NEW ZEALAND JOURNAL OF ECOLOGY

RESEARCH

New Zealand Environmental Data Stack (NZEnvDS): A standardised collection of spatial layers for environmental modelling and site characterisation

James K. McCarthy[1]*, John R. Leathwick[2], Pierre Roudier[3,4], James R. F. Barringer[1], Thomas R. Etherington[1], Fraser J. Morgan[4,5], Nathan P. Odgers[1], Robbie H. Price[6], Susan K. Wiser[1] and Sarah J. Richardson[1]

[1]Manaaki Whenua – Landcare Research, Lincoln 7640, New Zealand
[2]john.leathwick@gmail.com
[3]Manaaki Whenua – Landcare Research, Palmerston North 4442, New Zealand
[4]Te Pūnaha Matatini, A New Zealand Centre of Research Excellence, Auckland 1142, New Zealand
[5]Manaaki Whenua – Landcare Research, Auckland 1072, New Zealand
[6]Manaaki Whenua – Landcare Research, Hamilton 3216, New Zealand
*Author for correspondence (Email: mccarthyj@landcareresearch.co.nz)

Published online: 18 June 2021

Abstract: Environmental variation is a crucial driver of ecological pattern, and spatial layers representing this variation are key to understanding and predicting important ecosystem distributions and processes. A national, standardised collection of different environmental gradients has the potential to support a variety of large-scale research questions, but to date these data sets have been limited and difficult to obtain. Here we describe the New Zealand Environmental Data Stack (NZEnvDS), a comprehensive set of 72 environmental layers quantifying spatial patterns of climate, soil, topography and terrain, as well as geographical distance at 100 m resolution, covering New Zealand's three main islands and surrounding inshore islands. NZEnvDS includes layers from the Land Environments of New Zealand (LENZ), additional layers generated for LENZ but never publicly released, and several additional layers generated more recently. We also include an analysis of correlation between variables. All final NZEnvDS layers, their original source layers, and the R-code used to generate them are available publicly for download at https://doi.org/10.7931/m6rm-vz40.

Keywords: climate surfaces, environmental variables, geographical distance, landform, macroecology, open access data, terrain, topography, soil

Introduction

The term macroecology was first used to describe the study of relationships between organisms and their environment at large spatial scales by Brown and Maurer (1989). Since then, the use of quantitative spatial data to explore and explain ecological patterns has proliferated. Concurrent advances in our ability to collect (e.g. iNaturalist; https://inaturalist.nz/), store (e.g. National Vegetation Survey Databank, https://nvs.landcareresearch.co.nz/; Wiser et al. 2001), and share (e.g. Global Biodiversity Information Facility; https://www.gbif.org/) biodiversity data; to describe the abiotic environment at fine scales (Viscarra Rossel & Bui 2016); and to interrogate data sets of increasing size and complexity (Weigand et al. 2020) have allowed ecologists to tackle analyses of ever-increasing complexity and geographical scope. While species distribution models (SDMs) are probably the most common application of spatial analysis and prediction in ecology, other spatial patterns are also analysed and explored along a similar theme. Examples from New Zealand include investigations of species richness patterns (Lehmann et al. 2002), forest loss (Ewers et al. 2006; Walker et al. 2006; Perry et al. 2012), potential forest distributions (Leathwick 2001), species refugia (Buckley et al. 2010, McCarthy et al. 2021), ecosystem services (Ausseil et al. 2013), production forest productivity (Palmer et al. 2009), and forest carbon uptake (Whitehead et al. 2001). One common feature across these studies, however, is their reliance on high-quality, collated, and curated spatial data characterising the abiotic environment.

Spatial layers used for environmental analyses typically include some combination of variables describing climate, soil, topography, and disturbance. These are commonly used for broad-scale predictions, or, in the absence of site-level measurements, they can be used to derive values describing the abiotic conditions. Climate layers estimated using spatial interpolation methods, weather station data, and elevation data (Hutchinson 1995; Xu & Hutchinson 2010) are most commonly used for this purpose (e.g. Richardson et al. 2004; Pawson et al. 2008) and can provide a more realistic value than one measured at the nearest weather station if this is located some distance away (Tait et al. 2012). Soil and disturbance variables can be more challenging to measure and project across large areas, but advances are being made, ranging from national contributions (McNeill et al. 2014; Viscarra

DOI: https://dx.doi.org/10.20417/nzjecol.45.31

Rossel et al. 2015) to global initiatives (the Global Soil Map project; Arrouays et al. 2017). In recent decades, climate and soil variables have been supplemented with remotely sensed variables from satellites, aeroplanes and drones, which capture spectral and structural attributes in ever-increasing resolution and extent (Verrelst et al. 2015). Variables can be continuous (e.g. rainfall, temperature) or discrete (e.g. soil type); cover current, future or past time periods; and capture means, extremes, seasonality, and variability at a range of resolutions and extents. Individual variables can be highly correlated, but a particular variable may be considered more proximally relevant to a species' ecology than a more commonly used, correlated alternative (Austin 2002). For example, when analysing the environmental variables driving the distribution of an alpine herb, mean minimum temperature of the coldest month may be more appropriate than mean annual temperature, even though these two predictors are highly correlated (Pearson's correlation = 0.89; Appendix S2 in Supplementary Materials).

