.1} \lambda^{0.7} \quad (21)$$

The model is inverted by solving one of the equations for $\varepsilon \tan \theta$ and substituting it into the other. The resulting equation is then solved for ks to obtain the estimate of surface roughness:

$$ks = \frac{\sigma_{hh}^{1.3691}}{\sigma_{vv}^{0.8334}} 10^{1.7900} (\cos \theta)^{0.4465} (\sin \theta)^{3.3454} \lambda^{-0.375} \quad (22)$$

The resulting estimate of ks is then inserted into the equation for σ_{hh}^0 to retrieve the estimate of the dielectric constant:

$$\varepsilon' = \frac{\log_{10}(\sigma_{hh}^0) 10^{2.75} (\cos \theta)^{-1.5} (\sin \theta)^5 (ks \sin \theta)^{-1.4} \lambda^{-0.7}}{0.028 \tan \theta} \quad (23)$$

Finally, the estimate of the real portion of the dielectric constant is converted to volumetric soil moisture using an empirically-derived third order polynomial relation

described by Topp et al. (1980):

$$m_v = 5.3 \times 10^{-2} + 2.92 \times 10^{-2} \varepsilon' - 5.5 \times 10^{-4} \varepsilon'^2 + 4.3 \times 10^{-6} \varepsilon'^3 \quad (24)$$

Curtis (2001) evaluated the model and found it to agree well with experimental data for soil moisture contents of less than 40%.

3.3.3 Physical models

The most widely used physical model for soil moisture retrieval is the integral equation model (IEM; Fung et al., 1992; Fung and Chen, 2004). Because of its complex formulation (see Baghdadi and Zribi, 2006), it is commonly applied by running a number of forward iterations with a range of roughness and dielectric characteristics to populate lookup tables rather than inverting the model to predict soil moisture (Thoma et al., 2006). The IEM has a wider range of applicability than the Oh and Dubois models, and recent improvements have proposed a calibration of the model to fit remotely sensed data more closely (Baghdadi and Zribi, 2006). However, the model is highly sensitive to roughness, and uses correlation length, a difficult parameter to measure (Álvarez-Mozos et al., 2009). The main difficulty with physical models lies in finding closed solutions for their inversion (Pierdicca et al., 2009).

4 Study area

The study area is located northeast of the village of Casselman, Ontario, Canada (45°22'N 75°56'W; Figure 1). The area is composed of approximately equal areas of forest and agricultural land.

The crops planted in the study area included wheat, corn and soybean, under a

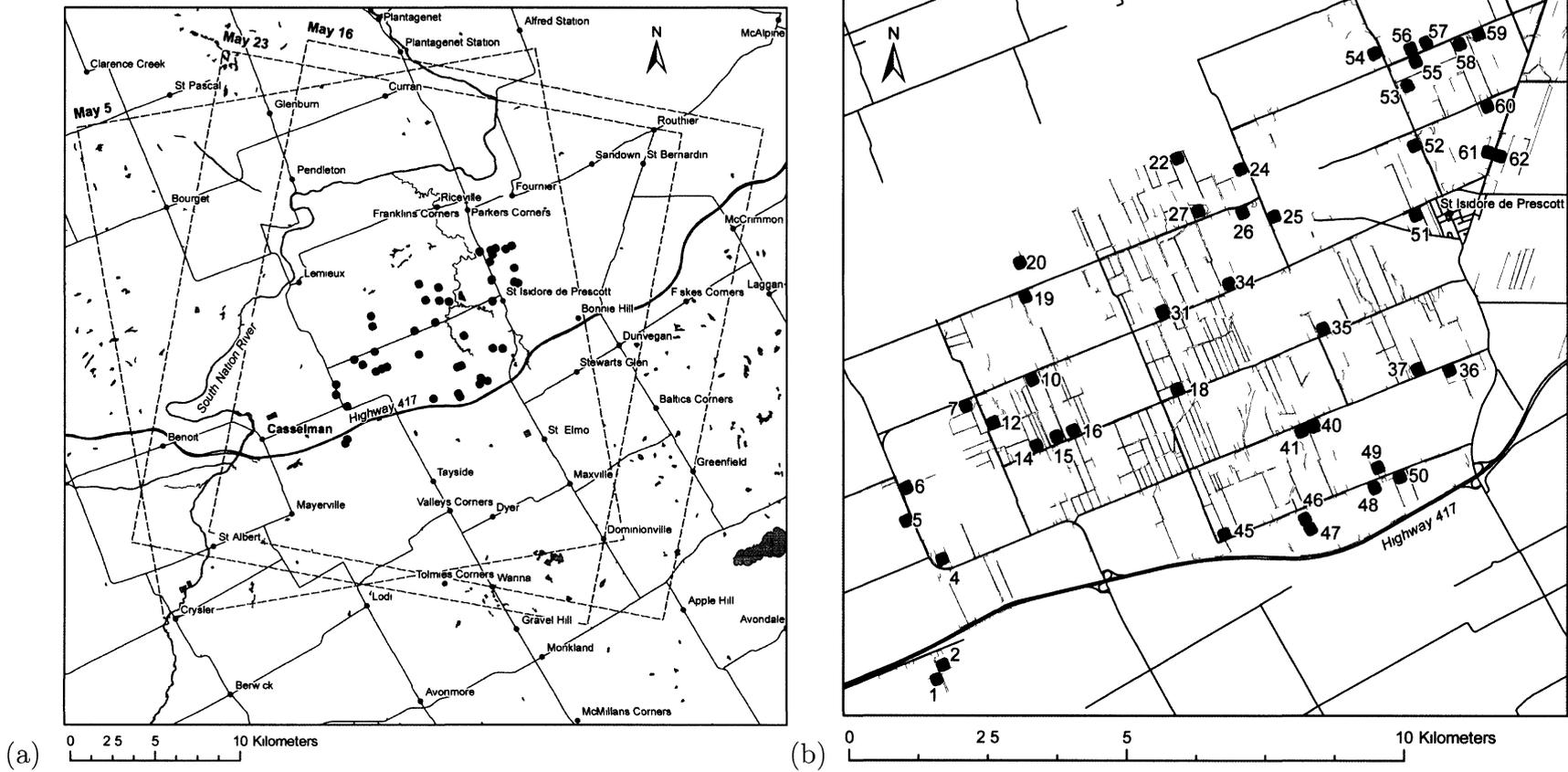


Figure 1: (a) Map of study sites (red points) showing the South Nation River, Highway 417, major roads and villages. Imagery footprints are shown in dashed orange lines. Lower-order streams and roads have been omitted for clarity. (b) Detail of field boundaries and sites. Each site contains 16 sampling points as outlined in Figure 3.

variety of tillage conditions. Most fields were managed under crop rotation and had a different crop the previous year. Only some corn fields and a few soy fields had the same crops replanted in 2008. All fields were seeded on or before May 15, 2008. Sites where wheat emergence had begun (between May 16 and May 23) were not included in the analysis of moisture in the May 23 scene. No wheat emergence was recorded on May 16. Tillage methods varied across the study area, from fall tillage and seeding (no spring tillage), to minimal or conventional spring tillage. Only two fields had no tillage operations whatsoever, and approximately a third of the fields were rolled after seeding (a roller compacts the soil to ensure good contact with the seeds).

Hydrological features include most notably the South Nation River, which drains north into the Ottawa. The entire area of study is contained within its watershed.

Soil types vary in a north-west to south-east gradient, from sandy soils in the north-west of the study area to clays more suitable to agriculture, and back to mostly sandy soils in the south-east.

According to the Russell weather station data (Environment Canada, 45°15N, 75°21W, roughly 30 km WSW of the center of the study area), the snow pack had melted completely by April 16, 2008. The area received no precipitation from April 14 until April 23. On April 23, there was 4.4 mm of rain, 13.2 mm of rain on April 26 and 1.2 mm per day on May 2nd and 3rd, before the May 5 acquisitions (Environment Canada). A weather station installed in the area for this experiment recorded 5.5 and 2.0 mm of of rain on May 6 and 7, followed by a 6 day drying period ending with traces of precipitation on May 15 and 1.5 mm precipitation before the May 16 acquisition. This was followed by a re-wetting period with 21.4 mm of precipitation accumulated between the May 16 image and the May 23 acquisition. It should be noted that the values given below are not fully representative of the conditions at each site; rather, they are given to provide context for the soil moisture conditions

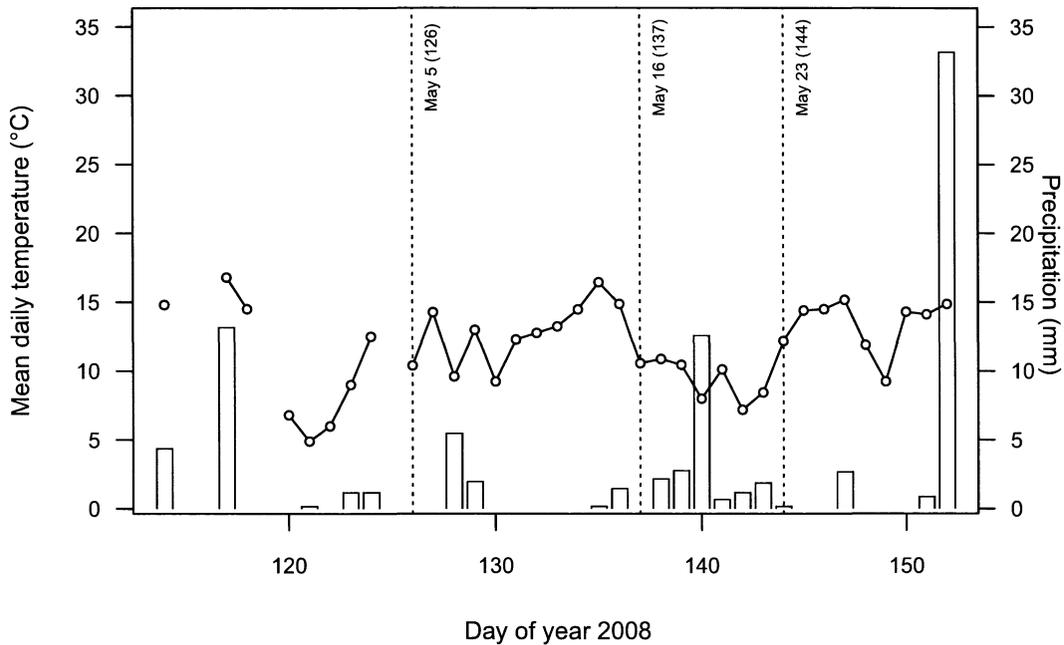


Figure 2: Daily mean temperature (line) and precipitation (bars) from April 23 (DOY = 114) to May 31 (DOY = 152). Data from April 23 to May 5 are for the village of Russell (Environment Canada) and data for the remainder of May is from an AAFC weather station located just west of the village of St-Isidore. No data were recorded for April 24, 25 and 28 and May 4. Image acquisition dates shown by dashed lines.

during the experiment. Figure 2 gives an overview of the precipitation and mean daily temperature during the experiment.

5 Data and field methodology

The data used in this study consisted of three RS-2 scenes accompanied by coincident field measurements of soil moisture and surface roughness. Weather data were also available for the duration of the experiment. Data access was arranged through a partnership with Agriculture and Agri-Food Canada (AAFC) as part of a larger

Table 2: Summary of imagery characteristics

Image Date (DOY 2008)	Incidence angle (degrees)	Orbit	Beam mode
May 5 (126)	38.3 - 39.8	Ascending	FQ19
May 16 (137)	30.2 - 32.0	Descending	FQ11
May 23 (144)	35.4 - 37.0	Descending	FQ16

project, and all the raw data were already collected and documented. All analyses for this study began with raw, unprocessed data. A brief description of the data and data collection methods is provided below; the reader can refer to McNairn et al. (2008) for more details on the data collection methodology. The data and field methodology is also described in more detail in Merzouki et al. (accepted 2010a) and Merzouki et al. (submitted 2010b).

Three single-look complex (SLC) fine quad-pol mode RS-2 scenes were acquired over the Casselman site (Table 2). The scenes cover an area of approximately 25 km x 25 km and have been re-sampled to a pixel size of 8 x 8 meters (this resampling is required because raw slant-range pixels are rectangular, the exact nominal pixel size is determined based on the beam mode and is given in the RS-2 product specification). The images were taken at low incidence angles, ensuring their suitability for soil moisture detection and maximum penetration depth. In order to acquire the most images in a short time span, the images could not be taken with uniform orbits and incidence angles.

Additionally, *in situ* soil moisture measurements were collected for 44 sites in 42 fields on each date (two larger fields had two sites) using ThetaProbe MLX-2 (Delta-T Devices, Cambridge, UK) soil moisture probes. The probes consist of three 6 cm “shield” probes in a triangular configuration with one central 6 cm “signal” probe. The probes measure the dielectric constant and effect a conversion to volumetric soil moisture using equations such as those in Topp et al. (1980). The conversion to

volumetric soil moisture is subject to an error of up to 5%. At each site, 16 sampling sub-sites were located at 30 meter intervals on a 4 x 4 square grid aligned on the magnetic north-south axis (Figure 3). At each site, the recorded soil moisture was the average of three replicates. Where deep furrows occurred, the replicates were taken at the top, middle and bottom of the furrow to fully account for moisture conditions. Even numbered points were sampled at a depth of 6 centimeters and odd-numbered points at a depth of 3 centimeters, yielding a dataset of 16 points per field at two depths on three dates for a total of 2,112 samples. To obtain the 6 cm depth measurements, the probe was inserted vertically, perpendicular to the soil surface. The 3 cm depth measurements were obtained by inserting the probe at a 30 degree angle. Soil moisture measurements were always near-coincident with the RS-2 acquisitions, occurring either the morning after a night or evening acquisition or in the day following an early morning acquisition.

Soil roughness measurements were also obtained using 1-meter pin profilometers. The profilometer was custom-fabricated for AAFC and consists of 200 0.01 cm wide pins with a spacing of 0.5 cm. Five one-meter transects were measured and photographed for subsequent automated analysis at one sample site in each field. The roughness data were collected from April 30, 2008 to May 23, 2008. Changes in field conditions were monitored and roughness measurements were repeated if necessary to ensure that the data corresponded as closely as possible to the RS-2 imagery. In some instances, specific sites were being tilled, seeded or fertilized during the RS-2 acquisition, invalidating prior soil roughness measurements. These sites/dates were excluded from the analysis. RMS height and correlation length were calculated for each site from the mean of the five measurements.

Finally, a weather station was installed to log complete meteorological and soil moisture data for the duration of the experiment. This included air temperature,

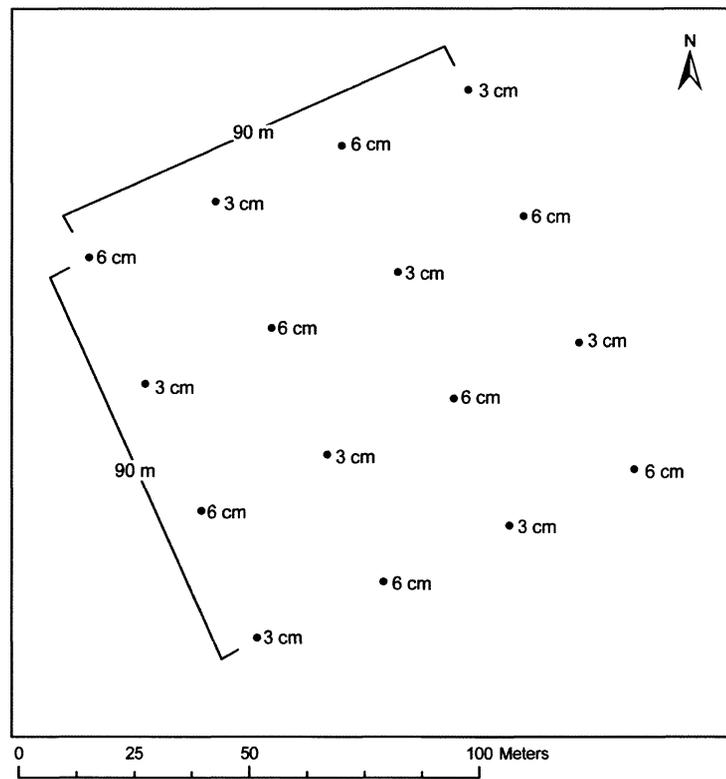


Figure 3: Sampling design for an example field. Note the alternating sampling depths and grid spacing.

