CSIRO PUBLISHING

RESEARCH ARTICLE
https://doi.org/10.1071/SR25127

SOIL RESEARCH

Designing and configuring monitoring sites to fulfill Australia's national soil monitoring objectives

Brendan Malone[A,*], Gerard Grealish[B] and Stirling Roberton[C]

For full list of author affiliations and declarations see end of paper

***Correspondence to:**
Brendan Malone
CSIRO Agriculture and Food, Black Mountain, ACT, Australia
Email: brendan.malone@csiro.au

Handling Editor:
Richard Harper

Received: 20 June 2025
Accepted: 24 October 2025
Published: 14 November 2025

Cite this: Malone B *et al*. (2025) Designing and configuring monitoring sites to fulfill Australia's national soil monitoring objectives. *Soil Research* **63**, SR25127. doi:10.1071/SR25127

© 2025 The Author(s) (or their employer(s)). Published by CSIRO Publishing.
This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND).

OPEN ACCESS

ABSTRACT

Context. Priority Action 1 of the *Australian National Soil Action Plan* underscores the need for a nationally agreed framework to measure, monitor, map, report, and share information on soil condition and trends. Such a framework supports best-practice soil management, informs evidence-based decision-making, and guides future investments in securing Australia's soil resources. **Aims.** This study aims to present a statistically robust and operationally feasible sampling design that provides the foundation for a national soil monitoring program. The design seeks to support consistent assessments of soil change, identify emerging threats and opportunities, and enable comparisons between managed and minimally disturbed reference soils. **Methods.** The framework applies a hierarchical, systematic approach to identify representative monitoring sites across Australia's major agricultural regions. The design centres on the establishment of *reference soil areas* (Genosoils) for all mapped combinations of soil and landscape types, enabling comparison with managed soils. Flexibility is embedded to accommodate logistical constraints and to integrate legacy datasets and existing monitoring programs where appropriate. **Key results.** The proposed network comprises 4775 sites nationally, of which 3463 fall within agricultural regions. These sites represent 343 unique soil and landscape entities (Pedogenons) and encompass more than 1000 distinct soil and land-use combinations. The framework builds on decades of collaborative, multi-organisational work to develop national soil, landscape, and land-use datasets, which underpin the digital foundation for monitoring. **Conclusions.** The design provides a statistically sound and adaptable foundation for long-term, national-scale soil monitoring in Australia. Its structure enables consistent comparison, integration with existing data resources, and scalability for diverse monitoring objectives. **Implications.** While developed for the Australian context, the approach is broadly transferable. With access to suitable soil and geospatial data, the principles and methodology can be adapted internationally to support comparable monitoring efforts, contributing to global initiatives in soil condition assessment and sustainable land management.

Keywords: Australia, digital soil mapping, Genosoil and Phenosoil physiographic regions, National Soil Monitoring Program, Pedogenon mapping, sampling design, soil condition assessment.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

Introduction

Soils are a vital resource for water, energy, and food security, underscoring the need for long-term coordination of soil monitoring networks. Such networks provide essential baseline information and track the status and trends of soil resources, enabling early-warning mechanisms to identify and delineate soil threats. Fundamentally, soil monitoring supports evidence-based policies that intend to incentivise sustainable soil management (Van Leeuwen *et al*. 2017).

The global literature extensively outlines the objectives and considerations for soil monitoring. For example, Arrouays *et al*. (2018) provide a comprehensive review on establishing and maintaining large-scale monitoring networks across regions and countries. Broadly paraphrasing de Gruijter *et al*. (2006), the primary objectives include the following:

1. Comprehensive coverage: ensuring sites adequately represent the range of climate, landforms, geology, soils, and land use/management interactions across a defined spatial extent.
2. Status monitoring: characterising and quantifying the status of soils and tracking changes over time, such as topsoil carbon content under varying land use and climatic conditions.
3. Trend/Effect monitoring: assessing the effects of pressures or drivers on soils to determine both the status and the underlying causes of observed changes.
4. Regulatory/Compliance monitoring: determining whether soils meet established standards or targets.

Australia has long recognised the need for a National Soil Monitoring Program (NSMP) under these broad objectives. Substantial investments and intellectual efforts over the years have explored how such a program might be implemented. Key contributions include the following:

1. McKenzie *et al.* (2002): Monitoring Soil Change: Principles and Practices for Australian Soil Conditions
2. Baldock *et al.* (2010): Building a Foundation for Soil Condition Assessment
3. Grealish *et al.* (2011): National Soil Condition Monitoring Program for Soil pH and Soil Carbon
4. Wilson *et al.* (2021): Designing a National Soil Monitoring Program

These efforts have delivered significant outcomes. For example, the Soil Carbon Research Program (SCaRP; Baldock *et al.* 2013) has enhanced understanding of soil carbon stocks across Australia's managed landscapes. National-scale digital mapping of soil organic carbon concentrations and stocks has relied heavily on SCaRP data (Román Dobarco *et al.* 2023*a*; Wadoux *et al.* 2023). Such data have not only provided insights into carbon retention across diverse landscapes but also quantified the potential for carbon gains through improved soil management practices (Karunaratne *et al.* 2024). This work holds critical implications for carbon economies, climate change mitigation, and key soil functions, including physical structure, water retention, and filtering. Current efforts are revisiting SCaRP sites to empirically assess land management impacts on soil organic carbon (Karunaratne *et al.*, in prep.).

A key lesson from these programs is that while establishing a monitoring network is a substantial task, maintaining long-term coordination for resampling and revisitation is even more demanding. Funding cycles, shifting priorities, and changing motivations often disrupt or abandon programs altogether. Similar challenges exist internationally. However, enduring examples such as the European Commission's LUCAS Soil Programme (Orgiazzi *et al.* 2018) and France's RMQS Soil Quality Monitoring Network (Jolivet *et al.* 2022) demonstrate the value of well-designed, sustained efforts.

The renewed emphasis on a NSMP in Australia began with the 2017 report, *Restore the Soil, Prosper the Nation*, by Australia's first National Soils Advocate, Major General the Honourable Michael Jeffery, AC, CVO, MC (Retd) (1937–2020). Concerned about soil security for future generations, the Advocate urged Australia to adopt national policies to protect and understand its soil resources. This vision led to the National Soil Action Plan 2023–2028 (DAFF 2023) – the first plan under the 20-year National Soil Strategy.

The Action Plan identifies four priority actions, including Priority Action 1: *Develop an agreed national framework to support measurement, monitoring, mapping, reporting, and sharing of soil state and trend information to inform best practice management, decision-making, and future investment.* This Priority Action has subsequently led to the establishment of a NSMP, funded through the Australian Government Natural Heritage Trust (Department of Agriculture, Fisheries and Forestry) in collaboration with CSIRO. It also aligns with the Australian Government's National Soil Strategy (DAWE 2021), specifically Goal 3: *Strengthen soil knowledge and capability*, and Objective 3b: *Measure the benefits of improved soil management by tracking changes in soil condition.*

This study presents a site configuration for a NSMP tailored to Australia's unique conditions. It aims to meet the core objectives of soil monitoring while accounting for the country's diverse soils, terrain, climate, and land use, as well as the logistical challenges of sampling across a vast continent dominated by remote and regional areas.

The guiding principles of this design include the following:

1. Systematic and transparent approach: site selection is data-driven, relying on consistent national-scale environmental and biophysical datasets, including national digital soil mapping products from the Soil and Landscape Grid of Australia (SLGA; Malone *et al.* 2025).
2. Flexibility: the design must adapt to logistical challenges, such as land access issues, without compromising core principles. It also allows for expert assessments to make on-the-ground adjustments.
3. Integration: the design establishes processes to incorporate existing monitoring programs, leveraging prior efforts, minimising duplication, and ensuring efficiency.

This study is structured as follows:

1. Key considerations for establishing the NSMP site configuration are outlined, with a clear scope of work to be conducted in this investigation.
2. General design principles of the NSMP are established.
3. Data-driven processes for identifying site locations across Australia for the NSMP are detailed.
4. Integration with existing monitoring networks is addressed, including a simple framework for local site relocation where primary sites are unsuitable while operators are in the field.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/17843658/sr25127.pdf by guest on 25 January 2026

Materials and methods

General design considerations for Australia's NSMP

The design phase of a soil monitoring program requires careful consideration of its structure, feasibility, and long-term value. Key decisions include sampling design, site revisitation strategies, data collection methods, and integration with existing initiatives. These considerations are outlined below, along with an assessment of whether each is within the scope of the present work.

Network design and sampling considerations

A fundamental design decision is the choice between design-based sampling and purposive sampling. Design-based sampling supports unbiased inferences about trends and spatial means, whereas purposive sampling prioritises spatial coverage and enables detailed mapping of soil variables and uncertainties. Grid-based purposive approaches, widely used in global soil monitoring networks (Arrouays *et al.* 2018), provide a practical balance of representation and efficiency.

The intensity and frequency of sampling depend on resource availability and program objectives. Tools such as power analysis and knowledge of variability in target variables (Chappell *et al.* 2013) can optimise site numbers, though these approaches remain underutilised in broad-scale programs.

With respect to sampling frequency, soil properties typically change gradually; however, abrupt shifts can occur in response to extreme events or changes in land management. A five-year interval is commonly regarded as a practical compromise – frequent enough to detect meaningful trends while remaining feasible under logistical constraints. Nonetheless, implementing this interval at scale is often limited by budgetary considerations and the direction of government policy.

The choice to revisit monitoring sites or rotate revisit locations within the same sampling strata significantly shapes program design. Revisiting the same sites strengthens temporal trend analysis by reducing variability from differences in soil properties and landscape settings. Rotating sites, on the other hand, improves spatial coverage by sampling new areas within similar soil–landscape contexts (de Gruijter *et al.* 2006). A hybrid strategy can combine these benefits through methods like serially alternating or supplemented panel designs (Brus and de Gruijter 2011). However, implementing such designs on a national scale, especially in a vast country like Australia, remains untested.

