South African Journal of Botany 2003, 69: 1–6

SOUTH AFRICAN JOURNAL OF BOTANY
ISSN 0254–6299

ACKDAT: a digital spatial database of distributions of South African plant species and species assemblages

MC Rutherford*, LW Powrie and GF Midgley

Kirstenbosch Research Centre, National Botanical Institute, Private Bag X7, Claremont 7735, South Africa
** Corresponding author, e-mail: rutherford@nbi.ac.za*

Received 12 November 2002, accepted in revised form 20 November 2002

The fieldwork that underpins John Acocks' work on the Veld Types of South Africa, was recorded in a set of field notebooks, which were preserved after Acocks' retirement. These are adequately geo-referenced to allow the construction of a relational database linking more than 7 000 species' presence, abundance and site location data at more than 3 000 sites throughout much of South Africa. The database is described with reference to data quality, representivity and spatial precision. Some appropriate and inappropriate types of uses of the data are illustrated.

Introduction

The description of the veld types of South Africa was always considered an unfinished work by John Acocks (Hoffman and Cowling 2003). This substantial contribution to South African botany and ecology was based on a career of field excursions, during which Acocks used his prodigious knowledge of South African plant species to record mostly the full complement of vascular plant species at selected sites throughout the country. Acocks collected species presence and abundance data at several thousand sites throughout South Africa (Figure 1). Many sites were sampled exhaustively, and others less intensively. He revisited about 10% of his sample sites in order to supplement his records for species at a site. The data were collected between 1936 and 1977, and by 1953 led to his publication of Veld Types of South Africa (Acocks 1953) that was reprinted with photographs added and names updated (Acocks 1975, 1988). His post 1953 data were used for a revision of the veld types in the western half of South Africa that was never published. Although the publication of Acocks (1953) records the first distillation of this almost unprecedented scope of data collection of its kind, Acocks' field notes (Figure 2) remained preserved in the archives of the National Botanical Institute (NBI) after his retirement. These originals are clearly legible and adequately spatially referenced to serve as the basis for a database of plant species distributions and abundance, now realised in the form of ACKDAT. The ACKDAT product was launched in 1994 (O'Callaghan *et al*. 1994) with Version 1.0 being made available on NBI's FTP site in 1997, together with documentation. This version is a DOS-based version using DataEase as a database platform. The period since 1997 has seen corrections and further verification, leading to a soon to be released Version 1.1 using Windows and MSAccess. In an earlier account of ACKDAT (O'Callaghan 2000), information and various statistics were provided that subsequent checks have shown need to be updated.

Database description

ACKDAT is a relational database, linking lists of vascular plant species to spatial locations, and the abundance of the species at the site. ACKDAT Version 1.0 does not include abundance data. The current database contains 289 414 records from 3 098 sites of which 1 311 are indicated by Acocks as not having a comprehensive species list. Site data in ACKDAT were collected from 1937 to 1977 with greatest number of records in 1959 (Figure 3). The database contains 7 415 species (binomials, i.e. no intraspecific taxa) within 1 523 genera. Some naturalised alien species are also included. The abundance of each species at each site is given in 20 categories of estimation that correspond to Acocks' symbols of abundance that translate to numbers of plant individuals per hectare. Using unmodified translation, these numbers range from 1 to 15 122 520 per hectare. The conversion of these numbers from those expressed per morgen to those per hectare was erroneous in Acocks (1975) but was corrected by Killick (1982) and corrected in Acocks (1988). Acocks also provides 16 habitat categories which have been included in the database. However, not applied in the database are his 'habitat reduction factors' for the 16 categories and three additional categories of patchiness. These arbitrarily reduce, for example, estimated abundance by a two on a west aspect slope, by three on a south aspect and by 1/100 on a rocky place. These reduction factors are still under review.

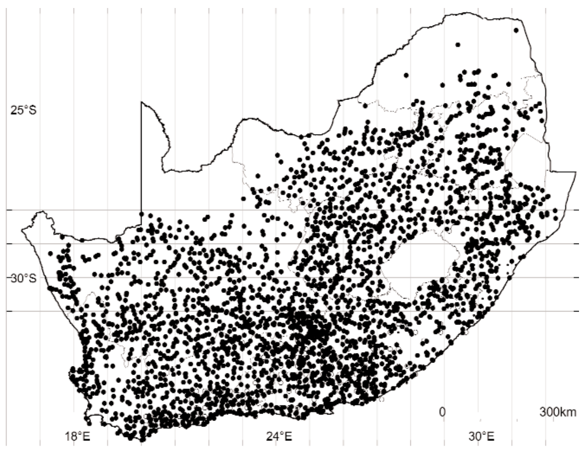

Figure 1: Distribution of Acocks' sampling sites (within South Africa) included in ACKDAT

Quality control

The original hand written notes (Figure 2) of Acocks were used by a data capturing company to help avoid misinterpretation of handwriting or changes made by Acocks that might have been less clear when using copies. Every record in the captured product was then carefully checked against the original by NBI teams of not less than two persons that always included a professional botanist. Other teams checked the locality of each sample site, as indicated by an unnumbered dot on his original set of 1:500 000 maps, against his original description of the locality of the site. A small number of sites could not be confidently confirmed and remained outside ACKDAT.

