Overcoming Key Challenges of Satellite-based Monitoring of Ecosystem Condition: A Continental-scale Example From Australia

Kristen J Williams[1], Simon Ferrier[1], Eric A Lehmann[2], Thomas D Harwood[1,3], Randall J Donohue[1], Kathryn M Giljohann[4], Roozbeh Valavi[4], Ning Liu[1], Chris Ware[5], Peter Lyon[6], Thomas G Van Niel[7], Tim R McVicar[1], Anna E Richards[8], Cassandra Malley[9]

1. Environment Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), GPO Box 1700, Canberra, Australian Capital Territory 2601, Australia

2. Data61 Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, Australian Capital Territory, Australia

3. Current address: Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom

4. Environment Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Melbourne, Victoria, Australia

4. Environment Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Hobart, Tasmania, Australia

6. Environment Information Strategy and Policy, Environment Information Australia, Australian Government Department of Climate Change, Energy, the Environment and Water (DCCEEW), Canberra, Australian Capital Territory, Australia

7. Environment Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Perth, Western Australia, Australia

8. Environment Research Unit, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Darwin, Northern Territory, Australia

9. Environment Data and Analysis Branch, Environment Information Australia, Australian Government Department of Climate Change, Energy, the Environment and Water (DCCEEW), Canberra, Australian Capital Territory, Australia

Corresponding author: kristen.williams@csiro.au

Abstract

Effective satellite-based monitoring of ecosystem integrity or condition needs to address four key challenges: (a) context dependency; (b) alternative ecological states; (c) short-term temporal ecosystem dynamics; and (d) scarcity of reference data where ecosystems retain high levels of integrity. Here we present a typology, and outline strengths and weaknesses, of different approaches to mapping and monitoring ecosystem integrity across entire regions or continents using time series satellite data. We then describe how one of these approaches, the Habitat Condition Assessment System (HCAS), addresses all of the above challenges, and provide an outline of the evolved method which includes annual outputs, and Australian continent applications. HCAS requires three readily available inputs (i.e., representative examples of relatively natural areas as reference sites, remotely sensed ecosystem characteristics, and environmental covariate data) and could be easily adapted and applied by other countries to provide an effective indicator of ecosystem integrity for nature-based decisions.

Key words

Ecosystem integrity, remote sensing, change detection, ecosystem accounting, biodiversity persistence

Globally, less than 25% of the world's terrestrial ecosystems are estimated to remain relatively free of direct anthropogenic disturbance and nearly 60% are under moderate or intense human pressure (Williams B. A. et al. 2020a). The cumulative and interacting effects of chronic and acute pressures are driving widespread ecosystem collapse (Bergstrom et al. 2021). In turn, consumers are using market behaviours to communicate concerns about risks to the environment (Bradshaw et al. 2021, White K. et al. 2019), businesses and financiers are evaluating their exposure to nature-negative outcomes (TNFD 2023), and shareholders are demanding corporate balance sheets include accounting for natural capital (Barker 2019, Unerman et al. 2018). In response, the United Nations has formulated an approach to ecosystem accounting to make visible the contributions of nature to the economy and people (United Nations et al. 2021), incorporating concepts of ecosystem condition and integrity (Roche and Campagne 2017), and countries have agreed to report against sustainable development goals related to life on land and to combat desertification (e.g., Sims et al. 2020, Sims et al. 2019) and to achieve the global vision of a world living in harmony with nature (CBD 2022a, United Nations 1992). Thus, due to systemic over-utilisation, the conservation of nature and natural resources has become a mainstream concern (Scott et al. 2022), requiring that governments and corporates institute monitoring for social license to operate (Brand et al. 2023).

To support the accelerating need for regular information about the status of ecosystems, pressures, drivers and impacts, scientists and regulators are collaborating to develop monitoring systems and reporting tools for tracking change (e.g., Tallis et al. 2012, Timmermans and Daniel Kissling 2023). A wide range of environmental indicators have been proposed for reporting on status and trends within different frameworks, with ecosystem condition and its synonyms such as integrity (e.g. Supplemental Material A Table S1) and antonyms such as pressures (via hemeroby) are a common thread (e.g., Cowie et al. 2018, UN 2015, UNCCD 2016, United Nations 1992, United Nations Forum on Forests Secretariat 2019). Under the Kunming-Montreal Global Biodiversity Framework (CBD 2022a), for example, Goal A seeks to substantially increase the extent of all natural ecosystems by 2050 by maintaining, enhancing or restoring their integrity, connectivity and resilience; thereby reducing the risk of further collapse (Nicholson et al. 2021). The extent and integrity (i.e. condition) of ecosystems sustained within a country has direct consequences for regional persistence of native species and genetic diversity, as well as human wellbeing (e.g., Ulrich et al. 2023).

It is not surprising therefore, that we see a renewed focus on methods for measuring ecosystem condition (here, synonymous with "integrity" for nature conservation, sensu Roche and Campagne 2017) as a leading indicator of the risk of ecosystem collapse and potential for species extinction (Hansen et al. 2021, Stevenson et al. 2021). There is also growing interest in the role of satellite remote sensing in ecosystem integrity assessment (Harwood et al. 2016, Lawley et al. 2016, Murray N. J. et al. 2018, Tehrany et al. 2017). It is therefore timely to reflect on the challenges of using remote sensing to monitor ecosystem condition (i.e. integrity) as an input to assessments of biodiversity outcomes (Ferrier et al. 2020). Because multiple terms with similar meaning and purpose are in common use (Supplemental Material A Table S1), we consider methods that aim to address the quality, integrity, condition, health, capacity, intactness or naturalness of ecosystems, as conceptually consistent for monitoring the quality of habitat for native species persistence.

Herein, we outline those challenges, present a typology of how well existing approaches address them, and detail recent advances in developing one of these approaches – the 'Habitat Condition Assessment System', since the framework for this approach was first published (Harwood et al. 2016). Our concept of ecosystem condition follows that outlined in the United Nations System of Environmental-Economic Accounting—Ecosystem Accounting (United Nations et al. 2021), wherein ecosystem condition is measured relative to areas of an ecosystem type considered to be in reference condition, with condition defined as *the system's capacity to maintain composition, structure, autonomous functioning and self-organisation over time using processes and elements characteristic for its ecoregion and within a natural range of variability* (Keith et al. 2020).

Challenges of using remote sensing to monitor ecosystem condition

The main ecological application challenges inherent to use of satellite remote sensing for estimating ecosystem condition were outlined by Harwood et al. (2016). These challenges all relate to the over-arching ambition to separate effects of exogenous, anthropogenic disturbances that modify ecosystems from endogenous regimes inherent to an environment that species and ecosystems have adapted to over evolutionary time scales—in order to correctly measure the ecological similarity of the current state of an ecosystem to its reference state (high levels of ecosystem integrity). Here we rephrase and extend those challenges, namely: (a) context dependency, (b) alternative ecological states, (c) short-term temporal ecosystem dynamics, and (d) scarcity of reference data where ecosystems retain high levels of integrity (broadly illustrated in Figure 1).

Context dependency. The ecological context of an ecosystem is its environment and historical legacies of disturbance (both natural and anthropogenic) that have shaped its characteristic structure, function and composition. The challenge of context dependency relates to making the correct ecological interpretation of remotely sensed land cover data. Different locations in substantially different environments can exhibit the same set of remotely sensed ecosystem characteristics, when viewed from satellites, but may actually have very different levels of ecosystem condition from an integrity perspective in nature conservation. This is because spatial variation in environmental factors shaping distribution of natural ecosystem types, and variability within types, can be conflated with anthropogenic processes that modify ecosystems. For example, from a remote sensing perspective, characteristics of an intact (natural) open grassy woodland might appear identical to a former closed forest which has long since been modified to promote grass growth and continues to be managed for livestock grazing (Figure 1). A correct interpretation requires additional information about the reference state of the ecosystem prior to industrial era anthropogenic influences. If not addressed, the ensuing error of detection is one of measuring differences in ecosystem characteristics from the wrong reference point or baseline, especially where anthropogenic influences altered those characteristics prior to acquisition of remote sensing imagery (Harwood et al. 2016).

Alternative ecological states. More than one type of ecosystem or biome can occur naturally at any given location depending on long-term endogenous disturbance regimes (over multiple decades to hundreds of years) such as fire or large vertebrate herbivory (Pausas and Bond 2020). For example, savanna and forest distribution in many parts of the world, including Africa and Australia, depends upon maintenance or removal of specific disturbance regimes that advantage one or other ecological state, which recognised as distinct types yet occupying the same environment (i.e. similar combinations of soil, climate, landform and hydrology) (Murphy and Bowman 2012, Staver et al. 2011). When viewed from satellites, these juxtaposed alternative ecological states will look quite different, but have the same ecosystem reference condition (Figure 1). A long-term view of remote sensing data (e.g., multi-decadal) is needed to clearly distinguish endogenous disturbance and recovery dynamics from anthropogenic influences. If not addressed, the ensuing error is again one of measuring ecosystem condition from the wrong reference point or baseline (Harwood et al. 2016).

Short-term temporal ecosystem dynamics. Within a given ecosystem type, different forms may persist for short periods (months to years) at a given location, due to periodic events such as fire and rainfall followed by biomass recovery, which may or may not be seasonal. For example, natural phenomenon of low leaf cover or bare ground can occur annually, such as deciduousness or dieback during regular dry or cold periods (Moore et al. 2016), or over several years (e.g., El Niño–Southern Oscillation - Wang et al. 1999). Such short-term drivers within the natural range of variability, do not affect the ecosystem's condition but do alter its appearance when viewed from satellites. A long time series is needed to rule out change in ecosystem characteristics being due to an exogenous driver, and not part of a natural short- or medium-term disturbance-response dynamic (Burton et al. 2020, Harwood et al. 2016).

Scarcity of reference data where ecosystems retain high levels of integrity. Ecosystem condition is typically estimated relative to reference levels of specified ecosystem characteristics (Czúcz et al. 2021). In transformed landscapes, high ecosystem integrity reference sites can be scarce or non-existent. Inadequate or poorly specified reference sites can result in shifting baselines when anthropogenically modified ecosystems are substituted for missing data, if the condition assessment methodology does not adjust for this scarcity (Harwood et al. 2016). Reference sites used in condition assessment need to represent, as far as possible, spatial-temporal variation in ecosystem characteristics and related environmental gradients. There are many different ways to conceptualise and establish the prevailing natural reference state for ecosystem condition assessment (e.g., Jakobsson et al. 2020, Keith et al. 2020, McNellie et al. 2020). Hansen et al. (2021), for example, related these choices to the degree of ecological representation of the natural reference state versus feasibility of data collection (see Figure 3 therein).

These four interacting challenges, inherent to the use of satellite remote sensing for monitoring ecosystem condition, are a worldwide problem. Approaches to addressing any one of these can have implications for one or more of the other challenges, and so they need to be addressed collectively (Harwood et al. 2016).

Strengths and weaknesses of different analytical strategies

A wide range of approaches have been developed to solve the multiple challenges of monitoring and mapping ecosystem condition across large spatial extents using satellite-based remotely sensed ecosystem characteristics. These strategies may use remote sensing directly or indirectly (e.g., via land use mapping). We arranged these different approaches into a high-level typology, with an emphasis on use of satellite data (Figure 2). Each approach has particular strengths and weaknesses with respect to the challenges introduced in the previous section, which we outline below with examples. The effectiveness of any given approach, however, depends on the precise focus of the application of interest, and the nature, quality and quantity of available data streams.

Condition inference from mapped pressures

The first major split in this typology is between approaches which infer or predict ecosystem condition indirectly from mapping of pressures (i.e. hemeroby), as opposed to approaches which estimate condition more directly from remote sensing. Deriving spatially-complete mapping of ecosystem condition—by overlaying best-available mapping of pressure indicators such as land use or tenure, human-population density, distance to roads or urban centres, etc—offers a means of estimating condition rapidly, and at low cost, across very large spatial extents, including globally (Purvis et al. 2018, Venter et al. 2016, Williams B. A. et al. 2020a). Another notable benefit of this approach is that it potentially allows consideration of the impact of pressures which are not readily detected directly from remote sensing, such as predicting relative impact of feral predators as a function of distance from roads and human settlements (Andrews 1990, Doherty et al. 2015, Forman and Alexander 1998, Schneider 2001). However, a potential weakness is inability to distinguish between situations in which a given level of pressure has already resulted in loss of ecosystem condition versus situations in which impacts from that pressure are yet to be realised (e.g., an area of intact habitat which is vulnerable to future transformation given its tenure, proximity to roads, etc). This means that any use of this approach for monitoring change in site condition over time will only reflect changes in pressures, not the realisation of those pressures in terms of actual impacts on condition. Another weakness is that the approach often uses pressure data drawn from multiple sources of varying temporal currency which reduces accuracy of change detection. Given the respective, largely complementary, strengths and weakness of pressure-based-versus-direct-remote-sensing approaches to ecosystem condition mapping, these two broad approaches are sometimes applied in combination (e.g., Grantham et al. 2020, Hansen et al. 2021, Love et al. 2020).

Estimation of condition as a direct function of remote-sensing variables

Among methods that estimate condition directly from remote sensing are those that estimate condition as a direct function of one or more remote sensing-derived variables. These methods assume that higher tree canopy density, above-ground biomass or productivity such as derived from normalized difference vegetation index (NDVI) and its modifications, for example, equates to higher condition. Naicker et al. (2024) developed modified NDVI-type indices to assist rangeland managers assessing change in vegetation condition in high density grasslands. Zelený et al. (2021) range-normalised three remote sensing variables—NDVI, at-satellite brightness temperature (BT) and vegetation surface heterogeneity (HG) derived from Sentinel-2 and Landsat 8 sensors—to rank land classes along an ecosystem integrity gradient; although even the best performing land class could still have low ecosystem integrity. Huang et al. (2020) used top 10% above-ground biomass obtained through remote sensing as a benchmark of naturalness for assessing ecosystem asset quality. Such applications are necessarily restricted to particular regions or ecosystems where assumptions remain valid. They are not transferable or generalisable to entire continents or countries because they do not inherently address challenges of context dependency or alternative ecological states.

Correlative modelling of condition from remote-sensing data

Efforts over the past 25 years to estimate, and thereby map, ecosystem condition more directly from remote sensing have mostly pursued one of two main analytical paradigms. The first of these involves treating the problem as a relatively standard correlative-modelling challenge, solvable through application of standard statistical-regression or machine-learning tools. Data on sample locations known to exhibit different levels of ecosystem condition (i.e., the response, or dependent, variable) are used to train a correlative model capable of predicting condition as a function of predictor (independent) variables derived from remote sensing (along with any other relevant environmental covariates), thereby allowing condition to be mapped predictively across the entire region of interest. Two variants of this approach are worth distinguishing. The most widely applied of these employs training data generated through site-based condition assessment activities—that is, the training data consist of a set of field sites at which condition has been assessed through ground-based observation (Newell et al. 2006, Zerger et al. 2009).

The other variant of this correlative-modelling approach employs spatially complete training data across the region of interest. These data typically consist of best-available mapping of ecosystem condition derived through some form of the pressure-based approach described above. The logic of this variant is that, by building a correlative model relating best-available mapping of ecosystem condition to remote sensing predictor variables, this model can then be used to estimate changes in condition directly from remote sensing (Hoskins et al. 2016, Keys et al. 2021). If the spatial resolution of the condition layer used as training data is much coarser than the resolution of the remote-sensing data used to fit, and make predictions from the model, then this process is essentially one of statistical downscaling.

A major advantage of mapping condition through standard regression or machine-learning techniques is that these techniques are both readily available and well proven through their extensive application across other domains. However, their application to mapping ecosystem condition from remote sensing needs to overcome a significant challenge rarely shared by other applications. This is the challenge of context dependency, and of alternative ecological states because of the deterministic nature of predictions. Unlike correlative modelling of more straightforward ecosystem characteristics such as tree height, biomass or primary productivity, different locations exhibiting precisely the same set of measurement values for a set of remotely sensed ecosystem characteristics can actually be at very different levels of condition depending on their ecological context; and vice versa with regard to ecological states. Strategies for addressing these challenges—for example, by fitting separate models for different natural ecosystem types or biomes, or by including contextual environmental variables as interacting covariates within fitted models—usually require sizeable amounts of training data to perform effectively, especially when this approach is applied across large spatial extents encompassing a wide range of ecosystem types and their alternative states (e.g., Spatial BioCondition - DES 2021).

Estimation of condition from difference between observed and expected vegetation characteristics

The other main analytical paradigm for estimating ecosystem condition from satellite remote sensing approaches the problem from a different angle. Rather than attempting to model, and thereby predict, condition directly as a correlative function of remote sensing variables, this paradigm instead focuses on predicting what the ecosystem at any given location would be ‘expected’ to look like (from a remote-sensing perspective) as if it had persisted with high levels of integrity. Then the actual condition of the ecosystem at that location and given time point is estimated as a function of the deviation in observed remotely sensed characteristics from this expectation. This paradigm is designed, from the outset, to deal directly with the challenge of context dependency outlined above. Three variants of this general approach have emerged over recent years.

The first of these involves comparing the spatial distribution of discrete land-cover classes mapped from remote sensing with the distribution of classes expected if the entire landscape were in reference condition (highest attainable integrity). Mapping of expected reference condition land-cover classes is essentially equivalent to mapping ‘potential natural vegetation’ classes which has been undertaken using a wide range of correlative, mechanistic, and expert-based modelling techniques (Bonannella et al. 2023, Hengl et al. 2018). The deviation between observed and expected (reference condition) land-cover classes can be assessed with varying degrees of rigour, ranging from binary match/mismatch analysis through to relatively sophisticated consideration of relationships between multiple ecosystem states within a state-and-transition modelling framework (Blankenship et al. 2015, Daniel et al. 2016, Richards et al. 2021). The use of discrete classes in this first variant confers a clear advantage in eliciting expert knowledge and communicating with policy and management practitioners. However, this same feature also brings with it some potential disadvantages, including the risk that any error in mapping the expected distribution of natural vegetation classes, or ecosystem types, can result in a spurious divergence between observed and expected remote-sensing characteristics, and therefore error in the estimation of condition.

The second variant of the observed-versus-expected paradigm works with continuous habitat or ecosystem variables rather than discrete land cover classes. Ecosystem condition can potentially be estimated by comparing remotely sensed spatial variation in ecosystem characteristics with that expected if the entire landscape were in reference condition, based on correlative or mechanistic modelling. Amongst the surprisingly few applications of this approach to date, the most prominent are those comparing observed and expected (mechanistically modelled) levels of Net Primary Production, including at global scale (Haberl et al. 2014). As for the comparison of observed and expected land-cover classes (above), any error in modelling the expected distribution of continuous ecosystem variables can again result in a spurious divergence between the observed and expected distribution of these variables. The highly deterministic nature of this approach may also be challenged by the existence of multiple valid alternative ecological states occurring at a given location within abiotic environmental space; for example, mosaics of woodland/open-forest and closed forest (rainforest) vegetation shaped largely by relatively stochastic patterns in the distribution of past fire events.

A related approach works with expected values from particular remotely sensed ecosystem characteristics of interest (e.g., persistent ground cover) as reference pixels within a neighbourhood around each pixel of interest (Bastin et al. 2012, Donohue et al. 2022, Pickup et al. 1994). This variant of “difference based on continuous remote sensing-derived habitat or ecosystem variables” (Figure 2) uses proximity in geographical space, sometimes also stratified by soil type, rather than directly using environmental space (or discrete vegetation types), to deal with context dependency. It was developed for rangeland environments where the local environment changes relatively slowly across geographical space to separate the effects of grazing from rainfall in assessing rangeland pasture condition to inform management practice sustainability. This approach is context limited and will not be applicable in environmentally heterogenous landscapes with steep climatic, soil and topographic gradients.

A more extensively applied approach to estimating condition based on mapping of continuous habitat or ecosystem characteristics involves comparing the value for each (at any given location) with that expected for the reference ecosystem type, or vegetation community, concerned. This comparison is typically based on knowing reference values for the relevant characteristics, obtained through field-based ecological survey at reference (benchmark) sites within each vegetation type (Cohen et al. 2001, DES 2021, Kocev et al. 2009, McNellie et al. 2021). While offering one of the most rigorous approaches to mapping ecosystem condition from remote sensing, this approach requires reliable mapping of natural ecosystem types (or vegetation communities), the existence of field-based survey data for benchmarking ecosystem characteristics within each type, and the ability to use remote sensing to accurately map the same ecosystem characteristics as those assessed in the field (including sub-canopy and ground-layer features which may be difficult to detect from satellite imagery). These methods may also not cope well if there’s significant natural variation in the remotely sensed ecosystem characteristics (due to local heterogeneity or multiple ecological states) within each class of ecosystem type.

The third and final variant of the observed-versus-expected paradigm, which includes the Habitat Condition Assessment System (HCAS) approach, makes no attempt to map land-cover classes, or to estimate ecosystem characteristics that precisely match those employed in field-based condition assessments. It instead works with best-available remotely sensed variables characterising continuous variation in overall ecosystem structure, function and composition (within contemporary technology limits of remote sensing). These remotely sensed ecosystem characteristics are viewed as forming a multidimensional space, within which comparisons between the observed state and expected reference state are made in terms of the ecological differences (i.e. distances) between any two points within that space. In this approach, ecosystems are characterised within a continuum rather than predefined discrete classes. While the analytical approach employed in HCAS is described in detail elsewhere (Harwood et al. 2016, Lehmann et al. 2021, Williams K. J. et al. 2023a, Williams K. J. et al. 2023c, Williams K. J. et al. 2021b, Williams K. J. et al. 2020b), it is worth briefly contrasting this particular approach with other potential options for using the analysis of ecological differences (or distances) within a multidimensional remote-sensing space to map ecosystem condition across large spatial extents.

The first of these options would involve working with a discrete set of natural ecosystem types, or vegetation classes. If a set of locations known to be in reference condition within each of these types can be identified, then the condition of all other locations (pixels) within that same type might be readily estimated from the difference between the ecological position of any given pixel in this multidimensional space and the positions of the reference sites. This option would, however, require reliable mapping of the distribution of natural ecosystem types, and ready availability of a representative set of reference locations within each of those types. Given that species composition varies continuously in space or time (Austin 2013, McGeoch et al. 2019), this approach risks quantifying natural within-class variation as variation in condition (i.e. the challenge of multiple ecological states). The HCAS approach relaxes this requirement by modelling the expected remote-sensing differences in ecosystem characteristics between locations as a continuous function of abiotic environmental differences between those locations rather than as a binary function of ecosystem-type membership.

Another variation on the HCAS approach might have been to directly compare the observed position of a pixel in multidimensional remote-sensing space with the expected position of that same pixel, as if it were in reference condition. However, this would not have allowed for the existence of alternative ecological states at any given position in abiotic environmental space. Hence the approach adopted in HCAS makes comparisons among the observed and expected differences in multi-dimensional remotely sensed ecosystem characteristic space between the position of a site (pixel) of interest and the positions of all relevant reference sites.

Benefits of the Habitat Condition Assessment System (HCAS) approach

Different approaches to using remote sensing to monitor ecosystem condition typically address one or other of the four challenges illustrated in Figure 1. HCAS was specifically configured to address all four challenges (Harwood et al. 2016). We are not aware of any other approach that has been designed to do this.

A schematic of HCAS 'mechanics' in Figure 3, summarises how ecosystem condition is assessed for the simplest case of one test site and one reference site. The test site is a location for which condition needs to be estimated from remotely sensed ecosystem characteristics data alone. The reference site is a location in a relatively similar abiotic environment to the test site, and is known to be in reference condition (i.e. has high ecosystem integrity from a nature conservation perspective). The challenge of context dependency is addressed using a model predicting the multivariate remote-sensing distance (or 'difference') expected between any two sites if both these sites are in reference condition. This remote-sensing difference is predicted as a function of the abiotic environmental characteristics (e.g., climate, soil, landform, hydrology) of the two sites concerned. The ecological difference, based on Manhattan distances, between the test site and reference site is derived for both the observed (d_{obs}) and predicted (d_{pred}) sets of remotely sensed ecosystem characteristics (Figure 3). Observed remotely sensed characteristics are selected (insofar as possible) to represent the structure, function and compositional features of ecosystems, for which their inter- and intra-annual variability has been summarised over a specified period (at least 10 years). These summaries aim to address spatial-temporal challenges of short- to medium-term seasonal dynamics.

The two dimensions of these multivariate ecological differences are plotted (observed and predicted reference states). The y-axis is the observed difference in remotely sensed ecosystem characteristics between the test site and a particular reference site (d_{obs}). The x-axis is the difference in remotely sensed characteristics expected between these two sites if both were in reference condition (d_{pred}), predicted as a function of their environmental difference. If a test site has high levels of ecosystem integrity then observed differences between this site and each reference site will, on average, fall close to the 1:1 line, approaching a maximum condition score of 1.0. However, if a test site has low levels of ecosystem integrity then observed differences between this sites and each reference site will, on average, fall further way from the 1:1 line, approaching a minimum condition score of 0.0 (Figure 3).Points that fall further along the predicted (d_{pred}) axis represent comparisons with reference sites that are increasingly dissimilar in predicted remotely sensed ecosystem characteristics from the test site of interest (i.e., could be classed as entirely different ecosystem types), and these therefore play less of a role, and/or carry less weight, in assessing condition than those closer to the origin. These more dissimilar ecosystems, however, may have a role in assessing condition when similar ecosystem reference sites to the test site no longer exist, thereby addressing the challenge of scarce reference sites (discussed below).

The challenge of accommodating alternative ecological states is addressed by further weighting the influence of reference sites which are not only similar in predicted ecosystem characteristics to the test site, but are also most similar in terms of their observed remotely sensed characteristics. Emphasising reference sites that are most similar to the test site in terms of both their predicted and observed remotely sensed ecosystem characteristics also helps address the challenge of seasonal dynamics (along with the longer time-series of remote sensing data), because observed remotely sensed characteristics of reference and test sites are expected to change in similar ways, especially if these are also selected to be in close geographical proximity to one another.

The challenge of scarcity of reference sites is addressed through the scatter plot of observed and predicted differences in remotely sensed ecosystem characteristics (Figure 3) which does not require every possible test site to have a reference site in an identical environment—as scaled by parameters of the reference ecosystem model which simulates the continuum in ecosystem types (a prior step in the HCAS workflow, Figure 4). The difference in remotely sensed ecosystem characteristics between a test and reference site is considered a function not only of condition, but is also an ecological legacy of its physical environment. Accounting for differences in environment (i.e., the challenge of context dependency) ensures condition is assessed using reference sites from environments most similar to the test site, while simultaneously adjusting for the effect that any deviation from exact environmental similarity is expected to have on the difference in predicted remotely sensed ecosystem characteristics observed between test and reference sites. A certain density of reference sites is needed to account for both short- and long-term dynamics of ecosystems expressed through alternative natural ecological states due to disturbance regimes such as fire, drought and flood. Identification and selection of reference sites that are representative, insofar as possible, of both environmental and remote sensing variability is therefore crucial to addressing the challenge of reference site scarcity.

Implementation of an enhanced HCAS across the Australian continent

Since the original conceptual framework and proof of concept by Harwood et al. (2016), the implementation of HCAS to the Australian continent has significantly advanced though successive updates, as summarised in Supplemental Material B Table S2. The method requires three types of input data: reference sites (the most intact examples of contemporary natural ecosystems), environmental covariates (relatively stable physical drivers of ecosystem distribution and diversity) and remote sensing variables (characterising as far as possible the structure, function and composition of ecosystems). Major steps in the workflow (Figure 4) include:

1) a model of the remotely sensed reference ecosystem characteristics, using a representative training sample of reference sites (inferred to have high levels of ecosystem integrity such that condition is approaching 1.0), to predict those characteristics across all locations of interest based on long-term stable environmental covariates; and
2) a benchmarking stage, involving:
    a) a process for selecting several reference sites that are most like the reference ecosystem characteristics of the test site of interest for estimating its condition; and

b) an algorithm for estimating proximity of each test site to the selected reference sites based on differences in their remotely sensed ecosystem characteristics.

A detailed schematic representation of the workflow is provided in Supplemental Material Figure S2. The status of these workflow components is summarised below, illustrated using results from the published HCAS version 2.3 (Harwood et al. 2023a, Williams K. J. et al. 2023c) applied at 250m grid resolution (see Supplemental Material C for method details, and Supplemental Material B Table S2 for a brief technical summary).

Reference sites

Reference sites are contemporary locations where we expect to find the most intact examples of natural ecosystems and their variants. HCAS assumes reference sites retain their status for the duration of the remote sensing period over which ecosystem condition is assessed. Reference sites serve two primary purposes: 1) as training data in the reference ecosystem model for predicting the continuous expectation of remotely sensed reference ecosystem characteristics across all sites, as used in the two-way scatterplot shown in Figure 3, and 2) as benchmarks for estimating the condition of test sites based on each test sites' proximity to a dynamic reference state (expressed through multiple reference sites and their remotely sensed ecosystem characteristics). Reference sites used as training data need to represent, as far as possible, the potential diversity of ecosystems; whereas reference sites used as benchmarks need to also represent contemporary alternative ecological states and seasonal dynamics.

Logical inference is the primary way reference sites are derived for use in HCAS, supplemented by expert knowledge and field observations (Supplemental Material C: *Reference Sites*). Multiple lines of evidence from existing mapped data are used to infer the location of reference sites applying methods similar to hemeroby but focussed only on identifying locations most likely to have retained high levels of ecosystem integrity. Spatially inferred reference sites therefore largely occur within protected areas or relatively natural areas where no significant land use prevails. They make up approximately 35% of the Australian continent (Figure 5). HCAS assumes inferred reference sites are accurate, especially those used as benchmarks. An index of native species proportions derived from Mokany et al. (2022b) provided supporting evidence for the multiple lines of evidence approach.

Inferred reference sites were representatively subsampled using a detailed ecological land classification derived from fine-scale mapping of ecological regions (Department of the Environment 2014) and native vegetation (DAWE 2020), resulting in more than 5000 units for continental Australia (Supplemental Material C: *Sub-sampling reference sites*). A stratified random sample of approximately 100,000 sites are used as training data and 200,000 sites as benchmarks, although numbers can vary in different HCAS versions (Supplemental Material B Table S2) depending on what is both computationally tractable and provides comprehensive coverage of ecosystem diversity. Optimal sampling methods for training and benchmark data are being explored as a future refinement.

Environmental covariates

We use the existence of a relationship between the reference state of remotely sensed ecosystem characteristics and environmental covariates to develop a predictive capacity – the HCAS reference ecosystem model (Figure 4). This statistical model describes the correlative relationship observed between a set of remote sensing (response) variables and a set of environmental (predictor) covariates for the training subsample of reference sites. The fitted model is then used to predict the reference state of the remotely sensed ecosystem characteristics based on the environmental covariates, for every location across continental Australia. For this purpose, the environmental covariates need to characterise the equilibrium reference states of the environment to which natural ecosystems have become adapted and diversified, over ecological and evolutionary time frames. Suitable data are compiled from multiple sources (Supplemental Material C – *Environmental covariates*). Exploratory data analyses then help to identify and reduce multicollinearity to derive a candidate set for use as potential predictors (Supplemental Material C Table S3).

Remotely sensed ecosystem characteristics

A multi-decadal assessment period is required of the remote sensing variables to distinguish natural ecosystem processes of within- and between-year seasonal dynamics from variability due to anthropogenic influences that cause a departure from this predictable behaviour; in order to avoid errors of interpretation. Satellite-based remote sensing is used as the ecosystem observatory, and therefore the choice of variables aims to encompass, as comprehensively as possible, the characteristic variability in structure, function and composition of all ecosystems. In practice, not all field observable features relevant to condition assessment can be detected from satellites, but this will improve over time with advances in sensor and satellite technology (Murray Cameron et al. 2022, Pettorelli et al. 2017). The accuracy of condition assessment is therefore necessarily limited to remote sensing data that meet minimum requirements of a multi-decadal time series and seasonal completeness—having relatively high frequency imagery to reduce missing data due to cloud and smoke. For this reason, data derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite was originally selected for use with HCAS, thereby restricting the output resolution to 250 m^2 (see rationale in Williams K. J. et al. 2021b). Alternative approaches based on long time-series Landsat data (Wulder et al. 2022) are being explored as a future refinement.

Remote sensing variables derive from four MODIS Collection 6 fractional cover products using satellite imagery generated between 1st January 2001 and 31st December 2018 (Supplemental Material C – *Remote sensing variables*). Persistent and recurrent green foliage fractions were derived from MOD13Q1 (Didan 2015) using the method of Donohue et al. (2009); and bare ground and litter cover fractions from MOD09A1 (Vermote 2015) using the method of Guerschman and colleagues (Guerschman 2019, Guerschman and Hill 2018). Persistent green cover fractions mainly characterise perennial plant species (e.g., non-deciduous shrubs and trees) whereas recurrent fractions mainly characterise annual species (e.g., grass and herbage, deciduous shrubs and trees). Litter fractions mainly characterise non-photosynthesising plant material and bare ground fractions are neither covered by litter nor green foliage. Collectively, these predominantly represent ecosystem structural characteristics and, to some extent, ecosystem function and composition. The long-term average and maximum statistics were derived from the 18-year time series, after first deriving the annual statistics. The maximum statistic for the persistent green cover fraction did not vary significantly from the mean and so it was not used, resulting in seven variables to characterise ecosystems. Each variable was then standardised (i.e., mean centred with a standard deviation of one) to ensure a common scaling for the calculation of Manhattan distances, used in the benchmarking algorithm.

Predicting reference ecosystem characteristics

Projection pursuit regression (PPR) was used to collectively model the seven standardised remotely sensed ecosystem characteristics to 29 candidate environmental covariates using a training data sample of around 100,000 reference sites (method detailed in Supplemental Material C: *Predicting reference ecosystem characteristics*). The resulting frequency distribution between observed versus predicted Manhattan distances for a random sample of *c.* 100,000 reference site-pairs is shown in Figure 6. This plot is indicative of the two-dimensional plot of differences used in condition benchmarking (Figure 3). The scatter of points is due to various processes such as alternative ecological states and seasonal variation for the same type of ecosystem, as well as inherent error in reference site assignments and other sources of model error. We expect more variability in observed remotely sensed ecosystem characteristics due to these natural dynamics than can be represented by their predictions based on stable environments.

Estimating ecosystem condition (benchmarking)

The approach to estimating ecosystem condition using observed and predicted remotely sensed ecosystem characteristics addresses the four challenges in an integrated way. The <u>predicted</u> remotely sensed ecosystem characteristics represent the reference state expected as a function of a stable natural environment (i.e., addressing the challenge of context dependency and scarcity of reference sites). The <u>observed</u> remotely sensed characteristics represent the ecosystem in its contemporary state which could be shaped by a combination of natural and/or anthropogenically-driven processes (i.e., addressing the challenges of alternative ecological states and seasonal dynamics). A multi-decadal remote sensing assessment period, over which the remote sensing variables are summarised, also helps to address the challenge of short- to medium-term temporal dynamics. Reference sites used as benchmarks then aim to characterise spatial and temporal variability among the alternative ecological states of any particular natural ecosystem in high integrity (i.e., reference condition).

The analysis is conducted using Manhattan distances derived from reference-reference and test-reference site-pairs (method detailed in Supplemental Material C: *Estimating ecosystem condition*). Two sets of Manhattan distances are first derived for each reference-reference site-pair using the training data: (1) observed and (2) predicted remotely sensed characteristics. A normalised two-dimensional frequency histogram of observed versus predicted distances is used to approximate a probability density surface of the ecosystem reference state. For each test site, 20 reference sites are selected that are the most relevant as benchmarks, and two sets of test-reference site Manhattan distances (observed and predicted) calculated. These distances are plotted over the density surface, to derive expected probabilities. Condition of the test site is then calculated as the predicted distance-weighted average of the 20 test-benchmark probabilities of being in reference condition (see Supplemental Material C Figure S20). The number of benchmarks is necessarily a trade-off between context dependency and the need to address the challenges of alternative ecological states and seasonal dynamics.

Calibrating ecosystem condition (0.0-1.0 scaling)

The output is calibrated and standardised in the range 0.0 (lowest – ecosystem removed) and 1.0 (highest – ecosystem integrity in reference condition). Calibration ideally draws on independent observations of site condition; however, such data are not readily available. While some land management agencies in Australia have implemented field protocols for estimating ecosystem condition to regulate native vegetation clearing—for example, the State of Queensland (Eyre et al.

2017, Eyre et al. 2015), the State of Victoria (DSE 2004, Parkes et al. 2003), Tasmania (Michaels 2006, Michaels et al. 2020), South Australia (DNR and NVC 2020), New South Wales (DPIE 2020, Oliver et al. 2021)—these have not been harmonised for consistent national use. Therefore, we developed a calibration approach using other sources of data (method detailed in Supplemental Material C: *Calibrating ecosystem condition*).

A piecewise linear rescaling algorithm with two inflection points was used to account for potential non-linearity. The x-axis coordinates for the inflection points were defined by the average uncalibrated condition values in areas of intensive land use (i.e., highly modified ecosystems) as of 2015–16 (ABARES 2022), and mapping of inferred reference sites *i.e. relatively natural areas), respectively (Table 1, Figure 5). The y-axis coordinates for condition scores were derived from a species-area relationship ($S=A^z$; for $z = 0.25$) transformation of PREDICTS project coefficients (i.e., the proportion of native species in an intact landscape which are found in paired modified habitats of that type) (Hudson et al. 2017) for 2015 global harmonised land use classes (LUH2 - Chini et al. 2020, Hurtt et al. 2020) that aligned with highly modified or relatively natural areas, and averaged using an area-weighting. The end points of the scaling (0, 1) were defined by minimum and maximum uncalibrated values, respectively (Figure 7). The calibrated result is shown in Figure 8. Implementation of calibration could alternatively use a monotonic spline (Dougherty et al. 1989).

Annual epochs of ecosystem condition

Annual epochs of ecosystem condition were derived using the same benchmarking process and calibration algorithm as the long-term epoch by substituting the observed long-term remotely sensed ecosystem characteristics with annual equivalents in test-benchmark comparisons.

Evaluating ecosystem condition

Validation was performed using two independent sources of ecosystem condition data derived through expert elicitation: (1) virtual transects (methods detailed in Supplemental Material) and (2) site condition assessments (White M. D. et al. 2023). Nine ecologists with extensive field experience in specific regions visually assessed condition at 11 evenly spaced points along one or two of 11 pre-defined virtual transects using Google Earth imagery. The transects traversed large swathes of the Australian continent. Twenty-one experts contributed 314 site condition assessments through the Habitat Condition Assessment Tool (Brenton et al. 2018), which included a method for expert cross-calibration enabling the results to be rescaled (White M. D. et al. 2023). A Major Axis Type-II regression (Legendre and Legendre 2012), which assumes error variances are approximately equal in the comparisons, demonstrated reasonable agreement between HCAS and each set of expert scores (Figure 9).

The calibrated HCAS scores were also compared with categorical mapping of native vegetation modification levels derived from a wide range of land use and land cover datasets for Australia (Lesslie et al. 2010) consistent with the Vegetation Assets, States and Transitions (VAST) narrative framework (Thackway and Lesslie 2006, 2008). The continuous HCAS scores were assigned to discrete VAST classes on the basis of elicited expert's condition scores (methods detailed in Supplemental Material) to enable a comparison of ordered categories. Concordance between the two datasets was qualitatively assessed using a confusion matrix. To approximate the temporal range of the VAST spatial data (1995-2006), the average of the six HCAS annual epochs of the ecosystem site condition index in the overlapping temporal range, 2001 to 2006, were used for the comparison (Supplemental Material C Figure S36). Overall concordance for the comparison of five common categories was 42% indicating moderate agreement and, when collapsed to two classes depicting relatively natural versus intensively modified areas, overall concordance was 87%, indicating high agreement.

Example applications

Two HCAS versions derived using MODIS remote sensing data have been published as continent-wide datasets (Harwood et al. 2023a, Harwood et al. 2021) along with several regional versions (listed in Supplemental Material B Table S2), and applied to both operational and research uses. For example, Mokany et al. (2022a) used HCAS as one of the primary inputs to habitat-based biodiversity assessment for ecosystem accounting in the extensive Murray-Darling Basin region of Australia. Giljohann et al. (2022) used HCAS as the main input to a continent-wide connectivity index for operational use in conservation planning and policy, and both are included in Australian Government performance reporting on environmental outcomes (DCCEEW 2023). Williams K. J. (2023) used HCAS to define three regions of Australia – intensive use, extensive use and relatively natural – for operational state of the environment reporting, and for reporting on average condition within a particular ecosystem (Williams K. J. et al. 2021a). Forbes et al. (2021) used ecosystem condition from the proof-of-concept version of HCAS (Harwood et al. 2016) along with other factors to model the drivers and risks of the infectious zoonotic disease, cryptosporidiosis. Nowrouzi et al. (2019) used that same earlier version of ecosystem condition from HCAS, combined with a model of native ant species compositional diversity in rainforest, to predict the impacts of climate change on effective habitat area. Ward et al. (2024) used HCAS as a line of evidence in assessing the impacts of forest harvesting and degradation on threatened species. Williams K. J. et al. (2023c) used a trend analysis over the 18 years of HCAS annual epochs to map locations of statistically significant change in condition, summarised as either increasing or decreasing extents (e.g., Table 2 and Figure 11). Giljohann et al. (2024) used HCAS annual epochs to track progress in providing habitat for threatened species over time and space. These applications are just a few of the multiple ways in which ecosystem condition from HCAS can be used.

HCAS as an indicator of ecosystem integrity

The HCAS conceptual framework and method addresses all evaluation criteria outlined by Hansen et al. (2021) for systematically monitoring and evaluating trends in ecosystem integrity (reproduced in Box 1). As far as possible, remote sensing variables are selected to represent common ecosystem characteristics of structure, function, and composition (**criterion 1,** Box 1). HCAS to date has largely been based on structural variables from remote sensing (e.g., fractional cover of visible ground and canopy properties), but the method can flexibly take advantage of new remote sensing products that provide greater coverage of ecosystem properties or improve on previous measures.

The HCAS has been successfully applied to the continent of Australia, which is representative of the majority of global biomes (**criterion 2,** Box 1). With inclusion of reference sites agreed by participating countries as indicative of their natural areas (Neugarten et al. 2024, Xiao et al. 2024), the approach can feasibly be applied across terrestrial ecosystems globally. The pilot application of HCAS used a 1 km grid to test method feasibility (Harwood et al. 2016) and then implemented using a 250 m grid across Australia (Harwood et al. 2023a, Harwood et al. 2021) (**criterion 3,** Box 1). The next phase of work at 90 m is underway (Munroe et al. 2024). The first published version of HCAS introduced a time series of annual epochs as a derivative of the long-term model (Harwood et al. 2021, Williams K. J. et al. 2021b). HCAS annual epochs have been used to estimate change in condition using statistical trend analyses considering serial correlation bias (Harwood et al. 2023a, Lehmann et al. 2023, Williams K. J. et al. 2023c), as demonstrated in Figure 11 (**criterion 3,** Box 1).

The HCAS method can be applied at any scale and region depending on suitable resolution input data. Being a site-level assessment across whole landscapes or continents, regional assessments are inherently comparable, and results can be aggregated at higher levels for reporting without introducing bias (**criterion 4,** Box 1). Alternatively, analyses of trends in annual epochs over a given time-series applied to individual pixels can be used to derive an average estimate of condition change per year. This estimate can be aggregated by summing pixel values to derive an area estimate (e.g., change in condition-hectares per year), and then multiplied by the number of years of the trend analysis, for any subsequent regionalisation, without introducing scaling bias (e.g., Figure 11).

The general approach to HCAS has been published (Harwood et al. 2016), and successive technical enhancements co-designed through stakeholder and science consultative processes, have also been peer reviewed and published, as outlined in Supplemental Material B Table S2 (**criterion**

5, Box 1). Detailed evaluations of the output were conducted using multiple lines of evidence, to inform a schedule of limitations and recommendations for continuous improvement (Williams K. J. et al. 2023c, Williams K. J. et al. 2021b). The data and metadata are publicly available using open standards – CC By licenses (e.g., Harwood et al. 2023a, Harwood et al. 2023b, Harwood et al. 2021).

The HCAS framework is inherently referenced to states characteristic of the climatic, geomorphic, and native community ecosystem (**criterion 6,** Box 1). The reference ecosystem model part of HCAS (Figure 4) uses climatic and geomorphic covariates of native ecosystems in their high integrity reference state to predict the reference state characteristic of ecosystems using satellite remote sensing as the observation platform. These predicted reference state characteristics are used in the benchmarking part of HCAS (Figure 4) to ensure correct and most effective use of scarce reference site data in estimating condition at a site of interest. Furthermore, HCAS has been designed to explicitly account for alternative ecological states among reference ecosystems. The resulting spatial data can be intersected with other map products, for example depicting type and extent of native ecosystems (Williams K. J. et al. 2023b), and for nuanced reporting on gains and losses with implications for biodiversity (Mokany et al. 2022a).

Future directions

Given access to the three input datasets (reference sites depicting relatively natural areas, remotely sensed ecosystem characteristics, and environmental covariate determinants of ecosystem diversity), the HCAS approach could be applied to any region of the world. Identification of suitable reference sites is the most limiting input (Harwood et al. 2016). However, using multiple lines of evidence from time series human footprint mapping (Grantham et al. 2020, Watson and Venter 2019, Williams B. A. et al. 2020a), contemporary natural areas can be inferred where there are low levels of human-influenced ecosystem conversions (Neugarten et al. 2024, Xiao et al. 2024); thereby enabling global or country-based applications to support consistent reporting on ecosystem condition and integrity under the CBD (CBD 2022b) and SEEA EA frameworks (United Nations et al. 2021), and for business reporting on environmental performance, nature-related risks or nature positive outcomes (e.g., zu Ermgassen et al. 2022).

Within Australia, continued investment in HCAS is addressing the limitations summarised in Williams K. J. et al. (2021b). Prioritised improvements include: 1) spatial resolution, accuracy and utility of HCAS by undertaking further work on reference sites, environmental covariates and remote sensing variables; 2) implementing uncertainty quantification to derive confidence intervals and guide appropriate use in decision-making; 3) revising the benchmarking algorithm to incorporate very low condition ('removed') sites (~0.0) in addition to high ecosystem integrity reference sites (~1.0) as the method originally envisaged (see Fig. 1 in Harwood et al. 2016); 4) introduction of parameter tuning in the benchmarking algorithm to address inherent trade-offs; 5) streamlining workflows and refactoring software for transparency, traceability and near real time generation of outputs; and 6) developing methods to support interpretation and attribution of condition change and trends including detection of abrupt versus gradual and other types of change (e.g., Bergstrom et al. 2021). The next phase of work on HCAS is being developed at 90 m pixel resolution utilising the long time-series of Landsat data (Commonwealth of Australia 2021) and an extensive 90 m compilation of environmental covariates filtered to remove signatures of anthropogenic land use (Searle 2023, Searle et al. 2022). In addition to terrestrial environments, the three inputs will be extend to improve depiction of land surface condition in wetland, riparian and floodplain environments (Munroe et al. 2024).

A key limitation in remote sensing of ecosystem condition is the ability to detect ecosystem characteristics below a closed canopy (Lawley et al. 2016, Tehrany et al. 2017); for example, where a canopy is intact, but the structure below has been modified by alien invasive browsing ungulates (Mitchell et al. 2017, Russo et al. 2023). These changes will not be evident from optical satellite sensors but may become evident with longer periods of monitoring from satellite-based radar and lidar detecting complex three-dimensional woody structures (Bergen et al. 2009, Mitchell et al.

2017). In the interim, therefore, field observation, expert opinion, empirical analysis and ecological theory will continue to be needed to infer processes impacting ecosystem integrity at the site level beyond that detected using satellite remote sensing (Cavender-Bares et al. 2020, Gao et al. 2020).

To fill some of these gaps in satellite-based remote detection of ecosystem characteristics that have been negatively impacted by surrounding pressures, we applied a simple proximity algorithm to infer diffuse local pressures that potentially negatively influence the realised condition at a site (Supplemental Material C: *Deriving ecosystem site condition*). The derived index of 'ecosystem site condition' (Figure 10) provides a slightly improved index for operational purposes such as ecosystem accounting (e.g., Richards et al. 2023). More comprehensive, context-specific edge effects analyses would be equally applicable (e.g., Ries et al. 2017, Zurita et al. 2012).

While a further novel aspect of HCAS is its ability to incorporate new remote sensing technologies as these become available, these also need to address the challenges illustrated in Figure 1. For example, the need for a multi-decadal time-series in order to understand and correctly separate effects of natural processes of response to disturbance regimes from anthropogenic drivers that remove (conversions) or modify ecosystems (Senf and Seidl 2021, Zhu et al. 2022). Therefore, as satellite technologies evolve, such as satellite-based radar/lidar and hyperspectral (Ustin and Middleton 2021), it may be several years before they can be effectively used for ecosystem condition assessment. A promising alternative is for new technologies to integrated with legacy time series to fill cloud gaps or sensor error in remotely sensed ecosystem characteristics (e.g., Myroniuk et al. 2023).

Conclusions

HCAS is unique among approaches that have been developed over recent decades to monitor ecosystem integrity using satellite remote sensing across whole regions and continents. It was specifically designed to address four challenges of satellite-based remote sensing: context dependency, alternative ecological states, seasonal dynamics and scarce reference data. The methodology has evolved significantly since the proof of concept was first published by Harwood et al. (2016) and this evolution continues as new data sets and approaches are incorporated.

HCAS outputs an index of ecosystem condition and the remotely sensed ecosystem characteristics are condition variables under the SEEA-EA ecosystem condition typology (Keith et al. 2020, United Nations et al. 2021). The conceptual framework is entirely consistent with the theoretical concepts outlined by Czúcz et al. (2021), wherein it would be feasible to reconfigure the HCAS workflow to additionally output individual remotely sensed indicators of ecosystem condition. Among the 43 headline biodiversity indicators under the Kunming-Montreal Global Biodiversity Framework (CBD 2022b), Indicator A.2: *Extent of natural ecosystems* is most relevant to ecosystem condition, because it can inform on thresholds for naturalness. The HCAS method is also suited for systematically monitoring and evaluating trends in ecosystem integrity, meeting all evaluation criteria outlined by Hansen et al. (2021).

The HCAS method has been successfully applied to the entire continent of Australia and used in a wide range of research and policy applications. The three inputs are readily available globally and the two-stage model to effectively addresses four key challenges of monitoring ecosystem condition (i.e. integrity) from space. The method can be easily adapted and applied to other countries and globally, to complement on-ground assessments, land use and ecosystem type mapping, and as an input to biodiversity assessment. HCAS can help refine our understanding about the global status of habitats for biodiversity as an input to land use planning and landscape management strategies.

REFERENCES

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

Andrews A. 1990. Fragmentation of Habitat by Roads and Utility Corridors: A Review. Australian Zoologist 26: 130-141.

Austin MP. 2013. Vegetation and Environment: Discontinuities and Continuities. Pages 71-106. Vegetation Ecology.

Barker R. 2019. Corporate natural capital accounting. Oxford Review of Economic Policy 35: 68-87.

Bastin G, Scarth P, Chewings V, Sparrow A, Denham R, Schmidt M, O'Reagain P, Shepherd R, Abbott B. 2012. Separating grazing and rainfall effects at regional scale using remote sensing imagery: A dynamic reference-cover method. Remote Sensing of Environment 121: 443-457.

Bergen KM, Goetz SJ, Dubayah RO, Henebry GM, Hunsaker CT, Imhoff ML, Nelson RF, Parker GG, Radeloff VC. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences 114.

Bergstrom DM, et al. 2021. Combating ecosystem collapse from the tropics to the Antarctic. Global Change Biology 27: 1692-1703.

Blankenship K, Frid L, Smith JL. 2015. A state-and-transition simulation modeling approach for estimating the historical range of variability. AIMS Environmental Science 2: 253.

Bonannella C, Hengl T, Parente L, de Bruin S. 2023. Biomes of the world under climate change scenarios: increasing aridity and higher temperatures lead to significant shifts in natural vegetation. PeerJ 11: e15593

Bradshaw CJA, et al. 2021. Underestimating the Challenges of Avoiding a Ghastly Future. Frontiers in Conservation Science 1.

