Geoderma Regional 1 (2014) 31–47

Contents lists available at ScienceDirect

Geoderma Regional

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

A model for the identification of terrons in the Lower Hunter Valley, Australia

Brendan P. Malone *, Philip Hughes, Alex B. McBratney, Budiman Minasny

Faculty of Agriculture and Environment, The University of Sydney, Room 115, Biomedical Building, Australia Technology Park, Eveleigh, NSW 2015, Australia

ARTICLE INFO

Article history:
Received 28 March 2014
Received in revised form 1 August 2014
Accepted 4 August 2014
Available online 8 August 2014

Keywords:
Terroir
Terron
Digital soil mapping
Wine region
Soil-landscape characteristics
WRB
Luvisol
Calcisol
Gleysol
Acrisol
Cambisol
Regosol

ABSTRACT

Terroir in the Lower Hunter Valley – a prominent wine-producing region of Australia – is an appealing concept for landowners both in terms of enabling strategic land management and for positioning businesses favourably in a competitive consumer market. To those ends, preliminary steps are made in this study to establish common soil and landscape entities in the area using the terron concept that was proposed by Carré and McBratney (2005). Here soil and landscape variables were assembled and then used to generate 12 terrons or continuous soil-landscape units which are described quantitatively in terms of their distinguishing characteristics. Each terron is characterised by landscape variables (derived from a digital elevation model) and soil variables which include soil pH, clay percentage, soil mineralogy (clay types and presence of iron oxides), continuous soil classes, and presence or absence of marl. This study demonstrates a number of soil inferential techniques used for assembling the soil terron variables, based on common or easy to measure soil properties that populate most soil information databases. The approach is the first step in describing terroir; the next step will be to match the terrons with different grape varietals.

1. Introduction

Terron, a soil and landscape concept, with an associated model, was described by Carre and McBratney (2005). The embodiment of terron is a continuous soil-landscape unit or class which combines soil knowledge, landscape information, and their interactions together. In some ways comparable to agroecozones (Liu and Samal, 2002), land systems (Chrisitan and Stewart, 1953), and terroir (Barham, 2003), terrons (plural) can be described in terms of their soil and landscape attributes. A distinguishing characteristic of the terron concept, however, is that it has an underlying model, which is also based on the continuity of soil cover and landscape (Carre and McBratney, 2005). Thus terrons are objectively realised, are observed as continuous entities across the landscape, and their defining attributes can be described quantitatively. This objective and quantitative evaluation of soil and landscape is therefore quite amenable for environmental assessment, such as determining land capability (or land evaluation) (Stewart, 1968), assessing enterprise suitability (Kidd et al., 2012), and defining terroir-like zones of management (Taylor, 2004).

Given an available soil information database, and information regarding the landscape and digital modelling approaches, this study intends to explore the terron concept further as an initiating effort to indentifying terroir in the Lower Hunter Valley — a famous viticultural and wine producing region of Australia.

A value-laden concept (Vaudour, 2002), terroir by definition refers to an area or terrain, usually small, whose soil and micro-climate impart distinctive qualities to food products (Wilson, 1998; Macqueen and Meinert, 2006; Dougherty, 2012). Terroir is usually associated with the production of wine, but is just as applicable to other agricultural domains and their related food products (Barham, 2003). Because the Lower Hunter Valley is situated in a predominantly viticultural zone, the concept is terroir has considerable appeal. From a land management perspective, a terroir (or terron map in this case) map may be easier to interpret and more useful than just a soil map alone (Carre and McBratney, 2005). Additionally, terroir can be used for the determination of different agricultural areas or zones, and their suitability to produce or grow a given agricultural enterprise. In the case of wine production, identifying terroir could be reduced down to defining areas where viticulture can be practiced, what varieties of grape can be grown and where, and what wine making styles are appropriate (Unwin, 2012). From a business perspective such as the advertising and labelling of food and wine products, the identification with and belongingness to a

* Corresponding author.
E-mail address: brendan.malone@sydney.edu.au (B.P. Malone).

http://dx.doi.org/10.1016/j.geodrs.2014.08.001
2352-0094/

particular terroir could be an attractive marketing tool. Here the importance of terroir is about establishing a point of difference or positioning the region or a cooperative of primary producers favourably in a competitive and discerning marketplace (Feagan, 2007).

The establishment of terroir, however, is more than some sort of classification based on what Carey et al. (2002) termed stable natural factors—soils, geology, landforms and climate. And this is apparent in Bonfante et al. (2011) and Carey et al. (2009) who used mechanistic and empirical approaches respectively by also taking into account dynamic terroir variables such as crop variety, crop micro-climate, water availability, and oenological factors. Nonetheless, landscape classification into relatively homogeneous areas of soil, landform, geology, climate combinations and their interactions is generally a first investigative step to identifying terroir. Depending on the spatial extent of the study, this sort of classification could default to the realm of precision management and the defining of site-specific management zones (Taylor, 2004) if one is considering terroir within a field or a vineyard for example (Green, 2012). The availability of high resolution environmental information gained from proximal sensing systems such as finely resolved elevation models from an on-the-go GPS tracking system, electromagnetic induction, and gamma radiometrics, makes the realisation of site-specific management zones possible (Bramley and Hamilton, 2007).

Contrastingly, at larger spatial extents, the luxury of very detailed information is more limiting. Yet there have been many improvements in digital technologies over time, and will likely to continue into the future, particularly that of digital elevation models. There are also new developments in remote sensing, airborne gamma radiometrics, and climatic information that could assist in the identification of preliminary terroir at large spatial extents, such as that of a landscape, region or district.

A notable omission from these available data sources is spatial soil information. Soil information may be available per se, but its scale and how it is depicted may not be suitable or properly reflect the natural variations of the soil under study. In previous efforts, conventional polygon-based soil maps and modal map unit soil information have been included into the landscape classification process for identifying preliminary terroir or suitability thereof e.g. Carey et al. (2008) and Jones et al. (2004). While a pragmatic solution, with new information technologies, there is a lot that could be done in this space in terms of using digital soil mapping approaches as a complimentary tool to identify terroir at larger spatial extents.

Digital soil mapping is embodied within the terron concept. Carre and McBratney (2005) developed 18 terrons in the La Rochelle area of France using quantitative soil attribute information extracted from a large database, and landform variables derived from a digital elevation model. The soil variables included attributes such as colour, organic matter concentration, texture, and others. Their method involved non-hierarchical clustering methods to define a fixed number of terrons from a given number of observed sites. Terron-landform rules were then determined from a regression kriging analysis, which was followed by interpolation of these rules to create a spatial map of terrons for their study area.

Given environmental variables, an important question in the process of identifying terrons is what types of soil attributes should be included into a terron model? In terms of viticulture, but also agriculture in general, White et al. (2007), Burns (2012), and similarly Lanyon et al. (2004) provided much detail on what soil properties might be considered. A main consideration is assessing first what soil information is available within a given database. Intuitively, there are considerations about soil physical and drainage properties which are generally not directly available, but could be estimated or inferred from data that is available via pedotransfer functions (e.g. Taylor and Minasny, 2006 or McBratney et al., 2002). Considerations about soil nutrients are plausible, as is soil mineralogy. Generally for soil mineralogy, this type of information is sparse and can be costly to determine via conventional methods of XRD analysis. With Vis-NIR technology however, there is a scope to make intelligible inference of clay minerals and iron oxides from soil spectra (Viscarra Rossel et al., 2009; Mulder et al., 2013).

The first aim of this study is to use digital soil mapping methods to create soil maps for inclusion into a terron model. This involves extracting information from an available soil and spectral soil database, and using these data directly or indirectly to map soil attributes of interest that would generally be useful for the preliminary identification of terroir. The second aim is to then assemble the portfolio of the soil and other landscape variables (such as that derived from a digital elevation model) with the intention of deriving continuous terron classes through the use of a non-hierarchical fuzzy clustering algorithm.

2. Materials and methods

2.1. Study area

The area of this study is the Hunter Wine Country Private Irrigation District (HWCPID), situated in the Lower Hunter Valley, NSW (32.83°S 151.35°E), and covers an area of approximately 220 km^2. The HWCPID is approximately 140 km north of Sydney, NSW, Australia (Fig. 1). Climatically, the HWCPID is situated in a temperate climatic zone, and experiences warm humid summers, and relatively cooler yet also humid winters. Rainfall is mostly uniformly distributed throughout the year. On average the HWCPID receives just over 750 mm of rainfall annually (Bureau of Meterology, 2014). Topographically, this area consists mostly of undulating hills that ascend to low mountains to the south-west. The underlying geology of the HWCPID includes predominantly Early Permian siltstones, marl, and some minor sandstone (Hawley et al., 1995). Other extensive parent material includes Late Permian siltstones, and Middle Permian conglomerates, sandstones and siltstones. As to be indicated below, the soils, or composition thereof is quite variable, but in general terms are weathered kaolinitic–smectitic type soils, ranging from light to medium texture grade. In terms of landuse, an expansive viticultural industry is situated in the area and is possibly most widespread of rural industries, followed by dryland agricultural grazing systems.

The workflow of this study involves three main tasks: 1) assemblage of terron variables, 2) identification of terrons (continuous soil and landscape classification), and 3) terron description.

2.2. Assemblage of terron variables: summary

The model for the identification of terrons in the HWCPID is based entirely on soil and landform variables for this study. This decision was made under the consideration that available data indicate that climate is more-or-less uniform across the area. Similarly available geological information or spatial information regarding variations in parent materials, such as gamma-radiometric data, were unsuitable, as detailed surveys have not been conducted yet.

As an initial step, a decision is required to determine an appropriate number of terron classes to generate within the study area. This could be determined by any number of means — objectively and/or subjectively, but in the case of this study, and similarly in Carre and McBratney (2005), it was decided to have parity between the number of soil types that have been observed (during successive soil surveys of the area), and the number of terrons to generate. In nearly all cases, soils in the study area have been observed to the sub-order level of the Australian Soil Classification (Isbell, 2002). Given some consolidation of infrequently observed soil classes (to the sub-order level) into more taxonomically broader classes, there are 12 'main' observed soil types—and by default, 12 terrons to generate.

The assemblage of terron variables were generated either from a digital elevation model (DEM) and its derivatives directly, or from digital soil mapping efforts using available soil data (points). Seventeen soil and landform variables were used in this study. As described below in more detail, 16 of the variables were apportioned for non-hierarchical cluster analysis. As the remaining variable, the presence/absence of marl was used as a stratifying variable to which separate clustering processes were performed. More formally, two terrons were generated