Environmental spatial layers are most easily used when they employ a consistent grid (of given resolution and extent) and projection system, allowing for easy extraction of point data for model parameterisation and interpretation, and subsequent prediction. Although such layers are available at a global scale (e.g. WorldClim, CHELSA), these initiatives are, by necessity of their coverage, constrained to a 1 km resolution (Hijmans et al. 2005; Karger et al. 2017). Over large parts of New Zealand, complex topography and mountain ranges produce steep gradients of climate and soil fertility that are scarcely captured at a 1 km scale. A single 1 km grid square in the Southern Alps can span lowland rainforest to alpine grassland (500–1200 m elevation), underscoring the need for finer-scale environmental spatial layers for ecological modelling in New Zealand.

In 2002 the Land Environments of New Zealand (LENZ) was published, which included seven climate layers, seven soil layers (ordinal categories), and a measure of slope (Leathwick et al. 2002a; Leathwick et al. 2003). These 15 layers were generated at 100 m resolution, downscaled to 25 m resolution, and used to classify New Zealand's land environments. Several additional climate layers were also generated as part of the LENZ project, but did not inform the classification exercise. The LENZ classifications and 15 environmental variables are available online (at 25 m resolution; e.g. Land Information New Zealand 2020a), but the additional climate layers were never publicly released. While these layers have been used in spatial analyses (e.g. Watt et al. 2010; Perry et al. 2012), the number of climate variables available is much reduced compared to modern standards (e.g. WorldClim; Hijmans et al. 2005), and topographic variables (Amatulli et al. 2018) are lacking.

Here we describe Version 1.0 of the New Zealand Environmental Data Stack (NZEnvDS), a package of 72 environmental layers comprising 41 climate variables, eight soil variables, 18 topographic/terrain variables, and six geographical distance variables. Currently we do not include layers depicting land use and land cover, or remotely sensed variables depicting vegetation properties, but these could be included in future versions because they are known to be good predictors of biodiversity patterns (Müller et al. 2015; Dymond et al. 2019). Layers are provided at 100 m resolution and comprise all the existing layers from LENZ, and additional layers calculated from the source climate variables, including equivalents to all 19 WorldClim variables. Topographic variables have been derived from a 100 m version of the New Zealand National Digital Elevation Model (NZDEM; Barringer et al. 2002), which is also supplied (Appendix S1). Additional variables have been derived from a range of New Zealand spatial data layers. The layers are ready to use as input data in a range of environmental models and are available for download from the Manaaki Whenua – Landcare Research DataStore at https://doi.org/10.7931/m6rm-vz40.

Methods

Source layers

The environmental layers provided here were largely derived from previously unreleased layers generated as part of the LENZ project (Leathwick et al. 2002a; Leathwick et al. 2002b). Three of the geographical distance layers (horizontal distance to nearest river, distance to nearest road, vertical distance to river) were derived using the LENZ spatial grids and shape files downloaded from the Land Information New Zealand (LINZ) Data Service (Land Information New Zealand 2020b, 2020c). The LENZ data included climate variables capturing humidity, water balance, precipitation, solar radiation, sun hour ratios, temperature, vapour pressure deficit, and wind speed; soil variables capturing phosphorous levels, calcium levels, age, drainage, induration (hardness), and particle size; and slope. Many of the climate variables were originally presented as annual values, but were, in fact, also generated monthly (Leathwick et al. 2002b).

The LENZ climate layers were derived using ANUSPLIN software, which uses climate station and elevation data, and thin plate smoothing splines (Hutchinson & Gessler 1994) to predict across an entire surface, in this case, all of New Zealand's three main islands and surrounding inshore islands (Leathwick et al. 2002a). All climate data available from the New Zealand Meteorological Service at the time were included, primarily from the period 1950–1980. The LENZ soil layers were derived from the New Zealand Fundamental Soil Layers (FSL) (https://soils.landcareresearch.co.nz/soil-data/fundamental-soil-layers/; Newsome et al. 2008), which were themselves generated by expert interpretation of two major soil data sources: the New Zealand Land Resource Inventory (NZLRI), and the National Soils Database (NSD). While these data sources suffer from inconsistencies and bias towards lowland agricultural areas (see pp. 18–24 of Leathwick et al. 2002a for a discussion of the LENZ soil layer reliability), they remain the only spatially complete, quantitative layers describing a range of soil properties for New Zealand. A recently released, quantitative layer describing soil pH between 0 and 10 cm is also included (Roudier et al. 2020). The LENZ slope gradient layer was derived using 5 × 5 cell averaging filters from a 25 m digital elevation model (DEM), and resampled to 100 m (Leathwick et al. 2002a).

Post processing

All data management and analyses were completed using R 4.0.2 (R Core Team 2020) and the raster R package (Hijmans 2020), GRASS GIS version 7.8 (Neteler et al. 2012), and SAGA GIS version 7.3.0 (Conrad et al. 2015). Appendix S1 lists and describes the environmental layers, with details of their derivation, and includes unmodified versions of the LENZ data layers. The full suite of 19 BIOCLIM variables were calculated using monthly precipitation, minimum temperature, and maximum temperature layers from LENZ, using the biovars function in the dismo R package (Hijmans et al. 2017); however, two of the layers (mean annual temperature, bio1; total annual precipitation, bio12) were already included in LENZ, so we provide those original versions. Mean annual humidity, mean annual daily sunshine ratio, annual water deficit, mean annual vapour pressure deficit, and mean annual windspeed were calculated as the mean of their respective monthly LENZ layers. Normalised minimum winter temperature was calculated following Leathwick (2001) as the degree to which winter temperatures vary from that expected based on their mean annual temperature. This layer provides a measure of continentality, with high values in oceanic (coastal and mountainous) areas and low values in continental (lowland and interior) settings (Appendix S1). This layer has a very low correlation with many other commonly used temperature layers, such as mean annual temperature (Fig. 1; Appendix S2), allowing them to be used together in a model without confounding their relative effects (Leathwick 1995). Annual temperature amplitude, another measure of continentality with low values at the coast and high values in the interior (Appendix S1), was calculated as the maximum monthly mean temperature minus the minimum monthly mean temperature. Two growing degree days layers (5°C and 16°C base) were calculated using the monthly minimum and maximum temperature layers following the methods in Coops et al. (2001). We also include a range of layers describing water balance (Appendix S1).