Brand V, Lacey J, Tutton J. 2023. Social Licence as a Regulatory Concept: An Empirical Study of Australian Company Directors. Univeristy of New South Wales Law Journal 46: 111-142.

Brenton P, Pirzl R, Raisbeck-Brown N, Dickson F, White M, Warnick A, Williams KJ. 2018. The Habitat Condition Assessment Tool: BioCollect data collection portal hosted by the Atlas of Living Australia, https://biocollect.ala.org.au/hcat. Canberra, Australia: Atlas of Living Australia.

Burton PJ, Jentsch A, Walker LR. 2020. The Ecology of Disturbance Interactions. BioScience 70: 854-870.

Cavender-Bares J, Gamon JA, Townsend PA. 2020. The Use of Remote Sensing to Enhance Biodiversity Monitoring and Detection: A Critical Challenge for the Twenty-First Century. Pages 1-12 in Cavender-Bares J, Gamon JA, Townsend PA, eds. Remote Sensing of Plant Biodiversity. Cham: Springer International Publishing.

CBD. 2022a. Decision adopted by the Conference of the Parties to the Convention on Biological Diversity: 15/4. Kunming-Montreal Global Biodiversity Framework. Paper presented at Conference of the Parties to the Convention on Biological Diversity, Fifteenth meeting – Part II, Montreal, Canada, 7-19 December 2022, Agenda item 9A, Montreal, Canada.

—. 2022b. Monitoring framework for the Kunming-Montreal Global Biodiversity Framework, Conference Of the Parties to the Convention on Biological Diversity Fifteenth meeting (CBD/COP/DEC/15/5). Paper presented at Conference of the Parties to the Convention on Biological Diversity, Fifteenth meeting – Part II, Montreal, Canada, 7-19 December 2022, Agenda item 9B; 19 December 2022, Montreal, Canada.

Chini LP, Hurtt G, Sahajpal R, Frolking S. 2020. GLM2 Code (Global Land-use Model 2) for generating LUH2 datasets (Land-Use Harmonization 2). Zenodo.

Cohen WB, Maiersperger TK, Spies TA, Oetter DR. 2001. Modelling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. International Journal of Remote Sensing 22: 2279-2310.

Commonwealth of Australia. 2021. Geoscience Australia Landsat Analysis Ready Data Collection 3. Canberra, Australia: National Computing Infrastructure.

Cowie AL, et al. 2018. Land in balance: The scientific conceptual framework for Land Degradation Neutrality. Environmental Science & Policy 79: 25-35.

Czúcz B, Keith H, Maes J, Driver A, Jackson B, Nicholson E, Kiss M, Obst C. 2021. Selection criteria for ecosystem condition indicators. Ecological Indicators 133: 108376.

Daniel CJ, Frid L, Sleeter BM, Fortin M-J. 2016. State-and-transition simulation models: a framework for forecasting landscape change. Methods in Ecology and Evolution 7: 1413-1423.

DAWE. 2020. Australia - Pre-1750 Major Vegetation Subgroups - NVIS Version 6.0 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

DCCEEW. 2023. Annual Report 2022-23. Canberra, Australia: Australian Government Department of Climate Change, Energy, the Environment and Water, Canberra.

Department of the Environment. 2014. Interim Biogeographic Regionalisation for Australia (IBRA), Version 7 (Subregions). Canberra, Australia: Australian Government Department of the Environment.

DES. 2021. Spatial BioCondition: Vegetation condition map for Queensland. Brisbane, Australia: Queensland Government Department of Environment and Science.

Didan K. 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. https://lpdaac.usgs.gov/: NASA EOSDIS Land Processes DAAC.

DNR, NVC. 2020. Native Vegetation Council (NVC) Bushland Assessment Manual. Adelaide, Australia: Department of Natural Resources, Government of South Australia.

Doherty TS, Dickman CR, Nimmo DG, Ritchie EG. 2015. Multiple threats, or multiplying the threats? Interactions between invasive predators and other ecological disturbances. Biological Conservation 190: 60-68.

Donohue RJ, McVicar TR, Roderick ML. 2009. Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. Global Change Biology 15: 1025-1039.

Donohue RJ, Mokany K, McVicar TR, O'Grady AP. 2022. Identifying management-driven dynamics in vegetation cover: Applying the Compere framework to Cooper Creek, Australia. Ecosphere 13: e4006.

Dougherty RL, Edelman A, Hyman JM. 1989. Positivity-, monotonicity-, or convexity-preserving cubic and quintic Hermite interpolation. Mathematics of Computation 52: 471-494.

DPIE. 2020. Biodiversity Assessment Method 2020 Operational Manual – Stage 1. Sydney, Australia: State of NSW through the Department of Planning, Industry and Environment.

DSE. 2004. Vegetation Quality Assessment Manual–Guidelines for applying the habitat hectares scoring method. Version 1.3. Melbourne, Australia: Victorian Government Department of Sustainability and Environment.

Eyre TJ, Kelly AL, Neldner VJ. 2017. Method for the Establishment and Survey of Reference Sites for BioCondition. Version 3. Brisbane, Australia: Queensland Herbarium, Department of Science, Information Technology and Innovation.

Eyre TJ, Kelly AL, Neldner VJ, Wilson BA, Ferguson DJ, Laidlaw MJ, Franks AJ. 2015. BioCondition: A Condition Assessment Framework for Terrestrial Biodiversity in Queensland. Assessment Manual. Version 2.2. Brisbane, Australia: Queensland Herbarium, Department of Science, Information Technology, Innovation and the Arts.

Ferrier S, Harwood TD, Ware C, Hoskins AJ. 2020. A globally applicable indicator of the capacity of terrestrial ecosystems to retain biological diversity under climate change: The bioclimatic ecosystem resilience index. Ecological Indicators 117: 106554.

Forbes O, Hosking R, Mokany K, Lal A. 2021. Bayesian spatio-temporal modelling to assess the role of extreme weather, land use change and socio-economic trends on cryptosporidiosis in Australia, 2001–2018. Science of The Total Environment 791: 148243.

Forman RTT, Alexander LE. 1998. Roads and their major ecological effects. Annual Review of Ecology and Systematics 29: 207-231.

Gao Y, Skutsch M, Paneque-Gálvez J, Ghilardi A. 2020. Remote sensing of forest degradation: a review. Environmental Research Letters 15: 103001.

Giljohann KM, Mokany K, Ferrier S, Harwood TD, Ware C, Williams KJ. 2024. Accounting for Threatened Species: Estimating Historical and Recent Change in Habitat. Ecological Indicators in press: 1-27.

Giljohann KM, Drielsma M, Love J, Pinner L, Lyon P, Harwood TD, Williams KJ, Ferrier S. 2022. Technical description for a National Connectivity Index version 2.0 based on HCAS version 2.1. Canberra, Australia: CSIRO.

Giljohann KM, et al. 2023. Experimental ecosystem accounts and supplementary data by ecosystem types: ecosystem condition in the Flinders, Norman and Gilbert river catchments. In: Experimental ecosystem accounts and supplementary data for the Flinders, Norman and Gilbert river catchments in Queensland. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Grantham HS, et al. 2020. Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. Nature Communications 11: 5978.

Guerschman JP. 2019. Fractional Cover - MODIS, CSIRO algorithm. Version 3.1 (Dataset): Terrestrial Ecosystem Research Network (TERN).

Guerschman JP, Hill MJ. 2018. Calibration and validation of the Australian fractional cover product for MODIS collection 6. Remote Sensing Letters 9: 696-705.

Haberl H, Erb K-H, Krausmann F. 2014. Human Appropriation of Net Primary Production: Patterns, Trends, and Planetary Boundaries. Annual Review of Environment and Resources 39: 363-391.

Hansen AJ, et al. 2021. Toward monitoring forest ecosystem integrity within the post-2020 Global Biodiversity Framework. Conservation Letters 14: e12822.

Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G, White M. 2016. Habitat Condition Assessment System: A new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. Methods in Ecology and Evolution 7: 1050-1059.

Harwood TD, et al. 2023a. 9-arcsecond gridded HCAS 2.3 (2001-2018) base model estimation of habitat condition and general connectivity for terrestrial biodiversity, ecosystem site condition, annual epochs and 18-year trends for continental Australia. Data Collection. Canberra, Australia: CSIRO.

Harwood TD, et al. 2023b. Account-ready data: ecosystem site condition in the Murray-Darling Basin. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Harwood TD, et al. 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. v7. Data collection. Canberra, Australia: CSIRO.

Hengl T, Walsh MG, Sanderman J, Wheeler I, Harrison SP, Prentice IC. 2018. Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential. PeerJ 6: e5457.

Hoskins AJ, Bush A, Gilmore J, Harwood TD, Hudson LN, Ware C, Williams KJ, Ferrier S. 2016. Downscaling land-use data to provide global 30″ estimates of five land-use classes. Ecology and Evolution 6: 3040-3055.

Huang B, et al. 2020. A new remote-sensing-based indicator for integrating quantity and quality attributes to assess the dynamics of ecosystem assets. Global Ecology and Conservation 22: e00999.

Hudson LN, et al. 2017. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecology and Evolution 7: 145-188.

Hurtt GC, et al. 2020. Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. Geosci. Model Dev. 13: 5425-5464.

Jakobsson S, et al. 2020. Setting reference levels and limits for good ecological condition in terrestrial ecosystems – Insights from a case study based on the IBECA approach. Ecological Indicators 116: 106492.

Keith H, Czúcz B, Jackson B, Driver A, Nicholson E, Maes J. 2020. A conceptual framework and practical structure for implementing ecosystem condition accounts. One Ecosystem 5.

Keys PW, Barnes EA, Carter NH. 2021. A machine-learning approach to human footprint index estimation with applications to sustainable development. Environmental Research Letters: 1-22.

Kocev D, Džeroski S, White MD, Newell GR, Griffioen P. 2009. Using single- and multi-target regression trees and ensembles to model a compound index of vegetation condition. Ecological Modelling 220: 1159-1168.

Lawley V, Lewis M, Clarke K, Ostendorf B. 2016. Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. Ecological Indicators 60: 1273-1283.

Legendre P, Legendre L. 2012. Numerical ecology. Number 24 in Developments in Environmental Modelling. 3rd edition. Amsterdam: Elsevier.

Lehmann EA, Ferrier S, Williams KJ. 2023. Trend test for serially correlated time series data. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2023-3255. Canberra, Australia: CSIRO.

Lehmann EA, Williams KJ, Harwood TD, Ferrier S. 2021. A not-too-technical introduction to the HCAS v2.x mechanics: a revised method for mapping habitat condition across Australia. Publication number EP211609. Canberra, Australia: CSIRO.

Lesslie R, Thackway R, Smith J. 2010. A national-level Vegetation Assets, States and Transitions (VAST) dataset for Australia (version 2.0) Canberra, Australia: Bureau of Rural Sciences, Australian Government.

Love J, Drielsma M, Williams KJ, Thapa R. 2020. Integrated model-data fusion approach to measuring habitat condition for ecological integrity reporting: Implementation for habitat condition indicators. Sydney, Australia: NSW Department of Planning, Industry and Environment.

McGeoch MA, et al. 2019. Measuring continuous compositional change using decline and decay in zeta diversity. Ecology 100: e02832.

McNellie MJ, Oliver I, Dorrough J, Ferrier S, Newell G, Gibbons P. 2020. Reference state and benchmark concepts for better biodiversity conservation in contemporary ecosystems. Global Change Biology 26: 6702-6714.

McNellie MJ, Oliver I, Ferrier S, Newell G, Manion G, Griffioen P, White M, Koen T, Somerville M, Gibbons P. 2021. Extending vegetation site data and ensemble models to predict patterns of foliage cover and species richness for plant functional groups. Landscape Ecology 36: 1391-1407.

Michaels K. 2006. A Manual for Assessing Vegetation Condition in Tasmania, Version 1.0. Hobart: Resource Management and Conservation, Department of Primary Industries, Water and Environment.

Michaels K, Panek D, Kitchener A, eds. 2020. TASVEG VCA Manual: A manual for assessing vegetation condition in Tasmania, Version 2.0 Hobart, Australia: Natural and Cultural Heritage, Department of Primary Industries, Parks, Water and Environment, Tasmania.

Mitchell AL, Rosenqvist A, Mora B. 2017. Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. Carbon Balance and Management 12: 9.

Mokany K, Ware C, Harwood TD, Schmidt RK, Tetreault-Campbell S, Ferrier S. 2022a. Habitat-based biodiversity assessment for ecosystem accounting in the Murray-Darling Basin. Conservation Biology 36: e13915.

Mokany K, McCarthy J, Falster D, Gallagher R, Harwood T, Kooyman R, Westoby M. 2022b. Harmonised Australian Vegetation Plot dataset (HAVPlot). v1. Data Collection. . Canberra, Australia: CSIRO.

Moore CE, et al. 2016. Reviews and syntheses: Australian vegetation phenology: new insights from satellite remote sensing and digital repeat photography. Biogeosciences 13: 5085-5102.

Munroe S, Malley C, Tetreault Campbell S, Hazelwood L, Gunawardana D, Perrott L, Williams KJ. 2024. Redesign kicks off for the Habitat Condition Assessment System. Pages 2. Canberra, Australia: CSIRO.

Murphy BP, Bowman DMJS. 2012. What controls the distribution of tropical forest and savanna? Ecology Letters 15: 748-758.

Murray C, Larson A, Goodwill J, Wang Y, Cardace D, Akanda AS. 2022. Water Quality Observations from Space: A Review of Critical Issues and Challenges. Environments 9: 125.

Murray NJ, Keith DA, Bland LM, Ferrari R, Lyons MB, Lucas R, Pettorelli N, Nicholson E. 2018. The role of satellite remote sensing in structured ecosystem risk assessments. Science of The Total Environment 619-620: 249-257.

Myroniuk V, Zibtsev S, Bogomolov V, Goldammer JG, Soshenskyi O, Levchenko V, Matsala M. 2023. Combining Landsat time series and GEDI data for improved characterization of fuel types and canopy metrics in wildfire simulation. Journal of Environmental Management 345: 118736.

Naicker R, Mutanga O, Peerbhay K, Odebiri O. 2024. Estimating high-density aboveground biomass within a complex tropical grassland using Worldview-3 imagery. Environmental Monitoring and Assessment 196: 370.

Neugarten RA, et al. 2024. Mapping the planet’s critical areas for biodiversity and nature’s contributions to people. Nature Communications 15: 261.

Newell GR, White MD, Griffioen P, Conroy M. 2006. Vegetation condition mapping at a landscape-scale across Victoria. Ecological Management & Restoration 7: S65-S68.

Nicholson E, et al. 2021. Scientific foundations for an ecosystem goal, milestones and indicators for the post-2020 global biodiversity framework. Nature Ecology & Evolution 5: 1338-1349.

Nowrouzi S, Bush A, Harwood T, Staunton KM, Robson SKA, Andersen AN. 2019. Incorporating habitat suitability into community projections: Ant responses to climate change in the Australian Wet Tropics. Diversity and Distributions 25: 1273-1288.

Oliver I, Dorrough J, Seidel J. 2021. A new Vegetation Integrity metric for trading losses and gains in terrestrial biodiversity value. Ecological Indicators 124: 107341.

Parkes D, Newell G, Cheal D. 2003. Assessing the quality of native vegetation: The 'habitat hectares' approach. Ecological Management & Restoration 4: S29-S38.

Pausas JG, Bond WJ. 2020. Alternative Biome States in Terrestrial Ecosystems. Trends in Plant Science 25: 250-263.

Pettorelli N, et al. 2017. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. Remote Sensing in Ecology and Conservation 4: 71-93.

Pickup G, Bastin GN, Chewings VH. 1994. Remote-sensing-based condition assessment for nonequilibrium rangelands under large-scale commercial grazing. Ecological Applications 4: 497-517.

Purvis A, et al. 2018. Modelling and Projecting the Response of Local Terrestrial Biodiversity Worldwide to Land Use and Related Pressures: The PREDICTS Project. Advances in Ecological Research 58: 201-241.

Richards AE, Lucas R, Clewley D, Prober SM, Schmidt RK, Tetreault-Campbell S, Ware C. 2021. Assessing extent of ecosystem types and condition states at Gunbower-Koondrook-Perricoota Forest Icon Site. A technical report for the Land and Ecosystem Accounts Project. . Canberra, Australia: CSIRO.

Richards AE, et al. 2023. Experimental ecosystem accounts for the Flinders, Norman and Gilbert river catchments. A synthesis report from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO.

Ries L, Murphy SM, Wimp GM, Fletcher RJ. 2017. Closing Persistent Gaps in Knowledge About Edge Ecology. Current Landscape Ecology Reports 2: 30-41.

Roche PK, Campagne CS. 2017. From ecosystem integrity to ecosystem condition: a continuity of concepts supporting different aspects of ecosystem sustainability. Current Opinion in Environmental Sustainability 29: 63-68.

Russo NJ, Davies AB, Blakey RV, Ordway EM, Smith TB. 2023. Feedback loops between 3D vegetation structure and ecological functions of animals. Ecology Letters 26: 1597-1613.

Schneider MF. 2001. Habitat loss, fragmentation and predator impact: spatial implications for prey conservation. Journal of Applied Ecology 38: 720-735.

Scott A, Holtby R, East H, Lannin A. 2022. Mainstreaming the Environment: Exploring pathways and narratives to improve policy and decision-making. People and Nature 4: 201-217.

Searle R. 2023. HCAS Optimised SLGA Products. Canberra, Australia: Terrestrial Ecosystem Research Network.

Searle R, Malone B, Wilford J, Austin J, Ware C, Webb M, Roman Dobarco M, Van Niel T. 2022. TERN Digital Soil Mapping Raster Covariate Stacks. Canberra, Australia: CSIRO. Data Collection.

Senf C, Seidl R. 2021. Mapping the forest disturbance regimes of Europe. Nature Sustainability 4: 63-70.

Sims NC, Barger NN, Metternicht GI, England JR. 2020. A land degradation interpretation matrix for reporting on UN SDG indicator 15.3.1 and land degradation neutrality. Environmental Science and Policy 114: 1-6.

Sims NC, England JR, Newnham GJ, Alexander S, Green C, Minelli S, Held A. 2019. Developing good practice guidance for estimating land degradation in the context of the United Nations Sustainable Development Goals. Environmental Science & Policy 92: 349-355.

Staver AC, Archibald S, Levin SA. 2011. The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. Science 334: 230-232.

Stevenson SL, Watermeyer K, Caggiano G, Fulton EA, Ferrier S, Nicholson E. 2021. Matching biodiversity indicators to policy needs. Conservation Biology 35: 522-532.

Tallis H, et al. 2012. A Global System for Monitoring Ecosystem Service Change. Bioscience 62: 977-986.

Tehrany MS, Kumar L, Drielsma MJ. 2017. Review of native vegetation condition assessment concepts, methods and future trends. Journal for Nature Conservation 40: 12-23.

Thackway R, Lesslie R. 2006. Reporting vegetation condition using the Vegetation Assets, States and Transitions (VAST) framework. Ecological Management & Restoration 7: S53-S62.

—. 2008. Describing and mapping human-induced vegetation change in the Australian landscape. Environmental Management 42: 572-590.

Timmermans J, Daniel Kissling W. 2023. Advancing terrestrial biodiversity monitoring with satellite remote sensing in the context of the Kunming-Montreal global biodiversity framework. Ecological Indicators 154: 110773.

TNFD. 2023. Recommendations of the Taskforce on Nature-related Financial Disclosures. online: tnfd.global: Taskforce on Nature-related Financial Disclosures (TNFD).

Ulrich W, et al. 2023. From biodiversity to health: Quantifying the impact of diverse ecosystems on human well-being. People and Nature 5: 69-83.

UN. 2015. Transforming our world: the 2030 Agenda for Sustainable Development. online https://sustainabledevelopment.un.org/post2015/transformingourworld: United Nations, Department of Economic and Social Affairs.

UNCCD. 2016. Land in Balance - The Scientific Conceptual Framework for Land Degradation Neutrality: Science-Policy Brief 02 - September 2016. Available at: https://www.unccd.int/publications/land-balance: United Nations Convention to Combat Desertification.

Unerman J, Bebbington J, O'dwyer B. 2018. Corporate reporting and accounting for externalities. Accounting and Business Research 48: 497-522.

United Nations. 1992. Convention on biological diversity. online http://www.cbd.int/doc/legal/cbd-en.pdf: United Nations.

United Nations et al. 2021. System of Environmental-Economic Accounting—Ecosystem Accounting (SEEA-EA). White cover publication, pre-edited text subject to official editing. www.unstats.un.org: United Nations Committee of Experts on Environmental-Economic Accounting, United Nations Statistical Division.

United Nations Forum on Forests Secretariat. 2019. Global forest goals and targets of the UN strategic plan for forests 2030. New York, USA: United Nations Department of Economic and Social Affairs.

Ustin SL, Middleton EM. 2021. Current and near-term advances in Earth observation for ecological applications. Ecological Processes 10: 1.

Venter O, et al. 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nature Communications 7: 12558.

Vermote E. 2015. MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006 [Data set]. https://lpdaac.usgs.gov/: NASA EOSDIS Land Processes DAAC.

Wang H-J, Zhang R-H, Cole J, Chavez F. 1999. El Niño and the related phenomenon Southern Oscillation (ENSO): The largest signal in interannual climate variation. Proceedings of the National Academy of Sciences 96: 11071-11072.

Ward M, et al. 2024. Shifting baselines clarify the impact of contemporary logging on forest-dependent threatened species. Conservation Science and Practice 6: e13185.

Watson JEM, Venter O. 2019. Mapping the Continuum of Humanity's Footprint on Land. One Earth 1: 175-180.

White K, Habib R, Hardisty DJ. 2019. How to SHIFT Consumer Behaviors to be More Sustainable: A Literature Review and Guiding Framework. Journal of Marketing 83: 22-49.

White MD, et al. 2023. Towards a continent-wide ecological site condition database using calibrated expert evaluations. Ecological Applications 33: e2729.

Williams BA, et al. 2020a. Change in Terrestrial Human Footprint Drives Continued Loss of Intact Ecosystems. One Earth 3: 371-382.

Williams KJ. 2023. Land use zones derived from HCAS v2.1 and IBRA 7 as used in the Land Chapter of the Australia State of the Environment 2021 report. Canberra, Australia: CSIRO. Data Collection.

Williams KJ, Hunter B, Schmidt RK, Woodward E, Cresswell I. 2021a. Australia State of the Environment 2021 - Land. Independent report to the Australian Government Minister for the Environment, Commonwealth of Australia.

Williams KJ, Harwood TD, Giljohann KM, Ferrier S, Lehmann EA, Ware C, Stewart SB, Tetreault-Campbell S, Schmidt RK. 2023a. Extended methods for developing account-ready data: ecosystem site condition in the Murray-Darling Basin. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2022-5750. Canberra, Australia: CSIRO.

Williams KJ, Liu N, Giljohann KM, Ferrier S, Harwood TD, Lehmann EA, Richards AE, Stewart SB, Tetreault-Campbell S, Schmidt RK. 2023b. Methods for compiling accounts: ecosystem condition in the Western Australian Wheatbelt. In: Experimental ecosystem accounts and supplementary data for the Western Australian Wheatbelt. A data collection from the Regional Ecosystem Accounting Pilot projects. CSIRO. Data Collection.

Williams KJ, et al. 2023c. Extended methods used in developing the Habitat Condition Assessment System (HCAS) version 2.3, ecosystem condition account-ready data and experimental accounts for two mixed-use landscapes. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2023-1426. Canberra, Australia: CSIRO.

Williams KJ, et al. 2021b. Habitat Condition Assessment System (HCAS version 2.1): Enhanced method for mapping habitat condition and change across Australia. Canberra, Australia: CSIRO.

Williams KJ, et al. 2020b. Habitat Condition Assessment System: developing HCAS version 2.0 (beta). A revised method for mapping habitat condition across Australia. Canberra, Australia: Publication number EP21001. CSIRO Land and Water.

Wulder MA, et al. 2022. Fifty years of Landsat science and impacts. Remote Sensing of Environment 280: 113195.

Xiao Y, Wang Q, Zhang HK. 2024. Global Natural and Planted Forests Mapping at Fine Spatial Resolution of 30 m. Journal of Remote Sensing 4: 0204.

Zelený J, Mercado-Bettín D, Müller F. 2021. Towards the evaluation of regional ecosystem integrity using NDVI, brightness temperature and surface heterogeneity. Science of The Total Environment 796: 148994.

Zerger A, Gibbons P, Seddon J, Briggs S, Freudenberger D. 2009. A method for predicting native vegetation condition at regional scales. Landscape and Urban Planning 91: 65-77.

Zhu Z, Qiu S, Ye S. 2022. Remote sensing of land change: A multifaceted perspective. Remote Sensing of Environment 282: 113266.

zu Ermgassen SOSE, Howard M, Bennun L, Addison PFE, Bull JW, Loveridge R, Pollard E, Starkey M. 2022. Are corporate biodiversity commitments consistent with delivering ‘nature-positive’ outcomes? A review of ‘nature-positive’ definitions, company progress and challenges. Journal of Cleaner Production 379: 134798.

Zurita G, Pe’er G, Bellocq MI, Hansbauer MM. 2012. Edge effects and their influence on habitat suitability calculations: a continuous approach applied to birds of the Atlantic forest. Journal of Applied Ecology 49: 503-512.

TABLES

Table 1. Summary statistics for the uncalibrated HCAS version 2.3 score in each of the areas shown in Figure 5.

Dataset	Minimum	First quartile	Median	Mean	Third quartile	Maximum
Relatively natural areas (inferred reference sites)	0.00000	0.01461	0.01535	0.01506	0.01589	0.01900
Highly modified areas (intensive land use)	0.00001	0.00617	0.01049	0.00939	0.01284	0.01869

Table 2. Ecosystem condition trend: extent (hectares) and condition-weighted extent (in condition-hectares) in each ecosystem condition trend class by catchments, 2001-02 to 2018-19, for the Flinders, Norman and Gilbert river catchments in Queensland (FNG). Source: adated from Table 'FNG_HCAS23LCEX_EC_RCA_S02' in Giljohann et al. (2023). Derived from HCAS version 2.3 (Harwood et al. 2023a).

Catchments and coastal	**Ecosystem condition trend class**	**Extent (hectares)**	**Average ecosystem condition index change per year (trend slope coefficient)**	**Total change in condition-weighted extent (condition-hectares)**
Flinders river catchment area	significant increase (> 0)	186,968	0.0025	8,329
	non-significant (= 0)	9,847,448		
	significant decrease (< 0)	912,804	0.0035	57,862
	unclassified	3,743		
Gilbert river catchment area	significant increase (> 0)	137,023	0.0020	4,848
	non-significant (= 0)	4,070,130		
	significant decrease (< 0)	433,635	0.0032	25,015
	unclassified	274		
Norman river	significant increase (> 0)	98,761	0.0020	3,538

Catchments and coastal	**Ecosystem condition trend class**	**Extent (hectares)**	**Average ecosystem condition index change per year (trend slope coefficient)**	**Total change in condition-weighted extent (condition-hectares)**
catchment area	non-significant (= 0)	4,507,252		
	significant decrease (< 0)	437,078	0.0028	22,078
	unclassified	1,332		
Subtotal	significant increase (> 0)	422,752	0.0065	16,715
	non-significant (= 0)	18,424,830		
	significant decrease (< 0)	1,783,517	-0.0095	-104,955
	unclassified	5,349		
Flinders, Norman and Gilbert river catchments		20,636,448	**NA**	**NA**

Box 1. Evaluation criteria listed by Hansen et al. (2021) for systematically monitoring and evaluating trends in ecosystem integrity.

1. A direct measure of a specific aspect of ecosystem structure, function, or composition.
2. Biome to global extent with spatial resolution sufficiently fine to allow for management relevance and subnational assessment (≤1 km).
3. Temporal resolution to allow assessment at annual to 5-year periods.
4. Ability of the indicator to be aggregated from subnational to national to global without introducing bias.
5. Known credibility through validation and peer review, data and metadata are publicly available, adheres to open data standards.
6. Potential to be referenced to states characteristic of the climatic, geomorphic, and native community ecosystem.

FIGURES

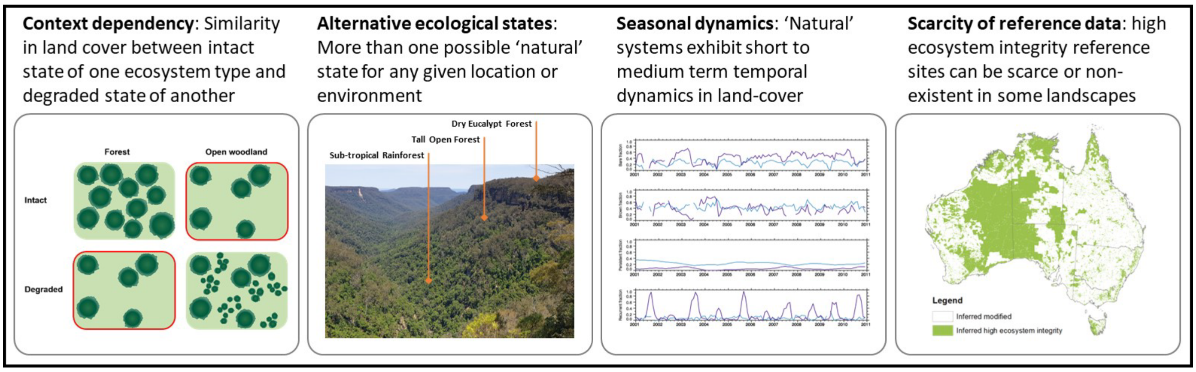


Figure 1. An illustration of three of the four main types of ecological application challenges inherent to the use of satellite remote sensing for estimating ecosystem condition. Adapted from Harwood et al. (2016).

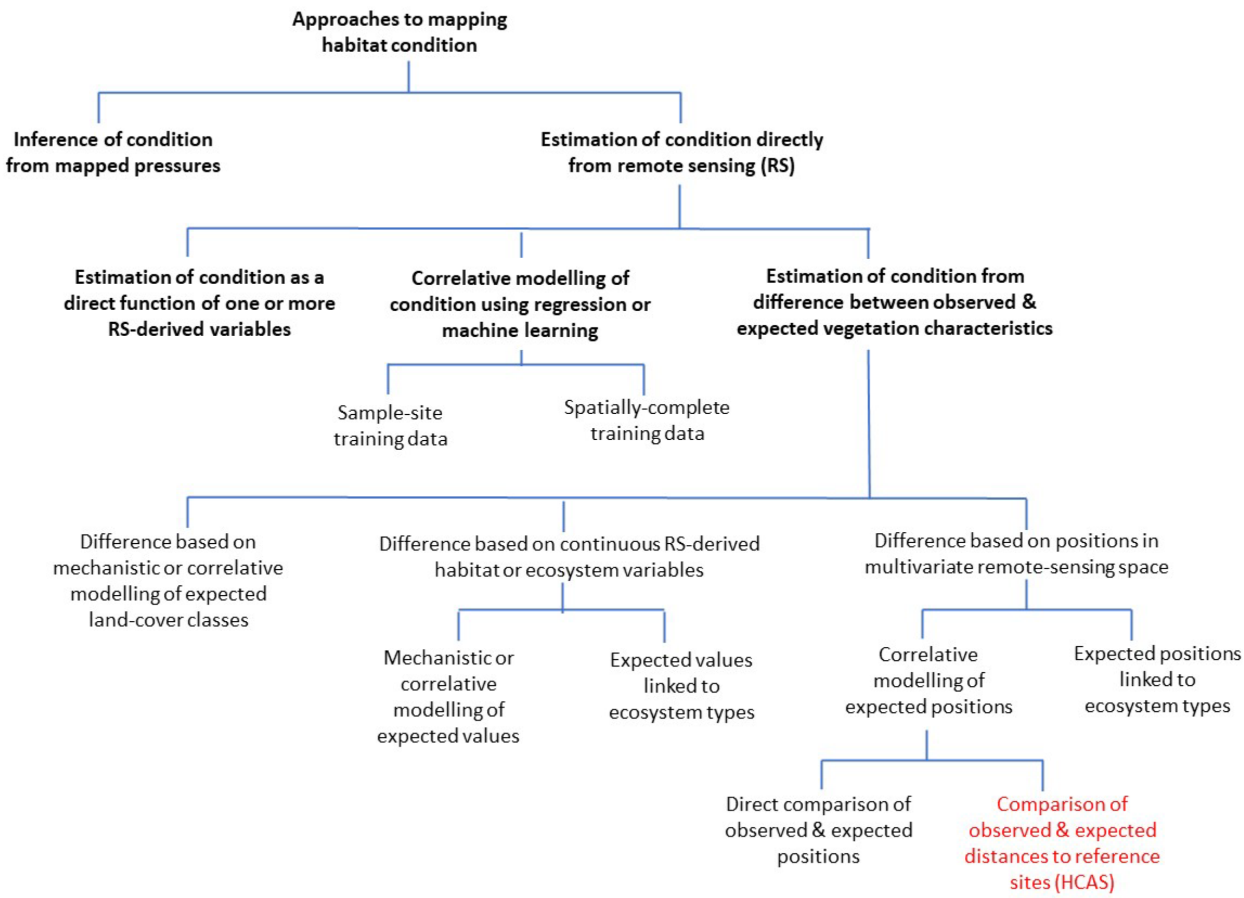


Figure 2. Typology of approaches to mapping habitat condition across large spatial extents using satellite-based remote sensing.

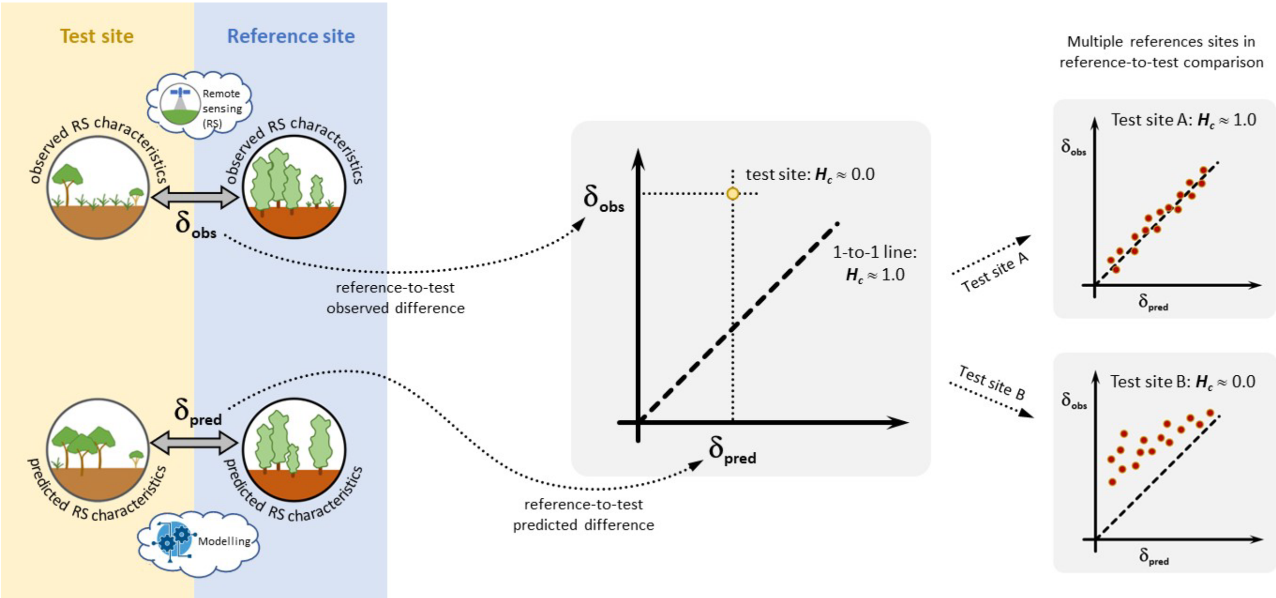


Figure 3. Habitat Condition Assessment System (HCAS) mechanics – an overall schematic of how ecosystem condition is benchmarked using reference sites, showing the case where the test site condition is closer to 1.0 or 0.0.

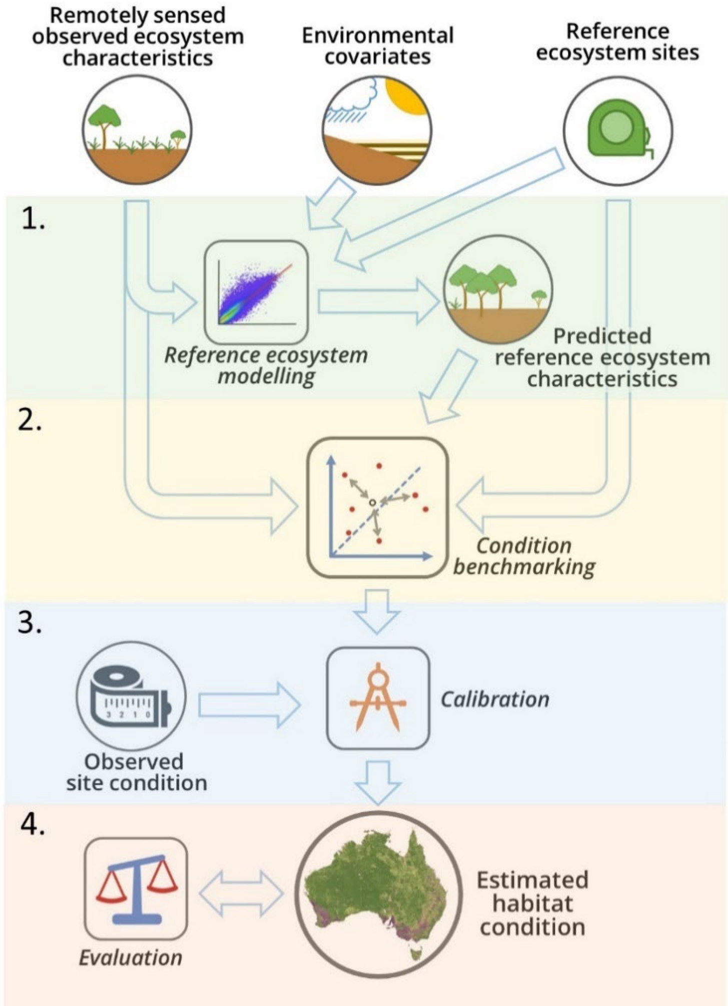


Figure 4. Summary of HCAS model workflow structure. The workflow hinges on two main processing stages (shown as steps 1 and 2). First, a multivariate regression model is developed (labelled 'Reference ecosystem modelling') to predict ecosystem characteristics (using satellite-observed remotely sensed ecosystem characteristics) from a set of predictors (environmental covariates such as climate, soil, landform and hydrology) for sites in reference condition (having high levels of ecosystem integrity). The reference ecosystem model is used to predict ecosystem characteristics at every site of interest. The second stage (labelled 'condition benchmarking') calculates differences between predicted and observed ecosystem characteristics, and uses sites in reference condition (this time as 'benchmarks') to derive the initial uncalibrated habitat condition index, indicating similarity to reference conditions for every test site. Subsequent steps calibrate and standardise the estimates to values between 0.0 and 1.0, and compares results with other land information datasets to inform interpretation and use.

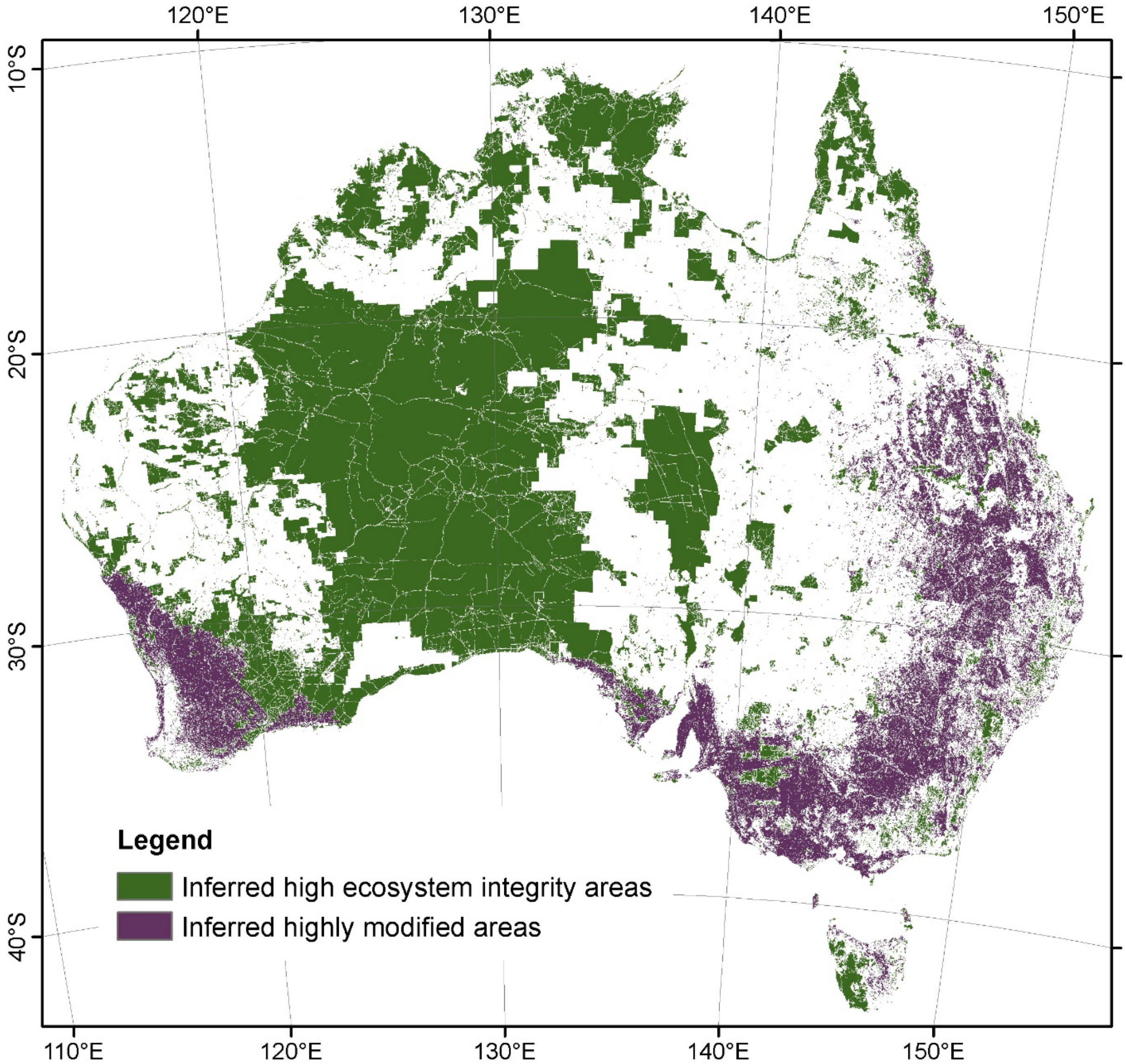


Figure 5. Distribution of inferred high ecosystem integrity areas (i.e., reference sites) and inferred highly modified areas (potentially removed ecosystems) used in HCAS version 2.3 benchmarking and scaling algorithms. White areas are intermediate modified areas. Projection: Australian Albers GDA 1994.

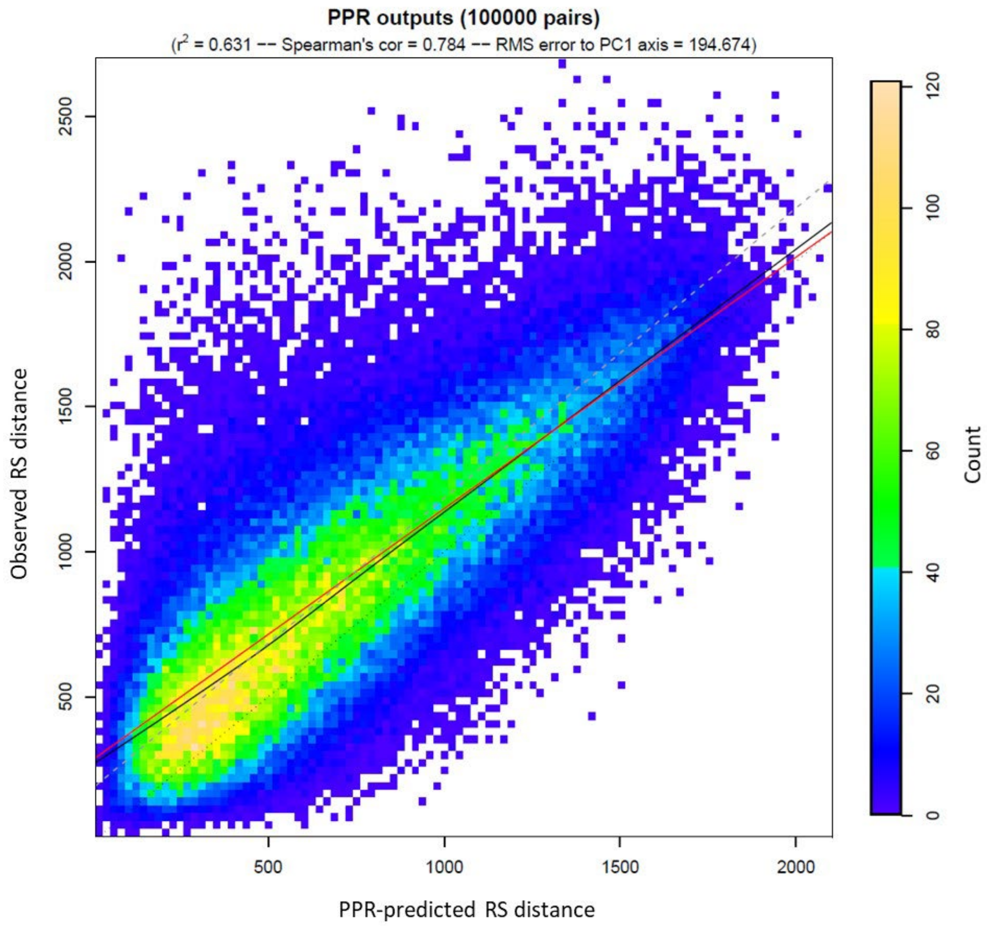


Figure 6. Projection Pursuit Regression model fit in terms of observed versus predicted remote sensing principal component Euclidean distances used in HCAS version 2.3. A random sample of 100,000 reference site-pairs (of the N×(N-1)/2 combinations, N = 101,686) are used for computational tractability. Red line is a linear model fit of the data; black line is a smoothing fit of the data; dashed grey line is the diagonal.

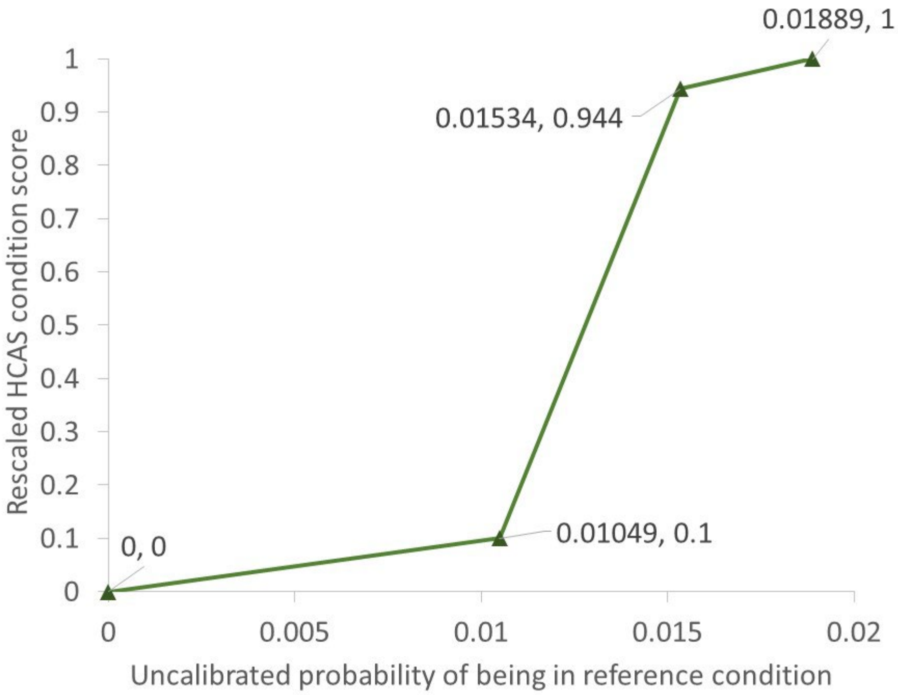


Figure 7. Piecewise linear rescaling coordinates used in HCAS version 2.3 to derive a calibrated and standardised index ranging from 0.0 (removed) to 1.0 (reference condition). The two inflection points are for the respective medians of uncalibrated scores in highly modified areas (left) versus high ecosystem integrity areas (right); areas as mapped in Figure 5.

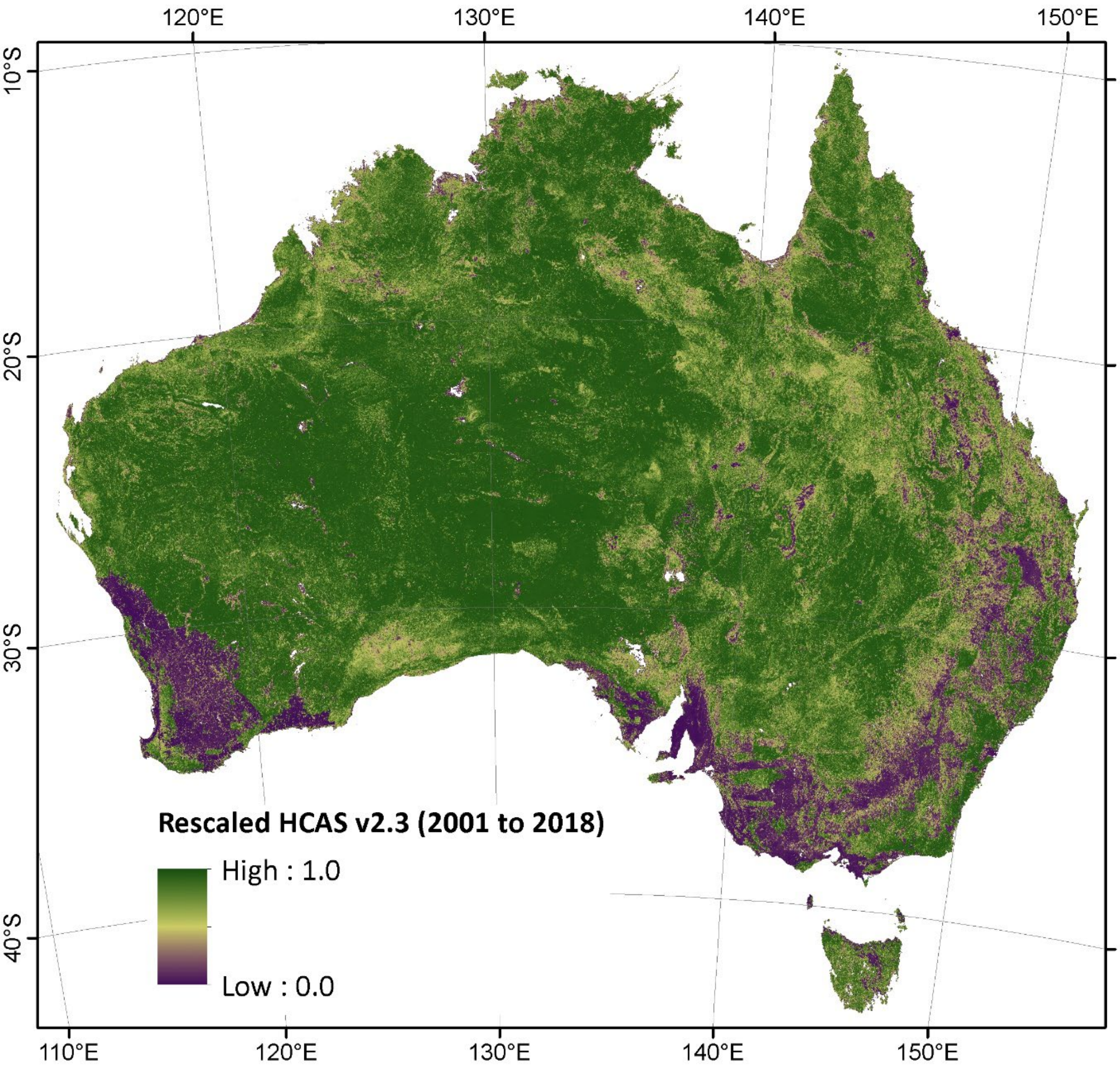
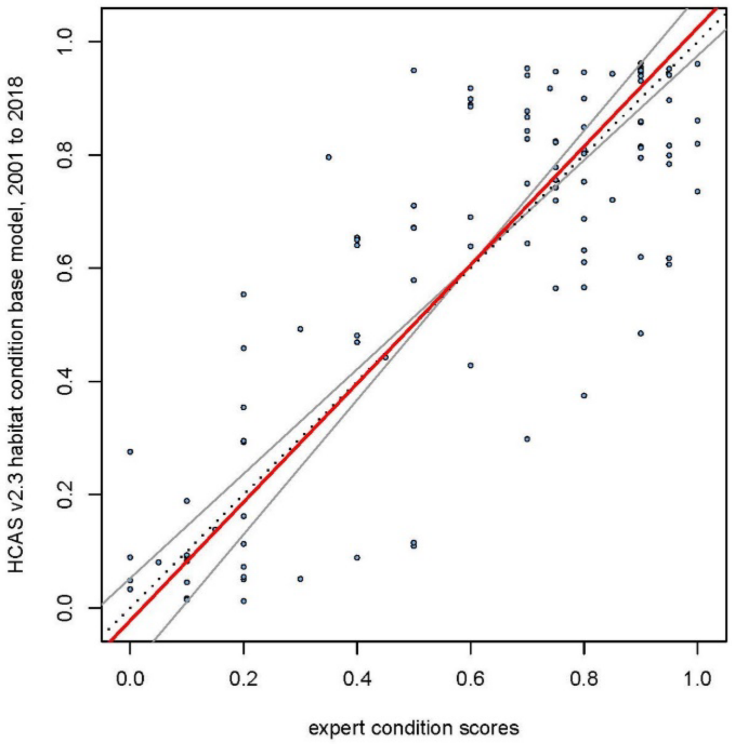


Figure 8. Calibrated HCAS version 2.3 for the base model (2001-2018). Data: Harwood et al. (2023a). Projection: Australian Albers, GDA 1994.



n = 111
Intercept = −0.02 [−0.11 ... 0.05]
Slope = 1.05 [0.92 ... 1.19]
Angle = 46.32 deg.
RMSOE = 0.13

Figure 9. Type II regressions between HCAS version 2.3 habitat condition and expert condition scores from eleven virtual transects. The 'Intercept' results are the estimated intercept using the Type II regression (with confidence interval, CI, range); the 'Slope' results are the estimated slope coefficient (with CI) – best when closest to 1.0; the 'Angle' result is the estimated angle of the fitted line (best when closest to 45°); RMSOE is the "bespoke" orthogonal RMSE between the data points and the Type II regression line ('bespoke' in the sense that it is not really a standard metric of modelling error, but provides some insight into the "orthogonal" variability of the data points from the regression line, i.e., in the spirit of the Type II analysis).