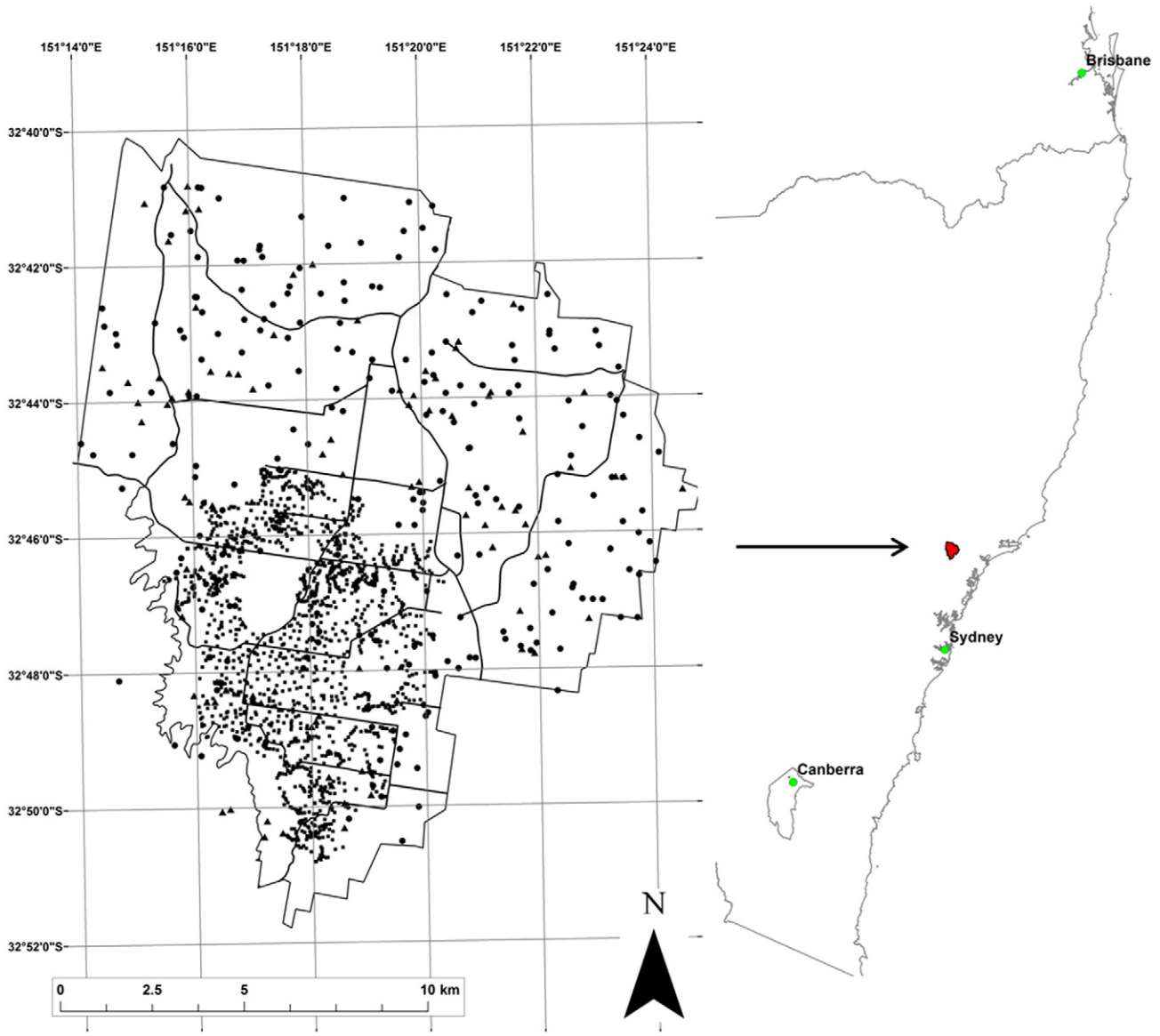

Fig. 1. Locality map of the HWCPID in reference to the eastern coastline of Australia and associated capital cities. Locality map shows existing soil survey sites conducted throughout the HWCPID, beginning 2001 through to 2012. Lines on locality map indicate road network and HWCPID boundary.

for areas where the presence of marl was predicted, and 10 terrons generated where marl is believed to be absent. Marl in the HWCPID typically occurs as loose, earthy deposits (indurated marine deposits from the Permian) consisting chiefly of an intimate mixture of clay and calcium carbonate. Subsequently, the soils where marl is observed are naturally described as calcareous, that from a viticultural perspective, are highly valued the world over (White et al., 2007).

For the landform variables, a 25 m resolved DEM was used. This is currently the finest resolution DEM available for the study area. From the DEM the following 8 variables were extracted:

1) Elevation (ELV): The given elevation from the DEM
2) Slope angle (SLP): Measured in degrees, is the first derivative of ELV in the direction of greatest slope
3) Topographic wetness index (TWI): A secondary landform parameter which estimates for each pixel, its tendency to accumulate water (Quinn et al., 1995).
4) Multi-resolution valley bottom flatness index (MVF): Is derived using slope and elevation to classify valley bottoms as flat, low areas (Gallant and Dowling, 2003). This is accomplished through a series of neighbourhood operations at progressively coarser resolutions with the goal of identifying both small and large valleys. MVF has been used extensively for the delineation and grading of valley floor units corresponding to areas of alluvial and colluvial deposits.
5) Vertical distance to channel network (VDC): Difference between ELV and an interpolation of a channel network base level elevation. Knowledge of the spatial distribution of channel networks (lines) is therefore necessary for this parameter.
6) Mid-slope position (MSP): A relative slope position parameter which gives a classification of the slope position in both valley and crest positions.
7) Diffuse incoming solar radiation (ISR): Measure of potential incoming solar radiation, and used as a parameter for evaluating the positional aspect effect. This parameter was evaluated over the duration of a single calendar year with a 5 day time step.
8) Catchment area (CTA): A parameter which describes the amount of potential water that will drain through a pixel, and derived from the ELV using theoretical flow direction algorithms.

Algorithms for deriving these variables were implemented using the SAGA GIS software (http://www.saga-gis.org/en/index.html).

Eight soil variables (besides the presence of marl) used in this study included:

1) The whole-soil averaged pH (to 1 m) (Bishop et al., 1999)
2) The average sub-soil clay percentage (0.4–1 m)
3) Soil drainage potential index (Malone et al., 2012)
4) Soil mineralogy, specifically:
   4.1. Likelihood of kaolinite presence
   4.2. Likelihood of the presence of a kaolinite and smectite mixture
   4.3. Likelihood of iron oxide presence
5) Classified soil classes, specifically:
   5.1. 1st principal component of continuous soil class predictions
   5.2. 2nd principal component of continuous soil class predictions.

These soil variables were selected for two main reasons: 1) to achieve a reasonable cross-section of attributes describing the physical, chemical and mineralogical properties of the soil, and 2) to use most effectively the data that are stored within the available soil information database. Described in more detail below, the database used in this study has many observations of soil classes, soil pH, clay percentage and colour, with limited observations of other usually commonly observed attributes such as soil organic carbon and exchangeable cations. Thus the intention overall was to make the most use out of what data was 'plentiful'. In terms of the soil mineralogical variables, Generally this information is scarce in soil databases, but because a number of the soils examined have vis-NIR spectra, some estimate that their mineral compositions could be approximated from these, for which is also described in more detail below.

In terms of available soil data for the calibration of spatial soil prediction functions, there is, in total, 1691 individual locations where soil has been observed and described in one form or another (Fig. 1). These data have been collected over successive years beginning in 2001 to the present time, and more detailed descriptions of them can be found in Malone et al. (2011) and Odgers et al. (2011). Essentially, all these data are compiled into a single database table which contains separate columns for different soil attributes (in addition to columns for labels, coordinates, soil depths and horizon nomenclature, etc.). Each row is an observation made for a particular location at a particular genetic horizon, or sometimes a specified depth interval, and in many cases, the individual observations have a measured Vis-NIR spectrum as well. Yet this database is by no means fully populated with a full complement of measured soil attributes or soil spectra at every sample. Therefore, for each of the soil terron variables of interest in this study, filtering steps were performed to remove missing data (instances where no data was recorded for a whole profile), and other data anomalies such as missing spatial coordinates.

The soil variables were mapped using the digital soil mapping *scorpan* approach (McBratney et al., 2003):

$$S_c[x,y,\sim t] \text{ or } S_p[x,y,\sim t] = f(s[x,y,\sim t], c[x,y,\sim t],\ o[x,y,\sim t], r[x,y,\sim t],\ \ p[x,y,\sim t], a[x,y,\sim t],\ n) \quad (1)$$

where:

S_c	soil class
S_p	soil property
s	soils, other attributes of the soil at a point
c	climate, climatic properties of the environment at a point
o	organisms, vegetation, or fauna, or human activity
r	topography, landscape attributes
p	parent material, lithology
a	age, the time factor
n	space, spatial position
t	time (where t is defined as an approximate time)
x,y	the explicit spatial coordinates
f	function or soil spatial prediction function (SSPF)

2.3. Assemblage of terron variables: specific description of soil variable mapping

The soil terron variables are mapped in Figs. 2 and 3. Following now is the description of the pertinent information regarding how these variables were mapped across the HWCPID to 25 m resolution (the same resolution and extent as for the landform variables). This information is also summarised in Table 1.

2.3.1. Estimation of marl presence and whole-soil pH average

After data filtering, 1399 soil profiles remained that contained measurements of soil pH (1:5 soil:water suspension) (Rayment and Higginson, 1992). These measurements, taken at various times, were either made directly in the field at the time of soil description or later in a laboratory. In order to proceed with DSM, a continuous mass-preserving spline depth function (Bishop et al., 1999) was fitted to each soil pH profile. This function was then integrated to return the average for the 0–1 m depth interval. If a soil profile was not observed to reach beyond 1 m, the average for the whole profile was recorded.

Regression kriging was used as the spatial prediction model for mapping. Data were split into calibration and validation datasets (70% and 30% respectively). Cubist regression tree models (Quinlan, 1992) were used for regressing the averaged soil pH values against a suite of co-located environmental covariates. Some of the covariates have been described above, in addition to further terrain variables plus other environmental covariates such as normalised difference vegetation index (NDVI) and other spectral combinations or ratios derived from Landsat ETM+ imagery e.g. Band 3/Band 2, Band 3/Band 7, and Band 5/Band 7 were used. Further information about soil enhancement ratios can be found in Boettinger et al. (2008). The Landsat ETM+ imagery used in this work was collected during the southern hemisphere spring (September) of 2009, which coincides with a time where rainfall is at its long-term low, and when exposed soil cover is expected to be greatest.

Variography of the residuals (from the Cubist modelling) was used to examine their spatial correlation structure. Given the number of data available, it was deemed suitable to use locally fitted variograms based on the nearest 200 points for spatial interpolation of the residuals across the HWCPID. As a final step to regression kriging, the two maps—the one derived from regression modelling and the other from residual kriging, were summed (Fig. 2). The validation dataset indicated that the root mean square error of prediction for the soil pH was 0.76 pH units, which is similar to that observed in Malone et al. (2011).

Mapping the presence of marl involved application of expert knowledge, traditional GIS data analysis, and digital soil mapping. From the knowledge developed in the study area, a high soil pH in the sub-soil at a relatively high elevation and in areas that do not accumulate water flow, is generally an indicative criterion for detecting marl presence in a soil. Firstly using the spline fitted depth functions of soil pH at each profile, described previously, integration was performed to retrieve the average soil pH between 0.4 and 1 m depth. Then using this data and their collocated values of ELV and TWI, each soil profile was assessed for the presence or absence of marl. From some iteration, the criteria ultimately used was averaged sub-soil pH > 6.85, ELV > 90 m, and TWI > 12. The threshold criteria were not optimised; instead, a number of realistic combinations were tried before deciding that the final result – number of profiles detected as having marl present – did not vary that much overall. From this process there were 122 and 1277 soil profiles indicating either the presence or absence of marl respectively. Keeping the calibration and validation data configuration (70/30 random split) used for the spatial prediction of whole-profile soil pH, a binary logistic model was used to regress the presence/

