Generalisable Methods for Vegetation Classification Using Computer-Generated Pseudoquadrats

Liam Butler
Newcastle University, l.butler3@newcastle.ac.uk

Roy A. Sanderson Dr
Newcastle University, roy.sanderson@newcastle.ac.uk

iEMSs International Congress on Environmental Modelling and Software

9th International Congress on Environmental Modelling and Software
Fort Collins, Colorado, USA, Mazdak Arabi, Olaf David, Jack Carlson, Daniel P. Ames (Eds.)
https://scholarsarchive.byu.edu/iemssconference/2018/

Generalisable Methods for Vegetation Classification Using Computer-Generated Pseudoquadrats

Liam Butler [a], **Roy A. Sanderson** [a]
[a] *School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE17RU* (l.butler3@newcastle.ac.uk, roy.sanderson@newcastle.ac.uk)

Abstract: Many countries have developed phytosociological classifications of their vegetation to describe semi-natural and natural vegetation communities. Can methods be developed for any classification system to allocate newly surveyed quadrats, i.e. data gathered from new field surveys, into the most likely vegetation community? Algorithms or software already exist to allocate quadrats for some classifications, but these are not generalisable to any system. We test the robustness of several generalisable approaches to allocate quadrats to an existing phytosociological classification, using the British National Vegetation Classification (NVC) as a case study. Vegetation from 167 quadrats was classified using two-way indicator species analysis (TWINSPAN) and the resultant groups allocated to communities within the NVC using the NVC-specific 'MAVIS' software. Sets of artificial 'pseudoquadrats' for potential communities were computer-generated based on either the published NVC community descriptions or from the subset of species surveyed. Distance in ordination space of observed quadrats from pseudoquadrats was used to predict community type. The conventional NVC-specific MAVIS classification produced 11 sub-communities at the site, and this was assumed to be the most reliable descriptor of the vegetation communities. There was a close match between the pseudoquadrat-based community predictions and the MAVIS predictions, although pseudoquadrats based on the subset of species observed at the site appeared to be slightly more reliable. Our results demonstrate that the use of pseudoquadrats provides a flexible, generalisable means to allocate objectively vegetation quadrats into any extant classification system.

***Keywords**:* NVC; MAVIS; ordination; pseudoquadrats; sub-community.

1 INTRODUCTION

Phytosociological vegetation classification systems have been developed in numerous countries to describe semi-natural and natural vegetation communities or relevés. These include the regional European Vegetation Archive (EVA) currently being developed for Europe and neighbouring countries (EVS; European Vegetation Survey, 2018), the Irish Vegetation Classification (IVC: Biodiversity Ireland, 2018), the United States National Vegetation Classification (USNVC, 2016), the Canadian National Vegetation Classification (CNVC; Canadian National Vegetation Classification, 2013), the New Zealand National Vegetation Survey Databank (NVS; Landcare Research, 2016), and the Great Britain National Vegetation Classification (NVC; Rodwell, 1998a, b, 2006). There are broad similarities in the structure of some of these systems, for example hierarchical classes of different vegetation relevés into 'communities' and 'sub-communities' (Great Britain NVC) or broader 8-level 'formations' through to 'associations' (USNVC; CNVC). Some countries have collaborated to produce standard methods for vegetation database management, e.g. TurboVeg recommended by the EVS and International Association for Vegetation Science.

Due to the different range of habitats and ecosystems encompassed by these vegetation classifications, they differ in both their recommended field survey methods, and in the techniques used to allocate quadrats to vegetation classes. For example the British NVC recommends a minimum of five quadrats per relevé, precluding the use of many historical datasets where vegetation was surveyed on a per-quadrat level, whereas the Irish IVC can use single quadrats. Some classifications do not provide formal methods to allocate new quadrats to classes, whilst software has been developed for some national systems, e.g. ERICA for the IVC. The British NVC is derived from a two-way indicator species analysis (TWINSPAN, Hill 1979), and the resultant published NVC handbooks provided paper-based keys, analogous to binomial taxonomic keys, to allocate quadrats to communities. Given the practical difficulties of using such manual keys, computer-based methods to allocate field quadrats have been developed including MATCH (Malloch, 1998; Curreli et al. 2013), TABLEFIT (Hill, 1989; Dodd et al., 1994; CEH, 2014) and most recently the Modular Analysis of Vegetation Information System (MAVIS; MAVIS Plot Analyser, Version 1.04, UK; Morecroft et al., 2009; Smart et al., 2016). The latter also links outputs to plant traits under Grime’s et al. (1988) Competitor, Stress-tolerator and Ruderal (CSR) and Ellenberg scores (Morecroft et al., 2009, Smart et al., 2016).