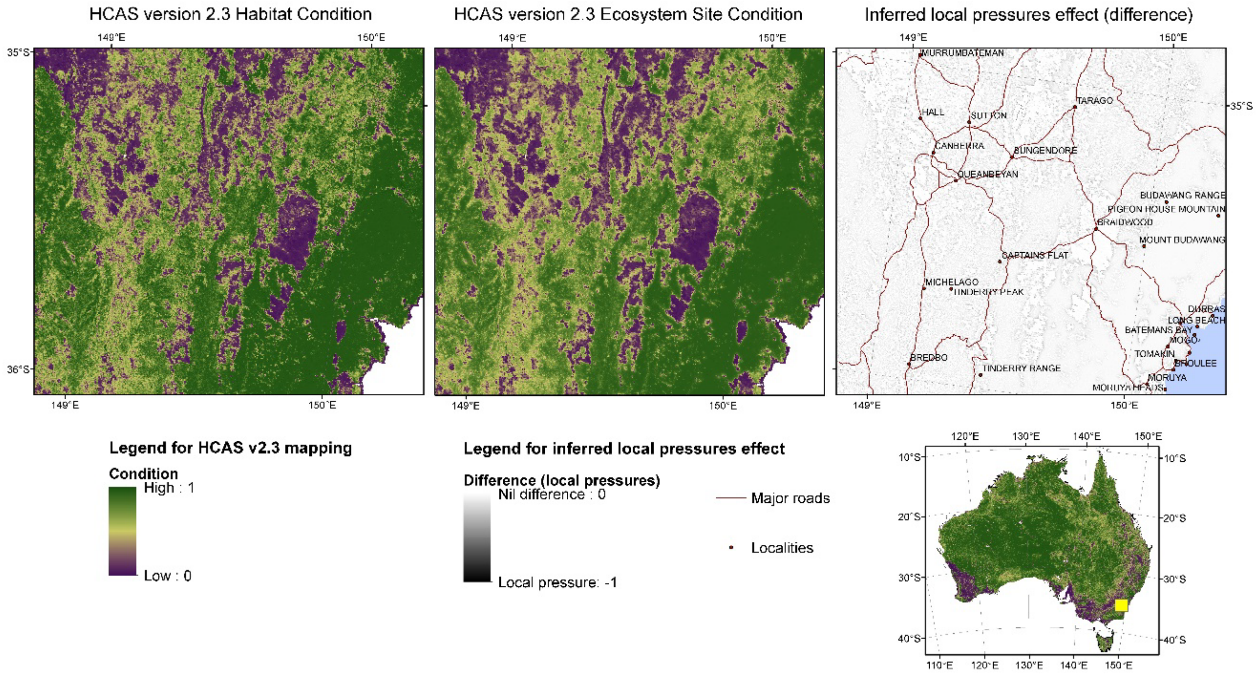


Figure 10. Local comparison (Canberra region of Australia – see inset for location) showing the inferred local pressurs effect (top-right) as the difference between HCAS version 2.3 habitat condition (far-left) and ecosystem site condition (middle) for the 2001-2018 long term epoch. Data: Harwood et al. (2023a). Projection: Australian Albers, GDA 1994.

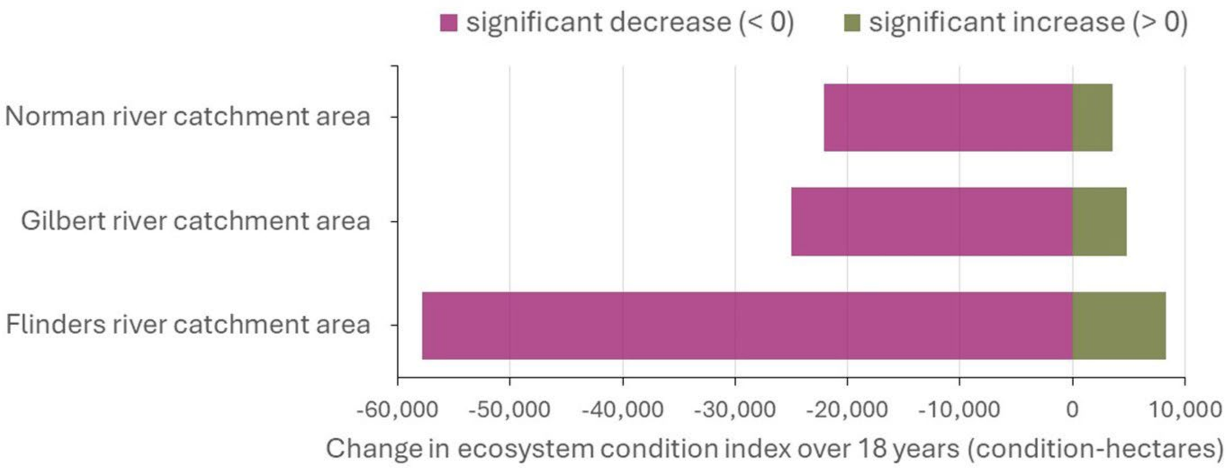


Figure 11. Change in ecosystem condition over 18 years by catchments, 2001-02 to 2018-19, for the Flinders, Norman and Gilbert river catchments in Queensland (FNG). Source: Table 'FNG_HCAS23LCEX_EC_RCA_S02' in Giljohann et al. (2023). Derived from HCAS version 2.3 (Harwood et al. 2023a).

SUPPLEMENTAL MATERIALS

Supplemental Material A – Common use terms with similar meaning to ecosystem condition and integrity

Supplemental Material B – A technical comparison of HCAS versions

Supplemental Material C – Methods used in developing HCAS version 2.3: A continental scale example from Australia

BIOGRAPHICAL NARRATIVE

Kristen J Williams (Kristen.williams@csiro.au), Simon Ferrier, Eric A Lehmann, Thomas D Harwood, Randall J Donohue, Kathryn M Giljohann, Roozbeh Valavi, Ning Liu, Chris Ware, Thomas G Van Niel, Tim R McVicar and Anna E Richards are affiliated with the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia. Within CSIRO, all except Eric A Lehmann (Data61 Research Unit) are affiliated with the Environment Research Unit. Kristen J Williams, Simon Ferrier, Eric A Lehmann, Thomas D Harwood, Randall J Donohue and Timothy R McVicar are located in Canberra, Australian Capital Territory; Kate M Giljohann and Roozbeh Valavi are located in Melbourne, Victoria; Chris Ware is located in Hobart, Tasmania; Thomas G Van Niel is located in Perth, Western Australia; and Anna E Richards is located in Darwin, Northern Territory. Peter Lyon and Cassandra Malley are affiliated with Environment Information Australia, Australian Government Department of Climate Change, Energy, the Environment and Water (DCCEEW), in Canberra, Australian Capital Territory, Australia. Thomas D Harwood's current affiliation is the Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom.

ACKNOWLEDGMENTS

The research reported here was conducted over several projects between 2015 and 2023 in partnership and with funding through the Australian Government Department of Climate Change, Energy, the Environment and Water. Ethics clearance for reuse of data obtained through expert elicitation via CSIRO Social Science and Human Research Ethics Committee, applications 048/21, 007/21, 196/23 and 197/23 (individual acknowledgments in Supplemental Material C). We thank State agency officers of South Australia, Tasmania, Victoria and Queensland for facilitating access to systematic field surveys of habitat condition. We gratefully acknowledge # anonymous peer reviewers whose insights improved an earlier version of this manuscript.

DATA AVAILABILITY

HCAS data collections have been lodged with data.csiro.au.

Supplemental Material A – common use terms with similar meaning to ecosystem condition or integrity

This document provides supplemental material for the manuscript: *Overcoming Key Challenges of Satellite-based Monitoring of Ecosystem Condition: A Continental-scale Example From Australia*

Table S1. Example terms in common use with a similar meaning or intent as for ecosystem condition or integrity, and that are generally applicable across terrestrial, freshwater and marine realms

Term (source)	**Purpose**	**Definition**	**Reference or baseline concept**
Condition, Ecological (Jakobsson et al. 2020, Jakobsson et al. 2021)	Environmental Reporting	The state and trends of structures and functions (incl. productivity) in an ecosystem; as characterised by key aspects of the biodiversity, structure, and functioning of the ecosystem.	Intact ecosystems, understood as nature not significantly affected by human-driven pressures in the industrial era, characterized by recent historical biodiversity and normal climates (1961–1990).
Condition, Ecological (Stoddard et al. 2006, USEPA 2022)	Environmental Reporting	The state of ecological systems, which includes their physical, chemical, and biological characteristics and the processes and interactions that connect them.	Intact ecosystems with respect to recent natural or semi-natural biodiversity and ecosystem functioning.
Condition, Ecosystem (Czúcz et al. 2021, United Nations et al. 2021)	Ecosystem accounting	The quality of an ecosystem measured in terms of its abiotic and biotic characteristics.	The condition against which past, present and future ecosystem condition is compared to in order to measure relative change over time.
Condition, Ecosystem (Keith et al. 2020)	Ecosystem accounting	The quality of an ecosystem that may reflect multiple values, measured in terms of its abiotic and biotic characteristics across a range of temporal and spatial scales.	The natural state of intact native ecosystems, in terms of ecosystem characteristics at their natural condition, allowing for dynamic ranges.
Condition, Ecosystem (Rendon et al. 2019)	Ecosystem services	The overall quality of an ecosystem unit, in terms of its biological, physical	An ecosystem unit at its maximum capacity

Term (source)	Purpose	Definition	Reference or baseline concept
		and chemical characteristics underpinning its capacity to generate ecosystem services	to generate ecosystem services.
Condition, Habitat (Harwood et al. 2016)	Biodiversity persistence	A measure of the difference between two sets of dynamic ecological states: one resulting from the natural regime of disturbance and recovery processes; and the other consisting of modified states resulting from anthropogenic perturbations.	The dynamic ecological states resulting from the natural regime of disturbance and recovery processes, and that vary continuously along environmental gradients.
Condition, Vegetation (Gibbons and Freudenberger 2006, Gibbons et al. 2008)	Native vegetation management	Based primarily on structure and/or composition in which reference conditions (relatively unmodified sites) are often used as the benchmark for assessment	Variation in native vegetation exhibiting relatively little evidence of modification by humans since European settlement (pre-industrial era in Australia)
Effectiveness, Habitat (https://www.lawinsider.com/dictionary/habitat-effectiveness)	Species persistence	The degree to which a habitat or its components fulfill specific habitat functions; the degree to which a species or population is able to continue using a habitat for a specific function.	A maximum potential habitat function for a given individual or population of a species
Health, Ecosystem (Andel and Aronson 2006, Andel et al. 2012, Society for Ecological Restoration International Science and Policy Working Group 2004)	Sustainability management	The state or condition of an ecosystem in which its dynamic attributes are expressed within the normal ranges of activity relative to its ecological stage of development	An ecological development stage
Health, Ecosystem (Costanza R. 1992, Costanza Robert and Mageau 1999)	Sustainability management	An ecological system is healthy and free from 'distress syndrome' if it is stable and sustainable – that is, if it is active and maintains its organization and autonomy over time and is resilient to stress; based on a system's characteristic levels of	A healthy system is one that possesses adequate resilience, vigour, and organization, to survive various small-scale perturbations

Term (source)	Purpose	Definition	Reference or baseline concept
		vigour, organization, and resilience.	
Health, Ecosystem (Lausch et al. 2018, Rapport et al. 1998)	Sustainability management	Vigorous, diverse systems that are characterized by a high resilience, that is, the ability to quickly return to an initial state following an external disturbance and thus to withstand negative impacts from external influences.	An initial state prior to external disturbance
Intactness, Biodiversity (Scholes and Biggs 2005) (Hudson et al. 2017, Newbold et al. 2016)	Biodiversity persistence	The proportion of the original number of species that remain and their abundance in any given area, despite human impacts.	The number and diversity of species at near-undisturbed sites.
integrity, Ecological (Mansourian 2005, Wurtzebach and Schultz 2016)	Forest landscape restoration	To maintain the diversity and quality of ecosystems, and enhancing their capacity to adapt to change and provide for the needs of future generations.	Capacity to maintain natural ecological and evolutionary processes
Integrity, Ecological (McGarigal et al. 2018)	Biodiversity persistence and ecosystem function	The ability of an area to support native biodiversity and the ecosystem processes necessary to sustain that biodiversity over the long term; and accommodates the modification or adaptation of systems (in terms of biotic composition and structure) over time to changing environments.	The ecological functions necessary to confer ecological integrity, using quantile scaling to rate sites relative to each other within a given region.
Integrity, Ecological (Parrish et al. 2003)	Protected area management	The ability of an ecological system to support and maintain a community of organisms that has species composition, diversity, and functional organization comparable to those of natural habitats within a region.	Dominant ecological characteristics (e.g., elements of composition, structure, function, and ecological processes) occur within their natural ranges of variation and can withstand and recover from most perturbations imposed

Term (source)	Purpose	Definition	Reference or baseline concept
			by natural environmental dynamics or human disruptions.
Integrity, Ecosystem (SER 2002, Society for Ecological Restoration International Science and Policy Working Group 2004)	Ecological restoration and management	The state or condition of an ecosystem that displays the biodiversity characteristic of the reference, such as species composition and community structure, and is fully capable of sustaining normal ecosystem functioning	The state of an ecosystem that is fully capable of sustaining normal ecosystem functioning
Integrity, Ecosystem (United Nations et al. 2021, WCS 2021)	Ecosystem accounting	The ecosystem's capacity to maintain its characteristic composition, structure, functioning and self-organisation over time within a natural range of variability	Natural or historic range of variability in composition, structure and function
Integrity, Habitat (Thompson 2018)	Biodiversity persistence	The capacity of a place to support indigenous species with the resources necessary to complete their life cycle.	Resources necessary for indigenous species to complete life cycles.
Integrity, Landscape (Perkl 2017)	Landscape management	A measure of the landscape's naturalness, or its inverse, the level of human modification. (Related to the human footprint approach)	The 'standard' or 'natural' baseline of the landscape. (Related to ecological integrity.)
Naturalness (Dengler et al. 2008, Machado 2004)	Conservation value	Naturalness, or its reciprocal concept, hemeroby, ranks communities by the strength of human influence and consequent alterations of species composition, structure, and ecological processes. (Related to wilderness.)	Maximum state of naturalness wherein all ecological components and processes are present and natural (native and intact), without human influence.
Naturalness, Ecological (Dussault 2016)	Sustainability management	The ecological normality allowing for the ability of a species to live in accordance with Callicott's principle of harmony with nature. (Related to ecosystem health).	The ecological normality of a region or place.

Term (source)	Purpose	Definition	Reference or baseline concept
Quality, Ecological (Shaoqiang et al. 2019)	Regional biodiversity and ecosystem function monitoring	The stability, adaptability and resilience of an ecosystem. It is the comprehensive sum of the characteristics of ecosystem elements (a.k.a. composition), structures and functions within a certain time and space, embodying the status, production capacity, structural and functional stability, adaptability and resilience of ecosystems.	Threshold criteria related to ecosystem functions for regulating, supporting, and maintaining biodiversity
Quality, Ecological (Rina et al. 2019)	Monitoring the stability of dryland ecosystems for sustainable development	The comprehensive characteristics of the structures and functions of the ecosystem within a certain spatial–temporal range.	Threshold criteria
Quality, Ecosystem (Verones et al. 2020, Woods et al. 2018)	Life cycle assessment	The area of protection that accounts for impacts on the natural environment. The endpoint unit used here is potentially disappeared fraction of species (PDF). This metric accounts for a fraction of species richness that may be potentially lost due to an environmental mechanism.	The current situation, relating the change either to a zero effect, a preferred state (e.g., environmental targets) or a prospective future state.
Quality, Habitat (Hall et al. 1997)	Species persistence	The ability of the ecosystem to provide conditions appropriate for individual and population (wildlife) persistence.	Ecosystem conditions appropriate for individual and population persistence.
Quality, Habitat (Johnson 2007, Zlinszky et al. 2015)	Biodiversity persistence	The ability of the environment to provide conditions appropriate for individual and species population persistence.	Environmental conditions appropriate for individual and species population persistence
Quality, Vegetation (Parkes et al. 2003)	Biodiversity persistence	The degree to which the current vegetation differs from mature and apparently long-	Average characteristics of a mature and apparently long-undisturbed

Term (source)	Purpose	Definition	Reference or baseline concept
		undisturbed stands of the same vegetation community.	stand of the same vegetation community.
Resilience, Ecosystem (Costanza Robert and Mageau 1999)	Sustainability management	Ability of a system to maintain its structure and pattern of behaviour in the presence of stress. (A component of ecosystem health.)	A resilient ecosystem possesses adequate vigour, and organization, to survive various small-scale perturbations

References

Andel Jv, Aronson J, eds. 2006. Restoration Ecology: The New Frontier. First ed. Oxford, UK: Blackwell Publishing.

Andel Jv, Grootjans AP, Aronson J. 2012. Chapter 2: Unifying Concepts. Pages 9-22 in Andel Jv, Aronson J, eds. Restoration Ecology: The New Frontier, Second version. Oxford, UK: Wiley-Blackwell.

Costanza R. 1992. Toward an operational definition of health. Pages 239–256 in Costanza R, Norton B, Haskell B, eds. Ecosystem Health: New Goals for Environmental Management. Washington DC, USA: Island Press.

Costanza R, Mageau M. 1999. What is a healthy ecosystem? Aquatic Ecology 33: 105-115.

Czúcz B, Keith H, Driver A, Jackson B, Nicholson E, Maes J. 2021. A common typology for ecosystem characteristics and ecosystem condition variables. One Ecosystem 6.

Dengler J, Chytrý M, Ewald J. 2008. Phytosociology. Pages 516-527 in Fath B, ed. Encyclopedia of Ecology (Second Edition). Oxford: Elsevier.

Dussault AC. 2016. Ecological Nature: A Non-Dualistic Concept for Rethinking Humankind's Place in the World. Ethics and the Environment 21: 1-37.

Gibbons P, Freudenberger D. 2006. An overview of methods used to assess vegetation condition at the scale of the site. Ecological Management & Restoration 7: S10-S17.

Gibbons P, Briggs SV, Ayers DA, Doyle S, Seddon J, McElhinny C, Jones N, Sims R, Doody JS. 2008. Rapidly quantifying reference conditions in modified landscapes. Biological Conservation 141: 2483-2493.

Hall LS, Krausman PR, Morrison ML. 1997. The habitat concept and a plea for standard terminology. Wildlife Society Bulletin (1973-2006) 25: 173-182.

Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G, White M. 2016. Habitat Condition Assessment System: A new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. Methods in Ecology and Evolution 7: 1050-1059.

Hudson LN, et al. 2017. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecology and Evolution 7: 145-188.

Jakobsson S, et al. 2020. Setting reference levels and limits for good ecological condition in terrestrial ecosystems – Insights from a case study based on the IBECA approach. Ecological Indicators 116: 106492.

Jakobsson S, et al. 2021. Introducing the index-based ecological condition assessment framework (IBECA). Ecological Indicators 124: 107252.

Johnson MD. 2007. Measuring habitat quality: A review. The Condor 109: 489-504.

Keith H, Czúcz B, Jackson B, Driver A, Nicholson E, Maes J. 2020. A conceptual framework and practical structure for implementing ecosystem condition accounts. One Ecosystem 5.

Lausch A, Bastian O, Klotz S, Leitão PJ, Jung A, Rocchini D, Schaepman ME, Skidmore AK, Tischendorf L, Knapp S. 2018. Understanding and assessing vegetation health by in situ species and remote-sensing approaches. Methods in Ecology and Evolution 9: 1799-1809.

Machado A. 2004. An index of naturalness. Journal for Nature Conservation 12: 95-110.

Mansourian S. 2005. Overview of Forest Restoration Strategies and Terms. Pages 8-13. Forest Restoration in Landscapes: Beyond Planting Trees. New York, NY: Springer New York.

McGarigal K, Compton BW, Plunkett EB, DeLuca WV, Grand J, Ene E, Jackson SD. 2018. A landscape index of ecological integrity to inform landscape conservation. Landscape Ecology 33: 1029-1048.

Newbold T, et al. 2016. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. Science 353: 288-291.

Parkes D, Newell G, Cheal D. 2003. Assessing the quality of native vegetation: The 'habitat hectares' approach. Ecological Management & Restoration 4: S29-S38.

Parrish JD, Braun DP, Unnasch RS. 2003. Are we conserving what we say we are? Measuring ecological integrity within protected areas. BioScience 53: 851-860.

Perkl RM. 2017. Measuring landscape integrity (LI): development of a hybrid methodology for planning applications. Journal of Environmental Planning and Management 60: 92-114.

Rapport DJ, Costanza R, McMichael AJ. 1998. Assessing ecosystem health. Trends in Ecology & Evolution 13: 397-402.

Rendon P, Erhard M, Maes J, Burkhard B. 2019. Analysis of trends in mapping and assessment of ecosystem condition in Europe. Ecosystems and People 15: 156-172.

Rina W, Weiwei C, Yonghua L, Siyao L, Dongfang W, Zhiqing J, Feng W. 2019. The Scientific Conceptual Framework for Ecological Quality of the Dryland Ecosystem: Concepts, Indicators, Monitoring and Assessment. Journal of Resources and Ecology 10: 196-201, 196.

Scholes RJ, Biggs R. 2005. A biodiversity intactness index. Nature 434: 45-49.

SER. 2002. The SER primer on ecological restoration: Society for Ecological Restoration Science and Policy Working Group (SER).

Shaoqiang W, et al. 2019. A National Key R&D Program: Technologies and Guidelines for Monitoring Ecological Quality of Terrestrial Ecosystems in China. Journal of Resources and Ecology 10: 105-111, 107.

Society for Ecological Restoration International Science, Policy Working Group. 2004. The SER International primer on ecological restoration (version 2). www.ser.org & Tucson USA: Society for Ecological Restoration International.

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH. 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications 16: 1267-1276.

Thompson LJ. 2018. Habitat Integrity in Kolb RW, ed. The SAGE Encyclopedia of Business Ethics and Society, SAGE Publications.

United Nations et al. 2021. System of Environmental-Economic Accounting—Ecosystem Accounting (SEEA-EA). White cover publication, pre-edited text subject to official editing. www.unstats.un.org: United Nations Committee of Experts on Environmental-Economic Accounting, United Nations Statistical Division.

USEPA. 2022. Report on the Environment: Ecological Condition. https://www.epa.gov/report-environment/ecological-condition)

Verones F, et al. 2020. LC-IMPACT Version 1.0: A spatially differentiated life cycle impact assessment approach. online, https://lc-impact.eu: LC-Impact project.

WCS. 2021. Frequently Asked Questions: Ecosystem Integrity and International Policy: Wildlife Conservation Society (WCS).

Woods JS, et al. 2018. Ecosystem quality in LCIA: status quo, harmonization, and suggestions for the way forward. Int J Life Cycle Assess 23: 1995-2006.

Wurtzebach Z, Schultz C. 2016. Measuring Ecological Integrity: History, Practical Applications, and Research Opportunities. BioScience 66: 446-457.

Zlinszky A, Heilmeier H, Balzter H, Czúcz B, Pfeifer N. 2015. Remote Sensing and GIS for Habitat Quality Monitoring: New Approaches and Future Research. Remote Sensing 7: 7987-7994.

Supplemental Material B – A technical comparison of HCAS versions

This document provides supplemental material for the manuscript: *Overcoming Key Challenges of Satellite-based Monitoring of Ecosystem Condition: A Continental-scale Example From Australia*

From **Supplemental Material A**: *common use terms with similar meaning to ecosystem condition or integrity*

- Table S1. Example terms in common use with a similar meaning or intent as for ecosystem condition or integrity, and that are generally applicable across terrestrial, freshwater and marine realms.

Narrative summary of the comparisons

Implementation of the Habitat Condition Assessment System (HCAS) methodology has evolved since the proof of concept (HCAS v1.0) was published by Harwood et al. (2016). Detailed technical documentation is provided in a series of reports associated with each successive version. A summary of major changes and enhancements is provided in Table S2.

For details about HCAS v1.0 we refer readers to Donohue et al. (2013) and Harwood et al. (2016) and supplementary material provided with Harwood et al. (2016).

For details about HCAS v2.0, see Williams et al. (2020); for HCAS v2.1, see Williams et al. (2021b); for HCAS v2.2, see Williams et al. (2023a); and for HCAS v2.3, see Williams et al. (2023b).

HCAS v1.0

The HCAS v1.0 was a proof of concept at 1 km grid resolution to develop and test an implementation based on the conceptual framework outlined in Donohue et al. (2013). The conceptual framework identified two types of reference sites – one set being the most intact sites with high levels of ecosystem integrity enduring in the landscape, and the other set being the most modified and removed ecosystems without capacity to provide supporting habitat for the original biodiversity to persist. The proof of concept, however, was pragmatically framed solely around the most intact sites of high ecosystem integrity, using the contemporary boundaries of Australia's national reserve system to delineate those places.

The generalised dissimilarity modelling (GDM) method was used to predict patterns of compositional turnover in Australia's ecosystems based on remotely sensed characteristics (combinations of MODIS and AVHRR) to represent the reference state of the continent's ecosystem distribution and diversity. The method used nine of 15 principal components of remotely sensed variables. A random sample of one million reference site-pairs for the GDM response variable were derived using ecological regions (Australia's bioregions) as strata, weighted so that 75% were between region site-pairs and the remainder within region. The GDM approach enforces monotonic relationships between the response variable and predictors (15 environmental covariates depicting long-term stable patterns of climate, soil and landform) so that ecological dissimilarity increases with environmental distance.

Any of the assumed high ecosystem integrity reference sites as delineated by the reserve system, after excluding property boundary potential edge effects, were available for selection as benchmarks. The GDM predictions provided a basis for selecting those reference sites to be used as benchmarks that are most ecologically similar to a test site. Manhattan distances between reference site pairs and test-reference site pairs for the observed and predicted sets of remotely sensed ecosystem characteristics, were used in the calculation of condition, as the average of the likelihood of the test site being in reference condition weighted by ecological similarity as defined by the 20 selected test-reference predicted distances (for benchmarking condition). The output was scaled in the range 0 (removed) to 1 (intact) using the average from major urban centres to set the 0 end point and then linearly by the maximum value, but enforcing a value of 1 for all inferred reference sites (aligned with protected area boundaries) in the output.

HCAS v2.0

The HCAS v2.0 was developed at 250 m grid resolution as an experimental implementation to test the suitability of the method for operational use cases. The method built on and improved the HCAS v1.0 mechanics of Harwood et al. (2016). Reference sites were inferred to be the most intact regions where ecosystems persisted in their most natural state, including undeveloped lands as well as protected areas, using multiple lines of evidence. Remotely sensed ecosystem characteristics derived from two MODIS fractional cover products at 500 m and 250 m grid resolution and AVHRR at 1 km, oversampled to match the 250 m grid. Projection pursuit regression (PPR) models replaced GDM as the method for predicting the remotely sensed reference state (all principal components) of intact ecosystems based on a wider range of environmental covariates, without imposing monotonicity.

Intact reference sites were randomly sampled across two strata representing largely intact or largely modified ecosystems weighted toward a much larger number of samples from modified regions. This sample was also used as benchmarks in the condition calculation. As for HCAS v1.0, two-sets of Manhattan distances (observed vs. predicted) were derived for each of reference-reference site pairs and test-reference site pairs. A half-Cauchy weighting on the similarity of test-reference site distances was introduced to the benchmarking calculation for estimating condition, and a geographic distance limit around each test site was introduced to guide appropriate selection of the final 20 reference sites to use as benchmarks. The output was linearly scaled by the maximum value to range between 0 (removed) and 1 (intact).

HCAS v2.1

The HCAS v2.1 was developed at 250 m grid resolution to improve upon HCAS v2.0 and provide a change assessment. The inferred reference sites from HCAS v2.0 were randomly sampled from within each of nearly 5000 ecological land units to derive c.100,000 subsamples to use as training data and the process repeated to derive c.200,000 samples to use as benchmarks. Remotely sensed ecosystem characteristics were derived from two MODIS fractional cover products at 500 m and 250 m grid resolution, but imagery was first filtered to remove water and snow pixels.

As for HCAS v2.0, PPR models were used to predict the remotely sensed reference state (all principal components) of intact ecosystems. The same basic method of estimating condition as for HCAS v2.0 was used, except a limited degrees of confidence (LDC) calculation was introduced to account for potential invalid reference sites in the test-benchmark site-pairs. LDC adds an additional weight to the single test-reference site-pair found to have the highest probability of being in reference condition.

A piecewise linear rescaling algorithm with two inflection points was introduced in the calibration step to simulate non-linearity between the output of the HCAS algorithm and expected values ranging from 0 to 1. Coordinates for the inflection points were determined using highly modified land use and relatively natural areas, respectively, and empirical data from global studies of biodiversity intactness in different land uses. A time series of condition was derived from the lineage of annual remote sensing variables using the same benchmarking and scaling algorithms by substituting the observed long-term with annual remote sensing PCs in the selection of benchmarks and test-benchmark comparisons.

HCAS v2.2

The HCAS v2.2 derives from HCAS v2.1 and varies only in minor updates to the reference sites and scaling algorithms. Data sources used in the multiple lines of evidence approach to inferring reference sites were updated to improve currency and extend sources of data used to exclude potentially modified areas. The stratified random subsampling of reference sites used as benchmarks (c. 200,000) was repeated and extended to include palustrine wetlands and salt lakes. The source data and method used to define coordinates for the two inflection points of the scaling algorithm was slightly revised. Annual epochs were derived as for HCAS v2.1.

HCAS v2.3

The HCAS v2.3 derives from HCAS v2.1 and builds upon the HCAS v2.2 improvements. The inferred reference sites were updated to include expert nominated inclusions and exclusions, and additional data on potentially modified areas used to exclude areas. The stratified random subsampling of reference sites used as benchmarks (c. 200,000) was repeated as for v2.2. The scaling algorithm followed the method developed for v2.2, except the revised set of inferred reference sites served as the extent of relatively natural areas. Annual epochs were derived as for HCAS v2.1.

109

110 Table S2. A technical summary comparing implementations of the Habitat Condition Assessment System (HCAS) methodology from proof of concept,
111 HCAS v1.0, to operational, HCAS v2.3.

Workflow component	**HCAS v1.0**	**HCAS v2.0**	**HCAS v2.1**	**HCAS v2.2**	**HCAS v2.3**
Primary purpose	Proof of concept	Experimental implementation for trialling in operational use cases	Publishable implementation suitable for use in State of the Environment reporting and a time-series to show change	Ecosystem accounting using annual time-series	Ecosystem accounting using annual time-series
Spatial resolution	0.01 degree GDA94 (approx. 1 x 1 km pixels)	9 arc second GDA94 (approx. 250 x 250 m pixels)	9 arc second GDA94 (approx. 250 x 250 m pixels)	9 arc second GDA94 (approx. 250 x 250 m pixels)	9 arc second GDA94 (approx. 250 x 250 m pixels)
Reference sites - inferred	The 'core areas' at 1 km^2 resolution of Australia's nature-based protected areas as of 2010 (DCCEEW 2023a) by eroding raster boundaries to remove edge effects (772,160 reference sites).	Multiple lines of evidence (c.2012 to 2016) combining native vegetation clearing (Department of the Environment 2014), land use other than 'conservation and natural environments' (ABARES 2016a, b), road networks (Geoscape Australia 2020) and settlement patterns (ABS 2014) to exclude all potential modified locations, to identify sites mainly within protected areas as of 2016 (DCCEEW 2023b) or relatively natural areas	As for v2.0.	Updated multiple lines of evidence to exclude all potential modified locations combining latest data on: land use as of 2015-16 (ABARES 2022); roads, railways, infrastructure and other human modified sites identified using Open Street Map (OSM) data, current to 20 April 2022 (OpenStreetMap Contributors 2022, Ramm 2022); and the 2022 update of global-scale mining polygons dataset (Maus et al. 2020, Maus et al. 2022). Inferred sites	As for v2.2 with inclusions and exclusions nominated by experts in each of two pilot regions (Flinders, Norman and Gilbert River catchments in Queensland and the Southwest Australian Wheatbelt). However, experts or inferred reference sites that overlapped with mapped infrastructure (as for v2.2) or mapped road networks (Geoscape Australia 2020) (as for v2.0) buffered by c.250m were removed (resulting in 38,773,526

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
		(Department of the Environment 2014). The output was eroded by c.250 m to remove edge effects and ensure only ‘core areas’ were included (37,046,447 reference sites of total 111,304,074 test sites).		mainly within protected areas as of June 2020 (DAWE 2021) and subsequent additions to Indigenous Protected Areas (DCCEEW 2022a), and remnant native vegetation mapped in NVIS v6.0 extant major vegetation groups (DCCEEW 2023c). Outputs were eroded by c.250 m to remove edge effects (resulting in 39,685,172 reference sites of total 110,936,913 test sites)	reference sites of total 110,936,913 test sites).
Reference sites - training sample	Stratified by IBRA 7.0 regions (Australia’s ecoregions) (DCCEEW 2023d), 1 million randomly sampled reference site-pairs made up of 25% within and 75% between regions, utilising 425,156 reference sites.	Two strata derived from IBRA 7.0 regions (DCCEEW 2023d): (1) relatively intact with ≥ 50% reference site coverage, and (2) relatively fragmented with < 50% reference site coverage. The training data was a random sample of 100,000 sites with 20 times more drawn from relatively fragmented regions.	4961 ecological land units with at least one reference site present defined based on IBRA 7.0 subregions (DCCEEW 2023e), and NVIS present major vegetation sub groups version 5.1 (DAWE 2018a), excluding categories suggesting a water body, salt lake or modified vegetation type, resulting in a population of 35,485,829 reference sites. Up to 25 sites were	As for v2.1.	As for v2.1.

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
			randomly drawn from each stratum to derive a training sample of 101,686 reference sites.		
Reference sites – benchmarking sample	All 772,160 inferred reference sites, filtered by the condition algorithm to derive 20 reference site-pairs for each test site of interest.	Same as training sample.	As for training sample, except up to 55 sites were randomly drawn from each stratum to derive a benchmark sample of 200,278 reference sites, wherein some sites may be the same as the training data due to limited options.	As for v2.1 training sample except the benchmarking stratification used NVIS v6.0 pre-1750 extent of major native vegetation subgroups (DAWE 2020) with IBRA 7.0 subregions (DCCEEW 2023e), instead of the 'present' extent, resulting in 5579 strata with at least one reference site. Up to 50 sites were randomly sampled, and then combined with 28 expert identified reference sites from a previous ecosystem accounting case study (Harwood et al. 2021a, Harwood et al. 2021b), resulting in a total of 208,856 reference sites as benchmarks.	As for v2.2, repeated using the updated inferred reference sites, selecting up to 50 sites within the 5481 strata containing at least one site, resulting in a total of 202,515 reference sites as benchmarks.
Environmental covariates	15 predictors (5 climate, 8 soil, 2 landform) as listed in Table 1 of Donohue et al. (2013).	21 predictors (9 climate, 10 soil, 1 landform, 1 surface water) as listed in	23 predictors (9 climate, 11 soil, 2 landform, 1 surface water) as listed in Table 7 of Williams et al.	As for v2.1	As for v2.1

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
		table 5 of Williams et al. (2020).	(2021b). Same candidates as for v2.0, except a MODIS-derived, alpha-NDVI water algorithm (Donohue et al. 2022) replaced the Water Observations from Space equivalent (Mueller et al. 2016).		
Remote sensing variables	The first 10 principal components (PCs) of 15 variables from five products. The 11-year averages, 2001-2011, of annual means, maximums and standard deviations from monthly values for (1) bare ground and (2) brown (litter) fractional land cover data derived from the c.500 m MODIS Collection 5 MCD43A4.005 product (Guerschman et al. 2009), plus 10-year averages, 2001-2010, of annual means, maximums and standard deviations from monthly values of c.1 km AVHRR-derived (3) surface albedo (Donohue et al. 2008) and (4)	All principal components of 6 variables from three products. The 16-year averages, 2001-2016, of annual mean and intra-annual range (maximum minus minimum) of monthly values for (1) surface albedo (Donohue et al. 2008) from c.1 km AVHRR and (2) persistent green fractional vegetation cover from c.250 m MODIS collection 5 MOD13Q1 (Donohue et al. 2009), and annual mean and maximum of monthly values for (3) recurrent green fractional vegetation cover also from MODIS collection 5 MOD13Q1 (Donohue et al. 2009).	All principal components of 7 variables from four products, for which source imagery was filtered to remove surface water associated with dynamic water bodies and tidal coastlines, and seasonal snow cover. The 18-year averages, 2001-2018, of annual means and maximums from 16-day aggregated data for (1) bare ground and (2) brown (litter) fractional land cover data derived from the c.500 m MODIS Collection 6 MCD43A4 product (Guerschman 2019, Guerschman and Hill 2018), and (3) recurrent green fractional vegetation cover from	As for v2.1	As for v2.1

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
	persistent and (5) recurrent green fractional vegetation cover data (Donohue et al. 2009).		c.250 m MODIS collection 6 MOD13Q1 (Donohue et al. 2009), and only annual means for (4) persistent green fractional vegetation cover also from MODIS collection 6 MOD13Q1 (Donohue et al. 2009). The lineage of annual remote sensing variables and their principal components were used as annual epochs.		
Reference ecosystem model	(GDM) Generalized dissimilarity modelling (Ferrier S. et al. 2007) of compositional turnover using scaled Manhattan distances capped at 1.0 (most different) of 10 remote sensing PCs (each rescaled 0-1) for one million reference site-pairs as the response variable, and 15 environmental covariates as predictors (explained 61.5% of model deviance, resulting in an r^2 of 0.683 and Spearman's correlation of 0.823 for	(PPR) Projection Pursuit Regression (Friedman and Stuetzle 1981) which simultaneously models the six PCs of remote sensing response variables as the sum of nonlinearly transformed linear combinations of the 21 environmental predictor variables, for 100,000 reference sites, resulting in an r^2 of 0.762 (Spearman's correlation = 0.865) for the observed vs. predicted Manhattan distances. Fit statistics based on a random subsample of	(PPR) As for v2.0, but using 7 PCs of remote sensing response variables and 23 environmental predictor variables with 101,686 reference sites, resulting in an r^2 of 0.631 (Spearman's correlation = 0.784), for the observed vs. predicted Manhattan distances, based on a random subsample of 1,000 site-pairs.	As for v2.1	As for v2.1

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
	the observed vs. predicted Manhattan distances, based on a random subsample of 1,000 site-pairs). Dissimilarity was converted to similarity.	1,000 site-pairs. This method handles inter-relationships between response variables.			
Condition algorithm (benchmarking)	Two sets of Manhattan distances derived for each reference-reference and test-reference site-pairs from 425,156 reference sites as benchmarks and 6.9 million test sites: (1) observed and (2) predicted using just the first nine remote sensing PCs (dropping the 10th). Condition is the average of the 20 test-reference site-pairs that minimised both distances, of an initial 100 site-pairs that minimised the predicted distance, weighted by the test-reference predicted distances.	Two sets of Manhattan distances were derived for each reference-reference site-pair using the reference site training data: (1) observed and (2) predicted 6 remote sensing PCs. A two-dimensional frequency histogram of these observed versus predicted distances simulates a probability density surface of the ecosystem reference state, using a bin size of 0.025, normalised within each bin of the x- axis (predicted distances) and truncated to remove irrelevant large distances. For each of the 111 million test sites, two sets of test-reference site distances are calculated using the sample of benchmark reference sites.	As for HCAS v2.0, except the bin size for the 2D simulated probability density surface was reduced to 0.005 to improve granularity and counts were smoothed using bilinear interpolation (Moore neighbourhood at 0.005) to fill gaps and edge effects. Additionally, a limited degrees of confidence calculation (weight 0.5) was used to combine the highest performing test site with the weighted average across the final 20 reference sites. This was introduced to account for uncertainty among the test-benchmark site-pairs.	As for v2.1, except substituting with the v2.2 updated benchmark reference sites.	As for v2.1, except substituting with the v2.3 updated benchmark reference sites.

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
		These distances are plotted over the reference site frequency histogram to derive expected probabilities. Condition is the half-Cauchy decay (Shaw 1995) (on the predicted distance) weighted average of the cumulative probability density of 20 test-benchmark site-pairs that minimise observed distance, of an initial 50 site-pairs that minimise the predicted distance for reference sites within 200 km circumference of a test site.			
Calibration model	The average of the condition algorithm output for 6 major urban centres was used to set the minimum score (=0), and then linearly rescaling to derive a score ranging up to 1, but all reference sites were given an inferred value of 1.	Linearly rescaled by the maximum value to derive a score ranging between 0 and 1.	A piecewise linear rescaling algorithm with two inflection points was used to simulate non-linearity. The inflection points were defined by the average uncalibrated condition values coincident with mapping of intensive land use as of December 2018 (ABARES 2019), and relatively natural areas as	As for v2.1, except the inflection points were updated using a species-area relationship (z = 0.25) back-transformation of PREDICTS project coefficients (i.e. the proportion of native species in an intact landscape which are found in paired modified habitats of that type) (Hudson et al. 2017) for land use classes	As for v2.2, except the expert and inferred reference sites were used as the relatively natural areas and the area-weighted average of the back-transformed PREDICTS project coefficients (Hudson et al. 2017) were recalculated to match. Maximum values were not truncated.

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
			of 2012 (Department of the Environment 2014), respectively. The corresponding condition scores were derived from (Chaudhary and Brooks 2018) using the weighted average of all taxon average affinities for the corresponding land use (weighting based on areal proportion of corresponding land use in Australia). The near maximum and absolute minimum uncalibrated condition values were associated with scores of 1 and 0 respectively. Maximum values were slightly truncated due to higher frequencies of reference-reference site-pairs.	aligned with highly modified or relatively natural areas. Areal proportions of corresponding Australian land uses derived from national-level land use mapping as of 2015–16 (ABARES 2022). Disaggregation of the agricultural crop category was based on the Land Use Harmonisation version 2 dataset for 2015 (LUH2 - Chini et al. 2020, Hurtt et al. 2020). Relatively natural areas were based on NVIS present major vegetation groups, version.6.0 (DCCEEW 2023c) for which proportions of primary, secondary and rangeland vegetation were derived from the LUH2 dataset for 2015, with a livestock density adjustment for Australian rangelands.	
Annual epochs	NA	NA	Annual epochs of ecosystem condition were derived using the same	As for v2.1, but applying v2.2 benchmarking and scaling adjustments.	As for v2.1, but applying v2.3 benchmarking and scaling adjustments.

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
			benchmarking process and scaling algorithm as the long term epoch by substituting the observed long-term with annual remote sensing PCs in test-benchmark comparisons. Some different reference sites used as benchmarks may be selected for each epoch.		
Uncertainty quantification	NA	Mapped training and benchmark data coverage as qualitative indicators of uncertainty due to reference site scarcity in some regions.	As for v2.0, mapped benchmark data coverage as a qualitative indicator of uncertainty due to reference site scarcity in some regions.	As for v2.1: updated the mapped benchmark data coverage as a qualitative indicator of uncertainty due to reference site scarcity in some regions.	As for v2.1: updated mapped benchmark data coverage as a qualitative indicator of uncertainty due to reference site scarcity in some regions.
Validation method	Linear comparison with 16,967 habitat hectares field assessments for the State of Victoria, filtered for 1km comparability.	Type II linear regression comparison with expert site assessments from 11 virtual transects, using Google Earth imagery, representing major Australian biomes and ecological gradients, each with 11 survey points at regular intervals; resulting in an r^2 of 0.51, intercept 0.03, slope 43.8°. Retrospective analysis by Williams et al. (2021b) using expert site	As for v2.0, Type II linear regression comparison with expert site assessments from 11 virtual transects; resulting in an r^2 of 0.63, intercept 0.09, angle 42.3°. Plus two other Type II comparisons: (1) expert site assessments derived from the habitat condition assessment tool, HCAT (Pirzl Rebecca et al. 2018, White et al. 2023), using polygon centroids for	Nil, expected to be same as v2.1.	Developed a method for validating inferred reference sites using the Harmonised Australian Vegetation plot (HAVplot) dataset (Mokany et al., 2022), which is a compilation of field data from many studies across Australia (1900–2020). As for v2.1, Type II analysis using two independent sources of data, except HCAT expert site assessments (Pirzl

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
		assessments derived from the habitat condition assessment tool (Pirzl Rebecca et al. 2018, White et al. 2023), resulting in an r^2 of 0.19, intercept 0.25, angle 31.6°.	assessments within specified spatial-temporal and certainty bounds; resulting in an r^2 of 0.42, intercept 0.18, slope 36.1°. (2) An aggregation of c. 17,000 field observations of habitat condition from four Australian states; resulting in an r^2 of 0.17, intercept -0.47, angle 64.8°.		Rebecca et al. 2018, White et al. 2023) included all applicable sites, not just polygon centroids. Results: (1) virtual transects (r^2 of 0.69, intercept -0.02, angle 46.3°); (2) HCAT assessments (r^2 of 0.67, intercept 0.13, slope 39.3°).
Evaluation methods	Comparison with two categorical land modification datasets – binary natural areas (Department of the Environment 2014) and VAST v2 (Lesslie et al. 2010).	Visual comparative assessment using auxiliary data and a series of case studies to derive a schedule of limitations.	As for v2.0, with additional case studies, visual comparisons to derive a schedule of limitations assessed for improvements over those listed under v2.0. Plus qualitative comparisons with categorical maps of land modification: the regional Landscape Health Stress Index (Morgan 2001), VAST v2 (Lesslie et al. 2010) and NVIS present major vegetation groups version 5.1 (DAWE 2018b); and two regional predictions of habitat condition: Victoria (Newell et al. 2006) and	Nil, expected to be same as v2.1.	Visual comparisons as for v2.1, with further additions to the case studies to derive a schedule of limitations compared with v2.1; plus semi-quantitative comparisons of annual epochs averaged 2001-2006 with VAST v2 (Lesslie et al. 2010) using confusion matrices and concordance assessments (e.g. 87.5% for binary comparison of relatively natural and intensively utilised categories) and a Type II analysis for five comparable categories (r^2

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
			New South Wales (Love et al. 2020).		of 0.28, intercept -0.16, slope 50.6°).
Extensions – derived from HCAS outputs	NA	NA	Experimental annual (2001 to 2018), 5 and 10-year epochs (overlapping by 5 years); National Connectivity Index v2.0 (DCCEEW 2022b, Giljohann et al. 2022). Linear regression trend analysis applied to annual epochs.	Incorporating the potential negative effects of local pressures using a distance-weighted average of site condition within 2 km circumference of the test site, modelled as an exponential decline; then recombined with HCAS condition using a geometric average to derive 'ecosystem site condition' (Williams et al. 2023a). Refined the linear trend analysis to account for temporal auto-correlation effects (Lehmann et al. 2023) applied to annual epochs.	Ecosystem site condition using the method as for v2.2. Combined expert (ecosystem state condition) and data driven (ecosystem site condition) assessments to derive 'ecosystem condition'. Applied the linear trend analysis to account for temporal auto-correlation effects (Lehmann et al. 2023) to annual epochs. Aggregated continuous condition scores into categories aligned with the VAST narrative framework (Thackway and Lesslie 2006, 2008), informed by expert elicitation.
Data publication	Not published.	Not published.	Continental Australia (Harwood et al. 2021c); Gunbower-Koondrook-Perricoota Forest (GKP) Icon Site (Harwood et al. 2021b).	Murray Darling Basin extent (Harwood et al. 2023a).	Continental Australia (Harwood et al. 2023b); Flinders, Norman and Gilbert River Catchments (Giljohann et al. 2023d); Western Australia Wheatbelt (Giljohann et al. 2023c).

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
Technical publications	Donohue et al. (2013) and Harwood et al. (2016)	Williams et al. (2020) and Lehmann et al. (2018)	Williams et al. (2021b) and Lehmann et al. (2021)	Williams et al. (2023a) and Lehmann et al. (2023).	Williams et al. (2023b)
Applications	Experimental use in research applications (Forbes et al. 2021, Nowrouzi et al. 2019)	Experimental integration into workflows depicting matters of national environmental significance managed by the Australia Government (unpublished). Demonstration use in habitat-based biodiversity assessments (Mokany et al. 2018).	Further embedded into workflows depicting matters of national environmental significance managed by the Australia Government (unpublished). Australia State of the Environment reporting 2021 Land chapter (Williams 2023, Williams et al. 2021a). Australian Government Annual report 2021/22 (DAWE 2022). Experimental ecosystem condition account for the GKP icon site (Harwood et al. 2021a). Input to landscape connectivity analysis (DCCEEW 2022b, Giljohann et al. 2022), and habitat-based biodiversity assessments (Mokany et al. 2021, Mokany et al. 2022). Demonstration application of the Bioclimatic Ecosystem Resilience Index (UN CBD indicator)	Incorporated into ecosystem condition (Williams et al. 2023a) and biodiversity (Mokany et al. 2023) components of ecosystem accounts for the Murray-Darling Basin.	Updates v2.1 in workflows depicting matters of national environmental significance managed by the Australian Government (unpublished). Incorporated into ecosystem condition (Williams et al. 2023b) and biodiversity components (Giljohann et al. 2023a, Giljohann et al. 2023b) of ecosystem accounts for two pilot mixed use landscapes in north-east and south-west Australia.