A range of topographic/terrain indices and geographical distances were generated using the NZDEM resampled at 100 m resolution (Barringer et al. 2002; Land Information New Zealand 2020d, 2020e). Measures of slope, aspect, northness, and eastness (and also northness and eastness incorporating slope gradient) were generated using the terrain function in the raster R package (Hijmans 2020), and following advice from Amatulli et al. (2018). Indices of topographic position, roughness, and ruggedness, and the flow direction of water (see Appendix S1 for details) were also generated using the terrain function from dismo. A range of other terrain derivatives were generated in SAGA GIS from the NZDEM; namely, the topographic wetness index (SAGA wetness index), valley depth (a measure of vertical height below a summit), normalised height (a measure of position along a slope), and wind exposition. Other layers were derived from the NZDEM in GRASS GIS: annual and winter potential solar radiation using the NZDEM and the r.sun.daily add-on (Petras et al. 2015), geomorphons using the r.geomorphon module (Jasiewicz et al. 2013; Jasiewicz & Stepinski 2013), and latitude and longitude using the r.latlong module. Distance to coast, roads, and rivers was generated using the distance function from raster. Vertical distance to river was calculated using the NZDEM and the SAGA Channels module in SAGA GIS. Additional information is provided in Appendix S1.

All layers were processed at the original LENZ grid resolution (100 m), and projection (New Zealand Map Grid; NZMG). Layers were also reprojected to New Zealand Transverse Mercator (NZTM; see 'Data accessibility', below). The original layers, source layers, and all code (R and Bash scripts) used to generate the NZEnvDS layers are provided for reproducibility and to allow the community to generate additional layers from the source data. Since the NZEnvDS layers have been generated and supplied with consistent resolution and extent (Fig. 1), they are ready for use for these types of analyses without the need for further processing. Additional variables can, however, be created as required by ecologists using the source layers, which are also supplied.

Data accessibility

All source layers, the derived/final NZEnvDS layers, and the scripts used to generate them are available for download from the Manaaki Whenua – Landcare Research DataStore (as TIFF files in NZMG and NZTM projections), where they have been archived with a DOI: https://doi.org/10.7931/m6rm-vz40.

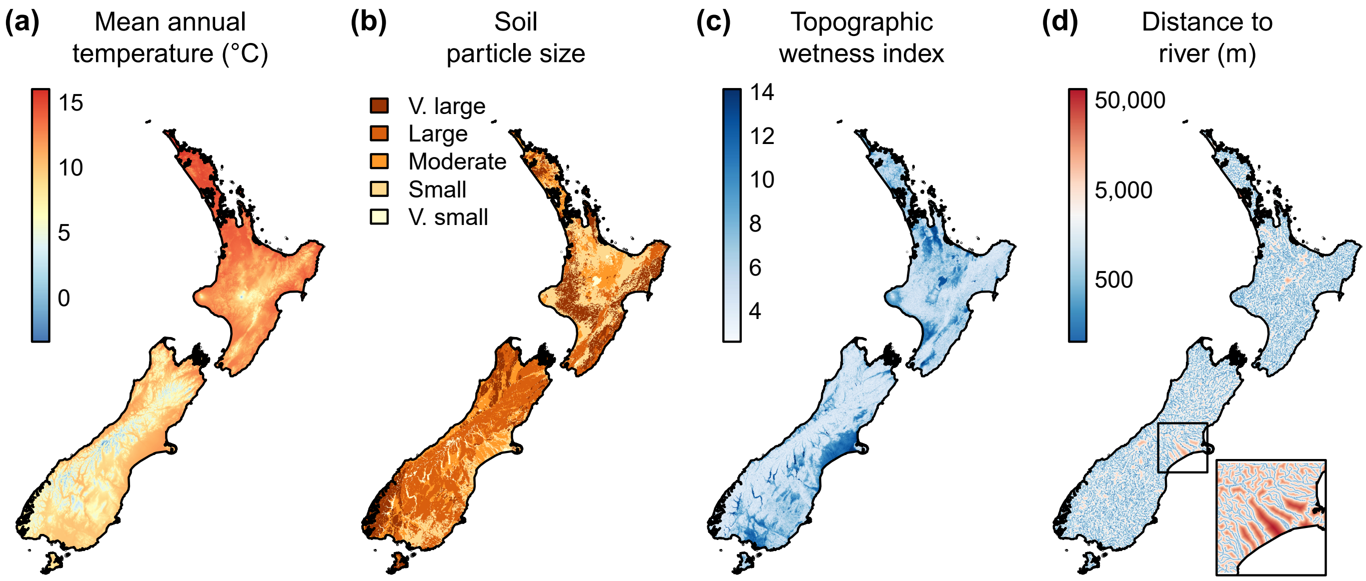

Figure 1. Example layers from the four broad environmental categories. (a) mean annual temperature (climate), (b) soil particle size (soil), (c) topographic wetness index (terrain), and (d) distance to river (geographical distance). Note that distance to river (d) is presented on a $\log_{10}$ scale to aid interpretation. Maps for all variables are included in Appendix S1.