Whilst such computer-based systems have been shown to be invaluable aids for vegetation scientists, several problems remain. First, their design may preclude the allocation of some historical vegetation records within a national classification due to differences in survey methods, for example the need for a recommended minimum of five random quadrats per relevé for the British NVC. Quadrats may have been placed at regular intervals along transects, across survey grids, or at random, none of which accords with standard NVC survey methods. Subsequently ecologists, especially when utilising published data from historical surveys, may wish to incorporate such quadrats into a national system and therefore need to be able to resolve differences in survey methods. Second, such computer-based systems are not available for many national systems, and even where software is available, outputs are usually restricted to tables that indicate the highest predicted vegetation classes. In the real-world, semi-natural and natural vegetation generally exists in a continuum, therefore it would be useful to be able visualise this continuum in an ordination diagram that can be readily interpreted by practicing vegetation scientists, whilst still obtaining the advantages of a national classification system. Such a system has to be relatively simple, so that it could be generalised to any country's vegetation classification.

Our primary aims were therefore to 1) develop and test methods to create vegetation 'pseudoquadrats' that were representative of vegetation quadrats for known vegetation communities in an established vegetation classification; 2) place these pseudoquadrats within an ordination framework so that the relationships between the published vegetation communities could be observed on a continuum; 3) allocate field quadrats into this ordination space to predict community membership. Two different methods of generating vegetation pseudoquadrats were tested, and we used the British NVC as our classifier, but the method could be applied to any system.

The British National Vegetation Classification (NVC – Rodwell, 1998a, b) is a phytosociological classification system that is hierarchically structured. It defines broad vegetation groups (e.g. ‘U’ – upland vegetation/acid grasslands, ‘MG’ – mesotrophic grasslands, ‘H’ – heath) within which are defined communities (e.g. ‘U2’ *Deschapsia flexuosa* grassland, ‘MG6’ *Lolium perenne-Cynosurus cristatus* grassland) which are in turn subdivided into sub-communities (e.g. ‘U2b’ *Vaccinium myrtillus* sub-community and ‘MG6b *Anthoxanthum odoratum* sub-community; Rodwell, 1998b). Each community is associated with a characteristic set of climatic, physical and biological factors, and the NVC provides coarse-scale 10km dot-distribution maps of where they are recorded (Rodwell, 1998a, b; Dodd et al., 1994; Smart et al., 2016).

2 METHODOLOGY

2.1 Study site

Vegetation was collected from the ‘Ashtrees Dipper’ heft within the Rede catchment at Redesdale, Northumberland National Park, UK (Figure 1), a 96 ha area of sheep-grazed semi-natural vegetation, ranging in altitude from about 250 to 350m (Sanderson et al., 1995b). Vegetation was surveyed from 167 1-metre quadrats positioned across a grid 75m apart (Figure 1) in 2001, recording the percentage cover of all species in each quadrat. This heft is dominated by grassland and moorland communities (Rushton et al., 1992). This area is listed as a priority habitat in the Primary Habitats Inventory (PHI) land cover as part of the Biodiversity Action Plan (BAP UK), later succeeded by the UK post-2010 Biodiversity Framework as part of the ‘Strategic Plan for Biodiversity 2011-2020’ addressing targets set in the EU Biodiversity Strategy (EUBS; JNCC, 2016). The main soil types are a mixture of surface-water gleys, podzols and raw peats (SSEW, 1983).