Workflow component	HCAS v1.0	HCAS v2.0	HCAS v2.1	HCAS v2.2	HCAS v2.3
			(Ferrier Simon et al. 2020, Harwood et al. 2022) for the Australian Government (2021 unpublished)		

Other contributors

Table S3 lists contributors to HCAS development from v1.0 to v2.3, other than those listed as authors in this publication.

Table S3. Other contributors to HCAS development (alphabetical by first name)

Contributor and affiliation (at the time of the contribution)	Role and HCAS version
Dwaipayan Deb, Director, Australian Government Department of Climate Change, Energy, the Environment and Water	Sponsored development of HCAS v2.4 (Williams et al. 2023c), as an update of HCAS v2.3 (Williams et al. 2023b)
Fiona Dickson, Assistant Director, Australian Government Department of Climate Change, Energy, the Environment and Water	Advised alignment with government programs and sourced funding for HCAS v2.0 (Williams et al. 2020)
Geoff R. Hosack, Research Scientist, CSIRO	Contributed to design of expert elicitation of VAST condition scores in HCAS v2.3 (Williams et al. 2023b)
Glenn Newnham, Research Scientist, CSIRO	Contributed to development of HCAS v2.4 (Williams et al. 2023c), as an update of HCAS v2.3 (Williams et al. 2023b)
Graeme Newell, Research Scientist, Victoria State Government Arthur Rylah Institute	Collaborated on development of pilot application HCAS v1.0 (Harwood et al. 2016)
Helen T. Murphy, Research Scientist, CSIRO	Contributed to development of expert elicited ecosystem state condition reported in (Williams et al. 2023b)
Jenet Austin, Experimental Scientist, CSIRO	Implementation of up-scaling method for aggregating environmental covariates (Gallant and Austin 2015)
John Gallant, Research Scientist, CSIRO	Developed the up-scaling method for aggregating environmental covariates (Gallant and Austin 2015) used in HCAS v2.1 (Williams et al. 2021b)
Karel Mokany, Research Scientist, CSIRO	Liaised with the HCAT project to acquire some of the data used in Section 4.3 of Williams et al. (2021b)
Luke Pinner, Spatial Analyst, Australian Government Department of Climate Change, Energy, the Environment and Water	Contributed to evaluation of HCAS reported in Sections 5.3-5.4, and analysis for applications in Section 12.3 of Williams et al. (2021b)
Matt Bolton, Assistant Director, Australian Government Department of Climate Change, Energy, the Environment and Water	Advised alignment of the HCAS concept (v1.0) with government programs (Donohue et al. 2013)
Matt Paget, Research Scientist, CSIRO	Contributed to development of HCAS v2.4 (Williams et al. 2023c), as an update of HCAS v2.3 (Williams et al. 2023b)
Matt White, Research Scientist, Victoria State Government Arthur Rylah Institute	Collaborated on development of pilot application HCAS v1.0 (Harwood et al. 2016)

Contributor and affiliation (at the time of the contribution)	Role and HCAS version
	and contributed agency data used in Section 3.9.4 in Williams et al. (2021b)
Randal J. L. Storey, Spatial Analyst, Australian Government Department of Climate Change, Energy, the Environment and Water	Contributed to some of the data underpinning inferred reference sites method reported in Section 4.2.2 of Williams et al. (2020)
Rebecca K. Schmidt, Research Scientist, CSIRO	Contributed to plain English communication of HCAS v2.1 (Williams et al. 2021b), and Leader of ecosystem accounting project funding Williams et al. (2023b)
Robert Lesslie, Assistant Director, Australian Bureau of Agricultural and Resource Economics and Sciences	Advised alignment of the HCAS concept (v1.0) with government programs (Donohue et al. 2013)
Sally Tetreault-Campbell, Experimental Scientist, CSIRO	Supported management of ecosystem accounting project funding Williams et al. (2023b)
Shuvo Bakar, Research Scientist, CSIRO	Developed the experimental stratified optimal sampling method reported in Williams et al. (2021b)
Suzanne M. Prober, Research Scientist, CSIRO	Contributed to development of expert elicited ecosystem state condition reported in Williams et al. (2023b)

Acknowledgments

The elicitation of new expert knowledge of ecosystem condition used in developing HCAS v2.3 was approved by CSIRO’s Social Science Human Research Ethics Committee (CSHREC) in accordance with the National Statement on Ethical Conduct in Human Research (2007, updated 2018): ethics clearance 115/22.

Ethics clearance (196/23) for reuse of expert’s ecosystem condition data from prior projects was provided through CSHREC, as follows:

- “HCAS expert opinion blitz data (original ethics clearance, 025/18)” - for the analysis of *Expert assessments provided through the rapid transects approach*. Details about this method are provided in (Williams et al. 2020). We acknowledge Garry Cook, Tanya Doody, Michael Drielsma, Rod Fensham, Simon Ferrier, Justin Perry, Suzanne Prober, Chris Ware, and Matt White for their contributions.
- “Habitat condition data using expert elicitation (original ethics clearance, 004/17)” – for the analysis of *Expert assessments provided through the Habitat Condition Assessment Tool*. Details about this method are provided in (Pirzl R. et al. 2019). We acknowledge the 28 participants from the ecological science and natural resource management practitioner community for their contributed data.
- “Land and Ecosystem Accounts Project (LEAP): implementation phase 1 of the Valuing Parks Case Study Project (the Project) – oversight of ecology sub-project (original ethics clearance, 204/19)” – for use as reference sites in developing HCAS, refer Expert elicitation Section of this report. We acknowledge Kate Bennetts, Jean Dind, Doug Frood, Megan Good, Jamie

Hearn, Paul McInerney, Gavin Rees and Genevieve Smith for their contributed data on the condition of ecosystem condition states.

References

ABARES. 2016a. Catchment Scale Land Use of Australia - Update May 2016 (digital raster data) in (ABARES) ABoAaREaS, ed. Canberra, Australia: Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), Australian Government Department of Agriculture and Water Resources.

—. 2016b. Land use of Australia 2010–11 V5 (digital data). Pages 57. Canberra, Australia: Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), Australian Government Department of Agriculture, Water and the Environment.

—. 2019. Catchment scale land use of Australia - Update December 2018. Canberra, Australia: Australian Government of Department of Agriculture, Water and the Environment.

—. 2022. Land use of Australia 2010–11 to 2015–16, 250 m. Canberra, Australia: Australian Bureau of Agricultural and Resource Economics and Sciences.

ABS. 2014. 1270.0.55.007 - Australian population grid, 2011. Canberra, Australia: Australian Bureau of Statistics.

Chaudhary A, Brooks TM. 2018. Land Use Intensity-Specific Global Characterization Factors to Assess Product Biodiversity Footprints. Environmental Science & Technology 52: 5094-5104.

Chini LP, Hurtt G, Sahajpal R, Frolking S. 2020. GLM2 Code (Global Land-use Model 2) for generating LUH2 datasets (Land-Use Harmonization 2). Zenodo.

DAWE. 2018a. Australia - Present Major Vegetation Subgroups - NVIS Version 5.1 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2018b. Australia - Present Major Vegetation Groups - NVIS Version 5.1 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2020. Australia - Pre-1750 Major Vegetation Subgroups - NVIS Version 6.0 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2021. Collaborative Australian Protected Areas Database (CAPAD) 2020 - Terrestrial. Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2022. Department of Agriculture Water and the Environment Annual Report 2021–22. Canberra, Australia: Australian Government Department of Agriculture, Forestry and Fisheries on behalf of DAWE.

DCCEEW. 2022a. Indigenous Protected Areas (IPA) - Dedicated. Canberra, Australia: Department of Climate Change, Energy, the Environment and Water.

—. 2022b. National Connectivity Index 2.0 (2001-2018). online, 'find environmental data': Department of Climate Change, Energy, the Environment and Water (DCCEEW).

—. 2023a. Collaborative Australian Protected Areas Database (CAPAD) 2010 - Terrestrial. Canberra, Australia: Australian Government Department of Climate Change, Energy, the Environment and Water.

—. 2023b. Collaborative Australian Protected Areas Database (CAPAD) 2016 - Terrestrial in Manager PD, ed. Metadata. Canberra, Australia: Australian Government Department of Climate Change, Energy, the Environment and Water.

—. 2023c. Australia - Present Major Vegetation Groups - NVIS Version 6.0 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Climate Change, Energy, the Environment and Water.

—. 2023d. Interim Biogeographic Regionalisation for Australia (IBRA) Version 7 (Regions) - 2012. Canberra, Australia: Australian Government Department of Climate Change , Energy, the Environment and Water.

—. 2023e. Interim Biogeographic Regionalisation for Australia (IBRA) Version 7 (Subegions) - 2012. Canberra, Australia: Australian Government Department of Climate Change , Energy, the Environment and Water.

Department of the Environment. 2014. Natural areas of Australia - 100 metre (digital dataset). Canberra, Australia: Australian Government Department of the Environment.

Donohue RJ, Roderick ML, McVicar TR. 2008. Deriving consistent long-term vegetation information from AVHRR reflectance data using a cover-triangle-based framework. Remote Sensing of Environment 112: 2938-2949.

Donohue RJ, McVicar TR, Roderick ML. 2009. Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. Global Change Biology 15: 1025-1039.

Donohue RJ, Mokany K, McVicar TR, O'Grady AP. 2022. Identifying management-driven dynamics in vegetation cover: Applying the Compere framework to Cooper Creek, Australia. Ecosphere 13: e4006.

Donohue RJ, Harwood TD, Williams KJ, Ferrier S, McVicar TR. 2013. Estimating habitat condition using time series remote sensing and ecological survey data: preliminary investigations. Canberra, Australia: CSIRO Earth Observation and Informatics Transformational Capability Platform.

Ferrier S, Manion G, Elith J, Richardson K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13: 252-264.

Ferrier S, Harwood TD, Ware C, Hoskins AJ. 2020. A globally applicable indicator of the capacity of terrestrial ecosystems to retain biological diversity under climate change: The bioclimatic ecosystem resilience index. Ecological Indicators 117: 106554.

Forbes O, Hosking R, Mokany K, Lal A. 2021. Bayesian spatio-temporal modelling to assess the role of extreme weather, land use change and socio-economic trends on cryptosporidiosis in Australia, 2001–2018. Science of The Total Environment 791: 148243.

Friedman JH, Stuetzle W. 1981. Projection Pursuit Regression. Journal of the American Statistical Association 76: 817-823.

Gallant JC, Austin JM. 2015. Derivation of terrain covariates for digital soil mapping in Australia. Soil Research 53: 895-906.

Geoscape Australia. 2020. PSMA - Transport & Topography - Street Network (Line) August 2020. Pages 70. Canberra, Australia: Geoscape Australia (formerly PSMA Australia).

Giljohann KM, Ware C, Mokany K, Ferrier S, Tetreault-Campbell S, Schmidt RK. 2023a. Methods for developing account-ready data: biodiversity in the Flinders, Norman and Gilbert river catchments. In: Account-ready data: biodiversity in the Flinders, Norman and Gilbert river catchments. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Giljohann KM, Ware C, Mokany K, Ferrier S, Stewart SB, Tetreault-Campbell S, Schmidt RK. 2023b. Methods for developing account-ready data: biodiversity in the Western Australian Wheatbelt. In: Account-ready data: biodiversity in the Western Australian Wheatbelt. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Giljohann KM, Drielsma M, Love J, Pinner L, Lyon P, Harwood TD, Williams KJ, Ferrier S. 2022. Technical description for a National Connectivity Index version 2.0 based on HCAS version 2.1. Canberra, Australia: CSIRO.

Giljohann KM, et al. 2023c. Account-ready data: ecosystem condition in the Western Australian Wheatbelt. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Giljohann KM, et al. 2023d. Account-ready data: ecosystem condition in the Flinders, Norman and Gilbert river catchments in Queensland. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Guerschman JP. 2019. Fractional Cover - MODIS, CSIRO algorithm. Version 3.1 (Dataset): Terrestrial Ecosystem Research Network (TERN).

Guerschman JP, Hill MJ. 2018. Calibration and validation of the Australian fractional cover product for MODIS collection 6. Remote Sensing Letters 9: 696-705.

Guerschman JP, Hill MJ, Renzullo LJ, Barrett DJ, Marks AS, Botha EJ. 2009. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. Remote Sensing of Environment 113: 928-945.

Harwood TD, Love J, Drielsma M, Brandon C, Ferrier S. 2022. Staying connected: assessing the capacity of landscapes to retain biodiversity in a changing climate. Landscape Ecology 37: 3123–3139.

Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G, White M. 2016. Habitat Condition Assessment System: A new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. Methods in Ecology and Evolution 7: 1050-1059.

Harwood TD, Richards AE, Williams KJ, Mokany K, Schmidt RK, Ware C, Ferrier S, Prober SM. 2021a. Assessing condition of ecosystem types at Gunbower-Koondrook-Perricoota Forest Icon Site. A technical report for the Land and Ecosystem Accounts Project. Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

Harwood TD, et al. 2021b. Ecosystem condition for the Gunbower-Koondrook-Perricoota Forest Icon Site v03.04.2021. Data Collection, v9. Canberra, Australia: CSIRO.

Harwood TD, et al. 2023a. Account-ready data: ecosystem site condition in the Murray-Darling Basin. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Harwood TD, et al. 2023b. 9-arcsecond gridded HCAS 2.3 (2001-2018) base model estimation of habitat condition and general connectivity for terrestrial biodiversity, ecosystem site condition, annual epochs and 18-year trends for continental Australia. Data Collection. Canberra, Australia: CSIRO.

Harwood TD, et al. 2021c. 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. v7. Data collection. Canberra, Australia: CSIRO.

Hudson LN, et al. 2017. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecology and Evolution 7: 145-188.

Hurtt GC, et al. 2020. Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. Geosci. Model Dev. 13: 5425-5464.

Lehmann EA, Ferrier S, Williams KJ. 2023. Trend test for serially correlated time series data. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2023-3255. Canberra, Australia: CSIRO.

Lehmann EA, Harwood TD, Williams KJ, Ferrier S. 2018. HCAS Activity 3a(1): Environmental Space Modelling – Methods Document. Canberra, Australia: CSIRO, Document Number EP184899.

Lehmann EA, Williams KJ, Harwood TD, Ferrier S. 2021. A not-too-technical introduction to the HCAS v2.x mechanics: a revised method for mapping habitat condition across Australia. Publication number EP211609. Canberra, Australia: CSIRO.

Lesslie R, Thackway R, Smith J. 2010. A national-level Vegetation Assets, States and Transitions (VAST) dataset for Australia (version 2.0) Canberra, Australia: Bureau of Rural Sciences, Australian Government.

Love J, Drielsma M, Williams KJ, Thapa R. 2020. Integrated model-data fusion approach to measuring habitat condition for ecological integrity reporting: Implementation for habitat condition indicators. Sydney, Australia: NSW Department of Planning, Industry and Environment.

Maus V, Giljum S, Gutschlhofer J, da Silva DM, Probst M, Gass SLB, Luckeneder S, Lieber M, McCallum I. 2020. A global-scale data set of mining areas. Scientific Data 7: 289.

Maus V, da Silva DM, Gutschlhofer J, da Rosa R, Giljum S, Gass SLB, Luckeneder S, Lieber M, McCallum I. 2022. Global-scale mining polygons (Version 2): PANGAEA.

Mokany K, Ware C, Harwood TD, Schmidt RK, Tetreault-Campbell S, Ferrier S. 2021. Biodiversity in the Gunbower-Koondrook-Perricoota Forest Icon Site and the Murray-Darling Basin. A technical report for the Land and Ecosystem Accounts Project. Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

Mokany K, Ware C, Harwood TD, Schmidt RK, Tetreault-Campbell S, Ferrier S. 2022. Habitat-based biodiversity assessment for ecosystem accounting in the Murray-Darling Basin. Conservation Biology 36: e13915.

Mokany K, Harwood TD, Ware C, Williams KJ, King D, Nolan M, Ferrier S. 2018. Enhancing landscape data: capacity building for GDM analyses to support biodiversity assessment. Canberra, Australia: CSIRO Land and Water.

Mokany K, Giljohann KM, Ware C, Harwood TD, Ferrier S, Tetreault-Campbell S, Schmidt RK. 2023. Methods for developing account-ready data: biodiversity in the Murray-Darling Basin. In: Account-ready data: biodiversity in the Murray-Darling Basin. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Morgan G. 2001. Landscape health in Australia: a rapid assessment of the relative condition of Australia's bioregions and subregions. Canberra: Environment Australia, Australian Government.

Mueller N, et al. 2016. Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. Remote Sensing of Environment 174: 341-352.

Newell GR, White MD, Griffioen P, Conroy M. 2006. Vegetation condition mapping at a landscape-scale across Victoria. Ecological Management & Restoration 7: S65-S68.

Nowrouzi S, Bush A, Harwood T, Staunton KM, Robson SKA, Andersen AN. 2019. Incorporating habitat suitability into community projections: Ant responses to climate change in the Australian Wet Tropics. Diversity and Distributions 25: 1273-1288.

OpenStreetMap Contributors. 2022. OpenStreetMap database [download.geofabrik.de/australia-oceania/australia.html, updated to 20 April 2022]. Available online: https://planet.openstreetmap.org/planet/full-history/: OpenStreetMap (OSM).

Pirzl R, White M, Williams KJ, Dickson F, Lyon P, Raisbeck-Brown N, Warnick A, Sinclair S. 2018. A novel approach to developing a view of national ecological condition using expert elicitation. Paper presented at Ecological Society of Australia Annual Conference 2018: Ecology in the Anthropocene, 25-29 November 2018, Royal International Convention Centre in Brisbane, Queensland, Australia.

Pirzl R, et al. 2019. A National Reference Library of Expert Site Condition Assessments: Development and evaluation of method. Report to the Department of the Environment and Energy. Canberra, Australia: CSIRO.

Ramm F. 2022. OpenStreetMap Data in Layered GIS Format (Free shapefiles – 2022-03-31). available online, http://download.geofabrik.de/australia-oceania/australia.html: geofabrik.

Shaw MW. 1995. Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proceedings of the Royal Society of London. Series B: Biological Sciences 259: 243-248.

Thackway R, Lesslie R. 2006. Reporting vegetation condition using the Vegetation Assets, States and Transitions (VAST) framework. Ecological Management & Restoration 7: S53-S62.

—. 2008. Describing and mapping human-induced vegetation change in the Australian landscape. Environmental Management 42: 572-590.

White MD, et al. 2023. Towards a continent-wide ecological site condition database using calibrated expert evaluations. Ecological Applications 33: e2729.

Williams KJ. 2023. Land use zones derived from HCAS v2.1 and IBRA 7 as used in the Land Chapter of the Australia State of the Environment 2021 report. Canberra, Australia: CSIRO. Data Collection.

Williams KJ, Hunter B, Schmidt RK, Woodward E, Cresswell I. 2021a. Australia State of the Environment 2021 - Land. Independent report to the Australian Government Minister for the Environment, Commonwealth of Australia.

Williams KJ, Harwood TD, Giljohann KM, Ferrier S, Lehmann EA, Ware C, Stewart SB, Tetreault-Campbell S, Schmidt RK. 2023a. Extended methods for developing account-ready data: ecosystem site condition in the Murray-Darling Basin. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2022-5750. Canberra, Australia: CSIRO.

Williams KJ, et al. 2023b. Extended methods used in developing the Habitat Condition Assessment System (HCAS) version 2.3, ecosystem condition account-ready data and experimental accounts for two mixed-use landscapes. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2023-1426. Canberra, Australia: CSIRO.

Williams KJ, et al. 2021b. Habitat Condition Assessment System (HCAS version 2.1): Enhanced method for mapping habitat condition and change across Australia. Canberra, Australia: CSIRO.

Williams KJ, et al. 2023c. Habitat Condition Assessment System (HCAS): developing HCAS v2.4 with annual epochs updated to 2021. Publication number EP2023-4902. Canberra, Australia: CSIRO.

Williams KJ, et al. 2020. Habitat Condition Assessment System: developing HCAS version 2.0 (beta). A revised method for mapping habitat condition across Australia. Canberra, Australia: Publication number EP21001. CSIRO Land and Water.

Supplemental Material C – Methods used in developing HCAS version 2.3: A continental scale example from Australia

This document provides supplemental material for the manuscript: *Overcoming Key Challenges of Satellite-based Monitoring of Ecosystem Condition: A Continental-scale Example From Australia*

From Supplemental Material A: *Common use terms with similar meaning to ecosystem condition or integrity*

- **Table S1. Example terms in common use with a similar meaning or intent as for ecosystem condition or integrity, and that are generally applicable across terrestrial, freshwater and marine realms.**

From Supplemental Material B: *A technical comparison of HCAS versions*

- **Table S2. A technical summary comparing implementations of the Habitat Condition Assessment System (HCAS) methodology from proof of concept, HCAS v1.0, to operational, HCAS v2.3.**

Contents

Introduction 3
Glossary of technical terms 3
Simplified workflow 4
Detailed workflow 6
Reference sites 8
Inferred reference sites 8
Expert nominated reference and non-reference sites 13
National extent of HCAS v2.3 inferred reference sites 17
Validating the inferred reference sites data 18
Summary of findings and caveats – validating inferred reference sites 23
Sub-sampling reference sites 24
Training data (HCAS v2.1-3) 24
Benchmark data (HCAS v2.3) 26
Environmental covariates (HCAS v2.1-3) 28
Remote sensing variables (HCAS v2.1-3) 31
Principal components of remote sensing variables (HCAS v2.1-3) 32
Spatial mask for input data 35
Predicting reference ecosystem characteristics (HCAS v2.1-3) 35
Estimating ecosystem condition 37
Similarity index for reference sites used as benchmarks in HCAS v2.3 41
Calibrating ecosystem condition (scaling, 0-1) 42
PREDICTS database 43
HCAS scaling algorithm 44
Annual epochs of ecosystem condition 49
Deriving ecosystem site condition 50
Evaluating ecosystem condition 53
Expert elicitation 54
Existing maps of ecosystem modification levels 59
Visual assessments 67
Acknowledgments 68
References 69

Introduction

The Habitat Condition Assessment System (HCAS) methodology has evolved since the proof of concept (HCAS v1.0) was published by Harwood et al. (2016a). Implementation of HCAS v2.3 (Williams et al. 2023b) builds on methods used in developing HCAS v2.0 (Williams et al. 2020), HCAS v2.1 (Williams et al. 2021) and HCAS v2.2 (Williams et al. 2023a). Methods used in developing HCAS v2.3 are summarised here, drawing together relevant material from four technical reports, including example results.

Supplemental Material B Table S2 summarises how methods evolved with successive HCAS versions up to version 2.3. HCAS v2.3 derives from the HCAS v2.1 base model and method. Revisions were made to the reference sites used as benchmarks and calibration algorithm. A method for validating the inferred reference sites was introduced and additional methods for evaluating the output condition score. An inferred local pressures index was introduced to derive ‘ecosystem site condition’ as a second-order output for use in ecosystem accounts. All else remained the same.

Common inputs and processes used in HCAS v2.1, HCAS v2.2 and HCAS v2.3 are shortened here to ‘HCAS v2.1-3’.

Glossary of technical terms

Core technical terms introduced in describing the methods used in developing HCAS v2.1-3 are provided in Box S1. A comprehensive glossary is provided in Williams et al. (2021).

Box S1 Key terms and definitions used in the HCAS base model, epoch, trend and change framework

HCAS base model – An implementation of the HCAS modelling framework that is technically complete in that both the statistical model and condition algorithm were developed using the same multi-decadal remote sensing assessment period. The base model provides the HCAS ‘best estimate of ecosystem condition’ for terrestrial native biodiversity continent-wide for a valid assessment period of at least 10 years.

HCAS epoch – An epoch uses the same model components and parameters as the HCAS base model but applies those in the benchmarking algorithm to observed remotely sensed ecosystem characteristic variables summarised over a shorter period within the timeframe of the base model. The minimum duration is one year, and may be longer, depending on how the short-term epoch is generated from the source data for compatibility with the base models’ long-term epoch.

A **best estimate** of condition results from an implementation of the HCAS base model using an assessment period of appropriate length (ideally multi-decadal).

A **derived estimate** of condition results from applying the base model to a remote sensing epoch other than the base model epoch (usually within the assessment period of the base model).

Reference sites are inferred locations representing ecosystems in reference condition with high integrity (i.e., least modified examples of their type) used as training and/or benchmark data.

Proximity to reference is the method used to estimate condition, scaled from 0.0 to 1.0.

HCAS condition trend is the linear or monotonic regression fit to observations across a time series of short-term epochs of ecosystem condition or derivatives and, ideally, encompasses at least 10 years.

HCAS condition change is estimated as the difference between two epochs with different assessment periods of length, ideally, averaged over 10 years or more, or via a trend analysis.

Simplified workflow

The HCAS workflow (Figure S1) was formulated to take both long- and short-term views by summarising the remote sensing characteristics of ecosystem dynamics from multiple decades of Earth observation imagery. Land cover products derived from satellite imagery for use in HCAS are selected to represent, as far as possible, the multiple features of an ecosystem's composition, structure and function. Statistical summaries of the multi-decadal intra-annual and seasonal variability derived from these land products are inputs to the HCAS model.

Context dependency is addressed by modelling and predicting the remote sensing signal observed for a representative set of reference sites (the natural state of an ecosystem with high levels of integrity) as a function of their abiotic environmental descriptors (e.g., climate, soils, landforms, surface water). Multiple reference sites are selected as similar ecosystem benchmarks for each site of interest to account for context dependency and for alternative expressions of an ecosystem in its reference state. HCAS ecosystem condition is measured as the weighted average of its proximity to the reference sites, and then scaled between 0.0 (ecosystem integrity extinguished) to a maximum of 1.0 (ecosystem integrity in reference condition) using empirical data to inform the calibration.

The HCAS method was specifically designed to also address a fourth challenge in monitoring ecosystem condition from space; that of inherently scarce reference site data, especially in transformed landscapes. HCAS uses a predictive framework—the reference ecosystem model—so that data gaps can be filled by statistical interpolation. This predictive framework is further leveraged in the benchmarking algorithm and by using multiple reference sites for estimating proximity to reference condition for each site of interest.

Being founded on the use of remote sensing as the observatory for monitoring ecosystem condition, the HCAS approach is necessarily limited by the present ability of remote sensing products to fully inform the structure, function and composition of ecosystems. Therefore, the HCAS condition score is considered a partial estimate of ecosystem condition, when compared with on ground observations (e.g., BioCondition - Eyre et al. 2015).

A novel feature of the HCAS framework is its adaptability to incorporate advances in satellite monitoring and analytic technologies (e.g., new or enhanced input data streams, a more dynamic modelling approach, integration with threat-based assessment products) as these become available. Therefore, each HCAS implementation is a new version that incrementally improves one or more components of the system by building on the science and technology learning of previous iterations, and back casts the time series.

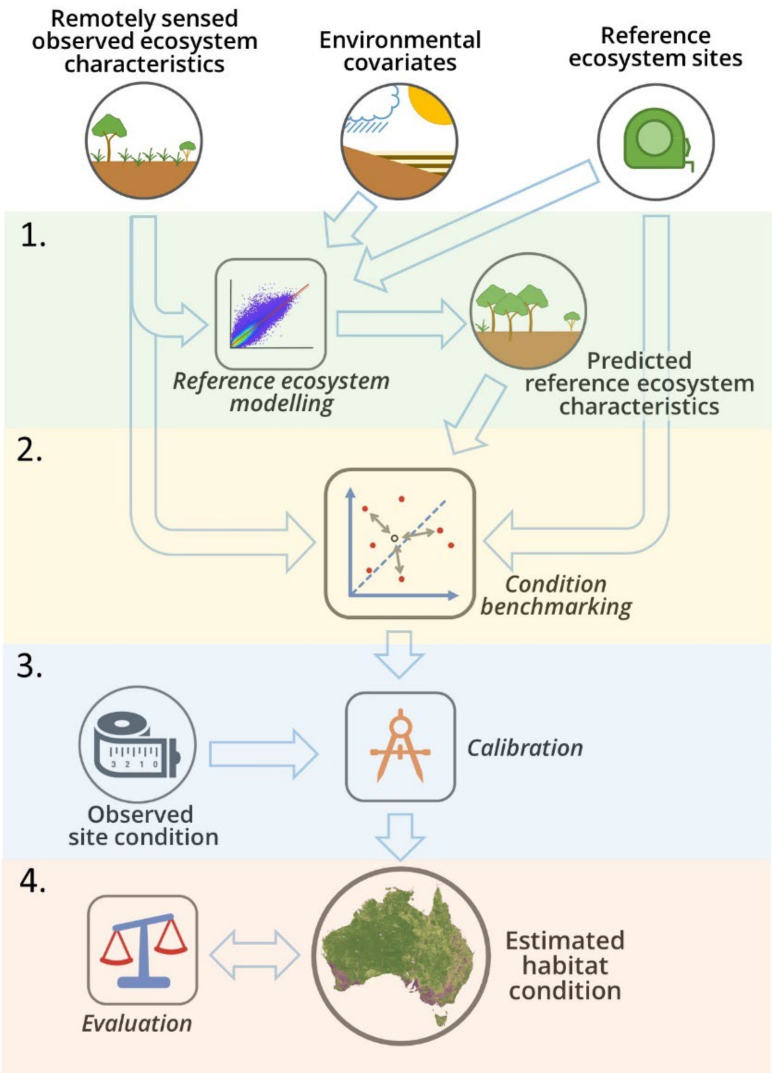


Figure S1. Summary of HCAS model workflow structure.

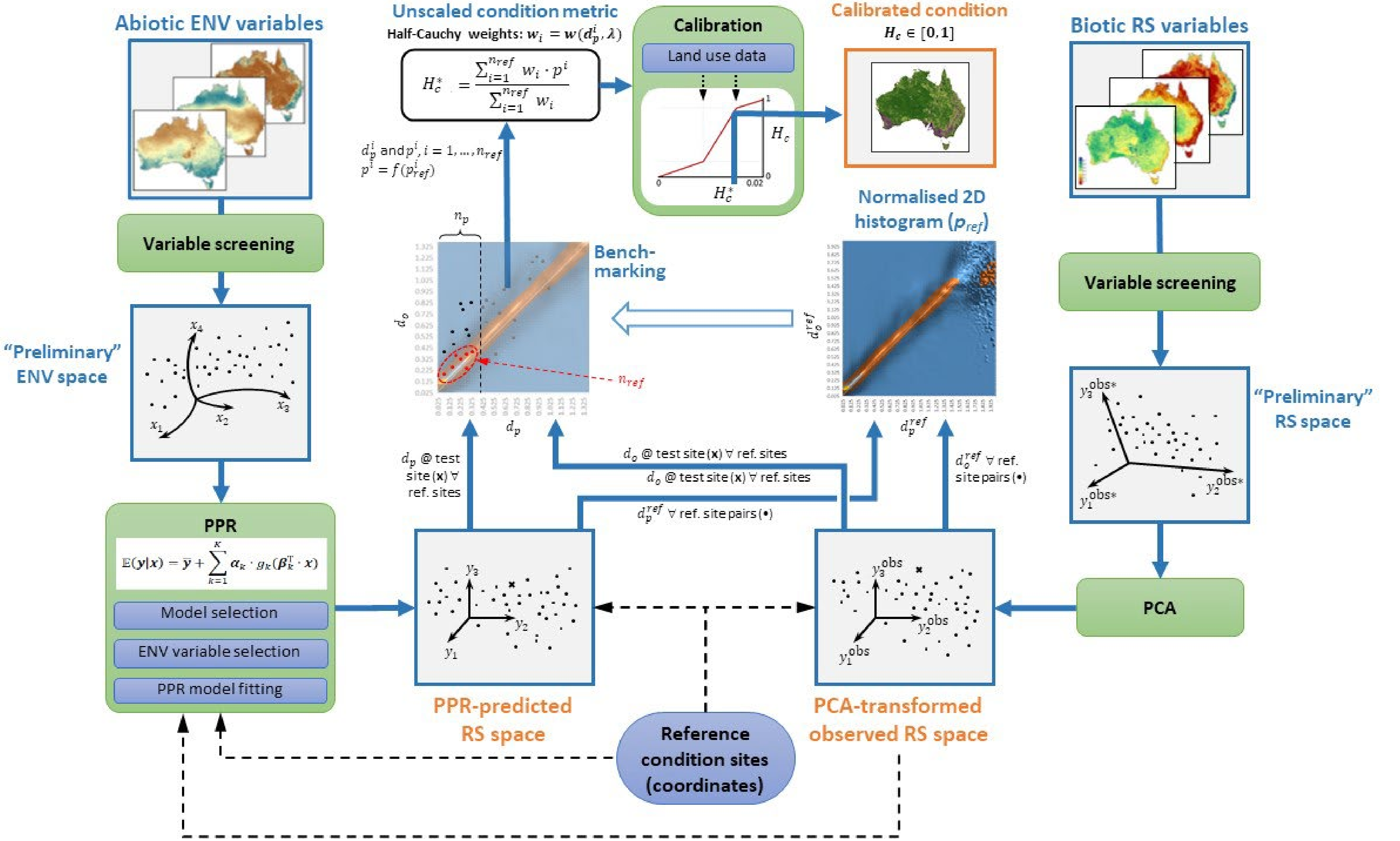
The workflow hinges on two main processing stages (shown as steps 1 and 2). In the first stage, a multivariate regression model is developed (labelled 'Reference ecosystem modelling') to predict ecosystem characteristics (using satellite-observed remotely sensed ecosystem characteristics) from a set of non-remote sensing based abiotic predictors (environmental covariates such as climate, soil, landform and surface water) for sites in reference condition (having high levels of ecosystem integrity). The reference ecosystem model is used to predict ecosystem characteristics at every site of interest. The second stage (labelled 'condition benchmarking') calculates differences between predicted and observed remotely sensed ecosystem characteristics at each site, and uses sites in reference condition (this time as 'benchmarks') to derive the initial uncalibrated habitat condition index, indicating the similarity to reference ecosystem characteristics for every test site. Subsequent steps calibrate and standardise estimates to values between 0.0 and 1.0, and compares results with other land information datasets to inform interpretation and use. Source: adapted from Figure 4 in Williams et al. (2021).

Detailed workflow

A detailed summary of the workflow used in developing HCAS v2.1-3 is provided in Figure S2 and described in Box S2.

Box S2. Plain English technical description of the HCAS v2.1-3 workflow (Figure S2)

1. Take a set of long-term remote sensing variables, transform them into a 'remote sensing space' using a principal component analysis (PCA) to ensure comparable scaling for a proper calculation of Manhattan distances in the benchmarking stage.
2. Model the observed remote sensing space as a function of environmental covariates, using Projection Pursuit Regression (PPR), for a training sample of observed reference sites. This results in a predicted reference vegetation signal derived by nonlinear transformation of the environmental space, which is directly comparable to the PCA-transformed observed remote sensing space.
3. Calculate the Manhattan distance between each pair of reference sites, separately for the observed and predicted PCA-transformed remote sensing spaces. Plot the x-axis as the predicted distance d_p and the y-xis as the observed distance d_o. Split the two distance axes into equal-sized bins Z and plot the frequency (density) of a given bin combination (i) of predicted d_p^i and observed d_o^i distances to represent the likelihood of combination of distances (d_p, d_o) for sites in reference condition. Use bilinear interpolation to compensate for under-sampling between bins, leading to a smoother surface.
4. Normalise this smoothed surface within each bin (i) along the predicted distance axis, d_p^i, to give the probability (p_{ref}) of any observed distance (d_o) for a given predicted distance (d_p), to derive the reference-distance density surface, P_{ref}.
5. To assess the condition at a test site, first select a set of reference sites that are geographically proximal to the test site, using a constant radius, R. For each test site – benchmark reference site combinations, calculate the predicted d_p^i and observed d_o^i distances from this test site to each reference site i (site-pairs).
6. Plot the position of the test-reference site pairs observed and predicted distance combinations (d_p^i, d_o^i) on the reference-distance density surface and select a subset of these test and reference site-pairs as up to n_p site-pairs with minimum predicted distance d_p^i, representing the reference sites potentially most ecologically similar to the test site, potentially suitable as benchmarks.
7. From the n_p test- reference site-pairs on the reference-distance density surface, select a further subset of n_{ref} site-pairs with maximum likelihood p_{ref}^i within the bin Z^i defined by the predicted distance d_p^i. The selected n_{ref} test- reference site-pairs define the most relevant set of 'benchmark' reference sites for the test site.
8. For each, now, test-benchmark site-pair i, $i = 1, \ldots, n_{\text{ref}}$, extract the probability score p^i from the reference-distance density surface.
9. Calculate the uncalibrated condition score H_c^{LDC} as a predicted distance d_p-weighted average (Half-Cauchy) of the probability scores p^i, calculated for all n_{ref} test-benchmark site pairs, i, using a limited degrees of confidence (LDC) algorithm to account for reference site uncertainty as suitable benchmarks.
10. Calibrate the preliminary, uncalibrated condition score H_c^{LDC} using observations of condition or other inference, that range between 0.0 and 1.0 to produce the final condition output H_c.



189

190 **Figure S2. Diagrammatic HCAS v2.1-3 workflow illustrating key concepts of the reference ecosystem modelling and benchmarking components. RS – remote sensing;**
191 **ENV – environmental; PCA – Principal Component Analysis; PPR – Projection Pursuit Regression. The workflow is described in Box S2 (H_c^* is the same as H_c^{LDC}). Note:**
192 **Lehmann et al. (2021) schematically described how the benchmarking algorithm works.**

Reference sites

The HCAS is sensitive to the correct location of reference sites. The accuracy of HCAS scores can be improved by excluding invalid reference sites and including valid reference sites that fill gaps (i.e., addressing errors of omission and commission among reference sites). Valid reference sites for the purpose of HCAS are locations where dynamic variants of the ecosystem reference state retain ecosystem integrity for the duration of the remote sensing epoch of interest. Invalid reference sites are those in which ecosystems have been modified or converted due to anthropogenic influences, including mixtures of reference and modified ecosystems at the resolution of grid cells (pixels) used in HCAS, and at any time during the period of the base model's epoch (e.g., 2001–18 in the case of HCAS v2.3).

Inferred reference sites

Reference sites used in HCAS v2.3 largely derive from logical inference, supplemented by expert knowledge and field observations. In summary, the most up-to-date national datasets depicting remnant native vegetation extent (e.g., DCCEEW 2023) and protected areas (e.g., DAWE 2021, DCCEEW 2022) provide a starting point. Then datasets depicting pressures, such as land use (e.g., ABARES 2022), settlement and infrastructure networks (e.g., ABS 2023, Geoscape Australia 2020, OpenStreetMap Contributors 2022, Ramm 2022), and mining disturbance (e.g., Maus et al. 2020, Maus et al. 2022, Werner et al. 2020) are used to exclude all potentially modified locations. Potential mixtures of reference and modified sites are also removed. Expert nominated reference sites to include, or modified locations to remove, update the output unless there is clear evidence otherwise.

Starting with the 250 m raster of spatially inferred reference sites from HCAS v2.0 (Williams et al. 2020), which were also used in HCAS v2.1 (Williams et al. 2021), potential new reference sites were included from:

- the recently gazetted Ngadju and Ngururrpa Indigenous Protected Areas (DCCEEW 2022)
- additions to the national reserve system as of 2020 (DAWE 2021)
- areas of remnant terrestrial native vegetation based on a reclassification of the present major vegetation groups provided in NVIS v6.0 (DCCEEW 2023) (see Table S3).

Using the updated 250 m raster derived above, existing and potential new inferred reference sites and remnant native vegetation were retained where they overlapped with potential reference land uses from the 2015–16 update of the Land Use of Australia dataset (ABARES 2022) (Table S4). Existing inferred reference sites and remnant native vegetation were excluded where they overlapped any other land use types that suggested at least some degree of modification, except for parts of the national reserve system (including Indigenous Protected Areas) that overlapped with grazing native vegetation or production native forests land use types (Table S4). Local information about land use and management history is needed to make decisions about which parts of recently gazetted protected areas should be excluded from consideration as reference sites. In the absence of this local information, we assumed all recent additions to the national reserve system and Indigenous protected areas were in reference condition. Open water land cover types (e.g., lakes, estuaries) were excluded from consideration because remote sensing variables used in HCAS v2.1-3 were not designed to detect condition of open water. Rivers were included because the majority are narrow linear features that may be surrounded by riparian vegetation and gallery forest that are often detectable remotely.

Potential reference sites were then excluded where they coincided with roads, railways, infrastructure and other human modified sites identified using Open Street Map (OSM) data, current to 20 April 2022 (OpenStreetMap Contributors 2022, Ramm 2022). OSM data were filtered to exclude natural features (Table S5). All retained features were first buffered by 200 m to ensure complete conversion to 250 m raster as modified land types. The 2022 update of the global-scale mining polygons dataset (Maus et al. 2020, Maus et al. 2022) was used to exclude mining sites not captured in the OSM data.

As a final step, the extent of all inferred reference condition patches was reduced (eroded) by a 250 m wide band (one grid cell wide). This was undertaken to remove cells where the remote sensing signal may include a mix of both reference and non-reference characteristics.

The overall workflow is shown in Figure S3, and results shown in Figure S4. Retained original reference sites are among those that were also reference sites in the dataset developed by Williams et al. (2020). New remnant native vegetation are additional reference sites derived from the updated NVIS v6.0 extant major vegetation groups (DCCEEW 2023), largely due to the (now) inclusion of vegetated aquatic systems such as ephemeral lakes, floodplains and palustrine wetlands (previously excluded from HCAS v2.1 due to concerns about the ability of remote sensing variables to accurately detect dynamics associated with periodic flooding). New national reserves are additional reference sites derived from the national reserve system database as of June 2020 (DAWE 2021) and subsequent additions to the Indigenous Protected Areas (DCCEEW 2022). New exclusions as modified or removed native vegetation are not specifically shown, but would include locations that are no longer considered to be in reference condition due to contrary evidence provided by land use and infrastructure datasets.

Table S3. NVIS 6.0 Major Vegetation Groups (MVGs) classified as 'remnant native vegetation' in the remap column contributed to the update of inferred reference condition sites.

MVG sort order	MVG name	Remap name
1	Rainforests and Vine Thickets	Remnant native vegetation
2	Eucalypt Tall Open Forests	Remnant native vegetation
3	Eucalypt Open Forests	Remnant native vegetation
4	Eucalypt Low Open Forests	Remnant native vegetation
5	Eucalypt Woodlands	Remnant native vegetation
6	Acacia Forests and Woodlands	Remnant native vegetation
7	Callitris Forests and Woodlands	Remnant native vegetation
8	Casuarina Forests and Woodlands	Remnant native vegetation
9	Melaleuca Forests and Woodlands	Remnant native vegetation
10	Other Forests and Woodlands	Remnant native vegetation
11	Eucalypt Open Woodlands	Remnant native vegetation
12	Tropical Eucalypt Woodlands/Grasslands	Remnant native vegetation
13	Acacia Open Woodlands	Remnant native vegetation
14	Mallee Woodlands and Shrublands	Remnant native vegetation
15	Low Closed Forests and Tall Closed Shrublands	Remnant native vegetation
16	Acacia Shrublands	Remnant native vegetation
17	Other Shrublands	Remnant native vegetation
18	Heathlands	Remnant native vegetation

MVG sort order	MVG name	Remap name
19	Tussock Grasslands	Remnant native vegetation
20	Hummock Grasslands	Remnant native vegetation
21	Other Grasslands, Herblands, Sedgelands and Rushlands	Remnant native vegetation
22	Chenopod Shrublands, Samphire Shrublands and Forblands	Remnant native vegetation
23	Mangroves	Remnant native vegetation
24	Inland Aquatic - freshwater, salt lakes, lagoons	Aquatic
25	Cleared, non-native vegetation, buildings	Cleared
26	Unclassified native vegetation	Modified
27	Naturally bare - sand, rock, claypan, mudflat	Remnant native vegetation
28	Sea and estuaries	SeaEstuaries
29	Regrowth, modified native vegetation	Regrowth
30	Unclassified forest	Modified
31	Other Open Woodlands	Remnant native vegetation
32	Mallee Open Woodlands and Sparse Mallee Shrublands	Remnant native vegetation
99	Unknown/no data	Unknown

Table S4. Categories of the Australian Land Use and Management (ALUM) classification version 8 (ABARES 2016) – a line of evidence for inferring reference condition.

Land use type	Secondary code	Tertiary code (raster values)	Potential reference condition
Nature conservation	1.1	110–117	Reference
Managed resource protection	1.2	120–125	Reference
Other minimal use	1.3	130–134	Reference
Grazing native vegetation	2.1	210	Modified
Production native forests	2.2	220–222	Modified
River	6.3	630–631	Reference
Marsh/wetland	6.5	650–651, 654	Reference

Table S5. OpenStreetMap (OSM) data layers and filters (FCLASS and NAME) used to select relevant infrastructure categories (OpenStreetMap Contributors 2022, Ramm 2022). Names that include '_a' indicate polygon format, data was otherwise in point or line format. OSM datasets 'places', 'landuse' and 'natural' were not used because these categories were provided by other datasets.

OSM data	Filename	Filter: fclass	Filter: name
Buildings	gis_osm_buildings_a_free_1.shp	all included	NA
Railways	gis_osm_railways_free_1.shp	all included	NA
Roads	gis_osm_roads_free_1.shp	all included	NA
Traffic	gis_osm_traffic_a_free_1.shp	all included except waterfall	NA
	gis_osm_traffic_free_1.shp	all included except waterfall	NA
Transport	gis_osm_transport_a_free_1.shp	all included	NA
	gis_osm_transport_free_1.shp	all included	NA
Worship	gis_osm_pofw_a_free_1.shp	all included	NA
	gis_osm_pofw_free_1.shp	all included	NA
Water	gis_osm_water_a_free_1.shp	dock, reservoir	NA
	gis_osm_waterways_free_1.shp	canal, drain	NA
Pois	gis_osm_pois_a_free_1.shp	all included except archaeological, attraction, viewpoint (these classes are	NA

OSM data	Filename	Filter: fclass	Filter: name
		covered in gis_osm_pois_free_1.shp)	
	gis_osm_pois_free_1.shp	all except archaeological, attraction, viewpoint (filters below)	NA
		Archaeological	All that indicate human modified sites (e.g., Aboriginal art, shelters, abandoned copper mine, cottage, chinatown). Natural features omitted (e.g., tree, island, cave).
		Attraction	All that indicate human modified sites. Natural features omitted (e.g., plants, animals).
		Viewpoint	All that indicate human modified sites (including all instances of lookout, outlook, viewpoint, viewing platform, bird hide). Natural features omitted (e.g., plants, animals)

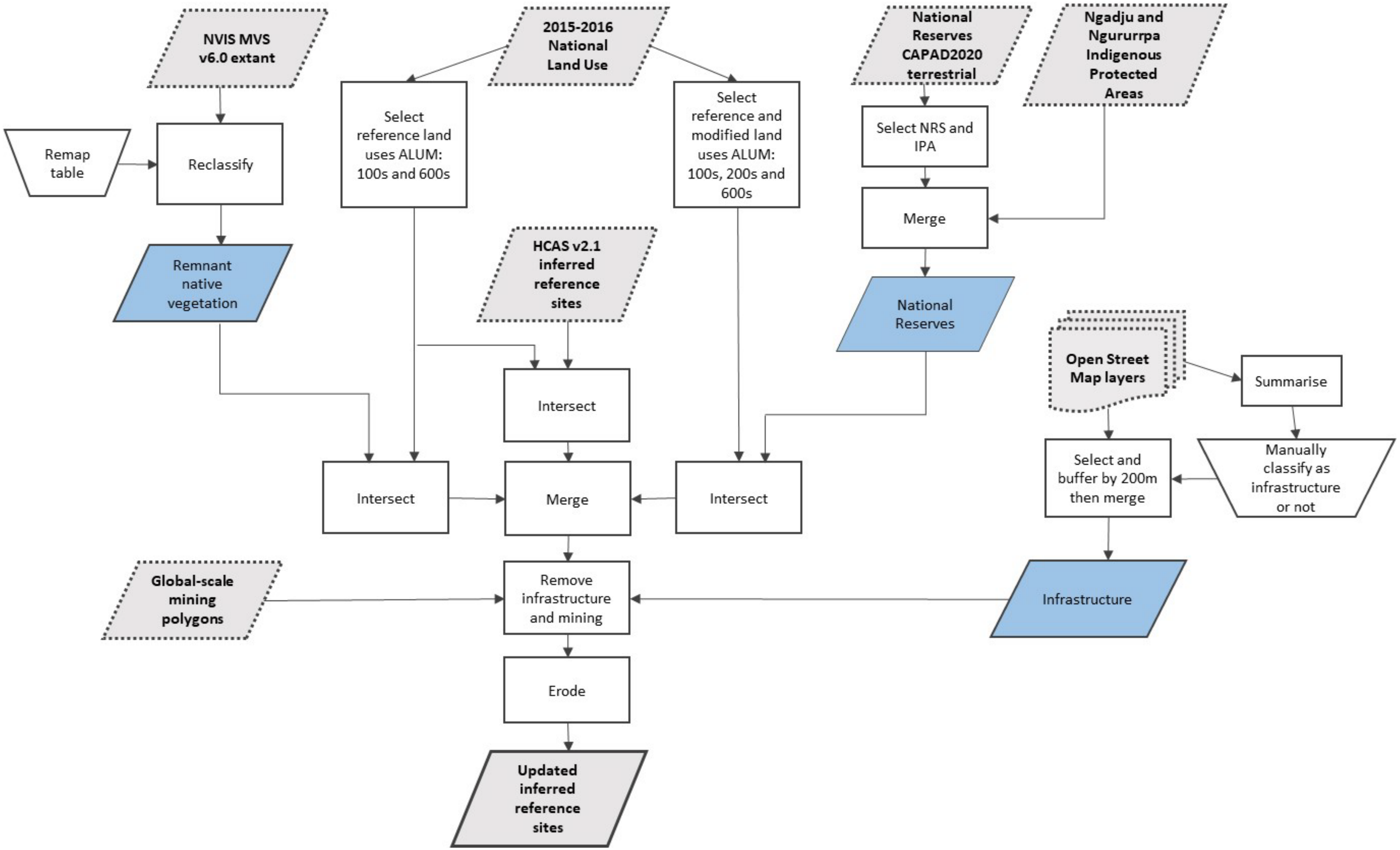


Figure S3. Schematic workflow of the multiple lines of evidence approach used to infer reference sites. Dotted lines indicate source input spatial datasets, bold lines indicate the final spatial layer of inferred reference sites.

