




Lessons learned from compiling a flora checklist for the Cape Peninsula, South Africa





Authors:

Anthony G. Rebelo^{1,2} 

Patricia M. Holmes^{3,4} 

Dian Spear^{2,5,6} 

Ronell R. Klopper^{7,8} 

Nicola J. van Wilgen^{4,5} 

Affiliations:

¹Threatened Species Research Unit, South African National Biodiversity Institute, Cape Town, South Africa

²Department of Biological Sciences, University of Cape Town, Cape Town, South Africa

³Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

⁴Centre for Invasion Biology, School of Climate Studies, Stellenbosch University, Stellenbosch, South Africa

⁵Cape Research Centre, South African National Parks, Cape Town, South Africa

⁶Centre for Sustainability Transitions, Stellenbosch University, Stellenbosch, South Africa

⁷Research and Scientific Services, Foundational Biodiversity Sciences, South African National Biodiversity Institute, Pretoria, South Africa

⁸H.G.W.J. Schweickerdt Herbarium, Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa

Checklists form an important component of biodiversity conservation, underpinning species monitoring, conservation planning and management prioritisation. Developing an accurate and taxonomically up-to-date plant checklist for a protected area requires the integration of diverse datasets, verification of species names and careful data management. Using the Cape Peninsula, South Africa, as a case study, we outline key steps and considerations in curating a comprehensive checklist for protected area management. We compiled data from multiple sources, including herbaria, museum collections, local conservation agencies, non-governmental organisations, universities, private conservancies, historical surveys and citizen science platforms such as iNaturalist. Key recommendations for checklist development include: (1) defining the geographic and taxonomic scope of the checklist, (2) identifying data sources, (3) optimising database design with standardised data collection and essential metadata fields, (4) having a verifiable taxonomic backbone, and (5) a clear workflow for working through each data source. In this process, it is important to retain, but flag erroneous records rather than deleting them, make provision for correctly assigning status information to extralimital and alien species, and use a local taxonomic expert to assist in decision-making required for resolving errors. Challenges encountered during the compilation of the checklist include resolving taxonomic inconsistencies, handling misidentifications, addressing orthographical errors in plant names and filtering out cultivated records from naturally occurring species – particularly in iNaturalist data. Our methodology provides practical guidelines to minimise these challenges, aligning with international best practices for checklist compilation and maintenance. By ensuring data completeness, accuracy and taxonomic consistency, we offer a framework that can benefit future biodiversity monitoring and conservation efforts.

Conservation implications: Accurate species checklists are crucial for informed conservation decisions. Standardised protocols for data validation and taxonomic accuracy enhance the reliability of biodiversity assessments, ultimately improving conservation outcomes in protected areas.

Keywords: alien plants; GBIF; indigenous plants; iNaturalist; taxonomy; taxonomic backbone; World Heritage Site.

Introduction

Accurate species lists are essential for research, monitoring and decision-making in protected areas. They are fundamental to understanding the flora of a locality (Droegge, Cyr & Larivée 1998) and can be used to prioritise focal species for monitoring (Rebelo et al. 2011a), identify endemic species that could be vulnerable to climate change (Pomoim et al. 2022), or provide warning of alien species that should be controlled (Foxcroft et al. 2009). Compiling species lists requires the collation of a diversity of distribution data (see Spear et al. 2023), which can be detailed within, extend beyond, or be totally outside of a target protected area.

Creating a comprehensive and taxonomically up-to-date plant list for a protected area requires that all available datasets be accessed, collated and cleaned. Species names in these datasets need to be checked against a taxonomically up-to-date taxonomic backbone, which includes

Corresponding author: Nicola van Wilgen, nicola.bredenkamp@sanparks.org

Dates: Received: 17 Mar. 2025 | Accepted: 27 Sept. 2025 | Published: 30 Nov. 2025

How to cite this article: Rebelo, A.G., Holmes, P.M., Spear, D., Klopper, R.R. & van Wilgen, N.J., 2025, 'Lessons learned from compiling a flora checklist for the Cape Peninsula, South Africa', *Koedoe* 67(1), a1856. <https://doi.org/10.4102/koedoe.v67i1.1856>

Copyright: © 2025. The Authors. Licensee: AOSIS. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

Note: Additional supporting information may be found in the online version of this article as Online Appendix 1.

Read online:



Scan this QR code with your smart phone or mobile device to read online.

taxonomic status information and accepted names for synonyms (Costello & Wieczorek 2014; Grenié et al. 2023; Spear et al. 2023). For South Africa, the South African National Plant Checklist (SANPC) can be used for plants, and the South African Animal Checklist can be used for some groups of animals. These lists are maintained by the South African National Biodiversity Institute (SANBI) and are extracts from the taxonomic tables of the Botanical Database of Southern Africa (BODATSA, running on BRAHMS software; Victor et al. 2023) and the Zoological Database of Southern Africa (ZODATSA, running on Specify software; Daly & Ranwashe 2023). The use of such taxonomic backbones is particularly important to deal with issues of synonyms in distribution data. They enable the identification of cases where a species has been named differently in different places, times or sources. A useful taxonomic backbone with unique identifiers that can be used for all groups in all countries globally is the taxonomic backbone of the Global Biodiversity Information Facility (GBIF), which is easily searchable using the GBIF species matching tool (<https://www.gbif.org/tools/species-lookup>).

A curated, local, annotated checklist is an essential component of the process needed to resolve the issue of synonymy during the collation of species distribution data and the development of species lists. Existing presence-absence data is rarely adequate to meet data needs for research and management, but such distribution data is invaluable to validate and refine checklists. The finer the location accuracy of the checklists or distribution data, the more useful they are.

Several challenges are known to exist with the collation of taxon data, including data cleaning (Coetzer & Hamer 2019), matching names (Grenié et al. 2023), taxonomic issues (Godfray 2002) and inaccurate locality data and identifications (López-Guillén et al. 2024). This article provides the methodology that was used to produce a species checklist for the flora of the Cape Peninsula, which is part of the Cape Floral Region Protected Areas World Heritage Site, Western Cape, South Africa (Rebello et al. 2025). We outline the challenges that were encountered during this process, and how they can be overcome, providing guidelines for similar projects.

Because of the importance of accurate checklists for research and management (Klopper 2025; Spear et al. 2023), there is a global movement towards creating species checklists for especially conservation areas where such lists are absent, or where existing checklists are not based on verified information (Cantú-Salazar, Gaston & Rouget 2013). In the past few decades, several updated checklists have been produced for countries, national parks and other protected areas in Africa (see, for instance, Abdelaal et al. 2018; Gosline et al. 2023; Hammanjoda et al. 2025; Houghnon et al. 2021; Meddour & Sahar 2021; Monteiro et al. 2022; Steyn, Bester & Bezuidenhout 2013; Van Wyk, Bezuidenhout & Jürgens 2024). We hope that the lessons learned through compiling a

checklist for the Cape Peninsula, South Africa, detailed here, will be useful to guide similar endeavours in Africa and other biodiverse regions of the world where floristic inventories are lacking.