Table 3: Descriptive statistics for observed soil moisture

Statistic	May 5			May 16			May 23		
	<i>3 cm</i>	<i>6 cm</i>	<i>All</i>	<i>3 cm</i>	<i>6 cm</i>	<i>All</i>	<i>3 cm</i>	<i>6 cm</i>	<i>All</i>
Mean	20.8	21.8	21.3	13.4	16.5	15.0	22.3	24.0	23.2
Std. Deviation	5.4	5.4	5.4	4.4	4.4	4.6	4.1	4.6	4.5
<i>n</i>	266	265	531	344	344	688	276	276	552

wind direction and speed, precipitation, and soil moisture at various depths. Two sub-stations were also installed at different locations to record soil temperature, soil moisture and precipitation. The soil moisture measurements made at the weather stations were invalid due to equipment malfunctions and could not be used here.

6 Methods

6.1 Pre-processing

6.1.1 Ground control point collection

Fifteen ground control points (GCPs) per image were collected manually using the raw imagery and a vector road network for eastern Ontario (CanVec v1.1, Natural Resources Canada, 2010). Any GCPs collected for a single image that were situated within the footprint of another image were also transferred to that image, acting as tie points between the images, further improving accuracy. However, some points occurring in more than one image footprint were not clearly locatable, despite occurring in the two images, and were rejected. Ground control points are not absolutely required to perform the orthorectification (because tie points are also included with the imagery), but their inclusion increases positional accuracy and allowed the calculation of RMS statistics to quantify the accuracy of the georeferencing. Table 4 shows that an overall RMS error of less than 1 pixel was achieved for each scene.

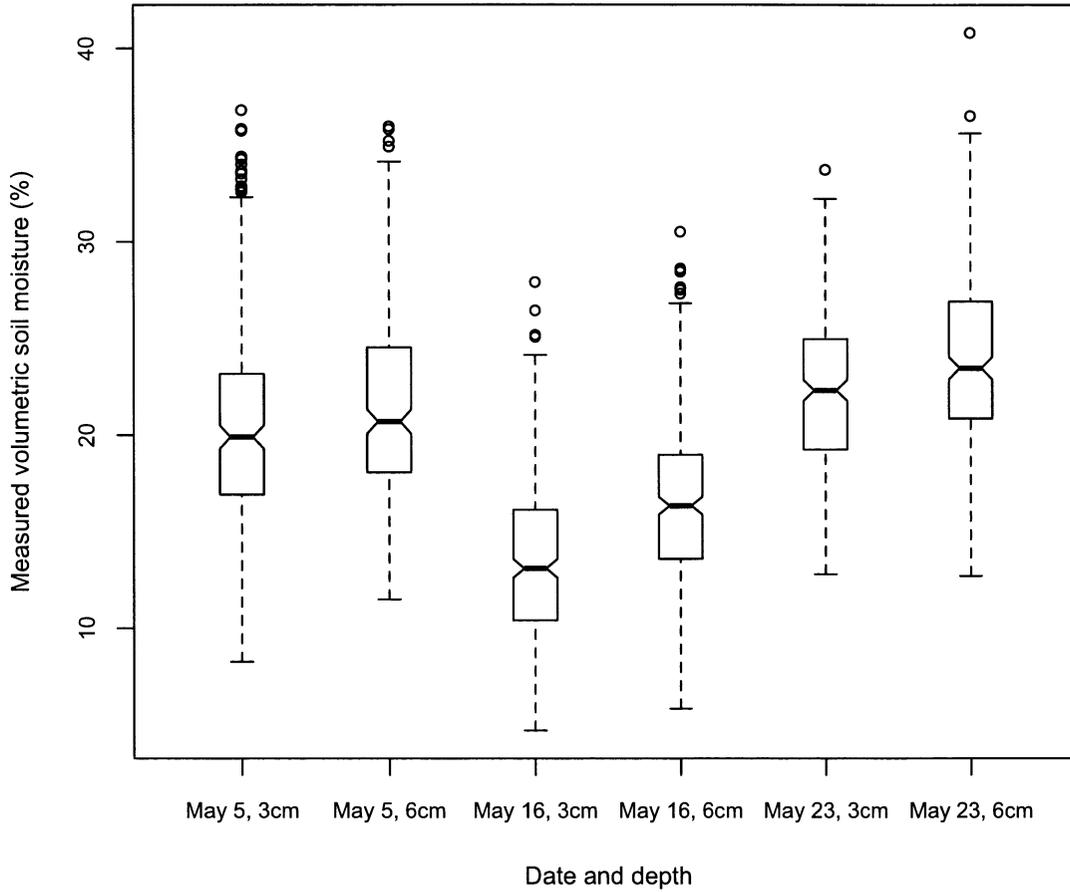


Figure 4: Box-and-whisker plot showing measured soil moisture data at 3 and 6 cm depth on May 5, 16 and 23. Boxes indicates the spread of the second to third quartile about the median. The notch approximates a 95% confidence interval about the median (non-overlapping notches represent statistically significantly different medians). Whiskers extend out to 1.5 times the interquartile range. Outliers are also shown (circles). Note that volumetric moisture content is always higher at the 6 cm depth, and that soil moisture decreased between May 5 and May 16, then returning to slightly higher levels after a rainfall event (21.4 mm over 6 days) between May 16 and May 23.

Table 4: Registration error statistics (in pixels).

Image Date (DOY 2008)	Total RMS	X RMS	Y RMS	Number of GCPs
May 5 (126)	0.91	0.75	0.52	15 + 3
May 16 (137)	0.97	0.82	0.52	15 + 7
May 23 (144)	0.99	0.76	0.63	15 + 13

6.1.2 Filtering, resampling and orthorectification

Once the GCPs were collected, OrthoEngine 10.3 software (PCI Geomatics, Gatineau, Canada) was used to process the three scenes concurrently by: i) resampling the imagery to 8 x 8 meter pixels, ii) performing the slant-range to ground-range conversion, iii) applying a speckle filter (this step is discussed in more detail below) and iv) projecting to UTM coordinates. Because the slant-range to ground range correction, georeferencing and orthorectification are performed in one step, the process is much faster than when the steps are performed individually, which requires intermediate images to be calculated and stored. Applying the speckle filter in this manner eliminates the need for two operations: instead of using cubic interpolation (or another interpolation method) to resample pixels to the projected and corrected version, the average filter both filters and resamples at the same time.

A 1:250,000 scale 3 arc-second digital elevation model (DEM) with 1 meter vertical accuracy (Canada Digital Elevation Model v.3, Natural Resources Canada, 2010) was also used in the orthorectification process to further increase accuracy. The use of a DEM for orthorectification is also optional, especially in areas of low relief, but was performed in order to increase the radiometric and geometric accuracy of the imagery. Local incidence angle affects backscatter considerably, even in relatively flat areas, and in this case, the DEM showed an elevation range of more than 30 meters across the scene, showing that the area is not truly “flat” and can benefit from corrections. An additional consideration in support of using a DEM to perform orthorectification is that the three scenes were not collected with the same viewing geometry.

It should be noted that orthorectification was performed prior to running the soil moisture models in order to avoid having to interpolate the soil moisture estimates, which would alter them undesirably. The rationale for this choice was twofold: if

orthorectification occurs after the model estimates are generated, i) the influences of topography on backscatter are not corrected as they are when they are generated from DEM-orthorectified imagery, ii) the estimates are generated from the unreferenced imagery and must be resampled into square pixels of the final resolution after they are generated, thus requiring some form of averaging which modifies the values of the estimates. It is necessary that the estimates be orthorectified in order to produce a useful map of estimates and to extract values at the specific sample site locations, which are recorded in geographic coordinates. Figure 5 shows the resulting orthorectified imagery.

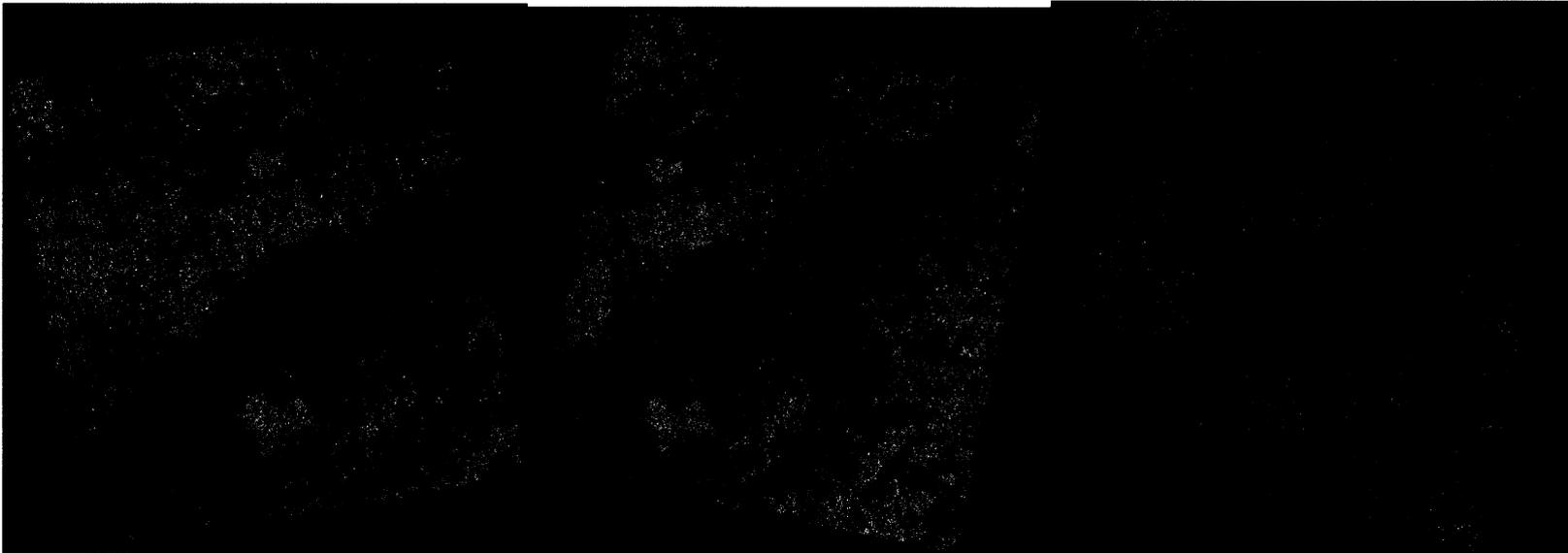
In order to assess the effects of the commonly used boxcar filter and filter size, the images were georectified with boxcar filters ranging from 3 x 3 to 21 x 21 (a more detailed discussion follows in Section 6.4).

6.1.3 Incidence angle map

An incidence angle map was also extracted from the unfiltered imagery and georectified with no filtering, using cubic interpolation (this interpolation method was selected in order to produce smooth gradients in the incidence angle values without altering them). This incidence angle map is used to generate the soil moisture estimates using semi-empirical models.

6.1.4 Backscatter calibration

Many studies using semi-empirical soil moisture retrieval models apply a non-linear backscatter calibration. The models are first run in forward mode and the estimates of backscatter compared to the actual image backscatter. Non-linear relations between the two can then be used to “correct” or calibrate backscatter values, potentially leading to improvements in soil moisture estimates. Another route for calibration is



RADARSAT-2 Data and Products © MacDONALD, DETTWILER AND ASSOCIATES LTD., 2008 - All Rights Reserved

Figure 5: Three orthorectified (unfiltered) RS-2 Fine-quad SLC scenes acquired on May 5, 16 and 23 2008. The images are presented as R:HH, G:HV and B:VV. Casselman is visible in the lower left, as is the South Nation River and Highway 417 (center, running east-west). Differences in backscatter are due to changes in incidence angle, soil moisture and surface roughness. Scenes are approximately 25 x 25 km, north is up.

the calibration of the model's coefficients directly, rather than calibration of the SAR backscatter, such as demonstrated by D'Urso and Minacapilli (2006).

For this study, calibrations were not applied, in part due to methodological constraints. The analysis carried out in Section 6.4 addresses the determination of the optimal speckle filter size, and required multiple model runs with different parameters; calibrating backscatter for each would have been highly time consuming. Instead, the basic models in their original formulation were evaluated, without calibration or corrections and with their originally stated ranges of validity as to provide a baseline of error and bias characteristics to which altered versions of the models could be subsequently compared. Furthermore, calibration is not absolutely required because the comparisons effected between models and dates were relative, not absolute.

6.2 Application of the soil moisture models

The OH92, OH04 and DU95 soil moisture estimates were calculated based on the filtered and orthorectified imagery. A total of 90 soil moisture images were generated (3 dates, 3 models at 10 filter sizes) using a Python (version 2.6) script (Appendix A) to i) import the filtered and orthorectified images, ii) convert to ENVI format, iii) call the PolSARpro module to retrieve the estimate of volumetric soil moisture and iv) output the result to PCIDSK format for viewing.

PolSARpro (ESA v.4.1, 2010) is open-source SAR analysis software created by the European Space Agency (ESA). The source code for the three surface inversion models was evaluated and found to conform to the original model specifications and ranges of validity, with two exceptions (which are addressed further in Section 8), one pertaining to the ranges of valid soil moisture estimates, and the other to the model used to convert between dielectric permittivity and volumetric soil moisture.

6.3 Extraction of estimates at sample sites and pre-extraction averaging

For the model estimates generated at each filter size, a second Python script (Appendix B) was written that i) imported the model estimates generated in the previous section, ii) read the coordinates of sample sites from a CSV file and iii) extracted the value at the sample site with some level of pre-extraction averaging. Pre-extraction averaging consisted of taking a n by n average of non-zero values around the sample site, ranging from no average (the exact value of the pixel under the sample site) to 3 x 3 to 21 x 21 kernel sizes. The Python script then computed RMSE, MBE, Pearson's correlation coefficient and the number of sample sites with non-zero values (after the pre-extraction averaging was applied) separately for sample sites at the 3 and 6 cm depths. The result of the analysis is a database of RMSE, MBE, Pearson's r and the number of sample sites with non-zero values for 3 dates, 3 models, 10 speckle filter sizes, 11 pre-extraction averaging sizes and 2 sampling depths, for a total of 1,980 different combinations of the above.

6.4 Determination of optimal speckle filter size, pre-extraction averaging and field sampling depth

Speckle filtering is an essential step in obtaining estimates of soil moisture using semi-empirical models. Estimates derived without prior speckle filtering show obvious artifacts, such as valid pixels (pixels within the model's ranges of validity for roughness, soil moisture, etc.) within nonsensical areas such as water bodies, roads or forested areas, and also show much less continuity in valid areas. Instead, isolated pixels are found throughout the image. However, in altering the backscatter signature to reduce speckle, the speckle filters that lead to improved visual quality may

not necessarily led to improved model performance, especially in filters that apply different modifications to different categories of pixels. For example, the Enhanced Lee filter provides good speckle reduction for visual interpretation at the expense of adding artifacts to the image (Sumantyo and Amini, 2008). These patterns can be particularly troublesome because they are maintained and replicated in the resulting soil moisture estimates, introducing artificial spatial patterns not related to the landscape.

In a SAR modelling context, a large number of independent returns are required to reduce the influence of speckle, thus requiring large window sizes. Many methods can be employed to “average” multiple returns, including multi-looking, filtering and resampling to larger pixels. Resampling is not the optimal solution because it creates a significant reduction of spatial resolution. If the desired outcome is a map of soil moisture at the pixel level, this is not a good option. Filtering is a good option because it (potentially) maintains the spatial resolution of the original image. The selection of a filter size and of the filter type is however a mostly arbitrary decision, partly due to the fact that the scale of soil moisture patterns is not known *a priori*.

For this study, a standard “boxcar” or average filter was used. This consists of a moving window in which the central pixel is assigned the average of surrounding pixels in a square neighbourhood. This filter is frequently used in radar remote sensing, and often serves as a benchmark for comparing the performance of other filters.

For a filter size to be considered “optimal”, it should lead to maps which meet the following criteria: i) have low RMSE and MBE near zero (which would indicate low overall error and little bias towards over- or underestimation), ii) have high positive Pearson’s r (indicating high agreement and correlation) between the observed and predicted soil moisture contents and iii) have a high number of pixels within the ranges of validity of the models. Not all these requirements will be met simultaneously, thus

a balance is required in determining the “optimal” size.