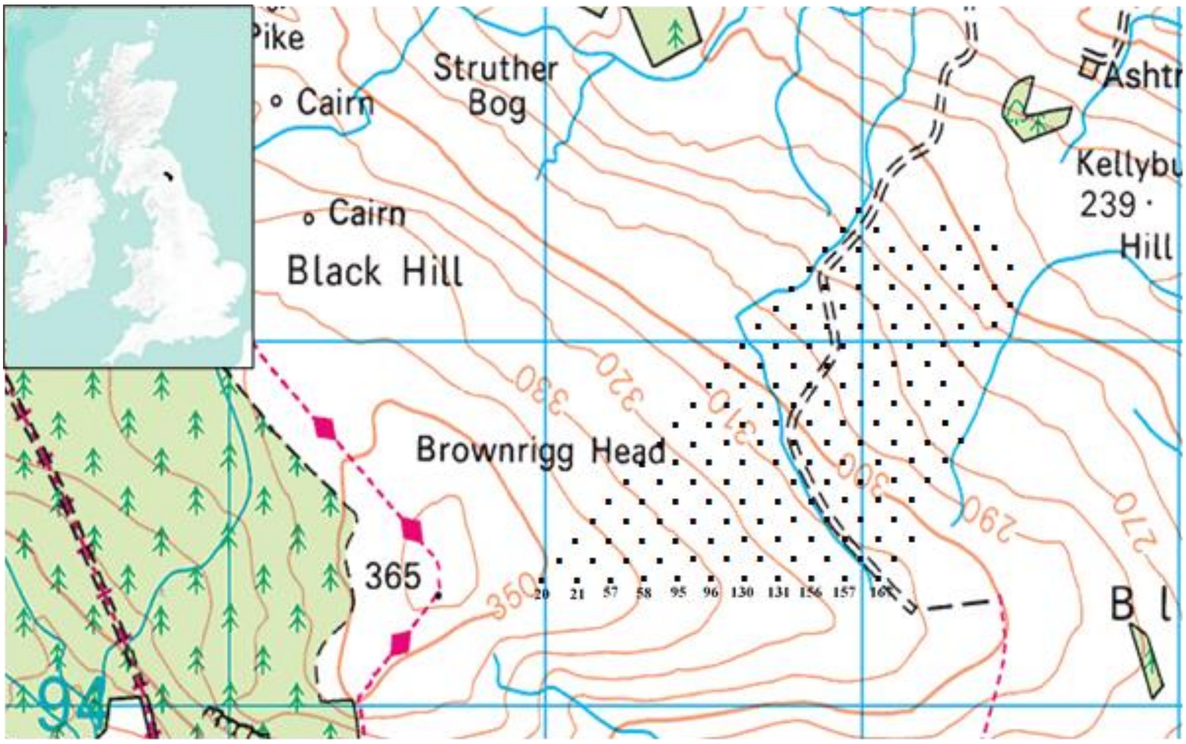

Figure 1. Location of 167 surveyed quadrats (points) at Ashtrees Dipper, Redesdale, Northumberland National Park, Northumberland, UK. 1km grid squares. Ordnance Survey backdrop provided through EDINA Digimap.

2.2 Initial classification of surveyed vegetation data using standard methods

An initial classification of the vegetation was undertaken via MAVIS, to try and produce the most accurate classification that might be considered an observed 'baseline' against which to compare the new pseudoquadrat techniques. Outputs from MAVIS are most accurate when inputs consist of sets of at least five quadrats of similar vegetation composition, rather than on individual quadrats (CEH, 2014; Smart et al., 2016). Therefore, prior to using MAVIS the vegetation (percentage cover abundance) was classified via two-way indicator species analysis (TWINPAN – Hill, 1979) to create sets of quadrats. TWINSPAN was chosen as this is the classification method used during the original development of the British NVC. The NVC community\sub-community classifications derived from MAVIS, supplemented by other field data from 2017, meant that each quadrat could be reliably allocated to an NVC sub-community. Eleven sub-communities were identified at the Ashtrees Dipper using this approach (Table 1) and these eleven sub-communities were used as the subset from which to create and test the generalisable pseudoquadrat method.

2.3 Generation of pseudoquadrats from (i) literature community data and (ii) surveyed species

Two complementary approaches were used to generate the pseudoquadrats for a sub-community, the first based on the entire set of species published in the NVC handbooks for a sub-community (henceforth denoted LIT), whilst the second was restricted to utilise only the set of species recorded in the Ashtrees field survey (ASH). The latter was used to analyse a subset of species that are representative of true field surveys for better comparison with community data found in the NVC handbooks. The method described here is a refinement of that originally described in Sanderson et al. (1995a), such that it can be more readily generalised to any country's vegetation classification system.