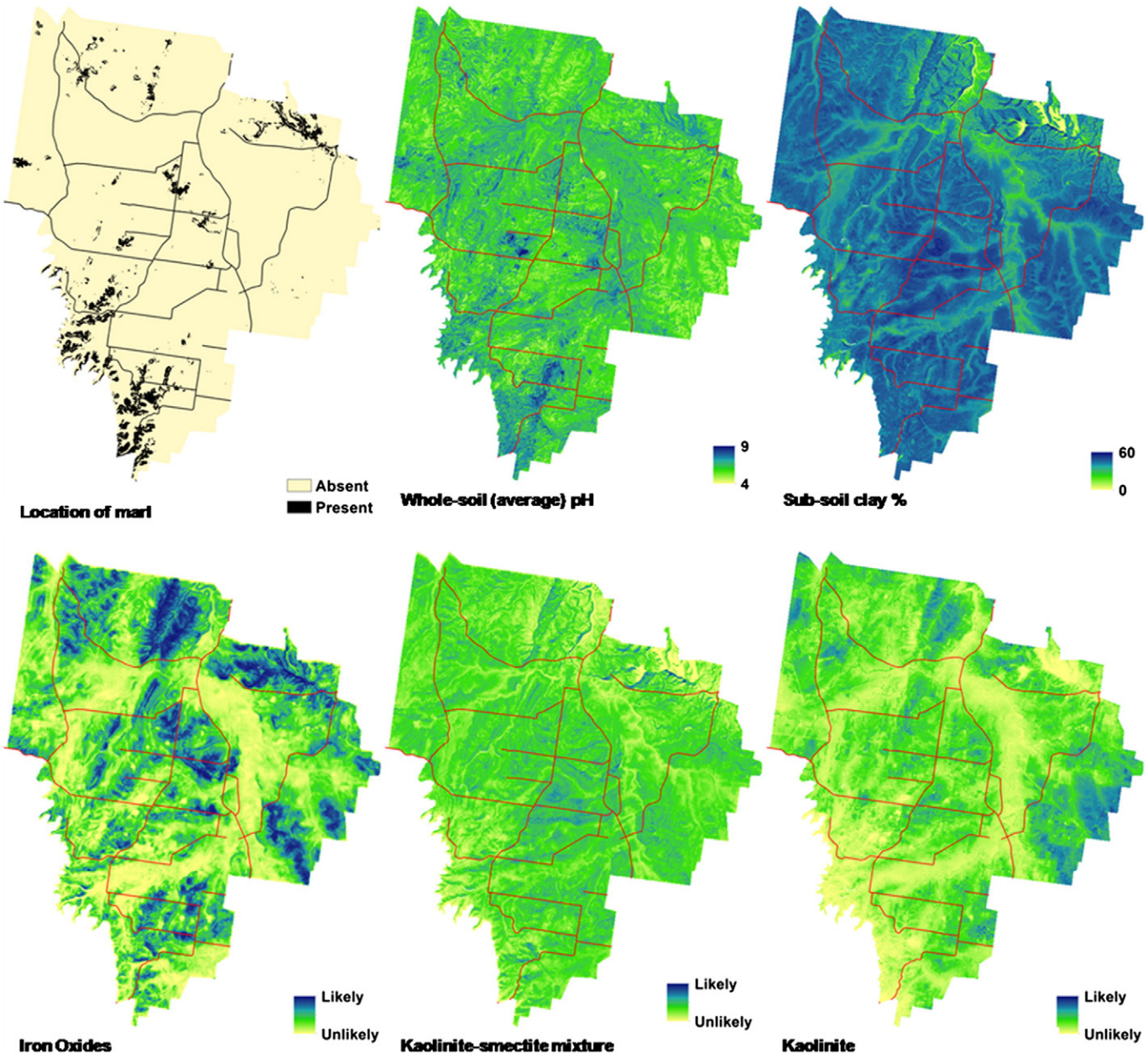

Fig. 2. Maps of soil variables for inclusion into the terron model. Beginning from the top left to right: Presence of marl, whole-soil average pH, sub-soil clay percentage. Bottom left to right: Likeliness of iron oxide presence, likeliness of kaolinite–smectite mixture presence, likeliness of kaolinite presence.

absence of marl data with the aforementioned environmental covariate predictors. From the validation dataset, an overall accuracy of 92% was achieved. Producer's accuracy (error of omission) for predicting the presence of marl was 41%, while the user's accuracy (error of commission) was 44%.

The map showing the spatial distribution of marl presence is shown in Fig. 2. This map was qualitatively validated in a field survey in 2012. From this, the map was found to be reliable in delineating areas of appreciable amounts of marl in the sub-soil.

2.3.2. Sub-soil clay content

After data filtering, 1532 soil profiles remained for which there were observations of soil texture, specifically that of clay content. These data are all field texture observations that had been assigned to texture grades (Northcote, 1979). Consequently, they were converted to clay percentage estimates based on a lookup table adapted from Taylor and Minasny (2006). As with soil pH, a continuous mass-preserving spline depth function was fitted to each profile, from which the average clay content between 0.4 and 1 m was evaluated. Similarly, as with soil pH, regression kriging (with Cubist models) was used to model and spatially predict sub-soil clay content across the HWCPID. 30% of the data was withheld for model validation purposes. The Cubist model was able to explain 20% of the variation in the observed data. Modelling of the spatial structure of the residuals marginally improved the predictions. Using the validation dataset, the RMSEP of the regression kriging predictions was 9%.

2.3.3. Soil drainage index

The method for estimating soil drainage is given in Malone et al. (2012), and its application within the HWCPID is summarised therein as well. Soil drainage is described as an index on a scale of 100 (very well drained) to 0 (very poorly drained), and is determined on the basis of the weighted combination of matrix soil colours for the whole soil profile, excluding the A horizon. It is well established

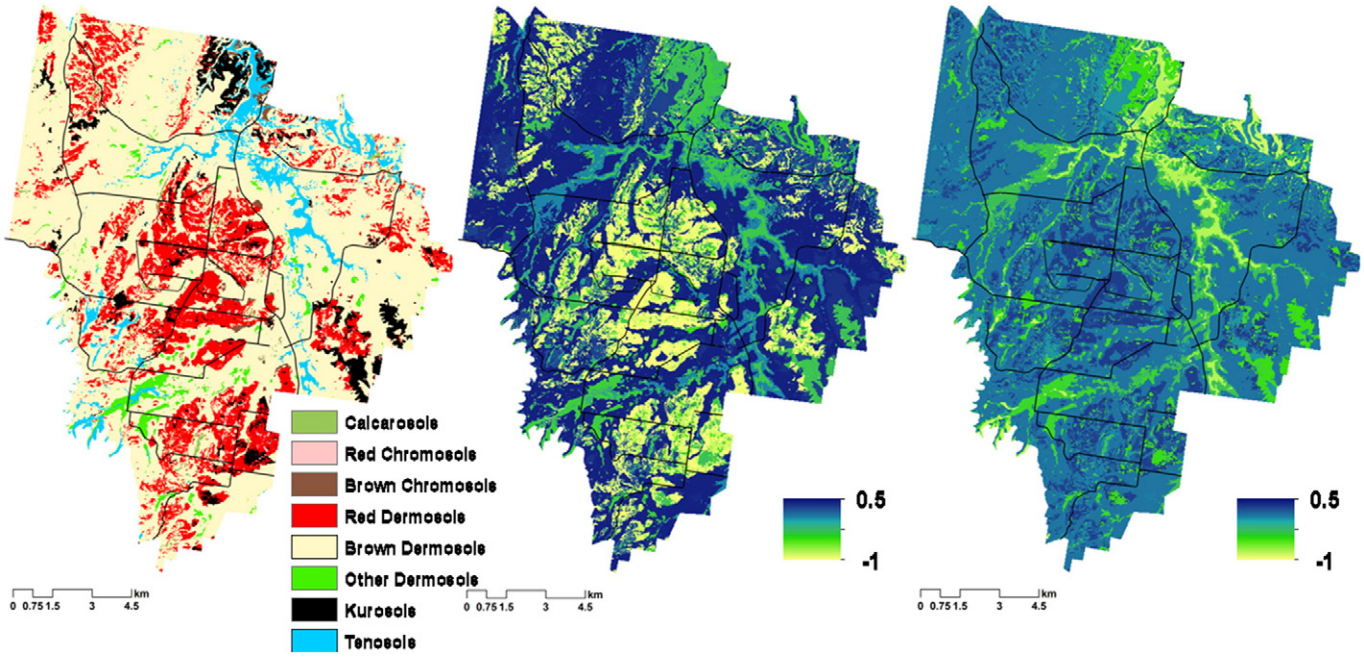

Fig. 3. Soil class map of the HWCPID and associated first two principal component maps of the prediction confidences from See5 modelling.

that soil colour can be used as a general purpose indicator of internal soil drainage (Evans and Franzmeier, 1988). In the HWCPID, a red → brown → yellow → grey → black matrix soil colour sequence is often indicative of soils ranging from well draining (red) to poorly draining (black) (Kovac and Lawrie, 1990).

Because Munsell® HVC is not directly conducive for quantitative studies; the drainage index calculation of Malone et al. (2012) is first initiated by converting soil colour descriptions to the CIELAB colour space (Robertson, 1976; CIE, 1978). Next, archetypal colour/s to each soil of the drainage sequence is ascertained empirically from the available soil colour dataset. A Euclidean distance is calculated to compare each observation with each archetypal soil colour. The distances are then used for a fuzzy classifier to assign each soil a fuzzy membership to each soil colour class. An estimate of soil drainage is then calculated through weighted averaging of the fuzzy memberships and the associated value assigned to each archetypal colour. In this case, the archetypal red is assigned a value of 100, brown (75), yellow (50), grey (25), and black (0).

From Malone et al. (2012), statistical relationships were found between the soil drainage index and a number of terrain variables, notably, the distance to a channel network and topographic wetness index in the HWCPID. Independently validated, the spatial model (regression kriging using Cubist models) of soil drainage index revealed a RMSEP of 22 points. The map in Fig. 2 shows the spatial variation of the soil drainage index across the HWCPID.

2.3.4. Soil mineralogy

Vis-NIR spectroscopy was used for the detection of clay mineral species and iron oxides as an analytical step prior to mapping their spatial distribution across the study area. Clay minerals and iron oxides absorb at specific wavelengths in the vis-NIR range of the electro-magnetic spectrum (Clark et al., 1990). Subsequently, previous studies have demonstrated the use of soil vis-NIR spectra in soil compositional studies with notable success (Viscarra Rossel et al., 2009; Brown et al., 2006). In the present work, 354 soil profiles (1571 soil horizons) have vis-NIR spectra for each of their recorded depths. Spectroscopic measurement was made with an Agrispec portable spectrophotometer with a contact probe attachment (Analytical Spectral Devices, Boulder, Colorado). This particular instrument has a spectral range between 350 and 2500 nm. Soil samples were in air-dry condition and ground to <2 mm prior to scanning. To reduce signal-to-noise ratios of the spectra, five scans of each sample were performed, from which an averaged reflectance spectrum was derived. Calibration of the instrument was made with a Spectralon white tile and was re-calibrated after every 15 scans or 3 samples.

Reference mineral spectra of the following end member clay minerals and iron oxides were retrieved from U.S. Geological Survey digital spectral library (Clark et al., 2007): kaolinite (*KGa-2*), illite (*GDS4*), smectite (*SWy-1*), kaolinite–smectite 50/50 mixture (*H89-FR-2*), goethite (*GDS240*), and hematite (*GDS576*). Clark et al. (2003) document the diagnostic wavelengths of these end member specimens which are summarised in Table 2. The idea behind assessing the mineral composition of soils with vis-NIR spectroscopy is to compare the reflectance of the diagnostic wavelengths from the reference spectra with the reflectance at the same wavelengths of the soil samples. To initiate this, for each of the reference and soil spectra, the wavelength ranges specifically diagnostic to each clay mineral and iron oxide specimen were isolated. Each range was then normalised separately by fitting a convex hull to it, followed by computation of the deviation from the fitted hull (Clark and Roush, 1984). Fig. 4 shows an example of the fitted convex hulls to the diagnostic reflectance ranges from the reference spectra for smectite and kaolinite and the associated spectra once the continuum is removed.

Once all the soil spectra have been normalised, the likely presence of each clay mineral and iron oxide in each soil is estimated when it is compared to that of the normalised reference spectra. The method used in this study is based on the approach developed by the U.S. Geological Survey and implemented in their Tetracorder decision making framework (Clark et al., 2003). Fundamentally, Tetracorder uses a shape-fitting algorithm, which essentially reduces down to a least-squares fit between the reference spectra and the observed soil spectra. The correlation coefficient of this fit (F) is a quantitative estimate of the shape similarity between the reference and soil spectra. The importance of this measure can be exemplified in Fig. 4 where both smectite and kaolinite have similar diagnostic wavelength ranges yet quite different reflectance behaviour—kaolinite with the distinctive doublet feature,

Table 1
Summary of soil terron variables and their inference, method of spatial mapping and associated result of validation using a withheld dataset. SSPF (soil spatial prediction function), RMSE (root mean square error), OA (overall accuracy).