Compiling a flora checklist

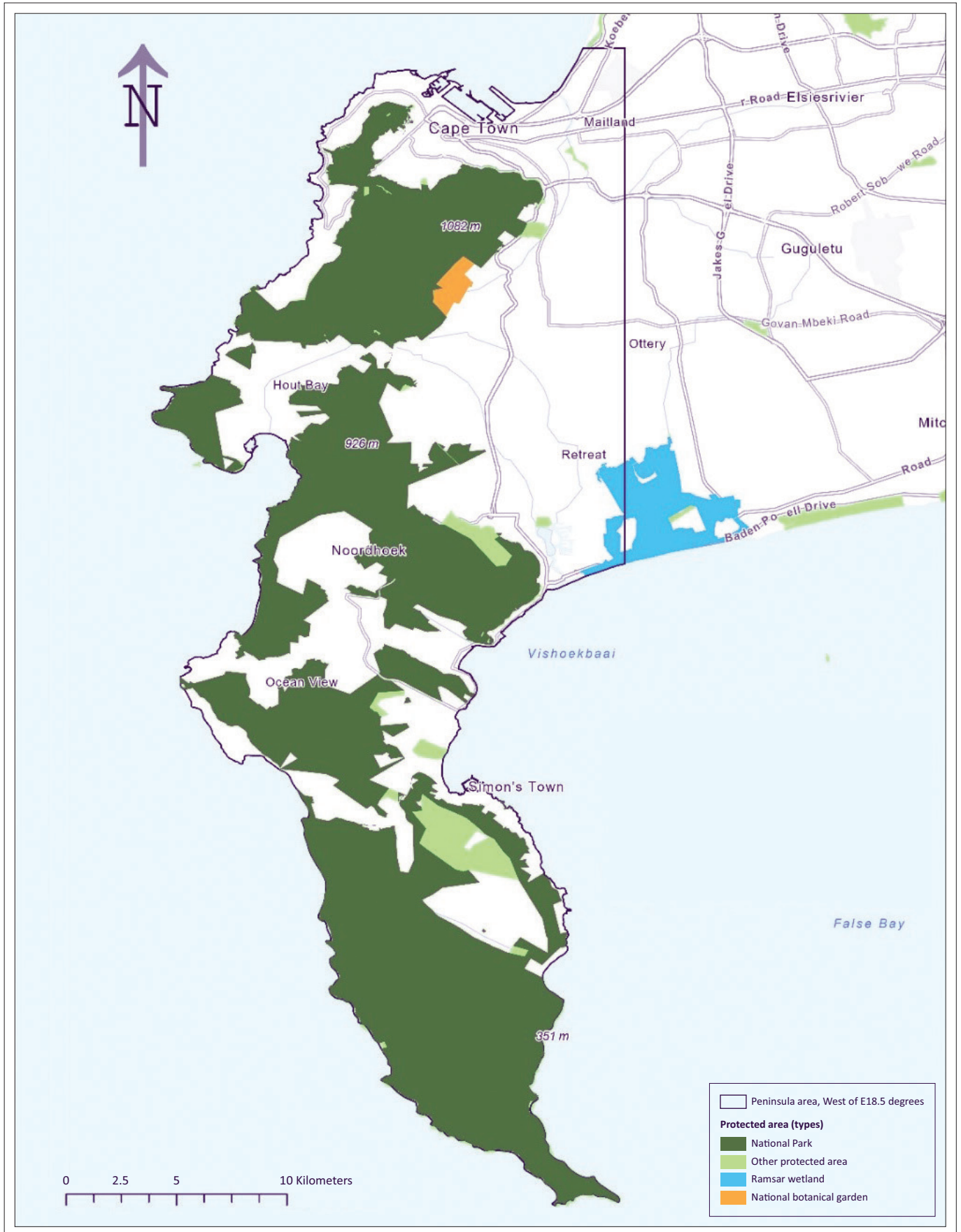
Location and scope

Table Mountain National Park (TMNP) is located on the Cape Peninsula, at the southwestern tip of South Africa (Figure 1). It is a topographically diverse mountainous area, home to a rich diversity of endemic plants and invertebrates (Cowling, MacDonald & Simmons 1996). Historical work (Adamson & Salter 1950) estimated that around 2285 plant species occur in the area, 158 of which are endemic (Helme & Trinder-Smith 2006), with more than 500 alien plants on the iNaturalist checklist for the Cape Peninsula and 291 alien invasives reported in TMNP (Spear et al. 2011). The peninsula falls within the City of Cape Town, which has a population of nearly 5 million people (4772846 in 2022 census; <https://census.statssa.gov.za/#/province/1/2>). The TMNP (currently ~25 km²) was established in 1998 to consolidate the management of conservation-worthy land on the peninsula, previously managed by various municipal and provincial authorities (SANParks 2015).

For the purposes of this assessment, species data were collected from relevant management authorities within the City of Cape Town, which include previous managers of the land (inter alia, Cape Nature and the City of Cape Town). To match the spatial scope of the previous definitive species list of the area (Adamson & Salter 1950), the spatial scope of the Cape Peninsula is here defined as the area south of -33.8°S and west of 18.5°E (Figure 1). This includes the areas directly adjacent to the park that represent the local biogeographic context. It is likely that some of the species in the surrounding areas will not occur in the park, but more importantly, it allows for park expansion and the identification of species that should be conserved in the park, and those that cannot be. This study focuses on terrestrial and aquatic (non-marine) higher plants, and includes data published and available up to June 2022.

Collation of flora occurrence data for the Cape Peninsula

All available distribution data on species and infraspecific taxa, including subspecies and varieties (hereafter referred to as 'taxa'), were collated. These data included specimen locality data from herbaria and museums, data from conservation agency databases, citizen science platforms, historical checklists for sites in the park, and survey and study plots, including research and student study plots from the University of Cape Town (UCT). Where lists contained relevant notes and distributional data, these were included (see Online Appendix 1: Supplementary Material 1 for full details of data curation steps). A summary of the main components of the data required and the sources is shown in Figure 2.



Source: Basemap from Google OpenStreetMap (Esri, TomTom, Garmin, FAO, NOAA, USGS, and the GIS user community).

FAO, Food and Agriculture Organization of the United Nations; NOAA, National Oceanic and Atmospheric Administration; USGS, the United States Geological Survey; GIS, Geographic Information System.

FIGURE 1: Map showing the Table Mountain National Park (dark green), other protected areas (light green and blue), and the Cape Peninsula as defined by the Flora of the Cape Peninsula. This area was used as the study domain (effectively the area west of 18.50°E and south of 33.89°S). The Table Mountain National Park covers approximately 50% of the study area.

Creating a checklist using a database

Although checklists can be created in MS Excel spreadsheets, it is beneficial to use a database, such as MS Access, to process data and involve someone in the process who is competent in designing and querying such a database (Figure 2). Additionally, it is essential to involve a biodiversity specialist familiar with the target taxa, who understands nomenclatural procedures and can check the data and provide insight into any taxonomic issues that arise (see Online Appendix 1: Figure S1). The Cape Peninsula flora data were collated in MS Access and the team compiling the data included botanists and taxonomists familiar with the taxa and databases.