Certainly not all of Acocks data were accepted by the ACKDAT database. The 289 414 records make up 84% of the number of records with confirmed geo-referencing. Most of the remaining 54 840 records had only the genus name listed and the species were thus mostly unknowable and were excluded. Some could be traced through being a monotypic genus or through clear cut and unique distribution patterns of the species of the genus. Other exclusions from the database were records with any plant names with an appended question mark, brackets, or any other indicator of uncertainty.

ACKDAT attempted to update all Acocks plant names to current valid names using the PRECIS (National Herbarium, Pretoria (***PRE***) **C**omputerised **I**ndex **S**ystem (Arnold and De Wet 1993)) database that represents the computerised holdings of the National Herbarium (Gibbs Russell 1985a) and is continually being updated. In the limited number of cases where a species had subsequently been split into two or more species, it was not included in ACKDAT. Also excluded were several records where Acocks' plant names (e.g. *Galenia humifusa* and *Trachypogon plumosus*) could not be matched with any names or with synonyms in PRECIS. This does not detract from Acocks' own diligent quality control where his approximately 7 500 recorded species in ACKDAT

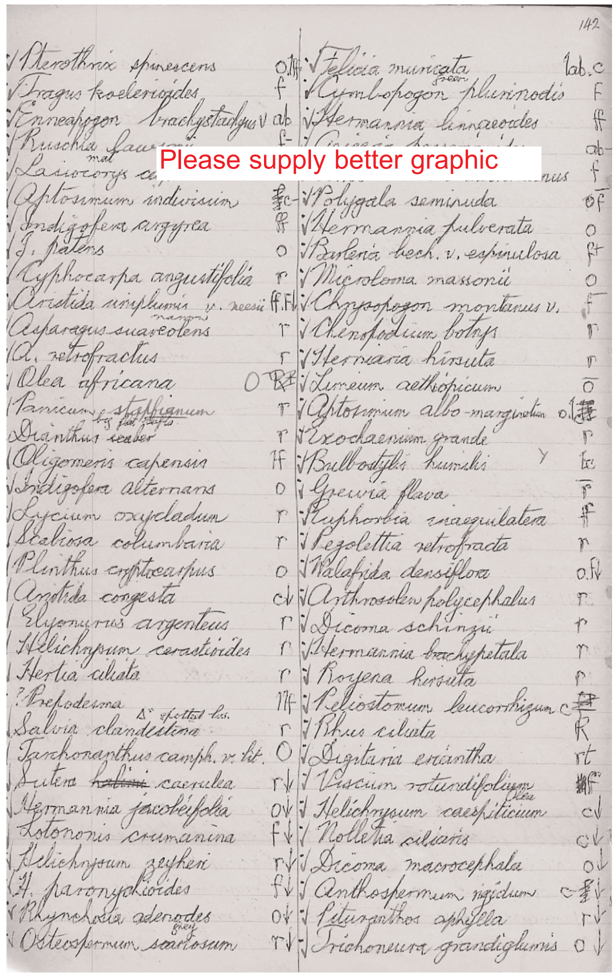

Figure 2: An example of a page of Acocks' field notes with his recorded species and symbols indicating their estimated abundance and habitat type for part of a sampling site

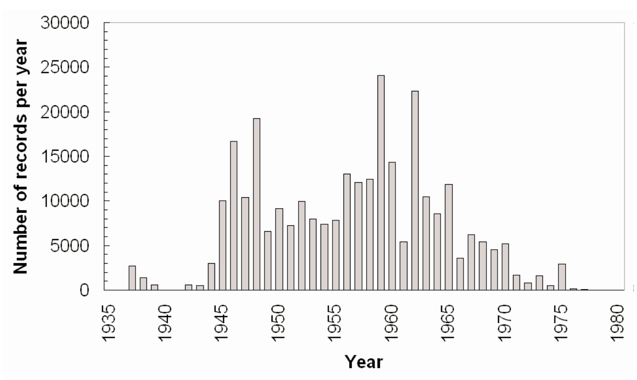

Figure 3: Acocks' sampling effort over time, as reflected by the number of records in ACKDAT per year

compare to over 25 000 collections that he made (Killick 1980). Further work on relating information from the voucher specimens to these unknown names might lead to their