MBE measures the average direction of prediction error, and can detect under- and over-estimations of soil moisture:

$$MBE = \frac{1}{N} \sum_{i=1}^N (P_i - O_i) \quad (25)$$

where N is the number of cells, P is the predicted soil moisture and O is the observed soil moisture. The RMSE statistic serves as an aggregate measure of the model’s residual error and gives the overall magnitude of the error:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2} \quad (26)$$

Finally, the Pearson correlation coefficient represents the statistical strength of the relation between observed and predicted values and is given by:

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (27)$$

where x and y are the predicted and observed values and n the number of valid pairs of those values.

Additionally, the effect of measurement depth was measured separately to verify if better model performance could be achieved by using either measurements at the 3 cm depth or 6 cm depth. Using the measurements from the two depths combined always leads to poorer agreement over selecting a single set of data because the statistics for the combined data sets are the average of those obtained for the two subdatasets. Thus, any indication of better agreement with either dataset would be helpful in further analysis.

The RMSE, MBE, Pearson’s r and number of non-zero pixel values generated

at the 1,980 different combinations of date, model, speckle filter size, pre-extraction averaging size and sampling depth were analyzed graphically using three dimensional surface plots. General trends were identified and verified using two-dimensional plots, and for each date and model, an “optimal” speckle filter size and pre-extraction averaging size combination was identified. All subsequent analyses were performed using those filtering levels.

6.5 Production of soil moisture maps at “optimal” filter combination

Once the “optimal” filtering combinations for each model were identified, a Python script was used to apply the selected pre-extraction averaging described earlier to the (already) speckle-filtered imagery. This consisted of a customized moving average filter that returned the average of only non-zero pixels within the specified kernel size. The maps of all three models were colour coded with the same colour ramp for ease of comparison, from a range of 0 to 35% volumetric soil moisture, divided into 70 classes, and with pixels having a value of 0 coded in black. The colour ramp was a continuous colour spectrum through magenta, red, orange, yellow, green, turquoise and blue. Finally, a mask of wooded areas (Ontario Ministry of Natural Resources, 2009) was applied to facilitate interpretation (some pixels in forested areas are within the model’s ranges of validity but are not true bare or lightly vegetated pixels as the models require), leaving mostly pixels in agricultural areas and other bare surfaces.

6.6 Analysis of trends in relative error of estimates

6.6.1 Tillage methods

Three multivariate analyses of variance (MANOVA) were performed to assess the effects of tillage methods on the relative error of the three models on the three dates. For May 5, tillage conditions were either minimal fall tillage, conventional fall tillage, spring tillage, no tillage or fall seeded. All fields were tilled (except no-till and fall seeded) and seeded prior to the May 16 image. On May 16 and 23, the tillage conditions observed were fields with fall seeding, spring tillage unrolled, spring tillage rolled or fall tillage minimal. The difference in tillage method between the two dates is due to the application of spring tillage and rolling (in some case) to most fields, with the exception of those seeded in the fall or with minimal fall tillage and direct spring seeding.

6.6.2 Soil properties

Fine-scale soil properties data were obtained from AAFC. The data was a finer-scale version of the Soil Landscapes of Canada version 3.3.1 (SLC, AAFC) than is currently available to the public. The SLC data was revised in 2007 and is distributed in polygon format. The data used consisted of five numerical variables: three soil fraction variables (percentage of sand, silt and clay) plus an additional variable for the percentage of sand consisting of very fine sand, and a variable giving the percentage of organic material. Three additional categorical variables were also examined: texture (clay loam, clay, fine sandy loam, etc.), order (gleysolic, podzolic, etc.) and drainage (perfect vs. imperfect). For the numerical values, simple bivariate correlations were performed to search for significant and strong relations between the variables and relative error. For the categorical variables, a multivariate analysis of

variance (MANOVA) was used to assess the effects of the variable's levels on the relative error of the three models for each date.

7 Results

7.1 Effects of speckle filter size, pre-extraction averaging and measurement depth

RMSE, MBE, Pearson's r and the number of valid pixels were computed for each model and date at a combination of filter sizes. The models were run on imagery pre-processed with boxcar speckle filters ranging from 3 x 3 to 21 x 21 pixels. The model estimates were retrieved either by sampling the pixel directly at the sampling site coordinates or by taking an average of surrounding non-zero pixels ranging from 1 x 1 - no averaging, to 21 x 21. Specific trends in model error (RMSE), bias (MBE), correlation (Pearson's r) and the number of pixels with valid estimates will be discussed below, then summarized in Section 7.1.5. The resulting soil moisture maps are presented in Section 7.2.

7.1.1 Number of sampling sites with non-zero values

A critical consideration for mapping soil moisture is the number of pixels with non-zero solutions, that is, the number of pixels with usable data for generating a map. Figure 6 shows the number of sampling sites for which non-zero values were obtained, either through a model estimate or via pre-extraction averaging.

For the OH92 model, increasing the speckle filter size led to more pixels falling outside the model's range of validity, while increasing pre-extraction filtering increased the number of valid pixels. At a speckle filter size of 3 x 3 and pre-extraction sizes

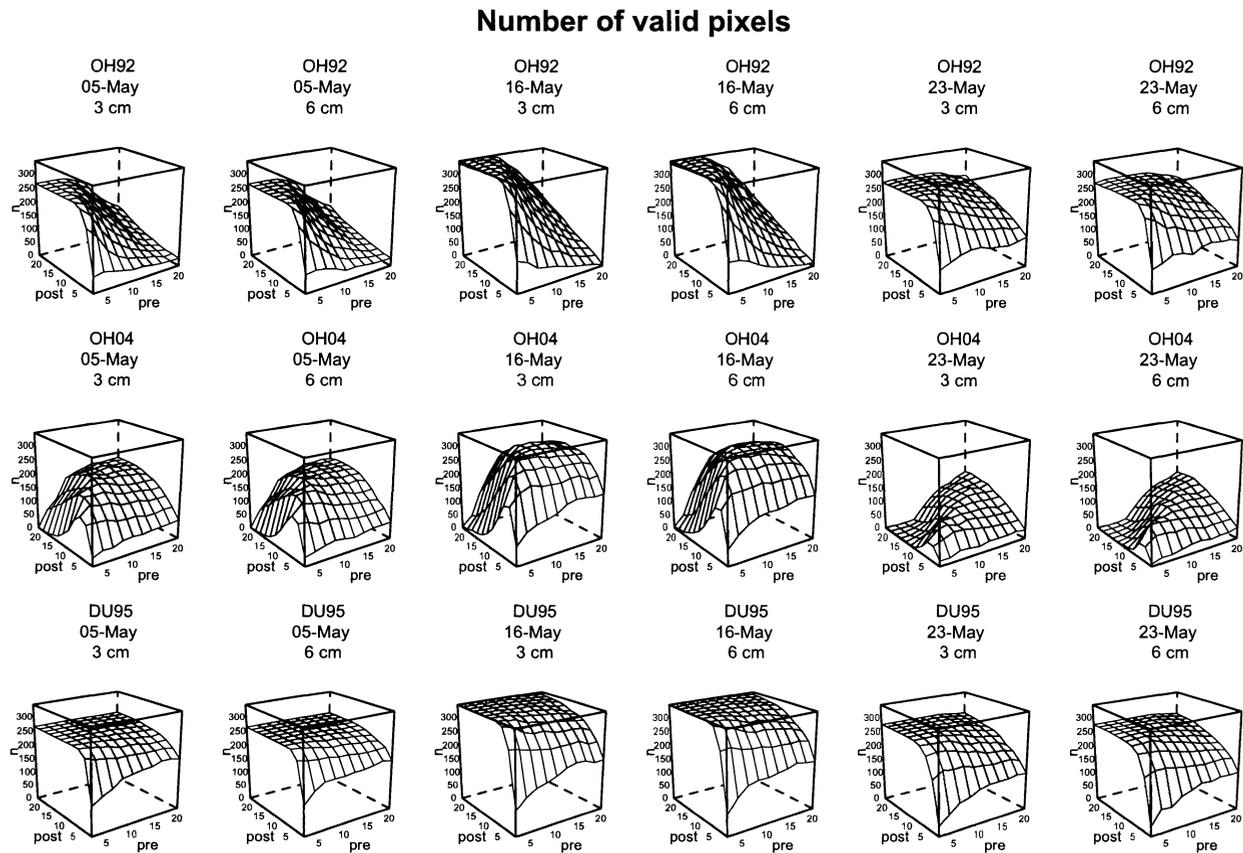


Figure 6: Number of sites with valid model solutions for the three models, two depths and various combinations of speckle filter sizes (indicated as “pre-inversion”) and pre-extraction average filter sizes (indicated as “post-inversion”). Each date has a different maximum number of valid pixels possible, according to the number of sites with valid field soil moisture measurements.

of more than 7×7 , all sites had valid solutions; further pre-extraction averaging was required at larger speckle filter sizes.

The OH04 model had a much different behaviour, generally reaching the maximum number of pixels with speckle filters of 7×7 or more and with pre-extraction averages of 7×7 or more. With low or no pre-extraction averaging, or with small speckle filter sizes, the proportion of valid pixels was extremely low. Note that for this model, some combinations of filter sizes gave no pixels with valid solutions; these will be reflected in other graphs (correlations, RMSE, MBE) as missing values.

Finally, the DU95 had a high proportion of valid pixels when pre-extraction averaging of 5×5 or more was applied, irrespective of the speckle filter size. Except when no pre-extraction filter was applied, increasing the size of the speckle filter at a selected pre-extraction size actually led to a decrease in the number of valid pixels, as it did with the OH92 model.

7.1.2 Model error (RMSE)

For the OH92 model, increasing the size of the speckle filter led to higher RMSE values, though this trend was more muted with the May 16 image, which had lower overall soil moisture conditions and a steeper incidence angle (Figure 7). Averaging prior to extracting the values from the imagery for comparison with the field measurements led to an additional reduction of approximately 1-1.5% in RMSE. Depending on the size of the speckle filter, pre-extraction averaging could actually increase RMSE (May 16 and May 23, for example, where RMSE increases for speckle filter sizes of 9×9 or more). The largest reductions with pre-extraction averaging occurred at small filter sizes, with 3×3 , 5×5 or 7×7 kernels; past these filter sizes the improvements were much smaller. Hence, the lowest RMSE values observed were with a speckle boxcar filter of 3×3 and a pre-extraction 7×7 average. This combination of filters

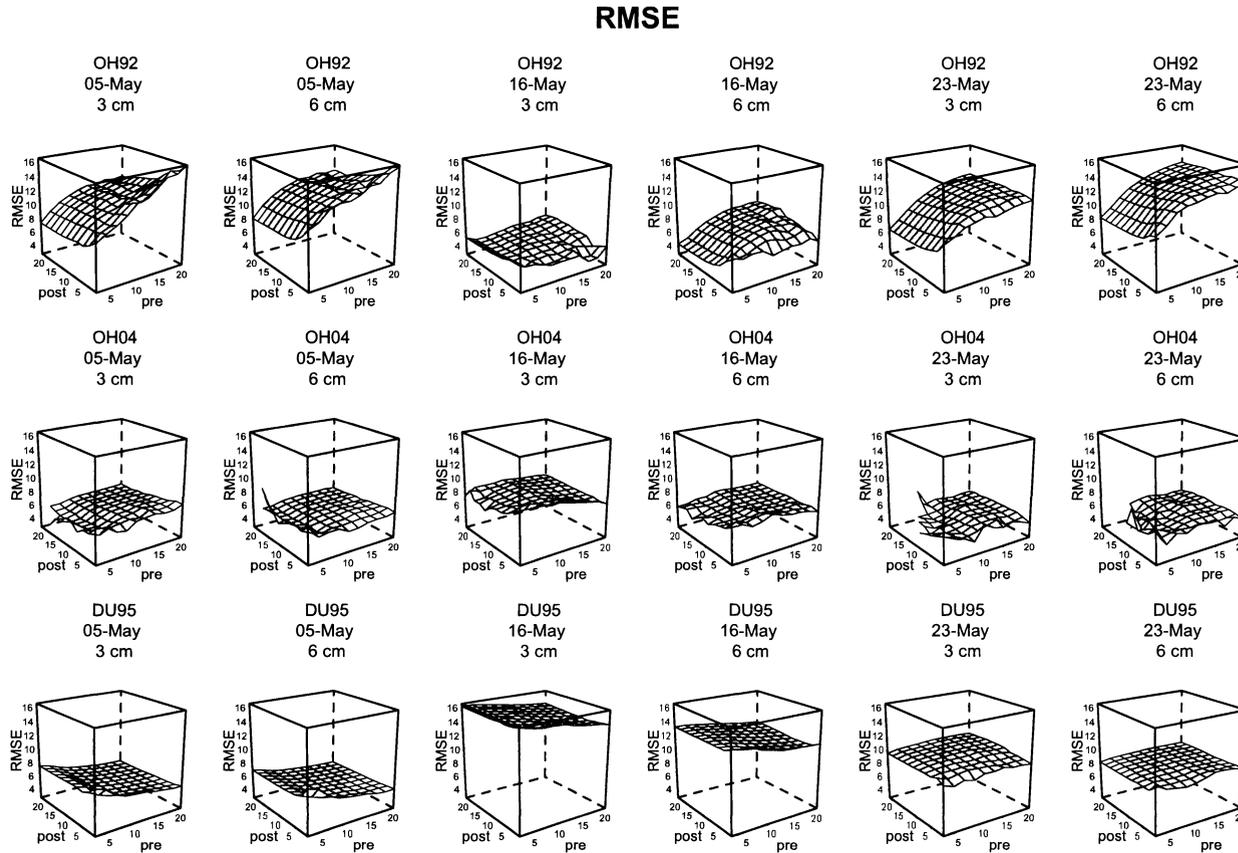


Figure 7: Root mean squared error (RMSE) for the three models, two depths and various combinations of speckle filter sizes (indicated as “pre-inversion”) and pre-extraction average filter sizes (indicated as “post-inversion”). OH04 data should be interpreted carefully according to the number of valid pixels. Filter combinations with low numbers of valid pixels yield unreliable RMSE values.

also leads to solutions for a majority of pixels (Figure 6). For May 5 and May 23, RMSE was lower by approximately 1-1.5% for the 3 cm depth. In the May 16 image, RMSE is lower with the 6 cm data at a 3 x 3 speckle filter, but at larger filter sizes, error is also lower with the 3 cm data. Given these trends we can infer that a good combination is a 5 x 5 speckle filter and 7 x 7 pre-extraction averaging, giving RMSE values of 9.5, 5.1 and 8.4% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 3 cm depth data. Although larger pre-extraction filters would have given slightly smaller RMSE values, they significantly alter the maps produced due to the increased averaging of values; smaller filter sizes are thus desirable.

Given the small number of valid pixels at small speckle filter sizes and small pre-extraction average sizes, the values of RMSE calculated for the OH04 model are more highly variable and inaccurate compared to those of the other models. The general trend (Figure 7) shows that RMSE is not highly affected by either speckle filter size or pre-extraction averaging. For May 5 and May 16, RMSE was lower when comparing with the 6 cm depth data, while error values were similar for both depths on the May 23 image. At a speckle filter size of 7 x 7 and a pre-extraction 7 x 7 average, RMSE values were 5.8, 6.7 and 5.1% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 6 cm depth data. With the OH04 model, error for May 16 was actually higher than for the two wetter dates, but the values are both lower and much more similar from date to date when compared to OH92.

Finally, with the DU95 model, RMSE either decreased or remained stable when increasing speckle filter size. For all dates, RMSE was lower at the 6 cm depth. Additionally, pre-extraction filtering gave no additional reduction in RMSE beyond a 3 x 3 or 5 x 5 average. Due to the decreasing rather than increasing RMSE with increased speckle filter size, the “optimal” combination for the DU95 model differs

from those of the two other models. Good results were obtained with a 7 x 7 speckle filter and 5 x 5 pre-extraction averaging. At that combination, RMSE values were 6.3, 13.0 and 8.6% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 6 cm depth data. Similarly to the OH04 model, error was higher for the driest date, May 16.