Setting parameters and designing a database

Geographic area of interest

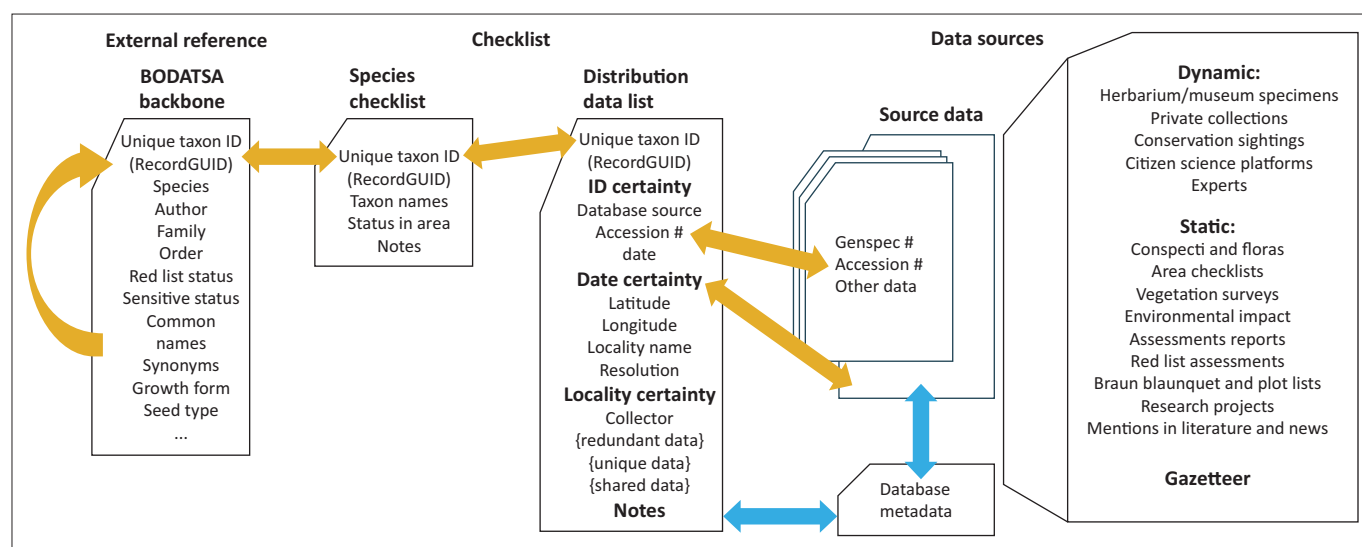
The first step in the process of developing a checklist is to determine the geographical area of interest. One may need several definitions (e.g. a bounding box, a specific mapped polygon, such as a nature reserve or wetland, a magisterial outer boundary including 'Surveys and Mapping' cadastres, Quarter Degree Square map units) to integrate different data types. Part of determining the geographical area of interest is determining whether your intended list is going to be exclusive, i.e. only relevant to the core area, or if the core area needs to be considered as part of a larger unit (e.g. vegetation type, phytocorion, biogeographic zone). For example, does a nature reserve represent purely itself, or should it conserve a representative sample of the surrounding regions (its domain), or include potential expansion plans? For the Cape Peninsula plant checklist, the historical area of the Cape Peninsula was included. However, it is unlikely that the national park will expand to this extent, and most of the area to the east is transformed (Rebelo et al. 2011b).

Identification of previous checklists

The next step in the process is to determine whether there are any existing checklists for the area of interest. If there is an existing checklist or published list, then this can be used as the starting point. For the Cape Peninsula, the City of Cape Town Conservation Unit had compiled a checklist in 2011 (Rebelo et al. 2011b). Although it included areas beyond the Cape Peninsula, this list was used as the starting point for creating an updated checklist. Two other existing checklists for the Cape Peninsula were available: Adamson and Salter (1950), which had been updated in the 1990s by Terry Trinder Smith for the Bolus Herbarium's Cape Peninsula collection, and the Friends of Silvermine Nature Area (FOSNA) which compiled a species identification aid to the Cape Peninsula in the early 2000s and has kept it current (see the FloraDoc application at <https://play.google.com/store/apps/details?id=za.co.koya.floradoc&gl=ZA>).

Database design

Before collating any data, a database needs to be designed that includes a table for the species checklist (checklist), a table for the distributional data (occurrence or specimen data), and a backbone table (for instance, the SANPC, for plants; Online Appendix 1: Figure S1, Supplementary Material 2). These can be linked relationally through a unique identifier (e.g. the BODATSA field RecordGUID). It should be noted that the Unique Taxon Identifier field (RecordGUID) is a unique identifier for taxon names and not observations within occurrence data. The backbone table is essential, especially if the list is to be curated to remain topical: this allows the latest nomenclatural and taxonomic changes to be automatically updated as needed based on the updates to the backbone (e.g. SANPC updates), without the entire checklist having to be manually processed. Bear in mind that citizen science data are increasing exponentially, and regular downloads will be required if the list is to remain useful – a static list will rapidly become obsolete, not least because of the updates in identification of observations.



BODATSA, Botanical Database of Southern Africa; GUID, Globally Unique Identifier; ID, identifier.

FIGURE 2: The main components of a taxon (i.e. species and infraspecific taxa) checklist database setup, including links to both an external taxonomic reference and to source data. The checklist comprises the taxon names and associated distribution data. The relative interactions of the components are shown (the metadata cross-referencing both components and elements). The gazetteer is especially important where localities need to be digitised.

Use of a taxonomic backbone

Scientific names in a checklist should be checked against an authoritative taxonomic backbone (Costello & Wieczorek 2014). For plants in South Africa, the latest annual version of the SANPC can be acquired from SANBI at <http://hdl.handle.net/20.500.12143/6880.3> (see also SANBI, Klopper & Winter 2025). Important fields in this dataset include: the TaxStatus field that indicates taxonomic status of a name (values include 'accepted', 'inclusive' [an accepted name at species level that has accepted infraspecific taxa, including subspecies or varieties or formas], 'synonym', and 'invalid'); and the AcceptedNameGUID field (which links synonyms to the currently accepted name through RecordGUID). The database tables are relationally linked through the Unique Taxon Identifier field (here RecordGUID).

Identification of relevant sources

Before being able to collate all relevant occurrence data, all suitable data sources for distribution data in the area need to be identified. Data sources include relevant national databases, citizen science (e.g. iNaturalist, Virtual Museums), literature sources (e.g. publications, checklists, surveys, field guides), and others such as Facebook, blogs, personal lists and non-accessioned nature reserve herbaria. In instances where online data are aggregated (too coarse), the raw data can often be obtained directly through the curator (e.g. special application for Virtual Museum data). Where possible, taxon specialists should check identifications online before downloading data from platforms like iNaturalist. In addition, care must be taken with choosing Research Grade or Needs ID observations: it is recommended that the Needs ID be checked by specialists using the curation tool on iNaturalist before downloading data. Rare, threatened and endemic species should be prioritised. Alternatively, one can download all data, but include their identification status. Note that obscured data are given 'false' localities on iNaturalist, and for smaller reserves may consequently not be included in the delimiting polygons, and these deprecated location data need to be filtered out of distribution lists.