Soil terron variable	Depth	Source data	Nos. of soil profiles	Soil inference	SSPF	Predictors	Validation		Notes
							RMSE	OA	
Marl	Sub-soil (0.4–1 m)	Field and lab pH measurements	1399	Subsoil pH threshold, and thresholds for ALT and TWI	Binary logistic regression	Selected topographic and Landsat ETM + spectral imagery.		92%	
Soil pH	Whole-soil (1 m)	Field and lab pH measurements	1399	Mass-preserving splines	Cubist regression and residual kriging	Terrain variables: (as described in the text) ALT, SLP, TWI, MVF, VDC, MSP, ISR. Further variables included hillshading, multi-resolution ridge top flatness, and terrain ruggedness index	0.76 pH units		
Clay percentage	Sub-soil (0.4–1 m)	Field texture grades	1532	Clay texture lookup table and mass-preserving splines			9%		
Soil drainage	Sub-soil (0.4–1 m)	Moist soil colours	1486	Inferred soil colour-drainage sequence			22 drainage index units		
Likelihood of kaolinite presence	Sub-soil (0.4–1 m)	Vis-NIR spectra	354	Spectral shape matching with reference material at diagnostic wavelengths	Ordinal regression model (3 ordinal classes)	Landsat variables: NDVI, band ratios b3/b2, b3/b7, and b5/b7		80%	Probabilities of class 3 mapped
Likelihood of the presence of a kaolinite and smectite mixture	Sub-soil (0.4–1 m)	Vis-NIR spectra	354					44%	
Likelihood of iron oxide presence	Sub-soil (0.4–1 m)	Vis-NIR spectra	354					75%	
Continuous soil classes	Whole-soil	Field classified soil classes	1679	Soil class consolidation	C5 classification trees			38%	First two principal components of C5 model prediction confidences mapped.

which is absent from smectite. Other criteria used for estimating similarity to the reference spectra are the relative depth and relative area of the soil spectral feature (to the reference). The depth of a spectral feature is calculated by:

$$D = 1 - \frac{R_b}{R_c}. \tag{2}$$

D is the depth of the spectral feature, and R_b is the reflectance value of the raw isolated spectra where the minimum reflectance value of the continuum-removed spectra is observed. R_c is the minimum reflectance in the continuum-removed feature. The relative depth is calculated as the ratio of the spectral feature depth of a soil sample over that of the given reference spectra. This is the same as for relative areas, where areas (of the spectral features) are estimated by the conventional area calculation method. The 'relative abundance' of a clay mineral or iron oxide in a particular soil sample is derived by:

$$M_{ab} = F \times r_D \times r_A \tag{3}$$

where M_{ab} is the relative abundance of a given soil mineral, r_D is the relative depth of the spectral feature for the diagnostic wavelength for a given reference mineral species and, finally, r_A is the relative spectral feature area. In the case where there is more than one diagnostic spectral feature for a clay mineral (illite) or iron oxide (both hematite and goethite), M_{ab} is derived by:

$$M_{ab} = \sum_{i=1}^{n} c_i \times F_i \times r_{D_i} \times r_{A_i} \tag{4}$$

where n is the number of diagnostic spectral features, and c_i is the proportional area of the reference spectral feature i to the total summed area of the (reference mineral) spectral features. Fig. 5 shows that selected continuum removed soil spectra for the diagnostic wavelengths of each mineral species. For each, the minimum, maximum, first, second and third quartiles of M_{ab} were selected, followed by plotting the corresponding continuum removed spectra in association with the reference spectra (bold dark lines Fig. 5).

With some estimate of the likely presence of clay minerals and iron oxide in each soil sample, spatial predictive models were fitted for the intended purpose of generating maps of their spatial variation across the HWCPID, specifically that of the subsoil (0.4–1 m). As for some of the other soil-based terron variables, the mass-preserving spline depth function was used to generate continuous soil profile representations from a limited number of observations. From the fitted splines of each soil profile, for each mineral species, the average abundance for the 0.4–1 m depth interval was derived.

Spatial modelling of the soil mineral species involved ordinal logistic regressions. A preparatory step to this was to derive four ordinal classes of 'mineral abundance' for each soil mineral. Cut points were made at the 0.25 spaced intervals. In other words, the abundances for each class were: Class 1 (0–0.25), Class 2 (0.26–0.50), Class 3 (0.51–0.75), and Class 4 (0.76–1). For most soil mineral species, very few samples had relative abundances in Class 4. Therefore, these few exceptions were collapsed down into Class 3. The occurrence of illite was rare, meaning no attempt was made to build a predictive model of their spatial distribution. Additionally, goethite and hematite have quite similar absorbance features, making their individual identification difficult. The approach in this situation was to combine the estimated 'mineral abundances' of these iron oxides (by taking the maximum), to form an agglomerated iron oxide class for the spatial modelling. In summary the soil minerals used for spatial modelling in this study were kaolinite, smectite, kaolinite–smectite mixture, and iron oxides.

As for the other DSM efforts already discussed, validation of these models was performed via a random-holdback procedure. Here 25% of

Table 2
Diagnostic wavelength ranges of clay minerals and iron oxides investigated in this study.

Mineral	Diagnostic wavelength range/s
Kaolin	2079–2277 nm
Smectite	2118–2287 nm
Kaolin-smectite 50–50 mixture	2128–2258 nm
Illite	2155–2266 nm, 2306–2385 nm
Goethite	457–563 nm, 776–1266 nm
Hematite	455–612 nm, 765–1050 nm

the available data for each model was withheld from the model fitting process.

Specifically in terms of the spatial modelling, ordinal logistic regression is an extension of logistic regression such that regression coefficients are estimated (typically by maximum likelihood estimation) to predict the probability of a discrete outcome (such as a class) occurring (Hastie et al., 2001). This is performed using a logit link function, which is the natural logarithm of the odds (log-odds) of a particular class or event being observed. The proportional odds model (POM) is the most popular model for ordinal logistic regression (Agresti, 1984), and is an extension of logistic regression because it allows ordered data to be modelled by analysing it as a number of dichotomies. For example, a binary logistic regression model compares one dichotomy (e.g. the presence or absence of marl model), whereas the proportional odds model compares a number of dichotomies by arranging the ordered categories into a series of binary comparisons. A binary logistic regression models a single logit; the POM on the other hand models several cumulative logits. Therefore, in the case of having 3 ordinal classes, 2 logits are modelled, one for each of the following class cut points: Class 1 vs. Class 2, Class 3 vs. Class 1, and Class 2 vs. Class 3. From simple inference of these models, the probability of each ordinal class occurring can be made. Agresti (1984) describes the theoretical framework of POMs in greater detail. From validation (n = 89), the overall accuracy of the ordinal models was: iron oxides — 75%, kaolinite — 80%, smectite — 40%, and kaolinite–smectite mixture — 44%. In consideration of the terron model, the 3 most accurately predicted variables for inclusion were selected. Mapping the three soil minerals (sub-soil iron oxides, kaolinite, and kaolinite–smectite mixture) involved mapping the probabilities of

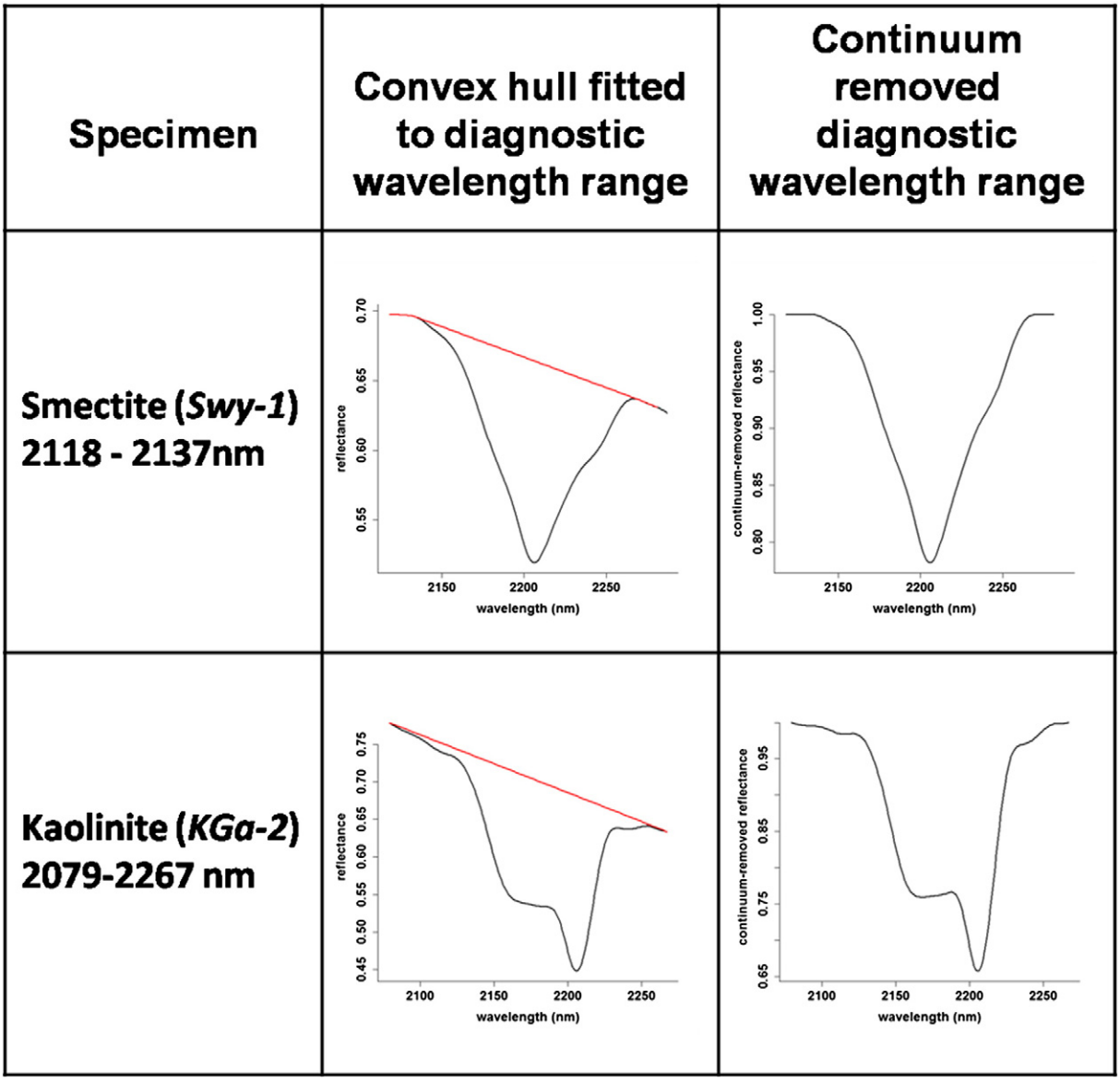

Fig. 4. Diagnostic wavelength ranges for smectite and kaolinite. To prepare for analyses, spectra need to be normalised by fitting a convex hull. Soil spectra and reference spectra are then compared on the basis of the continuum removed spectra.