7.1.3 Model bias (MBE)

For the OH92 model, MBE values increased with increasing speckle filter sizes, irrespective of the initial bias, whether positive or negative (Figure 8). Pre-extraction averaging either increased or had no effect on MBE values. Thus, in terms of bias, an “optimal” combination is a small 3 x 3 speckle filter and a 3 x 3 pre-extraction averaging, but this combination provides few valid pixels. Using the same combination as used for obtaining the lowest RMSE values (3 x 3 speckle filter and 7 x 7 pre-extraction averaging) yields better results, with MBE values of 5.6, -2.8 and 5.4% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 3 cm depth data. It should be noted that in some cases, pre-extraction averaging can be somewhat detrimental if the estimates are positively biased.

The OH04 model behaves differently than the OH92 with respect to bias. Increasing the size of the speckle filter decreases the value of bias for May 5 and May 23, and increases it somewhat for May 16. However, the impact of speckle filter size is sufficiently small that using a 3 x 3 filter is still the best option. Again, considering the combinations leading to a high number of pixels with non-zero values, a speckle filter size of 7 x 7 and pre-extraction filter of 7 x 7 were selected, giving biases of -1.1, -3.7 and 1.0% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 6 cm depth data.

Finally, the DU95 model, unlike the OH92 and OH04, consistently showed high

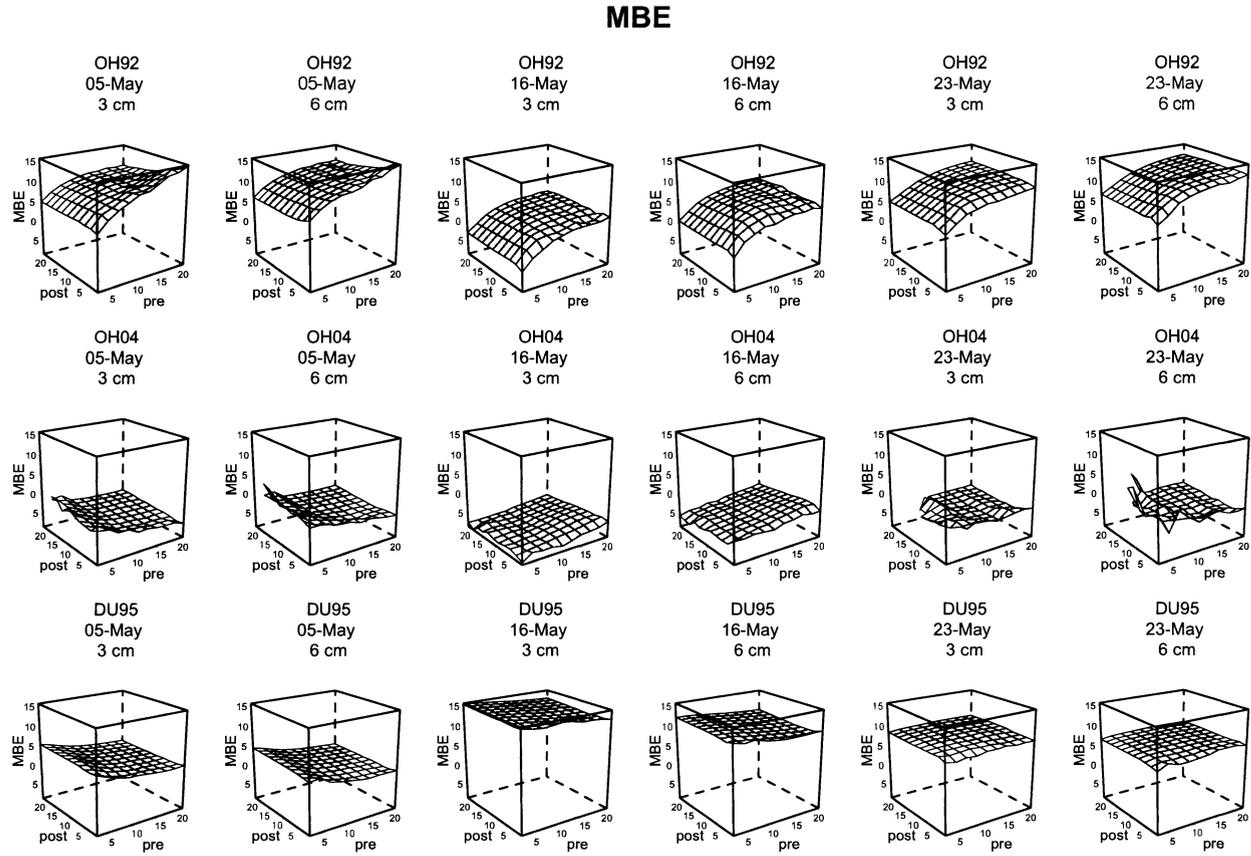


Figure 8: Mean bias error (MBE) for the three models, two depths and various combinations of speckle filter sizes (indicated as “pre-inversion”) and pre-extraction average filter sizes (indicated as “post-inversion”).

positive bias. Increasing the size of the speckle filter led to marginally smaller negative biases, while increasing the size of the pre-extraction averaging window had no significant effect. Biases were lower (closer to zero) when comparing with the 6 cm data. Given this behaviour, the same filtering combination as used to obtain the lowest RMSE values was used, a 7 x 7 speckle filter combined with 5 x 5 pre-extraction averaging, yielding biases of -2.8, -11.8 and -7.2% volumetric soil moisture for May 5, 16 and 23, respectively, when compared with the 6 cm depth data.

7.1.4 Correlation between model estimates and field measurements

Unlike the analysis of RMSE and MBE, analysis of Pearson's correlation coefficient at various filter sizes was more difficult to interpret. Notably, correlations were extremely low, and many were not statistically significant, and for most model/dates no clear relations emerged (Figure 9).

For the OH92 model, the highest correlations were observed at large speckle filter sizes and with large amounts of pre-extraction averaging, but this is likely related to the lower number of valid pixels obtained at those sizes. At smaller filter sizes, correlation coefficients lower than or equal to 0.3 were obtained, and it was difficult to determine if correlations were better at either of the sampling depths. For May 5, most correlations were significant at $p = 0.05$. The correlations that were not significant occurred either when using the smallest 3 x 3 speckle filter (except if using very large pre-extraction averages), or when using no pre-extraction averaging at larger speckle filter sizes. For May 16 and 23, nearly none of correlations were significant, and those that were showed no regular pattern. At the same combination of filtering used to obtain good RMSE and MBE values (3 x 3 speckle filtering and 7 x 7 pre-extraction averaging), Pearson's r values of 0.06, -0.05 and 0.08 were obtained for May 5, 16 and 23, respectively, when correlating with the 3 cm depth data. As mentioned earlier,

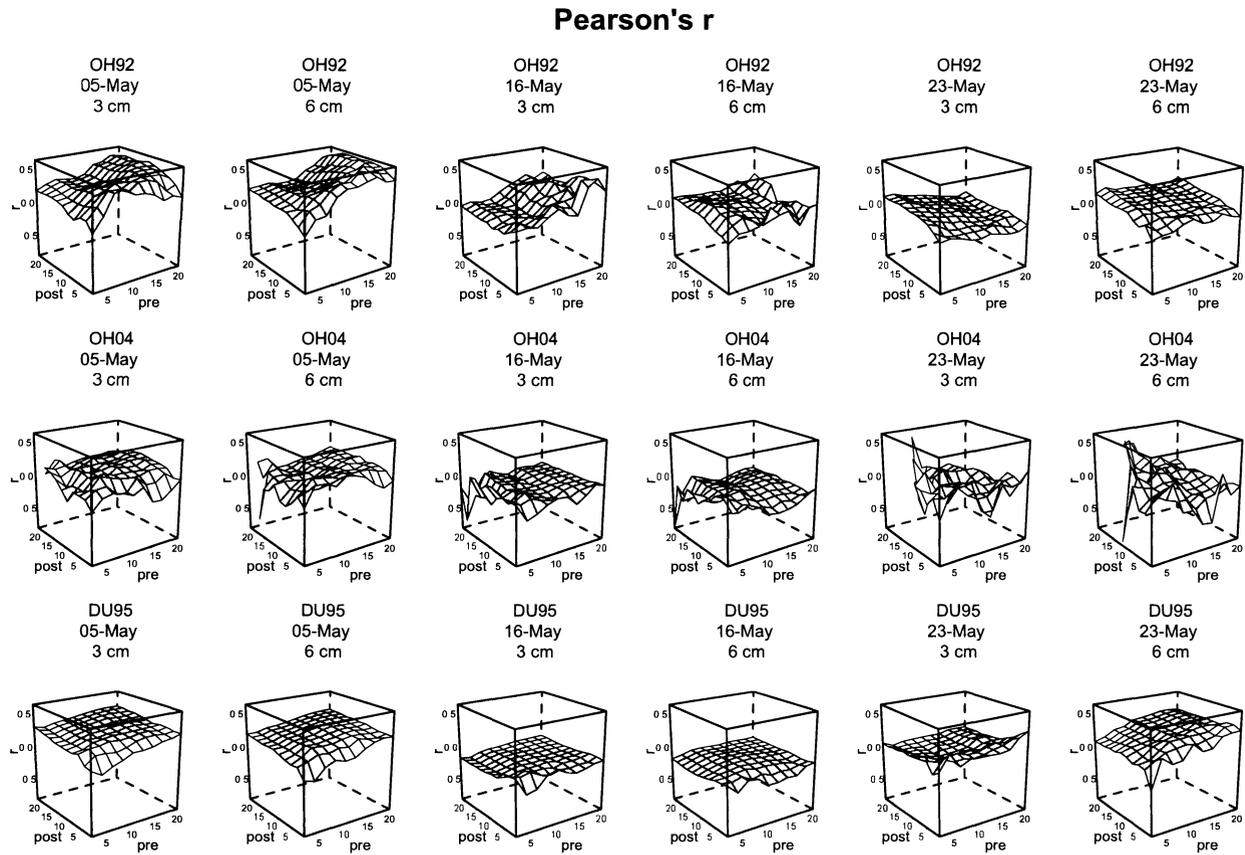


Figure 9: Pearson's correlation coefficient (r) for the three models, two depths and various combinations of speckle filter sizes (indicated as "pre-inversion") and pre-extraction average filter sizes (indicated as "post-inversion").

using a 5 x 5 speckle filter and 7 x 7 pre-extraction average improves correlations considerably, and the correlations become significant; however, this also increases bias and error.

The OH04 model's correlations were difficult to interpret due to combinations of filter sizes leading to little or no valid pixels. Correlations were only significant at either very large speckle filter sizes and/or pre-extraction average filter sizes. At the same combination used for lowering error and bias (7 x 7 speckle filter and 7 x 7 pre-extraction averaging), non-significant correlations of 0.17, -0.04 and 0.05 were obtained for May 5, 16 and 23, respectively, when correlating with the 6 cm depth data.

Finally, the DU95 model had the most uniform response to filter sizes, with correlation increasing for increased speckle filter size and either increasing or remaining stable with increased pre-extraction averaging. Similarly to the OH92 model, most correlations were significant for the May 5 data, except when using no pre-extraction averaging or when using very small speckle filter sizes (3 x 3). For May 16, small speckle filter sizes required large pre-extraction averaging to obtain significance, while for May 23 correlations tended to be significant only for correlations at with the 6 cm depth data. Using the same filter sizes as for RMSE and MBE values (7 x 7 speckle filter and 5 x 5 pre-extraction averaging), correlation coefficients of 0.26, -0.09 and 0.22 were observed for May 5, 16 and 23, respectively when correlating with the 6 cm depth data. Only the correlations (two-tailed) for May 5 and May 23 were significant at $p = 0.05$.

7.1.5 Summary of model statistics and optimal filter sizes

Based on the trends described above, the requirements of the three models differ. To minimize error and bias and maximize correlation and the number of valid pixels, the

best combinations of filters for OH92 was a small, 3 x 3 speckle filter with a 7 x 7 pre-extraction average, for the OH04 model, 7 x 7 speckle filter and pre-extraction average and for the DU95, a 7 x 7 speckle filter and 5 x 5 pre-extraction average. The OH92 model performed better when compared to the 3 cm depth data, while the OH04 and DU95 models had better performance when compared to the 6 cm depth data. Table 5 presents a summary of RMSE, MBE, r and number of valid pixels for the three models with these filter combinations.

7.2 Soil moisture maps at “optimal” filter sizes

Figures 10, 11 and 12 show the soil moisture maps produced with the above “optimal” filter combinations.

A major issue highlighted by these maps is that the model’s bias is not constant on the three dates despite using the same filter combinations. The DU95 maps, for instance, show increasing soil moisture whereas the field data indicates equal soil moisture on May 5 and 23 and lower soil moisture on May 16. If we subtract the biases (2.8, 11.8 and 7.2%) from the average estimated soil moisture, we obtain a more exact representation. However, the biases are not usually known *a priori*, rendering the models ineffective. Section 7.4 discusses a potential solution to this problem.

The issue of insufficient pixels with non-zero solutions which was discussed earlier is also visible in map form, noticeably in the OH04 maps. The May 23 map does not have sufficient pixels to permit a visual analysis of patterns, unlike the two other models which show clear large scale and within-field trends. Additionally, the distribution of values appears to be much more uniform with OH04 than with the other models (i.e. values are closer to the overall mean estimated soil moisture), which show more variability.

Table 5: Summary of model statistics and “optimal” combinations of speckle filter and pre-extraction average size.

Statistic	Model								
	OH92			OH04			DU95		
	<i>May 5</i>	<i>May 16</i>	<i>May 23</i>	<i>May 5</i>	<i>May 16</i>	<i>May 23</i>	<i>May 5</i>	<i>May 16</i>	<i>May 23</i>
RMSE (m_v , %)	9.5	5.1	8.4	5.8	6.7	5.1	6.3	13.0	8.6
MBE (m_v , %)	5.6	-2.8	5.4	-1.1	-3.7	1.0	2.8	11.8	7.2
Pearson's r	0.06	-0.05	0.08	0.17	-0.04	-0.05	0.26*	-0.09	0.22*
Non-zero sites (%)	100	100	100	74	86	35	96	92	89
Best depth	3 cm			6 cm			6 cm		
“Optimal” filtering	3 x 3, 7 x 7			7 x 7, 7 x 7			7 x 7, 5 x 5		

*Significant at $p = 0.05$

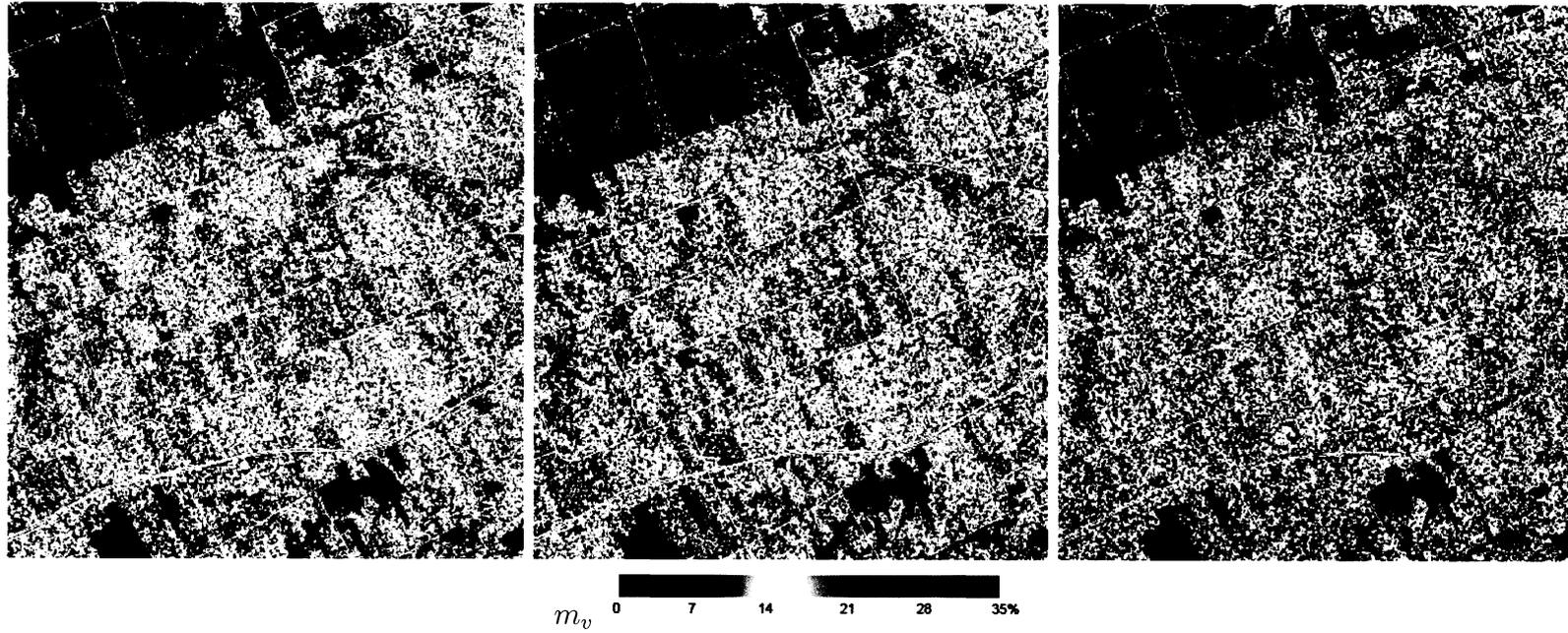


Figure 10: Maps of volumetric soil moisture content estimates from the **OH92** model for May5 (left), 16 (center) and 23 (right). Black pixels are either masked wooded areas or pixels outside the model's range of validity.