Twenty-five pseudoquadrats were generated per sub-community, with the number of species within a pseudoquadrat being determined by two different randomisation methods. In the LIT approach the number of species per pseudoquadrat was derived to select a number between the published minimum to maximum expected numbers of species per quadrat in a sub-community, whilst for the ASH approach this was based on the minimum and maximum number of species per quadrat recorded across all the field quadrats at Ashtrees Dipper. This thus means that the minimum and maximum number of species a quadrat can contain is not less or more than the number of species recorded in any given quadrat at Ashtrees Dipper. Species were selected at random by drawing uniform distribution deviate (a uniform probability between 0 and 1) which, in simple terms, means that the number of species per quadrat was likely to reflect that in field quadrats. A named species was then drawn at random: in the LIT method the random species was selected from any of the possible species published for that sub-community, whilst for the ASH approach the published list of species was restricted to the subset of species actually recorded at the field. The percentage cover abundance for that species in that sub-community reported in the NVC handbooks was then allocated to the species. This procedure was repeated, drawing species at random (without repetition) until the total number of species in a pseudoquadrat reached that initially drawn at random. Finally, the 275 pseudoquadrats (11 sub-communities, 25 pseudoquadrats each) were ordinated by detrended correspondence analysis, a type of multivariate ordination analysis, (DCA – Hill, 1979) and the mean centroids and standard errors of each sub-community visualised in ordination space (Figures 2 and 3), for both the LIT and ASH sets of 167 pseudoquadrats. Ordination analysis, in its simplest form and in the context of this research, produces graphs where the axes summarise community variation, and in which quadrats of similar species composition will be plotted as points close to each other.

Table 1. A summary of the eleven sub-communities derived from TWINSPAN plus MAVIS together with the number of quadrats in each sub-community as well as the relevant vegetation community and sub-community codes, name of each community and sub-community and the number of quadrats within each sub-community found at Ashtrees

Community code	Community type	Sub-community code	Sub-community type	Number of quadrats in sub-community
H12	*Calluna vulgaris-Vaccinium myrtillus* heath	H12a	*Calluna vulgaris* sub-community	10
H9	*Calluna vulgaris-Deschampsia flexuosa* heath	H9e	*Molinia caerulea* sub-community	30
M15	*Scripus cespitosus-Erica tetralix* wet heath	M15d	*Vaccinium myrtillus* sub-community	12
M19	*Calluna vulgaris-Eriophorum vaginatum* blanket mire	M19a	*Erica tetralix* sub-community	13
MG10	*Holcus lanatus-Juncus effusus* rush pasture	MG10a	Typical sub-community	6
MG6	*Lolium perenne-Cynosurus cristatus* grassland	MG6b	*Anthoxanthum odoratum* sub-community	5
U2	*Deschampsia flexuosa* grassland	U2b	*Vaccinium myrtillus* sub-community	26
U4	*Festuca ovina-Agrostis capillaris-Galium saxatile* grassland	U4b	*Holcus lanatus-Trifolium repens* sub-community	3
		U4d	*Luzula multiflora-Rhytidiadelphus loreus* sub-community	4
U5	*Nardus stricta-Galium saxatile* grassland	U5a	Species-poor sub-community	40
U6	*Juncus squarrosus-Festuca ovina* grassland	U6b	*Carex nigra-Calypogeia trichomanis* sub-community	18

H: Heaths; **M**: Mires; **MG**: Mesotrophic grasslands; **U**: Calcifugous grasslands/montane communities

Rodwell, J. S. (1998a). British Plant Communities Volume 2 Mires and heaths. Cambridge, UK, Cambridge University Press.
Rodwell, J. S. (1998b). British Plant Communities Volume 3 Grassland and montane communities. Cambridge, UK, Cambridge University Press.

2.4 Prediction of field surveyed data in pseudoquadrat ordinations

Field quadrats were placed within the pseudoquadrat DCA ordination space as 'passive samples', i.e. the data from the field quadrats were placed into the resulting ASH and LIT detrended correspondence analysis ordination plot, so that their positions were based on the pseudoquadrat ordination scores, whilst ensuring that the field quadrats had no effect on the original ordination (Hill, 1979; Sanderson et al., 1995b). It was assumed that the shorter the Euclidean distance between a field quadrat and a sub-community mean centroid the higher the probability that the field quadrat belonged to that sub-community. The following equation, adapted equation from Sanderson et al., (1995a), was used to determine the probability that a field quadrat belonged to a sub-community:

$$P_i = 1 - \frac{d_i}{\sum_{t=1}^{k} d_t} 100$$

where,
P_i = probability that field quadrat belongs to community i
d = distance of field quadrat to mean centroid of community
k = total number of vegetation communities

All analyses were undertaken in R (Version 3.4.2, Vienna, Austria) plus the vegan (Dixon, 2003) and TwinspanR packages (Roleček et al., 2009).