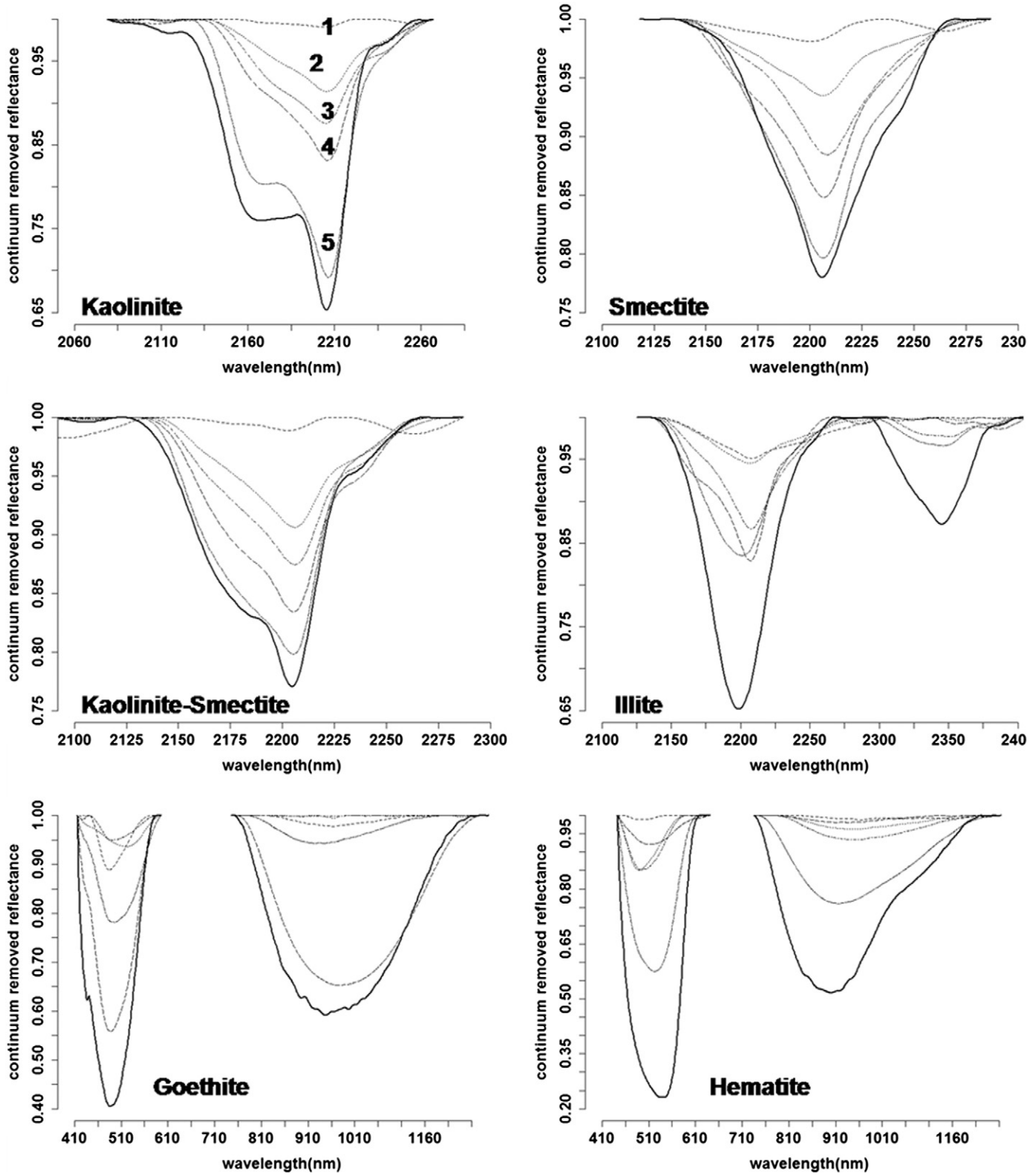

Fig. 5. Continuum removed soil spectra (dashed lines) with comparison to reference spectra at their given diagnostic wavelength ranges. The dashed lines represent (as numbered for kaolinite) the spectra for soils where M_{ab} is 1) the minimum, 2) first quartile, 3) second quartile, 4) third quartile, and 5) maximum of the given M_{ab} distribution for each mineral species.

class 3 for each. The maps are shown in Fig. 2. These indicate that where the probabilities are closer to 1, there is a high likelihood of the presence (in appreciable concentration) of a given soil mineral. Conversely, low probabilities indicate either the presence of a soil mineral in a relatively minor abundance or the absence entirely from the soil.

2.3.5. Soil classes

The inclusion of soil class information into the terron model is reasoned on the fact that soil classes embody much useful information about pedogenesis and soil processes that cannot be satisfactorily or easily be explained by measurable soil attributes alone. 1679 soil

profiles in the available database have been classified to the sub-order level of the Australian Soil Classification System (Isbell, 2002). A general guide on their World Reference Base (WRB) (IUSS Working Group WRB, 2007) and Soil Taxonomy (Soil Survey Staff, 2010) counterparts can be found in Morand (2013) or Isbell et al. (1997). At the Sub-Order level of the Australian Soil Classification, there have been 53 different soil types observed in the HWCPID. However 72% or 1201 are one of either four soils: Brown Dermosol or Chromosol, or Red Dermosol or Chromosol (WRB Luvisols or Lixisols). As may be surmised, many of the soil types have been observed once or only a few times. Therefore to facilitate soil class modelling, soil classes were consolidated into the following generalised soil types (their proportion of the total observed population in brackets): Calcarosols (Calcisols) (2%), Red Chromosols (8%), Brown Chromosols (12%), other Chromosols (2%), Brown Dermosols (26%), Red Dermosols (26%), other Dermosols (7%), Hydrosols (Gleysols, Stagnosols, Fluvisols) (2%), Red Kurosols (Acrisols) (5%), other Kurosols (Acrisols) (4%), Tenosols (Regosols, Cambisols) (4%), and Kandosols (Cambisols) (2%).

After intersecting these observations with a suite of environmental covariates, followed by splitting the data into calibration–validation datasets (75%–25%), Quinlan's C5.0 algorithm (Quinlan, 1993) was used to fit a classification tree model to the calibration data. As described in Moran and Bui (2002), classification trees (or the algorithm that builds them) partition a dataset into successionally more homogeneous subsets. Nodes are where trees branch or split the dataset; terminal nodes are called leaves. At each node of the tree, C5.0 splits the data in such a way that most effectively results in subsets enriched in one class or another, and is more formally expressed as the normalised information gain or difference in entropy (Quinlan, 1993). The attribute with the highest normalised information gain is chosen to make the decision. Validation of the C5.0 model fitted to the soil class data resulted in an overall accuracy of 38%. The kappa (k) statistic for this model was of fair agreement ($k = 0.2$).

The C5.0 model was applied for the whole study area, resulting in the map of Fig. 5 which displays the soil classes as Australian soil classes. Possibly due to the relative proportions of each soil class, four of the soil classes were not predicted using this C5.0 model. As can be seen, it is quite clear of the dominance of Brown and Red Dermosols across the HWCPID. This map however, because of its categorical nature is not suitable to be included into the terron model, which can only permit continuous variables in its current form. To avert this issue, the prediction confidence values that are determined from the C5.0 model were investigated. The confidence values of each class are given for every case or prediction location, and range between 0 (no confidence) and 1 (undoubtable confidence), and sum to unity across all classes. If a case is classified by a single leaf of a decision tree, the confidence value is the proportion of training cases at that leaf that belong to the predicted class (pers. comm. R. Quinlan). If more than one leaf is involved (i.e., one or more of the attributes on the path has missing values), the value is a weighted sum of the individual leaves' confidences. Principal component analysis was used to reduce the variable dimension of the class confidences (12) to just a few (2) main variables. The first two principal components captured 82% of the variation in the prediction confidences. These maps are displayed in Fig. 3, and by all accounts display the spatial variation patterns of the predicted soil classes.

2.4. Identification of terrons

Non-hierarchical fuzzy cluster analysis was performed on the fully assembled suite of soil and landform variables. This was performed separately for the areas where marl was predicted as being present and where it was absent. For the area with marl present, the fuzzy k-means algorithm (Bezdek et al., 1984) was used to generate two classes based on the 16 terron variables. Similarly, for the area with marl absent, 10 classes were generated. The fuzzy-k algorithm was implemented through the stand-alone FuzMe software (Minasny and McBratney, 2002). Besides setting the number of specified classes, Mahalanobis distance was used for evaluating distance metrics in the fuzzy-k algorithm – unlike Euclidean distance, the Mahalanobis distance takes into account the correlation between input variables – and a fuzzy exponent value of 1.2 was used. See McBratney and de Gruijter (1992) for a more detailed explanation of the algorithm. The outputs from both separate analyses, namely the membership classes, were re-assembled for mapping the spatial distribution of each terron.

3. Results

The two terrons for where marl was predicted to be present were labelled as HVT_001 and HVT_002. The other 10 terrons were labelled sequentially from HVT_003 to HVT_012. Maps of the terrons and their membership grades are shown in Fig. 6 with black colouring indicating a high membership to a particular terron. By assigning each grid cell to the terron of highest membership (i.e. hardening), a summary of each terron's soil and landscape attributes is provided in Table 3. The quantitative variables are summarised as a range of the 2.5% and 97.5% quantiles. Regarding the soil classes, instead of summarising the principal components of the soil class probabilities, they are summarised according to the most relevant consolidated Australian Soil Classes that are mapped/occur in the given terron. Fig. 7 displays the associated spatial distribution of the hardened terrons across the HWCPID.

As a further aid to interpreting the different terron classes, Ward's method (Ward, 1963) of clustering or agglomerative hierarchical clustering was performed to examine their taxonomic similarities and differences. The dendrogram depicted in Fig. 8 shows the result of this agglomeration, whereby the terron centroids (resultant from the fuzzy clustering procedures) were used for the analysis. To accompany Fig. 8, Table 4 shows to Mahalanobis distance (taxonomic distance matrix) between terrons, and which was used for the hierarchical clustering. Based on 1-way ANOVA tests between terron class and each input terron soil and landscape variable, a significant difference ($P < 0.001$) was observed. But as can be observed from Table 4 and Fig. 8, some terron classes can be easily distinguished, while others share similarities. Fig. 9 shows the score plots of the first two principle components of the terron input variables. The scores of each terron class are coloured separately. The first principle components explain 38% of the whole data variation while the second component explained a further 18%. In the principle component space it is possible to visually distinguish between the terrons, and also detect similarities. For example HVT_001 and HVT_002 occupy similar regions of the principle component space, but together are quite distinct from the other terrons. Similarly, HVT_005 and HVT_008 are similar, which are both quite distinct from either HVT_006 or HVT_007 etc.

3.1. Terron descriptions

Using the information contained in Tables 3 and 4, and Figs. 7, 8 and 9, a summary of the pertinent characteristics of each terron class is provided.

3.1.1. HVT_001

This terron is first of the marl terrons, and appears to be dominantly situated to the south-west of the study area. Soils are dominantly Red and Brown Dermosols. Some Calcarosols are also present which is to be expected given the calcareous nature of some of the parent material here. Subsequently the soils in this terron relative to other terrons have a higher pH. Soil clay content varies considerably in this terron which could be resultant from the nature of the parent material and the variability of the terrain. In terms of landform, this terron is situated in a relatively high elevation position and is quite variable in terms of the slope. Soils are generally moderate to well draining, based on the soil drainage