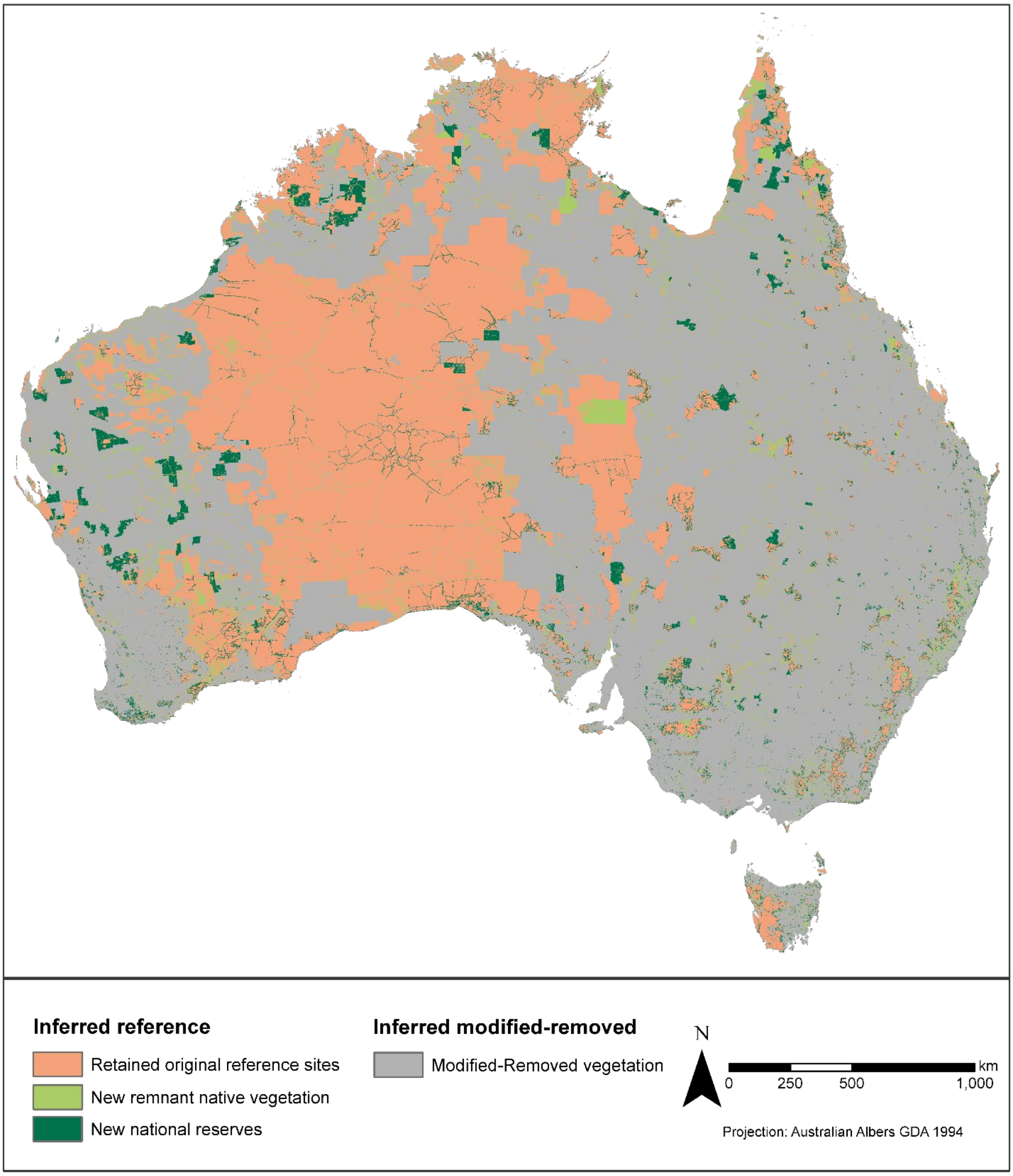


Figure S4. Rapid update of inferred reference sites resulting from multiple lines of evidence summarised in Figure S3.

Projection: geographic, GDA94. Retained original reference sites are those previously identified and published with the HCAS v2.1 data collection (Harwood et al. 2021).

Expert nominated reference and non-reference sites

Experts across two case study areas provided local knowledge about the location of reference sites based on field observations and, conversely, locations subsequently identified as modified. These data were used to update inferred modified sites to reference status and vice versa.

Expert knowledge was elicited as part of *Ecosystem State and Transition Modelling* workshops held in Townsville and Perth, Australia, during September 2022 (Prober et al. 2023, Richards et al. 2023) for the 'Flinders, Norman and Gilbert River Catchments' (FNG) and the 'Western Australian Wheatbelt' (WAW) regions, respectively. Experts were asked to propose sites they knew to be in reference condition. In the FNG case study, Queensland Herbarium experts provided condition assessments linked to spatial mapping of regional ecosystem types, and spatial polygons of areas known to be degraded (methods detailed below). In the WAW case study, experts provided point locations of field study sites and Bush Heritage Australia provided spatial mapping derived from on-ground condition assessments across five of their properties from which reference and non-reference areas could be identified (methods detailed below).

After updating the inferred reference sites database as identified by the experts, newly added sites that overlapped the buffered infrastructure mapping (Table S5) or buffered mapped road networks (Geoscape Australia 2020) as previously used in HCAS v2.0 (Williams et al. 2020), were removed (as shown in Figure S5).

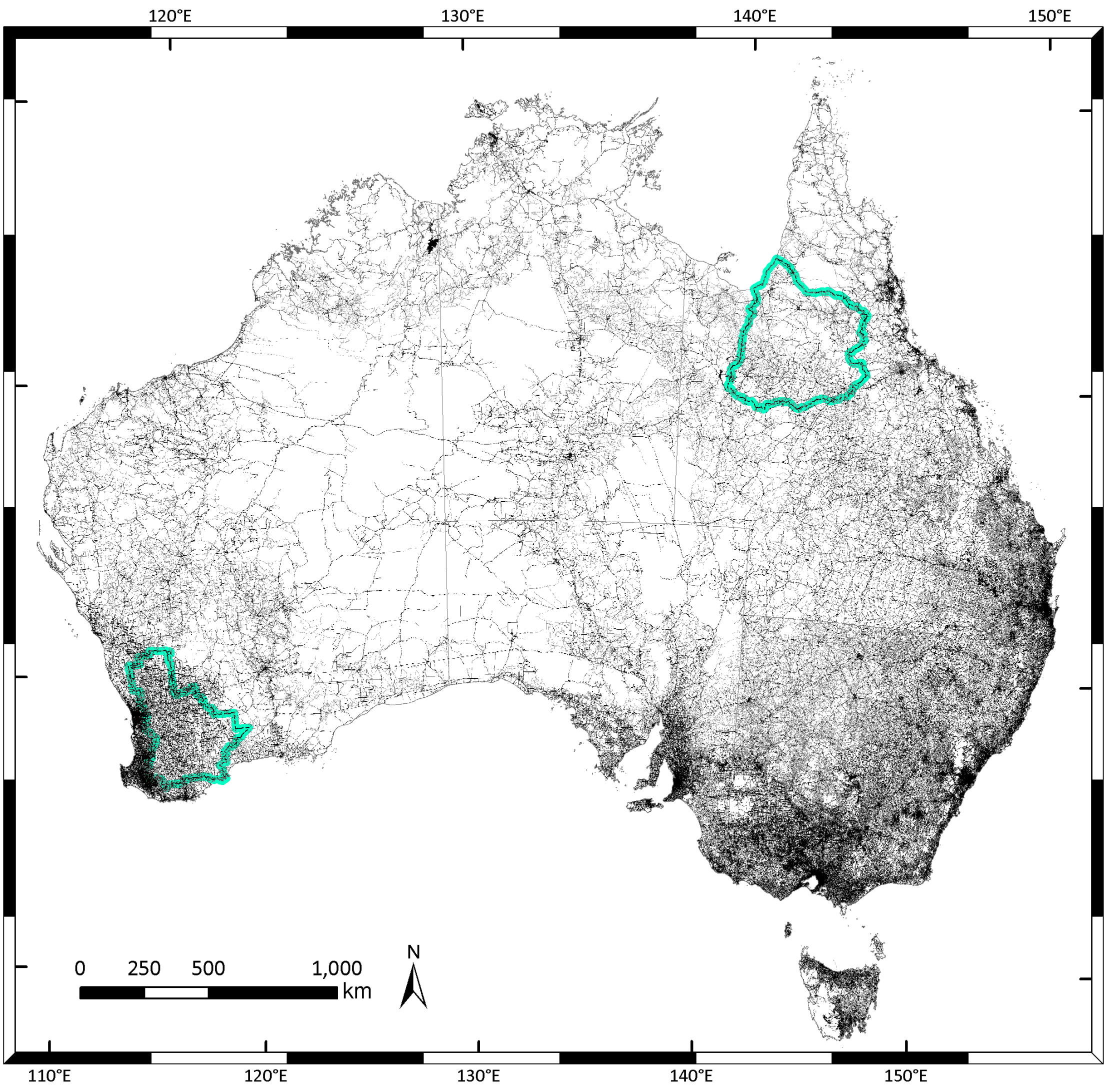


Figure S5. Location of 200 m buffered mapped infrastructure (Table S5) and transport networks (Geoscape Australia 2020) used to additionally screen and remove expert's or inferred reference sites potentially classed as modified prior to use in developing HCAS v2.3.

Blue outlined areas show the two case study regions: 'Flinders, Norman and Gilbert River Catchments' (FNG – top right) and the 'Western Australian Wheatbelt' (WAW – lower left). Projection: Australian Albers, GDA 1994.

Flinders, Norman and Gilbert River catchments in Queensland (FNG)

The Queensland Herbarium provided an assessment of the condition status of remnant native ecosystems (i.e., regional ecosystems) across the FNG case study region (‘Reference’ or ‘Reference (with caveats)’) or not in reference condition (see Appendix C in Williams et al. 2023b). Regional ecosystems (REs) assessed to be in reference condition were extracted from version 12.2 of the 2019 remnant REs for Queensland spatial data (Department of Environment and Science 2022). Each map unit provides the percentage occurrence for up to five REs. These percentages were summed for each map polygon to determine the coverage of reference condition ecosystems. Only map polygons composed entirely (100%) of either i) one or more ‘Reference’ condition RE types or ii) one or more ‘Reference (with caveats)’ condition RE types were retained. Polygons composed entirely of ‘Reference (with caveats)’ condition ecosystems were used to remove inferred reference sites coincident with potentially degraded areas.

Conversely, REs assessed to be not in reference condition were extracted from version 12.2 of the 2019 remnant regional ecosystems for Queensland spatial data (Department of Environment and Science 2022) (ie., blank fields in Appendix C in Williams et al. 2023b). All polygons composed of 5% or more ‘not reference’ condition ecosystem types were retained as indicative of the presence of modified (i.e., not reference) REs.

The resulting combination of inferred and expert-delimited reference sites for FNG is shown in Figure S6.

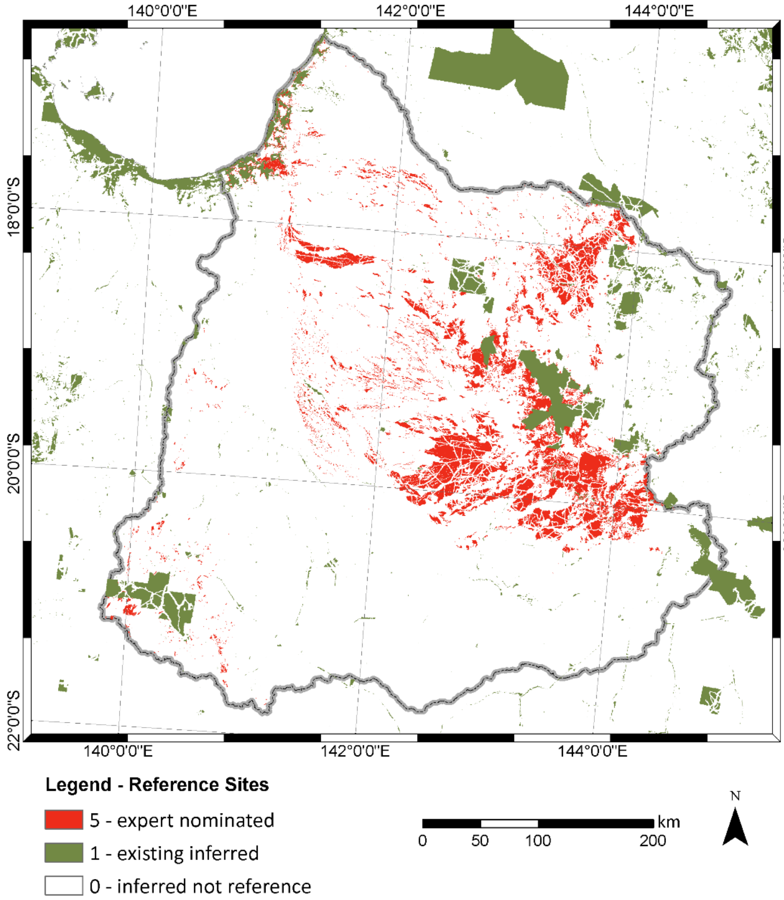


Figure S6. Location of inferred reference sites identified for the Flinders, Norman and Gilbert river catchments in Queensland (FNG) showing expert nominated sites (red colour).

FNG boundary shown in dark grey. Source: FNG_HCAS23_RC_INFERRED.tif, in the data collection (Giljohann et al. 2023, Williams et al. 2023c). Projection: Australian Albers, GDA 1994.

Western Australian Wheatbelt (WAW)

Experts provided coordinates of field study sites known to be in reference condition. All but one site, which was located in Samphire vegetation, were already represented in the inferred reference sites database (Figure S7). In addition, Bush Heritage Australia (BHA) provided spatial mapping of vegetation condition assessed locally across their properties Yarraweyah, Monjebup (north and south), Chereninup, Beringa and Red Moort Reserve (denoted 'target' in the spatial data). Areas assessed to be in reference condition, and conversely, areas assessed to be not reference, were extracted from within the BHA property boundaries (Figure S7). The resulting combination of inferred and expert delimited reference sites for FNG is shown in Figure S8.

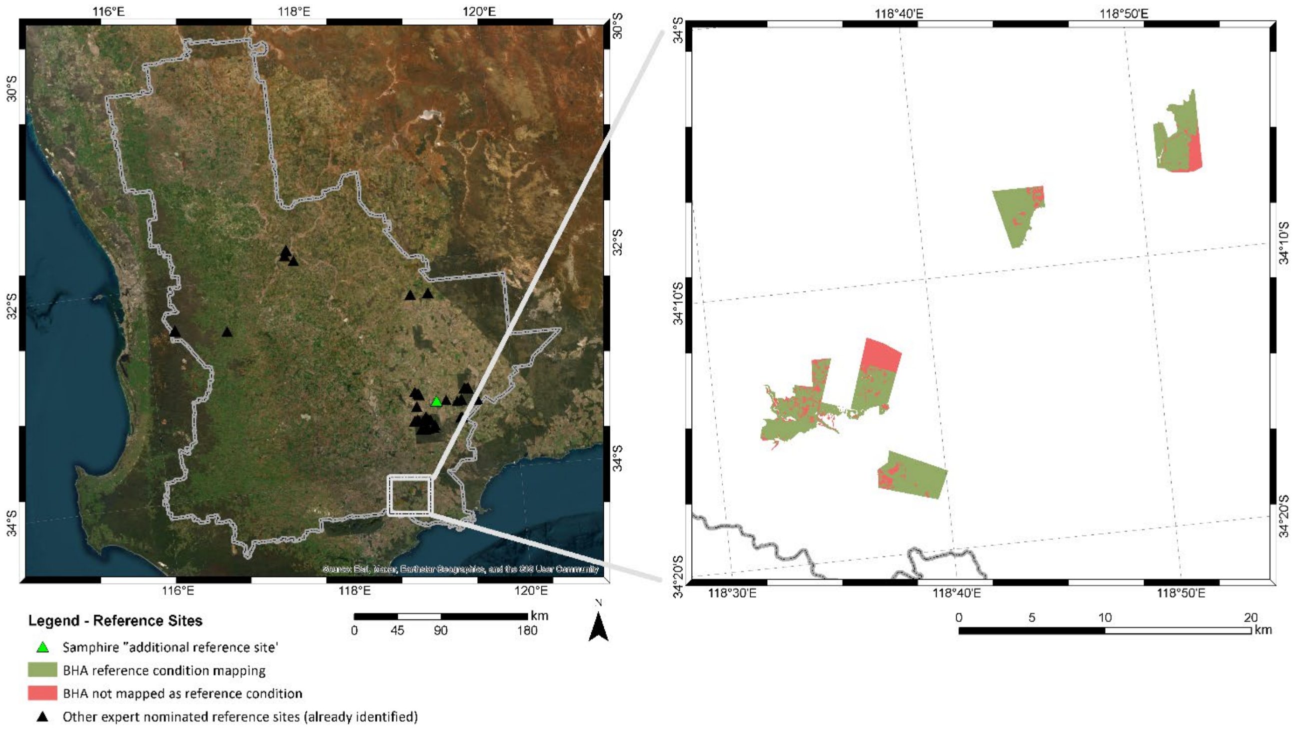


Figure S7. Expert-identified reference and not reference sites identified for the Western Australian Wheatbelt (WAW). WAW boundary shown in dark grey.
Zoom image shows the mapping of reference condition provided by Bush Heritage Australia (BHA) for their properties in the study area. Projection: Australian Albers, GDA 1994.

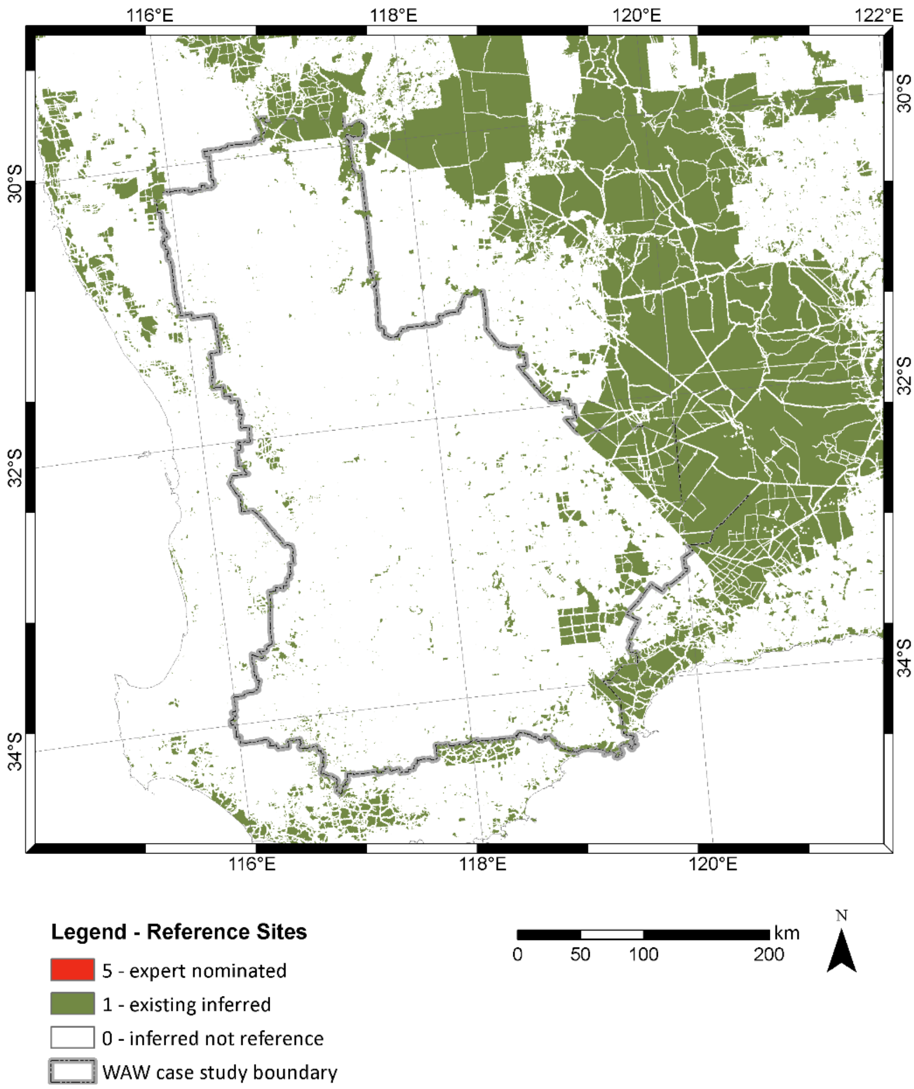


Figure S8. Location of inferred reference sites identified for the Western Australian Wheatbelt (WAW) showing expert nominated sites (red colour).

FNG boundary shown in dark grey. Source: FNG_HCAS23_RC_INFERRED.tif, in the data collection (Giljohann et al. 2023, Williams et al. 2023c). Projection: Australian Albers, GDA 1994.

National extent of HCAS v2.3 inferred reference sites

The national extent of inferred reference sites, totalling 38,773,526 pixels (250m resolution), developed for use in deriving HCAS v2.3 is shown in Figure S9, which is around 35% of all pixels.

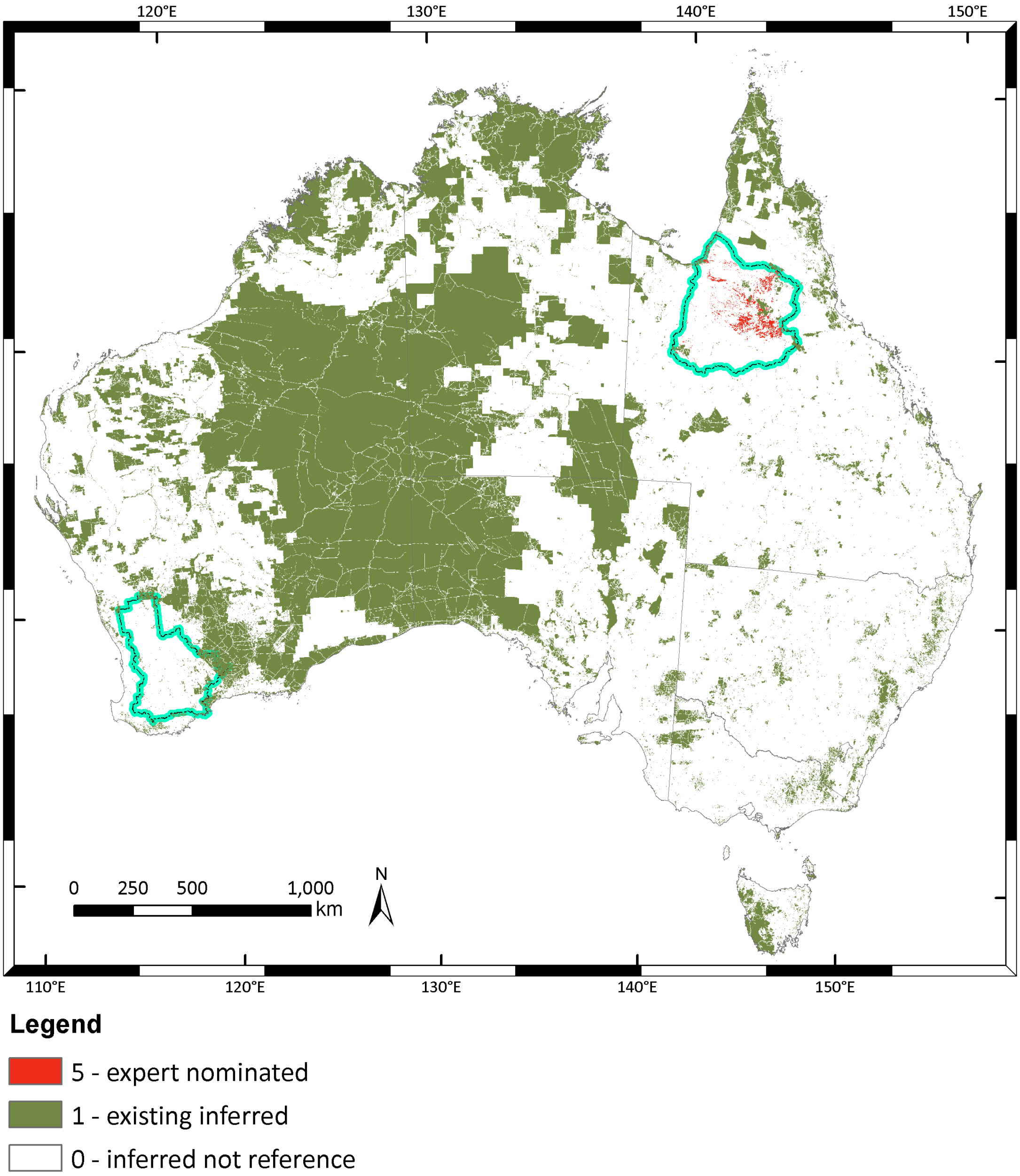


Figure S9. National extent of inferred reference sites developed for use as benchmarks in deriving HCAS v2.3 (before sub-sampling).

Blue outlined areas show the two case study regions: 'Flinders, Norman and Gilbert River Catchments' (FNG – top right) and the 'Western Australian Wheatbelt' (WAW – lower left). Projection: Australian Albers GDA 1994.

Validating the inferred reference sites data

Given the central role of reference sites in HCAS it is important to know whether the multiple lines of evidence approach, supplemented with data from experts, can identify sites that are actually in reference condition. To this end, we explored two approaches to validating HCAS reference sites using the Harmonised Australian Vegetation plot (HAVplot) dataset (Mokany et al., 2022): i) comparison of frequency distribution plots (histograms) and ii) a presence-only statistical model created using the MaxEnt algorithm (Phillips 2022).

The HAVplot dataset (Mokany et al., 2022) includes 219,552 sites from field-based floristic vegetation surveys undertaken between 1900 and 2020 across Australia. Plot areas range from 1m^2 to 4,000,000 m^2 (median = 400 m^2). Plot location is given as a point coordinate (latitude and longitude). We used only the most recent survey data at each plot location, comprising 206,472 sites (99% were ≥1970, 88% were ≥1990 and 62% were ≥2000). We used data on the proportion of species that are native in each plot as an indicator of habitat quality. Native species proportions approaching 100 percent were assumed to have a higher likelihood of occupying habitat that is close to reference condition.

As we were only interested in validating HCAS inferred reference sites (i.e., not whether some may be modified), we compared the subset of HAVplot sites coinciding spatially with HCAS inferred reference sites to all HAVplot sites. First, the HAVplot data subset (n = 206,472) was filtered to retain only the most recent survey (by year) within each 250 m grid cell coinciding with the HCAS analysis mask, irrespective of whether they coincided with inferred reference sites or not (n = 163,870). The subset of sites that also intersected the HCAS v2.3 inferred reference site spatial data were identified (n = 15,747).

In this way, two datasets were created from the filtered vegetation survey data: i) ‘all HAVplot’ ‘background’ sites (i.e., n = 163,870) and ii) the subset of ‘all HAVplot’ sites coinciding with HCAS inferred reference areas (i.e., n = 15,747). Around 10% of HAVplot data were found to coincide with inferred reference sites. The HAVplot data cover a wide range of survey years. The 50th and 75th quantiles of survey year for the background dataset (i) are, respectively, 1999 and 2008, and in the subset that also coincide with HCAS inferred reference areas (ii) are 1998 and 2005, respectively. It is not known if some of these sites are no longer in reference condition.

Comparison of frequency distribution plots

Histogram were used to visually assess whether the distribution of HCAS inferred reference sites is biased towards sites with a higher proportion of native species (i.e., higher likelihood of being in reference condition), and to compare the distribution of sites across the two datasets.

Histograms revealed that both of the HAVplot data subsets are strongly biased towards sites with a high proportion of native species (Figure S10). The subset of HAVplot data coinciding with HCAS inferred reference sites (orange bars) contains proportionally fewer lower quality sites (proportion of native species), but proportionally more higher quality sites than the ‘all sites’ HAVplot dataset (grey bars) (Figure S10).

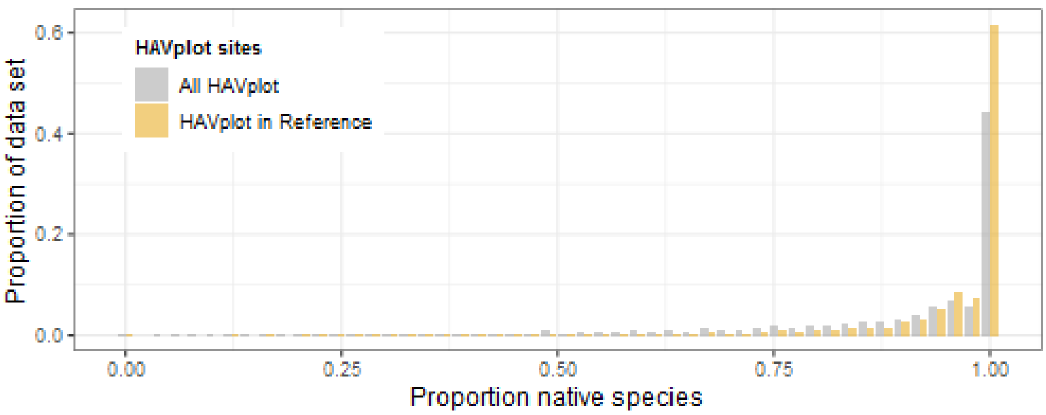


Figure S10. Proportion of background HAVplot sites (grey bars, *n* = 163,870) or proportion of those HAVplot sites that also coincide with HCAS inferred reference sites (orange bars; *n* = 15,747) by proportion of species that are native.

Histogram bins are ~0.02 wide.

Figure S11 shows that there is a larger proportion of presumed relatively natural vegetation survey sites (as determined by HAVplot proportion of native species approaching 100%) in the set of HAVplot data coinciding with HCAS inferred reference sites than in the 'all sites' HAVplot data subset. This indicates that the HAVplot data that coincide with HCAS inferred reference sites (i) are not simply a random sample of the 'all sites' HAVplot dataset (ii) but are biased towards sites containing a greater proportion of native species. If the HAVplot data coinciding with HCAS inferred reference sites were a random sample, the bars in Figure S11 would be equal across the histogram bins.

The HAVplot data that coincide with HCAS inferred reference sites has a higher frequency of native species proportions approaching 100% than the background HAVplot data, thereby qualitatively validating the multiple lines of evidence approach used to infer reference sites.

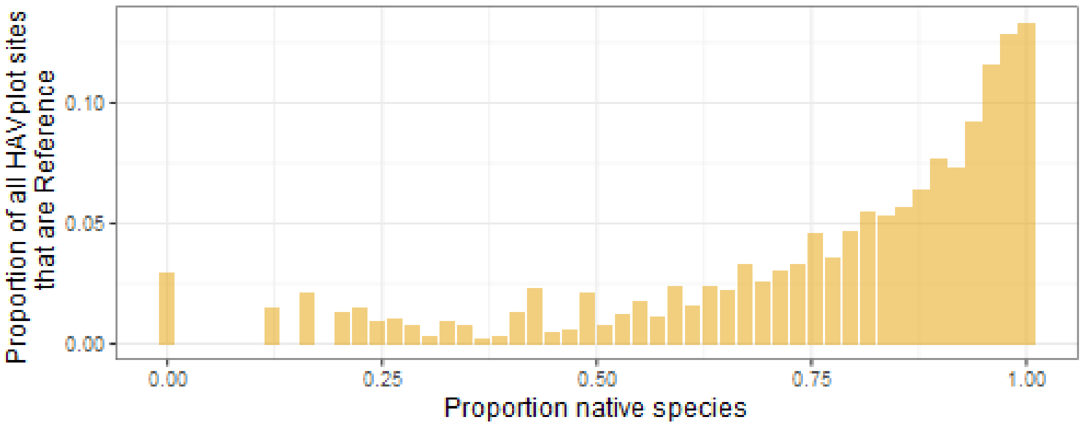


Figure S11. The subset of HAVplot sites that coincide with HCAS inferred reference sites (*n* = 163,870) as a proportion of background HAVplot sites (*n* = 163,870) for each of the 50 'proportion native species' bins. Histogram bins are ~0.02 wide.

MaxEnt model of reference site occurrence

MaxEnt (Phillips 2022) was used to derive a quantitative measure. HAVplot data coinciding with inferred reference sites provided the presence response variable (=1), and HAVplot background data (=0), to inform occurrence rates, which MaxEnt uses to characterise the environment (Phillips et al. 2006). The 'proportion of species that are native' was the sole predictor in the model. Models were fitted with either linear or a combination of linear and quadratic features, using spatial cross-validation with 5-folds. Folds were randomly allocated to 15 spatial blocks using the R package *blockCV* (Valavi et al. 2019) (Figure S12).

Model validation results are presented for each fold, and as the average and standard deviation across folds. Three statistics were used to evaluate model performance (see explanations below): the Area under the ROC curve (AUC), which measures the model's ability to discriminate the environment at withheld occurrence sites from those in the full set of background samples (training and validation); the continuous Boyce index (CBI); and the Akaike Information Criterion corrected for small sample sizes (AICc), which provides information on model quality given the data. The average model prediction is plotted using a clog-log transformation, which is considered to approximate occurrence probability (with assumptions) bounded by 0 and 1 (Phillips et al. 2017).

Models and statistics were implemented in R v4.2.1 (R Core Team 2022) using the ecological niche modelling and evaluation package *ENMeval* v2.0.3 (Kass et al. 2021) that calls MaxEnt from the *maxnet* package v0.1.4 (Phillips 2022).

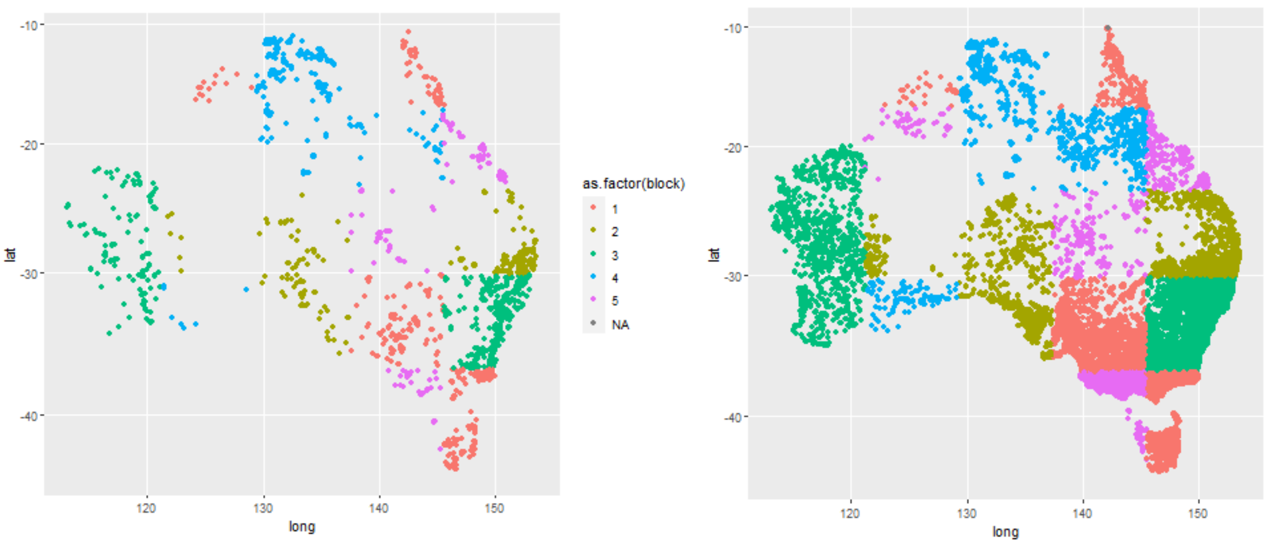


Figure S12. Spatial blocks for the two datasets used in the maxent model: HAVplot coinciding with HCAS inferred reference sites (left) and background HAVplot sites (right).

Area under the Receiver-Operator Characteristic (ROC) curve

The Area Under the ROC Curve (AUC) is a commonly used threshold-independent measure of predictive accuracy based on the ranking of locations. Originally developed for binary classified data (e.g., presence/absence), when applied to presence-only models, AUC is interpreted as the probability that a randomly chosen presence point is ranked higher than a randomly chosen background point (Merow et al. 2013). High AUC values indicate the model can distinguish between presences and background points. However, as the MaxEnt background sample also contains the presence points, and as there is no reason to expect all non-reference HAVplot sites to be poor quality (i.e., have a low proportion of native species) AUC is unlikely to be informative of the pattern we aim to detect.

The continuous Boyce index

The continuous Boyce index (CBI) is a presence-only and threshold-independent evaluator for species distribution models (Hirzel et al. 2006). It has been suggested to be the most appropriate way to evaluate predictions from presence-only models like MaxEnt (Di Cola et al. 2017, Hirzel et al. 2006). The Boyce index measures how much model predictions differ from a random distribution of the observed presences (Boyce et al. 2002) (i.e., the trend in the proportion of presences across classes of the predictions) and the CBI applies the Boyce index within a moving window across prediction gradient. It is considered the quantitative equivalent of the graphical presence-only calibration (POC) plot (Phillips and Elith 2010). The CBI is analogous to a Spearman correlation and varies between –1 and +1. Positive values indicate a model in which predictions are consistent with the distribution of presences in the evaluation data set, values close to zero mean that the model is not different from a random model, and negative values indicate counter predictions.

Another statistic

Another metric recently advocated for evaluating presence-only models is the area under the Precision-Recall Gain curve (AUC-PRG) (Sofaer et al. 2019, Valavi et al. 2022). In contrast to AUC (ROC) this metric specifically focuses on the accurate prediction of presences, not whether absences are correctly predicted; making it potentially more relevant for ecological cases where the costs of distinct error types are different (such as species distribution models in conservation prioritisation). However, as the AUC-PRG includes the number of false positives in its calculation (e.g., presumed relatively natural HAVplot vegetation survey sites that are not HCAS reference sites) it is not a useful metric for this situation and will not be considered further.

Results

The MaxEnt models revealed a strong association between HAVplot within HCAS inferred reference sites and habitat quality (i.e., higher proportion of species that are native). The probability (relative likelihood) of a HAVplot site also being a reference site increased from approximately 10% when half the species in a plot were native (proportion native = 0.5) to a maximum of approximately 70% when almost all the species in a plot were native (proportion >= 0.95) (Figure 3). Evaluation statistics were similar for the models containing linear or linear and quadratic features (Table S6). Both models had equal, albeit poor discrimination as measured by AUC (average AUC = 0.63). However, model predictions were strongly consistent with the distribution of the HAVplot evaluation data (average CBI = 0.88–0.89) with very strong agreement for three of the five cross-validation folds (CBI >=0.9). Using quadratic features in the model provided a better fit to the data (difference in AICc = 22).

Table S6. Statistics for the MaxEnt models fit with linear features or with linear and quadratic features.

Model features	Fold	AUC (standard deviation of the mean)	CBI standard deviation of the mean)	AICC
Linear	1	0.60	0.90	
	2	0.67	0.94	
	3	0.66	0.93	
	4	0.58	0.78	
	5	0.65	0.84	
	Average	**0.63 (0.04)**	**0.88 (0.07)**	**38573**
Linear and Quadratic	1	0.60	0.90	
	2	0.67	0.96	
	3	0.66	0.95	
	4	0.58	0.74	
	5	0.65	0.87	
	Average	**0.63 (0.04)**	**0.89 (0.09)**	**38551**

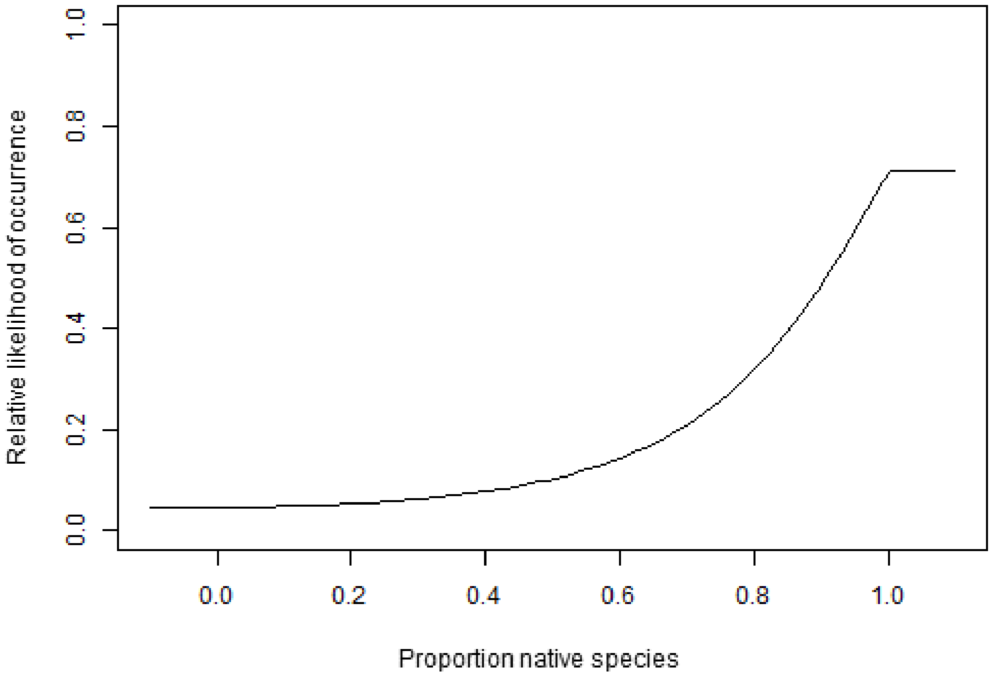


Figure S13. Modelled probability of a HAVplot site being also a HCAS v2.3 inferred reference site as a function of the HAVplot habitat quality predictor – proportion of native species.
Predictions are the average across the five folds from the model using both linear and quadratic features.

Summary of findings and caveats – validating inferred reference sites

Overall, the validation exercise supports the multiple lines of evidence approach to inferring reference sites for use in HCAS. The subset of HAVplot sites that also coincide with HCAS v2.3 inferred reference sites are encouragingly biased towards sites containing a higher proportion of native species (i.e., presumed relatively natural with higher ecosystem integrity). The MaxEnt models predicted that the relative likelihood of a HAVplot site being a HCAS inferred reference site increased with higher proportions of native species, with presence-only CBI statistic indicating strong model performance.

As was expected, given the large number of background (i.e., assumed non-reference) sites with high proportions of native species the models had poor discrimination as measured by AUC. For this validation exercise, AUC was not expected to perform well because it is calculated using both the prediction of occurrences and background. We expected there to be higher proportions of native species at vegetation survey sites that were not coincident with HCAS inferred reference areas (background data). This outcome was not of importance for our approach to validation.

However, there are important caveats to note.

1. The number of HAVplot sites that coincide with HCAS inferred reference sites represent only a very small fraction of all available HCAS reference sites (15,747/39,685,172). It is feasible that these HAVplot sites might not be a very representative sample of HCAS reference sites, and so drawing conclusions about the entire set of reference sites is not possible without first comparing sample structures.
2. We used proportion of species that are native in the HAVplot data as an indicator of likelihood in reference condition (i.e., presumed relatively natural). The implicit assumption

is that higher proportions of native species are positively correlated with high levels of ecosystem integrity. However, this may not be accurate as the measure is just one variable considered important for the estimation of condition (i.e., variables should encompass structure, function and composition characteristics of ecosystems).

3. HCAS v2.3 and HAVplot datasets do not address the same spatial footprint. Plot area in the HAVplot dataset ranges from 1 m^2 to 4,000,000 m^2 (median = 400 m^2) and plot boundaries are unknown (represented by a single point) and so do not necessarily align with the 250 m grids of the HCAS v2.3 reference layer. Further, the larger plot sizes are likely to be area-aggregated lists of native species only from a number of surveys rather than actual vegetation survey plots. In future analyses, a filter for plot size or other descriptor indicating the original method used to obtain the data should be included.

4. HCAS v2.3 and HAVplot datasets are not temporally aligned. The remote sensing data underpinning HCAS v2.3 spans the years 2001 to 2018, whereas HAVplot surveys range from 1900 to 2020. This could result in a mismatch between the remote sensing signal and the on-ground vegetation assessment. However, it is assumed that contemporary inferred reference sites were continually in high ecosystem integrity decades earlier than the earliest date of the remote sensing epoch in 2001. Therefore, spanning earlier survey dates is reasonable.

5. The subset of HAVplot survey sites used for validation were the most recent within a 250m grid cell, which is the resolution of the remote sensing data used in HCAS. Therefore, earlier HAVplot survey sites that may indicate mixtures of native versus introduced plant species or reinforce naturalness were not reflected in the comparisons. In future validation exercises, a more complete history of vegetation surveys could be used, for example by including covariates such as survey date or year, plot size and survey method in the analysis, for example.

Sub-sampling reference sites

The multiple lines of evidence approach resulted in proportionally larger number of reference sites in the more remote and central regions of arid Australia, where extensive areas are relatively undeveloped or form part of Australia's network of protected areas. The subsample of reference condition sites to use as training data or benchmarks needed to minimise bias toward the arid regions and characterise, as far as possible, the reference state diversity of Australian ecosystems.

Training sites are used in the reference ecosystem model, and benchmarks are used subsequently in the condition algorithm using proximity to reference state. The HCAS v2.3 applied the HCAS v2.1 reference ecosystem model, and so the training sites are the same as those derived for use in HCAS v2.1. The HCAS v2.3 benchmarks were derived using the updated multiple lines of evidence approach described above.

Training data (HCAS v2.1-3)

The inferred reference sites used were initially derived using an earlier multiple lines of evidence approach for use with HCAS v2.0 (Williams et al. 2020). Those inferred reference sites accounted for 28.8% (35,485,829 0.0025-degree grid cells) of continental Australia (i.e., 71.2% of lands were considered modified to some degree or aquatic and excluded from consideration).

An ecological land classification was derived by spatially intersecting the ‘present’ mapped extent of Australia’s 85 major vegetation subgroups (MVS) version 5.1 (DAWE 2018) with the 411 bioregional subregions version 7.0 (Department of the Environment 2014) to derive a stratification comprising 7039 discrete regions. This classification was rasterised to match the 9-arcsecond digital elevation model for Australia (Hutchinson et al. 2008). Several MVS categories that suggest the land unit is a waterbody or a modified vegetation type were excluded from consideration:

- ‘Salt lakes and lagoons’
- ‘Freshwater, dams, lakes, lagoons or aquatic plants’
- ‘Regrowth or modified forests and woodlands’
- ‘Regrowth or modified shrublands’
- ‘Regrowth or modified graminoids’
- ‘Regrowth or modified chenopod shrublands, samphire or forblands’
- ‘Unclassified forest’
- ‘Cleared, non-native vegetation, buildings’
- ‘Unknown/no data’.

This had the effect of further excluding reference that may otherwise have been included using the multiple lines of evidence described in Williams et al. (2020). The resulting ecological land classification comprises 4961 discrete types of which 4952 included at least one inferred reference site. Of these land units, 18% had 10 or fewer reference sites and 45% had 100 or fewer. To approximate equal representation of 4952 units (strata) in the training dataset, we randomly drew up to 25 reference site samples from each, to achieve around 100,000 samples for the projection pursuit regression (PPR) model. Approximately 26% (1312) of the ecological land units provided 24 or fewer reference sites, resulting in 101,686 reference sites for use as training data. This approach systematically sampled environmental diversity, approximated by 4961 ecological units.

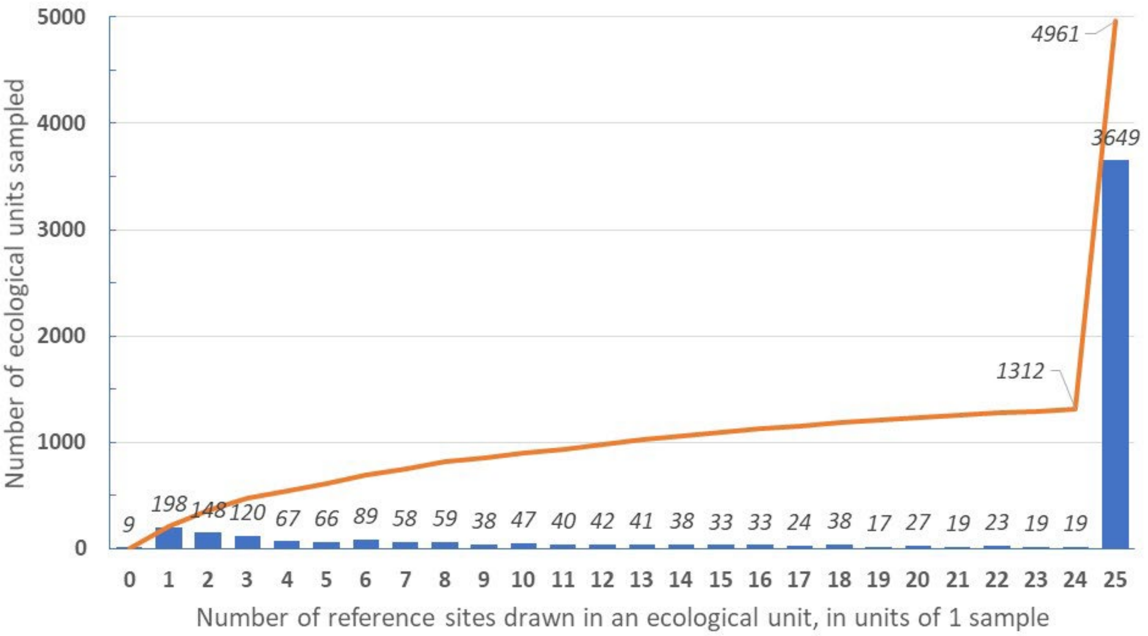


Figure S14. Reference site sampling within 4961 ecological land units, showing number of units with sampling below the target level of 25.

There are 9 units for which zero reference sites were available to be sampled. The orange line shows the cumulative count of units (26% have < 25 reference sites available to be sampled).

Benchmark data (HCAS v2.3)

A national benchmark dataset of 202,515 expert identified and inferred reference sites was derived by taking a random-stratified sample of up to 50 sites per strata, without replacement (Figure S15). Strata were spatially defined by the historic extent (pre-1750 mapping) of major native vegetation subgroups (pre-1750 NVIS version 6.0 - DAWE 2020) within bioregional subregions (IBRA version 7.0 - Department of the Environment 2014), as shown in Figure S16.

The frequency distribution of sampled reference sites within the 5481 strata containing at least one site is shown in Figure S17. Of these, 36% of strata (1988 in total) had less than 50 reference sites available for selection, resulting in all available reference sites being selected in those cases.

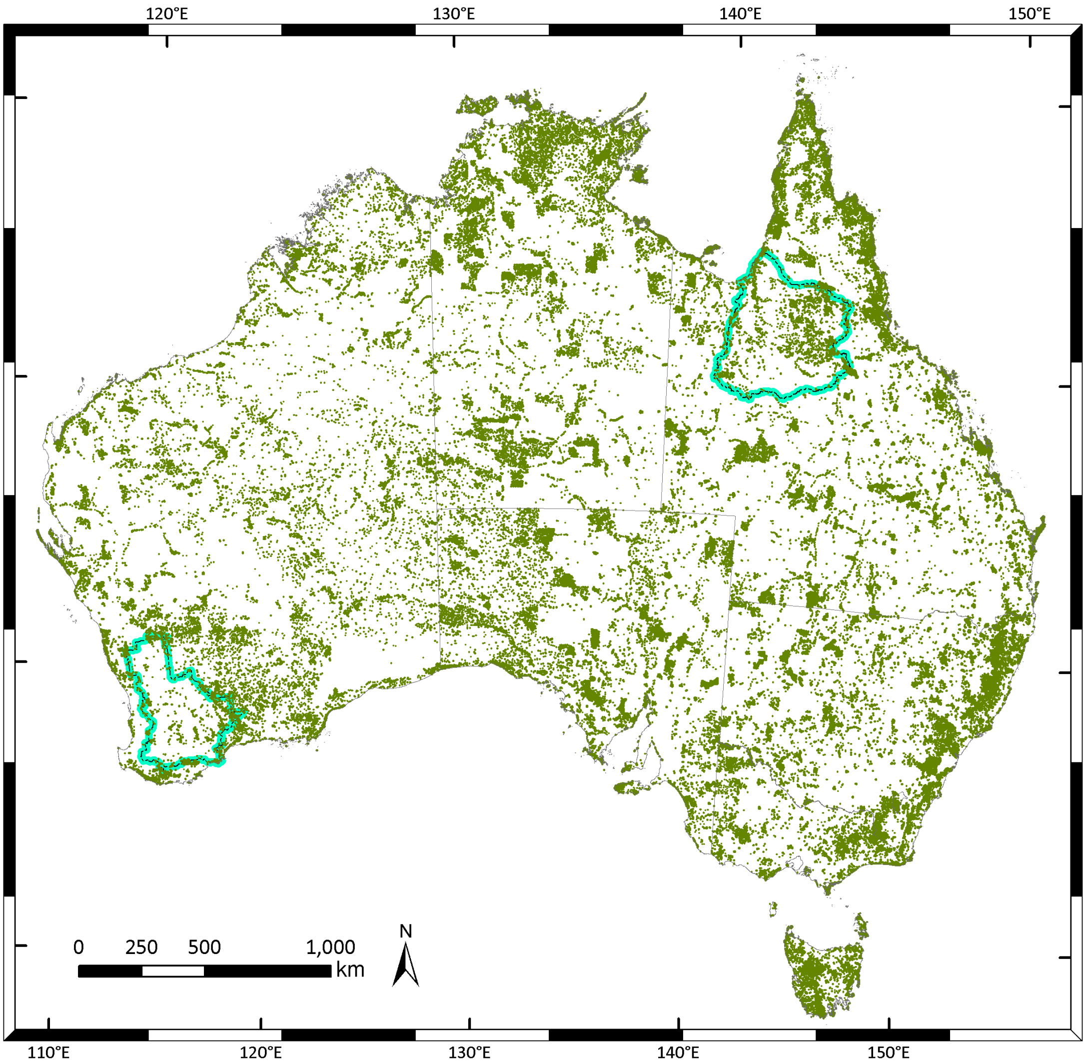


Figure S15. Spatial pattern of 202,515 sites sampled from 5481 ecological strata.
Blue outlined areas show the two case study regions: 'Flinders, Norman and Gilbert River Catchments' (FNG – top right) and the 'Western Australian Wheatbelt' (WAW – lower left). Projection: Australian Albers GDA 1994.