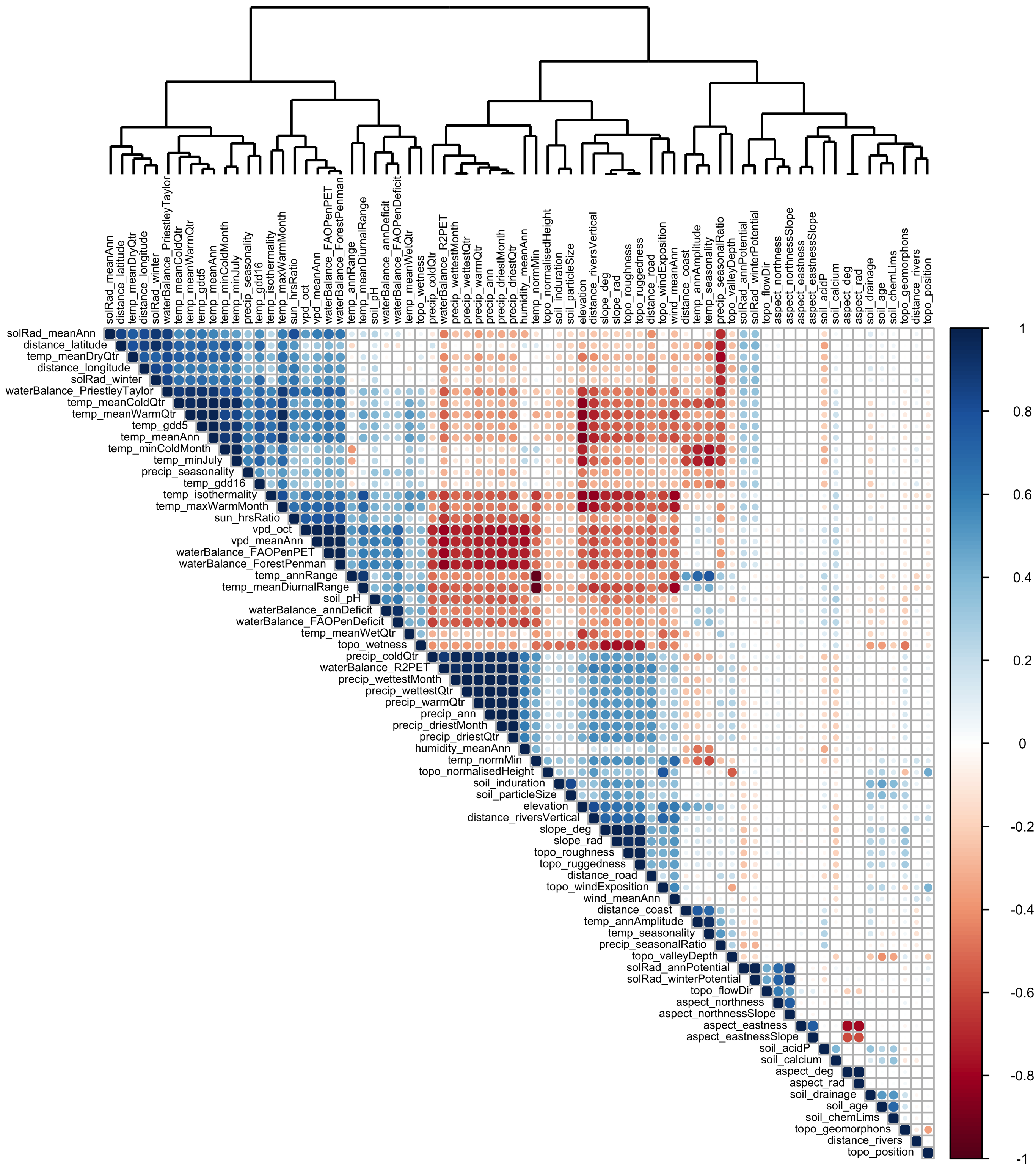

Figure 1. Pearson's correlation coefficients between all 72 environmental variables. Correlation coefficients were estimated using values extracted from 50,000 random locations, and variables are grouped into related variables using Euclidean hierarchical clustering, based on their correlations using the superheat function from the superheat R package (Barter & Yu 2020). Full descriptions of all variables are included as Appendix S1, and correlation coefficient values are presented in Appendix S2.

Attribution

Please cite the original LENZ dataset (Leathwick et al. 2002a) in addition to this article, to acknowledge the generation of many of the underlying data layers. When using the soil pH layer, please also cite Roudier et al. (2020).

Discussion

Collinearity and scale

In any modelling framework using environmental variables such as those included on NZEnvDS, their selection and use should be carefully considered (Williams et al. 2012). Many of the variables included in NZEnvDS exhibit various levels of Pearson's pairwise correlation, ranging from 0.01 (e.g. mean monthly temperature range and degrees of latitude) to 0.99 (e.g. precipitation of the wettest month and total annual precipitation) (Fig. 2; Appendix S2). The inclusion of two or more collinear variables in a statistical model can result in unstable models and poor estimates of regression parameters because it inflates their variance (Dormann et al. 2013). This can lead to the inclusion of irrelevant predictors in a model, which can be particularly problematic when extrapolating (Meloun et al. 2002). Colinear variables tend to be clustered into thematic groups such as precipitation, temperature, and soil (Fig. 2), but variables within the same broad category can also be very weakly correlated. For example, mean temperature of the coldest quarter and annual temperature variation share a Pearson's correlation of only –0.10 (Appendix S2). A common threshold is to select variables with a correlation <|0.7| (Dormann et al. 2013); however, other methods are often recommended to assess the influence of including colinear variables in a model (e.g. variance inflation factor; Burnham & Anderson 2002).