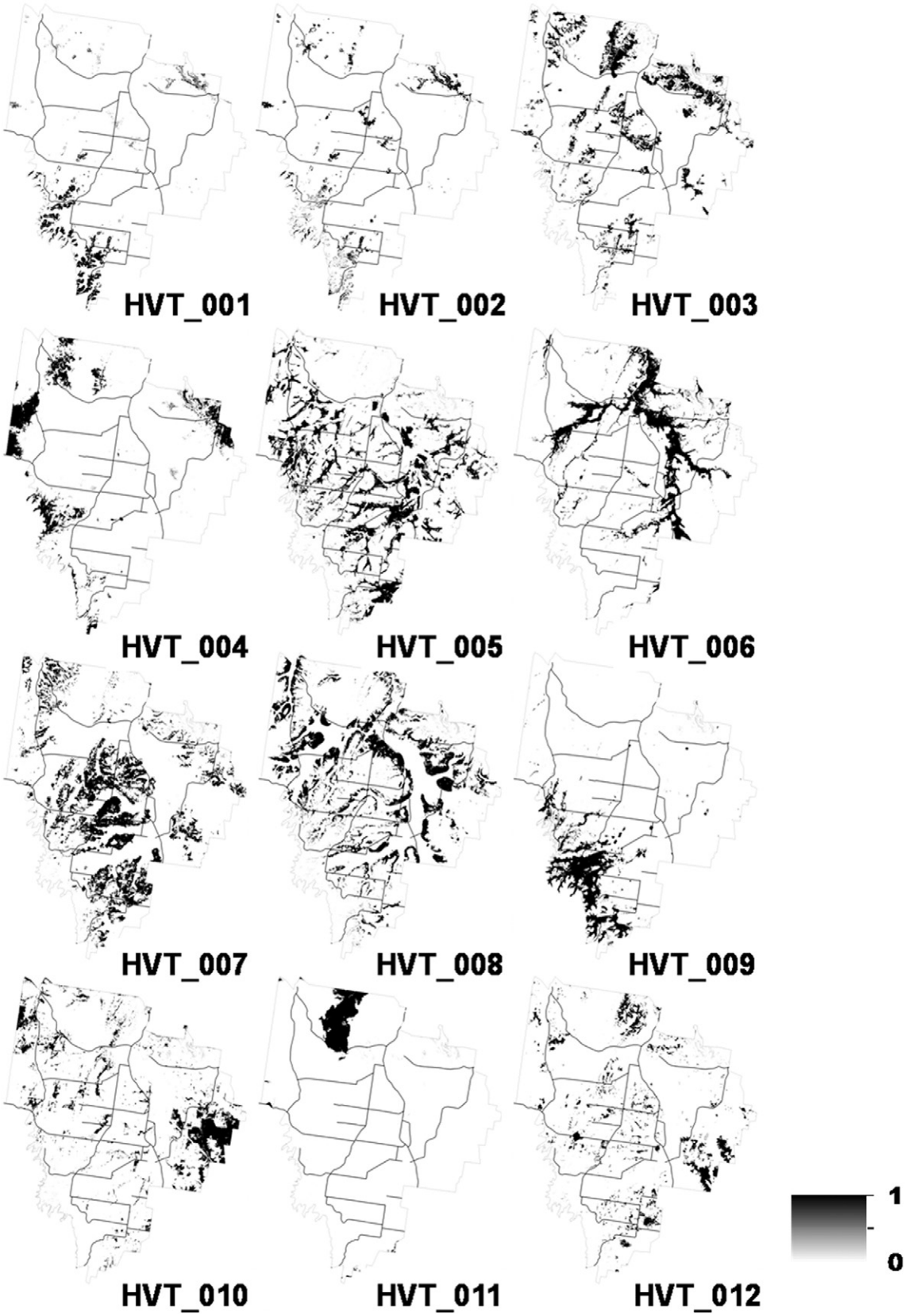

Fig. 6. Membership grades for each terron. This is a continuous grading from one (complete membership to terron) to zero (no membership to terron).

index distribution for this terron. Soil factors such as the prevalence of iron oxides in this terron is also another indication of relatively well draining soils.

3.1.2. HVT_002

This terron is the second of the marl terrons. As can be seen in Fig. 8, the two marl terrons are quite similar in terms of their separation or

Table 3
Soil and landform characteristics of each terron. Ranges are 2.5% and 97.5% quantiles of the distributions. Reported soil classes are those that cumulatively make up over 80% is the predicted soils types in the terron. The Australian Soil classes are: DEAB (Brown Dermosol), DEAA (Red Dermosol), DEXX (other Dermosols), CAXX (Calcarosols), CHAA (Red Chromosols), TEXX (Tenosols), and KUAA (Red Kurosols).

		Soil variables						Landform variables							
Terron	Area	pH	Clay %	Drainage index	Kaolin	Kaolin-smectite	Iron oxides	Australian soil classes	ELV	SLP	TWI	MVF	VDC	MSP	ISR	CTA
HVT_001	5 km^2 (3%)	5.2–8.0	14–50	46–77	0.00–0.20	0.01–0.55	0.03–0.83	DEAB, DEAA	139–224	4–24	8–11	0–0.5	21–90	0.11–0.84	751–1631	7–10
HVT_002	4 km^2 (2%)	5.4–8.1	1–50	54–80	0.00–0.35	0.00–0.70	0.22–0.97	DEAB, DEAA, CAXX	84–192	2–26	8–11	0–1	26–114	0.10–0.87	671–1601	7–11
HVT_003	18 km^2 (9%)	5.0–7.8	23–48	52–77	0.10–0.60	0.08–0.50	0.38–0.94	DEAB, DEAA	60–146	1–10	9–15	0–2	15–63	0.10–0.83	1025–1381	7–13
HVT_004	13 km^2 (6%)	4.8–7.4	25–48	43–67	0.01–0.44	0.04–0.44	0.03–0.81	DEAB	81–209	1–18	9–16	0–4	23–91	0.03–0.88	914–1413	7–13
HVT_005	32 km^2 (15%)	5.2–7.6	37–49	32–60	0.07–0.40	0.22–0.50	0.02–0.47	DEAB	49–121	0–4	14–21	2–6	0–25	0.25–0.92	1134–1270	11–18
HVT_006	22 km^2 (10%)	5.4–7.2	11–40	24–62	0.04–0.34	0.05–0.33	0.01–0.51	DEAB, TEXX	33–118	0–5	14–22	0–7	0–34	0.00–0.95	1117–1260	12–19
HVT_007	33 km^2 (16%)	4.8–7.0	42–52	59–78	0.14–0.54	0.22–0.60	0.12–0.85	DEAA	65–142	1–9	10–17	0–3	6–42	0.04–0.74	1034–1344	8–14
HVT_008	33 km^2 (16%)	5.0–7.4	33–51	37–64	0.10–0.43	0.17–0.54	0.05–0.61	DEAB	45–122	1–8	13–20	0–3	2–30	0.05–0.88	1038–1314	11–18
HVT_009	14 km^2 (7%)	4.9–7.7	33–51	30–50	0.02–0.31	0.08–0.47	0.00–0.56	DEAB, DEXX	72–203	1–17	9–19	0–4	0–45	0.15–0.94	897–1446	8–16
HVT_010	16 km^2 (8%)	4.7–6.6	36–50	31–64	0.26–0.71	0.16–0.45	0.05–0.77	DEAB	52–104	0–5	12–19	0–6	0–32	0.04–0.91	1103–1275	10–17
HVT_011	7 km^2 (3%)	5.0–7.0	33–48	31–59	0.12–0.49	0.12–0.44	0.05–0.73	DEAB	52–104	1–7	10–18	0–4	8–59	0.05–0.83	1077–1318	12–19
HVT_012	11 km^2 (5%)	4.6–6.7	36–51	59–77	0.25–0.72	0.19–0.56	0.21–0.91	KUAA	49–124	1–8	11–18	0–3	4–31	0.04–0.79	1038–1327	9–15

taxonomic distance, which together, are quite apart (dissimilar) from the other terrons. This terron does appear to be more concentrated to the north east of the study area, and shares more-or-less similar soil and landscape characteristics to HVT_001. A distinguishing characteristic between HVT_001 and HVT_002 is that currently HVT_002 supports less area of viticultural landuse. From the soil information, sub-soil clay textures appear to be more variable in this terron compared with HVT_001. Together HVT_001 and HVT_002 account for 5% of the land area in the HWCPID.

3.1.3. HVT_003

This terron (9% of land area) is most concentrated to the north of the study area and is composed dominantly of Red and Brown Dermosols. It is taxonomically most similar to HVT_004, HVT_005, and HVT_008 which share commonalities in relatively high clay subsoils. This terron occurs at moderate elevation on slightly to moderately sloping terrain. The soils appear moderate to well draining, and clearly indicate the presence of iron oxides in them. Kaolinite and kaolinite–smectite seem detectable in these soils in appreciable quantities as well.

3.1.4. HVT_004

This terron (6% of land area) is most concentrated to the peripheral zones of the study area that is best described as of higher elevation and more variably sloping than HVT_001. Dominant soils are Brown Dermosols. It is likely that soil pH in this terron can verge on being acidic. Variables such TWI and MVF indicate that there is quite some variation in the hydrology of this terron in terms of water movement and accumulation, which is reflected by a moderate soil drainage index.

3.1.5. HVT_005

Brown Dermosols are the most dominant soils of this terron. This terron is quite widespread spatially across the HWCPID where it occupies approximately 15% of the land area. It is situated in the lower parts of the area, on slightly to gently sloping land. Soils are generally not acidic, and while the general observation of higher sub-soil clay is apparent, the detection of kaolinite or kaolinite–smectite was relative to other terrons, moderate. These soils also appear not to have appreciable iron oxides, and given the estimated soil drainage for this terron and associated landform parameters, waterlogging of these soils occurs here on occasions.

3.1.6. HVT_006

This terron occupies one of the lowest parts of the HWCPID which covers the area of old and existing channel course lines (10% of land area). Brown Dermosols are dominant, as are Tenosols. The soils are best described as being non-acidic, can be susceptible to drainage problems, and do not, relative to other terrons, have significant quantities of the examined soil clay minerals or iron oxides. Lower sub-soil clay contents and position relative to water channels indicate that this terron is more alluvially dominated than the other terrons. Taxonomically, this terron is most similar to HVT_005, but distinguished in terms of differences in sub-soil clay content.

3.1.7. HVT_007

This terron is the only one where Red Dermosols are singly the most dominant soil type. Soils here are particularly clayey, but are generally moderate to well draining. Kaolinite, kaolinite–smectite and iron oxides all appear to be in appreciable quantities in these soils. Soil pH ranges from acidic to neutrality. Occupying 16% of the land area, this terron is mainly concentrated to the middle of the HWCPID. Topographically, it is moderately elevated and sloping. Hydrological indices indicate that water passes through this terron efficiently, rather than accumulates, which would cause problems for soil drainage.

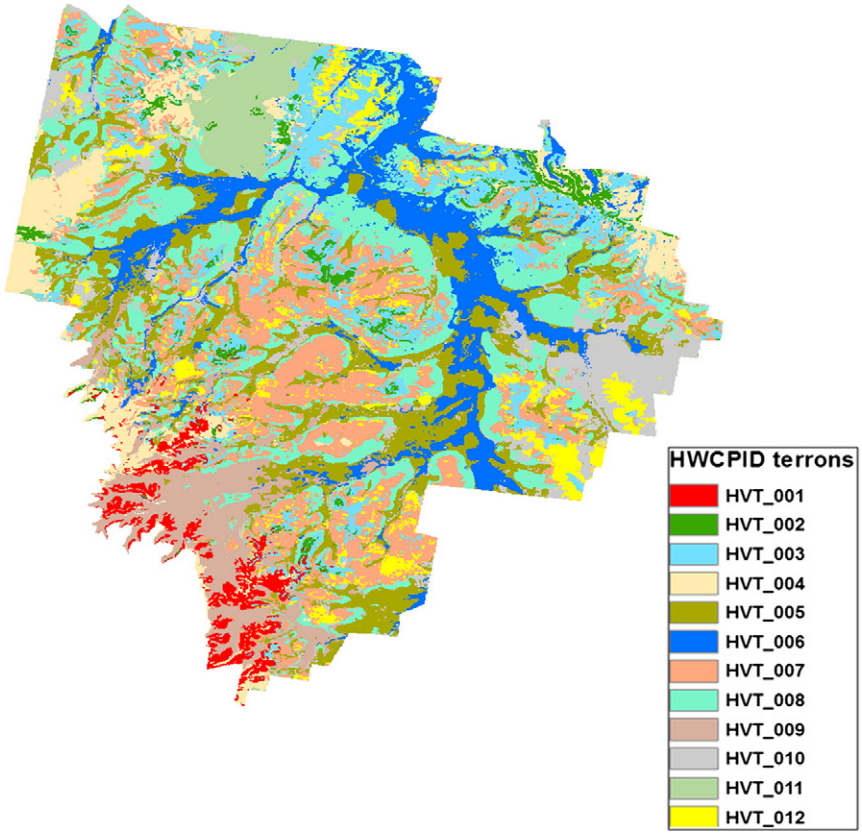

Fig. 7. Map of terrons (highest fuzzy membership) for the HWCPID.

3.1.8. HVT_008

Occupying 16% of the land area, Brown Dermosols are the most dominant of all soils in this terron. It is taxonomically similar to HVT_005, and occurs more-or-less geographically adjacent to HVT_006. Such that HVT_006 occurs predominantly around the creek lines, while this terron occupies a position slightly more elevated. As a result, the soils are heavily textured, and they contain appreciably more clay minerals and iron oxides. Soil pH does not appear to be that much of a limiting factor in this terron.