Capturing relevant soil data

Historically, national soil monitoring proposals in Australia have focused on a limited set of variables, such as soil carbon and pH, reflecting the primary priorities at the time. However, as the need to assess broader soil functions and ecosystem services increases, future programs must incorporate a more comprehensive set of soil indicators, which cover the chemical, physical, and biological domains. Although determining the specific attributes to be measured is beyond the scope of this study, it remains a crucial consideration for future program development. The design presented in this study is based on the assumption that monitoring will encompass multiple soil attributes, ensuring a more holistic approach to soil assessment.

Integration with existing monitoring programs

Building on existing soil monitoring efforts is essential to avoid duplication, leverage historical data, and enhance program efficiency. Many of Australia's current soil monitoring schemes, summarised by Wilson *et al.* (2021), provide a foundation for integration and provide a mix of national- and state-driven efforts. Close collaboration with these existing networks is necessary to harmonise protocols, for example, by aligning sampling depths or methods with NSMP requirements. Although this study outlines a strategy for integrating existing networks, resolving specific technical and procedural issues will require further detailed work.

Data management and soil archiving

Robust data management and archiving systems, along with protocols for landholder communication, secure storage of contact information, and appropriate data provision, are all critical to the long-term success of a soil monitoring program. These elements ensure continuity, support reporting, and enable evidence-based decision-making, including the facilitation of future site revisits. However, detailed treatment of these aspects – including related issues such as data ownership, open access, and intellectual property – falls outside the scope of the present study and will require attention in both planning and operational phases.

Establishing design principles for a purpose-built NSMP for Australia

Taking into account the overarching design considerations described above, the proposed NSMP for Australia does not deviate substantially in principle from the purposive approaches outlined by Baldock *et al.* (2010) and Grealish *et al.* (2011). To effectively assess the impacts of land management, the program prioritises monitoring in agricultural areas, peri-urban zones, and regions with identified vulnerable soils. It also extends to areas targeted for future agricultural expansion, including Australia's extensive but less intensively managed rangelands.

Baldock *et al.* (2010) proposed prioritising sampling within physiographic areas – regions characterised by shared climate, geology, and soil properties. This approach remains a logical framework for site selection, as these areas provide a coherent basis for interpreting soil changes in relation to geological processes, both current and historical climatic patterns, and other local landscape influences. Additionally, utilising these well-defined regions offers practical benefits, streamlining resource allocation and program implementation.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

Baldock *et al.* (2010) also proposed prioritising sites based on soil and landscape representativeness, land-use intensity, and resilience to change, with particular attention to soil carbon and pH. In the new design, the NSMP de-emphasises this specificity, placing greater emphasis on identifying regions characterised by diverse land-use practices, particularly those dominated by agriculture, that have undergone widespread detectable change or are expected to experience significant intensification in the future.

Site numbers, sampling approach, and logistical considerations

Currently, 55% of Australia's land area is under some form of agricultural enterprise. For context, France's Réseau de Mesures de la Qualité des Sols (RMQS) employs a 16 km square grid, which if applied to Australia, would equate to approximately 264,688 sites. In contrast, feasible site numbers for Australia's NSMP (based on project funding allocations) is in the range of 2000–5000, corresponding to sampling densities of one site per 2100 and 840 km^2, respectively.

At this scale, grid-based sampling is not necessarily the most efficient or representative approach, particularly when the number of sites is limited. Alternative methods such as conditioned Latin Hypercube (cLHC) sampling, as recommended by Baldock *et al.* (2010) and developed by Minasny and McBratney (2006), can provide more effective coverage of the primary climate, landscape, and land management combinations once the number of sites has been determined. However, cLHC is rigidly tied to maximising representation of the global feature space, sometimes at the expense of local conditions. This often results in the selection of sites in spatially rare environments, such as transition zones, which may not be representative of the dominant landscapes being monitored. While cLHC effectively captures variability, it is also less adaptable to needs such as site relocation, integration with existing monitoring sites, or adjusting sampling intensity to prioritise vulnerable areas or de-emphasise areas with low current and future land utility. For example, it would be logical to deploy more sites in intensively managed dryland and irrigated cropping areas, where changes are expected to occur rapidly, compared to rangeland areas that experience minimal management impacts.

Regarding logistical considerations, coordinating a national-scale hybrid design such as those detailed by Brus and de Gruijter (2011) also poses logistical issues. Flexibility is needed to account for issues such as site access, unexpected relocations, and other unforeseeable events. Additionally, integrating existing soil monitoring networks adds complexity to the process. A practical approach is to plan all monitoring sites during the initial design phase. After accounting for adjustments to site locations made during fieldwork to establish the network, all sites would be revisited after a 5-year interval. Subsequent resampling campaigns would then occur every 5 years as planned.

Frameworks for assessing soil changes and Pedogenon mapping and its role in site selection

The relationships between land management impacts and soil changes are not uniform, reflecting variations in soil properties and the landscape in which they are situated. A key principle is to select sites grouped by shared soil and landscape attributes, enabling comparisons between managed and minimally disturbed conditions. Frameworks such as the Soil Security Assessment Framework (Evangelista *et al.* 2023), the Soil Health Assessment Protocol and Evaluation Tool (Nunes *et al.* 2021), and the 'Soil Health Gap' concept specific to soil carbon (Maharjan *et al.* 2020) emphasise establishing a reference condition against which measured properties can be benchmarked. Román Dobarco *et al.* (2023*b*) propose comparing soils within the same setting, termed Phenosoils, against their minimally disturbed counterparts, termed Genosoils, to better infer land management impacts.

This approach hinges on Pedogenon mapping, a process that delineates clusters of homogeneous environmental variables representing stable soil-forming factors. These clusters, termed Pedogenon classes, correspond to quasi-steady-state soil systems under specific conditions (Román Dobarco *et al.* 2021). Within each Pedogenon, Genosoils represent minimally disturbed conditions, while Phenosoils reflect varying levels of anthropogenic pressure. Pedogenon mapping typically incorporates terrain and landform data derived from digital elevation models, long-term climate records, geological information that could be inferred from gamma radiometrics, and vegetation data to represent soil-forming factors.

Recent empirical work provides clear precedent for operationalising this approach. For example, Jang *et al.* (2023) demonstrated its validity in the Lower Namoi Valley in north-western New South Wales, one of Australia's most productive agricultural regions, showing that soil organic carbon losses could be robustly quantified even in intensively managed landscapes. At the continental scale, Styc *et al.* (2025) applied Pedogenon mapping nationally to compare Genosoils and Phenosoils, confirming that the framework can systematically detect and interpret soil organic carbon change across diverse environments.

Refinement of Pedogenon mapping processes is however necessary to accommodate established design principles, such as site numbers (2000–5000), differential sampling intensity, and integration with existing soil monitoring networks. Physiographic regions, as previously described, need also be incorporated as an additional dataset. The Habitat Condition Assessment System (HCAS; Harwood *et al.* 2021) is another tool with potential applications in identifying Genosoils (Román Dobarco *et al.* 2023*b*). By scoring land conditions on a scale from 0 (completely removed habitat) to 1 (habitat in optimal condition), HCAS provides a quantitative basis for identifying minimally modified ecosystems that may serve as proxies for the least disturbed soils. These data therefore complement the use of Pedogenon units for identifying both

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

reference (Genosoils) and impacted (Phenosoils) conditions within a given landscape.

In summary, while the principles outlined above provide a robust foundation, further work is required to refine Pedogenon mapping, account for specific logistical challenges, and ensure compatibility with Australia's unique environmental context and monitoring objectives. Integration of physiographic regions and careful alignment with existing soil monitoring efforts will enhance the feasibility and effectiveness of the program.

Taking these considerations into account, we present a novel sampling design for adoption in Australia's NSMP:

- –Follows a hierarchical approach,
- –Structured with meaningful spatial strata,
- –Focused on obtaining different land use and soil type combinations,
- –Can incorporate legacy monitoring sites,
- –Allows for relocation of sites in areas where site access is not possible,
- –Considers landscape change, focusing sampling efforts in areas that are most vulnerable to decline.

Principles for identification of sites for the NSMP

The design of the NSMP follows a hierarchical structure, with each level intended to represent increasing homogeneity in soil, landscape, and land-use characteristics. The terminal level is where soil monitoring sites are identified for establishment. Fig. 1 provides a schematic representation of this structure to illustrate the overall design.

Monitoring regions

Physiographic regions provide a pre-established framework for broadly classifying landscapes based on shared climate, geology, and topographic characteristics. These broad segmentations of Australia's land surface serve as an initial reference point for further refinement into more detailed landscape groupings. At subsequent levels of the hierarchical structure, these divisions help define areas with greater homogeneity in terms of soils, landscape features, and land-use patterns. Additionally, these broad groupings are labelled entities with descriptive attributes, making them valuable for managing and operationalising logistical components of the NSMP during its implementation.

As will be discussed further, the allocation of sample sites and their sampling density across the country are closely tied to the characteristics of physiographic regions, as well as their structure and the environmental conditions they encompass.

Monitoring units

The primary objective in defining monitoring units is to refine environmental variation into more homogeneous land parcels within a monitoring region. A key factor in the sampling design is the extent of partitioning that can be applied within a region. In cases where the allocated sampling density is low, a monitoring region may serve as the monitoring unit if site numbers fall below a defined threshold. Conversely, a high density of sampling locations would necessitate further partitioning of the region into smaller, relatively more homogenous land parcels.