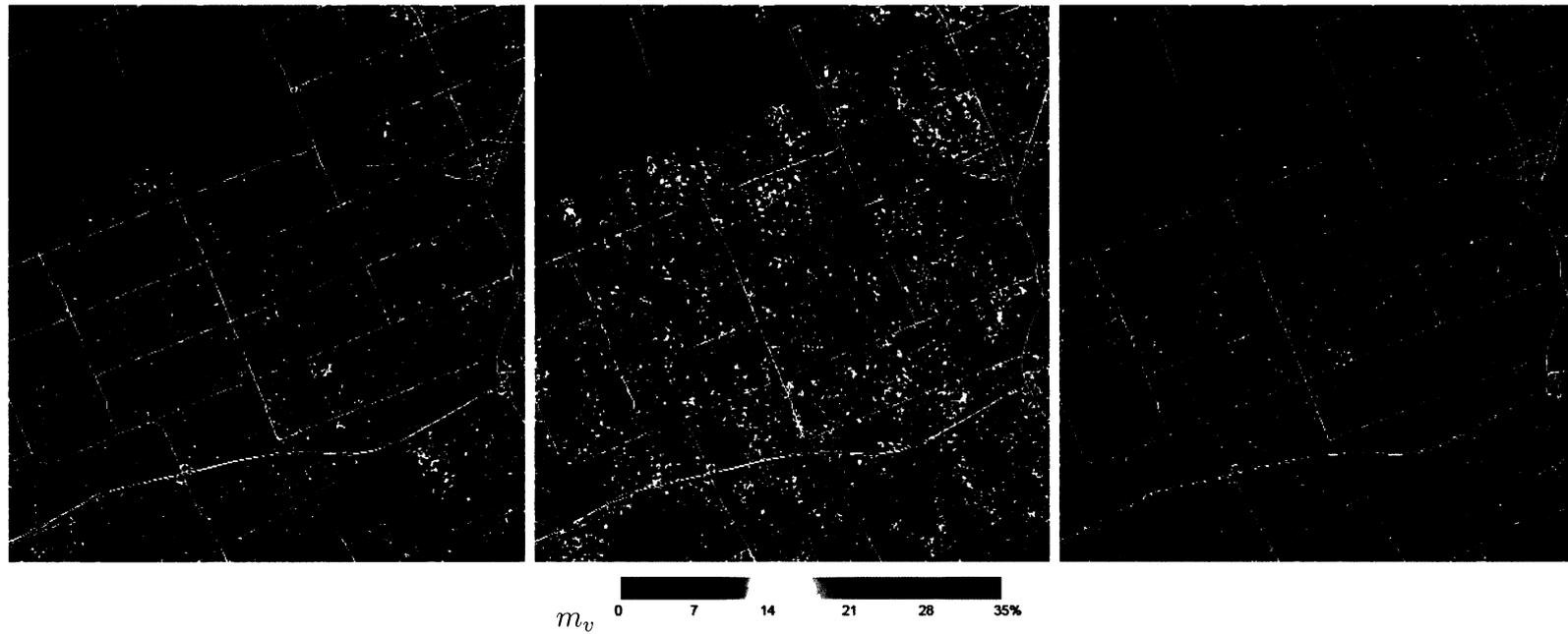


Figure 11: Maps of volumetric soil moisture content estimates from the **OH04** model for May5 (left), 16 (center) and 23 (right). Black pixels are either masked wooded areas or pixels outside the model's range of validity.

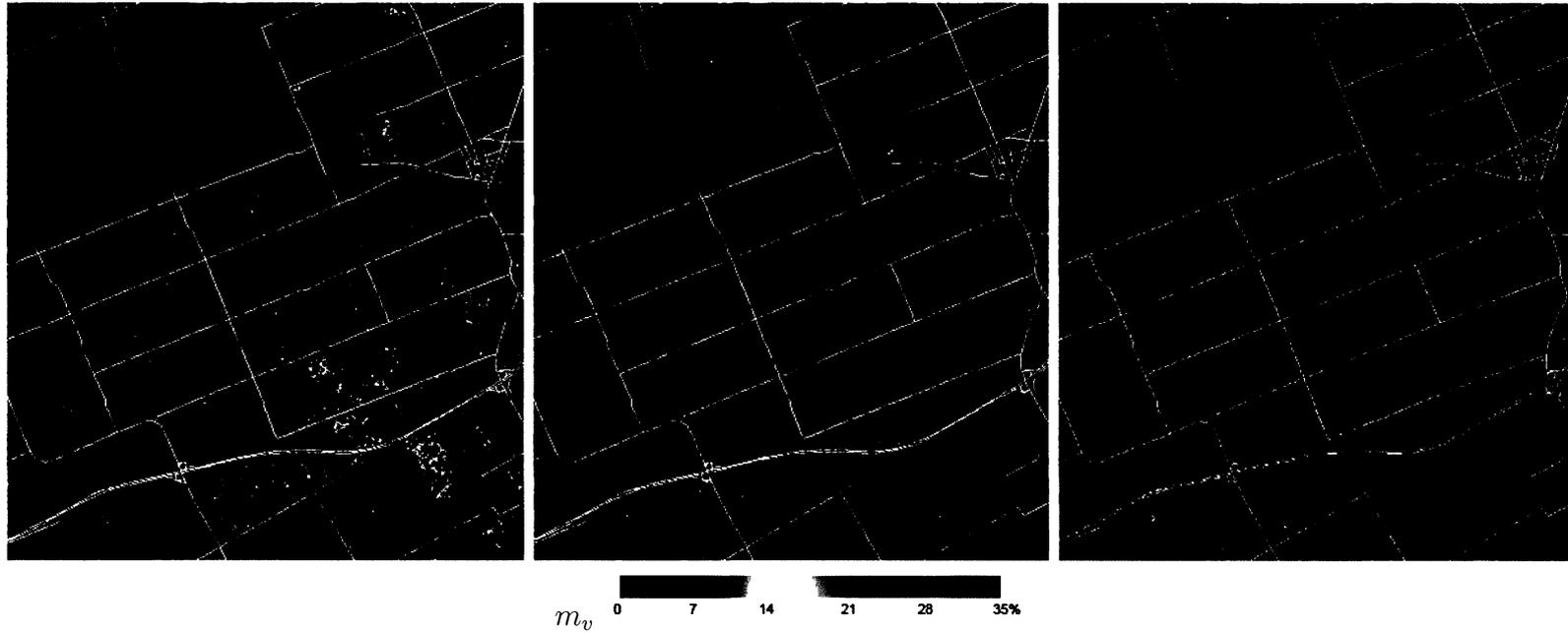


Figure 12: Maps of volumetric soil moisture content estimates from the **DU95** model for May5 (left), 16 (center) and 23 (right). Black pixels are either masked wooded areas or pixels outside the model's range of validity.

A third issue, which shall be discussed further below, is that the distribution of values is “clipped” in order to respect the model’s ranges of validity.

Finally, there is also some correlation between the three models. On all dates, the DU95 model was significantly correlated with the OH04 model ($r = 0.447, 0.671$ and 0.323 at $p = 0.01$ for May 5, 16 and 23, respectively). Additionally, the OH92 model was significantly correlated with the DU95 on May 5 ($r = 0.213, p = 0.01$) and with the OH04 model on May 16 ($r = -0.106, p = 0.01$).

7.3 Trends in relative error of estimates

The model estimates produced with the “best” combinations of speckle filters and pre-extraction filters were extracted for each site. The relative errors of the estimates were analysed for correlations with tillage methods (Section 7.3.1) and soil properties (Section 7.3.2).

7.3.1 Influence of tillage method

May 5 For all three models, the Wilks Lambda multivariate test of overall differences among tillage methods was statistically significant ($p < 0.001$) indicating that relative error values are significantly different according to tillage method (but not necessarily that it is different for all methods); the effect of tillage method was moderate, as indicated by the partial-eta squared score of 0.162, 0.106 and 0.083 for the OH92, OH04 and DU95 models, respectively (16.2, 10.6 and 8.3% of the variance in relative error can be attributed to tillage methods).

For the OH92 model, relative errors of sites with minimal fall tillage were not significantly different than sites with no-till methods or sites seeded in the fall. However, there was a significant difference with sites with conventional fall tillage or sites

which had had their spring tillage; additionally, the difference between conventional fall tillage and spring tillage was also significant. In summary, the effects of tillage methods on OH92 relative errors (Figure 13) can be grouped in three groups: i) sites with minimal fall tillage, no-till methods or seeded the previous fall, ii) sites with conventional fall tillage and iii) sites with spring tillage. Sites with spring tillage applied had the lowest relative error of all (a mean relative error of -10.5% - note that this is a percentage of the estimate, not a percent volumetric soil moisture), followed by sites having received conventional fall tillage (mean relative error of -15.8%). The mean relative error was nearly twice as high for sites having received minimal fall tillage, no tillage or having been seeded the previous fall (mean relative errors of -33.7%, -39.1% and -31.1%, respectively).

With the OH04 estimates, the same groups of tillage methods are observed: i) no-till, minimal fall tillage and fall seeding, the three of which are not significantly different from each other and which are different from the spring tillage and conventional fall tillage methods, and ii) spring tillage and conventional fall tillage methods which are not significantly different from each other. In opposition to the OH92 model, sites with spring tillage and conventional fall tillage had the highest mean relative errors (24.0 and 22.8%) while sites with minimal fall tillage, fall seeding or no tillage had lower relative errors (-2.1, 5.2 and -8.9%).

Finally, the DU95 estimates show the same grouping of tillage methods as the OH04 model: i) no-till, minimal fall tillage and fall seeded sites and ii) conventional fall tillage and spring tillage. Like the OH04 model, there is no statistically significant difference amongst methods in either group. The tillage methods with the lowest relative errors were no tillage, minimal fall tillage and fall seeded (3.9, 5.2 and 8.1%) and those with the highest relative errors, conventional tillage and spring tillage (25 and 30.1%). Interestingly, the mean relative errors of the OH04 and DU95 models

are not statistically different, such that the behaviour of relative error with regards to tillage treatment appear to be identical and opposite that of the OH92 model.

The groupings observed appear to be related to soil roughness characteristics. In the group with high relative error (fall tillage conventional and spring tillage), the soil is more disturbed with higher roughness than in the group with low relative error (no tillage, fall seeding and minimal fall tillage), which disturbs the soil less and for which the soil is unworked in the spring prior to the acquisition of imagery. The differences could also potentially be related to differences in evaporation and infiltration of the various tillage methods.

May 16 The effect of tillage method on relative error decreased with progression in the growing season, but was still significant for all models, with eta-squared values of 0.033, 0.068 and 0.035 (only 3.3, 6.8 and 3.5% of the variance in relative error can be accounted for by tillage methods). This decrease in the effect is partially due to the progression of seeding and spring tillage. By May 16, all sites were seeded. All sites except those seeded in the fall received fall tillage, followed by spring tillage and seeding (with the exception of two fields which received minimal fall tillage and were seeded directly in the spring). Additionally, some sites were rolled after seeding.

In the OH92 model, there is a significant difference in the relative errors of sites seeded in the fall and those with spring tillage, either rolled or unrolled. There was no significant difference in relative error between sites that were rolled after seeding and those that were not. The relative error of sites having received only minimal fall tillage was not significantly different than that of any of the other tillage methods; however only two fields used this method, making it difficult to ascertain the significance of this result. This results in two tillage groups: i) those sites having received spring tillage and ii) the sites seeded in the fall. Again, the sites seeded in the fall had lower

mean relative error (3.5%) compared to those with spring tillage and spring tillage plus rolled (21.8 and 31.4%).

With the OH04, a significant difference in relative error occurred between spring tillage rolled and all other methods, which are not significantly different from each other. Again, this results in two groups: i) spring tillage rolled with high mean relative error (77.9%) and ii) minimal fall tillage, fall seeded and spring tillage unrolled with low relative error (35.5, 38.7 and 42.6%).

Finally with the DU95 model, as with the OH04 model, the fall minimal tillage relative error was not significantly different from any other method, again, likely due to a lack of sufficient sites with this method. The relative error of fall seeded sites was not significantly different from that of sites having received spring tillage, but sites having received rolled spring tillage were significantly different from the two others. This creates the same two groups as with the OH04 model: i) sites with rolled spring tillage (high mean relative error at 133.8%) and ii) sites seeded in the fall or with unrolled spring tillage (with lower mean relative errors of 90.7 and 103.3%).

It should be noted that the mean relative errors observed for May 16 are much higher than those observed on May 5. In the DU95 and OH04 models, tillage methods which tend to decrease soil roughness (such as rolling) led to lower relative errors compared to techniques which increase soil roughness.

May 23 No tillage operations were applied between May 16 and May 23. The Wilks Lambda multivariate test of overall differences among tillage methods was not statistically significant ($p = 0.021$) indicating that relative error values were not significantly different according to tillage method on this date.

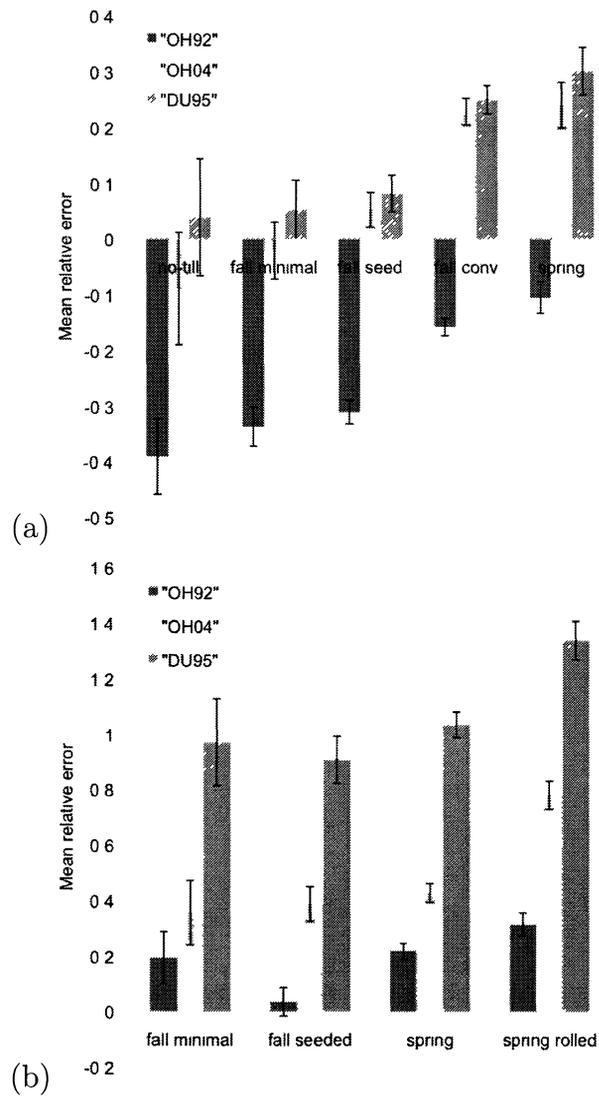


Figure 13: Mean relative errors for the three models and various tillage methods (a) May 5 and (b) May 16. Error bars represent the standard error of the mean.