We considered the following sources for plants of the Cape Peninsula: (1) BODATSA – specimens from the collections of the three SANBI herbaria (the National Herbarium [PRE], Compton Herbarium [NGB], and the KwaZulu-Natal Herbarium [NH]); (2) provincial conservation agency databases (CapeNature); (3) local conservation agency databases (City of Cape Town); (4) SANParks databases, including Cybertracker and C-more records; (5) non-governmental organisations (NGOs) and universities (especially the Bolus Herbarium [BOL] and library at the Botany Department of UCT); (6) private conservancies, stewardship, reserve and other land owners; (7) CREW (Custodians of Rare and Endangered Wildflowers) and national programmes (especially important for Red List taxa); (8) the National Vegetation Mapping project (SANBI), and any relevé or plot data collected for mapping the communities or vegetation types within reserves; (9) historical surveys: Protea Atlas Project (SANBI, Rebelo 1991;

and specialised plot data collected for monitoring or surveillance within the reserve e.g. Taylor 1983); and (10) iNaturalist data. The GBIF (<https://www.gbif.org/>) is another valuable source of distribution data. However, it was not used during this project, as we sourced distribution data directly from the original sources (e.g. BODATSA, iNaturalist), as these are the primary sources of data for South African plants on GBIF.

Determining the scope of the list

After identifying relevant sources, the scope of the list must be defined. This includes deciding which taxa to include (e.g. all plants, higher plants, flowering plants) and specifying the geographic extent, that is, whether to keep a full list for later spatial clipping, or to preclip an existing list for specific localities (using polygons if necessary). It also involves setting rules for excluding records outside the scope (e.g. planted records, alien species or data flagged as suspect). Then the relevant data fields need to be determined, including useful fields from the sources where they exist (e.g. altitude, population data, threats, notes). The institution and/or data owners can be contacted to request incorporation of the data into the database, specifying minimum fields and metadata (contact person, conditions of use, restrictions on data [including release dates for current academic studies], citation, etc.), timelines, and establish how issues with the data should be resolved (no action: data owner will fix and return updated data; fixes and notes made during processing to be returned to provider [or not], etc.).

Essential (compulsory) data fields that are needed for processing the distribution data for analyses need to be determined. Essential fields should include: (1) Data source, which assists in linking to contacts, documents, agreements, citation requirements (e.g. iNaturalist, City Biodiversity Database – see *Compiling metadata*); (2) Source accession: a unique accession number for each record, for verifying and updating data, as used by the data provider (e.g. 20576; B-1057); (3) Species name: names should be as complete as possible and for plants should include the author names; absence of author names in a dataset can result in problems where homonym names exist (identical names with different authors applied to different taxa where, for example, in plants, the younger homonym name is illegitimate, whereas for animals, it is invalid); (4) Latitude (in decimal degrees, here minus degrees to indicate South); (5) Longitude (in decimal degrees, here positive for East); (6) Location resolution, that is, accuracy or error, in metres radius (for checklists of areas, use half the maximum dimension in metres); (7) A location description, for checking location issues; (8) Collector(s) names; (9) Date (of observation or collection); and (10) Notes.

Compiling metadata

Next, it is vital to compile metadata for the data sources. The metadata should include (1) data source (provider); (2) institution; (3) type of records (distribution data, checklist, specimen-linked checklist, etc.); (4) number of records per type; (5) contact person; (6) restrictions on data use (including

obscured localities, time embargoes); (7) contracts if needed (including use limitations and confidentiality agreements); (8) turnaround times and procedures for fixing data queries; (9) citation requirements; and (10) use by dates (after which new data downloads should be done).

Defining the workflow for the process

Before starting to process data, a *modus operandi* needs to be developed. The purpose of the list will influence the workflow for checklist development. The assumption here is to compile a checklist and collate distribution data, but the purpose may be merely to compile a checklist without any distribution data. Essentially, the process is the same, except that already recorded species are not processed (for distribution information) from each dataset for a names-only checklist.

Each list from different sources can be processed individually, or all lists can be combined and processed together. The strategy and how much processing will be required will depend on the associated data, the size of the different databases and how clean the data are. As a rule, each list has its own challenges, and for larger datasets (over 1000 records), it is usually better to tackle each independently. It is advisable to process recent lists separately from those compiled before the 1970s, as the latter will have significant extra checking to be done, because of likely major taxonomic changes in the intervening period.

A decision needs to be made as to whether unresolved data (excluded records) will be discarded or retained in the distribution data with an 'exclude' flag. The problem with discarding data is that in any future updates, the data will have to be reprocessed. But if retained, future data updates can be filtered to include only records not already processed, and analyses can filter out 'discarded data'.

It is important to never delete a record. Incorrect species names not in the domain may arise from incorrect identifications, incorrect localities, incorrect updates and other causes. These need to be retained in the database; otherwise, these errors will recur in future updates and revisions and need to be fully processed again. Keep incorrect, old or synonym names within the database, and clearly flag and annotate them as such. These incorrect names will assist in future database curation, but they should not appear in outputs from the database, for example, updated checklists produced from the database.

A process needs to be established for how to deal with future updates. Herbarium and Virtual Museum data are dynamic since corrections are continuously made to locality and identification fields. Three options are available: (1) updating existing data before adding new data; (2) discarding the original distribution data and replacing it entirely with the current version of the data; or (3) ignoring changes to previous distribution data in future updates, adding only new data. This last option is not recommended, but unavoidable for static checklist data.

Distribution data often come with extra data fields that are superfluous to those requested. Usually, these extra data are never considered for future use, but may be useful. Decide if these fields will be discarded, summarised or linked. While these additional fields are seldom used, if they are not retained, they cannot be used. This includes redundant (and often historical) data best stored in the checklist or backbone (e.g. family, genus, Red List status, CITES [Convention on International Trade in Endangered Species of Wild Fauna and Flora] status). Additional information can be summarised in an 'other data' field. This allows easy access within the single data table and allows standardisation of fields, if possible. Additional information includes unique data (e.g. SANParks management block number) and shared fields (e.g. abundance and threats). Additional information fields in their original form can be linked to the distribution table, so that they can be accessed independently. This is only possible for data with a unique accession number for each record. It requires retaining and linking the original data within the database and requires retaining metadata of definitions and methods for each field.

Assembling the checklist

Updating existing checklists

Where there is an existing plant checklist, first check and update the taxonomy against the backbone (here SANPC) before processing any other datasets and annotate any changes. It is useful to have two fields, one with the original name from the source (verbatim name) and one for the updated accepted name according to the taxonomic backbone that you decide to use. Examples of other databases to use for checking the latest taxonomy for indigenous and exotic plant taxa in South Africa include the SANBI Biodiversity Advisor (<https://biodiversityadvisor.sanbi.org/>; this website is based on the SANPC, but is updated quarterly, whereas the SANPC is released annually), SANBI Red List (<http://redlist.sanbi.org/>), Plants of the World Online (<https://powo.science.kew.org/>), the World Flora Online (<http://www.worldfloraonline.org/>), the Taxonomic Name Resolution Service (<http://tnrs.iplantcollaborative.org/TNRSapp.html>), GBIF Species Look-up tool (<https://www.gbif.org/tools/species-lookup>), and iNaturalist (<https://www.inaturalist.org/>). Decide whether the checklist will only contain current names (in which case the SANPC will be used as an intermediate step in processing synonyms to their current names) or whether it will include synonyms (this will entail a much heavier workload, especially when processing new data, requiring internal updates whenever the SANPC is updated). A schematic illustrating the taxonomic checking process is shown in Online Appendix 1: Figure S1).