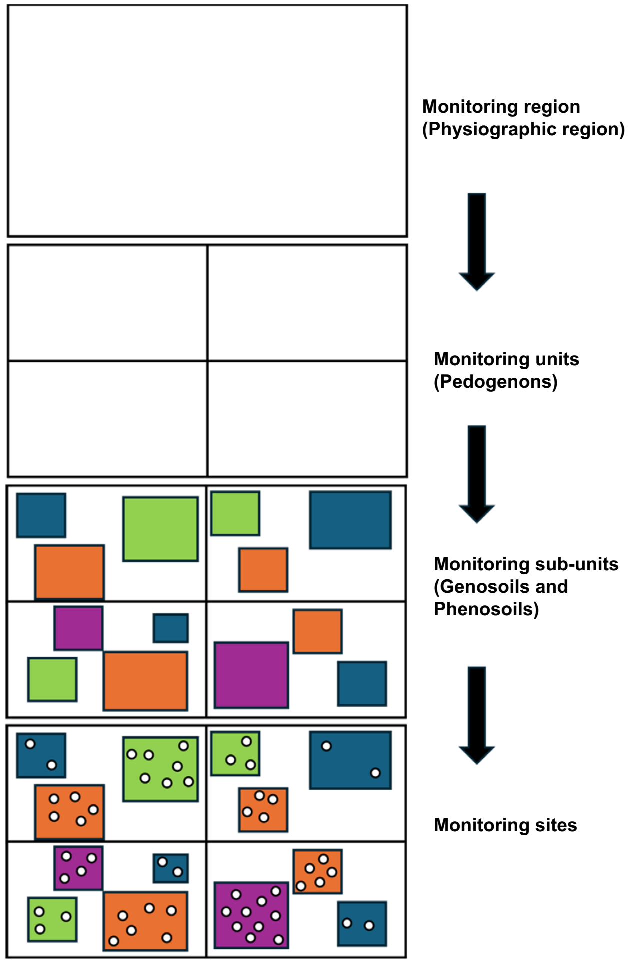

Fig. 1. Schematic illustration of the framework underpinning the proposed soil monitoring design. At the terminal level are the monitoring sites, selected from sampling frames representing soil monitoring sub-units. These sub-units are defined by similarities in soil and land use. Monitoring units and regions represent progressively broader divisions of the land surface, reflecting soil, landscape, climate, and geological patterns.

Equating to other existing definitions, a monitoring unit as we describe here is equivalent to a Pedogenon class and would be established using a clustering approach. Given the complexity and extent of available datasets (discussed later), unsupervised classification algorithms such as k-means clustering or similar are well-suited for this process.

Monitoring sub-units

Aligned with the delineation of monitoring units, further subdividing the landscape along a gradient – from minimal anthropogenic influence (i.e. Genosoils) to greater anthropogenic pressure (i.e. Phenosoils) – enhances the ability to

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

assess land management effects within specific soil and landscape contexts. This further subdivision results in non-contiguous monitoring areas, and prioritises the most prevalent combinations of soils and land uses. While smaller or less common combinations may be overlooked, focusing on dominant combinations offers a key advantage: it maximises the areal coverage of the monitoring scheme while ensuring alignment with both design objectives and budgetary constraints.

Monitoring sites

Monitoring sub-units serve as individual sampling frames from which monitoring sites are selected. Site selection within these frames can follow either a design-based (probability sampling) approach or a purposive strategy, such as cLHC sampling. Flexibility is built into the design by storing location information for all potential sites within each sampling frame, rather than only the relatively few sites initially selected through random or purposive methods.

As a result, every point within the sampling frame remains a viable monitoring site and can be selected as a replacement if needed, whether from a desktop-based review or directly in the field. This flexibility ensures that expert judgment can be incorporated into the final decision-making process when establishing a monitoring site. Additionally, existing monitoring sites that align with a specified sub-unit can seamlessly replace a site within the sampling frame, provided they correspond with the defined characteristics of the monitoring sub-unit.

Underpinning datasets used to establish the NSMP

Key datasets used in the soil monitoring design work are described below. Images of these datasets are displayed in Fig. 2.

Physiographic regions

Pain *et al.* (2011) provide a historical timeline of the efforts to map the physiographic regions of Australia. In their work, they produced an updated version, making use of improved data and advances in GIS technology. Similarly, in the present work we further modified the physiographic regions to achieve a more data-rich characterisation using a machine learning modelling approach (Malone *et al.*, in review). This new mapping incorporates multiple digital environmental datasets, but, as with earlier updates, the resulting changes in regional boundaries are relatively subtle. However, for this updated version, these boundaries are guided by machine learning inference and learned from the extensive data provided to these models.

There are 220 mapped physiographic regions across Australia and near-shore islands. Only one region (10101 – North Reefs) was excluded due to an absence of available spatial data.

Digital soil mapping

The SLGA provides consistent digital soil information across the country (Malone *et al.* 2025). It offers high-resolution (90 m) mapping of various soil attributes, with each map generated through customised geospatial modelling informed by soil observations collected by CSIRO, state and territory governments, universities, and other sources (Searle 2020).

For integration into the NSMP sampling design, digital maps of soil classification (at the Order level of the Australian Soil Classification; Searle 2021) and subsoil dominant soil colour (Malone 2022) were included. Together, these soil classification and colour datasets provide a multi-attribute characterisation of soil differences and similarities, reflecting key soil-forming factors and both present and past soil processes. Adoption of these layers over state or regional soil polygon maps allows for a nationally consistent sampling design.

Land-use mapping

National-scale land-use mapping is developed by ABARES (2022). Data used specifically were that of the 2015–2016 land-use classification. Land use is specified according to the Australian Land Use and Management (ALUM) Classification ver. 8; the classifications used in this work were the PRIMV8 and CL18 types. The native resolution of these data is 250 m and is distributed in raster format.

Habitat Condition Assessment System

The HCAS is a remote-sensing based algorithm for assessing the condition of habitats for native terrestrial biodiversity (Harwood *et al.* 2021). The HCAS was designed to differentiate when an ecosystem's condition results from natural dynamics through to anthropogenic influence, considering the temporal and ecological variability of natural ecosystems (Harwood *et al.* 2021). The HCAS uses as input abiotic environmental data (e.g. soil, landform, and climate), remote-sensing data, and reference sites condition data. The spatial ecological model is based on the notion that sites with similar abiotic environmental conditions would have a similar remote-sensing signal averaged over time. The reference sites are assumed to be the least modified for that habitat type and are identified based on explicit knowledge (field observations) or inferred from multiple spatial data sources (land tenure, land cover, and remote sensing). In the present work, HCAS ver. 2.1 (2001–2018; Harwood *et al.* 2021) was used. The native resolution of these data is 250 m. The HCAS scores range between 0 (completely removed habitat) and 1 (habitat in best possible condition).

Library of digital environmental data

A nationally consistent library of digital environmental data (Searle *et al.* 2022) was sourced with the primary intention of mapping or delineating the monitoring units across the country. These data were all obtained from a broad range of original data sources. Through a processing sequence

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

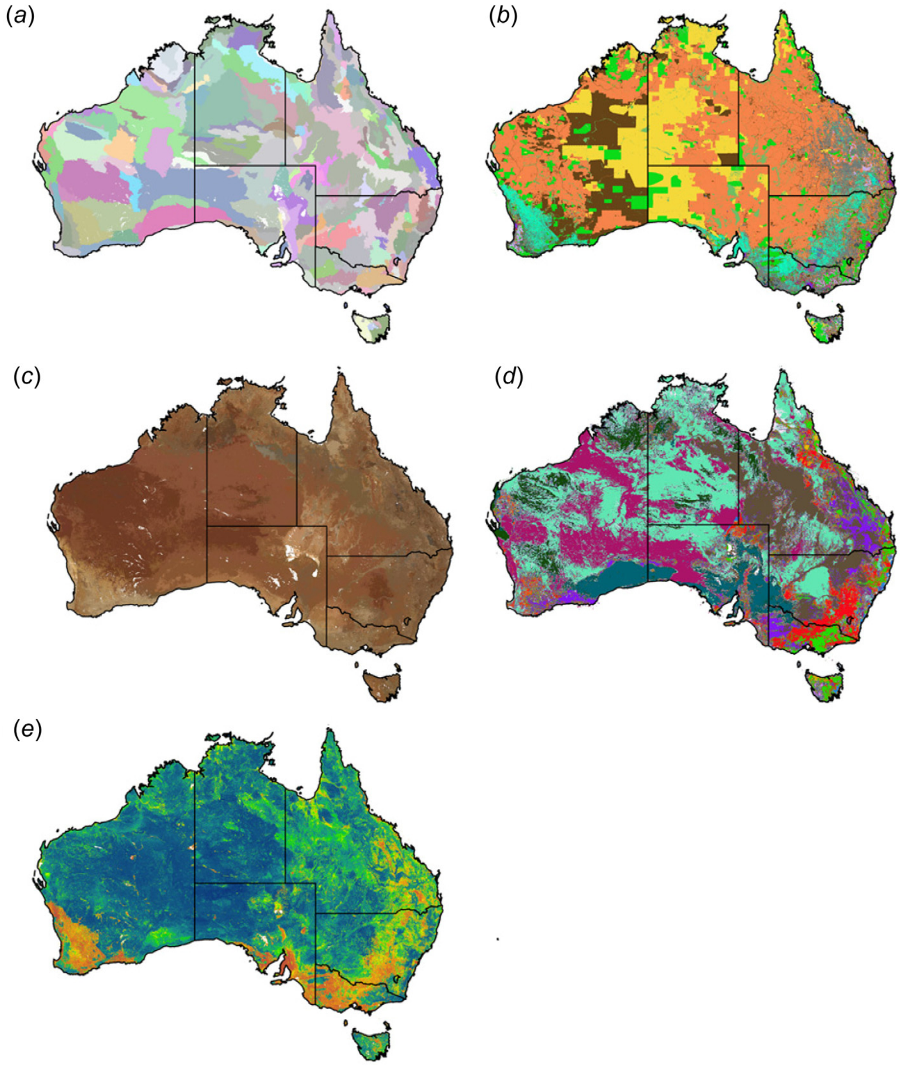

Fig. 2. Underpinning datasets that were used to guide the identification of site locations for the proposed National Soil Monitoring Program: (*a*) physiographic regions (updated from Pain *et al.* (2011) by Malone *et al.* (in review)), (*b*) ABARES Land Use Mapping (ABARES 2022), (*c*) subsoil colour (Malone 2022), (*d*) digital mapping of Australia Soil Classification Soil Orders (Searle 2021), and (*e*) Habitat Condition Assessment System mapping (Harwood *et al.* 2021). In addition to these data, a comprehensive collation of environmental datasets was also used (Searle *et al.* 2022).