All terrain variables included in NZEnvDS were generated using the 100 m LENZ grid and, while a resolution of 100 m is usually considered fine-scale (Guisan et al. 2007), the resolution of a DEM used to generate terrain indices (such as topographic roughness index or topographic ruggedness index) will alter the representation of micro- versus macro-topographic variation. Certainly, for the assessment of physical properties such as landslide risk, finer resolutions than 100 m are essential (Tarolli 2014), but the scale appropriate for ecological analyses is less clear. Similar questions over scale are also commonplace for climate layers, with evidence suggesting spatial resolutions tend to be too coarse to accurately represent biological pattern and process (Potter et al. 2013; Bütikofer et al. 2020). This question deserves further research focused on specific ecological questions (e.g. how trees are distributed along toposequences; how tree regeneration patterns are influenced by landslides). The generation and use of terrain indices using code provided here, and either a downscaled version of the DEM or a finer-resolution DEM such as the original 25 m version of NZDEM (Land Information New Zealand 2020e, 2020d), and subsequent resampling back to 100 m (or interpolation of other NZEnvDS layers to the finer scale DEM), would advance our understanding of how best to incorporate terrain variables into ecological modelling.

Future directions

While NZEnvDS comprises a relatively extensive set of 72 spatial layers, spatially complete environmental layers that explain the complicated disturbance and geological history of the New Zealand landscape (glaciation, earthquake, volcanic history, etc.) are lacking. These factors are known to be important drivers of ecological processes and species distributions, from fine spatial scales and decadal time scales, through to nationwide patterns over geological time scales (Wyse et al. 2018). For example, many tree species regenerate on disturbed soils (e.g. mānuka *Leptospermum scoparium*; Stephens et al. 2005) or recent landslides (e.g. southern rātā *Metrosideros umbellata*; Stewart & Veblen 1982), yet spatial layers capturing these fine-scale disturbance events are currently lacking, which has hampered efforts to predict the distributions of affected species (McCarthy et al. 2021). At larger temporal and spatial scales, New Zealand's turbulent history of glaciation has left a strong imprint on the current distributions of many plant and animal species. One example is the discontinuous distribution of species present on the upper North and South Islands, but absent in between due to the ancient separation of the two by a Pliocene sea strait (e.g. *Quintinia serrata*; McGlone 1985). Similar patterns are observed in the South Island's central Westland "beech gap", a stretch of podocarp–broadleaf forest on the West Coast flanked by beech forest to the north and south, likely to have been driven by glaciation during the Last Glacial Maximum (and low levels of disturbance since; McGlone et al. 1996). The widely distributed rove beetle *Brachynopus scutellaris* is also largely absent from this gap (Leschen et al. 2008), and also see Leathwick (1998) for an exploration of the effect of climate (using the layers provided here) on New Zealand beech gaps. Sourcing or generating layers describing these processes is an area of priority.

In terms of soils, NZEnvDS is superficially well represented with the inclusion of seven soil layers originally developed as part of LENZ (Leathwick et al. 2002a) and a recently generated pH layer (Roudier et al. 2020). While the LENZ soil layers are useful, their relatively coarse scale (all with two to five ordinal categories; Appendix S1) and issues with reliability of the underlying data should be considered when using these layers (see pp. 18–24, Leathwick et al. 2002a for further discussion). Other New Zealand-wide soil layers are available, such as the Fundamental Soil Layers (Newsome et al. 2008), from which the LENZ soil layers were derived. These polygon-based shape files are largely considered out of date (Lilburne et al. 2012), but still represent the best available data for a large portion of New Zealand. They also include categorial information and, while categorial variables can be incorporated into many modelling frameworks, they increase model complexity because each class is usually treated as a separate variable in the model (Gelman & Hill 2007), and can also make models more difficult to interpret. Categorial variables can be simplified by reducing their classes to those relevant to a particular study, but the consideration of continuous variables allows more analyses of quantitative relationships, consideration of non-linear relationships and interactions, and facilitation of more quantitative and generalisable predictions. There are plans to produce quantitative spatial layers for other soil properties, but at present soil pH is the only layer available. Efforts are underway to generate updated data products describing New Zealand soil types through increased sampling and delineation of soil types across landscapes, but these may never be spatially complete, will probably prioritise production landscapes, and still largely capture categorial data (e.g. S-map; Lilburne et al. 2004; Lilburne et al. 2012).

All climate layers in NZEnvDS were generated using climate data collected from 1950 to 1980 (Leathwick et al. 2002a). Layers characterising more recent climate measurements (1981–2010) have been generated for New Zealand, but these must be purchased, and were generated at a coarser 500 m resolution (NIWA 2021). While the LENZ-sourced layers were generated from older climate data, the relative magnitude of climate trends across the country and between locations remains relevant. They may also more closely match existing data records/observations, which often represent historical collections of long-lived organisms (e.g. multiple decades at least for shrubs and many herbaceous plants, centuries for trees). For example, plant observational records from the New Zealand National Vegetation Survey Databank (https://nvs.landcareresearch.co.nz/; Level 1 data) and the New Zealand Virtual Herbarium Network have an average collection year of 1976 (SD = 14.7 years) and 1967 (SD = 33.7 years), respectively (analysis not shown), which more closely match the LENZ climate layers than the more recent layers. Updated layers covering time periods more recent than those captured by LENZ would, however, be useful for analyses of biotic invasions, shorter-lived species, and a mechanistic understanding of species responses to climate instability. As such, wider availability of more-recent climate surfaces should remain a priority.