Minimum recommended fields in your database

It is recommended that there be a minimum set of curation fields in each of the Masterchecklist and Distribution data lists. Four fields need to be included in the checklist:

Unique Taxon Identifier, Curation Notes, Exclude and Residence Status. The Unique Taxon Identifier is the link to the current name in the taxonomic backbone (for the SANPC, this is RecordGUID); Curation Notes contain notes on problems and resolutions in dealing with the taxon; Exclude is used to flag taxa retained in the main list that do not form part of the checklist (e.g. errors in identification or distribution or taxonomy, because of taxon concept changes, misapplied names or synonyms); and Residence Status is the biogeographical status of the species to the focal area (e.g. endemic, indigenous, alien, naturalised, invasive). Keep national and provincial status separate from regional status if these are being considered. The following fields need to be included in the Distribution Data database: original species name (verbatim name from the source); accepted name; curation notes (ID issues for specimen); location notes (locality outside of known range, transcription or typing issues, recorded range extensions, and outliers); and date-related notes (missing or unlikely dates). The curation notes field is useful to annotate the processing of distribution data. These notes can also be incorporated into the checklist (Figure 2).

Dealing with alien species

There are numerous categories of alien species depending on their distribution and spread, including ecosystem transformers, invaders (spreading in natural areas), ruderals (spreading in croplands), naturalised (spreading, but only in disturbed areas), alien (established, not spreading), or extralimital (species native to South Africa, but documented outside of their historical native range). The latter are often assisted in their spread by people, but in some instances may spread naturally, particularly in modified environments. Note that for higher plants, the category 'naturalised' applies rather than 'extralimital' as humans generally introduce the species first before they spread, except in a few cases of natural dispersal.

A cut-off needs to be determined as to what will be considered (e.g. only transformer and invasive species or all alien species not planted, or including garden plants if regulations require these to be cleared within the area, etc.). Those species to be considered need to be added to the taxonomic backbone if they are new records. The SANPC includes indigenous and naturalised (including invasive) plant taxa occurring in South Africa. South African National Biodiversity Institute's Invasive Species Programme requests that potentially new aliens or names of alien plants that are not included in the backbone be reported to them for investigation and monitoring. Through this process, the coverage of naturalised aliens in the national checklist is being improved constantly. Note that checking an alien species' nomenclature and status takes much additional time, so if not required for the project, such species could either be excluded from the process or retained in the database without checking, but flagged as alien and excluded from analyses.

Adding new records to the species checklist

The addition of new records to the checklist requires manually checking if species being added to the checklist occur naturally in the domain. For example, many species in iNaturalist, Treemap, and some other datasets are planted, naturalised, or invasive. A field should be created to flag extralimital, alien, naturalised and invasive species. Check first to establish if species flagged in these categories are not merely misidentifications.

Document errors where species were historically recorded in the area, or routinely misidentified as species in the area, but have subsequently been revealed as identification or locality errors. Retain these in the list; otherwise, these records will continually turn up in future. Add these to the names backbone, with an explanatory note and set Exclude = True.

Send feedback to the curators of the taxonomic backbone (here, the SANPC curators) for any new records not in the backbone: As a rule, the SANPC is up-to-date, but especially new alien species may need to be added. Create a temporary accession number for these (and any other newly discovered species) and send the details to the Plant Checklist Coordinators for their attention. Consider publication of new records to the area, both indigenous and alien, as part of the checklist compilation process. This will provide published evidence that will assist the taxonomic coordinators and others compiling species lists or doing research on these plants to accurately reflect the distribution of the species.

Send data back to sources for cleaning and updating of specific information, if possible. Await updated data (or skip if updating is not part of the process). Ensure that any errors or problems are clearly included in the returned data, and that curators know what is required and the timelines involved and have agreed to these. Budget time and resources to follow up on these.

Maintaining the checklist

Even if compiling the checklist is a once-off process, provision should still be made to keep it taxonomically current. This can easily be achieved by periodically (annually or every 5 years) updating it according to the latest list from the backbone curators (here, the annually released SANPC update). Such updates will require three steps: Adding new names since the last update; Updating any new synonyms using the fields TaxStatus and AcceptedNameGUID to update RecordGUID (Unique Taxon Identifier); Relinking the data to the updated backbone (SANPC) through the Unique Taxon Identifier (RecordGUID). Where a different source from the SANPC is used for comparison, the relevant field names need to be identified to serve appropriate functionality, as well as possible changes in one or more of the standard field names in some sources (this may inevitably happen because of software or database structure updates of the source database).

Issues encountered and how to deal with them

During the assembly of data sources, it soon became clear that nomenclature and taxonomy were a major issue and that establishing a checklist of current names was a priority. However, even during the compilation of the Cape Peninsula checklist, some names became synonymised, requiring a dynamic checklist, linked to the SANPC. Issues that took time to resolve included (1) the need to clean source data, (2) some groups not being included in the taxonomic backbone, (3) the large number of planted species that were not indicated as planted in the iNaturalist data, (4) obscured locality data in iNaturalist, (5) needing to source subspecies information, (6) matching of species names, and (7) the mismatch in names between the South African Red List and SANPC.

Some of the problems that arose in the distribution data were related to difficulty in matching names to the taxonomic backbone, such as spelling errors, the use of outdated names, and the lack of names in the taxonomic backbone at the same taxonomic rank (see Online Appendix 1: Supplementary Material 3 for full details). Other problems included unassigned names, species only identified at a higher taxonomic rank, species that have been split, duplicate data and incorrectly identified species. These problems may be handled differently depending on the source and type of data: some need to be revisited or fixed later by the data owner. Despite the issues encountered, the bulk of the starting lists were relatively recent and clean (e.g. the list prepared for the City of Cape Town in 2017). Older lists, however, were associated with more issues. In total, across sources, there were 28 names that were irreconcilable (unfindable or ambiguous).

Matching scientific names

The use of standardised taxonomic names is vital when collating biodiversity data (Grenié et al. 2023; Spear et al. 2023). The issue of having a process to match scientific names was relatively easy to solve through the construction and use of 'friendly' species names – names without authors (although iNaturalist required a 'superfriendly' name using trinomials without rank), for linking to most of the data supplied. The checklist only needs to store the BODATSA RecordGUID as a way to link with the SANPC. Where the GBIF taxonomic backbone is used as source, the GBIF taxon key can be used as unique identifier. The following is an example of the use of 'friendly' and 'superfriendly' names that can be used for matching names in different lists. BODATSA includes a scientific name *Erica abietina* L. subsp. *atorosea* E.G.H.Oliv. & I.M.Oliv. or *Erica abietina* subsp. *atorosea* E.G.H.Oliv. & I.M.Oliv. The 'friendly' version of this name would be *Erica abietina* subsp. *atorosea* and the 'superfriendly' name would be *Erica abietina atorosea*. The latter 'superfriendly' version is only applicable for facilitating matching, as the name of an infraspecific plant should always include a rank denoting term (Art. 24.1; eds. Turland et al. 2025).