3 RESULTS

3.1 Initial classification of surveyed vegetation data

The eleven sub-communities derived from TWINSPAN plus MAVIS are summarised in Table 1, together with the number of quadrats in each community. The most well-defined communities within MAVIS were U6b, U5a and M19a with matching scores of 60.03%, 57.85% and 54.87% respectively. In contrast, the U4b sub-community was less well-defined by the TWINSPAN classification, with a MAVIS score of 47.96%. Overall accuracy for MAVIS when used without prior clustering of quadrats into sets via TWINSPAN on individual quadrats for comparison was 15.87% and a Kappa value of 0.004.

3.2 Generation of pseudoquadrats from (i) literature community data and (ii) surveyed species

Centroids of sub-communities for pseudoquadrats derived from the LIT method are summarised in Figure 2. DCA axis 1 represents a trend from the higher altitude, acid sub-communities (H and U) to mesotrophic grasslands (MG6b and MG10a). DCA axis 2 separated H12a *Calluna vulgaris*-*Vaccinium myrtilis* heath from the remaining acid grassland and heaths; H12a is totally dominated by *Calluna vulgaris* which can represent over 90% of the vegetation cover (Rodwell, 1998a). H12a and M15d *Scirpus cespitosus*-*Erica tetralix* wet heath produced the most variable pseudoquadrats, as represented by their error bars especially on axis 2 in Figure 2, whereas there was little variability amongst the pseudoquadrats generated by this method for U6b *Juncus squarrosus*-*Festuca ovina* grassland. The equivalent plot (Figure 3) of NVC sub-communities based on pseudoquadrats generated from the subset of species at the study-site (ASH) again shows a trend along DCA axis 1 from upland acid grasslands, heaths and mire sub-communities, through to mesotrophic grasslands. However, here axis 2 separates H9e from the remainder. This is the *Molinia caerulea* sub-community within H9 *Calluna-vulgaris*-*Deschampsia flexuosa* heath, and is relatively species-poor.

3.3 Prediction of field surveyed data in pseudoquadrat ordinations

It is clear from both the LIT and ASH methods (Figures 2 and 3 respectively) that many field quadrats were intermediate in species composition, being positioned between typical NVC sub-community centroids as defined by the pseudoquadrats. The highest probabilities for each field quadrat were compared to the 'observed' classification (initial TWINSPAN+MAVIS classification) and confusion matrices created to show the numbers of correctly predicted quadrats (Tables 2 and 3). These produced an overall accuracy of 20.36% for the LIT method, and 29.34% for the ASH method, and Kappa values of 0.109 and 0.178 respectively.

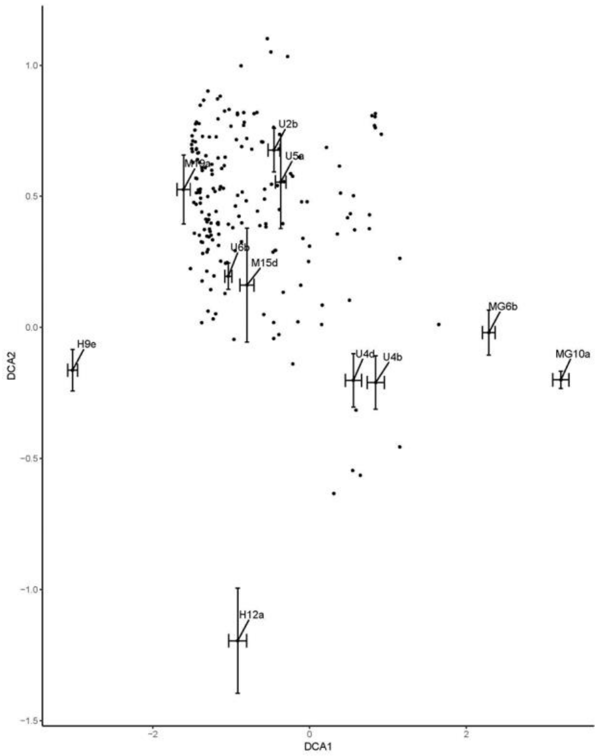

Figure 2. Sub-community centroids and standard errors of pseudoquadrats derived from all potential species for a sub-community reported in NVC handbook (LIT) in a detrended correspondence analysis. Field quadrats from Ashtrees Dipper (points) displayed as passive samples.