Finally, technological advances mean that indices derived from remotely sensed data, primarily depicting vegetation cover and growth (e.g. normalised difference vegetation index; NZVI) but also forest structure and texture, are becoming increasingly available. Satellite-based systems such as Landsat, Sentinel-2, GEDI and MODIS are readily available across various temporal and spatial scales, and resolutions, and airborne LiDAR coverage is continually expanding. Furthermore, the development and investigation of relationships between indices and biological processes is rapidly advancing (Xue & Su 2017). As these variables become even more commonly used for environmental modelling, it will warrant further inclusion of such variables in future layer packages for New Zealand.

Acknowledgements

Recent efforts to compile layers and write this manuscript were supported by the Ministry of Business, Innovation and Employment (MBIE; contract C09X1709) and MBIE's Strategic Science Investment Fund. We thank Matt McGlone for suggesting inclusion of the annual temperature amplitude layer, Rachelle Binny for help generating the growing degree day layers, and Ray Prebble for technical editing. The manuscript was greatly improved following comments from two anonymous reviewers.

Author Contributions

JKM compiled the dataset, generated several layers, and wrote the first draft of the manuscript; JRL and FJM generated the original LENZ layers, and contributed to revisions of the manuscript; PR generated several layers and contributed to revisions of the manuscript; JRFB, TRE, NPO, RHP, SKW, and SJR all contributed to the conceptualisation of the data package, and revisions of the manuscript.

References

Amatulli G, Domisch S, Tuanmu M-N, Parmentier B, Ranipeta A, Malczyk J, Jetz W 2018. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. Scientific Data 5: 180040.

Arrouays D, Leenaars JGB, Richer-de-Forges AC, Adhikari K, Ballabio C, Greve M, Grundy M, Guerrero E, Hempel J, Hengl T, Heuvelink G, Batjes N, Carvalho E, Hartemink A, Hewitt A, Hong S-Y, Krasilnikov P, Lagacherie P, Lelyk G, Libohova Z, Lilly A, McBratney A, McKenzie N, Vasquez GM, Mulder VL, Minasny B, Montanarella L, Odeh I, Padarian J, Poggio L, Roudier P, Saby N, Savin I, Searle R, Solbovoy V, Thompson J, Smith S, Sulaeman Y, Vintila R, Rossel RV, Wilson P, Zhang G-L, Swerts M, Oorts K, Karklins A, Feng L, Ibelles Navarro AR, Levin A, Laktionova T, Dell'Acqua M, Suvannang N, Ruam W, Prasad J, Patil N, Husnjak S, Pásztor L, Okx J, Hallett S, Keay C, Farewell T, Lilja H, Juilleret J, Marx S, Takata Y, Kazuyuki Y, Mansuy N, Panagos P, Van Liedekerke M, Skalsky R, Sobocka J, Kobza J, Eftekhari K, Alavipanah SK, Moussadek R, Badraoui M, Da Silva M, Paterson G, Gonçalves MdC, Theocharopoulos S, Yemefack M, Tedou S, Vrscaj B, Grob U, Kozák J, Boruvka L, Dobos E, Taboada M, Moretti L, Rodriguez D 2017. Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. GeoResJ 14: 1–19.

Ausseil AGE, Dymond JR, Kirschbaum MUF, Andrew RM, Parfitt RL 2013. Assessment of multiple ecosystem services in New Zealand at the catchment scale. Environmental Modelling & Software 43: 37–48.

Austin MP 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157: 101–118.

Barringer JRF, Pairman D, McNeill SJ 2002. Development of a high-resolution digital elevation model for New Zealand. Landcare Research Contract Report: LC0102/170. Lincoln, New Zealand, Manaaki Whenua – Landcare Research. 33 p.

Barter R, Yu B 2020. superheat: A graphical tool for exploring complex datasets using heatmaps. R package version 1.0.0.

Brown JH, Maurer BA 1989. Macroecology: The division of food and space among species on continents. Science 243: 1145–1150.

Buckley TR, Marske K, Attanayake D 2010. Phylogeography and ecological niche modelling of the New Zealand stick insect *Clitarchus hookeri* (White) support survival in multiple coastal refugia. Journal of Biogeography 37: 682–695.

Burnham KP, Anderson DR 2002. Model selection and multimodel inference. New York, Springer. 488 p.

Bütikofer L, Anderson K, Bebber DP, Bennie JJ, Early RI, Maclean IMD 2020. The problem of scale in predicting biological responses to climate. Global Change Biology 26: 6657–6666.

Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, Wehberg J, Wichmann V, Böhner J 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geoscientific Model Development 8: 1991–2007.

Coops N, Loughhead A, Ryan P, Hutton R 2001. Development of daily spatial heat unit mapping from monthly climatic surfaces for the Australian continent. International Journal of Geographical Information Science 15: 345–361.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré

G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 27–46.

Dymond RJ, Zörner J, Shepherd DJ, Wiser KS, Pairman D, Sabetizade M 2019. Mapping physiognomic types of indigenous forest using space-borne SAR, optical imagery and air-borne LiDAR. Remote Sensing 11: 1911.

Ewers RM, Kliskey AD, Walker S, Rutledge D, Harding JS, Didham RK 2006. Past and future trajectories of forest loss in New Zealand. Biological Conservation 133: 312–325.

Gelman A, Hill J 2007. Data analysis using regression and multi-level/hierarchical models. New York, Cambridge University Press. 625 p.

Guisan A, Graham CH, Elith J, Huettmann F, the NCEAS Species Distribution Modelling Group 2007. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13: 332–340.

Hijmans RJ 2020. raster: Geographic Data Analysis and Modeling. R package version 3.3-7. https://CRAN.R-project.org/package=raster.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965–1978.

Hijmans RJ, Phillips SJ, Leathwick JR, Elith J 2017. dismo: Species Distribution Modelling. R package version 1.1-4. https://CRAN.R-project.org/package=dismo.