Mismatch of names in the South African Red List and South African National Plant Checklist

Because of differences in workflow, there is a mismatch between the South African (SA) Red List and SANPC. About 190 taxa on the Cape Peninsula had a Red List assessment for a taxon that is not recognised by that name in the SANPC. This requires an additional step of updating the SA Red List data based on SANPC synonyms to acquire the SA Red List status for these species. This issue is being addressed by SANBI and is a temporary problem. However, the SA Red List will continue to include undescribed taxa, which do not necessarily exist in the SANPC, and will need to be added to the checklist. Similar issues are likely to exist in other countries where Red List information is curated separately from information relating to a national taxonomic backbone.

Taxonomic backbone information for relevant taxonomic groups

Mosses, liverworts, hornworts, lycophytes, lichens, fungi, algae and seaweeds were excluded from our checklist and download of the SANPC. Currently, the official yearly release of the SANPC does include mosses, liverworts, hornworts and lycophytes (as well as ferns, gymnosperms and angiosperms), while marine macro algae will be included in future. BODATSA does contain names of fungi and lichens only for use during curation of the mycological collections of SANBI. Checklists for fungi and lichens can be obtained from the Mycology Unit of the Agricultural Research Council (National Collection of Fungi). However, data providers did not always exclude these groups, and in some cases did not know or record them as non-vascular plants. Considerable time was spent checking these names, which could have been automatically excluded had we included these groups in the backbone download initially, even though they were not to be included in the checklist.

Taxonomic issues and typos

Some taxonomic issues that may be encountered and need to be resolved include: the matching of varieties and subspecies to current name records in the database (terminal taxa of the species or clade) that may be a species, matching alternate names (synonyms) to current (accepted) names, and the presence of hybrids. When a name match is only found at a higher taxonomic rank, then an annotation should be included to indicate this. The accepted name for a synonym can be obtained from the SANPC. If a name is updated, then this can be indicated in the curation notes. It is recommended to exclude unassigned names or names flagged as 'c.f.' or 'aff.' or species that are linked to distribution data as 'ID insufficient'.

When species with infraspecific taxa are only identified at the species rank, finer resolution data may be required. In many families, subspecies are of conservation significance and identification to the finest resolution should be attempted. A decision should be made regarding who will investigate this further: for example, the list provider or the compiler? Species with only one infraspecific taxon in the focal area may be

automatically resolved; otherwise, further data will be needed. An annotation of 'subsp./var. exist' can be used, or if resolved, the note could read 'subsp./var. exist – updated as only one variety *xyz* occurs in checklist area'.

While compiling the Cape Peninsula checklist, about 3300 errors were flagged across the 10 datasets submitted, excluding the issue of a lack of subspecific information. A total of 260 non-terminal taxa required subspecific information to be sourced (see Online Appendix 1: Supplementary Material 3).

In some instances, species have been split since their inclusion in older sources as species concepts change over time. Depending on the area, it may be possible to automatically resolve this where the different new species are geographically distinct. Otherwise, an effort should be made to re-identify the specimens. For data without reference to specimens or photographs, the data should be excluded (annotate as 'Unresolvable taxonomic change'). These changes are hard to detect – they are often only picked up when distribution data are analysed. Species A, present in the focal area, may be split into several species, of which all, some, or none may occur in the area. Depending on the date of the checklist, different datasets may call the same species by different names. New species that were split off older species often do not regard the original name as a synonym, so these names may simply appear as incorrect without the historical perspective.

Synonyms can be a problem if there are two or more current names for the plant in the dataset. This will require checking prior to processing. This is especially problematic with data collated before the 1970s, but it is less of an issue for records identified over the last 20 years.

Hybrids are usually excluded from checklists. Only formally published natural hybrids (nothotaxa) are routinely included in the SANPC. However, 'frankenflora' (i.e. invasive hybrids formed by introduced species, such as *Protea neriifolia*, from outside the Cape Peninsula, hybridising with the indigenous *Protea lepidocarpodendron*) may need to be added to the checklist, if potential hybrid and invasive zones need to be monitored, or if reliability of field identification between the hybrid and its parent species is an issue.

Inaccurate identifications

In some datasets, there are species that have probably been incorrectly identified (Spear et al. 2023). This is a known problem for some herbarium (Goodwin et al. 2015) and iNaturalist data (López-Guillén et al. 2024). These usually require evaluation, but odd records of species well out of range need to be double-checked. They should not just be discarded though, as they may include range-extensions, new species (sometimes incorrectly assigned to an existing species), or other significant finds. How to proceed depends on the source: for checklists, little can be done, but for herbarium specimens, or annotated checklists with cited specimens, checks can be requested from the data providers

(e.g. see Online Appendix 1: Figure S1), but it is wise to budget time and money to follow up on this.

Without information on who identified taxa, it is difficult to determine confidence in the identification. Except for some herbarium specimen records, there is seldom information available on who made identifications, how long ago they were made and whether the identifications were verified. For checklists, this can be particularly fragmentary, as even herbarium data do not routinely include determination ('det.') data, or a measure of certainty of the identification. Citizen science may be unreliable. However, citizen science platforms often include information that can be used to determine certainty. They may include a few agreements and disagreements, lists of identifiers per specimen, and even measures of competence or consistency of identifiers. These fields should be included when data is downloaded.

Sometimes information exists on specimens that have been collected, but for which identifications are outstanding. If there are significant collections, especially for range-restricted or threatened species, it may be worthwhile budgeting for their identification to be accelerated. Usually, these data need to be excluded.

Orthographical errors and duplicate data

Correctly spelt taxon names are vital for processing plant data (Wagner 2016). Spelling and other orthographical errors should be corrected, but the change should be annotated with the original spelling, as there are a few cases of very similar names for related taxa that might require attention later.

Another issue is duplicate data (Petersen et al. 2021; Ribeiro et al. 2022), which can be very hard to detect. It is not unusual for key herbarium specimens to be sent to several herbaria (e.g. all relevant provincial and national herbaria), or for citizen scientists to submit the same records to several sites, sometimes backed up with herbarium specimens. Such duplicates may not matter for some analyses, but in other cases, they may be of concern.