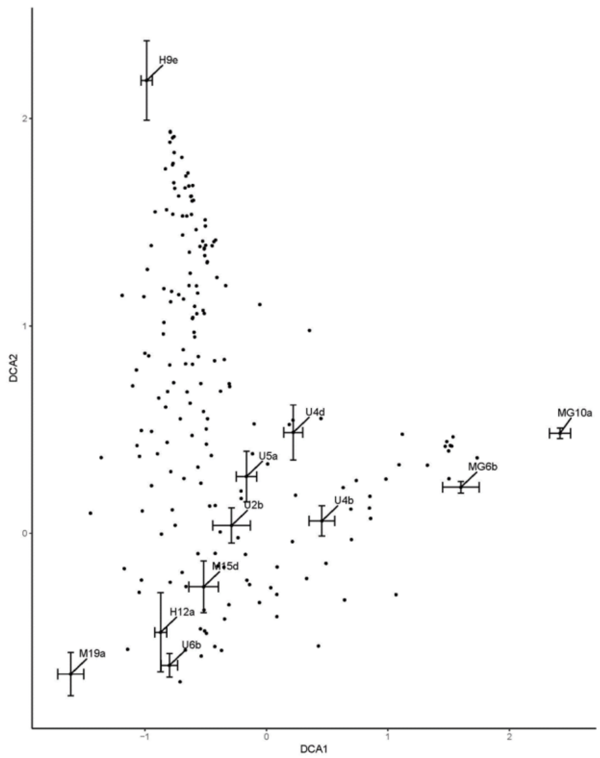

Figure 3. Sub-community centroids and standard errors of pseudoquadrats derived from subset of species recorded at the site (ASH) in a detrended correspondence analysis. Field quadrats from Ashtrees Dipper (points) displayed as passive samples.

Table 2. Confusion matrix of number of quadrats of highest probability corresponding to same vegetation sub-community classification. Columns: observed communities; rows: predicted via LIT pseudoquadrats.

	H12a	H9e	M15d	M19a	MG10a	MG6b	U2b	U4b	U4d	U5a	U6b	% Total
H12a	0	0	0	0	0	0	0	0	0	0	0	0.00
H9e	0	0	0	0	0	0	0	0	0	0	0	0.00
M15d	0	0	2	0	0	0	1	0	0	5	2	20.00
M19a	8	18	3	11	2	0	4	0	1	5	1	20.75
MG10a	0	0	0	0	0	0	0	0	0	0	0	0.00
MG6b	0	0	0	0	0	0	0	0	0	0	0	0.00
U2b	0	5	1	1	0	0	3	0	0	7	6	13.04
U4b	0	0	0	0	1	3	4	0	0	2	0	0.00
U4d	0	0	1	0	1	2	0	2	3	3	1	23.08
U5a	0	1	1	0	2	0	5	1	0	8	1	42.11
U6b	2	6	4	1	0	0	9	0	0	10	7	17.95
% Total	0.00	0.00	16.67	84.62	0.00	0.00	11.54	0.00	75.00	20.00	38.89	

Table 3. Confusion matrix of number of quadrats of highest probability corresponding to same vegetation sub-community classification. Columns: observed communities; rows: predicted via ASH pseudoquadrats.

	H12a	H9e	M15d	M19a	MG10a	MG6b	U2b	U4b	U4d	U5a	U6b	% Total
H12a	0	0	0	3	0	0	3	0	0	1	0	0.00
H9e	9	21	2	5	1	0	5	0	0	5	1	0.00
M15d	0	1	2	0	1	0	3	0	0	6	3	12.50
M19a	0	0	0	1	0	0	0	0	0	0	0	100.00
MG10a	0	0	0	0	0	0	0	0	0	0	0	0.00
MG6b	0	0	0	0	1	4	3	0	0	3	0	0.00
U2b	0	0	3	1	0	0	4	1	1	6	4	20.00
U4b	0	0	1	0	2	1	1	2	3	3	3	12.50
U4d	0	1	2	0	0	0	1	0	0	3	0	0.00
U5a	1	7	1	3	1	0	6	0	0	13	5	35.14
U6b	0	0	1	0	0	0	0	0	0	0	2	66.67
% Total	0.00	70.00	16.67	7.69	0.00	80.00	15.38	66.67	0.00	32.50	11.11	

4 DISCUSSION

This research tests the robustness of two generalisable approaches to allocate quadrats to an existing phytosociological classification. Both methods are simple to implement, and although their overall accuracy was relatively low (20% to 30%) it should be recognised that this was on a noisy and challenging set of field data. Furthermore, the accuracy for the ASH method (30%) was considerably better than when compared with equivalent predictions made on the level of individual quadrats by MAVIS, which achieved only 15.87% accuracy with the same data (recall that MAVIS is less reliable when used on single quadrats compared to sets of quadrats). The ability to visualise the pseudoquadrat communities in conventional ordination space is another advantage to the pseudoquadrat approach, as primary gradients affecting the subset of plant communities under investigation can be readily identified. Here, the main trends associated with elevation and soil acidity were the main drivers.