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

they were all co-registered to the same spatial resolution and coordinate reference system. Each of the data layers were grouped according to a soil forming factor-based classification in terms of climate, parent material, biology, relief, soil, and location. The present work only considered those relating to climate, parent material, and relief – and to balance numbers of layers per each grouping, principal component analysis was used. All subsequent data analyses using the environmental data were based on these derived components grouped by soil forming factor.

Spatial data processing

All data sources where necessary were resampled and reprojected to geographic coordinates in World Geodetic System 1984 (WGS84) and a 0.000833° cell size (~90 m grid). All the co-registered spatial data were then clipped by the boundary extents of each physiographic region, thus creating 219 individual libraries of spatial data.

Identification of site locations

This section describes the data analysis work following the general design principles just established and using and interrogating the various spatial datasets also described. It should be noted that to ensure that total sample size remained within the established 5000-site limit (overall sampling design specification), extensive experimentation and computational analysis were conducted. These analyses explored various combinations of adjustment to site densities, monitoring unit delineations, and monitoring sub-unit specifications – particularly in terms of defining dominant soil and land-use combinations.

Establishing a general sampling intensity

A maximum sample size of 5000 sites was established. For each physiographic region, the proportion of land under agricultural use was calculated using the PRIMV8 classification (ABARES 2022). Site numbers were then allocated to regions in proportion to their share of Australia's total agricultural land area.

Secondly, HCAS information was brought in, where the mean value for each physiographic region was derived. This value acted as a weighting factor, which was used to modify sample size numbers. Through several iterative assessments, a linear relationship with gradient −1 and *y*-intercept 1.5 was selected to modify the mean HCAS value to a weighting factor that was used to multiply with the allocated site number size in each physiographic region. This ensured that where there was an average value of HCAS (0.5), no change in the original allocation occurred. Where HCAS fell below and above 0.5, sample size allocation increased and reduced, respectively. From this analysis, the resulting sample size was reduced to 3500 sites.

The combination of land use and habitat condition in the initial site allocation process meant that physiographic regions containing large areas of agricultural land would receive greater allocation of sites compared with regions of lower coverage of agricultural land, but this was moderated based on habitat condition, where intact environments would not be allocated their original allocation. Where habitat conditions were less intact, these regions got a higher allocation of sites than they normally would if the analysis was only based on agricultural land coverage.

Consolidation of monitoring regions

An initial rationalisation of the monitoring regions was performed to assist with managing the number of sites, i.e. to keep total site numbers close to 5000. Where the site allocation to a physiographic region was less than 6, that region was dropped from further analysis. The value of 6 was chosen on the basis that within each monitoring unit there must be, at a minimum, one Genosoil monitoring sub-unit and two to three Phenosoil monitoring sub-units, each requiring a minimum of two to three sample sites for the monitoring design. This rationalisation resulted in the dropping of 79 monitoring regions from further analysis.

Mapping extents of monitoring units

A spatial data analysis was conducted for each of the 140 monitoring regions, beginning with an assessment of the number of sites allocated to each region. If a region contained more than 20 sites, fuzzy k-means classification was applied using a region-specific library of digital environmental data to delineate monitoring units. Consequently, monitoring units were not always contiguous. The number of monitoring units within a region was determined based on site allocation as follows:

- 6–19 sites: no clustering (monitoring region = monitoring unit)
- 20–29 sites: two monitoring units
- 30–39 sites: three monitoring units
- Increasing in increments of 10, up to 100–200 sites: 10 monitoring units.

This discretisation process aimed to create smaller land parcels with greater environmental homogeneity. However, in regions with lower site allocations, the level of discretisation was naturally constrained. This constraint is an inherent feature of the sampling design, ensuring that the total number of sites remains within the financial resource limitations of the program.

Independent of site allocation constraints, clustering could still be performed within a region to further subdivide areas into monitoring units. In such cases, the number of clusters can be guided by standard statistical approaches (e.g. the elbow method, which identifies points of diminishing returns in variance reduction) or by expert judgement, depending on the context. It should be noted, however, that increasing the number of clusters produces a near-multiplicative increase in sample size, which in turn affects subsequent discretisation based on soil and land-use associations.

In summary, this process sets the sample size allocation at the broadest organisational level (monitoring regions) and also guides how finely regions should be divided into more homogeneous units based on soil-forming factors. Where more localised assessments are required, clustering can be applied independently of site number constraints, allowing the creation of homogeneous regions tailored to specific study needs.

Identification of monitoring sub-units

Within each monitoring unit, an associated library of environmental data, soils data, and land-use information was compiled. The identification of Genosoil and Phenosoil monitoring sub-units followed a process that involved determining dominant combinations of soil classification, subsoil colour, and land-use type. While multiple combinations of these variables can exist within a monitoring unit, the focus on dominant combinations maximises the spatial extent of land that can be reliably monitored.

In practical terms, dominance was determined by selecting only those soil and land-use combinations with sufficient data coverage, defined as areas containing at least 2000 pixels (~1620 ha at 90 m × 90 m resolution). This threshold was chosen as a pragmatic balance: it ensures that each monitoring unit encompasses an appreciable land area while also providing enough scope to identify alternative sites if access or logistical issues arise. Although somewhat arbitrary, this criterion helps avoid very small or fragmented areas that would be less

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

useful for national monitoring, while still retaining flexibility in site selection.

Genosoils were distinguished from Phenosoils based solely on land-use type. According to the ABARES PRIMV8 classification (ABARES 2022), Genosoils were associated with conservation and natural environments, whereas Phenosoils were identified based on other land uses classified under CL18. These land uses included various forms of agriculture, such as dryland horticulture, dryland cropping, grazing modified pastures, grazing natural vegetation, irrigated cropping, irrigated horticulture, irrigated pastures, plantation forests, and production native forests.

Monitoring site establishment

Establishing monitoring site locations within each monitoring sub-unit began by randomly selecting the required number of sites from the given sampling frame. While randomisation effectively eliminates selection bias, it does not prevent the possibility of selecting implausible sites. These could range from obvious issues, such as sites positioned on roadways, buildings, or waterways, to more subtle discrepancies where land use (and sometimes soil type), as observed in aerial imagery, conflicts with the classification defined for the monitoring sub-unit.

To address these issues, each selected site was carefully checked for plausibility. Since this process could not be fully automated, a custom Shiny web-application was developed with mapping and visualisation capabilities to facilitate site-by-site assessments. If a site was deemed implausible for any reason, a new site was generated by selecting (at random) an alternative from the available sampling frame. This process was repeated until a plausible replacement site was identified.

Additional considerations were necessary, particularly in remote areas. For example, proximity to road networks was often an important selection criterion, with many sites in remote regions located >50 km from an unsealed road or farm track. In some cases, an additional constraint was applied by limiting the neighbourhood size around a rejected site from which a replacement could be selected. In this design phase, site rejections were managed internally using agreed processes, although detailed reasons for rejection were not systematically recorded. For future operational implementation, a formal system of documenting rejections and their justification will be necessary to ensure transparency, allow reassessment where required, and maintain consistency across field teams.

Further refinement of sample design to demarcate priority sample locations

Strategising the rollout of site visits and data collection efforts is essential for resource and labour planning. To rationalise this process, an initial assessment was conducted using rainfall isohyets, identifying monitoring units where the mean annual rainfall exceeded 350 mm. This pragmatic filter was chosen to prioritise higher-rainfall regions where soil change is more likely to be detectable and where field operations are generally more feasible, while de-emphasising very arid environments in the early stages of establishment. Post-analysis, however, revealed that this approach inadvertently excluded key areas, particularly those associated with irrigated agriculture, from priority site demarcation.

To address this, a revised listing was created by assessing the diversity of land-use types within each monitoring unit. Monitoring units with only a single land-use type (excluding conservation and protected areas) were excluded from priority status. This rationalisation of sample sites was not intended to eliminate potential monitoring locations but rather to optimise resource allocation and facilitate a possible phased implementation of the program.

It should be noted that this prioritisation inevitably reflects an agricultural productivity lens. Under the current resource and logistical constraints, regions with intensive agricultural land uses were emphasised to maximise efficiency, ensure alignment with national policy priorities, and increase the likelihood of detecting contemporary change. While this meant that many rangeland regions were assigned a lower priority, the framework is designed to be flexible and can readily incorporate additional rangeland sites in future phases of the program or through complementary projects.

Finally, we acknowledge that rationalisation may still result in some sites being geographically isolated, which could pose challenges for field operations. The optimisation of site visits and sampling trips (e.g. using road networks to define efficient sampling clusters) is considered beyond the scope of this design paper and will be addressed in subsequent operational work documents and publications.

Inclusion of other soil monitoring designs into design

Integrating existing monitoring sites into the national program is both desirable and challenging. Many sites have already generated valuable data, but their potential contribution depends on the soil parameters measured, the quality and consistency of those measurements, and the resources available to harmonise them within the national framework. Careful assessment of these factors is therefore required before integration can occur.