Planted species and obscured iNaturalist data

It is problematic when iNaturalist records do not indicate species that are planted (López-Guillén et al. 2024). The iNaturalist data contained numerous species planted in gardens, but not marked as 'planted'. The large number of planted species in the dataset resulted from the iNaturalist City Nature Challenge that took place during the lockdown in 2020, where contributions of garden plants featured prominently (almost 39% of 34 000 observations). Normally, these would be a small fraction of the data. Garden plants (i.e. only cultivated) are not included in the national checklist, and considerable time was spent checking these names during compilation of the Cape Peninsula checklist, where observations were not marked as planted on iNaturalist. This may be because of the reluctance of observers to mark observations as planted, as these are handled differently on

the platform (Richardson & Potgieter 2024), for example, cultivated species are removed from the 'Needs ID' queue, delaying time to identification. However, any reviewer can mark these as cultivated, and appointing an intern or student to clean the relevant iNaturalist data before it is downloaded would save considerable effort.

iNaturalist contains several 'obscured' observations without accurate locality information (Contreras-Díaz et al. 2023; Spear et al. 2023). Obtaining the actual localities turned out to be far more difficult than anticipated. Two types of obscurity exist: taxonomic obscurity based on the Red List, where actual localities should be easy to acquire in theory, and user obscurity, where the only ways to obtain these data are a direct request to the observer. Obtaining permission is a laborious process with no guarantee that the user will respond. In theory, if national curators are also the curators of data on iNaturalist (e.g. if the SANBI curated the southern African community on iNaturalist), such updates to locality information should be available biannually from the national curator.

Conclusion

Many of the challenges encountered during creating and updating a species checklist for the Cape Peninsula, South Africa, are not specific to this project and have inevitably been encountered by other similar projects in the past. Our methodology, and specifically the suggestions and guidelines we provide, can be applied to any future project where a species inventory is an output, regardless of the taxonomic, geographical, or temporal scope and will hopefully assist future project administrators to avoid or minimise the effect of these issues. The challenges we encountered and recommendations we propose are in line with other international guidelines that exist with regard to checklist compilation and maintenance. These recommendations include obtaining and validating all available data sources, correcting typographical and other errors, and ensuring the use of a suitable taxonomic backbone to provide unique identifiers and determine accepted names.

Acknowledgements

The following individuals (and organisations) are thanked for providing data: Trevor Adams, Chad Cheney (SANParks), Llewellyn Jacobs (CapeNature), Leighan Mossop (City of Cape Town), Corinne Merry (Friends of Silvermine Nature Reserve), Hannelie Snyman, Brenda Daly (SANBI), Ismail Ebrahim (CREW), Victoria Willman (Millennium Seed Bank), Rene Navarro (University of Cape Town), Jasper Slingsby (SAEON: South African Environmental Observation Network), iNaturalist. Zishan Ebrahim is thanked for creating the map of the Cape Peninsula. This work was funded by the Table Mountain Development Fund.

Competing interests

The author reported that they received funding from JRS Biodiversity Foundation and Table Mountain Fund, which

may be affected by the research reported in the enclosed publication. The author has disclosed those interests fully and has implemented an approved plan for managing any potential conflicts arising from their involvement. The terms of these funding arrangements have been reviewed and approved by the affiliated university in accordance with its policy on objectivity in research.

Authors' contributions

N.J.v.W and A.G.R. conceived of the presented idea. All authors attended a workshop on suggested species listing process and multiple meetings on processes related to list cleaning. A.G.R. and P.M.H. developed the workflow for working through each data source. R.R.K., P.M.H. and A.G.R. worked together to correctly assign taxonomy. R.R.K., D.S. and N.J.v.W. revised the suggested processes and protocols and compiled figures. All authors discussed the results and contributed to the final article.

Ethical considerations

This article followed all ethical standards for research without direct contact with human or animal subjects.

Funding information

Dian Spear was funded by the JRS Biodiversity Foundation (grant 60916), which is also supporting the development of species data management tools in SANParks. P.M.H. was supported for this work through a grant from the Table Mountain Fund (grant TM 5856).

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials, and will also be made available through the SANParks Biodiversity Information Management System: <https://bims.sanparks.org/>.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. They do not necessarily reflect the official policy or position of any affiliated institution, funder, agency or that of the publisher. The authors are responsible for this article's results, findings and content.

References

- Abdelaal, M., Fois, M., Fenu, G. & Bacchetta, G., 2018, 'Critical checklist of the endemic vascular plants of Egypt', *Phytotaxa* 360(1), 19–34. <https://doi.org/10.11646/phytotaxa.360.1.2>
- Adamson, R.S. & Salter, T.M., 1950, *Flora of the Cape Peninsula*, Juta & Co., Ltd., Cape Town.
- Cantú-Salazar, L., Gaston, K.J. & Rouget, M., 2013, 'Species richness and representation in protected areas of the Western Hemisphere: Discrepancies between checklists and range maps', *Diversity and Distributions* 19(7), 782–793. <https://doi.org/10.1111/ddi.12034>
- Coetzer, W. & Hamer, M., 2019, 'Managing South African biodiversity research data: Meeting the challenges of rapidly developing information technology', *South African Journal of Science* 115(3–4), 1–5. <https://doi.org/10.17159/sajs.2019/5482>