The greater accuracy of the ASH approach compared to the LIT method was unexpected, as the latter was based on the full set of species described in the standard NVC sub-community descriptions. The greater accuracy of the ASH method, restricting species to those observed at the study site, may simply have arisen because the resultant pseudoquadrats were more representative of species likely to be found in Northumberland, rather than the UK as a whole. This meant that the resultant sub-communities were also more representative local 'variants' of the sub-communities described in the NVC handbooks. In North East England it appeared that the field communities were more species poor than expected, when compared to the species listed for these communities in the British NVC handbooks. Whilst this might partly have been attributable to the use of smaller quadrats at Ashtrees (1-metre rather than the NVC standard 2-metre for such habitats), it is unlikely to be a major factor, especially given the large number of quadrats used in our survey. The 10km dot-distribution maps of several sub-communities in the NVC indicate that parts of North East England were not covered for those sub-communities in the original survey when the NVC was constructed. As such, the published sub-community descriptions may not necessarily describe accurately the species likely to be encountered in those same sub-communities in certain parts of the UK, since the data in the NVC handbooks was collected from different field sites not identical to Ashtrees. It is thus possible that some of the more species-poor community data we collected arose from this.

An additional factor that may account for the poorer accuracy of LIT pseudoquadrats is that since they use all the species potentially present within each sub-community, the random combination of species within each pseudoquadrats could have resulted in higher similarity between different sub-communities. This can be observed in some ‘M’ and ‘U’ communities positioned close to each other in ordination (Figures 2 and 3). Furthermore, the ASH method, by restricting species to those observed on-site, provided certainty that no ‘new’ species would occur in each ASH pseudoquadrat and are thus better comparable to observed data. This may explain the lower standard error of the sub-community centroids in Figure 3.

5 CONCLUSIONS

Pseudoquadrats can provide a reliable complimentary method to existing techniques to allocate quadrats into extant vegetation classification. The technique we describe is relatively simple, and could be easily modified to account for the characteristics of any national or regional classification. For example, if the classification provides data on the frequency of species, as well as abundance, this can easily be incorporated into the pseudoquadrat algorithm at the first step of selecting a species. Where high-quality vegetation cover data are available, the most abundant species for a pseudoquadrat could be selected first, and selection of additional species cease once total percentage cover has exceeded a threshold of e.g. 100%. Some national classification systems provide mean, minimum, maximum covers for species, and again these can be readily incorporated. Prediction of field quadrats within a pseudoquadrat community ordination-space could be weighted depending on the variability of pseudoquadrats within a community. For example, in our study H12a could be up-weighted, given that the standard error bars for the sub-community centroid derived from the pseudoquadrats are relatively large, suggesting considerable variability in the sub-community's underlying species composition. Conversely U6b, with small s.e. bars, could be down-weighted. Some national vegetation classifications provide information on abiotic data associated with particular vegetation classifications. This could be used in combination with, for example GIS maps etc., or published maps of species-distributions, to fine-tune the construction of pseudoquadrats depending on the end-user's requirements. Pseudoquadrats can also be considered as complimentary to traditional surveying methods especially with the high availability of small- and large-scale remotely-sensed data (abiotic and biotic) of different environments (Critchley et al., 2002). Again, it is this flexibility in customising the construction of pseudoquadrats, visualising them in ordination space, and using them to make predictions for new field quadrats that provides its attraction.

ACKNOWLEDGEMENTS

Sincere thanks go to Prof. S. Rushton at Newcastle University, UK for supplying additional surveyed data and his co-supervision in this research. The research work disclosed in this publication is fully funded by the ENDEAVOUR Scholarship Scheme – Group B – National Funds.

REFERENCES

Biodiversity Ireland (2018) Irish Vegetation Classification. [Online]. Available at: http://www.biodiversityireland.ie/projects/national-vegetation-database/irish-vegetation-classification/ (Accessed 2 February 2018).