3.1.9. HVT_009

This terron is taxonomically associated with HVT_005 and HVT_008. It is mainly concentrated to the south-west peripheral of the HWCPID, and is quite variable in terms of topography. Elevation ranges considerably, as do other terrain parameters such as SLP, MVF, and VDC. Generally soils have a moderate to high clay percentage. The most dominant soils here are Brown Dermosols. Other Dermosols (not red or brown) are also common in this terron. These included Yellow and Black Dermosols which indicate that internal soil drainage is an issue in the terron. This is corroborated by the relatively low drainage index values for this terron compared to others.

3.1.10. HVT_010

Soils here are best described as dominantly Brown Dermosols, being moderate to heavy textured, and have a tendency to be acidic; distinguishing it from HVT_008. Kaolinite is the most dominant clay mineral and iron oxides also appear to be present also in these soils. In terms of landform, this terron is slightly sloping and can be found to the lower landscape positions of the HWCPID.

3.1.11. HVT_011

This small terron (3% of land area) is almost exclusively situated to the northern locality of the HWCPID, where Brown Dermosols are the most dominant soils. This terron is taxonomically quite distinct, relative to other terrons, but is closest to HVT_008. Soils range from slightly acidic to neutral, and lightly to moderately textured.

3.1.12. HVT_012

This terron (5% of land area) is the only one where Kurosols are the most dominant soil class. Naturally soils will verge on being acidic. Kaolinite, kaolinite–smectite, and iron oxides are all present in appreciable quantities, in soils of moderate to heavy clay content. Generally soil

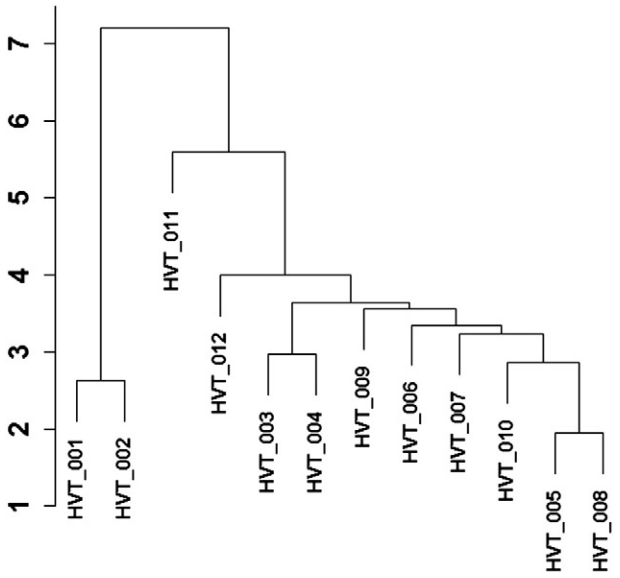

Fig. 8. Dendrogram resulting from Ward's hierarchical clustering, indicating taxonomic similarities and separations between terrons.

Table 4
Mahalanobis distances between each terron.

	HVT_001	HVT_002	HVT_003	HVT_004	HVT_005	HVT_006	HVT_007	HVT_008	HVT_009	HVT_010	HVT_011
HVT_002	2.65										
HVT_003	4.16	3.15									
HVT_004	3.77	3.41	2.97								
HVT_005	4.22	4.07	2.68	2.94							
HVT_006	4.83	4.41	3.31	3.60	2.57						
HVT_007	4.28	4.16	3.05	3.40	2.81	3.36					
HVT_008	4.48	4.19	2.73	2.98	1.95	3.04	2.85				
HVT_009	3.70	4.47	3.52	3.49	2.94	3.40	3.51	3.02			
HVT_010	4.79	4.73	3.22	3.51	2.50	3.43	3.22	2.77	3.50		
HVT_011	5.60	5.27	4.45	4.39	4.31	4.49	4.58	4.05	4.94	4.45	
HVT_012	4.96	4.84	3.42	3.84	3.67	3.44	3.70	3.39	3.91	3.40	5.03

drainage does not appear to be an issue here in this moderately sloping and elevated terron.

4. Discussion and further perspectives

The terrons that have been described embody the most complete knowledge of soil and landscape characteristics for this region, to this date. It would therefore be expected, that in consideration of the soil information that was used to populate the model, in addition to knowledge about the landforms, the terrons will be a good starting point to further study on realising terroir in the HWCPID.

A characteristic feature of this study was the novelty of enhancing the value of field collected data through the use of soil and spectral inference functions. The motive for this was essentially to estimate and map difficult and expensive to measure soil attributes, namely soil clay minerals, iron oxides and soil drainage. Or in other words, soil variables are often scarce in a legacy soil information database. Furthermore the treatment of soil classes as continuous variables by using the confidence estimates rather than using nominal class estimates provided a means to capture additional soil information for inclusion into the terron model, namely information about pedogenesis and soil processes. These processes are implicit to the soil classification but not necessarily available in terms of measured data.

The approach for the creation of terrons used in this study relies much on both pedometric and digital soil mapping techniques. The advantage (sometimes disadvantage) of this is that the modelling is objective, quantitative, and importantly validated. From the current study; clearly the results indicate that there's a need to have a systematic improvement in the accuracy of the soil spatial prediction models. There also exist a number of caveats and technical difficulties in making inference of the soil properties that were investigated in this study. For example the use of vis-NIR for soil clay mineral detection is complicated by the co-occurrence of absorption feature at similar wavelengths. This is notable for kaolinite and smectite which is illustrated in Fig. 3, and similarly for the iron oxides hematite and goethite as another example. In the case of soil drainage estimation, the implicit assumption that is made is that soil drainage and the corresponding soil colour are related. This phenomenon is a function of landscape position and weathering intensity, and there is scientific evidence to support this e.g. Bouma (1983).

Despite the issues with uncertainty, and implied assumptions it is difficult to ascertain whether statistically more accurate digital soil maps will drastically change the configuration and character of the terron units. However, some refinement should be expected naturally. Efforts to address the accuracy issues of the soil mapping include continued soil survey in the HWCPID (which occurs on an annual basis currently). Secondly, it is expected that over time, available predictive soil covariate information will improve and be rich enough to capture the complexity of soil cover across this region. The refinement of the terrons in the HWCPID also needs to be put into context of longer term research objectives in this region, which is the realisation of terroir. This incorporates not only refining the soil and landscape characteristics (the terrons), but also an assessment of management practices, detailed climate information, and wine grape varietal knowledge — a more-or-less holistic approach to define terroir.

One of these ongoing efforts for refining the terron includes preliminary investigations into the design and setting up of a weather monitoring programme to detect finer scale variations in climate conditions such as temperature, incidences of frost risk, and growing degree days as a few exemplar climatic variables. This monitoring network will require the installation of strategically placed weather recording systems from which data can be regularly retrieved from.

Another innovation is the detailed measurement of gamma-ray sources from the soil surface (radiometric data). Important information about the constituent properties of soil can be inferred with this technology (Taylor et al., 2002). Currently however, the radiometric information that is available is too coarse to use. Detailed proximal radiometric survey is currently in progress or in active planning (Stockmann et al., in review).

Another vital piece of data for realising terroir is the linking of grape varieties with soil and landscape information. Currently resources are being put in place to not only map varieties, but also populate a spatial database with other relevant factors such as year of planting and root stock type (pers comm. L. Riley).

On the point of the number of terrons, and what is the optimal number; one could surmise that the more terrons identified, the more complex the map will be, together with an increased difficulty in differentiating between them. Currently with the available soil and landscape information, 12 terrons may be optimal. Already with this number of terrons, taxonomically, a few are quite similar as indicated in Fig. 8 and Table 4. For example, both marl terrons (HVT_001 and HVT_002) are quite similar, as are HVT_005 and HVT_008. As additional landscape and/or soil information come to hand, and improve in accuracy, separation between the terrons may increase (or may even coalesce). Reducing the number of terrons could also be investigated in further work for the reason that greater generalisation may need to be made across the region. In this case, agglomerating classes based on taxonomical distance either through hierarchical clustering or some other method will provide the objectivity to this type of procedure.

5. Conclusions

Overall some important outcomes were achieved in this study:

- It identified, mapped and described 12 continuous soil and landscape entities or terrons for the HWCPID, which will be a preliminary outcome for the realisation of terroir in this region.
- Each terron is characterised by landscape variables (derived from a digital elevation model) and soil variables which include soil pH, clay percentage, soil mineralogy (clay types and presence of iron oxides), continuous soil classes, and presence or absence of marl.
- This study has been an example of harnessing a soil information database, and using innovative inference and predictive methods to derive useful soil information to be fed into a terron model.

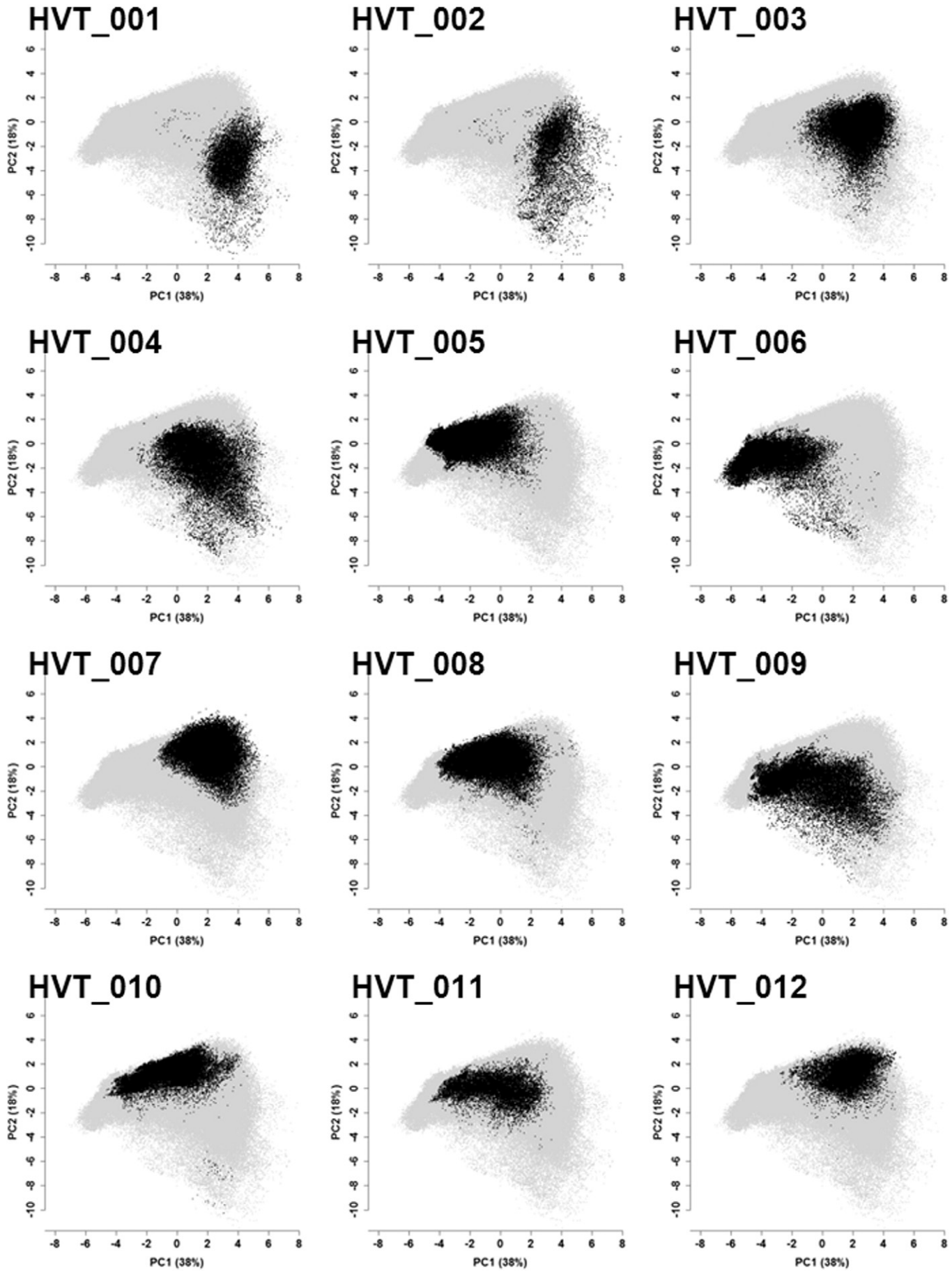

Fig. 9. Plots of the first two principle components of the soil and landscape terron variables with the scores highlighted for each hardened terron class.