- Contreras-Díaz, R.G., Nori, J., Chiappa-Carrara, X., Peterson, A.T., Soberón, J. & Osorio-Olvera, L., 2023, 'Well-intentioned initiatives hinder understanding biodiversity conservation: Cloaked iNaturalist information for threatened species', *Biological Conservation* 282, 110042. <https://doi.org/10.1016/j.biocon.2023.110042>
- Costello, M.J. & Wiczorek, J., 2014, 'Best practice for biodiversity data management and publication', *Biological Conservation* 173, 68–73. <https://doi.org/10.1016/j.biocon.2013.10.018>
- Cowling, R.M., MacDonald, I.A.W. & Simmons, M.T., 1996, 'The Cape Peninsula, South Africa: Physiographical, biological and historical background to an extraordinary hot-spot of biodiversity', *Biodiversity and Conservation* 5(5), 527–550. <https://doi.org/10.1007/BF00137608>
- Daly, B. & Ranwashe, F., 2023, 'South Africa's initiative toward an integrated biodiversity data portal', *Frontiers in Ecology and Evolution* 11, e1124928. <https://doi.org/10.3389/fevo.2023.1124928>
- Droege, S., Cyr, A. & Larivée, J., 1998, 'Checklists: An under-used tool for the inventory and monitoring of plants and animals', *Conservation Biology* 12(5), 1134–1138. <https://doi.org/10.1046/j.1523-1739.1998.96402.x>
- Foxcroft, L.C., Richardson, D.M., Rouget, M. & MacFadyen, S., 2009, 'Patterns of alien plant distribution at multiple spatial scales in a large national park: Implications for ecology, management and monitoring', *Diversity and Distributions* 15(3), 367–378. <https://doi.org/10.1111/j.1472-4642.2008.00544.x>
- Godfrey, H.C.J., 2002, 'Challenges for taxonomy. The discipline will have to reinvent itself if it is to survive and flourish', *Nature* 417(6884), 17–19. <https://doi.org/10.1038/417017a>
- Goodwin, Z.A., Harris, D.J., Filer, D., Wood, J.R. & Scotland, R.W., 2015, 'Widespread mistaken identity in tropical plant collections', *Current Biology* 25(22), R1066–R1067. <https://doi.org/10.1016/j.cub.2015.10.002>
- Gosline, G., Bidault, E., Van der Burg, X., Cahen, D., Challen, G., Condé, N. et al., 2023, 'A taxonomically-verified and vouchered checklist of the vascular plants of the Republic of Guinea', *Scientific Data* 10(1), a327. <https://doi.org/10.1038/s41597-023-02236-6>
- Grenié, M., Berti, E., Carvajal-Quintero, J., Dädlow, G.M.L., Sagouis, A. & Winter, M., 2023, 'Harmonizing taxon names in biodiversity data: A review of tools, databases and best practices', *Methods in Ecology and Evolution* 14(1), 12–25. <https://doi.org/10.1111/2041-210X.13802>
- Hammanjoda, S.A., Nulit, R., Yap, C.K., Nformi, U.B., Nodza, G.I., Saba, A.O. et al., 2025, 'Diversity assessments of native and alien vascular plants in Gashaka Gumti National Park, northeast Nigeria', *Phytotaxa* 717(2), 177–188. <https://doi.org/10.11646/phytotaxa.717.2.4>
- Helme, N.A. & Trinder-Smith, T.H., 2006, 'The endemic flora of the Cape Peninsula, South Africa', *South African Journal of Botany* 72(2), 205–210. <https://doi.org/10.1016/j.sajb.2005.07.004>
- Houghnon, A., Adomou, A.C., Dosling, W.D. & Adeonipekun, P.A., 2021, 'A checklist of vascular plants of Ewe-Adakplame Relic Forest in Benin, West Africa', *PhytoKeys* 175, 151–174. <https://doi.org/10.3897/phytokeys.175.61467>
- Klopper, R.R., 2025, 'Consensus classifications are crucial for conservation: How CITES utilises checklists', *Taxon* 74(4), 759–767. <https://doi.org/10.1002/tax.13348>
- López-Guillén, E., Herrera, I., Bensed, B., Gómez-Bellver, C., Ibáñez, N., Jiménez-Mejías, P. et al., 2024, 'Strengths and challenges of using iNaturalist in plant research with focus on data quality', *Diversity* 16(1), e42. <https://doi.org/10.3390/d16010042>
- Meddour, R. & Sahar, O., 2021, 'Floristic inventory of Djurdjura National Park, northern Algeria: A first checklist of its vascular flora', *Phytotaxa* 490(3), 221–238. <https://doi.org/10.11646/phytotaxa.490.3.1>
- Monteiro, F., Da Costa, E., Kissanga, R., Costa, J.C. & Catarino, L., 2022, 'An annotated checklist of the vascular flora of Quiçama National Park, Angola', *Phytotaxa* 557(1), 1–67. <https://doi.org/10.11646/phytotaxa.557.1.1>
- Petersen, T.K., Speed, J.D., Grøtan, V. & Austrheim, G., 2021, 'Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection', *Ecological Solutions and Evidence* 2(1), e12048. <https://doi.org/10.1002/2688-8319.12048>
- Pomoim, N., Hughes, A.C., Trisurat, Y. & Corlett, R.T., 2022, 'Vulnerability to climate change of species in protected areas in Thailand', *Scientific Reports* 12(1), e5705. <https://doi.org/10.1038/s41598-022-09767-9>
- Rebello, A., Holmes, P.M., Klopper, R., Spear, D. & van Wilgen, N.J., 2025, 'Plant checklist for Table Mountain National Park and surrounding areas', *Koedoe* 67(1), a1855. <https://doi.org/10.4102/koedoe.v67i1.1855>
- Rebello, A.G., 1991, *Protea Atlas Manual: Instruction booklet to the Protea Atlas Project*, University of Cape Town, Cape Town.
- Rebello, A.G., Holmes, P.M., Dorse, C. & Wood, J., 2011b, 'Impacts of urbanization in a biodiversity hotspot: Conservation challenges in Metropolitan Cape Town', *South African Journal of Botany* 77(1), 20–35. <https://doi.org/10.1016/j.sajb.2010.04.006>
- Rebello, T.G., Freitag, S., Cheney, C. & McGeoch, M.A., 2011a, 'Prioritising species of special concern for monitoring in Table Mountain National Park: The challenge of a species-rich, threatened ecosystem', *Koedoe: African Protected Area Conservation and Science* 53(2), 1–14. <https://doi.org/10.4102/koedoe.v53i2.1019>
- Ribeiro, B.R., Velazco, S.J.E., Guidoni-Martins, K., Tassarolo, G., Jardim, L., Bachman, S.P. et al., 2022, 'bdc: A toolkit for standardizing, integrating and cleaning biodiversity data', *Methods in Ecology and Evolution* 13(7), 1421–1428. <https://doi.org/10.1111/2041-210X.13868>
- Richardson, D.M. & Potgieter, L.J., 2024, 'A living inventory of planted trees in South Africa derived from iNaturalist', *South African Journal of Botany* 173(9), 365–379. <https://doi.org/10.1016/j.sajb.2024.08.006>
- SANParks, 2015, *Table Mountain National Park, park management plan 2015–2025*, SANParks, Pretoria.
- South African National Biodiversity Institute (SANBI), Klopper, R. & Winter, P., 2025, *South African National Plant Checklist: 2025 official yearly release, Version 2025 [Data set]*, Zenodo. <https://doi.org/10.5281/zenodo.15050847>
- Spear, D., McGeoch, M.A., Foxcroft, L.C. & Bezuidenhout, H., 2011, 'Alien species in South Africa's national parks: Checklist', *Koedoe: African Protected Area Conservation and Science* 53(1), 1–4. <https://doi.org/10.4102/koedoe.v53i1.1032>
- Spear, D., van Wilgen, N.J., Rebello, A.G. & Botha, J.M., 2023, 'Collating biodiversity occurrence data for conservation', *Frontiers in Ecology and Evolution* 11, 1037282. <https://doi.org/10.3389/fevo.2023.1037282>
- Steyn, H.M., Bester, S.P. & Bezuidenhout, H., 2013, 'An updated plant checklist for Tankwa Karoo National Park, South Africa', *South African Journal of Botany* 88, 247–251. <https://doi.org/10.1016/j.sajb.2013.07.018>
- Taylor, H., 1983, 'The vegetation of the Cape of Good Hope Nature Reserve', *Bothalia* 14(3/4), 779–784. <https://doi.org/10.4102/abc.v14i3/4.1241>
- Turland, N.J., Wiersema, J.H., Barrie, F.R., Gandhi, K.N., Gravendyck, J., Greuter, W. et al. (eds.), 2025, *International code of nomenclature for algae, fungi, and plants (Madrid Code)*, Regnum Vegetabile 162, The University of Chicago Press, Chicago, IL.
- Van Wyk, P., Bezuidenhout, H. & Jürgens, N., 2024, 'A checklist of indigenous flora in the Richtersveld National Park confirms high plant diversity in the arid north-western tip of South Africa', *Koedoe: African Protected Area Conservation and Science* 66(1), a1822. <https://doi.org/10.4102/koedoe.v66i1.1822>
- Victor, J.E., Klopper, R.R., Winter, P.J.D. & Le Roux, M.M., 2023, 'The plant checklist: Building the foundation of botanical knowledge in South Africa', *Taxon* 73(4), 943–948. <https://doi.org/10.1002/tax.13169>
- Wagner, V., 2016, 'A review of software tools for spell-checking taxon names in vegetation databases', *Journal of Vegetation Science* 27(6), 1323–1327. <https://doi.org/10.1111/jvs.12432>