7.3.2 Influence of selected soil properties

The MANOVAs indicated that neither soil order, soil texture nor soil drainage had significant effects on the relative error of each model/date. Correlations between soil fraction variables and relative error were significant, for the most part, but weak ($r < 0.3$) and not applicable to all model/dates.

7.4 Relation between model bias and incidence angle

The results in Table 5 appear to indicate a link between the MBE of estimates and incidence angle. To further investigate this link (and improve the power of the analysis), MBE and mean incidence angle were calculated for each field (16 sampling points), yielding 132 data pairs of average incidence angles and field bias for each model (44 sites per day on three dates). Figure 14 shows that the relation between incidence angle and bias (MBE) in the three models was strongly negatively and significantly correlated with $r = -0.634$, $r = -0.381$ and $r = -0.715$ all at $p = 0.01$ for the OH92, OH04 and DU95 models, respectively. A final result of interest is the x-intercept, the incidence angle which produces no bias. This was approximately 31, 41 and 43 degrees for the OH92, OH04 and DU95 models. This result is in part contrary to the notion that steeper incidence angles are better for soil moisture mapping. Despite the strength of these relations, it is important to note that soil moisture content and surface roughness varied for each date, and that these factors may also be responsible for this apparent relation. The May 5 and May 23 images, for example, differ only slightly in incidence angle, have nearly identical soil moisture contents, but roughness was lower on May 23 than on May 5. A more controlled experiment should account for changes in these factors individually.

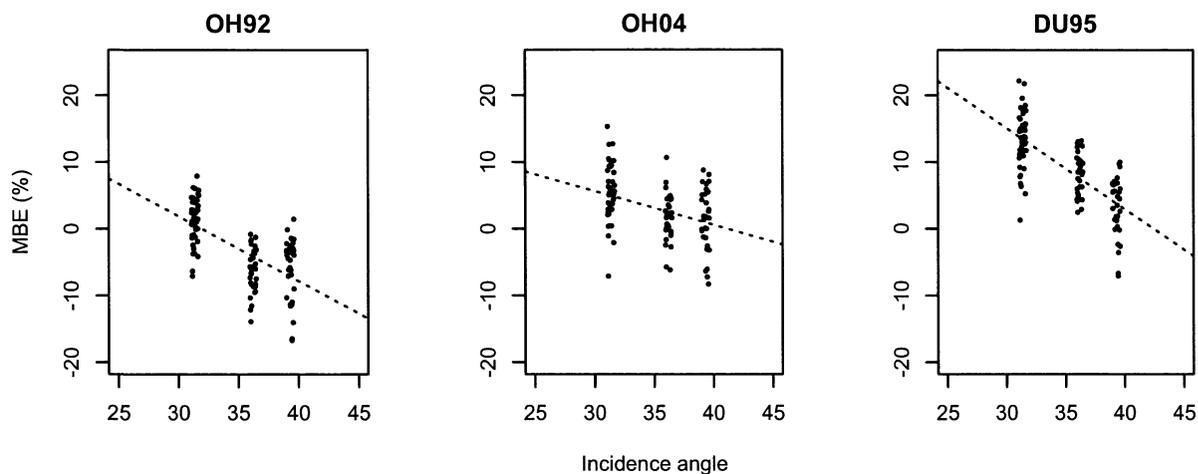


Figure 14: Relation between incidence angle and model bias at the field scale. Linear regression is shown in dashed line. (The May 5 image was acquired at approximately 39 degrees, the May 16 image at approximately 31 degrees and the May 23 image at approximately 36 degrees)

Table 6: Linear regression of incidence angle and model bias at the field scale.

Model	Intercept	Slope	R^2	Adjusted R^2
OH92 MBE	30.97	-0.97	0.402	0.397
OH04 MBE	20.80	-0.51	0.145	0.137
DU95 MBE	51.43	-1.21	0.511	0.507

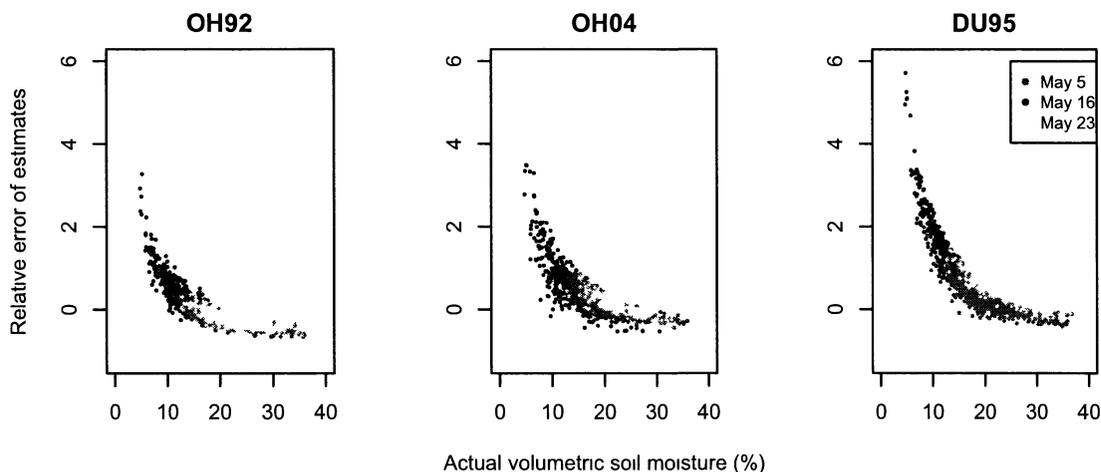


Figure 15: Relation between relative error of the three models' estimates to the measured volumetric soil moisture. Note the increased relative error values at lower soil moisture contents, also accompanied by an increase in the variability of those values.

7.5 Influence of actual soil moisture conditions

A plot of the relative error of soil moisture estimates versus actual measured soil moisture reveals a strong trend (Figure 15). Relative error was higher in low soil moisture conditions, and decreases logarithmically to eventually stabilize at higher soil moisture contents. The relation is stronger in the OH92 model, and is also more stable, with low relative errors from approximately 15 to 40%. In the OH04 model, relative error decreases more slowly with increasing soil moisture, and the range of relative errors at any given soil moisture content is more variable. Finally, the DU95 model had the highest relative error at low soil moisture content (due to apparent outliers), and unlike the OH92 model, had a very small region in which relative error was consistently low (from 25 to 40%). The graphs agree with the statistics obtained in Table 5 ; the position of the cloud of points about zero is related to model bias, and the range of values at a given moisture content to model error.

8 Discussion

8.1 Model implementation considerations

As discussed in Section 6.2, two issues in the PolSARpro implementation of the soil moisture models were found and addressed in the analysis, the first concerning ranges of validity for soil moisture and the second concerning the conversion of dielectric permittivity to volumetric soil moisture in the OH92 and DU95 models (volumetric soil moisture is estimated directly in the OH04 model).

Firstly, the PolSARpro implementation allowed for wider ranges of valid soil moisture contents than the original models: the OH92 and DU95 models allowed dielectric permittivity values of 0 to 20, which when converted with the Topp et al. (1980) equation equates to a volumetric soil moisture contents of 5.30 to 45.14%, well beyond the originally stated range of 9 to 31% for OH92 and 0 to 35% for the DU95 model. Similarly, the OH04 model PolSARpro implementation allowed soil moisture contents of 0 to 40%, while the originally stated range is only 4 to 29.1%. This issue was circumvented by using the PolSARpro output of dielectric permittivity directly for the OH92 and DU95 models, masking out values beyond the originally stated ranges of validity and applying the conversion to volumetric soil moisture using a Python script. Extending the ranges of validity can of course be a desirable approach, yielding more valid pixels for mapping, but the intention of this project was to provide an evaluation of the original models with their originally stated ranges of validity, therefore the ranges were not extended. Further analysis should compare the original and extended models, particularly in terms of error and bias. Unless model performance is similar or improved, extending the ranges of validity is not a desirable approach.

The selection of a model for converting between dielectric permittivity and volumetric soil moisture is important with regards to the accuracy of estimates. Some

other models for converting between the two are given by Hallikainen et al. (1985), and a simple equation is also suggested in the ThetaProbe documentation (Decagon Devices, 2006; this is the equation used when applying the conversion in the instrument itself). The Hallikainen model, however, requires information about the soil fractions (sand and clay), and is thus not the best option for use with remotely sensed data. Interestingly, the relations used by the ThetaProbe (one for mineral soil, the other for organic soils) differ from the Hallikainen equation (Figure 16). Using the Hallikainen equation may have led to an overall positive bias (equivalent to shifting the distribution of soil moisture values) which may be distinguishable in the MBE statistics calculated. It is difficult to argue in favour of one method over the other; however, using the same equation as used when collecting field data may have led to better agreement between the estimates and the field data. Two methods could have allowed for a better calibration of soil moisture values: i) calibrating the soil moisture probe conversion coefficients prior to data collection or ii) collecting soil samples for analytical determination of volumetric soil moisture and calibrating after data collection. The difference between these equations and the fact that the default calibration for mineral soils was used on the ThetaProbe could account for a portion of the bias observed in the OH92 and DU95 models. The OH04 model results are not affected by this because they calculate soil moisture directly.

A third issue, which was not addressed, is the choice of starting points for iteratively solving the OH92 and OH04 models. The PolSARpro implementation implements the Newton-Raphson iterative root finding method explicitly, and uses a starting point of 0.5 to determine the first soil moisture estimate in the OH04 model (Equation 14) and a starting point of 2 to determine Fresnel reflectivity in the OH92 model (Equation 9). The Newton-Raphson method is sensitive to the selection of starting points, and can stabilize at the incorrect solution when multiple solutions

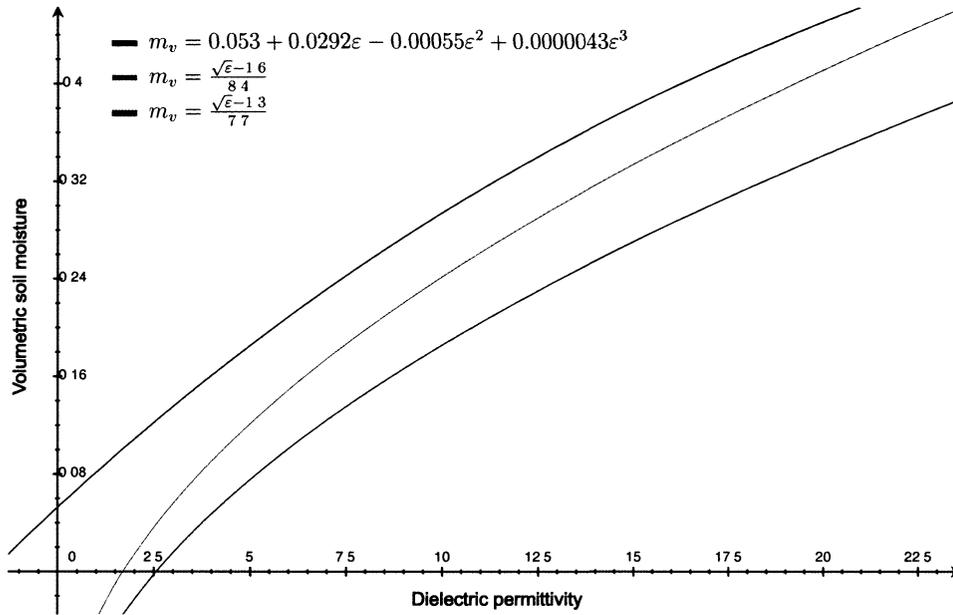


Figure 16: Comparison of models used to convert from dielectric permittivity to volumetric soil moisture. Showing Topp (red), and ThetaProbe curves for mineral soils (blue) and organic soils (green).

are present if the starting point is incorrectly chose. Given that the validity range for OH04 is from 4 to 29.1% volumetric soil moisture, a starting point of 0.15 would have been more appropriate, and more likely to select an appropriate solution when multiple solutions were present. This factor is a likely explanation for the low number of valid pixels returned by this model; selecting a high starting point led to rejecting solutions on the basis of them being outside of the model's range of validity. With this high starting point, the model also had a tendency to overestimate soil moisture by allowing the first estimate of soil moisture to be higher than the range of validity, subsequently reducing it to acceptable limits during the weighted averaging process. Similarly, the range of dielectric permittivity for the OH92 model is from 0 to 20, which gives Fresnel reflectivities of approximately 0.4 to 1 according to Equation 8, thus the selection of 2 as a starting point is inappropriate and likely to lead to solutions outside the range of validity; a value around 0.7 would have been a more

appropriate starting point. Further analysis would be required to determine the effects of starting points on the estimates and number of valid pixels for the OH92 and OH04 models. It should be noted that appropriate starting points for the iterative methods are not suggested by the authors of the two models.

Two other issues were identified with the OH04 model: i) the rounding of exponent coefficients and ii) the weighted averaging of the three soil moisture estimates. Firstly, an exponent term in the originally published OH04 model was rounded and should theoretically be represented by a fraction. This is also the case in the PolSARpro implementation. This occurs in Equations 13 and 14, where the 0.556 exponent should actually be given as $5/9$. This rounding error could become significant especially as it is carried through the iterative process and affects the first of the three estimates of soil moisture. Further analysis is required to determine if this omission has any significant effect. Secondly, the final estimate of soil moisture is obtained by using a weighted average of three estimates of soil moisture derived from different formulas. The first, as we have discussed earlier, tends to be very high (sometimes well beyond the model's range of validity). In cases where no solution is found by the iterative process, this first estimate is null, and not included in the weighted average. This is problematic because the final estimate relies on this first estimate (which tends to be higher) to counter-balance lower second and third estimates. Further analysis would be required to determine if there is a significant difference between final estimates derived from 2 or 3 estimates. Preliminary results indicate that since the first estimate is often inflated, excluding it from the weighted mean can dramatically reduce the final soil moisture estimate; this would tend to bias the overall estimates negatively. Additionally, the author of the model provides no discussion or analysis of the effects of the weights on the final estimate. (It should be noted that the weights for the calculation of the final soil moisture estimate are all 1, which makes the weighted

average identical to the arithmetic mean, but that the weights for estimating soil roughness are not all unity).

One final issue was found with the implementations of the OH92 and DU95 models. Dubois et al. (1995) suggest the pre-estimation masking of pixels with $NDVI > 0.4$ to estimate only for bare, non-vegetated soil covers. This is achieved using a relation between NDVI and the the cross-polarized ratio, described by the author (that an NDVI of 0.4 corresponds to a cross-polarized ratio of approximately -11 dB, depending on the incidence angle). However, this relation was found with L-band data, despite the model being valid for C-band. It is not known whether the relation holds for C-band, hence using this value as an initial vegetation mask is somewhat arbitrary until the link observed with L-band data can be replicated with C-band data. For an undetermined reason, PolSARpro also uses this relation to mask out non-vegetated pixels in the OH92 model despite its absence in the original formulation.

8.2 Optimal speckle filter size, pre-extraction averaging and field sampling depth

The analysis of optimal filter size demonstrated the importance and significant influence of speckle filtering on model estimates. The models performed best with boxcar filters of 3x3 for the OH92 model, and 7x7 for the OH04 and DU95. The finding that less speckle filtering gives better estimates for the OH92 model is unexpected, while the two other models perform as expected, with lower error and bias with increased filter size. The analysis also showed that pre-extraction averaging can reduce error by a few percent and further reduce bias, and that it improves the mappability of estimates substantially. The notion that pre-extraction filtering improves the accuracy of estimates agrees with the findings of many studies (i.e. Álvarez-Mozos et al.,

2005) that found higher accuracy in estimates at larger extents (field or watershed) when compared to the point scale, at which speckle is an issue, despite filtering. In this case, pre-extraction averaging serves as an aggregation method and improves the accuracy of the estimates, similarly to generating estimates at the field level.

The absence of significant correlations in the OH92 and OH04 models was unexpected, but can be attributed to many factors: i) some implementation issues discussed above, ii) differences in incidence angles among dates and the use of non-optimal incidence angles, iii) the low soil moisture content observed on May 16, which increased error and iv) the absence of model calibration. Álvarez-Mozos et al. (2005) also noted a decrease in correlation at finer scales (at the point or pixel scale) versus coarser scales (the field scale), and that the absence of significant correlations could be caused by differences in roughness in different tillage classes within the data. For instance, low-roughness tillage classes would have lower error and higher correlation, and vice-versa, but the overall effect results in the absence of a significant correlation.