Hutchinson MF 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems 9: 385–403.

Hutchinson MF, Gessler PE 1994. Splines — more than just a smooth interpolator. Geoderma 62: 45–67.

Jasiewicz J, Stepinski TF 2013. Geomorphons — a pattern recognition approach to classification and mapping of landforms. Geomorphology 182: 147–156.

Jasiewicz J, Stepinski T, GRASS Development Team 2013. Addon r.geomorphon. Geographic Resources Analysis Support System (GRASS) software, version 7.8. http://webgama.fsv.cvut.cz/grass/grass-cms/grass74/manuals/addons/r.sun.daily.html (accessed 23 September 2020).

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M 2017. Climatologies at high resolution for the earth's land surface areas. Scientific Data 4: 170122.

Land Information New Zealand 2020a. LENZ - Mean annual temperature. https://lris.scinfo.org.nz/layer/48094-lenz-mean-annual-temperature/ (accessed 24 September 2020).

Land Information New Zealand 2020b. NZ River Centrelines (Topo, 1:500k). https://data.linz.govt.nz/layer/50223-nz-river-centrelines-topo-1500k/ (accessed 9 January 2020).

Land Information New Zealand 2020c. NZ Road Centrelines (Topo, 1:250k). https://data.linz.govt.nz/layer/50184-nz-road-centrelines-topo-1250k/ (accessed 19 November 2020).

Land Information New Zealand 2020d. NZDEM North Island 25 metre. https://lris.scinfo.org.nz/layer/48131-nzdem-north-island-25-metre/ (accessed 23 September 2020).

Land Information New Zealand 2020e. NZDEM South Island 25 metre. https://lris.scinfo.org.nz/layer/48127-nzdem-south-island-25-metre/ (accessed 23 September 2020).

Leathwick JR 1995. Climatic relationships of some New Zealand forest tree species. Journal of Vegetation Science 6: 237–248.

Leathwick JR 1998. Are New Zealand's *Nothofagus* species in equilibrium with their environment? Journal of Vegetation Science 9: 719–732.

Leathwick JR 2001. New Zealand's potential forest pattern as predicted from current species-environment relationships. New Zealand Journal of Botany 39: 447–464.

Leathwick J, Morgan F, Wilson G, Rutledge D, McLeod M, Johnson K 2002a. Land environments of New Zealand: A technical guide. Wellington, Ministry for the Environment. 237 p.

Leathwick JR, Wilson G, Stephens RTT 2002b. Climate surfaces for New Zealand. Landcare Research Contract Report: LC9798/126. Hamilton, New Zealand, Manaaki Whenua – Landcare Research. 26 p.

Leathwick JR, Overton JM, McLeod M 2003. An environmental domain classification of New Zealand and its use as a tool for biodiversity management. Conservation Biology 17: 1612–1623.

Lehmann A, Leathwick JR, Overton JM 2002. Assessing New Zealand fern diversity from spatial predictions of species assemblages. Biodiversity & Conservation 11: 2217–2238.

Leschen RAB, Buckley TR, Harman HM, Shulmeister J 2008. Determining the origin and age of the Westland beech (*Nothofagus*) gap, New Zealand, using fungus beetle genetics. Molecular Ecology 17: 1256–1276.

Lilburne L, Hewitt A, Webb TW, Carrick S 2004 S-map – a new soil database for New Zealand. In: Singh B ed. Proceedings of SuperSoil 2004: 3rd Australian New Zealand Soils Conference. Sydney, Australia, 5–9 December 2004.

Lilburne LR, Hewitt AE, Webb TW 2012. Soil and informatics science combine to develop S-map: A new generation soil information system for New Zealand. Geoderma 170: 232–238.

McCarthy JK, Wiser SK, Bellingham PJ, Beresford RM, Campbell RE, Turner R, Richardson SJ 2021. Using spatial models to identify refugia and guide restoration in response to an invasive plant pathogen. Journal of Applied Ecology 58: 192–201.

McGlone MS 1985. Plant biogeography and the late Cenozoic history of New Zealand. New Zealand Journal of Botany 23: 723–749.

McGlone MS, Mildenhall DC, Pole MS 1996. The history and paleoecology of New Zealand *Nothofagus* forests. In: Veblen TT, Hill RS, Read J eds. *Nothofagus*: ecology and evolution. New Haven, Yale University Press. Pp. 83–130.

McNeill SJE, Golubiewski N, Barringer J 2014. Development and calibration of a soil carbon inventory model for New Zealand. Soil Research 52: 789–804.

Meloun M, Militký J, Hill M, Brereton RG 2002. Crucial problems in regression modelling and their solutions. Analyst 127: 433–450.

Müller M, Shepherd J, Dymond J 2015. Support vector machine classification of woody patches in New Zealand from synthetic aperture radar and optical data, with LiDAR training. Journal of Applied Remote Sensing 9: 095984.

Neteler M, Bowman MH, Landa M, Metz M 2012. GRASS GIS: A multi-purpose open source GIS. Environmental Modelling & Software 31: 124–130.

Newsome PFJ, Wilde RH, Willoughby EJ 2008. Land resource information system spatial data layers: Data dictionary. Palmerston North, Manaaki Whenua – Landcare Research. 75 p.