The practical approach for testing integration of existing networks into the national design was demonstrated using the TERN Surveillance AusPlots dataset (Sparrow *et al.* 2020). Existing site locations were intersected with the defined monitoring sub-units to identify matches, and only those sites falling within a sub-unit were considered suitable for integration. Where matches occurred, AusPlots sites were substituted directly for sites selected by the design process, ensuring that overall sample density was preserved. This example illustrates the broader process that can be applied to other networks and datasets held by state and territory agencies, universities, and other institutions.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/17843651/sr25127.pdf by guest on 25 January 2026

Results

Summary of NSMP site distribution

The primary level of organisation for the proposed design of the NSMP is based on physiographic regions. Below this, the design follows a flexible process that can be structured in either a 'global' or a 'local' manner, or as a combination of both. The 'global' method determines sample densities for each monitoring region based on the relationship between habitat condition and the proportion of agricultural land relative to the total area. This sets in motion the delineation of monitoring units and their composition, though it remains constrained by the total number of available sample sites. The 'local' method removes this constraint, instead prioritising the definition of homogeneous regions based on soil-forming factors. This approach is typically applied when specific monitoring regions are of particular interest.

Applying the 'global' approach in this study resulted in a total site allocation of 4000–5000 sites. The difficulty in fixing an exact number stems from the disconnection between the initial recommended allocation and the eventual combination of dominant soil and land-use types, which are only determined after monitoring units have been derived. In Tasmania, the 'global' approach led to relatively low site allocations despite the region's agricultural diversity. Compared to the rest of the country, agricultural land in Tasmania makes up a much smaller proportion of total land area, which resulted in lower site numbers. To account for this, the 'local' approach was applied, focusing on the four physiographic regions within the state. An initial target of 100 sites was set, but after assessing dominant soil and land-use combinations, 192 sites were ultimately selected.

To illustrate how sites were allocated across monitoring regions, Fig. 3 presents an abstraction of the data analysis used to determine sampling intensity through the 'global' approach. One example is Physiographic Region 10510 – Gunnedah Lowland, which covers just over 7000 km^2 and is relatively small compared to other regions. The average HCAS value for this region was 0.53, indicating a habitat condition that is approximately average. In the initial allocation, eight sites were assigned to this region, meaning further subdivision was not pursued. The dominant soil type here was Brown Vertosol, which constitutes a single monitoring unit. Within this unit, three dominant land uses were identified: cropping, grazing on modified pastures, and grazing on unimproved pasture/vegetation. To ensure that all dominant soil and land-use combinations met the minimum site requirement, the final allocation for this monitoring unit included one Genosoil reference monitoring sub-unit and three Phenosoil monitoring sub-units, bringing the total number of sites to 11.

Another example is Physiographic Region 20305 – Riverine Plain, which spans just over 105,000 km^2. The average HCAS value for this region was slightly above average (0.66), resulting in only a minor reduction in the

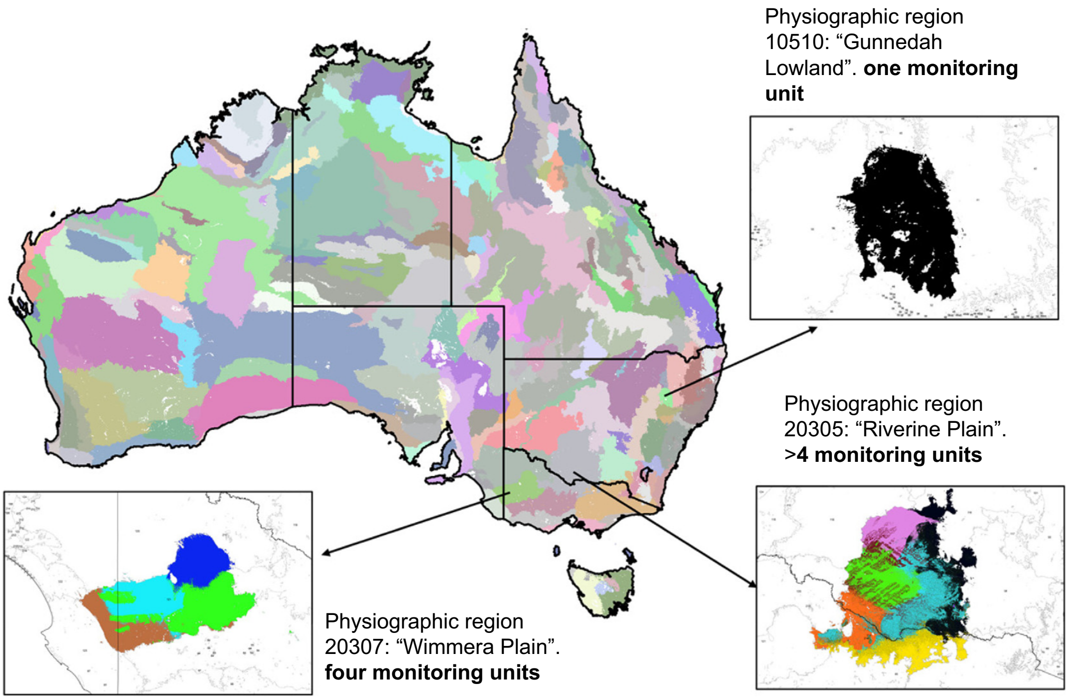

Fig. 3. Small selection of individual monitoring regions (physiographic regions) that were subjected to different processes of subdivision based on initial allocation of sites which were determined based on areal proportion of agricultural land-use coverage and HCAS index.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/17843657/sr25127.pdf by guest on 25 January 2026

initial site allocation. With agricultural land covering more than 85% of the region, this adjustment had little effect, and 92 sites were initially assigned. Clustering within the region produced nine monitoring units. Analysis of dominant soil and land-use combinations across these units identified 13 Genosoil and 53 Phenosoil monitoring sub-units, bringing the total site count to 185. The principal Australian Soil Classification orders represented were Vertosols (50%), Chromosols (35%), Sodosols (9%), and Dermosols (6%), with suborders dominated by Brown (94%) and Yellow (6%) variants. Land uses captured through the soil monitoring design were diverse, including dryland cropping (25%), grazing native vegetation (25%), grazing modified pastures (23%), irrigated cropping (17%), and irrigated pastures (11%).

The final example is Physiographic Region 20307 – Wimmera Plain, which covers just under 36,000 km^2, with approximately 82% of the land dedicated to agriculture. This region had a relatively low average HCAS value (0.15), which increased the site allocation from the original 33 to a revised total of 49. Clustering within the region produced four monitoring units. Analysis of dominant soil and land-use combinations identified eight Genosoil and 20 Phenosoil monitoring sub-units, giving a final site count of 80. The principal Australian Soil Classification orders represented were Sodosols (77%), Vertosols (13%), and Calcarosols (10%), with suborders dominated by Brown (73%) and Yellow (27%) variants. Land uses represented in the monitoring design included dryland cropping (44%), grazing modified pastures (37%), grazing native vegetation (14%), and plantation forestry (5%).

In total, the proposed NSMP consisted of 4775 sites, of which 3463 were designated as priority sites based on the established selection criteria defined earlier. Fig. 4 illustrates the distribution of these sites across the country. The non-priority sites were primarily located in the semi-arid and arid interior, where the dominant agricultural land use is grazing on unimproved pasture or vegetation.

An analysis of the 3463 priority sites shows that they are distributed across 343 monitoring units. Each of these includes a designated Genosoil monitoring sub-unit, along with an additional 846 Phenosoil monitoring sub-units, each representing a unique combination of soil and land-use types.

The distribution of sample numbers across different agricultural land-use types, and those situated in Genosoil reference areas, is summarised in Table 1. The present study utilised an Australian Soil Classification system that includes 14 mapped soil orders. However, this classification does not reflect the most recent update, which recognises 15 soil orders with the addition of the Arenosoil class (Isbell and National Committee on Soils and Terrain 2021). Of the 14

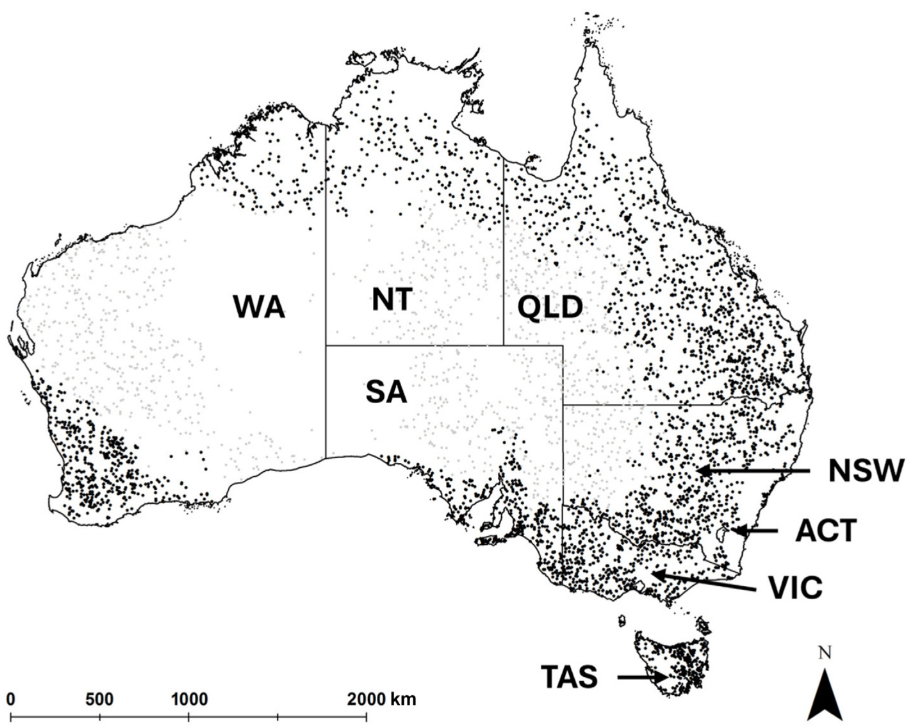

Fig. 4. Distribution of site locations that could constitute a national soil monitoring design. Black points represent the 3463 designated priority sites. Light grey points indicate non-priority sites which, together with the priority sites, comprise the full 4775-site configuration for the National Soil Monitoring Program. WA (Western Australia), NT (Northern Territory), SA (South Australia), QLD (Queensland), NSW (New South Wales), ACT (Australian Capital Territory), VIC (Victoria), TAS (Tasmania).