In terms of error and bias, the OH04 model had the best and most consistent performance, with an RMSE of about 6% and small biases. The problem with this model, however, is that a low proportion of pixels have non-zero values, and that the correlations between the observed and estimated soil moisture values are not significant. The OH92 model had error and bias levels higher than the OH04, but had solutions for all sites, despite having no significant correlations. The DU95 model had the only significant correlations (May 5 and May 23), but had higher error and bias for the May 16 image compared to the two other models. The RMSE values obtained for most models and dates were better or similar to those obtained in other studies (for example Baghdadi et al. (2006) report RMSE of 6-9% with ASAR data and a simple regression model between backscatter and soil moisture, and Jiankang et al. (1996) an RMSE of 6.1% with the DU95 model and EMISAR data). With the

exception of the DU95 model, RMSE was higher on the wetter dates due to the lower sensitivity of backscatter to soil moisture at higher moisture levels (Biftu and Gan, 1999). A final consideration is that RMSE cannot be lower than the error of field measurements (approximately 5% as discussed earlier), which is also included in the RMSE (Jiankang et al., 1996).

The three models show similar responses for the May 5 and May 23 images, which had similar moisture conditions but different incidence angles, and a different response for the May 16 image, which had lower soil moisture and a steep incidence angle. For the OH92, error and bias are lower on that date while they are higher for the two other models. Interestingly, bias appears to be related to the incidence angle of the imagery (this will be discussed in Section 8.3).

8.3 Relation between model bias and incidence angle

The analysis of optimal filter size revealed an apparent relation between bias and incidence angle. When further investigated, at the field scale, a strong relation emerged. The correlations derived in Section 7.4 were highly significant. Further analysis with a large number of images and field samples will be required to determine whether the relation was unique to this set of images or whether it is universal; additionally, the relation may not be as strong in other areas or at different points in the growing season. Ulaby et al. (1982) suggest that steeper incidence angles are preferable for soil moisture mapping but the results presented here suggest that incidence angles of approximately 30, 35 and 40 degrees give the least bias in the OH92, OH04 and DU95 models, respectively. Additionally, the correlation between bias and incidence angle was negative, with bias decreasing with increasing incidence angle and not increasing as would be expected. Unlike Ulaby et al. (1982), Baghdadi et al. (2008) found that

the optimal incidence angles were $< 37\%$, a figure in line with the mid-range optimal angles mentioned above.

Should other analyses confirm this trend in other datasets, the relations would allow for the removal of bias after deriving model estimates, effectively “tying” them with the actual values, albeit with some error. The only issue with this approach is that the distributions of soil moisture values would remain the same, only shifted, which would produce a sharp cutoff either at the start or end of the model’s range of validity. This issue could be prevented by generating the initial estimates with a wider range of validity (i.e. 0 to 45%), then applying the bias correction and masking out values outside the range of validity afterwards. With a bias correction effected, the only remaining influences between two nearby dates would be caused by changes in surface roughness or soil moisture.

A limitation of this relation is that it may be caused by other confounding effects, such as changes in roughness or moisture contents, and only partially be related to incidence angle. Baghdadi and Zribi (2006) suggest that the models are not highly sensitive to incidence angle, compared to their sensitivity to roughness and soil moisture, and further work would be required to determine how much of the variability in bias can be explained by incidence angle when also taking into account the effects of roughness and moisture content.

8.4 Trends in relative error of estimates

The analysis exploring potential relations between relative error and soil properties indicated no significant effects of soil order, texture, drainage, organic matter and soil fractions. The lack of a clear relation is likely related to the coarseness of the data, most sites having similar properties, and to high error in the model estimates

themselves. According to Oevelen (1998) this absence of significant correlation could be linked to a mismatch between the point observation of soil moisture (and relative error of soil moisture estimates) and the spatial coverage of the RS-2 imagery or be attributed to the partitioning of the data by model, date and field sampling depth which reduced the power of the analysis.

However, significant relations were found between relative error and tillage practices. Tillage had a strong effect on the May 5 relative error of estimates for all models, accounting for 16.2, 10.6 and 8.3% of the variance in the OH92, OH04 and DU95 models. Within types of tillage methods, significantly different groups emerged, the groupings appearing to be related the surface roughness produced by the tillage method. For May 5 these groups were i) no-till, fall minimal tillage and fall seeding (lower roughness methods) and ii) fall conventional tillage and spring tillage (more intensive, roughness increasing operations). For the OH04 and DU95 models, relative error was much higher for high-roughness treatments than for low-roughness treatments. Interestingly, the OH92 model had the same groupings, but inverse behaviour, with higher relative error in sites treated with low-roughness tillage methods. For May 16, the effects of tillage on relative error were reduced, but still significant, with only 3.3, 6.8 and 3.5% of the variance in relative error accounted for. Similar groups were found as for May 5, except that the mean relative errors of each method was extremely higher than the earlier date. For this date, however, the behaviour of the OH92 is not reversed, and all models show lower relative error for fall tillage and fall seeding versus higher error for spring tillage independently of whether the sites were rolled or not.

These trends in the influence of tillage methods can be used to assess the confidence levels of estimates in different fields, and to determine likely model performance under different conditions. Repeating this analysis with larger datasets would help

determine if the trends are transportable to other data and locations. The notion that relative error is related to roughness, indirectly, via tillage method is interesting. The fact that some tillage methods or classes of methods have high relative error translates to a large variability in the error of estimates. In the OH04 model on May 16, for example, sites with rolled spring tillage had a mean relative error of 133.8%; this astounding number essentially signifies that soil moisture estimates are on average 134% higher than the “real” value, compared to May 16 OH92 model sites seeded in the fall and having no spring tillage where the mean relative error is only 3.5%, this provides a good estimate of the potential variability of model estimates.

Similarly, the relation between relative error and actual soil moisture conditions suggests that the magnitude of estimate error is much larger when the actual soil moisture conditions are low, but also shows that some models have a larger “band” of soil moisture contents for which relative error is less variable, most notably in the OH92 model.

8.5 Limitations

The main limitation to estimating soil moisture with radar remote sensing is related to the fact that penetration depth varies according to soil moisture content (among other factors). This signifies that the radar return at every pixel is sensing the moisture in a different depth profile (i.e. some pixels may be representative of moisture from 0 to 4 cm and others from 0 to 6, etc.). Thus, the final soil moisture maps, while being relatively representative of overall conditions, do not represent uniform depth across the scene (Boisvert et al., 1997). This makes multirate comparisons somewhat more difficult, especially comparing soil moisture estimates at a per-pixel basis, which could in fact be values for different depths. An additional challenge is presented in

selecting a depth (or depths) for field sampling for validation work. While Ulaby et al. (1982) suggested that steep incidence angles were preferable for soil moisture mapping as they increased penetration depth and decreased the effects of roughness, the relation between bias and incidence angle described in Section 7.4 suggests that steeper incidence angles (and increased penetration depth) may not always be optimal.

Another severe limitation of the three models evaluated is the obvious need for calibration. As shown by the soil moisture maps produced, model bias is not constant from date to date (at least when imagery is acquired at different incidence angles), making it difficult to attach the estimates to absolute values. As discussed earlier, the trends (increases or decreases in moisture) are difficult or impossible to distinguish without the correction of this bias. More work would be required to determine if the relation found between bias and incidence angle holds for other imagery, areas and conditions. If so, this relation could be used to eliminate bias and have the estimates better match the actual conditions.

9 Conclusions

Field soil moisture measurements in 42 fields near Casselman, Ontario, were used to evaluate the performance of the OH92, OH04 and DU95 models.

The overarching finding is that model calibration is absolutely required for the model to perform as anticipated, especially if the imagery is not in the same incidence angle configuration for each date. Differences in incidence angle, and surface roughness made it difficult to understand actual changes in soil moisture from date to date, although the spatial patterns of soil moisture were still visible. Despite the fact that calibration was not used in this project, the findings do provide potential avenues for reducing bias using strong negative correlations between bias and incidence angle for the three models.

The effects of tillage on the relative error of estimates were also significant and were related to surface roughness characteristics, and decreased in strength as the season progressed. Relative errors were higher (for most models and dates) when higher-roughness tillage methods were used.

The optimal filtering and pre-extraction averaging size were found to be a 3 x 3 speckle filter and 7 x 7 pre-extraction average for the OH92 model, a 7 x 7 speckle filter and 7 x 7 pre-extraction averaging for the OH04 model and a 7 x 7 speckle filter and 5 x 5 pre-extraction averaging for the DU95 models. At these optimal filter sizes, RMSE errors comparable to those in the literature were obtained.

In a broader context, the results of this research support the idea that point-scale estimates of soil moisture from SAR imagery are not achievable with the current techniques and imagery (Wagner et al., 2007), and that the characterization of surface roughness remains the most significant issue in the use of the three models evaluated (Verhoest et al., 2008).

Funding

Data were provided by Agriculture and Agri-Food Canada and were originally collected as part of the Government Related Initiatives Program (GRIP) project of the Canadian Space Agency (CSA). The author also received funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

- Álvarez-Mozos, J., Casalí, J., Gonzalez-Audícana, M., and Verhoest, N. E. (2005). Correlation between ground measured soil moisture and RADARSAT-1 derived backscattering coefficient over an agricultural catchment of Navarre (North of Spain). *Biosystems Engineering*, 92(1):119–133.
- Álvarez-Mozos, J., Verhoest, N. E., Larrañaga, A., Casalí, J., and González-Audícana, M. (2009). Influence of surface roughness spatial variability and temporal dynamics on the retrieval of soil moisture from SAR observations. *Sensors*, 9(1):463–489.
- Baghdadi, N., Cerdan, O., Zribi, M., Auzet, V., Darboux, F., Hajj, M. E., and Kheir, R. B. (2008). Operational performance of current synthetic aperture radar sensors in mapping soil surface characteristics in agricultural environments: application to hydrological and erosion modelling. *Hydrological Processes*, 22(1):9–20.
- Baghdadi, N., Holah, N., and Zribi, M. (2006). Soil moisture estimation using multi-incidence and multi-polarization ASAR - data. *International Journal of Remote Sensing*, 27(10):1907–1920.
- Baghdadi, N., Paillou, P., Grandjean, G., Dubois, P., and Davidson, M. (2000). Relationship between profile length and roughness variables for natural surfaces. *International Journal of Remote Sensing*, 21(17):3375–3381.
- Baghdadi, N. and Zribi, M. (2006). Evaluation of radar backscatter models IEM, OH - and Dubois using experimental observations. *International Journal of Remote Sensing*, 27(18):3831–3852.
- Biftu, G. F. and Gan, T. Y. (1999). Retrieving near-surface soil moisture from Radarsat SAR data. *Water Resources Research*, 35(5):1569–1579.

- Bindlish, R. and Barros, A. P. (2001). Parameterization of vegetation backscatter in radar-based, soil moisture estimation. *Remote Sensing of Environment*, 76(1):130–137.
- Boisvert, J. B., Gwyn, Q. H. J., Chanzy, A., Major, D. J., Brisco, B., and Brown, R. J. (1997). Effect of surface soil moisture gradients on modelling radar backscattering from bare fields. *International Journal of Remote Sensing*, 18(1):153–170.
- Bruckler, L., Witono, H., and Stengel, P. (1988). Near surface soil moisture estimation from microwave measurements. *Remote Sensing of Environment*, 26(2):101–121.
- Bryant, R., Moran, M. S., Thoma, D. P., Collins, C. D. H., Skirvin, S., Rahman, M., Slocum, K., Starks, P., Bosch, D., and Dugo, M. P. G. (2007). Measuring surface roughness height to parameterize radar backscatter models for retrieval of surface soil moisture. *IEEE Geoscience and Remote Sensing Letters*, 4(1):137–141.
- Curtis, J. (2001). Moisture effects on the dielectric properties of soils. *IEEE Transactions on Geoscience and Remote Sensing*, 39(1):125–128.
- Dobson, M. and Ulaby, F. (1986). Active microwave soil moisture research. *IEEE Transactions on Geoscience and Remote Sensing*, 24(1):23–36.
- Dubois, P., van Zyl, J., and Engman, T. (1995a). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing*, 33(4):915–926.
- Dubois, P., vanZyl, J., and Engman, T. (1995b). Corrections to "Measuring soil moisture with imaging radars". *IEEE Transactions on Geoscience and Remote Sensing*, 33(6):1340.

- D'Urso, G. and Minacapilli, M. (2006). A semi-empirical approach for surface soil water content estimation from radar data without a-priori information on surface roughness. *Journal of Hydrology*, 321(1-4):297–310.
- Engman, E., Mattikalli, N., Ahuja, L., and Jackson, T. (1996). Application of microwave remotely sensed surface soil moisture for estimation of sub-surface soil hydraulic properties. In *Proceedings of the Geoscience and Remote Sensing Symposium*, volume 2, pages 1058–1060.
- Fung, A. and Chen, K. (2004). An update on the IEM surface backscattering model. *IEEE Geoscience and Remote Sensing Letters*, 1(2):75–77.
- Fung, A., Li, Z., and Chen, K. (1992). Backscattering from a randomly rough dielectric surface. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):356–369.
- Hallikainen, M., Ulaby, F., Dobson, M., El-Rayes, M., and Wu, L. (1985). Microwave dielectric behavior of wet soil - Part 1: Empirical models and experimental observations. *IEEE Transactions on Geoscience and Remote Sensing*, 23(1):25–34.
- Jiankang, J., van der Keur, P., Thomsen, A., and Skriver, H. (1996). Soil moisture retrieval using the Danish L- & C-band polarimetric SAR. In *Geoscience and Remote Sensing Symposium, 1996. IGARSS '96. 'Remote Sensing for a Sustainable Future.'*, *International*, volume 2, pages 1300–1302 vol.2.
- McNairn, H., Champagne, C., Gauthier, E., Jiao, X., Merzouki, A., Pacheco, A., Shang, J., Fitzmaurice, J., Howard, A., Jarvis, I., Yang, Z., Brisco, B., Charbonneau, F., Hawkings, R., Murnaghan, K., Nedelcu, S., and Touzi, R. (2008). Annual report to the Canadian Space Agency Gouvernement Related Initiatives Program (GRIP) - supporting agricultural risk management: An integrated passive-active microwave approach to identify soil moisture extremes. Technical report.