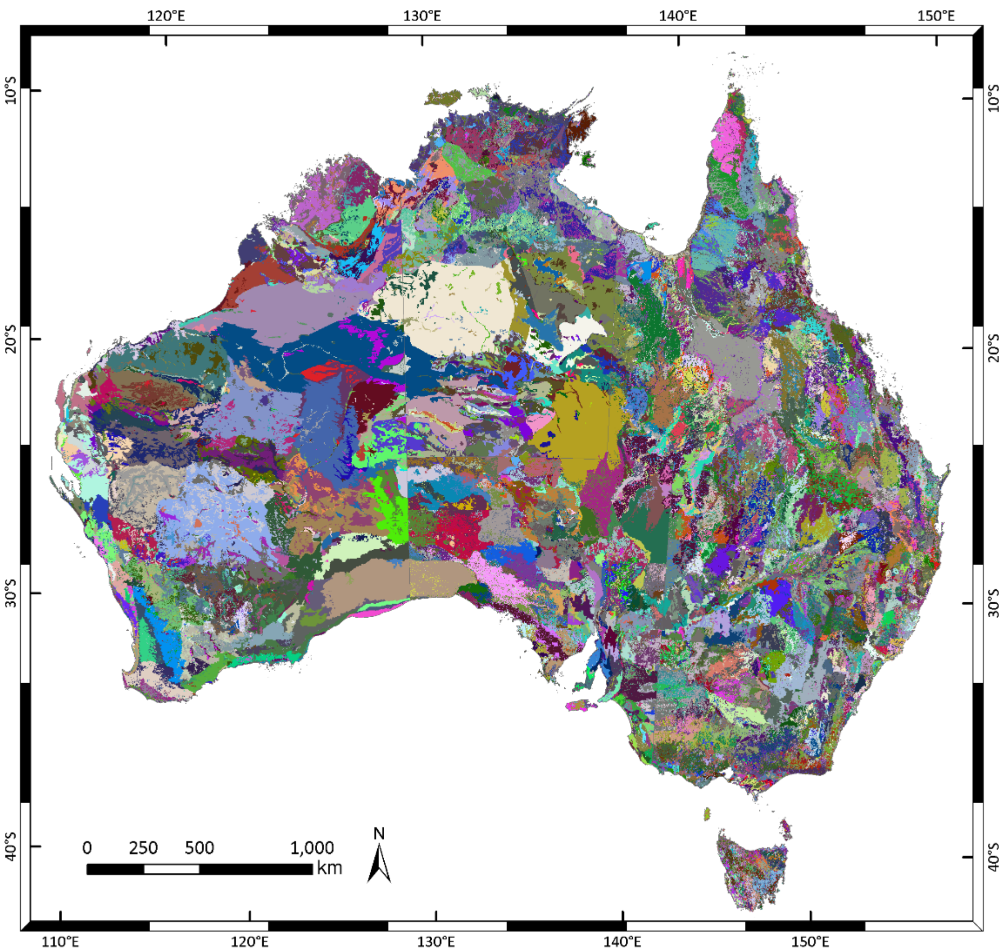


Figure S16. Spatial pattern of 7296 ecological land units defined by combining IBRA subregions version 7.0 with pre-1750 NVIS major vegetation groups version 6.0, of which c. 5481 strata contained at least one inferred reference site, derived for use as benchmarks in HCAS v2.3. Projection: Australian Albers GDA 1994.

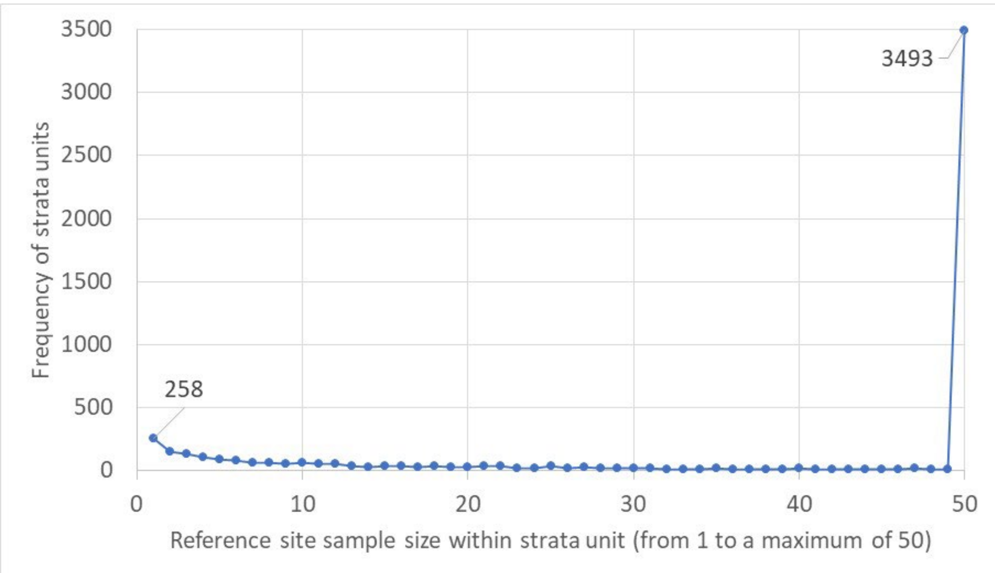


Figure S17. Frequency distribution of reference site sample sizes within ecological land units (strata). Up to 50 samples could be randomly selected from each of the 5481 strata units (36% of strata had <50 reference sites available for selection). Note, there are 1815 ecological land units with 0 reference sites.

Environmental covariates (HCAS v2.1-3)

Long-term variation in climate, soils, landforms and surface hydrology, interacting with disturbance regimes such as fire and other extreme events, are key environmental determinants of the distribution of Australia's terrestrial ecosystems. The conceptual model of drivers of natural vegetation growth, development and distribution, summarised by Guisan and Zimmermann (2000), provides a suitable basis for determining which environmental covariates to compile for biodiversity modelling (Williams et al. 2012).

The primary data sources are 30-year average (1976-2005) monthly climate variables derived using ANUCLIM version 6.1 (Xu and Hutchinson 2011, 2013) and the 0.0025 degree digital elevation model for Australia (Hutchinson et al. 2008) summarised into a series of statistics to represent long-term annual averages, extremes, and seasonality (Harwood et al. 2016b). This is complemented by data from the soil and land grid for Australia (Grundy et al. 2015, Viscarra Rossel R. A. et al. 2015). The 3-arcsecond gridded soil and terrain variables, were aggregated to 0.0025 degrees, taking into account soil depth limits (Gallant et al. 2018).

A MODIS-derived alpha-NDVI water algorithm was developed by Donohue et al. (2022) to identify inundation areas between 2001 and 2018—denoted NDVI_nfloods—as an environmental input and to mask values from the remote sensing variables when and where surface water was detected (see section on 'remote sensing variables' below). This algorithm was implemented for HCAS using the 500 m, 8-day MOD09A1 (Collection 6) reflectance data (Vermote 2015). Alpha-NDVI water grids were then aggregated from 8 to 16-day time-steps, and resampled (over-sampled) to 250 m resolution using GDALWARP with bilinear interpolation.

This resulted in an initial set of 54 environmental covariates, listed in Appendix D of Williams et al. (2020), of which 36 were considered suitable for use. An exploratory data analysis protocol, developed by Lehmann et al. (2018), was applied to these data to identify potential errors and resolve issues related to multicollinearity among related environmental covariates. This included examining pairwise correlations and choosing one of a highly correlated pair to take forward, guided by previous experience with the same data and variance inflation factors. Twenty-nine environmental covariates were selected as candidates for the projection pursuit regression variable selection and model fitting process (Table S7).

Table S7. The 36 abiotic environmental covariates (9-second gridded) considered for use in HCAS v2.1.
*** denotes the 29 candidate variables tested for inclusion in the project pursuit regression (PPR) model, 7 other variables were excluded during the exploratory data analysis. † denotes variables included in the PPR model.**

Label	Description	Units	Classification	Source
EAAS*	Budyko method mean annual atmospheric water balance	mm	Moisture	(Harwood et al. 2016b)
ADI	Aridity index - monthly minimum (precipitation/evaporation)	index	Moisture	(Harwood et al. 2016b)
ADX	Aridity index - monthly maximum (precipitation/evaporation)	Index	Moisture	(Harwood et al. 2016b)
ADM	Aridity index - monthly mean (precipitation/evaporation)	index	Moisture	(Harwood et al. 2016b)
WDI*	Precipitation deficit, monthly minimum (precipitation –	mm	Moisture	(Harwood et al. 2016b)

Label	Description	Units	Classification	Source
	evaporation with topographic adjustment)			
WDX*†	Precipitation deficit, monthly maximum (precipitation – evaporation with topographic adjustment)	mm	Moisture	(Harwood et al. 2016b)
WDA	Precipitation deficit, annual total (precipitation – evaporation with topographic adjustment)	mm	Moisture	(Harwood et al. 2016b)
PTI*†	Precipitation - monthly minimum	mm	Precipitation	(Harwood et al. 2016b)
PTX	Precipitation - monthly maximum	mm	Precipitation	(Harwood et al. 2016b)
PTA*†	Precipitation - annual total	mm	Precipitation	(Harwood et al. 2016b)
EPI*†	Evaporation - monthly minimum with topographic adjustment	mm	Evaporation	(Harwood et al. 2016b)
EPX*	Evaporation - monthly maximum with topographic adjustment	mm	Evaporation	(Harwood et al. 2016b)
EPA	Evaporation - annual total with topographic adjustment	mm	Evaporation	(Harwood et al. 2016b)
PTS1*	Precipitation seasonality - summer or winter dominated (inverse ratios)	index	Precipitation	(Harwood et al. 2016b)
PTS2*	Precipitation seasonality - spring or autumn dominated (inverse ratios)	index	Precipitation	(Harwood et al. 2016b)
TNI*	Minimum temperature - monthly minimum	°C	Temperature	(Harwood et al. 2016b)
TNX*†	Minimum temperature - monthly maximum	°C	Temperature	(Harwood et al. 2016b)
TNM	Minimum temperature - monthly mean	°C	Temperature	(Harwood et al. 2016b)
TXI*†	Maximum temperature - monthly minimum with topographic adjustment	°C	Temperature	(Harwood et al. 2016b)
TXX*†	Maximum temperature - monthly maximum with topographic adjustment	°C	Temperature	(Harwood et al. 2016b)
TXM	Maximum temperature - monthly mean with topographic adjustment	°C	Temperature	(Harwood et al. 2016b)
TRI*†	Diurnal range temperature - monthly minimum with topographic adjustment	°C	Temperature	(Harwood et al. 2016b)
TRX*†	Diurnal range temperature - monthly maximum with topographic adjustment	°C	Temperature	(Harwood et al. 2016b)
TRA	Diurnal range temperature - monthly maximum-minimum	°C	Temperature	(Harwood et al. 2016b)
NDVI_nfloods*†	Number of years inundation detected (2001-2018) using the MODIS-	Years/18	Landform	(Donohue et al. 2022)

Label	Description	Units	Classification	Source
	derived alpha-NDVI water algorithm			
TWI3S*†	Topographic wetness index (aggregated from 3-second version)	Index	Landform	(Gallant and Austin 2012c)
ELEVFR300*†	Elevation focal range within 300m moving window (aggregated from 3-second version)	m	Landform	(Gallant and Austin 2012b)
ELEVFR1000	Elevation focal range within 1000m moving window (aggregated from 3-second version)	m	Landform	(Gallant and Austin 2012a)
BDW*†	Bulk Density of the whole soil (including coarse fragments) in mass per unit volume by a method equivalent to the core method (spatially aggregated from 3-second version)	g/cm^3	Soil	(Viscarra Rossel Raphael et al. 2014j)
SOC*†	Organic Carbon as mass fraction by weight in the less than 2 mm soil material as determined by dry combustion at 900°C (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014i)
CLY*†	Clay content (2 µm mass fraction of the less than 2 mm soil material determined using the pipette method) (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014h)
SLT*†	Silt (2 - 200 µm mass fraction of the less than 2 mm soil material determined using the pipette method) (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014f)
SND	Sand (200 µm - 2 mm mass fraction of the less than 2 mm soil material determined using the pipette method) (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014k)
PHC*†	pH of 1:5 soil/0.01 m calcium chloride extract (aggregated from 3-second version)	-	Soil	(Viscarra Rossel Raphael et al. 2014e)
AWC*†	Available water capacity computed for each of the specified depth increments (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014d)
NTO*†	Total Nitrogen (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014c)
PTO*†	Total Phosphorus (aggregated from 3-second version)	%	Soil	(Viscarra Rossel Raphael et al. 2014g)
ECE*†	Effective Cation Exchange Capacity extracted using barium chloride ($BaCl_2$) plus exchangeable H + Al (aggregated from 3-second version)	meq/100g	Soil	(Viscarra Rossel Raphael et al. 2014a)

Label	Description	Units	Classification	Source
DER*†	Depth of Regolith - The regolith is the *in situ* and transported material overlying unweathered bedrock (aggregated from 3-second version)	m	Soil	(Wilford et al. 2015)
DES*†	Depth of soil profile (A & B horizons) (aggregated from 3-second version)	m	Soil	(Viscarra Rossel Raphael et al. 2014b)

Remote sensing variables (HCAS v2.1-3)

Seven remote sensing variables were derived as summaries of four MODIS Collection 6 vegetation products using satellite imagery generated between 1st January 2001 and 31st December 2018. These variables derive from four remote sensing products: persistent and recurrent green foliage fractions developed using the method of Donohue et al. (2009); and bare ground and litter cover fractions developed using the method of Guerschman and Hill (Guerschman 2019, Guerschman and Hill 2018). All variables have possible values between 0 and 1 as units of ground cover proportion. The persistent green cover fraction is mainly derived from perennial plant species (e.g., non-deciduous shrubs and trees) and the recurrent fraction is derived from annual species (e.g., grass and herbage, deciduous shrubs and trees). The litter fraction comprises non-photosynthesising plant material and the bare ground fraction is the ground not covered by litter or green foliage.

Persistent and recurrent green foliage fractions derive from the MODIS 16-day, 250 m NDVI data product (Collection 6), MOD13Q1 (Didan 2015). The original MODIS sinusoidal tiles were reprojected to geographics and compiled into a continental image using GDALWARP (Warmerdam et al. 2021). This was done using nearest neighbour resampling. The internal MODIS pixel reliability flag was used to remove any NDVI values deemed to be of low quality (that is, a reliability score of 2 or above). Following Roderick et al. (1999), total fractional cover (F) was derived by rescaling NDVI (V) between the bare ground value (V_n) and the full cover value (V_x) using equation 1.

$$F = 0.95 \frac{V - V_n}{V_x - V_n} \quad (1)$$

The full cover value was determined by identifying the maximum NDVI value for each pixel through the entire 2001-2018 period. The 95th percentile of these maximum values was set to be V_x. The bare ground value was determined by first identifying the minimum NDVI value for each pixel through the entire 2001-2018 period. To remove speckle in this minimum dataset, it was smoothed using a boxcar average with a width of 3 pixels. According to Montandon and Small (2008), real bare ground NDVI values can range between 0.05 and 0.40. Hence, all V_n values less than 0.05 were set to 0.05. However, inspection of the minimum NDVI grid across Australia showed that minimum NDVI values of 0.40 only occurred in places that have reasonably high woody foliage cover (and hence this value couldn't reasonably be expected to represent a pure bare ground signal). Hence, the upper soil NDVI value for Australia was identified as the largest value in the minimum NDVI grid from locations outside of woodlands and forests, as defined by NVIS present Major Vegetation Groups version 5.1 (DAWE 2018). This gave an upper limit to V_n of 0.225 and all values above this were set to 0.225. Total fractional green cover was split into its persistent and recurrent components using the method of Donohue et al. (2009). This effectively runs a low-pass filter through each pixel's 18-year timeseries, setting this to the persistent component. The difference between total and persistent becomes the recurrent component.

For both persistent/recurrent green foliage fractions and bare ground/litter cover fractions, data depicting surface water, snow cover and 'sea' (the latter in the coastal-land corridor) were used to mask pixels. This had the effect of removing these values from 8-day or 16-day time series resulting in some pixels having fewer samples when calculating the summary variables. The existing snow/ice QA flags of 4096 or 32768 present within the quality control attribute of the 8-day MOD09A1 500 m reflectance data (Vermote 2015), and any cloud above 1500 m elevation (cloud cover flag QA = 1024), were used to mask snow cover. The alpha-NDVI water algorithm (Donohue et al. 2022) was used to identify inundation zones in the 500 m, 8-day MOD09A1 reflectance data (Vermote 2015), and to mask surface water. The surface water and snow cover masks were aggregated to 16-days and resampled (over-sampled) to 250 m using GDALWARP (https://gdal.org/programs/gdalwarp.html#gdalwarp) with bilinear interpolation.

Each time-series cover fraction product was summarised using two statistics. The long-term average value was calculated from the annual means of the 16-day (or 8-day) values across the whole 18-year period. The average intra-annual maximum was calculated as the overall average of the maximum value recorded in each of the 18 years. The intra-annual maximum statistic was chosen because it is highly correlated with the intra-annual range whereas the minimum statistic is not. The mean and maximum statistics so derived for the persistent green foliage fraction were found to be 99% correlated, and therefore only the long-term average statistic was carried forward. Seven remote sensing variables were thus derived to characterise ecosystems (Table S8).

Table S8. Remote sensing time-series products (2001-2018) and summary variables used in HCAS v2.1-3.

Variable	Description	Summary metrics	Original spatial resolution	Source
Persistent green cover fraction	The fraction of ground covered by green foliage of persistent (~perennial) species	Long-term average	250 m	Donohue et al. (2009)
Recurrent green cover fraction	The fraction of ground covered by green foliage of recurrent (~annual) species	Long-term average Average intra-annual maximum	250 m	Donohue et al. (2009)
Litter cover fraction	The fraction of ground covered in non-photosynthesising plant material (litter)	Long-term average Average intra-annual maximum	500 m	(Guerschman 2019, Guerschman and Hill 2018)
Bare ground fraction	The fraction of ground not covered in green foliage or plant litter	Long-term average Average intra-annual maximum	500 m	(Guerschman 2019, Guerschman and Hill 2018)

Principal components of remote sensing variables (HCAS v2.1-3)

The HCAS benchmarking algorithm requires remote sensing variables of comparable scaling to ensure dimension consistency in calculating the Manhattan distances used in measuring ecosystem condition as the proximity to reference. The principal components (PCs) of all seven remote sensing variables were therefore derived using a principal components analysis (PCA). Since the units of all variables are proportions in the range 0-1 (Table S9), they were mean-centred but not range-standardised prior to running the PCA.

The PCA results are shown in Figure S18 and eigenvectors in Table S10. The first PC is dominated by the two bare cover fraction variables with the largest eigenvector values (maximum and mean ~ -0.58 each). The second PC is dominated by maximum litter and mean persistent green cover fractions and the third PC is dominated by maximum recurrent fraction. The mean recurrent fraction dominates the seventh and final PC in the series (-0.91 eigenvector value).

The order of the PCs reflects their numerical magnitudes (see Table S11), and therefore their relative influence in the Manhattan distance calculation of the condition algorithm, after first dividing by the scaling factor of 1000. The first two PCs carry much more weight than the latter 5 PCs (represented by the scree curve in Figure S18). The purpose of the PCA is not to reduce input dimensions. All PCs were used in the Manhattan distance calculation.

Table S9. Summary grid statistics for the seven remote sensing variables used in HCAS v2.1-3

Units are fractional cover (0-1) scaled by 1000 as integer data.

Label	Variable	Minimum	Maximum	Mean	Standard deviation
Mean_per	Long-term average persistent green cover fraction	0	930	117.14	134.23
Mean_rec	Long-term average recurrent cover fraction	0	420	81.46	62.03
Max_rec	Average intra-annual maximum recurrent cover fraction	0	866	207.18	158.52
Mean_ltt	Long-term average litter cover fraction	0	1000	424.34	100.71
Max_litt	Average intra-annual maximum litter cover fraction	0	1000	572.54	129.96
Mean_bare	Long-term average bare ground fraction	0	1000	339.15	182.49
Max_bare	Average intra-annual maximum bare ground fraction	0	1000	470.88	188.53

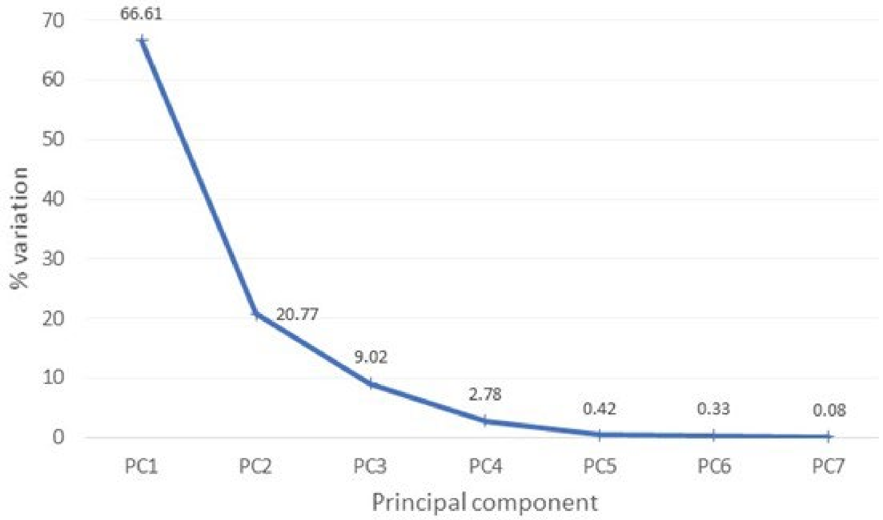


Figure S18. Scree plot for the seven principal components showing the variation captured in each

Table S10. Eigenvectors for the principal components of the seven remote sensing variables

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Average intra-annual maximum bare ground fraction	-0.5797052	-0.2485049	0.2457619	0.4641782	0.0112112	0.5617683	-0.1030418
Long-term average bare ground fraction	-0.5833303	-0.1144457	0.1556429	0.0271099	-0.2652325	-0.7421303	-0.0237185
Average intra-annual maximum litter cover fraction	0.2700794	-0.5292743	-0.1656900	0.5404358	0.4591000	-0.3140426	0.1341872
Long-term average litter cover fraction	0.1395264	-0.3841712	-0.5287682	0.0944034	-0.6897477	0.0964259	-0.2436947
Long-term average persistent cover fraction	0.2443486	0.5906739	0.1405578	0.6851888	-0.2766536	-0.1266414	-0.0978908
Average intra-annual maximum recurrent cover fraction	0.3824232	-0.3703122	0.7207694	-0.0710623	-0.3618178	0.0265347	0.2458592
Long-term average recurrent cover fraction	0.1590705	-0.1067798	0.2627970	-0.0910344	0.1885521	-0.0948476	-0.9172770

Table S11. Summary statistics for the seven remote sensing principal components.
Input variables were in units of fractional cover (0-1) scaled by 1000 as integers.
Data were rescaled prior to calculation of Manhattan distances.

Principal component	Minimum	Maximum	Mean	Standard deviation
PC1	-10130.07	802.11	-0.02	307.91
PC2	-646.57	1061.88	0.00	171.93
PC3	-736.05	833.34	0.00	113.29
PC4	-656.13	765.52	0.00	62.88
PC65	-294.44	426.37	-0.03	24.39
PC6	-246.22	474.96	0.00	21.55
PC7	-180.62	163.02	-0.01	10.69

Spatial mask for input data

The remote sensing variables used in HCAS v2.1-3 were designed to characterise only terrestrial environments. Therefore, all inputs (environmental and remote sensing variables, reference sites) were masked to ensure consistent removal of semi-permanent or permanent water bodies – treated as 'no data' cells. An aggregated water presence threshold of >80% was derived from the 25 m annual Landsat water observations from space (WOFS) dataset (Mueller et al. 2016), summarised as the frequency over the period 2001-2014 (Geoscience Australia 2015), then majority resampled to match the geographic 0.0025 degree grid and datum adopted for HCAS (GDA 94). This dataset was used to mask all input data, and as the extent layer for all spatial processing operations. The base grid derives from the 9-arcsecond digital elevation model (Hutchinson et al. 2008).

Predicting reference ecosystem characteristics (HCAS v2.1-3)

Projection pursuit regression was used to model the fit of the seven remote sensing PC response variables to the 29 candidate environmental covariates using the training data sample of 101,686 reference sites. Standard steps of PPR model selection and candidate covariates testing were applied on the basis of a prediction error metric (10-fold cross-validation residual sum of squares). The k-fold cross-validation helped determine which smoothing algorithm performs best in general and, therefore, which value of the smoothing parameter leads to the best overall model as a function of the number of PPR terms, as explained in Lehmann et al. (2018).

Covariate selection was performed in a forward, backward and bidirectional manner, for a total of 250 tested models, each starting from a random set of candidate environmental covariates. Of the 23 environmental predictors included in the best performing PPR model, nine were climate variables, two were terrain features, one surface water, and 11 soil attributes (Table S12 lists their ranked importance). The best overall PPR model (Table S13) was used to predict the seven remote sensing PCs as a function of the 23 environmental predictors to characterise the ecosystem reference state for each land pixel in the analysis mask for the Australian continent. Appendix G in Williams et al. (2021) shows the pattern of residuals and mapped outputs for each PC.

The resulting frequency distribution between observed versus predicted distances for a random sample of 100,000 reference site pairs is shown in Figure S19. When calculating predicted distances, the noise component of the model results in an overall bias factor with non-zero mean, causing the offset parallel to the 45 degree line. This offset is simply a mathematical by-product from the formula used to calculate distances. Each remote sensing PC variable is modelled to match the observed PC values, but has some additive noise as a result of the modelling process. That is: $PCvar = f(ENVvars) + \varepsilon$, where ε is the noise component (assumed zero-mean Gaussian). When calculating distances (Euclidean or Manhattan), the formula uses the square (or absolute value) of PCvar. On the basis of this, it can be shown that when PCvar is modelled as above, the noise component (ε) will result in a bias term that, as a result, leads to the offset in the distance plot.

Conceptually, some of the noise component, ε, likely can be attributed to specific processes such alternative ecological states and seasonal variation for the same environment, as well as inherent error in reference site assignments and error in other inputs. Overall, we expect more variability in observed remote sensing PCs due to natural ecosystem dynamics, such as alternate ecological states and seasonal dynamics, than can be represented by predicted PCs from environmental covariates that represent a long-term steady state.

Table S12. The 23 selected abiotic environmental covariates included in the PPR model.
Covariates (predictors) are ordered by relative importance based on cumulative (absolute) sum of loadings (% loadings).

Label	Description and units	Ranked relative importance
TXI	Maximum temperature – monthly minimum with topographic adjustment (°C)	1
EPI	Potential evaporation – monthly minimum with topographic adjustment (mm)	2
TNX	Minimum temperature – monthly maximum with topographic adjustment (°C)	3
TXX	Maximum temperature – monthly maximum with topographic adjustment (°C)	4
PTA	Precipitation – annual total (mm)	5
SOC	Organic Carbon as mass fraction by weight in the less than 2 mm soil material as determined by dry combustion at 900°C (%)	6
WDX	Precipitation deficit, monthly maximum (precipitation – evaporation with topographic adjustment) (mm)	7
TRX	Diurnal range temperature – monthly maximum with topographic adjustment (°C)	8
PTI	Precipitation – monthly minimum (mm)	9
PHC	pH of 1:5 soil/0.01 m calcium chloride extract (index)	10
BDW	Bulk Density of the whole soil (including coarse fragments) in mass per unit volume by a method equivalent to the core method (g/cm^3)	11
TRI	Diurnal range temperature – monthly minimum with topographic adjustment (°C)	12
NTO	Total Nitrogen (%)	13
DES	Depth of soil profile (A & B horizons) (m)	14
SLT	Silt (2 – 200 µm mass fraction of the less than 2 mm soil material determined using the pipette method) (%)	15
ECE	Effective Cation Exchange Capacity extracted using barium chloride ($BaCl_2$) plus exchangeable H + Al (meq/100g)	16
ELEVFR300	Elevation focal range within 300m moving window (m)	17
NDVI_nfloods	Number of years detected inundation for the period 2001-2018 (years) using the MODIS-derived alpha-NDVI water algorithm	18
CLY	Clay content (2 µm mass fraction of the less than 2 mm soil material determined using the pipette method) (%)	19
PTO	Total Phosphorus (%)	20
DER	Depth of Regolith – The regolith is the *in situ* and transported material overlying unweathered bedrock (m)	21
AWC	Available water capacity computed for each of the specified depth increments (%)	22
TWI3S	Topographic wetness index (index)	23

Table S13. Summary fit statistics for the seven remote sensing PCs observed versus PPR predicted values.

Principal component	R-squared	Pearson's correlation coefficient
1	0.837	0.915
2	0.839	0.916
3	0.291	0.540
4	0.372	0.610
5	0.173	0.418
6	0.123	0.351
7	0.160	0.400
Overall	0.631	0.784

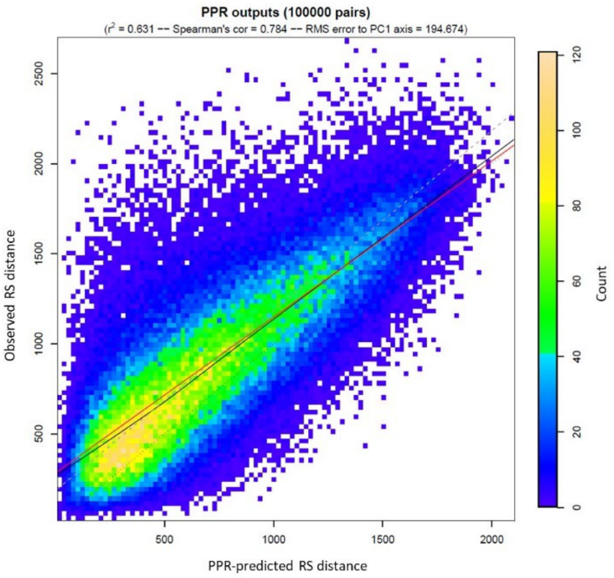


Figure S19. PPR model fit in terms of observed versus predicted remote sensing principal component Euclidean distances for HCAS v2.1-3.
A random sample of 100,000 reference site-pairs (of the N×(N-1)/2 combinations, N = 101,686) are used for computational tractability. Red line is a linear model fit of the data; black line is a smoothing fit of the data; dashed grey line is the diagonal. X and Y axis units are in multiples of 1000 corresponding with the integer rescaling of the input remote sensing variables.

Estimating ecosystem condition

A founding principle of HCAS is that we expect considerable variation in the remote sensing characteristics of natural ecosystems in reference condition within the same abiotic environments, due to alternative ecological states, seasonal dynamics and various stages of recovery following natural disturbances and site history. However, it is challenging to comprehensively represent all possible reference dynamics because physical observations in time and space are limited. Due to this natural variability, reference sites within any given abiotic environment can support quite different natural ecosystems despite similar abiotic environments; and, conversely, ecosystems in different abiotic environments can share similar remote sensing signatures. For example, closed canopy vegetation can look quite consistent regardless of canopy height, and so forests and closed heathlands share many of the same remotely sensed ecosystem characteristics. Further, at 250 m grid resolution, sharp discontinuities in ecosystem structure, function, and composition may be mixed within a single pixel. The HCAS modelling framework is designed to tackle both types of inherent variation and limited reference data by using distance-comparison measures in selecting benchmarks and estimating condition. The approach taken doesn't require every site of interest to have benchmark reference sites in an equivalent abiotic environment.

Ecosystem condition is estimated as the proximity to reference condition using Manhattan distances derived from reference-reference and test-reference site-pairs. The expected condition is established using a database of reference-reference site pairs, and then condition is estimated using the position of test-reference site-pairs on the database of reference-reference site pairs (see Lehmann et al. (2021) for a schematic description of how the benchmarking algorithm works).

For the expected condition surface, two sets of Manhattan distances are derived for each reference-reference site-pair using the training data attributed with values of the seven 1) observed and 2) predicted remote sensing PCs. The PCs were first rescaled to their proper dimensions by dividing by 1000. A two-dimensional frequency histogram of these observed (d_o, y-axis) versus predicted (d_p, x-axis) distances simulates a probability density surface of the ecosystem reference state. A convenient bin size, Z, is selected to approximate a 600 x 600 matrix depending on the distance range (e.g., 0.005 in the case of HCAS v2.1-3) within which the frequency of site-pairs is summarised as counts (i.e., likelihoods of being in reference condition). The counts within the reference-distance density surface are normalised within each bin of the x-axis (d_p, predicted distances), then smoothed using bilinear interpolation (Moore neighbourhood at 0.005) to fill gaps due to scarce data, and finally truncated to remove irrelevant large distances (d_p, d_o) to approximate a 400 x 400 matrix. The expected condition values provided by the reference-distance density surface (p_{ref}) are here termed 'probabilities' but are not true probabilities in the statistical sense; they are normalised frequency counts as a density surface.

To estimate condition for each test site (approx. 111 million in HCAS v2.1-3), two sets of test-reference Manhattan distances are first calculated using the sample of reference sites for testing as benchmarks (B_{ref}) attributed with values of the seven 1) observed and 2) predicted remote sensing PCs, after rescaling to their proper dimensions. These observed (d_o, y-axis) and predicted (d_p, x-axis) distances are plotted over the reference-distance density surface (described above) to derive expected probabilities. The next step involves determining which and how many reference sites are relevant to use as benchmarks for each test site. A nested set of parameters guided these decisions. These parameters are: 1) the maximum geographic radius (R, km) around each test site used to search for relevant reference sites, 2) within that radius, the maximum number of reference sites (n_p) closest in distance to the test site based on test-reference predicted distances (d_p, x-axis), 3) the maximum number of reference sites (n_{ref}) closest in distance to the test site, of previously selected n_p, based on test-reference observed distances (d_o, y-axis) with highest reference probability density values (p_{ref}) to use as benchmarks (henceforth test-benchmarks comparisons), 4) a half-Cauchy distance-decay function (Shaw 1995) using the median Manhattan distance ($\boldsymbol{\lambda}$) to down-weight selected n_{ref} with increasing test-benchmark predicted distances (d_p, x-axis), and 5) the confidence parameter, ω, used in a limited degree of confidence calculation (LDC) with the maximum probability value, P_{max}, of the selected n_{ref}, to deal with potential uncertainty in reference site validity as suitable benchmarks. Global parameter values were determined following exhaustive, iterative exploration of different settings, reported in Section 6.6 of Williams et al. (2020) and Section 3.7.4 in Williams et al. (2021).

The above parameters were applied as shown in Equations 2 and 3, and as listed in Table S14. To summarise (see also Box S2), 50 benchmarks (n_p) within 200 km radius (R) of each test site are initially selected on the basis of their predicted distance to the test site (d_p, x-axis) being minimised (i.e., likelihood of being of the same ecosystem type to address context dependency), from which 20 benchmarks reference sites (n_{ref}) are selected that maximise the likelihood of actually being a reference site based on their position on the reference-distance density surface (p_{ref}). Condition of the test site is then calculated as the predicted distance half-Cauchy decay ($\boldsymbol{\lambda}$ = median of d_p) weighted average of the 20 test-benchmark (n_{ref}) probabilities (p_i) of being in reference condition using a half-weight (ω = 0.5) LDC algorithm uncertainty (Figure S20). The sample size of 20 benchmarks represents a trade-off between context dependency and the need to account for multiple expressions of an ecosystem reference state (i.e., the challenge of alternative ecological states and seasonal dynamics). The output probabilities of being in reference condition have a numerical range influenced by the bin size of the reference distance density surface, and needs to be calibrated between 0.0 (lowest – ecosystem integrity extinguished) and 1.0 (highest – ecosystem integrity in reference condition).

$$w_i = \frac{1}{\pi\left(1 + \left(\frac{d_p}{\boldsymbol{\lambda}}\right)^2\right)} \tag{2}$$

$$H_c^{LDC} = \omega \cdot \left(\left(\frac{\sum p_i w_i}{\sum w_i}\right) + p_{max}\right) \text{ for } \omega \equiv 0.5 \tag{3}$$

Table S14. Equation notations and definitions used in HCAS v2.1-3 benchmarking algorithm

Parameter notation	Parameter value	Definition
$\boldsymbol{d_p}$	NA	The predicted Manhattan distance between all pairs of reference sites B_{ref} (c. 200,000) used to construct the reference-distance density surface (and subsequently also calculated for relevant benchmark reference-test site comparisons).
$\boldsymbol{d_o}$	NA	The observed Manhattan distance between all pairs of reference sites B_{ref} (c. 200,000) used to construct the reference-distance density surface (and subsequently also calculated for relevant benchmark reference-test site comparisons).
$\boldsymbol{B}_{\mathbf{ref}}$	NA	The set (representative sample) of reference sites (c. 200,000) as dynamic benchmarks; used in the calculation of the condition metric for a test site (benchmarking, using a subset of B_{ref}).
$\boldsymbol{Z}$	0.005	The distance bin size used in the reference-distance density surface applied equally to the x-axis and y-axis to derive the reference-distance density surface.
$\boldsymbol{i}$	NA	Denotes an individual benchmark (reference) site used for the calculation of the condition metric at a test site, $i = 1, \ldots, n_{\mathrm{ref}}$. The (predicted and observed) distances between a reference site and test site are denoted d_p^i and d_o^i.
$\boldsymbol{p}_{\mathbf{ref}}$	NA	The reference-distance density surface in which the x-axis is defined by predicted distance d_p and the y-axis is defined by observed distance d_o between pairs of reference sites; normalised within bins of the predicted distance d_p axis.
$\boldsymbol{n_p}$	50	The initial most analogous reference sites (smallest predicted distance d_p) within a geographic radius R (km), selected on the basis of their predicted distance d_p to the test site.
$\boldsymbol{R}$	200	A constant geographical search radius (km) from the test site within which reference sites are selected for assessment as benchmarks.

Parameter notation	Parameter value	Definition
$\boldsymbol{n}_{\mathbf{ref}}$	20	The final set of most analogous reference sites (benchmarks) selected (from the initial n_p reference sites) as the subset with highest probability values on the reference-distance density surface, p_{ref}, and used to benchmark the condition of the test site.
$\boldsymbol{w}$	NA	The weights (relative contribution) of each of the benchmark-to-test site probabilities to the calculation of the (uncalibrated) condition of the test site, calculated based on a half-Cauchy decay (Shaw 1995) with the median predicted distance, d_p, from the reference-distance density surface, p_{ref}.
ω	0.5	The confidence parameter (0.5) used in the Limited Degree of Confidence (LDC) calculation applied with $\boldsymbol{p}_{\mathrm{max}}$.
$\boldsymbol{\lambda}$	2.0	The median Manhattan distance on the predicted distance d_p axis used in the half-Cauchy decay calculation of weights ($\boldsymbol{w}$) applied to test-reference site comparisons determining contributions in calculating proximity to reference.
$\boldsymbol{p}$	NA	The pairwise benchmark-to-test site comparison probability, calculated for each of n_{ref} (20) benchmark site comparisons with a single test site.
$\boldsymbol{p}_{\mathbf{max}}$	NA	The maximum probability value from the reference-distance density surface, p_{ref}, achieved among the n_{ref} pairwise benchmark-to-test site comparisons; this value is given half the weight in the LDC calculation of the condition metric for the test site.
$\boldsymbol{H_c^{LDC}}$	NA	The initial uncalibrated condition score of the test site being in reference condition, representing the (half-Cauchy) weighted mean of the normalised probabilities calculated on the basis of n_{ref} benchmark-to-test site comparisons, and incorporating a Limited Degree of Confidence (LDC) calculation.

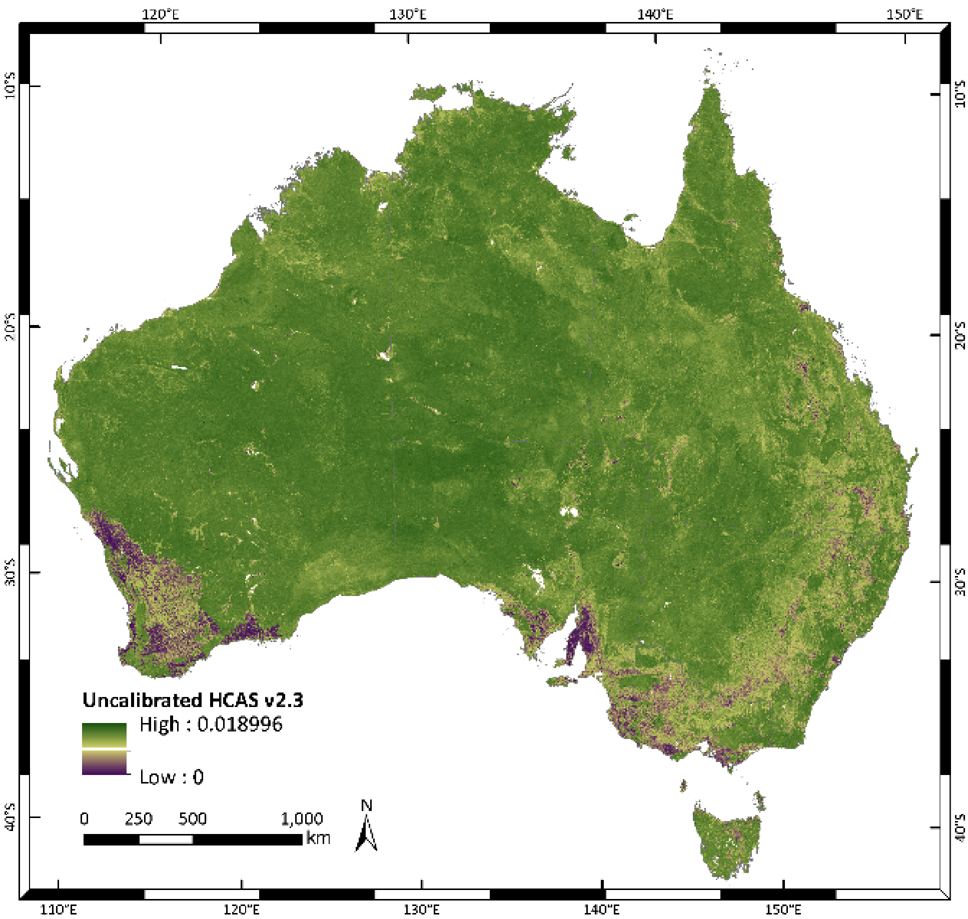


Figure S20. Uncalibrated HCAS v2.1-3 for the base model (2001-2018). Projection: Australian Albers, GDA 1994.

Similarity index for reference sites used as benchmarks in HCAS v2.3

A relative index of certainty (as opposed to uncertainty) was derived for each site of interest (250 m raster pixel) that integrates how 'nearby' and how relevant were the 20 reference sites used as benchmarks to estimate HCAS ecosystem condition. For each test site, the much larger sample of reference sites were distilled into a final set of 20 benchmark sites predicted to be the most ecologically relevant to the test site and therefore of similar ecosystem type. Their relative similarity to the test site informs their weighting as individual contributions to the empirical benchmark used in calculating the test sites' condition.

The contribution of the probability that the test site is in reference condition (calculated from the probability density surface for each reference site comparison) is weighted by a half-Cauchy distribution (median $d_p = 2$). The sum of these weightings represents the cumulative environmental similarity of all reference sites used in the calculation of the (unscaled) condition probability index for each test site. As such, this is a measure of confidence in the final selected set of test-benchmark comparisons, since higher weight is given to more similar sites (lower d_p). This summed weighting from the base model was recorded for each test site and is best interpreted as a relative rather than an absolute measure. Given the very long tail of the summed weightings' output distribution, for all locations continent-wide, the natural logarithm of the resulting dataset was used to increase resolution at lower summed weightings (Figure 21). This measure shows regions where the model is most limited by reference sites, using the current sample structure.

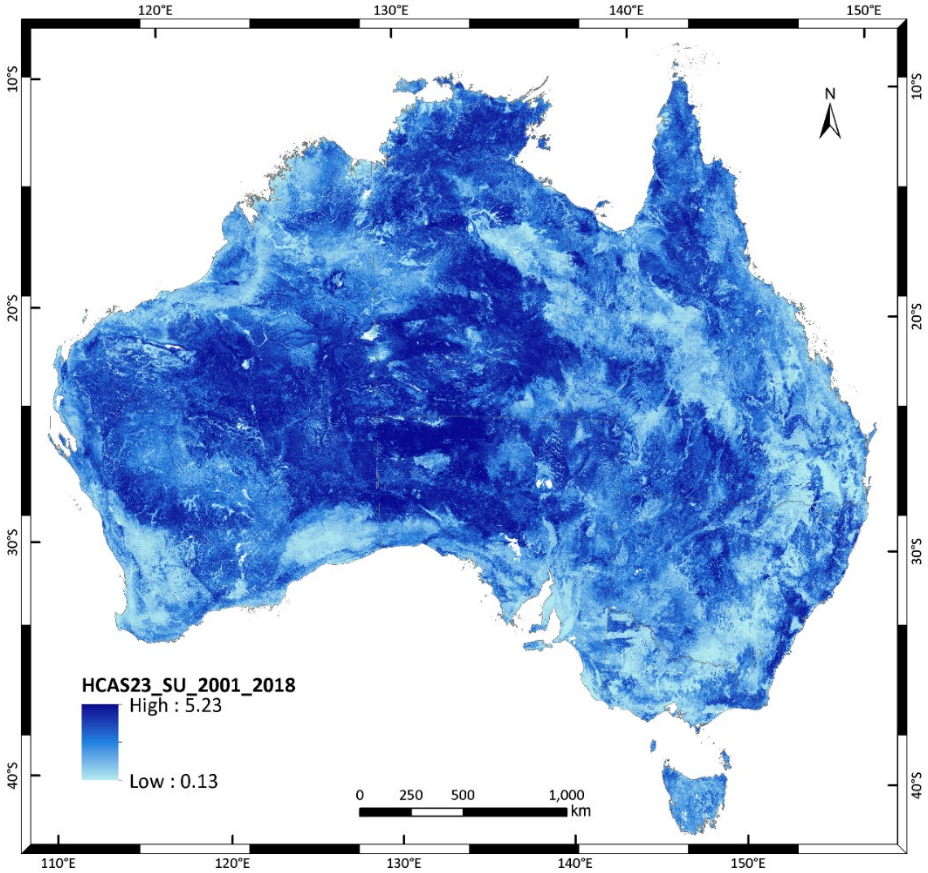


Figure 21. HCAS v2.3 test-benchmark similarity index (scaled as certainty).
The ESRI legend stretch 'histogram equalise' is used, which spreads values across the histogram equally to emphasise heterogeneity. This analysis shows the sum of environmentally similar weightings for 20 reference sites used in the HCAS v2.3 condition calculation for each site of interest. The output is rescaled by the natural logarithm of values. Lower values imply benchmark reference sites are less similar to the site of interest. Projection: Australian Albers, GDA 1994.

Calibrating ecosystem condition (scaling, 0-1)

Calibration requires independent data to inform scaling of the initial HCAS output between 0.0 and 1.0. Previewing the uncalibrated index (Figure S20) compared with land use mapping (Figure S22), it is clear that condition values are higher than expected in areas of intensive land use (e.g., dryland cropping). Scaling should result in lower condition scores in intensive land use areas, relative to natural areas (e.g., nature conservation and managed resource protection). The approach to scaling also needs to be conceptually consistent with use of the data as an input to habitat-based biodiversity assessments. For that purpose, each location (i.e., 250 x 250 m pixels in HCAS v2.1-3) represents the effective proportion of habitat available to biodiversity as if it were in reference condition; where a condition score of 0.10 in a 250 x 250 m area is treated as equivalent to 25 x 25 m habitat in reference condition, on average (e.g., see application by Mokany et al. 2022). In this context, HCAS condition scores can be viewed as an 'area' axis of the species-area relationship (i.e., x-axis). The species-area relationship (SAR) then describes how, as area of (assumed) contiguous, intact habitat increases to a maximum (i.e., all locations are in reference condition), the number of species that can persist in that type of habitat increases (Rosenzweig 1995). We therefore consider the role of the species-area relationship, along with other lines of evidence such as meta-analyses of species compositional responses to disturbance (e.g., Chaudhary and Brooks 2018, Hudson et al. 2017, Newbold et al. 2012), in our approach to HCAS calibration.

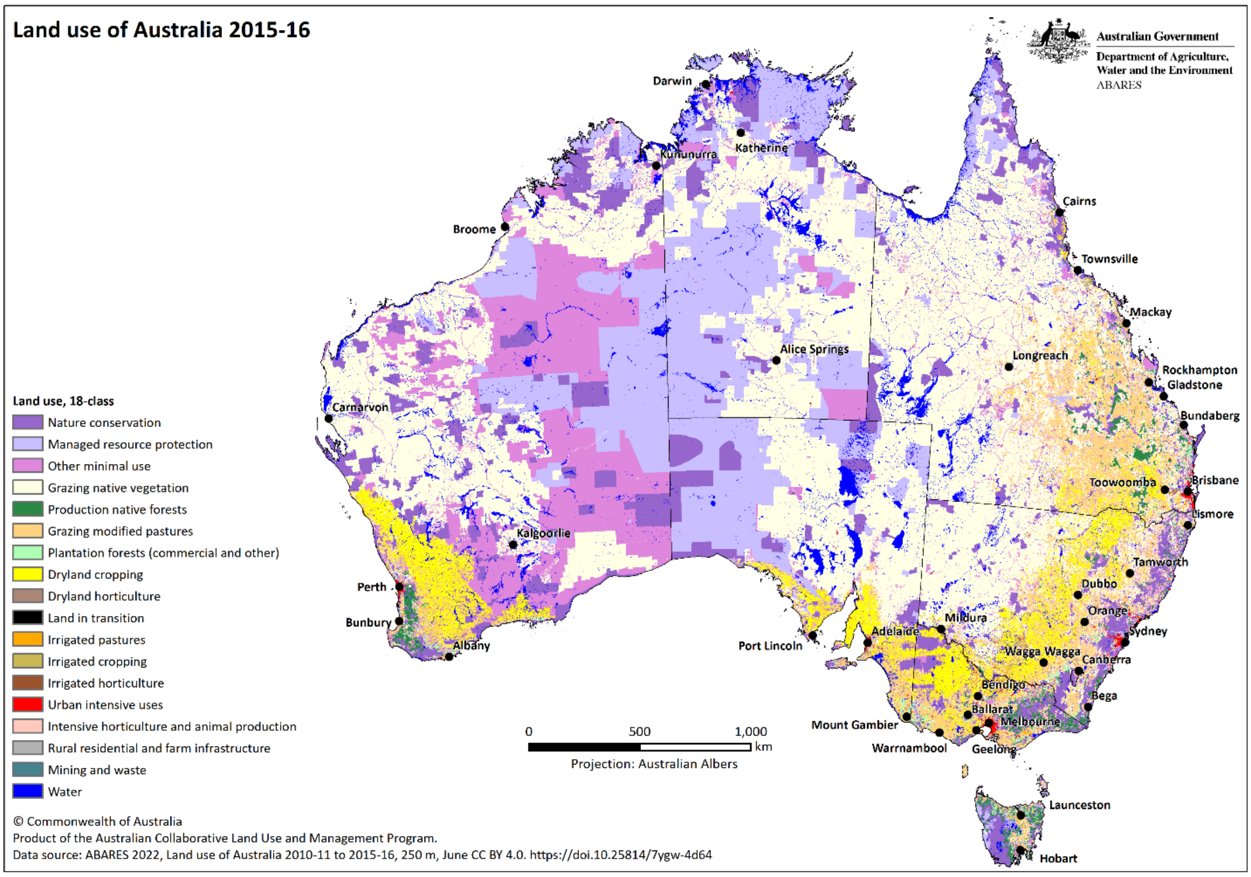


Figure S22. National level land use of Australia for 2015–16 (ABARES 2022), summarised into 18 classes.