- The map of terrons is more useful than a soil class map alone (it actually includes it), or a map of a given soil attribute. The embodiment of soil and landscape will provide a more complete picture of the natural environment. From a landowner's or manager's perspective, this comprehensive data will ideally lead to more informed decisions about how land management plans are implemented.
- The approach detailed in this paper is the first step in describing terroir for the HWCPID. Over time this work will be bolstered with the acquisition of new and additional environmental variables including gamma radiometrics, and temperature and climate related indices. Furthermore, the matching of terrons with different grape varietals will be a significant step to the full realisation of terroir in this prominent Australian wine region.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.geodrs.2014.08.001. These data include the Google map of the most important areas described in this article.

References

Agresti, A., 1984. Analysis of Ordinal Categorical Data. Wiley, New York.

Barham, E., 2003. Translating terroir: the global challenge of French AOC labelling. J. Rural. Stud. 19 (1), 127–138.

Bezdek, J.C., Ehrlich, R., Full, W., 1984. The fuzzy c-means clustering algorithm. Comput. Geosci. 10, 191–203.

Bishop, T.F.A., McBratney, A.B., Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma 91 (1–2), 27–45.

Boettinger, J.L., Ramsey, R.D., Bodily, J.M., Cole, N.J., Kienast-Brown, S., Nield, S.J., Saunders, A.M., Stum, A.K., 2008. Landsat spectral data for digital soil mapping. In: Hartemink, A.E., McBratney, A.B., Mendonca-Santos, M.L. (Eds.), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 193–203.

Bonfante, A., Basile, A., Langella, G., Manna, P., Terribile, F., 2011. A physically oriented approach to analysis and mapping of terroirs. Geoderma 167–68, 103–117.

Bouma, J., 1983. Hydrology and soil genesis of soils with aquic moisture regimes. In: Wilding, L.P., Smeck, N.E., Hall, G.F. (Eds.), Pedogenesis and Soil Taxonomy. 1. Concepts and Interactions. Elsevier Science Publishers, The Netherlands, pp. 253–281.

Bramley, R.G.V., Hamilton, R.P., 2007. Terroir and precision viticulture: are they compatible? Int. J. Vine Wine Sci. 41 (1), 1–8.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132 (3–4), 273–290.

Bureau of Meterology, 2014. Climate data online: monthly rainfall. Retrieved from: http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=061329.

Burns, S., 2012. The importance of soil and geology in tasting terroir with a case history from the Willamette Valley, Oregon. In: Dougherty, P.H. (Ed.), The Geography of Wine. Springer, Netherlands, pp. 95–108.

Carey, V.A., Archer, E., Saayman, D., 2002. Natural Terroir Units: What Are They? How Can They Help the Wine Farmer? WineLand.

Carey, V.A., Saayman, D., Archer, E., Barbeau, G., Wallace, M., 2008. Viticultural terroirs in Stellenbosch, South Africa. I. The identification of natural terroir units. Int. J. Vine Wine Sci. 42 (4), 169–183.

Carey, V.A., Archer, E., Barbeau, G., Saayman, D., 2009. Viticultural terroirs in Stellenbosch, South Africa. III. Spatialisation of viticultural and oenological potential for cabernet-sauvignon and sauvignon blanc by means of a preliminary model. Int. J. Vine Wine Sci. 43 (1), 1–12.

Carre, F., McBratney, A.B., 2005. Digital terron mapping. Geoderma 128 (3–4), 340–353.

Chrisitan, C.S., Stewart, G.A., 1953. General report on survey of Katherine–Darwin region, 1946. Land Research Series No. 1Commonwealth Scientific and Industrial Research Organisation, Melbourne.

Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. J. Geophys. Res. 89 (NB7), 6329–6340.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. J. Geophys. Res. Solid Earth Planets 95 (B8), 12653–12680.

Clark, R.N., Swayze, G.A., Livo, K.E., Kokaly, R.F., Sutley, S.J., Dalton, J.B., McDougal, R.R., Gent, C.A., 2003. Imaging spectroscopy: earth and planetary remote sensing with the USGS Tetracorder and expert systems. J. Geophys. Res. Planets 108 (E12), 1–44.

Clark, R.N., Swayze, G.A., Wise, R., Livo, K.E., Hoefen, T.M., Kokaly, R.F., Sutley, S.J., 2007. USGS Digital Spectral Library splib06a. U.S. Geological Survey, Data Series. 231.

Commission Interntionale de l'Eclairage, 1978. Recommendations on Uniform Color Spaces, Color-Difference Equations, and Psychometric Color Terms, Supplement No. 2 to Publication CIE No. 15(E-1.3.1)1971/(TC-1.3)1978. Bureau Central de la CIE, Paris.

Dougherty, P.H. (Ed.), 2012. The Geography of Wine: Regions, Terroir and Techniques. Springer, Netherlands (256 pp.).

Evans, C.V., Franzmeier, D.P., 1988. Color index values to represent wetness and aeration in some Indiana soils. Geoderma 41, 353–368.

Feagan, R., 2007. The place of food: mapping out the 'local' in local food systems. Prog. Hum. Geogr. 31 (1), 23–42.

Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39 (12).

Green, D.R., 2012. Geospatial tools and techniques for vineyard management in the twenty-first century. In: Dougherty, P.H. (Ed.), The Geography of Wine. Springer, Netherlands, pp. 227–245.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer Science, New York.

Hawley, S.P., Glen, R.A., Baker, C.J., 1995. Newcastle Coalfield Regional Geology 1:100 000, 1st Edition. Geological Survey of New South Wales, Sydney, Australia.

Isbell, R., 2002. The Australian Soil Classification. CSIRO Publishing, Melbourne.

Isbell, R., McDonald, W.S., Ashton, L.J., 1997. Concepts and Rationale of the Australian Soil Classification. ACLEP, CSIRO Land and Water, Canberra.

IUSS Working Group, W.R.B., 2007. World reference base for soil resources 2006. First update 2007. World Soil Resources Report No. 103. FAO, Rome.

Jones, G.V., Snead, N., Nelson, P., 2004. Geology and wine. 8. Modelling viticultural landscapes: a GIS analysis of the terroir potential in the Umpqua Valley of Oregon. Geosci. Can. 31 (4).

Kidd, D.B., Webb, M.A., Grose, C.J., Moreton, R.M., Malone, B.P., McBratney, A.B., Minasny, B., Viscarra Rossel, R.A., Cotching, L.A., Sparrow, L.A., Smith, R., 2012. Digital soil assessment: guiding irrigation expansion in Tasmania, Australia. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and Beyond. CRC Press, The Netherlands, pp. 3–8.

Kovac, M., Lawrie, J.M., 1990. Soil Landscapes of the Singleton 1:250 000 Sheet. Soil Conservation Service of NSW, Sydney, Australia.

Lanyon, D.M., Cass, A., Hansen, D., 2004. The effect of soil properties on vine performance. CSIRO Land and Water Technical Report No. 34/04. CSIRO Land and Water.

Liu, M.Q., Samal, A., 2002. A fuzzy clustering approach to delineate agroecozones. Ecol. Model. 149 (3), 215–228.

Macqueen, R.W., Meinert, L.D. (Eds.), 2006. Fine wine and terroir. The Geoscience perspective. Geological Association of Canada, Geoscience Canada Reprint Series. 9 (246 pp.).

Malone, B.P., de Gruijter, J.J., McBratney, A.B., Minasny, B., Brus, D.J., 2011. Using additional criteria for measuring the quality of predictions and their uncertainties in a digital soil mapping framework. Soil Sci. Soc. Am. J. 75, 1032–1043.

Malone, B.P., McBratney, A.B., Minasny, B., 2012. Digital mapping of a soil drainage index in the Lower Hunter Valley, NSW. Proceedings of the Joint SSA and NZSSS Soil Science Conference: Soil Solutions for Diverse Landscapes, Hobart, Tasmania.

McBratney, A.B., de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. J. Soil Sci. 43, 159–175.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer functions to soil inference systems. Geoderma 109 (1–2), 41–73.

McBratney, A.B., Mendonca-Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.

Minasny, B., McBratney, A.B., 2002. FuzME version 3.0. Australian Centre for Precision Agriculture, The University of Sydney, Australia.

Moran, C.J., Bui, E.N., 2002. Spatial data mining for enhanced soil map modelling. Int. J. Geogr. Inf. Sci. 16 (6), 533–549.

Morand, D.T., 2013. The World Reference Base for Soils (WRB) and Soil Taxonomy: an appraisal of their application to the soils of the Northern Rivers of New South Wales. Soil Res. 51 (3), 167–181.

Mulder, V.L., Plotze, M., de Bruin, S., Schaepman, M.E., Mavris, C., Kokaly, R.F., Egli, M., 2013. Quantifying mineral abundances of complex mixtures by coupling spectral deconvolution of SWIR spectra (2.1–2.4um) and regression tree analysis. Geoderma 207–208, 279–290.

Northcote, K.H., 1979. A Factual Key for the Recognition of Australian Soils, 4th edition. Rellim Technical Publications, Glenside, South Australia.

Odgers, N.P., McBratney, A.B., Minasny, B., 2011. Bottom-up digital soil mapping. I. Soil layer classes. Geoderma 163, 38–44.

Quinlan, R., 1992. Learning with continuous classes. Proceedings of the 5th Australian Joint Conference On Artificial Intelligence, Hobart, Tasmania, pp. 343–348.

Quinlan, R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Francisco.

Quinn, P.F., Beven, K.J., Lamb, R., 1995. The ln(a/tan b) index: how to calculate it and how to use it within the TOPMODEL framework. Hydrol. Process. 9 (2), 161–182.

Rayment, G.E., Higginson, F.R., 1992. Australian Laboratory Handbook of Soil and Water Chemical Methods. Inkarta Press, Melbourne.

Robertson, A.R., 1976. The CIE 1976 colour-difference formulae. Color. Res. Appl. 2, 7–11.

Soil Survey Staff, 2010. Keys to Soil Taxonomy, 11th edition. United States Department of Agriculture, Natural Resources Conservation Service, Washington, DC.

Stewart, G.A., 1968. Land evaluation. In: Stewart, G.A. (Ed.), Land Evaluation. Papers of a CSIRO Symposium (CSIRO/UNESCO, Canberra). Macmillan, Melbourne.

Stockmann, U., Malone, B.P., McBratney, A.B., 2014. Landscape-scale exploratory radiometric mapping using proximal soil sensing. Geoderma (in review).

Taylor, J.A., 2004. Digital Terroirs and Precision Agriculture: Investigations into the Application of Information Technology in Australian Vineyards. (PhD thesis) The University of Sydney.

Taylor, J.A., Minasny, B., 2006. A protocol for converting qualitative point soil pit survey data into continuous soil property maps. Aust. J. Soil Res. 44 (5), 543–550.

Taylor, M.J., Smettem, K., Pracilio, G., Verboom, W., 2002. Relationships between soil properties and high-resolution radiometrics, central eastern Wheatbelt, Western Australia. Explor. Geophys. 33 (2), 95–102.

Unwin, T., 2012. Terroir: at the heart of geography. In: Dougherty, P.H. (Ed.), The Geography of Wine. Springer, Netherlands, pp. 37–48.
Vaudour, E., 2002. The quality of grapes and wine in relation to geography: notions of Terroir at various scales. J. Wine Res. 13 (2), 117–141.
Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. Geoderma 150 (3–4), 253–266.
Ward, J.H., 1963. Hierarchical grouping to optimise an objective function. J. Am. Stat. Assoc. 58, 236–244.
White, R., Balachandra, L., Edis, R., Chen, D., 2007. The soil component of terroir. Int. J. Vine Wine Sci. 41, 9–18.
Wilson, J.E., 1998. Terroir: The Role of Geology, Climate and Culture in the Making of French Wines. Mitchell Beazley, London.