- Merzouki, A., McNairn, H., and Pacheco, A. (in press 2010a). Evaluation of radar backscatter models Dubois, Oh, and IEM over agricultural fields using polarimetric C-band RADARSAT-2 SAR image data. *Canadian Journal of Remote Sensing*.
- Merzouki, A., McNairn, H., and Pacheco, A. (in press, 2010b). Mapping soil moisture using RADARSAT-2 data and local autocorrelation statistics. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- Moran, M. S., Peters-Lidard, C. D., Watts, J. M., and McElroy, S. (2004). Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing*, 30(5):805–826.
- Oevelen, P. J. V. (1998). Soil moisture variability: a comparison between detailed field measurements and remote sensing measurement techniques. *Hydrological Sciences Journal*, 43(4):511–520.
- Oh, Y. (2004). Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3):596–601.
- Oh, Y., Sarabandi, K., and Ulaby, F. (1992). An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):370–381.
- Oh, Y., Sarabandi, K., and Ulaby, F. (1994). An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observation. In *Proceedings of the Geoscience and Remote Sensing Symposium*, volume 3, pages 1582–1584.
- Oh, Y., Sarabandi, K., and Ulaby, F. (2002). Semi-empirical model of the ensemble-

- averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 40(6):1348–1355.
- Pierdicca, N., Pulvirenti, L., and Bignami, C. (2009). Soil moisture estimation over vegetated terrains using multitemporal remote sensing data. *Remote Sensing of Environment*.
- Schaber, G. G. and Breed, C. S. (1999). The importance of SAR wavelength in penetrating blow sand in Northern Arizona. *Remote Sensing of Environment*, 69(2):87–104.
- Shoshany, M., Svoray, T., Curran, P. J., Foody, G. M., and Perevolotsky, A. (2000). The relationship between ERS-2 SAR - backscatter and soil moisture: generalization from a humid to semi-arid transect. *International Journal of Remote Sensing*, 21(11):2337–2343.
- Sokol, J., McNairn, H., and Pultz, T. J. (2004). Case studies demonstrating the hydrological applications of C-band multipolarized and polarimetric SAR. *Canadian Journal of Remote Sensing*, 30(3):470–483.
- Sumantyo, J. T. S. and Amini, J. (2008). A model for removal of speckle noise in SAR images (ALOS PALSAR). *Canadian Journal of Remote Sensing*, 34(6):503–515.
- Thoma, D. P., Moran, M. S., Bryant, R., Rahman, M., Holifield-Collins, C. D., Skirvin, S., Sano, E. E., and Slocum, K. (2006). Comparison of four models to determine surface soil moisture from C-band radar imagery in a sparsely vegetated semiarid landscape. *Water Resources Research*, 42.
- Topp, G. C., Davis, J. L., and Annan, A. P. (1980). Electromagnetic determination

- of soil water content: Measurements in coaxial transmission lines. *Water Resources Research*, 16(3):574–582.
- Ulaby, F., Batlivala, P., and Dobson, M. (1978). Microwave backscatter dependence on surface roughness, soil moisture, and soil texture: Part I-Bare soil. *IEEE Transactions on Geoscience Electronics*, 16(4):286–295.
- Ulaby, F. T., Dubois, P. C., and van Zyl, J. (1996). Radar mapping of surface soil moisture. *Journal of Hydrology*, 184(1-2):57–84.
- Ulaby, F. T., Moore, R. K., and Fung, A. K. (1982). Microwave remote sensing - active and passive. volume II - Radar Remote Sensing and Surface Scattering and Emission Theory, pages 816–921.
- Verhoest, N. E., Lievens, H., Wagner, W., Alvarez-Mozos, J., Moran, M. S., and Mattia, F. (2008). On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. *Sensors*, 8(7):4213–4248.
- Wagner, W., Blöschl, G., Pampaloni, P., Calvet, J., Bizzarri, B., Wigneron, J., and Kerr, Y. (2007). Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. *Nordic Hydrology*, 38(1):1.
- Zribi, M. and Dechambre, M. (2003). A new empirical model to retrieve soil moisture and roughness from C-band radar data. *Remote Sensing of Environment*, 84(1):42–52.

A Script for generating model estimates

Note: the Python code is provided as-is without guarantees under a Creative Commons Attribution-NonCommercial 3.0 Unported License. Other implementations will require modifications and analysis of the scripts as some variables, filenames and file types are hard-coded. Be aware that the “\” backslash character at the end of a line indicates that a command is continued on the following line.

```

import scipy
from scipy import ndimage
import os, subprocess, csv, math
from osgeo import gdal
from gdalconst import *

def generate_inang (pixname):
    # Open the incidence angle image
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(pixname) + ".pix"
    ds = gdal.Open(fn, GA_ReadOnly)
    inangband = ds.GetRasterBand(1)
    xsize = ds.RasterXSize
    ysize = ds.RasterYSize
    inang = inangband.ReadAsArray(0, 0, xsize, ysize)
    ds = None
    inangband = None

    # Export it to ENVI format for PolSARPro
    driver = gdal.GetDriverByName('ENVI')
    fn = str(pixname) + ".bin"
    ds = driver.Create(fn, xsize, ysize, 1, GDT_Float32, \
        ['SUFFIX=ADD', 'INTERLEAVE=BSQ'])
    ds.GetRasterBand(1).WriteArray(inang, 0, 0)
    ds = None
    inang = None

def importpix (pixname):

    # Open the file
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(pixname) + ".pix"
    ds = gdal.Open(fn, GA_ReadOnly)

    # Get the band objects from the dataset
    c11band = ds.GetRasterBand(1)
    c22band = ds.GetRasterBand(2)
    c33band = ds.GetRasterBand(3)

    # Get some metadata
    global xsize, ysize
    xsize = ds.RasterXSize
    ysize = ds.RasterYSize

    # Read the data into Numeric arrays
    c11 = c11band.ReadAsArray(0, 0, xsize, ysize)
    c22 = c22band.ReadAsArray(0, 0, xsize, ysize)
    c33 = c33band.ReadAsArray(0, 0, xsize, ysize)
    c11 = pow(c11, 2)
    c22 = pow(c22, 2)
    c33 = pow(c33, 2)

```

```

# Delete the band and dataset objects
ds = None
c11band = None
c22band = None
c33band = None

# Write the files to ENVI format
driver = gdal.GetDriverByName('ENVI')

ds = driver.Create('c11.bin', xsize, ysize, 1, GDT_Float32,\
['SUFFIX=ADD', 'INTERLEAVE=BSQ'])
ds.GetRasterBand(1).WriteArray(c11, 0, 0)
ds = None

ds = driver.Create('c22.bin', xsize, ysize, 1, GDT_Float32,\
['SUFFIX=ADD', 'INTERLEAVE=BSQ'])
ds.GetRasterBand(1).WriteArray(c22, 0, 0)
ds = None

ds = driver.Create('c33.bin', xsize, ysize, 1, GDT_Float32,\
['SUFFIX=ADD', 'INTERLEAVE=BSQ'])
ds.GetRasterBand(1).WriteArray(c33, 0, 0)
ds = None

# Write the accompanying config.txt file
file = open('config.txt', 'w')
text = "Nrow\n" + str(ysize) + "\n-----\nNcol\n" + \
str(xsize) + "\n-----\nPolarCase\nmnonostatic\n-----\n\
PolarType\nfull\n"
file.write(text)
file = None

def OH92(pixname):

# Run the OH92 module in PolSARpro to generate .bin files
cwd1 = 'C:\Program Files\PolSARpro_v4.1.5\Soft\data_process_sngl'
cmd1 = ['surface_inversion_oh_C3.exe', "G:/imagery/Ortho", \
"G:/imagery/Ortho", "G:/imagery/Ortho/may23_ortho_inang.bin", \
'0', '0', '4280', '4045', '0']
subprocess.Popen(cmd1, shell=True, stdout=subprocess.PIPE, \
cwd=cwd1).wait()

# Write the .bin.hdr file for the mv layer
file = open('oh_er.bin.hdr', 'w')
text = "ENVI\ndescription={\n  File Imported into ENVI.}\nnsamples = " \
+ str(xsize) + "\nlines   = " + str(ysize) + "\nbands   = 1\nheader\
offset = 0\nfile type = ENVI Standard\ndata type = 4\ninterleave = \
bsq\nsensor type = Unknown\nbyte order = 0\nband names = \
{\n oh_mv.bin}\n"
file.write(text)
file = None

```



```

# Export the estimates to PCIDSK format
driver = gdal.GetDriverByName('PCIDSK')
fn = str(pixname) + "_OH04.pix"
ds = driver.Create(fn, xsize, ysize, 1, GDT_Float32)
ds.GetRasterBand(1).WriteArray(mv, 0, 0)
ds = None
mv= None

def DU95(pixname):

# Run the OH92 module in PolSARpro to generate .bin files
cwd1 = 'C:\Program Files\PolSARpro_v4.1.5\Soft\data_process_sngl'
cmd1 = ['surface_inversion_dubois_C3.exe', "G:/imagery/Ortho", \
"G:/imagery/Ortho", "G:/imagery/Ortho/may23_ortho_inang.bin", \
'0', '0', '4280', '4045', '5.404999242769673', '0', '0', '1.0']
subprocess.Popen(cmd1, shell=True, stdout=subprocess.PIPE, \
cwd=cwd1).wait()

# Write the .bin.hdr file for the mv layer
file = open('dubois_er.bin.hdr', 'w')
text = "ENVI\ndescription={\n  File Imported into ENVI.}\nnsamples\
= " + str(xsize) + "\nlines   = " + str(ysize) + "\nbands   = \
1\nheader offset = 0\nfile type = ENVI Standard\nndata type = \
4\ninterleave = bsq\nsensor type = Unknown\nbyte order = 0\nband\
names = {\n dubois_er.bin}\n"
file.write(text)
file = None

# Re-import the ENVI mv file and write to array
driver = gdal.GetDriverByName('ENVI')
ds = gdal.Open('dubois_er.bin', GA_ReadOnly)
erdata = ds.GetRasterBand(1)
er = erdata.ReadAsArray(0, 0, xsize, ysize)
# Conversion using good version of Topp model
mv = (-0.053 + 0.0292*er - 0.00055*pow(er,2) + \
0.0000043*pow(er, 3)) * 100
mv = (mv>0)*1*mv # Filter out negative values
erdata = None
er= None
ds = None

# Export the estimates to PCIDSK format
driver = gdal.GetDriverByName('PCIDSK')
fn = str(pixname) + "_DU95.pix"
ds = driver.Create(fn, xsize, ysize, 1, GDT_Float32)
ds.GetRasterBand(1).WriteArray(mv, 0, 0)
ds = None
mv= None

```

B Script for extracting values, averaging and statistics

Note: the Python code is provided as-is without guarantees under a Creative Commons Attribution-NonCommercial 3.0 Unported License. Other implementations will require modifications and analysis of the scripts as some variables, filenames and file types are hard-coded. Be aware that the “\” backslash character at the end of a line indicates that a command is continued on the following line.

```

from numpy import *
import scipy
from scipy import ndimage
import os, subprocess, csv, math
from osgeo import gdal
from osgeo.gdalconst import *

def samplemv(date, filtsize, model, depth, outfilt):

    # Open the orthorectified version to get the projection information
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(date) + "_ortho_boxcar_" + str(filtsize) + "x" + str(filtsize)\
    + ".pix"
    ds = gdal.Open(fn, GA_ReadOnly)
    xsize = ds.RasterXSize
    ysize = ds.RasterYSize
    transform = ds.GetGeoTransform()
    xOrigin = transform[0]
    yOrigin = transform[3]
    pixelWidth = transform[1]
    pixelHeight = transform[5]
    ds = None

    # Open the .pix file and write to array
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(date) + "_ortho_boxcar_" + str(filtsize) + "x" + str(filtsize)\
    + "_" + str(model) + ".pix"
    ds = gdal.Open(fn, GA_ReadOnly)
    mvdata = ds.GetRasterBand(1)
    xsize = ds.RasterXSize
    ysize = ds.RasterYSize
    mv_obs = mvdata.ReadAsArray(0, 0, xsize, ysize)
    ds = None
    mvdata = None

    # Import x,y values to sample at, then sample
    x= []
    y= []
    name = []
    field = []
    coordsfile = str(date) + "_coords" + str(depth) + "cm.csv"
    coordReader = csv.reader(open(coordsfile), delimiter=',')
    for row in coordReader:
        x.append(float(row[0]))
        y.append(float(row[1]))
        name.append(row[2])
        field.append(float(row[3]))
    coordReader = None

    # Calculate offsets in pixel line format
    xOffset = []

```

```

yOffset = []
for i in range(0, len(x)):
    xOffset.append(int((x[i] - xOrigin) / pixelWidth))
    yOffset.append(int((y[i] - yOrigin) / pixelHeight))

observed = list(nan_to_num(abs(mv_obs[yOffset,xOffset])))

# Take a x by x average when extracting
if outfilt > 1:
    kernelavg = []
    for i in range(0, len(observed)):
        yStart = yOffset[i] - ((outfilt-1)/2)
        yEnd = yOffset[i] + ((outfilt-1)/2) + 1)
        xStart = xOffset[i] - ((outfilt-1)/2)
        xEnd = xOffset[i] + ((outfilt-1)/2) + 1)
        kernelsum = sum(mv_obs[yStart:yEnd, xStart:xEnd])
        kernelnum = sum((mv_obs[yStart:yEnd, xStart:xEnd]>0)*1)
        if kernelsum/kernelnum > 0:
            kernelavg.append(kernelsum/kernelnum)
        else:
            kernelavg.append(0)
    observed = kernelavg

# Eliminate zero rows
while observed.count(0)>0:
    i = observed.index(0)
    del observed[i]
    del field[i]
    del name[i]
    del xOffset[i]
    del yOffset[i]

#Calculate RMSE
diff = array(field) - array(observed)
rmse = math.sqrt(sum(pow(diff,2))/len(field))

# Calculate MBE
mbe = sum(diff)/len(field)

# Calculate Pearson's r
top1 = len(observed) * sum(array(observed) * array(field))
top2 = sum(array(observed)) * sum(array(field))
bottom1 = (len(observed) * sum(pow(array(observed), 2)) -\
pow(sum(array(observed)), 2))
bottom2 = (len(field) * sum(pow(array(field), 2)) -\
pow(sum(array(field)), 2))
r = (top1 - top2) / (math.sqrt(bottom1) * math.sqrt(bottom2))

# Write out a .csv file with the locations and values
#fn = str(date) + "_ortho_boxcar_" + str(filtsize) + "x" +\
str(filtsize) + "_" + str(model) + "_" + str(depth) + "cm_values_"\
+ str(outfilt) + "x" + str(outfilt) + "avg.csv"

```

```

#file = open(fn, 'w')
#firstrow = "name, x, y, field, estimate\n"
#file.write(firstrow)
#for i in range(0, len(observed)):
#    row = str(name[i]) + ", " + str(xOffset[i]+1) + ", " + \
str(yOffset[i]+1) + ", " + str(field[i]) + ", " + str(observed[i]) + "\n"
#    file.write(row)
#file = None

return rmse, mbe, r, len(observed), outfilt

# Extract the stats
fn = "allstats.csv"
file = open(fn, 'w')
firstrow = "date, filter_size, model, depth, avg_out_filter, rmse,\
mbe, r, n\n"
file.write(firstrow)

for date in ['may05', 'may16', 'may23']:
    for filtsize in ['3', '5', '7', '9', '11', '13', '15', '17', '19', '21']:
        for model in ['OH92', 'OH04', 'DU95']:
            for depth in ['3', '6']:
                for outfilt in [1,3,5,7,9,11,13,15,17,19,21]:
                    rmse, mbe, pearsonsr, numval, outfilt = \
samplemv(date, filtsize, model, depth, outfilt)
                    row = str(date) + ", " + str(filtsize) + ", " + \
str(model) + ", " + str(depth) + " cm, " + str(outfilt)\
+ ", " + str(rmse) + ", " + str(mbe) + ", " + str(pearsonsr)
                    file.write(row)

file = None

```

C Script for pre-extraction averaging of entire map

Note: the Python code is provided as-is without guarantees under a Creative Commons Attribution-NonCommercial 3.0 Unported License. Other implementations will require modifications and analysis of the scripts as some variables, filenames and file types are hard-coded. Be aware that the “\” backslash character at the end of a line indicates that a command is continued on the following line.

```

from numpy import *
import scipy
from scipy import ndimage
import os, subprocess, csv, math
from osgeo import gdal
from osgeo.gdalconst import *

def postfilter(date, filtsize, model, outfilt):
    # Open the .pix file and write to array
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(date) + "_ortho_boxcar_" + str(filtsize) + "x" + str(filtsize)\
    + "_" + str(model) + ".pix"
    ds = gdal.Open(fn, GA_ReadOnly)
    mvdata = ds.GetRasterBand(1)
    xsize = ds.RasterXSize
    ysize = ds.RasterYSize
    global mv_obs
    mv_obs = mvdata.ReadAsArray(0, 0, xsize, ysize)
    ds = None
    mvdata = None

    # Take the average
    def myfilt(invalues):
        if (sum(invalues) == 0) or (sum((invalues > 0)*1) == 0):
            return 0
        else:
            return sum(invalues) / sum((invalues > 0)*1)
    mv_filt = ndimage.generic_filter(mv_obs, myfilt, size=outfilt)

    # Write the file
    driver = gdal.GetDriverByName('PCIDSK')
    fn = str(date) + "_ortho_boxcar_" + str(filtsize) + "x" + str(filtsize)\
    + "_" + str(model) + "_" + str(outfilt) + "x" + str(outfilt) + "outfilt.pi"
    ds = driver.Create(fn, xsize, ysize, 1, GDT_Float32)
    ds.GetRasterBand(1).WriteArray(mv_filt, 0, 0)
    ds = None

postfilter('may05', '3', 'OH92', 7)
postfilter('may16', '3', 'OH92', 7)
postfilter('may23', '3', 'OH92', 7)

postfilter('may05', '7', 'OH04', 7)
postfilter('may16', '7', 'OH04', 7)
postfilter('may23', '7', 'OH04', 7)

postfilter('may05', '7', 'DU95', 5)
postfilter('may16', '7', 'DU95', 5)
postfilter('may23', '7', 'DU95', 5)

```