PREDICTS database

We used the global meta-analysis by Hudson et al. (2017) for consistency with use of that data in global habitat-based biodiversity assessments (Ferrier et al. 2020, Hoskins et al. 2020). The PREDICTS global database (Hudson et al. 2017) summarises observations of native species numbers found at a location in a modified ecosystem state (due to anthropogenic land use) compared with similar locations in a reference state (i.e., primary vegetation), from a wide range of field experimental and observational studies in different ecosystems (Purvis et al. 2018). The data has been standardised as proportions of native species occurring in a particular land use type (relative to the original set of high integrity ecosystems). Hoskins et al. (2020) used PREDICTS data to rescale globally harmonised land use (Hurtt et al. 2011) and derived a global condition index based on proportion of native species in reference condition, compared with number of native species in each land use class, for various taxonomic groups (noting that different native species may occur in land use and reference).

The PREDICTS project aims to quantify effects of land use on species richness (De Palma et al. 2021, Purvis et al. 2018). Summaries of this data as proportion of native species relative to reference by land use class can be extracted from the PREDICTS database for each of 12 globally harmonised land use classes version 2 (LUH2: Chini et al. 2020, Hurtt et al. 2020). LUH2 data underpin the Shared Socio-Economic Pathways (SSPs) used in global integrated biodiversity modelling and climate change modelling impact analyses (Popp et al. 2017).

Species-level composition data are only one of many ecosystem attributes commonly observed as indicators of condition (Parkes et al. 2003) and, alone, do not address many important habitat specific measures considered significant for assessing condition for biodiversity. Despite these limitations, species compositional outputs from the PREDICTS database have been used as a proxy of habitat quality for biodiversity (Ferrier et al. 2020, Hoskins et al. 2020). For this purpose, we first back-transform the PREDICTS coefficients using the species-area relationship (SAR) with z-value of 0.25 to approximate the scaling of a condition score (Table S15). A z-value of 0.25, as shown in Figure S23, ensures consistency with global applications of PREDICTS data (Ferrier et al. 2020, Hoskins et al. 2020) as a generic measure for biodiversity at the site level.

The higher than expected SAR-transformed PREDICTS coefficients associated with urban land uses (0.69 in Table S15) could reflect global variability in urban and peri-urban environments that support retention of, or attract, some native biodiversity. The PREDICTS project team are continuing to aggregate source data and revise their analysis, including greater granularity and alignment with global and regional land use mapping. Therefore, scaling parameters for HCAS may be updated in line with revisions to PREDICTS project's outputs.

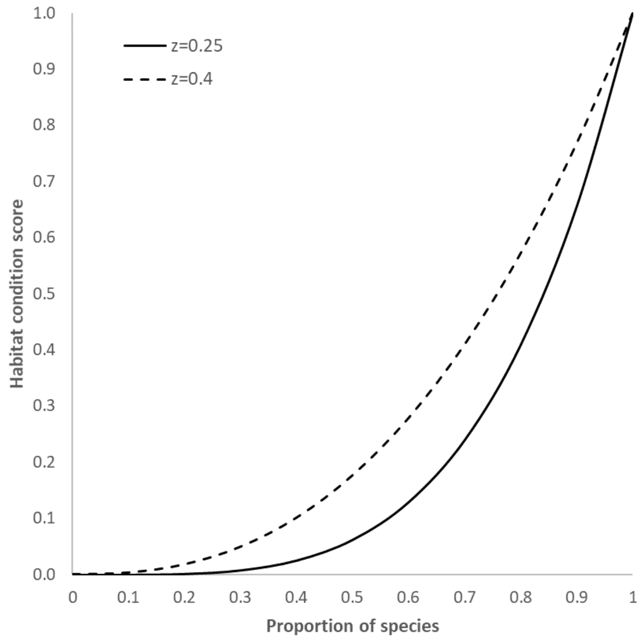


Figure S23. Schematic showing how habitat (ecosystem) condition relates to the proportion of native species (*p*) raised to the power of 1/z ($p^{1/z}$) for two typical z values (0.25, 0.4), founded on the species-area relationship (here, inverted).

Table S15. Species-area relationship back-transformation of PREDICTS coefficients (i.e., proportion of native species in an intact landscape which are found in paired modified habitats of that type) (Hudson et al. 2017) to derive a PREDICTS condition score for each of the 12 land use types (transformed using a *z*-value of 0.25). Condition scores marked with * were converted into spatially continuous surfaces based on the age of secondary vegetation and grazing density. Land use classes derive from the Global Harmonised Land Use version 2 dataset for 2015 (LUH2) (Chini et al. 2020, Chini et al. 2021a, Chini et al. 2021b, Hurtt et al. 2020).

Id	Global land use class (LUH2)	PREDICTS coefficient	Condition score (z = 0.25)
1	Primary vegetation	1.00	1.00
2	Secondary mature vegetation	0.91	*0.70**
3	secondary intermediate vegetation	0.79	*0.38**
4	secondary young vegetation	0.76	*0.33**
5	Rangelands	0.74	*0.30**
6	C3 perennial crop	0.69	0.23
7	C4 perennial crop	0.69	0.23
8	Urban	0.69	0.22
9	Pasture	0.57	*0.10**
10	C3 annual crop	0.53	0.08
11	C4 annual crop	0.53	0.08
12	C3 nitrogen-fixing crop	0.51	0.07

HCAS scaling algorithm

A piecewise linear rescaling algorithm with two inflection points was used to simulate non-linearity and derive a calibrated HCAS index between 0.0 and 1.0 (Figure S24). The x-axis coordinates were defined by median uncalibrated condition values in areas of intensive land use (i.e., highly modified, Table S16) as of 2015–16 (ABARES 2022), and mapping of inferred reference sites as indicative of relatively natural areas (Figure S25). The y-axis coordinates for condition scores were derived from the species-area relationship (SAR, z = 0.25) back-transformation of PREDICTS project coefficients (Hudson et al. 2017) for 2015 global harmonised land use classes (LUH2 - Chini et al. 2020, Hurtt et al. 2020) that aligned with mapping of highly modified or relatively natural areas for Australia. The area-weighted average of SAR-transformed PREDICTS scores in highly modified land areas summed to 0.1001 (Table S16), and for relatively natural areas this was 0.944 (Table S17).

In relatively natural areas, establishing the land use area-weighting for each PREDICTS condition score was more challenging (Table S17). The predominant secondary level land use classes for Australia (ABARES 2016) within relatively natural areas are ‘Grazing native vegetation’, ‘Other minimal use’, ‘Managed resource protection’, and ‘Nature conservation’. These provide a poor match with the three relatively natural global land use (LUH2) classes of Primary vegetation, Secondary vegetation (mature, intermediate, young) and Rangelands (Chini et al. 2020, Hurtt et al. 2020). We therefore used proportions of these LUH2 classes within inferred reference sites dataset (relatively natural areas) to weight the PREDICTS condition scores.

For rangelands, it was necessary to account for global bias toward more intensive grazing in PREDICTS biodiversity composition coefficients for that land use class, because locations of those source data are skewed towards higher grazing densities worldwide. PREDICTS coefficients would therefore overestimate the level of degradation within Australian rangelands, as shown by mapping of Australian rangeland condition in Newbold et al. (2016). To correct for this global bias, for an application to Australia, the Gridded Livestock of the World dataset (Gilbert et al. 2018) was used to calculate ruminant-only grazing pressure following the Tropical Livestock Units (TLU) methodology of Njuki et al. (2011), applied at 1-km resolution. Ruminants were choosen because these are most prominant among introduced grazing livestock in Australia and most non-ruminant grazing species are native to Australia (e.g., excepting introduced species such as feral horses and rabbits). By fitting a trendline to the PREDICTS conditon scores (SAR-transformed coefficients) as a function of ruminant TLU index for all PREDICTS metanalysis locations in the LUH2 rangelands land use class available from the database, a 1 km resolution spatially-continuous global raster surface of condition scores for rangelands in 2015 was derived. The adjusted SAR-transformed PREDICTS condition score was then summarised as a proportion-weighted average within the rangeland land use extent for 2015 (from the LUH2 database), delimited to the extent of relatively natural areas defined within Australia (Table S17). Note that Australian rangelands include a proportion of primary as well as mature secondary vegetation. Therefore this adjustment results in higher average condition scores than the original PREDICTS data for rangelands shown in Table S15.

For the ‘Secondary vegetation’ land use class, the PREDICTS database provides a range of coefficients for young, intermediate and mature secondary vegetation. Spatial mapping and definitions of secondary vegetation ages from the 25 km by 25 km pixel resolution of the LUH2 dataset, for classes used by the PREDICTS analysis, were combined with age-specific SAR-transformed, PREDICTS-derived condition scores (from Table S15: young, intermediate, mature) to generate a spatially continuous raster surface of condition scores for the year, 2015. The condition score for secondary vegetation was then summarised as a proportion-weighted average, delimited to the extent of relatively natural areas (Table S17). In this way the mean PREDICTS condition scores shown in Table S17 were revised for consistency with Australian land uses, and differ from the original global scores in Table S15.

The median uncalibrated HCAS v2.3 score for high ecosystem integrity areas (i.e., inferred reference sites) was found to be 0.01535, and for highly modified areas, this was 0.01049 (Table S18, Figure S24). Uncalibrated data were therefore rescaled to derive an index between 0.0 and 1.0, as shown in Figure S26. The resulting site-level condition score can be further adjusted to account for local neighbourhood pressures (e.g., to derive ecosystem site condition – see method detailed below), and as an input to a general connectivity analysis incorporating the effects of landscape fragmentation (e.g., Drielsma et al. 2022, Giljohann et al. 2022).

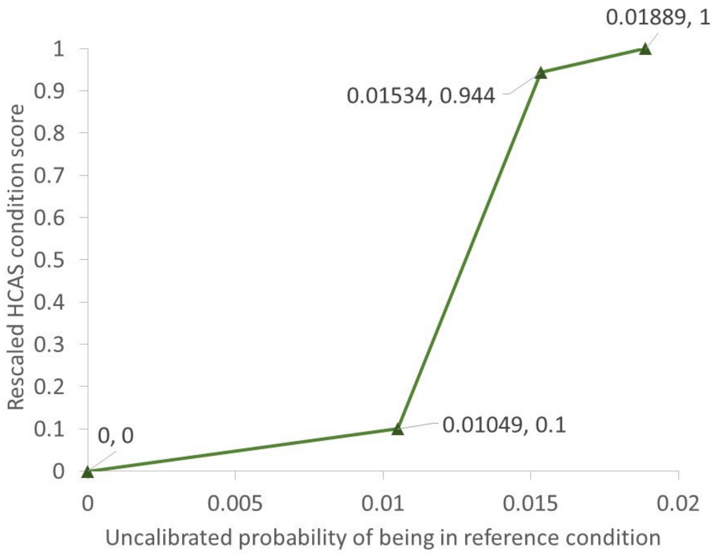


Figure S24. Piecewise linear rescaling coordinates used in HCAS v2.3 to derive an index ranging from 0.0 to 1.0.

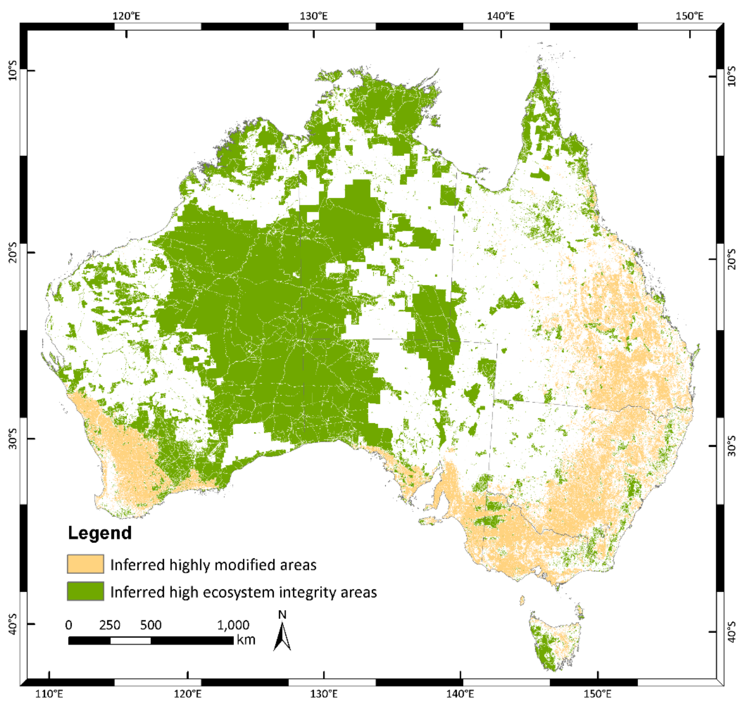


Figure S25. Spatial pattern of relatively natural areas from inferred high ecosystem integrity areas (i.e., reference sites) and inferred highly modified areas as of 2015–16 (ABARES 2022) used in the HCAS v2.3 updated scaling algorithm.
Intermediate areas are shown in white. Projection: Australian Albers GDA 1994.

Table S16. Using PREDICTS project data to derive condition coordinates for intensive land use areas of Australia.

Global Land Use Harmonisation version 2 (LUH2) (Chini et al. 2020, Hurtt et al. 2020) classes assumed to align with highly modified areas of Australia. PREDICTS biodiversity composition coefficients converted to condition scores (derived from Table S15). Grouping of Australian secondary land uses (ABARES 2016) into three types aligned with the Hudson et al. (2017) global meta-analyses of land use intensity impacts on biodiversity. Areal percentage of Australian land use (of those listed) derived from national-level land use mapping as of 2015–16 (ABARES 2022). Disaggregation of the agricultural crop category (total 60%) was based on the area proportions of the Land Use Harmonisation version 2 dataset for 2015 (LUH2 - Chini et al. 2020, Hurtt et al. 2020). Area-weighted condition scores are multiples of the mean PREDICTS condition scores by areal proportion of intensive land uses for each category, summing to 0.1001.

LUH2 intensive land use grouping	Mean predicts condition score (z=0.25)	Percentage (%) of intensive land use types in Australia	Australian land use and management classification (ALUM) (secondary)	Area weighted condition score
Pasture	0.102	31.00	3.2 Grazing modified pastures; 4.2 Grazing irrigated modified pastures	0.032
Agriculture (C3 and C4 perennial crops)	0.233	1.05	3.3 Cropping; 3.4 Perennial horticulture; 3.5 Seasonal horticulture; 4.0 Production from irrigated agriculture and plantations; 4.3 Irrigated cropping; 4.4 Irrigated perennial horticulture 4.5 Irrigated seasonal horticulture; 4.6 Irrigated land in transition; 5.1 Intensive horticulture; 5.2 Intensive animal production	0.002
Agriculture (C3 and C4 annual crops)	0.080	53.30		0.042
Agriculture (C3 nitrogen fixing crops)	0.066	5.70		0.004
Urban	0.22	9.00	5.0 Intensive uses; 5.3 Manufacturing and industrial; 5.4 Residential and farm infrastructure; 5.5 Services; 5.6 Utilities; 5.7 Transport and communication; 5.8 Mining; 5.9 Waste treatment and disposal	0.02

Table S17. Using the PREDICTS project data to derive the condition coordinate for relatively natural areas of Australia, as defined by the extent of inferred reference sites.

Global Land Use Harmonisation version 2 (LUH2) (Chini et al. 2020, Hurtt et al. 2020) classes in 2015 assumed to align with relatively natural areas of Australia from NVIS version.6.0 (DCCEEW 2023). PREDICTS biodiversity composition coefficients converted to condition scores (derived from Table S15) with rangelands and primary vegetation adjustments for Australian land types. Areal percentage of LUH2 categories were determined within the extent of inferred reference sites and used to average the mean predicts score, with spatial adjustments for age of secondary vegetation and grazing density. Area-weighted PREDICTS condition scores are multiples of the mean PREDICTS condition scores by areal proportion of extensive land uses for each category, summing to 0.944.

LUH2 extensive land use grouping	Adjusted mean predicts condition score (z=0.25)	Percentage (%) of extensive land use types in Australia	Global harmonised land use version 2 (LUH2) categories	Area weighted condition score
Primary Vegetation	1.000	75.79%	Primary Vegetation	0.7579
Rangelands	0.784	21.53%	Rangelands	0.1688
Secondary Vegetation	0.646	2.67%	Mature Secondary Vegetation; Intermediate Secondary Vegetation; Young Secondary Vegetation	0.0173

Table S18. Summary statistics for the uncalibrated HCAS v2.3 score in each of the areas shown in Figure S25.

Dataset	Minimum	First quarter	Median	Mean	Third quarter	Maximum
Relatively natural areas (inferred reference sites)	0.00000	0.01461	0.01535	0.01506	0.01589	0.01900
Highly modified areas (intensive land use)	0.00001	0.00617	0.01049	0.00939	0.01284	0.01869

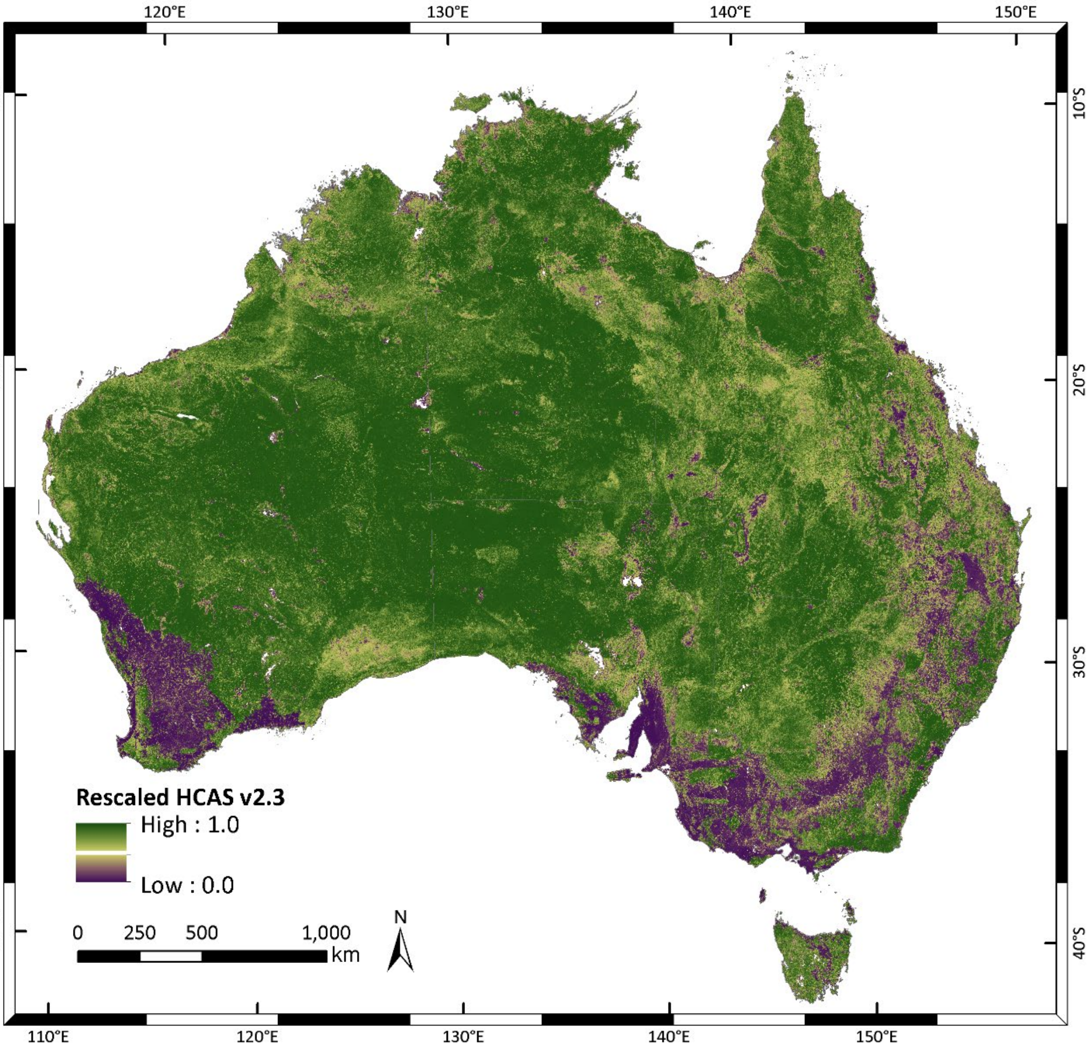


Figure S26. Calibrated HCAS v2.3 for the base model (2001-2018).
Projection: Australian Albers, GDA 1994.

Annual epochs of ecosystem condition

Annual epochs of ecosystem condition, from 2001 to 2018, were derived using the same benchmarking process and scaling algorithm as the long term epoch for the HCAS base model by substituting observed long-term with annual remote sensing PCs in test-benchmark comparisons. Annual epochs of remote sensing variables (listed in Table S19) formed part of the lineage used in deriving the long-term epochs. The PCs of these annual remote sensing variables were derived by first converting to their proper scaling by dividing by 1000, then standardised by subtracting the grid mean and dividing by the standard deviation so that all variables have the same mean (=0) and variance (=1), before running the principal components analysis.

Table S19. Annual remote sensing summary variables used in HCAS v2.1-3.

Variable	Description	Summary metrics	Original spatial resolution	Source
Persistent green cover fraction	The fraction of ground covered by green foliage of persistent (~perennial) species	Annual average	250 m	Donohue et al. (2009)
Recurrent green cover fraction	The fraction of ground covered by green foliage of recurrent (~annual) species	Annual average Annual maximum	250 m	Donohue et al. (2009)
Litter cover fraction	The fraction of ground covered in non-photosynthesising plant material (litter)	Annual average Annual maximum	500 m	(Guerschman 2019, Guerschman and Hill 2018)
Bare ground fraction	The fraction of ground not covered in green foliage or plant litter	Annual average Annual maximum	500 m	(Guerschman 2019, Guerschman and Hill 2018)

Deriving ecosystem site condition

Calibrated HCAS condition scores, ranging from 0.0 (habitat removed) to 1.0 (habitat intact), represent a partial measure of condition due to limitations in available remote sensing products to characterise all facets of ecosystem structure, function and composition, including beneath closed canopies, relevant to habitat quality assessment. As a satellite-based site-level estimate, the HCAS output also does not account for local edge effects of fragmentation that negatively influence site quality due to surrounding land uses, especially in highly modified, fragmented landscapes. The site condition in these fragmented landscapes is expected to be lower compared with sites within larger or more contiguous areas of habitat, such as relatively natural landscapes.

Here we distinguish two types of landscape context analysis applicable to condition assessment: 1) impacts of the surrounding landscape on condition of the site, and 2) contributions that condition of the site make to overall effectiveness of a landscape through connected habitat for biodiversity. The first represents the landscape context component of a condition assessment, and the second is a component of a subsequent biodiversity assessment; for example related to how metapopulations connect and interact at different spatial and temporal scales (Drielsma et al. 2022). Therefore, we developed a method to incorporate local landscape contexts related to fragmentation and edge effects into an overall measure of ecosystem condition at the site level using HCAS as an input dataset, to enhance its local applicability and reduce bias due to gaps in ecosystem quality characterisation.

As a proxy for the general effect of local pressures, we used a local neighbourhood proximity algorithm which gives a rapidly diminishing influence on site condition with distance. The approach is similar to that used to model human impacts on forest integrity (Grantham et al., 2020) or to model connectivity of habitat for biodiversity persistence applied locally (Drielsma et al. 2007, Drielsma et al. 2022).

Specifically, the inferred cumulative impact of multiple diffuse pressures (p) on site condition was modelled as an exponential decline with distance from the site of interest (represented by cells within a raster grid), truncated at 2 km, expressed mathematically as follows:

$$p_{i,j} = \begin{Bmatrix} \exp(-\lambda d_{i,j}), & d_{i,j} \leq 2\text{km} \\ 0, & d_{i,j} > 2\text{km} \end{Bmatrix} \qquad (4)$$

Where λ is the exponential decay constant and d is the Euclidean distance (m) between raster grid cells i and j. We set λ equal to 1/250 m, which is broadly consistent with previous studies (e.g., Alignier and Deconchat 2013, Grantham et al. 2020, Laurance 1991) and results in pressures declining 50% within 250 m and approaching zero by 1500 m (as approximated in Figure S27).

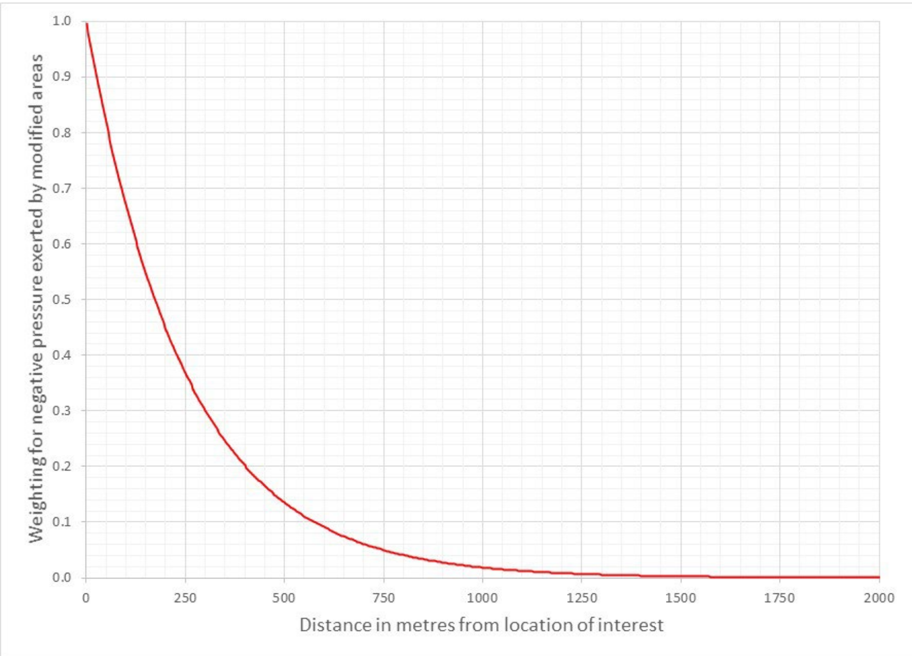


Figure S27. Exponential decay function for the 250 m distance parameter over which local neighbourhood pressure effects are inferred.

The total impact of inferred pressures (P) on site condition of raster grid cell i from all n cells within 2 km range (with j=1…n) is calculated as a distance-weighted average:

$$P_i = \sum_{j=1}^{n} C_j p_{i,j} \Big/ \sum_{j=1}^{n} p_{i,j}. \qquad (5)$$

Where C is the site condition of raster grid cell j (from HCAS), which is weighted by p, the distance function (from Equation 1).

Since P_i is a weighted average of landscape context condition, it is measured in units of condition with values between 0.0 and 1.0, consistent with C_i, and is negatively correlated with pressures exerted by the surrounding local neighbourhood.

As can be seen from equation 3 above, a site of interest with a low C_i in a surrounding landscape of higher condition (C_j) can potentially result in $P_i > C_i$. The inferred local pressures analysis aims to describe negative impacts of local neighbourhoods through processes that reduce the condition of a site of interest, as opposed to positive effects for constituent biodiversity of connectedness with quality habitat. Consequently the effects of P_i on C_i were limited to negative impacts by limiting P_i as follows:

$$P_i^{limited} = \begin{Bmatrix} P_i, & P_i \leq C_i \\ C_i, & P_i > C_i \end{Bmatrix} \tag{6}$$

The same parameters were used to derive inferred local pressure outputs for both HCAS v2.3 long term and annual epochs.

Then the ecosystem site condition index (*SCI*) is derived from the original HCAS condition index and inferred local pressures index as the geometric mean with equal weights as follows:

$$SCI_i = \sqrt{C_i P_i}. \tag{7}$$

The resulting adjusted index of ecosystem site condition (e.g., Figure S28) represents the contribution that a given site (grid cell) makes to effective area of habitat remaining within any given spatial reporting unit, expressed as a proportion of the contribution made by a site in reference condition.

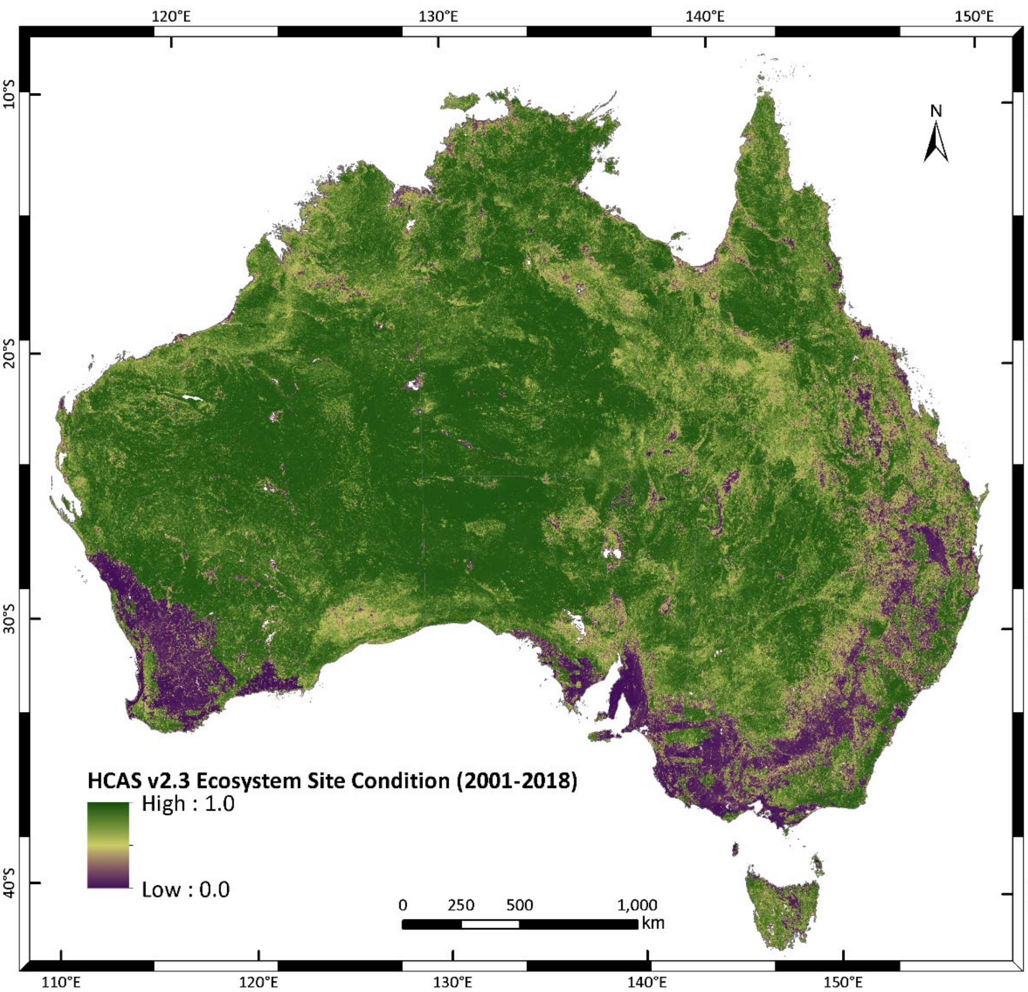


Figure S28. Example of the national extent of an ecosystem site condition subindex derived by combining the HCAS v2.3 base model sub-subindex (2001-2018) and its local pressures sub-subindex.
This represents the best overall estimate of ecosystem site condition based on the long-term remote sensing epoch, 2001 to 2018. Projection: Australian Albers, GDA 1994.

Evaluating ecosystem condition

Both qualitative and quantitative approaches were used to evaluate performance of HCAS output. Direct validation requires independent field observations of site condition, which are not consistently available across the Australian continent. While some land management agencies in Australia have implemented field protocols for estimating habitat/ecosystem condition; for example, the State of Queensland - (Eyre et al. 2017, Eyre et al. 2015), the State of Victoria - (DSE 2004, Parkes et al. 2003), Tasmania - (Michaels 2006, Michaels et al. 2020), South Australia - (DNR and NVC 2020), New South Wales - (DPIE 2020, Oliver et al. 2021); these are customised for local regulation of native vegetation clearing and have not been harmonised for consistent National use. While national field assessment methods have been scoped (McCallum et al. 2023), these are yet to be widely implemented. Therefore, a multiple lines of evidence approach was used to evaluate how well HCAS outputs compare with expectations.

Expert elicitation

Validation of calibrated HCAS v2.3 long-term base model dataset was performed using two independent sources of ecosystem condition data derived through expert elicitation: 1) virtual transects and 2) site condition assessments (White et al. 2023). Analyses were performed using a Major Axis Type II regression (Legendre and Legendre 2012) which assumes both response and predictor variables are random and measured with error (Carroll et al. 2006, Schennach 2016). This method was selected because all variables are expressed in the same physical units (ecosystem condition, dimensionless), and so error variances can reasonably be assumed approximately equal.

Virtual transects method

Nine ecologists with extensive field experience in specific regions visually assessed condition at 11 evenly spaced points along one or two of 11 pre-defined virtual transects using Google Earth imagery. The transects traversed large swathes of the Australian continent (Figure. S29, Table S20).

Transect locations were chosen iteratively. They were initially located to cover a representative sample of ecosystem types and major land uses across Australia. Then, in consultation with experts, transects were relocated (or added) to best cover regions they were most familiar with. Start and finish points of each transect were chosen subjectively. A straight line between these points was divided into 10 equal parts, giving a total of 11 survey points along each transect. At each point, the centre of the closest 250 x 250 m grid cell was identified. A circle of 125 m radius was placed around that cell centre whose perimeter was tangential to the cell edges (Figure 38). Hence, each survey location consisted of an area of just under 5 ha each, for compatibility with the HCAS spatial reference and grid size.

This activity was approved by the CSIRO Social Science and Human Research Ethics Committee (original clearance: 025/18; data reuse clearance 048/21 and 196/23).

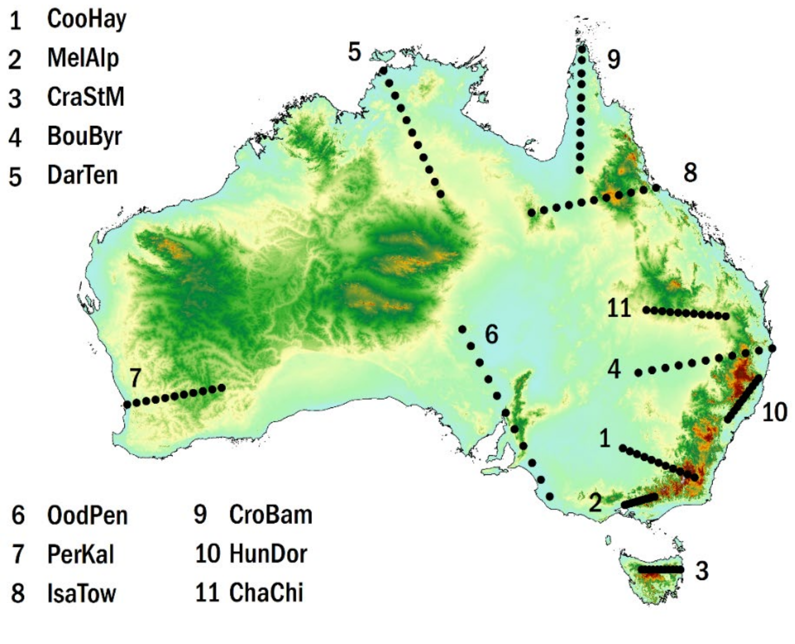


Figure. S29 Map of the 11 virtual transects established for the rapid expert assessment of ecosystem condition
Each transect is described in Table 8. The background shows elevation (0 – 2220 m) from Gallant et al. (2011)

Table S20. Descriptions of the 11 transects established for the rapid expert assessment of ecosystem condition. Transect names and numbers correspond to those in Figure. S29

No	Name	Start	Finish	Represents
1	CooHay	Centre of Hay	Farmland north-west of Cooma	Multiple land uses through former Grassy Woodland
2	MelAlp	Melbourne CBD	Alpine plateau in Alpine National Park	An elevational and land use gradient through forests, woodlands and grasslands
3	CraStM	Cradle Mountain Lodge	Coastal forest near St Marys	Spans designated wilderness areas, intensive agriculture and native forestry
4	BouByr	Farmland north-west of Bourke	Peri-urban area outside Byron Bay	Strong climatic gradient from semi-arid to sub-tropical; includes irrigated agriculture
5	DarTen	Suburb of Darwin	Tennant Creek	Strong rainfall gradient through tropical savannas
6	OodPen	Oodnadatta	Plantations near Penola	Arid to Mediterranean climate gradient, including intensive agriculture
7	PerKal	Perth CBD	Grazing land south of Kalgoorlie	City-urban-cropping-grazing gradient along a rainfall gradient
8	IsaTow	Mt Isa	Townsville	Multiple land uses and a rainfall gradient, crosses the Mitchell Grasslands
9	CroBam	Farmland east of Croydon	Savanna south of Bamaga	Spans Cape York Peninsula
10	HunDor	Forest in Yengo National Park	Forest near Dorrigo National Park	Coastal forests, intersects national parks, production forests and cleared farmland
11	ChaChi	Warrego River at Charleville	Farmland east of Chinchilla	A rainfall gradient through the Brigalow Belt

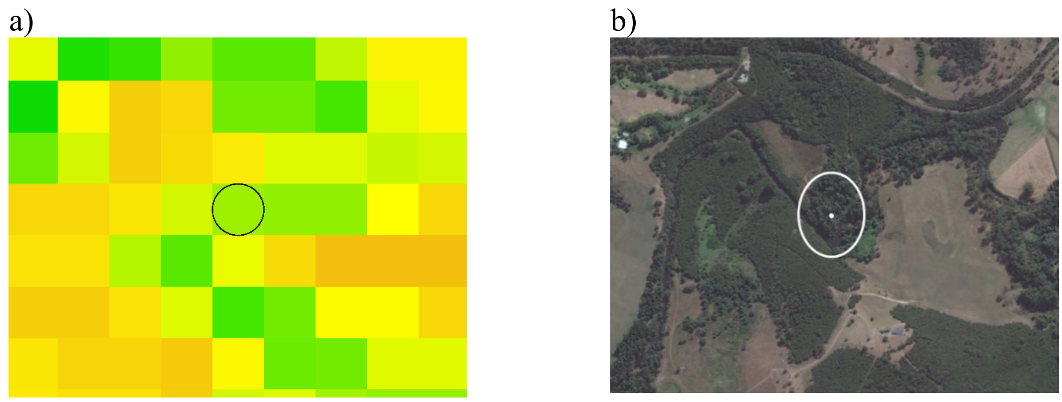


Figure S30. Example of a polygon at a survey location along a virtual transect. This is shown (a) in relation to the underlying HCAS grid cell alignment and (b) on a Google Earth image.
The grid in plot a) is unprojected (that is, the x and y coordinates are simple longitude and latitude coordinates) and so the 250 x 250 m polygon appears as a circle. The image in plot b) is projected (Plate Carree projection) and so the same polygon appears as an ellipsoid. The image on b) is what the experts view when undertaking the survey.

Experts had access to Google Earth Pro for conducting the virtual transect surveys. Each expert was sent an information pack that included the consent form, participant information and instructions, a recording spreadsheet and a Google Earth kml file. This kml file contained ellipsoids of their particular transect. Opening the kml takes each expert directly to the transect location. No contextual information was included other than underlying high-resolution satellite imagery within Google Earth to ensure experts relied only on their personal knowledge or insight about each site. The imagery currency was unknown, as made available for Australia by Google Earth in early 2018.

Brief guidance was given on how to define condition, as follows.

- Condition was to be scored between 0.0 (lowest) and 1.0 (highest).
- A ‘0.0’ condition score would only apply to vegetation that had been so altered or removed that it could no longer support its original indigenous biodiversity.
- A ‘1.0’ condition score would represent vegetation considered to be in an ‘intact’ state and able to support a full complement of indigenous biodiversity that would normally persist there.

The instruction for each expert was to make and record their best estimate of condition within each ellipsoid, representing an average for the 2001-2016 period. Then, if there had been any substantial changes in condition since 2001, experts were asked to record when, to the best of their knowledge, this occurred and the condition prior to change. Finally, experts were asked to record how confident they were with each estimate, ranging from 1 (“I’m really just guessing – I don’t know the area”) to 5 (“I know this area and its ecology very well”). A column was provided for them to add any notes they wished to include.

Site condition assessments method

Among known challenges of using remote sensing for assessing ecosystem condition is the paucity, limited spatial coverage and inconsistency of field observations (i.e., in situ) for training the interpretation of remote sensing images. With insufficient spatial coverage of field ecosystem condition assessments, it is possible to misinterpret remote sensing data—for example, by mistaking a highly modified habitat for a natural habitat in reference condition. The expert site assessments methodology applied using the Habitat Condition Assessment Tool (HCAT) (Brenton et al. 2018) was designed to test a process of systematically gathering ecological expert knowledge to inform training, calibration, and validation of HCAS workflow components, among other uses.

Twenty-one experts contributed 314 site condition assessments via HCAT, which included a method for expert cross-calibration enabling results to be rescaled (White et al. 2023). Expert site assessments covered a range of ecosystems and geographies (Figure S31). The majority (66%) of these contributed sites were given condition scores above 0.5 (Figure S32).

This activity was approved by the CSIRO Social Science and Human Research Ethics Committee (original clearance: 004/17; data reuse clearance 007/21 and 196/23).

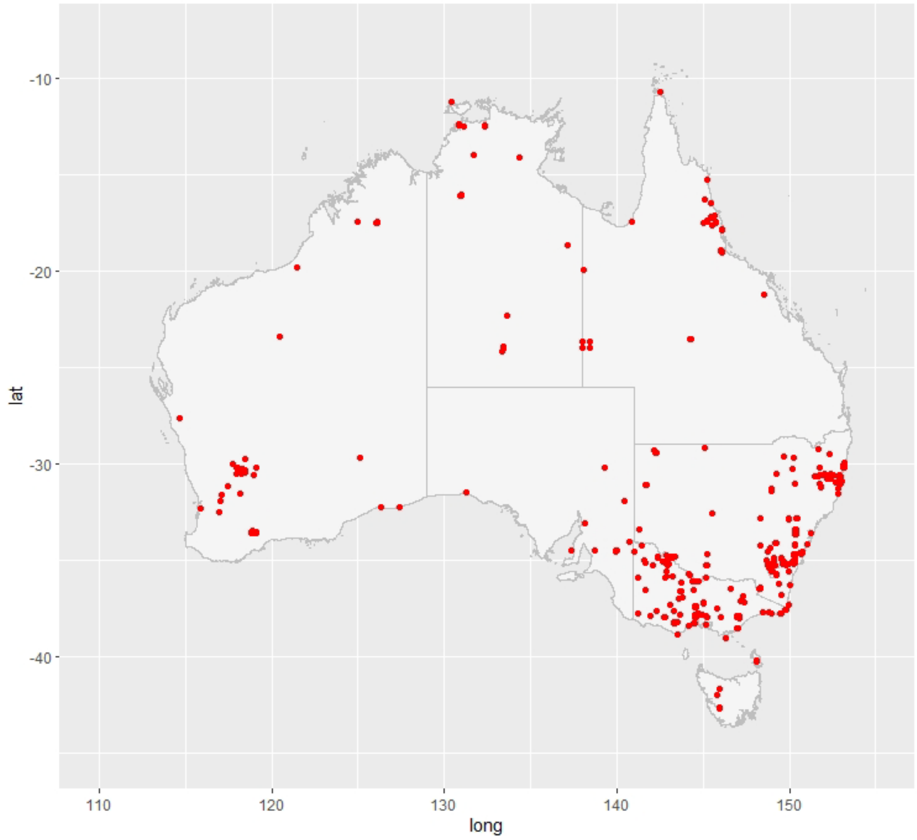


Figure S31. Geographic distribution of 314 sites, across 21 experts, for which site condition assessments were provided through the HCAT expert elicitation process
Source: White et al. (2023).

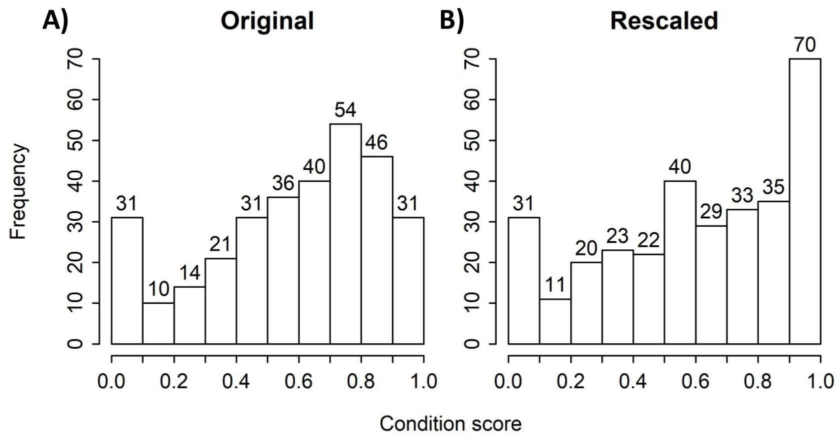


Figure S32. Frequency distribution of original (A) and rescaled (B) condition scores provided by experts for 314 site condition assessments contributed by experts via HCAT
Source: White et al. (2023).

Two types of information were elicited from experts via HCAT: 1) calibration image assessments and 2) site condition assessments (see data collection: White et al. 2019). Experts were first asked to provide an ecological condition score for a set of calibration images suited to their geographic and vegetation class expertise, of ecosystems in a range of condition states. These were photographs of different types of ecosystems in various reference and modified condition states for which the location was known so that the image could be broadly grouped by vegetation type and geographic region (see image data collection: Warnick et al. 2019). Experts provided scores to quantify ecosystem condition (naturalness) of those sites on a scale between 1.0 (a site in its most natural form) and 0.0 (a completely transformed site). Participants were guided to consider ecological condition as the *capacity of an area to support the plants and animals that would exist at that location in a natural state*. Each calibration image was scored by multiple experts with similar (self-nominated) regional expertise and a rescaling analysis was completed to quantify the tendency for positive or negative personal biases of each contributor.

Experts were then asked to consider specific sites that they knew well and provide condition scores for those. They could also detail which disturbance factors had influenced their score. Their scores were rescaled according to calibration parameters derived by analysing the calibration image assessments to enable site assessments from multiple contributors to be compiled into a single coherent dataset (White et al. 2023).

Calibration image collection of Australian ecosystems (Warnick et al. 2019) comprised 777 images of ecosystems in various condition states from a range of sources (private and public), of which 77 were deployed during the pilot, returning 278 calibration image scores. During a 10-week campaign (September to November 2018), 314 site condition assessments were contributed by 21 expert participants. The calibration and rescaling method, which adjusted the experts' scores against a consensus or collective opinion defined by a general linear model, successfully dampened individual scoring biases (Figure S32).

All 314 expert site assessments were selected for comparison with the HCAS v2.3 data. Polygons were converted to a 100m grid, then centroids were exported to a point file resulting in 143,831 sites for the Type II analysis. The HCAS v2.3 data resolution was 250m resulting in some duplication, but this was even across all comparisons.

Type II analysis

With a Type II regression approach, a perfect (even if noisy) correlation between two variables being analysed can be expected to yield a fitted line close to the 1-to-1 line (something that cannot be expected with ordinary least squares regression, OLS). This in turn translates to a slope of 1.0 with 0.0 intercept, and a slope angle of 45°. Also, the standard RMSE metric (which, in OLS, measures deviations from data points from the fitted line vertically; that is, parallel to the response variable's axis) becomes meaningless with a Type II regression – here, none of the variables were assumed to be a response to another, considered as a predictor variable. Instead, a similar measure of dispersion can be achieved by measuring average residual error (distance) *perpendicular* to the fitted Type II line, a metric which we labelled Root Mean Square Orthogonal Error (RMSOE), for which lower values are preferred.

The Major Axis Type-II regression demonstrated reasonable agreement between the HCAS v2.3 base model data (2001−18) and each set of expert's condition scores (Figure S33). Agreement was higher when using the virtual transect method, possibly because that approach was developed specifically for validating HCAS scores using remote sensing imagery for the assessment. The HCAT expert site assessments method relied upon direct field experience and was intended to provide data for a range of potential applications. The experts scores contributed through the virtual transects approach, however, could not be cross-calibrated because a mechanism to do this was not part of that method. Future application of virtual transects would ideally involve more than one expert scoring sites for each transect.

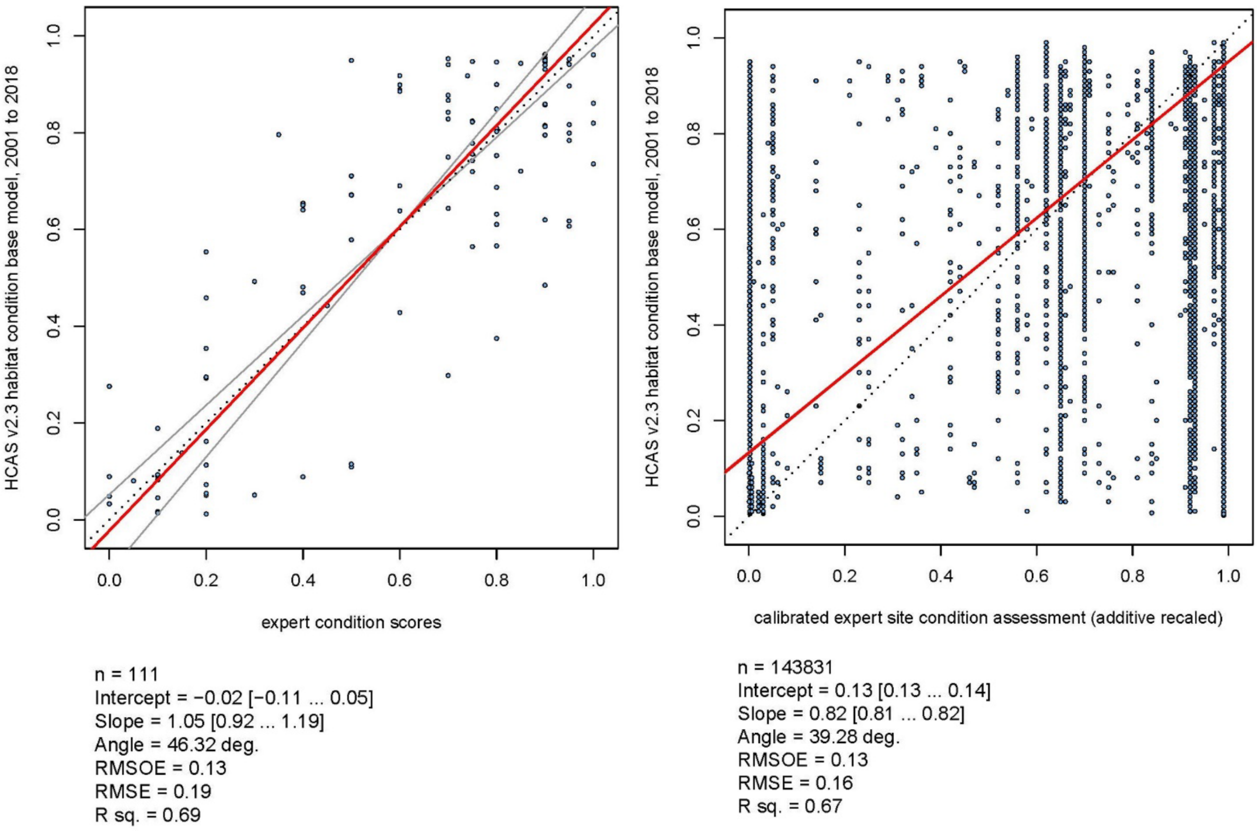


Figure S33. Type II regressions between HCAS v2.3 ecosystem condition and expert condition scores from virtual transects (left) and rescaled HCAT expert site condition 'best' scores (right)

The 'Intercept' results are the estimated intercept using the Type II regression (with confidence interval, CI, range); the 'Slope' results are the estimated slope coefficient (with CI) – best when closest to 1.0; the 'Angle' result is the estimated angle of the fitted line (best when closest to 45°); RMSE and R-squared are the standard metrics from a normal linear regression (i.e., not a Type II regression); and RMSOE is the "bespoke" orthogonal RMSE between the data points and the Type II regression line ('bespoke' in the sense that it is not really a standard metric of modelling error, but provides some insight into the "orthogonal" variability of the data points from the regression line, i.e., in the spirit of the Type II analysis).