Canadian National Vegetation Classification (2013.) *CNVC Overview*. [Online]. Available at: http://cnvc-cnvc.ca/page.cfm?page=2228 (Accessed 2 February 2018).

CEH (2014) *TABLEFIT and TABLCORN*. [Online]. Available at: https://www.ceh.ac.uk/services/tablefit-and-tablcorn (Accessed 2 February 2018).

Critchley, C.N.R., Chambers, B.J., Fowbert, J.A., Sanderson, R.A., Bhogal, A. and Rose, S.C. 2002 Association between lowland grassland plant communities and soil properties. *Biological Conservation*, 105(2), pp. 199-215.

Curreli, A., Wallace, H., Freeman, C., Hollingham, M., Stratford, C., Johnson, H. & Jones, L. (2013) Eco-hydrological requirements of dune slack vegetation and the implications of climate change. *Science of the total environment*. 443910–919.

Dixon, P. (2003) VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. 14 (6), pp. 927–930.

Dodd, M.E., Silvertown, J., McConway, K., Potts, J. and Crawley, M. 1994 Application of the British National Vegetation Classification to the communities of the Park Grass Experiment through time. *Folia Geobotanica,* 29(3), pp. 321-334.

European Vegetation Survey (2018) *European Vegetation Archive (EVA) – European Vegetation Survey*. [Online]. Available at: http://euroveg.org/eva-database (Accessed 2 February 2018).

Grime, J.P., Hodgson, J.G. & Hunt, R. (1988) Comparative Plant Ecology: A Functional Approach to Common British Species, Unwin Hyman. *London, England, UK*.

Hill, M.O. (1979) A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. *TWINSPAN*.

Hill, M.O. (1989) Computerized Matching of Relevés and Association Tables, with an Application to the British National Vegetation Classification. *Vegetatio*. 83 (1/2), 187–194.

JNCC 2016. *UK Upland Habitats*. Available at: http://jncc.defra.gov.uk/page-1436.

Landcare Research (2016) - *National Vegetation Survey*. [Online]. Available at: https://nvs.landcareresearch.co.nz/ (Accessed 2 February 2018).

Malloch, A.J.C. (1998) A computer program to aid the assignment of vegetation data to the communities and subcommunities of the National Vegetation Classification. *University of Lancaster, UK*.

Morecroft, M.D., Bealey, C.E., Beaumont, D.A., Benham, S., Brooks, D.R., Burt, T.P., Critchley, C.N.R., Dick, J., Littlewood, N.A., Monteith, D.T. and Scott, W.A., 2009. The UK Environmental Change Network: emerging trends in the composition of plant and animal communities and the physical environment. *Biological Conservation*, *142*(12), pp.2814-2832.

R Core Team (2016). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. Available at: https://www.R-project.org/

Rodwell, J. S. 1998a. British Plant Communities Volume 2 Mires and heaths, Cambridge, UK, Cambridge University Press.

Rodwell, J. S. 1998b. British Plant Communities Volume 3 Grassland and montane communities, Cambridge, UK, Cambridge University Press.

Rodwell, J.S. (2006) *National vegetation classification: users' handbook*. Joint nature conservation committee.

Roleček, J., Tichý, L., Zelený, D. & Chytrý, M. (2009) Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation Science*. 20 (4), 596–602.

Rushton, S.P., Byrne, J.P. and Young, A.G. 1992. Redesdale: a preliminary test catchment for the NERC/ESRC Land Use Programme. *ITE symposium*.

Sanderson, R.A., Rushton, S.P., Cherrill, A.J. and Byrne, J.P., 1995a. Soil, vegetation and space: an analysis of their effects on the invertebrate communities of a moorland in north-east England. *Journal of Applied Ecology*, pp.506-518.

Sanderson, R.A., Rushton, S.P., Pickering, A.T. and Byrne, J.P., 1995b. A preliminary method of predicting plant species distributions using the British National Vegetation Classification. *Journal of environmental management*, *43*(3), pp.265-288.

Smart, S., Goodwin, A., Wallace, H. and Jones, M. 2016. MAVIS (Ver 1.03) User Manual.

SSEW (1983) Soil Survey of England and Wales, Lawes Agricultural Trust, Rothamsted Experimental Station, UK.

USNVC [United States National Vegetation Classification]. 2016. United States National Vegetation Classification Database, V2.01. Federal Geographic Data Committee, Vegetation Subcommittee, Washington DC. [usnvc.org] (accessed 01 February 2018).