NIWA 2021. National and regional climate maps. https://niwa.co.nz/climate/research-projects/national-and-regional-climate-maps (accessed 31 March 2021).
Palmer DJ, Höck BK, Kimberley MO, Watt MS, Lowe DJ, Payn TW 2009. Comparison of spatial prediction techniques for developing *Pinus radiata* productivity surfaces across New Zealand. Forest Ecology and Management 258: 2046–2055.
Pawson SM, Brockerhoff EG, Meenken ED, Didham RK 2008. Non-native plantation forests as alternative habitat for native forest beetles in a heavily modified landscape. Biodiversity and Conservation 17: 1127–1148.
Perry GLW, Wilmshurst JM, McGlone MS, Napier A 2012. Reconstructing spatial vulnerability to forest loss by fire in pre-historic New Zealand. Global Ecology and Biogeography 21: 1029–1041.
Petras V, Petrasova A, GRASS Development Team 2015. Addon r.sun.daily. Geographic Resources Analysis Support System (GRASS) software, version 7.8. http://webgama.fsv.cvut.cz/grass/grass-cms/grass74/manuals/addons/r.sun.daily.html (accessed 23 September 2020).
Potter KA, Arthur Woods H, Pincebourde S 2013. Microclimatic challenges in global change biology. Global Change Biology 19: 2932–2939.
R Core Team 2020. R: A Language and Environment for Statistical Computing. Vienna, R Foundation for Statistical Computing. http://www.R-project.org/.
Richardson SJ, Peltzer DA, Allen RB, McGlone MS, Parfitt RL 2004. Rapid development of phosphorus limitation in temperate rainforest along the Franz Josef soil chronosequence. Oecologia 139: 267–276.
Roudier P, Burge OR, Richardson SJ, McCarthy JK, Grealish GJ, Ausseil A-G 2020. National scale 3D mapping of soil pH using a data augmentation approach. Remote Sensing 12: 2872.
Stephens JMC, Molan PC, Clarkson BD 2005. A review of *Leptospermum scoparium* (Myrtaceae) in New Zealand. New Zealand Journal of Botany 43: 431–449.
Stewart GH, Veblen TT 1982. Regeneration patterns in southern rata (*Metrosideros umbellata*) – kamahi (*Weinmannia racemosa*) forest in central Westland, New Zealand. New Zealand Journal of Botany 20: 55–72.
Tait A, Sturman J, Clark M 2012. An assessment of the accuracy of interpolated daily rainfall for New Zealand. Journal of Hydrology (NZ) 51: 25–44.
Tarolli P 2014. High-resolution topography for understanding Earth surface processes: Opportunities and challenges. Geomorphology 216: 295–312.
Verrelst J, Camps-Valls G, Muñoz-Marí J, Rivera JP, Veroustraete F, Clevers JGPW, Moreno J 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS Journal of Photogrammetry and Remote Sensing 108: 273–290.
Viscarra Rossel RA, Bui EN 2016. A new detailed map of total phosphorus stocks in Australian soil. Science of The Total Environment 542: 1040–1049.
Viscarra Rossel RA, Chen C, Grundy MJ, Searle R, Clifford D, Campbell PH 2015. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. Soil Research 53: 845–864.
Walker S, Price R, Rutledge D, Stephens RTT, Lee WG 2006. Recent loss of indigenous cover in New Zealand. New Zealand Journal of Ecology 30: 169–177.
Watt MS, Palmer DJ, Kimberley MO, Höck BK, Payn TW, Lowe DJ 2010. Development of models to predict *Pinus radiata* productivity throughout New Zealand. Canadian Journal of Forest Research 40: 488–499.
Weigand A, Abrahamczyk S, Aubin I, Bita-Nicolae C, Bruelheide H, I. Carvajal-Hernández C, Cicuzza D, Nascimento da Costa LE, Csiky J, Dengler J, Gasper ALd, Guerin GR, Haider S, Hernández-Rojas A, Jandt U, Reyes-Chávez J, Karger DN, Khine PK, Kluge J, Krömer T, Lehnert M, Lenoir J, Moulatlet GM, Aros-Mualin D, Noben S, Olivares I, G. Quintanilla L, Reich PB, Salazar L, Silva-Mijangos L, Tuomisto H, Weigelt P, Zuquim G, Kreft H, Kessler M 2020. Global fern and lycophyte richness explained: How regional and local factors shape plot richness. Journal of Biogeography 47: 59–71.
Whitehead D, Leathwick JR, Walcroft AS 2001. Modeling annual carbon uptake for the indigenous forests of New Zealand. Forest Science 47: 9–20.
Williams KJ, Belbin L, Austin MP, Stein JL, Ferrier S 2012. Which environmental variables should I use in my biodiversity model? International Journal of Geographical Information Science 26: 2009–2047.
Wiser SK, Bellingham PJ, Burrows LE 2001. Managing biodiversity information: development of New Zealand's National Vegetation Survey databank. New Zealand Journal of Ecology 25: 1–17.
Wyse SV, Wilmshurst JM, Burns BR, Perry GLW 2018. New Zealand forest dynamics: a review of past and present vegetation responses to disturbance, and development of conceptual forest models. New Zealand Journal of Ecology 42: 87–106.
Xu T, Hutchinson M 2010. ANUCLIM Version 6.1 User Guide. Fenner School of Environment and Society, The Australian National University, Canberra.
Xue J, Su B 2017. Significant remote sensing vegetation indices: A review of developments and applications. Journal of Sensors 2017: 1353691.

Received: 12 January 2021; accepted: 19 April 2021
Editorial board member: George Perry

Supplementary Material

Additional supporting information may be found in the online version of this article:

Appendix S1. Spatial layers included in NZEnvDS.

Appendix S2. Correlation between variables.

The New Zealand Journal of Ecology provides online supporting information supplied by the authors where this may assist readers. Such materials are peer-reviewed and copy-edited but any issues relating to this information (other than missing files) should be addressed to the authors.