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

Table 1. Proportion of monitoring sites allocated to land-use types and Australian Soil Classification orders.

Category	Proportion of 3463-site network
Land class	
Grazing on unimproved/native vegetation	0.31
Production native forests	0.05
Grazing modified pastures	0.22
Plantation forests	0.04
Dryland cropping	0.16
Irrigated pastures	0.01
Irrigated cropping	0.01
Irrigated horticulture	0.01
Reference areas/Genosoils	0.19
Soil order	
Calcarosol	0.08
Chromosol	0.19
Dermosol	0.09
Ferrosol	0.02
Kandosol	0.14
Kurosol	0.02
Podosol	0.02
Rudosol	0.01
Sodosol	0.18
Tenosol	0.06
Vertosol	0.19

This table shows the proportion of the 3463-site monitoring network allocated to broad land-use categories and Australian Soil Classification orders. Land classes, except Reference areas/Genosoils, align with those defined in the land-use ALUM Classification ver. 8.

mapped soil orders, 11 are represented in the NSMP. The three excluded soil orders – Anthroposols, Organosols, and Hydrosols – are absent due to their limited spatial extent and low prevalence in agricultural land-use contexts. The breakdown of sites by Australian Soil Classification Order is detailed in Table 1.

Further breakdowns of site distribution include those categorised by state and territory government jurisdiction, as shown in Table 2, as well as sample densities (expressed as the number of sites per km^2) and sample sizes as a function of land-use type.

Integration of existing monitoring networks and other key soil datasets

As outlined in Section 2.5.7, the TERN Surveillance AusPlots network (Sparrow *et al.* 2020) was used as a representative case study to test the integration of existing monitoring sites into the NSMP design. AusPlots is a nationally recognised

Table 2. Summary of site density and sample size by land class and jurisdiction.

Land class	First quartile (km^2/site)	Median (km^2/site)	Third quartile (km^2/site)	Sample size
New South Wales and ACT				
Dryland cropping	12	20	31	142
Grazing modified pastures	14	20	31	175
Grazing on unimproved/native vegetation	12	26	39	211
Irrigated cropping	7	8	14	24
Irrigated pastures	6	7	9	6
Plantation forests	5	7	9	12
Production native forests	4	8	14	15
Reference Sites	11	15	21	124
Total (New South Wales and ACT)				709
Northern Territory				
Grazing on unimproved/native vegetation	24	33	39	117
Plantation forests	1	1	1	3
Reference sites	43	56	86	64
Total (Northern Territory)				184
Queensland				
Dryland cropping	7	10	22	72
Grazing modified pastures	12	32	52	219
Grazing on unimproved/native vegetation	25	39	61	425
Irrigated cropping	8	12	15	6
Plantation forests	4	4	4	3
Production native forests	7	11	16	63
Reference sites	9	16	30	210
Total (Queensland)				998
South Australia				
Dryland cropping	7	15	25	96
Grazing modified pastures	7	11	28	100
Grazing on unimproved/native vegetation	7	8	15	98
Irrigated horticulture	2	3	3	9
Plantation forests	5	6	16	12
Reference sites	11	17	40	66
Total (South Australia)				381
Tasmania				
Dryland cropping	6	8	11	15
Grazing modified pastures	11	14	23	44
Grazing on unimproved/native vegetation	10	14	23	40
Irrigated cropping	7	7	7	3
Irrigated pastures	6	6	9	9
Plantation forests	5	7	11	40

(*Continued on next page*)

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/17843665/sr25127.pdf by guest on 25 January 2026

Table 2. (*Continued*).

Land class	First quartile (km^2/site)	Median (km^2/site)	Third quartile (km^2/site)	Sample size
Production native forests	11	19	31	47
Reference sites	18	53	86	36
Total (Tasmania)				234
Victoria				
Dryland cropping	9	22	30	85
Grazing modified pastures	11	35	55	96
Grazing on unimproved/native vegetation	6	10	15	66
Irrigated cropping	6	6	6	3
Irrigated pastures	6	9	10	15
Plantation forests	7	9	11	39
Production native forests	8	14	15	33
Reference sites	9	28	49	74
Total (Victoria)				411
Western Australia				
Dryland cropping	15	31	49	149
Grazing modified pastures	8	11	20	125
Grazing on unimproved/native vegetation	7	14	32	133
Plantation forests	4	5	7	15
Production native forests	27	40	53	12
Reference sites	10	49	79	112
Total (Western Australia)				546

Values in the first three columns represent spatial density in km^2 per site; lower numbers indicate greater monitoring density. Sample size refers to the number of sites or observations within each land class. Land classes, except Reference sites, align with those defined in the ALUM Classification ver. 8.

ecological monitoring system with 946 sites, the majority of which are located in rangeland environments. Of these, approximately 357 intersect with priority sampling areas defined in the NSMP framework. Among the aligned sites, an estimated 54% have the potential to serve as reference (Genosoil) sites, while 43% fall on unimproved grazing land, and the remaining 3% are associated with other land-use categories. The spatial distribution of these intersecting AusPlots sites is shown in Fig. 5.

This exercise demonstrates the feasibility of incorporating existing monitoring networks into the NSMP without altering overall sampling density. More broadly, it shows that legacy datasets can be leveraged to reduce establishment costs, strengthen data reuse, and enhance long-term comparability of soil and ecological monitoring data. While AusPlots was used here as a test case, the same process can be applied iteratively to datasets maintained by state and territory agencies, universities, and other institutions as they become available.

Discussion

A key advantage of the design is the incorporation of both the 'global' and 'local' approaches, along with the ability to integrate existing soil monitoring networks. The overall structure is adaptable and can be adjusted based on an approximate total site number. For this design, an upper limit of 5000 sites was initially established. Through multiple iterations and refinements, including site selection based on dominant soil and land-use combinations and adjustments made for the 'local' work in Tasmania, the final total reached 4775 sites.

This design ensures coverage of the major soil types and agricultural land uses across the country. The density of monitoring sites is largely driven by environmental conditions, with higher site allocations in areas of intensive agricultural activity. Less intensive agricultural regions are also included in the NSMP, though with a lower site density.

The design is structured to be flexible, allowing discretion in both desktop and field settings when selecting site locations within defined sampling frames rather than being restricted to fixed locations. While having a predefined set of locations is useful for planning, these should not be regarded as rigid, must-visit sites but rather as indicative locations where monitoring can be established if field and logistical considerations are suitable.

This flexibility operates at two levels. The first, primarily desktop-based, is inherent in the design, as sampling frames generally contain thousands of potential candidate sites. Using a combination of random selection and geospatial assessment through the web-based application developed for this work, a well-distributed configuration of sites can be generated, even in remote areas. This systematic yet adaptable approach also facilitates the integration of existing soil monitoring networks without introducing unnecessary complexity.

The second level of flexibility, primarily field-based, involves the need for real-time site relocation during fieldwork, which is an expected challenge in any soil survey campaign. This is one reason why strict probability-based sampling designs are often unsuitable when there is uncertainty about the integrity of the sampling frame. In purposive sampling, alternative sites can be selected without significantly affecting the overall design parameters. Clifford *et al.* (2014) developed a flexible approach using cLHC sampling to accommodate such relocations. Similarly, Malone *et al.* (2019) implemented a method based on similarity indices, allowing alternative sites to be identified based on their resemblance to the originally planned locations.

Fig. 6 presents an example from Malone *et al.* (2019) illustrating a scenario for identifying alternative site locations. In practical terms, a site may meet the criteria of the sampling design but, for various reasons, cannot be established as a monitoring site. Using a distance- and similarity-based assessment approach, powered by available spatial data organised for

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

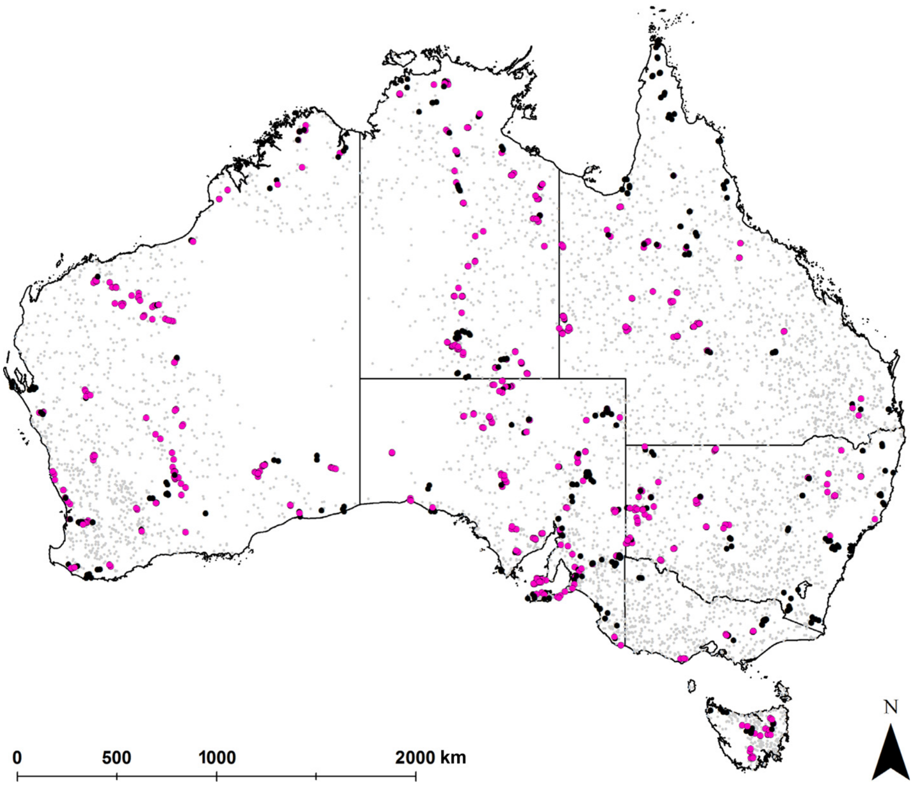

Fig. 5. TERN Surveillance sites overlaid on the 4775-site configuration of the National Soil Monitoring Program (NSMP). Pink coloured points indicate TERN sites that intersect with a delineated monitoring sub-unit, representing alignment with the NSMP sampling framework. Black coloured points represent TERN sites that fall outside of the defined monitoring sub-units.

each monitoring sub-unit, alternative locations with similar characteristics can be identified and proposed as replacement sites. A key advantage of this approach is that maps of alternative sites can be generated before fieldwork begins or even dynamically while in the field. The latter option would require a web-based application but would provide a data-driven method to support expert judgment and situational decision-making during site selection.