Existing maps of ecosystem modification levels

The calibrated HCAS version 2.3 scores were also compared with categorical mapping of native vegetation modification levels derived from a wide range of land use and land cover datasets for Australia (Lesslie et al. 2010) consistent with the Vegetation Assets, States and Transitions (VAST) narrative framework (Thackway and Lesslie 2006, 2008). Continuous HCAS scores were assigned discrete VAST classes on the basis of elicited expert's condition scores to enable a comparison of ordered categories. Concordance between two datasets was then qualitatively assessed using a confusion matrix. The experts condition scores were also assigned to VAST spatial data categories to generate an ordinal dataset used in a Type II comparison with HCAS continuous data to show the nature of the relationship. The VAST narrative framework is useful in this context because it enables broad categories of HCAS condition scores to be related to vegetation modification levels for interpretative and communication purposes.

The classification terminology originally developed by Thackway and Lesslie (2006) provides a values-neutral framework for general communication and reporting on vegetation condition (summarised in Table S21), which facilitates inclusion and discussion with diverse stakeholders having different world views and perspectives about the environment.

Table S21. The six categories of the vegetation assets, states and transitions (VAST) narrative framework (Thackway and Lesslie 2006, 2008) and the description of current regenerative capacity (one of several diagnostic criteria).

VAST category	VAST description of current regenerative capacity
Class 0: Residual Bare	Natural regenerative capacity unmodified— ephemerals and lower plants
Class I: Residual	Natural regenerative capacity unmodified
Class II: Modified	Natural regeneration tolerates or endures under past and or current land management practices
Class III: Transformed	Natural regenerative capacity limited or at risk under past and or current land use or land management practices. Rehabilitation and restoration possible through modified land management practice
Class IV: Replaced - Adventive	Regeneration of native vegetation community has been suppressed by ongoing disturbances of the natural regenerative capacity; limited potential for restoration
Class V: Replaced - Managed	Regeneration of native vegetation community lost or suppressed by intensive land management; limited potential for restoration
Class VI: Removed	Native vegetation community removed

Expert elicitation of condition scores for VAST vegetation modification levels

Twenty-six experts with field ecology survey and mapping expertise contributed condition scores for each of the VAST categories summarised in Table S21. Most also provided comments to explain their choices or reflections on the narrative framework descriptions, which was an optional part of the survey.

This activity was approved by the CSIRO Social Science and Human Research Ethics Committee (original clearance: 115/22; data reuse clearance 197/23 and 199/23).

Experts were asked to assign ecosystem condition scores as follows:

- an overall condition score for each VAST ecosystem category between 0 and 1
- a plausible upper and lower bound for their condition score
- their confidence (between 50 and 100%) that the interval described captures the true value of condition for that VAST ecosystem category.

Specific questions were provided in an Excel workbook with tables for each VAST category. Each table included the category description, diagnostic and examples as originally published by Thackway and Lesslie (2006). A registration page in the workbook asked participants to provide a survey code so that results could be anonymised but identifiable by the participant, and instructions on how to complete the survey. An introduction page outlined the method. A glossary defined the terms ‘site condition’ and ‘landscape context’ so that participants understood that their scores should only relate to site condition. Participants were also provided with a copy of the journal article for background.

In a follow up online workshop the aggregated results were presented, and experts invited to refine their scores, if they wished. One respondent revised their estimates in the second round.

Following consistency checks, scores for two respondents were not included in the final analysis due to an apparent misunderstanding of the exercise. Additionally, individual best estimates falling outside lower and upper bounds were excluded before pooling. The median statistic was used for pooled estimates (Table S22 and Figure S34) because it represents the typical response by experts and is preferentially used when data are skewed by outliers.

Table S22. Pooled estimates of experts' ecosystem condition scores for each VAST category, summarised as the median for the best, lower and upper bounds.

Vast category	Best median	Lower median	Upper median	Number of respondents
Residual bare	0.75	0.50	1.00	24
Residual	0.88	0.75	1.00	24
Modified	0.60	0.40	0.80	23
Transformed	0.40	0.28	0.60	24
Replaced-Adventive	0.25	0.10	0.40	24
Replaced	0.15	0.05	0.30	23
Removed	0.05	0.00	0.10	23

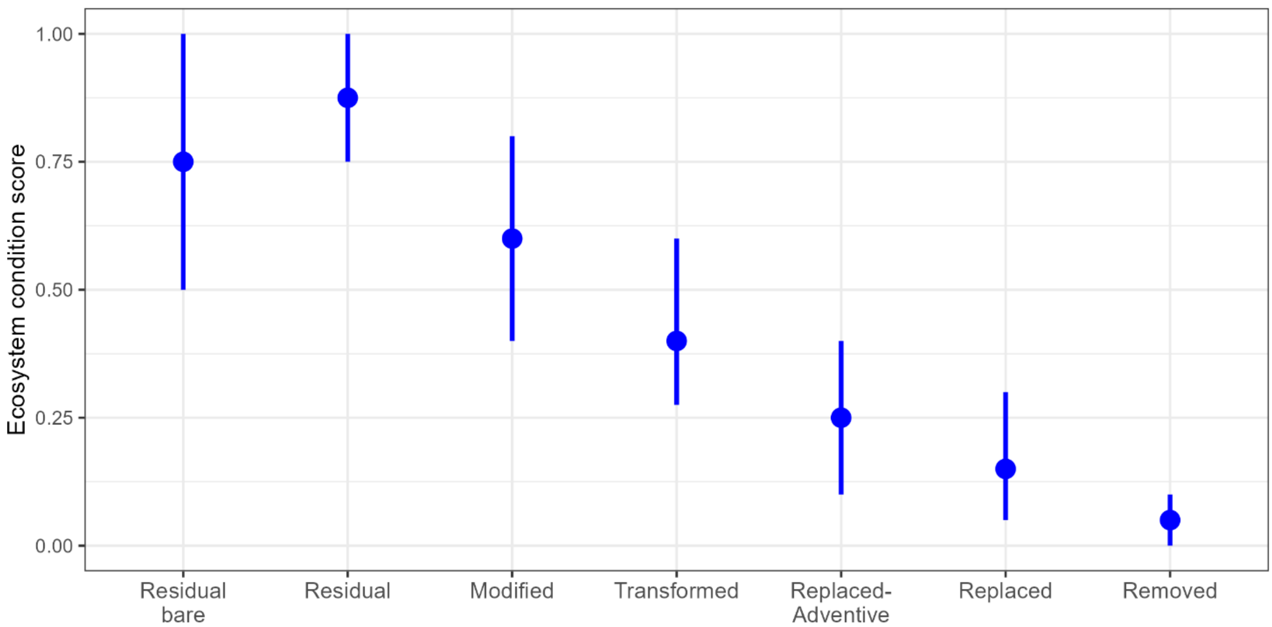


Figure S34. Pooled estimates of experts' ecosystem condition scores for each VAST category summarised as the median for the best, lower and upper bounds.
Respondents with best estimates falling outside the bounds did not contribute to the pooled estimates.

An unexpected result of the pooled expert estimates was the generally much lower scores given to the 'Residual-bare' category compared with the 'Residual' category (Table S22 and Figure S34). In the VAST framework (Thackway and Lesslie 2006, 2008), these two categories are considered equivalent reference condition levels, and differ only in the type of ecosystem. Naturally bare areas are defined as “Areas where native vegetation does not naturally persist and recently naturally disturbed areas where native vegetation has been entirely removed. (i.e., open to primary succession)”. Examples given included “Bare mud; rock; river and beach sand, salt and freshwater lakes, rockslides and lava flows”. Comments against responses from experts indicated that some considered climate variability and other pressures to be a factor in the occurrence of ‘Residual-bare’ natural areas and so condition scores below reference were frequently reported. The term ‘bare’ was often equated with disturbance-related degradation processes rather than a natural phenomenon. Some experts also referred to bare areas as not having much ‘value’, as a reason for their scores, and this may have further influenced lower scores than expected for these reference ecosystems. Variation in expert’s perceptions is further demonstrated by greater variability in individual estimates associated with ‘Residual-bare’ compared with other classes (Table S22 and Figure S34). Because of this confusion, ‘Residual-bare’ was dropped as a category for our purposes. We therefore use expert scores for ‘Residual’ as indicative of scores expected for ‘Residual-bare’.

Pooled medians of the expert’s best estimate scores for ecosystem condition based on six VAST classes were used to fit a curve for the relationship between 0 and 1 (as above, excluding ‘Residual-bare’ which was assumed equivalent to ‘Residual’ category for this purpose) (see Figure S35). The 0 and 1 condition extremes are halfway between VAST classes because the latter represent medians of experts best estimate for that class. The mid-points between scores along the fitted line were used to set class boundaries, and rounded (up or down) to the nearest 0.05 (Table S21) for discretising the ecosystem condition index. These class ranges are well within median aggregated upper and lower bounds based on individual expert’s condition scores (Figure S34).

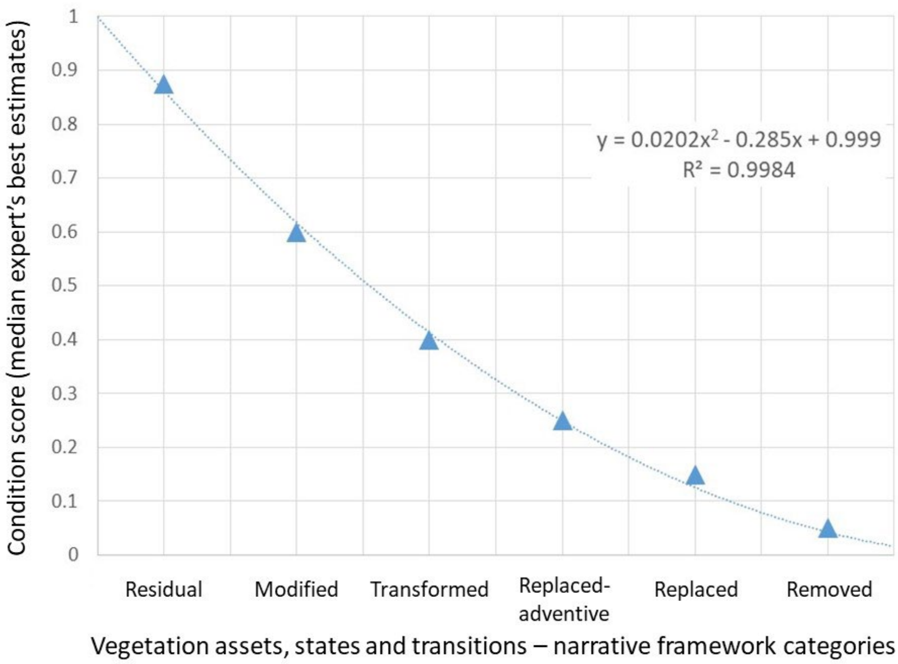


Figure S35. Curve fitted to median of expert’s pooled best estimates of condition scores (*n* = 23 for modified, replaced, removed; *n* = 24 for residual, transformed, replaced-adventive) interpreted for the VAST narrative framework (Thackway and Lesslie 2006, 2008).

Table S23. Discretisation of ecosystem condition into six ordinal categories aligned with the Vegetation Assets, States and Transitions (VAST) narrative framework (Thackway and Lesslie 2006, 2008), as informed by ecological experts (note that the 'Residual-bare' class is here equates with the 'Residual' class).

VAST category	Lower bound	Upper bound	Rounded lower	Rounded upper
1. Residual	0.73	1.00	0.75	1.00
2. Modified	0.51	0.73	0.50	0.75
3. Transformed	0.34	0.51	0.35	0.50
4. Replaced-adventive	0.18	0.34	0.20	0.35
5. Replaced (-managed)	0.08	0.18	0.10	0.20
6. Removed	0.00	0.08	0.00	0.10

Concordance analysis

The 'rounded' version of expert-informed VAST class condition scores in Table S23 was used to reclass the continuous ecosystem condition index. The reclassification was applied to the ecosystem site condition index output of the HCAS v2.3 base model (2001−18) and each of the annual epochs between 2001 and 2018 (data collection: Harwood et al. 2023). To approximate the temporal range of the VAST version 2 spatial data (1995-2006: Lesslie et al. 2010), the average of six HCAS 2.3 annual epochs of ecosystem site condition in the overlapping temporal range, 2001 to 2006, were used for comparison (Figure S36). Five categories were in common between the two datasets: residual-bare and residual classes were grouped as 'residual', and replaced-adventive and replaced-managed were grouped as 'replaced'.

Overall concordance between five common categories was 42% indicating moderate agreement (Table S24) and, when collapsed to two classes depicting relatively natural versus intensively modified areas, overall concordance was 87% indicating high agreement (Table S25).

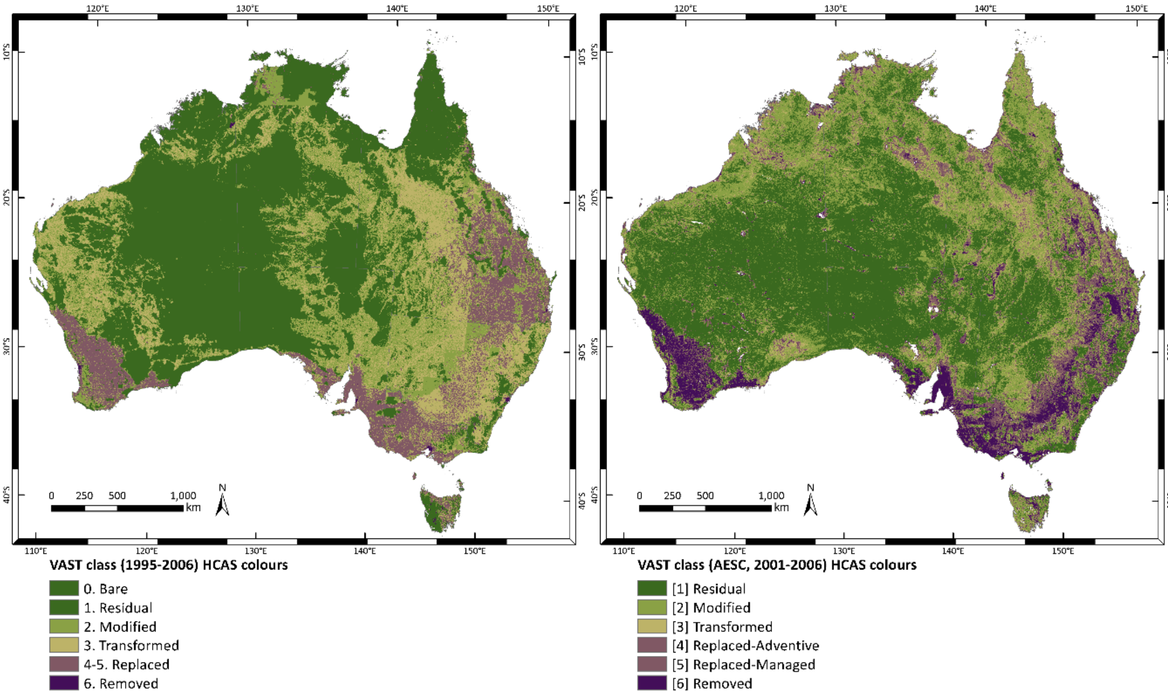


Figure S36. National VAST comparisons using cross-walked legend colours for the 5 common groupings. The VAST version 2 spatial dataset has grouped the two 'Replaced' classes (left) (Lesslie et al. 2010), whereas the expert-informed VAST classification of HCAS v2.3-derived ecosystem site condition (2001-2006 average of epochs) has grouped the two 'Residual' classes (right). Note: 'Bare' in the VAST narrative framework is a special case of the 'Residual' class.

Table S24. Five-class concordance assessment summary table (unbiased estimates) comparing the VAST version 2 spatial dataset and the HCAS v2.3 ecosystem site condition subindex averaged over 7 annual epochs, 2001 to 2006, for continental Australia (based on data in Table S26). The user's concordance is based on the proportion of HCAS and VAST pixels that are classified the same relative to the total number of HCAS pixels in that class (i.e., how consistent the classified HCAS data are compared with the VAST data and therefore the likelihood of concordance). Producer's concordance is the proportion of HCAS and VAST pixels classified the same relative to the total number of VAST pixels in that class (i.e., predicting how well new VAST-classified HCAS data would compare with the existing VAST spatial data). The standard error and 95% confidence intervals (CI) are for HCAS compared with VAST, frequency is number of 0.0025 degree data pixels.

Class (HCAS v2.3)	Standard error (frequency)	+/- 95% ci (frequency)	% users concordance (HCAS vs VAST)	% producers concordance (VAST vs HCAS)	% overall concordance
Residual	4,886	9,577	65.14	58.44	42.04
Modified	4,122	8,080	22.00	31.56	
Transformed	4,025	7,888	27.56	13.02	
Replaced	2,770	5,429	35.10	33.19	
Removed	374	733	1.18	67.06	

Table S25. Two-class concordance assessment summary table (unbiased estimates) comparing the VAST version 2 spatial dataset and the HCAS v2.3 ecosystem site condition subindex (averaged over 7 annual epochs, 2001 to 2006), for continental Australia (based on data in Table S27).

The user's concordance is based on the proportion of HCAS and VAST pixels that are classified the same relative to the total number of HCAS pixels in that class (i.e., how consistent the classified HCAS data are compared with the VAST data and therefore the likelihood of concordance). Producer's concordance is the proportion of HCAS and VAST pixels classified the same relative to the total number of VAST pixels in that class (i.e., predicting how well new VAST-classified HCAS data would compare with the existing VAST spatial data).

Class (HCAS v2.3)	% users concordance (HCAS vs VAST)	% producers concordance (VAST vs HCAS)	% overall concordance
Relatively natural	96.31%	25.73%	87.47
Intensively utilised	47.83%	74.27%	

For the concordance analysis, the VAST version 2 spatial dataset (Lesslie et al. 2010) was resampled from the original 0.01 degree grid to 0.0025 degrees to match the HCAS v2.3 ecosystem site condition data. Non data pixels in either dataset were excluded from consideration. The 5-class and binary confusion matrices are shown in Table S26 and Table S27. The corresponding concordance assessments (Table S24 and Table S25) applied the recommendations by Olofsson et al. (2013) for making better use of accuracy data in land change studies (see also NASA 2018a, NASA 2018b). Comparisons are reported as concordance assessments because neither dataset represents 'truth' for an accuracy assessment.

For interpretation, we use the Landis and Koch (1977) scale of observer agreement: a value greater than 0.80 (i.e., 80%) represents strong agreement; a value between 0.40 and 0.80 (i.e., 40–80%) represents moderate agreement; and a value below 0.40 (i.e., 40%) represents poor agreement.

Both user's (HCAS vs VAST) and producer's (VAST vs HCAS) concordances for the residual class are greater than 50% (Table S24). The HCAS v2.3 ecosystem site condition subindex shows some substantial areas of 'Residual' which correspond with VAST 'Modified' or 'Transformed' land types, and the VAST dataset shows some 'Residual' which correspond with 'Transformed' in the ecosystem site condition subindex. Overall the 'Removed' category is much more extensive in the HCAS v2.3 ecosystem site condition dataset, than in the spatial VAST dataset, whereas the 'Replaced' and 'Transformed' categories are much more extensive in the spatial VAST dataset than in the ecosystem site condition dataset.

Table S26. Confusion matrix percentages comparing VAST version 2 spatial dataset and HCAS v2.3 ecosystem site condition (ESC) subindex averaged over 7 annual epochs, 2001 to 2006, in 5 categories, for continental Australia.

The two replaced categories in the VAST classification of the ESC subindex were grouped, for consistency with that grouping in the VAST version 2 spatial dataset (Lesslie et al. 2010). Data are percentages of the total number of 0.0025 degree data pixels used in the analysis. Numbers for totals and diagonals are shown in Bold.

HCAS v2.3 ESC categories	VAST version 2 categories					
	Residual	Modified	Transformed	Replaced	Removed	Row total
Residual	**29.64**	9.28	6.00	0.58	0.00	**45.50**
Modified	14.98	**6.08**	5.27	1.29	0.01	**27.63**
Transformed	3.20	1.92	**2.38**	1.13	0.01	**8.64**
Replaced	2.36	1.63	3.11	**3.85**	0.03	**10.98**
Removed	0.54	0.35	1.51	4.75	**0.09**	**7.25**
Column total	**50.72**	**19.26**	**18.28**	**11.61**	**0.13**	**100.00**

Table S27. Confusion matrix percentages comparing VAST version 2 spatial dataset and HCAS v2.3 ecosystem site condition (ESC) subindex averaged over 7 annual epochs, 2001 to 2006, in 2 categories: relatively natural (Residual, Modified, Transformed) versus intensively utilised (Replaced, Removed), for continental Australia.

Data are percentages of the total number of 0.0025 degree data pixels used in the analysis. Numbers for totals and diagonals are shown in Bold.

HCAS v2.3 ESC categories	VAST version 2 categories		
	Relatively natural	Intensively utilised	Row total
Relatively natural	**78.75**	3.02	**81.77**
Intensively utilised	9.51	**8.72**	**18.23**
Column total	**88.26**	**11.74**	**100.00**

Type II analysis

The best median of expert's scores from Table S22 were used to convert VAST spatial data categories to an ordinal dataset for type II comparison with HCAS continuous data (Figure S37). The results suggest HCAS increasingly under-estimates condition below around 0.75 compared with expert's scores expressed through the VAST spatial data, and slightly over-estimates condition of the reference state (i.e., VAST 'residual' class).

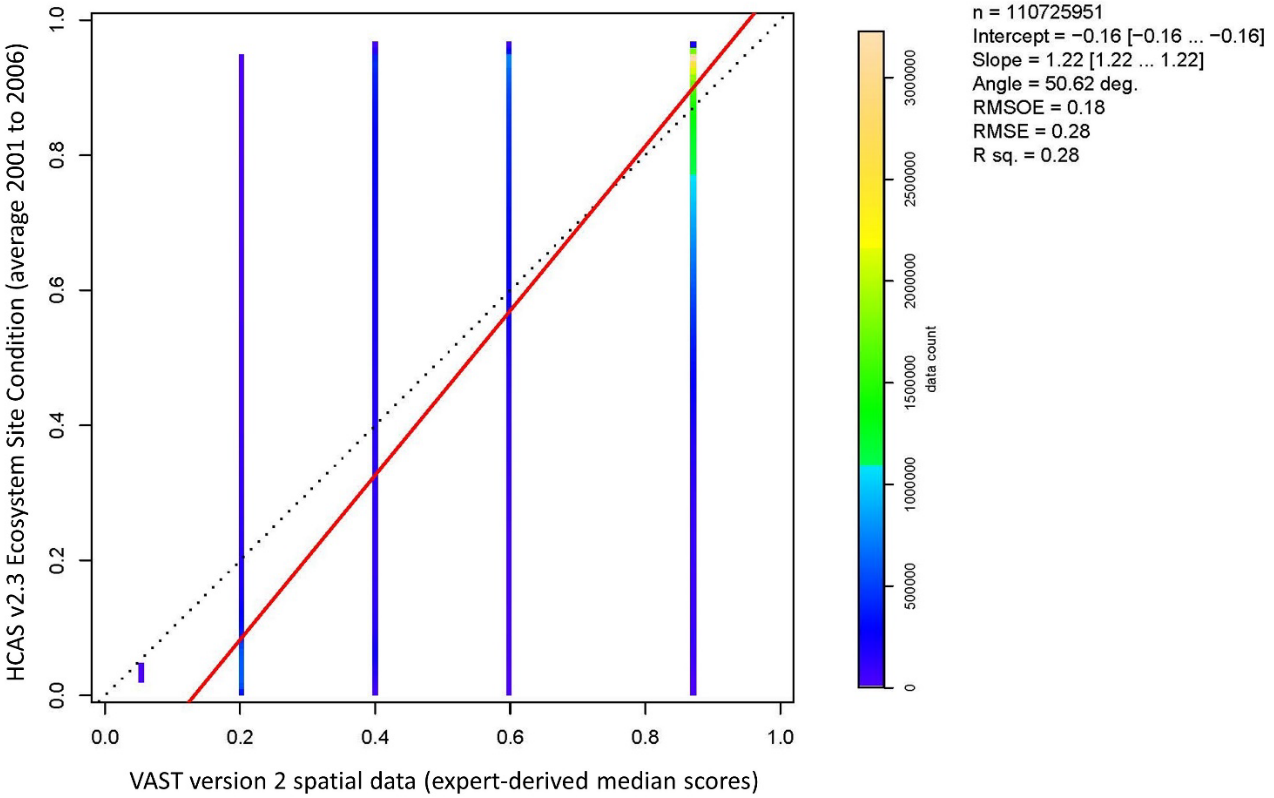


Figure S37. Type II regression between expert-derived spatial VAST median condition scores, 1995-2006 (x-axis) and HCAS v2.3 ecosystem site condition scores averaged over the seven annual epochs, 2001 to 2006 (y-axis). The 'Intercept' results are the estimated intercept using the Type II regression (with confidence interval, CI, range); the 'Slope' results are the estimated slope coefficient (with CI) – best when closest to 1.0; the 'Angle' result is the estimated angle of the fitted line (best when closest to 45°); RMSE and R-squared are the standard metrics from a normal linear regression (i.e., not a Type II regression); and RMSOE is the "bespoke" orthogonal RMSE between the data points and the Type II regression line ('bespoke' in the sense that it is not really a standard metric of modelling error, but provides some insight into the "orthogonal" variability of the data points from the regression line, i.e., in the spirit of the Type II analysis).

Visual assessments

A series of visualisation case studies were used to compare HCAS output against expectations of an ability to discriminate ecosystem condition where natural processes prevail, compared with known modified lands using national mapping of land use and land cover. These case studies focussed on areas of exotic species plantation forestry, inland regions of lateral water inflow and salt lakes, landscape dynamics associated with surface water and snow country, unique features in rugged landscapes, rapid development in urban areas, and surface mines – resource extraction. Details are presented in Williams et al. (2023b).

The remaining limitations relate mainly to: (i) incomplete characterisation of ecosystem structure, function, and composition based on the set of remote sensing input variables; (ii) an incomplete characterisation of landscape features in environmental covariates—especially those related to landscape heterogeneity and landscape water; and (iii) a need to further improve structure and quality of training and benchmark data used to represent locations with reference condition.

Many of the 65 recommendations by Williams et al. (2021) for improving the HCAS remain valid. Several of these have been addressed through incremental improvements up to HCAS v2.3.

Acknowledgments

The CSIRO and DCCEEW acknowledge the Traditional Owners of country throughout Australia and recognise their continuing connection to land, waters and culture. We pay our respects to their Elders past and present. View CSIRO's vision towards reconciliation and DCCEEW's Statement of Commitment to First Nations people..

We also acknowledge the incredible input provided by participants at the 'Developing state and transition models to support ecosystem accounting of agricultural and mixed-use landscapes in the Flinders, Norman and Gilbert river catchments' workshop from 12 to 13 September 2022 (and subsequent online condition workshop, 23 November 2022), and 'Developing state and transition models to support ecosystem accounting of agricultural and mixed-use landscapes in the Western Australian Wheatbelt' workshop from 19 to 20 September 2022 (and subsequent online condition workshop, 24 November 2022). The information provided at these workshops and surveys has been used in this document. Experts who generously provided their time and feedback, including data on reference sites, include:

Queensland: Brett Abbott, Chris Appelman, Paula Barry, Hari Boppudi, Don Butler, Ana Carla Leite de Almeida, Anthony Curro, Emily Larsen, Jane McNamara, Bernhard Morris, Bruce Murray, Olusegun Osunkoya, Geoff Penton, David Phelps, Noel Preece, Wayne Vogler, Nathan Waltham, Ian Watson, Wendy Williams, Zoe Williams, Maria Zann.

Western Australia: Brett Beecham, Dimity Boggs, Margaret Byrne, David Collard, Chris Curnow, Christophe D'Abbadie, Bonny Dunlop-Heague, Greg Durell, Paul Galloway, Carl Gosper, Mike Griffiths, Alex Hams, Peta Kelsey, Sarah Luxton, Nathan McQuoid, Helena Mills, Owen Nevin, Tim Overheu, Tina Parkhurst, Blair Parsons, Keith Pekin, Tom Picton-Warlow, Clinton Rakich, Debbi Slater-Lee, Melanie Strawbridge, Grant Wardell-Johnson, Renee Young.

We further acknowledge those experts and REAP team members who contributed data through online surveys of ecosystem condition for the vegetation, assets, states and transitions (VAST) narrative framework.

The elicitation of new expert knowledge of ecosystem condition used in developing HCAS v2.3 was approved by CSIRO's Social Science Human Research Ethics Committee (CSHREC) in accordance with the National Statement on Ethical Conduct in Human Research (2007, updated 2018): ethics clearance 115/22.

Ethics clearance (196/23) for reuse of expert's ecosystem condition data from prior projects was provided through CSHREC, as follows:

- "HCAS expert opinion blitz data (original ethics clearance, 025/18)" - for the analysis of *Expert assessments provided through the rapid transects approach*. Details about this method are provided in (Williams et al. 2020). We acknowledge Garry Cook, Tanya Doody, Michael Drielsma, Rod Fensham, Simon Ferrier, Justin Perry, Suzanne Prober, Chris Ware, and Matt White for their contributions.
- "Habitat condition data using expert elicitation (original ethics clearance, 004/17)" – for the analysis of *Expert assessments provided through the Habitat Condition Assessment Tool*. Details about this method are provided in (Pirzl et al. 2019). We acknowledge the 28 participants from the ecological science and natural resource management practitioner community for their contributed data.
- "Land and Ecosystem Accounts Project (LEAP): implementation phase 1 of the Valuing Parks Case Study Project (the Project) – oversight of ecology sub-project (original ethics clearance,

204/19)” – for use as reference sites in developing HCAS, refer Expert elicitation Section of this report. We acknowledge Kate Bennetts, Jean Dind, Doug Frood, Megan Good, Jamie Hearn, Paul McInerney, Gavin Rees and Genevieve Smith for their contributed data on the condition of ecosystem condition states.

We acknowledge other contributors to HCAS as listed in Supplementary Material B Table S3.

References

ABARES. 2016. The Australian Land Use and Management (ALUM) Classification Version 8. Prepared by the Australian Collaborative Land Use and Management Program Partners. Canberra,Australia: Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES), Department of Agriculture.

—. 2022. Land use of Australia 2010–11 to 2015–16, 250 m. Canberra, Australia: Australian Bureau of Agricultural and Resource Economics and Sciences.

ABS. 2023. Australian Population Grid 2022. Canberra, Australia: Australian Bureau of Statistics.

Alignier A, Deconchat M. 2013. Patterns of forest vegetation responses to edge effect as revealed by a continuous approach. Annals of Forest Science 70: 601-609.

Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA. 2002. Evaluating resource selection functions. Ecological Modelling 157: 281-300.

Brenton P, Pirzl R, Raisbeck-Brown N, Dickson F, White M, Warnick A, Williams KJ. 2018. The Habitat Condition Assessment Tool: BioCollect data collection portal hosted by the Atlas of Living Australia, https://biocollect.ala.org.au/hcat. Canberra, Australia: Atlas of Living Australia.

Carroll RJ, Ruppert D, Stefansky LA, Crainiceanu CM. 2006. Measurement Error in Nonlinear Models: A Modern Perspective, 2nd edition. Boca Raton , FL , U.S.A.: Chapman and Hall/CRC.

Chaudhary A, Brooks TM. 2018. Land Use Intensity-Specific Global Characterization Factors to Assess Product Biodiversity Footprints. Environmental Science & Technology 52: 5094-5104.

Chini LP, Hurtt G, Sahajpal R, Frolking S. 2020. GLM2 Code (Global Land-use Model 2) for generating LUH2 datasets (Land-Use Harmonization 2). Zenodo.

Chini LP, Hurtt GC, Sahajpal R, Frolking S, Goldewijk KK, Sitch S, Pongratz J, Poulter B, Ma L, Ott L. 2021a. LUH2-GCB2019: Land-Use Harmonization 2 Update for the Global Carbon Budget, 850-2019: ORNL Distributed Active Archive Center.

Chini LP, et al. 2021b. Land-use harmonization datasets for annual global carbon budgets. Earth Syst. Sci. Data 13: 4175-4189.

DAWE. 2018. Australia - Present Major Vegetation Groups - NVIS Version 5.1 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2020. Australia - Pre-1750 Major Vegetation Subgroups - NVIS Version 6.0 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

—. 2021. Collaborative Australian Protected Areas Database (CAPAD) 2020 - Terrestrial. Canberra, Australia: Australian Government Department of Agriculture, Water and the Environment.

DCCEEW. 2022. Indigenous Protected Areas (IPA) - Dedicated. Canberra, Australia: Department of Climate Change, Energy, the Environment and Water.

—. 2023. Australia - Present Major Vegetation Groups - NVIS Version 6.0 (Albers 100m analysis product). Canberra, Australia: Australian Government Department of Climate Change, Energy, the Environment and Water.

De Palma A, Hoskins A, Gonzalez RE, Börger L, Newbold T, Sanchez-Ortiz K, Ferrier S, Purvis A. 2021. Annual changes in the Biodiversity Intactness Index in tropical and subtropical forest biomes, 2001–2012. Scientific Reports 11: 20249.

Department of Environment and Science. 2022. Biodiversity status of 2019 remnant regional ecosystems - Queensland. Brisbane, Queensland: Queensland Herbarium, Department of Environment and Science.

Department of the Environment. 2014. Interim Biogeographic Regionalisation for Australia (IBRA), Version 7 (Subregions). Canberra, Australia: Australian Government Department of the Environment.

Di Cola V, et al. 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography 40: 774-787.

Didan K. 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. https://lpdaac.usgs.gov/: NASA EOSDIS Land Processes DAAC.

DNR, NVC. 2020. Native Vegetation Council (NVC) Bushland Assessment Manual. Adelaide, Australia: Department of Natural Resources, Government of South Australia.

Donohue RJ, McVicar TR, Roderick ML. 2009. Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. Global Change Biology 15: 1025-1039.

Donohue RJ, Mokany K, McVicar TR, O'Grady AP. 2022. Identifying management-driven dynamics in vegetation cover: Applying the Compere framework to Cooper Creek, Australia. Ecosphere 13: e4006.

DPIE. 2020. Biodiversity Assessment Method 2020 Operational Manual – Stage 1. Sydney, Australia: State of NSW through the Department of Planning, Industry and Environment.

Drielsma MJ, Ferrier S, Manion G. 2007. A raster-based technique for analysing habitat configuration: The cost-benefit approach. Ecological Modelling 202: 324-332.

Drielsma MJ, Love J, Taylor S, Thapa R, Williams KJ. 2022. General Landscape Connectivity Model (GLCM): a new way to map whole of landscape biodiversity functional connectivity for operational planning and reporting. Ecological Modelling 465: 109858.

DSE. 2004. Vegetation Quality Assessment Manual–Guidelines for applying the habitat hectares scoring method. Version 1.3. Melbourne, Australia: Victorian Government Department of Sustainability and Environment.

Eyre TJ, Kelly AL, Neldner VJ. 2017. Method for the Establishment and Survey of Reference Sites for BioCondition. Version 3. Brisbane, Australia: Queensland Herbarium, Department of Science, Information Technology and Innovation.

Eyre TJ, Kelly AL, Neldner VJ, Wilson BA, Ferguson DJ, Laidlaw MJ, Franks AJ. 2015. BioCondition: A Condition Assessment Framework for Terrestrial Biodiversity in Queensland. Assessment Manual. Version 2.2. Brisbane, Australia: Queensland Herbarium, Department of Science, Information Technology, Innovation and the Arts.

Ferrier S, Harwood TD, Ware C, Hoskins AJ. 2020. A globally applicable indicator of the capacity of terrestrial ecosystems to retain biological diversity under climate change: The bioclimatic ecosystem resilience index. Ecological Indicators 117: 106554.

Gallant JC, Austin J. 2012a. Relief - Elevation Range over 1000 m derived from 1" SRTM DEM-S v2. Data Collection data.csiro.au: CSIRO.

—. 2012b. Relief - Elevation Range over 300 m derived from 1" SRTM DEM-S (includes 3" product). Data Collection. data.csiro.au: CSIRO.

—. 2012c. Topographic Wetness Index (3" resolution) derived from 1" SRTM DEM-H. Data Collection. data.csiro.au: CSIRO.

Gallant JC, Dowling TI, Read AM, Wilson N, Tickle PK, Inskeep C. 2011. 1 second SRTM-derived Digital Elevation Models User Guide. Canberra, Australia: Australian Government Geoscience Australia.

Gallant JC, Austin J, Williams KJ, Harwood TD, King D, Nolan M, Mokany K. 2018. 9s soil and landform for continental Australia analysis of biodiversity pattern: aggregated from 3s data. Data Collection. data.csiro.au: CSIRO.

Geoscape Australia. 2020. PSMA - Transport & Topography - Street Network (Line) August 2020. Pages 70. Canberra, Australia: Geoscape Australia (formerly PSMA Australia).

Geoscience Australia. 2015. Landcover25 – Water (Water Observations from Space -WOfS) Product Description, V1.5 (D2015-8035). Canberra, Australia: Geoscience Australia.

Gilbert M, Nicolas G, Cinardi G, Van Boeckel TP, Vanwambeke SO, Wint GRW, Robinson TP. 2018. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. Scientific Data 5: 180227.

Giljohann KM, Drielsma M, Love J, Pinner L, Lyon P, Harwood TD, Williams KJ, Ferrier S. 2022. Technical description for a National Connectivity Index version 2.0 based on HCAS version 2.1. Canberra, Australia: CSIRO.

Giljohann KM, et al. 2023. Account-ready data: ecosystem condition in the Flinders, Norman and Gilbert river catchments in Queensland. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Grantham HS, et al. 2020. Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. Nature Communications 11: 5978.

Grundy MJ, Viscarra Rossel RA, Searle RD, Wilson PL, Chen C, Gregory LJ. 2015. Soil and Landscape Grid of Australia. Soil Research 53: 835-844.

Guerschman JP. 2019. Fractional Cover - MODIS, CSIRO algorithm. Version 3.1 (Dataset): Terrestrial Ecosystem Research Network (TERN).

Guerschman JP, Hill MJ. 2018. Calibration and validation of the Australian fractional cover product for MODIS collection 6. Remote Sensing Letters 9: 696-705.

Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135: 147-186.

Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G, White M. 2016a. Habitat Condition Assessment System: A new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. Methods in Ecology and Evolution 7: 1050-1059.

Harwood TD, Donohue RJ, Harman I, McVicar TR, Ota N, Perry J, Williams KJ. 2016b. 9s climatology for continental Australia 1976-2005: Summary variables with elevation and radiative adjustment. Data Collection. data.csiro.au: CSIRO.

Harwood TD, et al. 2023. 9-arcsecond gridded HCAS 2.3 (2001-2018) base model estimation of habitat condition and general connectivity for terrestrial biodiversity, ecosystem site condition, annual epochs and 18-year trends for continental Australia. Data Collection. Canberra, Australia: CSIRO.

Harwood TD, et al. 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. v7. Data collection. Canberra, Australia: CSIRO.

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199: 142-152.

Hoskins AJ, et al. 2020. BILBI: Supporting global biodiversity assessment through high-resolution macroecological modelling. Environmental Modelling & Software 132: 104806.

Hudson LN, et al. 2017. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecology and Evolution 7: 145-188.

Hurtt GC, et al. 2011. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Climatic Change 109: 117-161.

Hurtt GC, et al. 2020. Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. Geosci. Model Dev. 13: 5425-5464.

Hutchinson MF, Stein J, Stein J, Anderson H, Tickle P. 2008. GEODATA 9 second DEM and D8. Digital elevation model version 3 and flow direction grid. Gridded elevation and drainage data. Source scale 1:250 000. User guide (3rd ed). Canberra, Australia: Fenner School of Environment and Society, the Australian National University and Geoscience Australia, Australian Government.

Kass JM, Muscarella R, Galante PJ, Bohl CL, Pinilla-Buitrago GE, Boria RA, Soley-Guardia M, Anderson RP. 2021. ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. Methods in Ecology and Evolution 12: 1602-1608.

Landis JR, Koch GG. 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33: 159-174.

Laurance WF. 1991. Edge effects in tropical forest fragments: Application of a model for the design of nature reserves. Biological Conservation 57: 205-219.

Legendre P, Legendre L. 2012. Numerical ecology. Number 24 in Developments in Environmental Modelling. 3rd edition. Amsterdam: Elsevier.

Lehmann EA, Harwood TD, Williams KJ, Ferrier S. 2018. HCAS Activity 3a(1): Environmental Space Modelling – Methods Document. Canberra, Australia: CSIRO, Document Number EP184899.

Lehmann EA, Williams KJ, Harwood TD, Ferrier S. 2021. A not-too-technical introduction to the HCAS v2.x mechanics: a revised method for mapping habitat condition across Australia. Publication number EP211609. Canberra, Australia: CSIRO.

Lesslie R, Thackway R, Smith J. 2010. A national-level Vegetation Assets, States and Transitions (VAST) dataset for Australia (version 2.0) Canberra, Australia: Bureau of Rural Sciences, Australian Government.

Maus V, Giljum S, Gutschlhofer J, da Silva DM, Probst M, Gass SLB, Luckeneder S, Lieber M, McCallum I. 2020. A global-scale data set of mining areas. Scientific Data 7: 289.

Maus V, da Silva DM, Gutschlhofer J, da Rosa R, Giljum S, Gass SLB, Luckeneder S, Lieber M, McCallum I. 2022. Global-scale mining polygons (Version 2): PANGAEA.

McCallum K, Potter T, Cox B, Laws M, Bignall J, O'Neill S, Sparrow B. 2023. Condition Module, version 1.0. Adelaide, Australia: TERN.

Merow C, Smith MJ, Silander JA. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36: 1058-1069.

Michaels K. 2006. A Manual for Assessing Vegetation Condition in Tasmania, Version 1.0. Hobart: Resource Management and Conservation, Department of Primary Industries, Water and Environment.

Michaels K, Panek D, Kitchener A, eds. 2020. TASVEG VCA Manual: A manual for assessing vegetation condition in Tasmania, Version 2.0 Hobart, Australia: Natural and Cultural Heritage, Department of Primary Industries, Parks, Water and Environment, Tasmania.

Mokany K, Ware C, Harwood TD, Schmidt RK, Tetreault-Campbell S, Ferrier S. 2022. Habitat-based biodiversity assessment for ecosystem accounting in the Murray-Darling Basin. Conservation Biology 36: e13915.

Montandon LM, Small EE. 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. Remote Sensing of Environment 112: 1835-1845.

Mueller N, et al. 2016. Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. Remote Sensing of Environment 174: 341-352.

NASA. 2018a. Advanced Webinar: Accuracy Assessment of a Land Cover Classification, Feb 13-20, 2018, Exercise 1: Accuracy Assessment. Online webinar, http://arset.gsfc.nasa.gov/: NASA Applied Remote Sensing Training Program.

—. 2018b. Advanced Webinar: Accuracy Assessment of a Land Cover Classification, Feb 13-20, 2018, Exercise 2: Unbiased Error Estimation. Online webinar, http://arset.gsfc.nasa.gov/: NASA Applied Remote Sensing Training Program.

Newbold T, Hudson L, Purves DW, Scharlemann JPW, Mace G, Purvis A. 2012. call for data: PREDICTS: Projecting Responses of Ecological Diversity in Changing Terrestrial Systems. Frontiers of Biogeography 4: 155-156.

Newbold T, et al. 2016. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. Science 353: 288-291.

Njuki J, Poole J, Johnson N, Baltenweck I, Pali P, Lokman Z, Mburu S. 2011. Gender, livestock and livelihood indicators. Nairobi, Kenya: International Livestock Research Institute.

Oliver I, Dorrough J, Seidel J. 2021. A new Vegetation Integrity metric for trading losses and gains in terrestrial biodiversity value. Ecological Indicators 124: 107341.

Olofsson P, Foody GM, Stehman SV, Woodcock CE. 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment 129: 122-131.

OpenStreetMap Contributors. 2022. OpenStreetMap database [download.geofabrik.de/australia-oceania/australia.html, updated to 20 April 2022]. Available online: https://planet.openstreetmap.org/planet/full-history/: OpenStreetMap (OSM).

Parkes D, Newell G, Cheal D. 2003. Assessing the quality of native vegetation: The 'habitat hectares' approach. Ecological Management & Restoration 4: S29-S38.

Phillips SJ. 2022. Package ‘maxnet’: Fitting 'Maxent' Species Distribution Models with 'glmnet'. online: https://cran.r-project.org: CRAN-R Project.

Phillips SJ, Elith J. 2010. POC plots: calibrating species distribution models with presence-only data. Ecology 91: 2476-2484.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. 2017. Opening the black box: an open-source release of Maxent. Ecography 40: 887-893.

Pirzl R, et al. 2019. A National Reference Library of Expert Site Condition Assessments: Development and evaluation of method. Report to the Department of the Environment and Energy. Canberra, Australia: CSIRO.

Popp A, et al. 2017. Land-use futures in the shared socio-economic pathways. Global Environmental Change 42: 331-345.

Prober SM, Jordan R, Richards AE, Tetreault-Campbell S, Liedloff AC, Luxton S, MacFadyen S, Szetey K, Woodward E, Schmidt RK. 2023. Methods for developing an ecosystem classification and conceptual models for the Western Australia Wheatbelt. In: Ecosystem classification and conceptual models for the Western Australia Wheatbelt. A data collection from the Regional Ecosystem Accounting Pilot projects. Data Collection. Canberra, Australia: CSIRO.

Purvis A, et al. 2018. Modelling and Projecting the Response of Local Terrestrial Biodiversity Worldwide to Land Use and Related Pressures: The PREDICTS Project. Advances in Ecological Research 58: 201-241.

R Core Team. 2022. R: A language and environment for statistical computing. (v4.2.1) [Computer software]. https://www.R-project.org/: R Foundation for Statistical Computing.

Ramm F. 2022. OpenStreetMap Data in Layered GIS Format (Free shapefiles – 2022-03-31). available online, http://download.geofabrik.de/australia-oceania/australia.html: geofabrik.

Richards AE, Murphy H, Hayward J, Tetreault-Campbell S, Khan S, Giljohann K, Grainger D, Schirru E, Schmidt RK. 2023. Methods for developing ecosystem classification and conceptual models for ecosystem types in the Flinders, Norman and Gilbert river catchments. In: Ecosystem classification and conceptual models for the Flinders, Norman and Gilbert river catchments. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Roderick ML, Noble IR, Cridland SW. 1999. Estimating woody and herbaceous vegetation cover from time series satellite observations. Global Ecology and Biogeography 8: 501-508.

Rosenzweig ML. 1995. Species Diversity in Space and Time. Cambridge, UK: Cambridge University Press.

Schennach SM. 2016. Recent advances in the measurement error literature. Annual Review of Economics 8: 341-377.

Shaw MW. 1995. Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proceedings of the Royal Society of London. Series B: Biological Sciences 259: 243-248.

Sofaer HR, Hoeting JA, Jarnevich CS. 2019. The area under the precision-recall curve as a performance metric for rare binary events. Methods in Ecology and Evolution 10: 565-577.

Thackway R, Lesslie R. 2006. Reporting vegetation condition using the Vegetation Assets, States and Transitions (VAST) framework. Ecological Management & Restoration 7: S53-S62.

—. 2008. Describing and mapping human-induced vegetation change in the Australian landscape. Environmental Management 42: 572-590.

Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G. 2019. blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods in Ecology and Evolution 10: 225-232.

Valavi R, Guillera-Arroita G, Lahoz-Monfort JJ, Elith J. 2022. Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. Ecological Monographs 92: e01486.

Vermote E. 2015. MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006 [Data set]. https://lpdaac.usgs.gov/: NASA EOSDIS Land Processes DAAC.

Viscarra Rossel R, Chen C, Grundy M, Searle R, Clifford D, Odgers N, Holmes K, Griffin T, Liddicoat C, Kidd D. 2014a. Soil and Landscape Grid National Soil Attribute Maps - Effective Cation Exchange Capacity (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014b. Soil and Landscape Grid National Soil Attribute Maps - Soil Depth (3" resolution) - Release 1. v3. Data Collection. data.csiro.au: CSIRO.

—. 2014c. Soil and Landscape Grid National Soil Attribute Maps - Total Nitrogen (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014d. Soil and Landscape Grid National Soil Attribute Maps - Available Water Capacity (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014e. Soil and Landscape Grid National Soil Attribute Maps - pH - CaCl2 (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014f. Soil and Landscape Grid National Soil Attribute Maps - Silt (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014g. Soil and Landscape Grid National Soil Attribute Maps - Total Phosphorus (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014h. Soil and Landscape Grid National Soil Attribute Maps - Clay (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014i. Soil and Landscape Grid National Soil Attribute Maps - Organic Carbon (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014j. Soil and Landscape Grid National Soil Attribute Maps - Bulk Density - Whole Earth (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

—. 2014k. Soil and Landscape Grid National Soil Attribute Maps - Sand (3" resolution) - Release 1. v5. Data Collection. data.csiro.au: CSIRO.

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.

Warmerdam F, et al. 2021. gdalwarp: Image reprojection and warping utility. (29/12/2021 2021; https://gdal.org/programs/gdalwarp.html)

Warnick A, Raisbeck-Brown N, Mokany K, Williams KJ, White MD, Metcalfe D, Prober SM, Dickson F, Sparrow B, Pirzl R. 2019. Australian habitat image collection. v1. Data Collection. Canberra, Australia: CSIRO.

Werner TT, et al. 2020. A Geospatial Database for Effective Mine Rehabilitation in Australia. Minerals 10: 745.

White MD, Raisbeck-Brown N, Williams KJ, Warnick A, Mokany K, Brenton P, Sathya Moorthy S, Pirzl R. 2019. Habitat condition data for Australia from expert elicitation. v2. Data Collection. Canberra, Australia: CSIRO.

White MD, et al. 2023. Towards a continent-wide ecological site condition database using calibrated expert evaluations. Ecological Applications 33: e2729.

Wilford J, Searle R, Thomas M, Grundy M. 2015. Soil and Landscape Grid National Soil Attribute Maps - Depth of Regolith (3" resolution) - Release 2. v6. Data Collection. data.csiro.au: CSIRO. Report no.

Williams KJ, Belbin L, Austin MP, Stein J, Ferrier S. 2012. Which environmental variables should I use in my biodiversity model? International Journal of Geographic Information Sciences 26: 2009-2047.

Williams KJ, Harwood TD, Giljohann KM, Ferrier S, Lehmann EA, Ware C, Stewart SB, Tetreault-Campbell S, Schmidt RK. 2023a. Extended methods for developing account-ready data: ecosystem site condition in the Murray-Darling Basin. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2022-5750. Canberra, Australia: CSIRO.

Williams KJ, et al. 2023b. Extended methods used in developing the Habitat Condition Assessment System (HCAS) version 2.3, ecosystem condition account-ready data and experimental accounts for two mixed-use landscapes. A technical report from the Regional Ecosystem Accounting Pilot projects. Publication number: EP2023-1426. Canberra, Australia: CSIRO.

Williams KJ, et al. 2021. Habitat Condition Assessment System (HCAS version 2.1): Enhanced method for mapping habitat condition and change across Australia. Canberra, Australia: CSIRO.

Williams KJ, et al. 2023c. Methods for developing account-ready data: ecosystem condition in the Flinders, Norman and Gilbert river catchments in Queensland. In: Account-ready data: ecosystem condition in the Flinders, Norman and Gilbert river catchments in Queensland. A data collection from the Regional Ecosystem Accounting Pilot projects. Canberra, Australia: CSIRO. Data Collection.

Williams KJ, et al. 2020. Habitat Condition Assessment System: developing HCAS version 2.0 (beta). A revised method for mapping habitat condition across Australia. Canberra, Australia: Publication number EP21001. CSIRO Land and Water.

Xu T, Hutchinson MF. 2011. ANUCLIM Version 6.1 User Guide. Canberra: The Australian National University, Fenner School of Environment and Society.

—. 2013. New developments and applications in the ANUCLIM spatial climatic and bioclimatic modelling package. Environmental Modelling & Software 40: 267-279.