Limitations of the proposed design

The design has been structured with as much flexibility as possible to account for errors and uncertainties in the datasets used to create the monitoring network. Since the sampling frames for monitoring sub-units are based on mapped soil and land-use information, the reliability of these datasets plays a crucial role. We acknowledge that many of the inputs employed here – including SLGA products, HCAS, land-use mapping, and the newly derived Pedogenons – are model-based products that have not yet undergone systematic national-scale ground truthing. While their reported accuracies and general assessments provide some confidence, they inevitably contain artefacts and uncertainties. By focusing on dominant soil and land-use combinations, the design helps to mitigate this issue, as these are the most likely to occur in the vicinity of the designated sites. The Pedogenons used here were generated specifically for this work and are distinct from previously published versions, ensuring that they reflect the needs of the monitoring design.

Even though mapping is never entirely error-free, external model evaluations suggest that these products generally provide a reasonable representation of conditions on the ground. Where candidate sites prove anomalous relative to their expected soil or land-use classification, relocation to more representative positions is often feasible. Such decisions necessarily rely on expert judgement, underscoring the importance of training field teams to identify discrepancies and make appropriate adjustments.

Operational uncertainties must also be recognised. Sampling may be constrained by weather events, staff and equipment availability, and seasonal factors. Temporal clustering of site establishment is a potential risk, and while strict control of seasonal effects at national scale is not feasible, these can be managed through pragmatic scheduling (e.g. placing sites 'on hold' rather than abandoning them) and through subsequent modelling approaches to adjust for seasonal influences.

Finally, the proposed design does not in itself provide the statistical framework for quantifying soil change. Instead, it

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/17843665/sr25127.pdf by guest on 25 January 2026

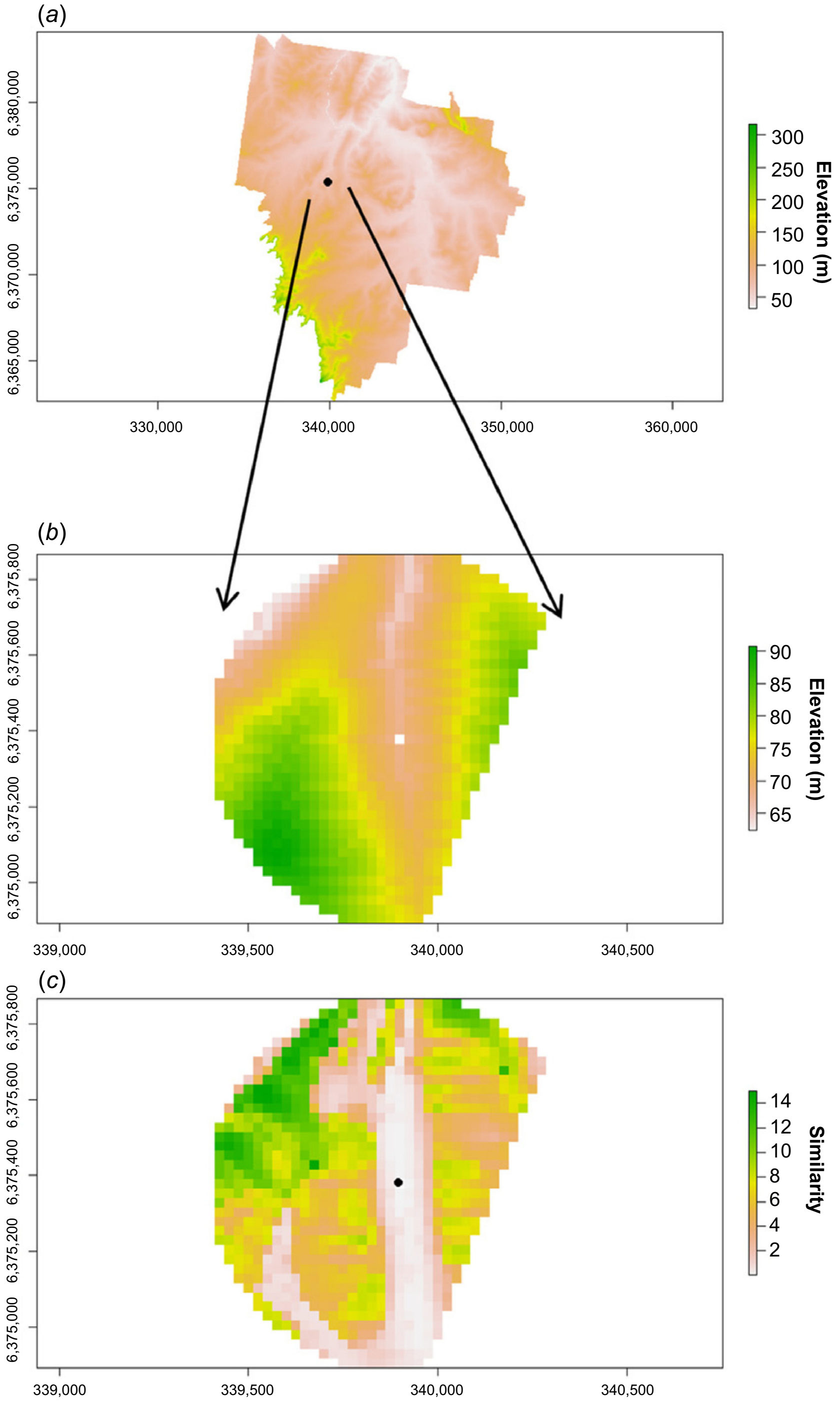

Fig. 6. Adapted from Malone *et al*. (2019), this figure illustrates the process for selecting alternative monitoring site locations when access to the originally designated site is not guaranteed. Panel (*a*) shows the location of the original sample site. Panel (*b*) displays a proximity buffer around the site, representing the permissible zone for selecting an alternative location. Panel (*c*) presents the taxonomic distance between the original site setting and its surrounding landscape, with lower values indicating greater similarity. This information can be used to identify nearby locations that preserve the intended design integrity and optimisation goals of the National Soil Monitoring Program. The key consideration in Panel (*c*) is the identification of areas with low taxonomic distance, which serve as viable substitutes with minimal impact on the sampling design.

establishes the sampling strata and structure upon which such analyses can be built. Given the sampling constraints and number of sites, the design is best regarded as a national sentinel system: it provides broad-scale coverage and early warning of abnormal or concerning changes in soil condition. These signals can then trigger more intensive, targeted investigations by government agencies or regional land management organisations. Importantly, the monitoring data generated through this framework will also serve as a valuable source of validation to improve existing predictive datasets over time, creating an iterative feedback loop between monitoring and modelling.

It should also be emphasised that this paper presents the general design framework rather than a detailed operational manual. Many of the finer-scale considerations – such as field protocols for plot establishment, procedures for handling temporal clustering, and strategies for managing seasonal variation – will be addressed in follow-up publications and

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

technical documents as the program transitions to its operational phase. The present focus is therefore on establishing a systematic and flexible national design, which can then be refined and implemented through subsequent technical guidance.

Conclusions

The design for a NSMP builds on previous proposals while incorporating recent advancements in soil assessment frameworks and the monitoring designs needed to support them. A key requirement for a NSMP is its ability to capture the full range of environmental variation and the diversity of soil and land-use combinations across the country. The proposed design provides a systematic approach for achieving this by leveraging high-quality and freely available digital soil and landscape data. Its hierarchical structure enables the identification and delineation of landscapes with similar environmental characteristics and soil properties. These monitoring units, or Pedogenons, can then be further subdivided based on soil and land-use combinations. One of these subdivisions consists of soils that remain in their natural state with minimal human impact. These areas serve as reference sites or Genosoils, providing baseline conditions against which soils from related monitoring sub-units with agricultural land use (Phenosoils) can be compared. Phenosoils include areas used for cropping or grazing on improved pastures. This hierarchical structure facilitates local assessments of soil change, as each Phenosoil can be compared to a corresponding local reference state.

A key strength of the design is its flexibility – particularly its ability to decouple from rigid sample size constraints. However, several levers must still be managed, requiring interactive decision-making to arrive at a suitable number of sampling sites within a defined budget. The described processes show how structured and adaptable approaches can be combined, as illustrated by the integration of localised monitoring in Tasmania within the broader national framework. Another notable feature is the capacity to incorporate existing soil monitoring networks, which enhances data reuse, reduces costs, and avoids duplicating efforts when establishing new sites. Additionally, data-driven frameworks support both desktop planning and field-based site selection, especially when candidate sites must be relocated. These tools facilitate an expert-guided process that offers structured guidance while allowing for expert-driven adjustments in the field.

References

ABARES (2022) Land use of Australia 2010–11 to 2015–16, 250 m. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, ACT.

Arrouays D, Marchant BP, Saby NPA, Meersmans J, Orton TG, Martin MP, Bellamy PH, Lark RM, Kibblewhite M (2018) Broad-scale soil monitoring schemes. In 'Pedometrics'. (Eds AB McBratney, B Minasny, U Stockmann) pp. 669–691. (Springer International Publishing: Cham, Switzerland)

Baldock JA, Grundy MJ, Griffin EA, Webb MJ, Wong MTF, Broos K (2010) Building a foundation for soil condition assessment. CSIRO Sustainable Agriculture Flagship Science Report.

Baldock J, Sanderman J, Macdonald L, Allen D, Cowie A, Dalal R, Davy M, Doyle R, Herrmann T, Murphy D, Robertson F (2013) Australian soil carbon research program. v2. CSIRO, Data Collection.

Brus DJ, de Gruijter JJ (2011) Design-based generalized least squares estimation of status and trend of soil properties from monitoring data. *Geoderma* **164**(3–4), 172–180. doi:10.1016/j.geoderma.2011.06.001

Chappell A, Baldock JA, Viscarra Rossel RA (2013) 'Sampling soil organic carbon to detect change over time.' (CSIRO: Australia)

Clifford D, Payne JE, Pringle MJ, Searle R, Butler N (2014) Pragmatic soil survey design using flexible Latin hypercube sampling. *Computers & Geosciences* **67**, 62–68. doi:10.1016/j.cageo.2014.03.005

DAFF (2023) National Soil Action Plan 2023 to 2028. Department of Agriculture, Fisheries and Forestry, Canberra, ACT.

DAWE (2021) National soil strategy. Department of Agriculture, Water and the Environment, Canberra, ACT.

de Gruijter JJ, Bierkens MFP, Brus DJ, Knotters M (2006) 'Sampling for natural resource monitoring.' (Springer) doi:10.1007/3-540-33161-1

Evangelista SJ, Field DJ, McBratney AB, Minasny B, Ng W, Padarian J, Román Dobarco M, Wadoux AMJ-C (2023) A proposal for the assessment of soil security: soil functions, soil services and threats to soil. *Soil Security* **10**, 100086. doi:10.1016/j.soisec.2023.100086

Grealish G, Clifford D, Wilson P, Ringrose-Voase A (2011) National soil condition monitoring program for soil ph and soil carbon: objectives, design, protocols, governance and reporting. CSIRO Publication, Canberra, ACT.

Harwood TD, Williams KJ, Lehmann EA, Ware C, Lyon P, Bakar S, Pinner L, Schmidt B, Mokany K, Van Niel TG, Richards AE, Dickson F, McVicar TR, Ferrier S (2021) 9 arcsecond gridded HCAS 2.1 (2001–2018) base model estimation of habitat condition for terrestrial biodiversity, 18-year trend and 2010–2015 epoch change for continental Australia. CSIRO, Canberra, Australia.

Isbell RF, National Committee on Soils and Terrain (2021) 'The Australian soil classification.' 3rd edn. (CSIRO Publishing: Melbourne, Australia)

Jang HJ, Dobarco MR, Minasny B, Campusano JP, McBratney A (2023) Assessing human impacts on soil organic carbon change in the Lower Namoi Valley, Australia. *Anthropocene* **43**, 100393. doi:10.1016/j.ancene.2023.100393

Jolivet C, Almeida Falcon J-L, Berch'e P, Boulonne L, Fontaine M, Gouny L, Lehmann S, Maitre B, Rati'e C, Schellenberger E, Soler-Dominguez N (2022) 'French soil quality monitoring network manual RMQS2: second metropolitan campaign 2016–2027, version 3.' (INRAE: Orleans, France)

Karunaratne S, Asanopoulos C, Jin H, Baldock J, Searle R, Macdonald B, Macdonald LM (2024) Estimating the attainable soil organic carbon deficit in the soil fine fraction to inform feasible storage targets and de-risk carbon farming decisions. *Soil Research* **62**(2), SR23096. doi:10.1071/SR23096

Maharjan B, Das S, Acharya BS (2020) Soil Health Gap: a concept to establish a benchmark for soil health management. *Global Ecology and Conservation* **23**, e01116. doi:10.1016/j.gecco.2020.e01116

Malone B (2022) Soil and landscape grid national soil attribute maps – soil colour (3″ resolution) – Release 1. v2. CSIRO, Data Collection.

Malone BP, Minansy B, Brungard C (2019) Some methods to improve the utility of conditioned Latin hypercube sampling. *PeerJ* **7**, e6451. doi:10.7717/peerj.6451

Malone BP, Searle R, Stenson M, McJannet D, Zund P, Román Dobarco M, Wadoux AMJ-C, Minasny B, McBratney A, Grundy M (2025) Update and expansion of the soil and landscape grid of Australia. *Geoderma* **455**, 117226. doi:10.1016/j.geoderma.2025.117226

McKenzie NJ, Henderson B, McDonald WS (2002) Monitoring soil change. Principles and practices for Australian conditions. CSIRO Land and Water Technical Report 18/02, Canberra, ACT.

Minasny B, McBratney AB (2006) A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences* **32**(9), 1378–1388. doi:10.1016/j.cageo.2005.12.009

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026

Nunes MR, Veum KS, Parker PA, Holan SH, Karlen DL, Amsili JP, van Es HM, Wills SA, Seybold CA, Moorman TB (2021) The soil health assessment protocol and evaluation applied to soil organic carbon. *Soil Science Society of America Journal* **85**(4), 1196–1213. doi:10.1002/saj2.20244
Orgiazzi A, Ballabio C, Panagos P, Jones A, Fernández-Ugalde O (2018) LUCAS Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil Science* **69**(1), 140–153. doi:10.1111/ejss.12499
Pain C, Gregory L, Wilson P, McKenzie N (2011) The physiographic regions of Australia – Explanatory notes 2011. Australian Collaborative Land Evaluation Program and National Committee on Soil and Terrain.
Román Dobarco M, McBratney A, Minasny B, Malone B (2021) A modelling framework for pedogenon mapping. *Geoderma* **393**, 115012. doi:10.1016/j.geoderma.2021.115012
Román Dobarco M, Wadoux AMJ-C, Malone B, Minasny B, McBratney AB, Searle R (2023*a*) Mapping soil organic carbon fractions for Australia, their stocks, and uncertainty. *Biogeosciences* **20**(8), 1559–1586. doi:10.5194/bg-20-1559-2023
Román Dobarco M, Padarian Campusano J, McBratney AB, Malone B, Minasny B (2023*b*) Genosoil and phenosoil mapping in continental Australia is essential for soil security. *Soil Security* **13**, 100108. doi:10.1016/j.soisec.2023.100108
Searle R (2020) TERN soil data federator. v1. CSIRO. Software Collection. Available at http://hdl.handle.net/102.100.100/480151?index=1
Searle R (2021) Soil and landscape grid national soil attribute maps – Australian soil classification map (3″ resolution) - Release 1. v3. CSIRO, Data Collection.
Searle R, Malone B, Wilford J, Austin J, Ware C, Webb M, Román Dobarco M, Van Niel T (2022) TERN digital soil mapping raster covariate stacks, v2. CSIRO, Data Collection.
Sparrow BD, Foulkes JN, Wardle GM, Leitch EJ, Caddy-Retalic S, van Leeuwen SJ, Tokmakoff A, Thurgate NY, Guerin GR, Lowe AJ (2020) A vegetation and soil survey method for surveillance monitoring of rangeland environments. *Frontiers in Ecology and Evolution* **8**, 157. doi:10.3389/fevo.2020.00157
Styc Q, Minasny B, Jang HJ, McBratney A (2025) Modelling soil organic carbon dynamics at the continental extent using pedogenon mapping. *European Journal of Soil Science* **76**, e70070. doi:10.1111/ejss.70070
van Leeuwen JP, Saby NPA, Jones A, Louwagie G, Micheli E, Rutgers M, Schulte RPO, Spiegel H, Toth G, Creamer RE (2017) Gap assessment in current soil monitoring networks across Europe for measuring soil functions. *Environmental Research Letters* **12**(12), 124007. doi:10.1088/1748-9326/aa9c5c
Wadoux AMJ-C, Román Dobarco M, Malone B, Minasny B, McBratney AB, Searle R (2023) Baseline high-resolution maps of organic carbon content in Australian soils. *Scientific Data* **10**(1), 181. doi:10.1038/s41597-023-02056-8
Wilson P, Malone B, Docherty C, Ringrose-Voase A, Karunaratne S, Gregory L, Clarkson T (2021) Designing a national soil monitoring program. A report to the Australian Government Department of Agriculture, Water and the Environment. CSIRO Report.

Data availability. The datasets analysed in this study are publicly available from the following sources: the *CSIRO Data Access Portal*, including the *Physiographic Regions of Australia* (https://doi.org/10.4225/08/579E72EA873CA) and the *Habitat Condition Assessment System* (*HCAS*) (https://doi.org/10.25919/a3h3-bs84); the *Soil and Landscape Grid of Australia* for digital soil and landscape information (https://esoil.io/TERNLandscapes/Public/Pages/SLGA/GetData-COGSDataStore.html); and the *Australian Bureau of Agricultural and Resource Economics and Sciences* (*ABARES*) for national land-use mapping (https://www.agriculture.gov.au/abares/aclump/land-use/land-use-of-australia-2010-11-to-2020-21).

Conflicts of interest. Brendan Malone is an Associate Editor for *Soil Research* but was not involved in the peer review or decision-making process for this paper. The authors declare no other conflicts of interest.

Declaration of funding. The National Soil Monitoring Program has been supported by funding through the Australian Government Natural Heritage Trust (Department of Agriculture, Fisheries and Forestry) in collaboration with CSIRO.

Acknowledgements. The National Soil Monitoring Program has been supported by funding through the Australian Government Natural Heritage Trust (Department of Agriculture, Fisheries and Forestry) in collaboration with CSIRO.

Author affiliations
[A]CSIRO Agriculture and Food, Black Mountain, ACT, Australia.
[B]CSIRO Agriculture and Food, Floreat, WA, Australia.
[C]CSIRO Agriculture and Food, St Lucia, Qld, Australia.

Downloaded from http://connectsci.au/sr/article-pdf/doi/10.1071/SR25127/1784365/sr25127.pdf by guest on 25 January 2026