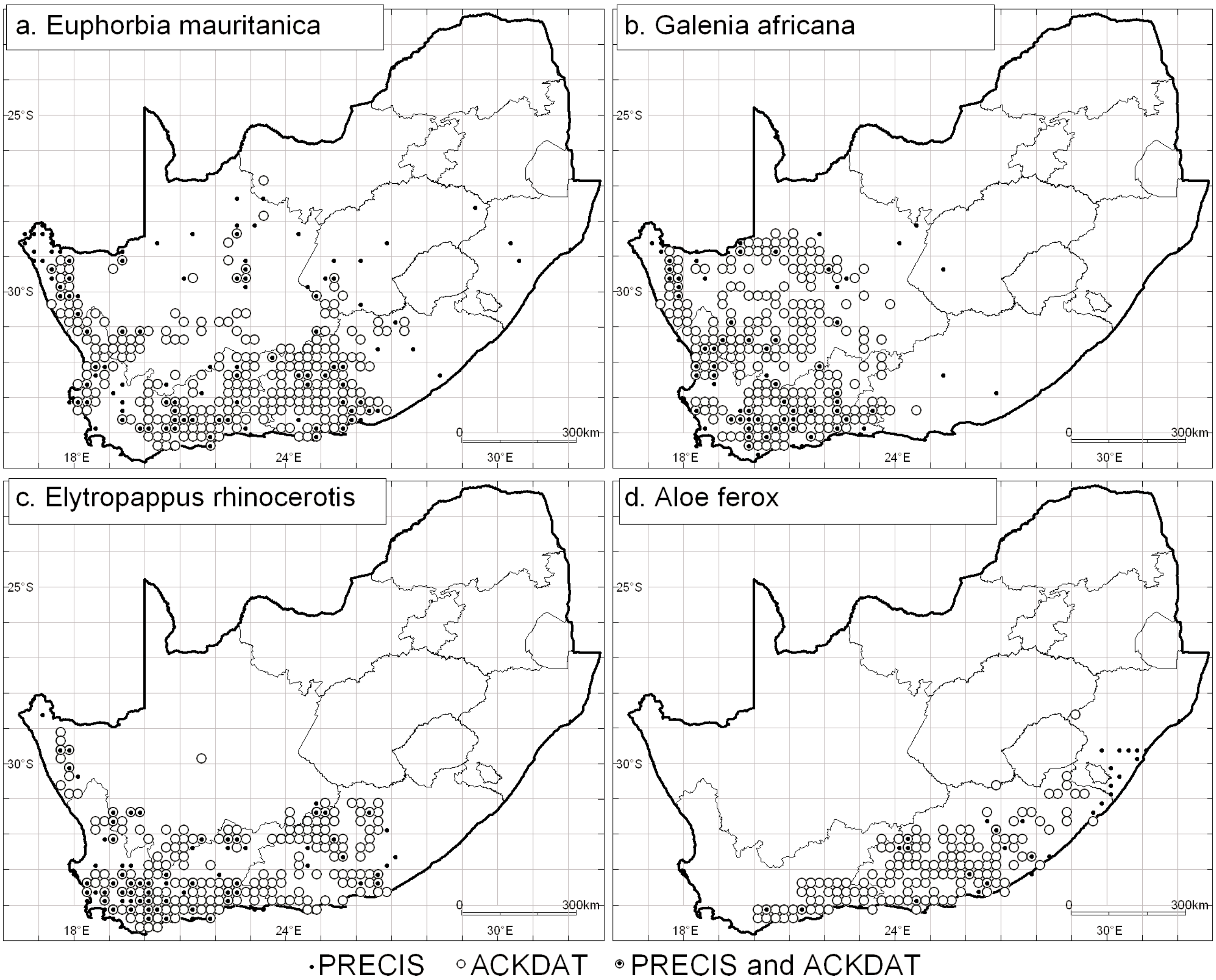

Figure 4: Recorded presence of selected species on a quarter degree grid according to PRECIS (Gibbs Russell 1985a) and ACKDAT databases

subsequent inclusion in ACKDAT. In an independent exercise of computerising Acocks records, Westfall and Greeff (1998) seem not to have experienced the same level of difficulty referred to above and apparently accept all records.

Representivity

The computerised records included those from South Africa and a very few from neighbouring areas in Namibia and Swaziland. Although Acocks sampling sites are spread around the country in 69 of the 70 veld types, his sampling intensity of regions was uneven. The number of sample sites per 10 000km^2 varied from 47.0 in the Eastern Cape Province to only 3.6 in Limpopo Province (Table 1). Acocks (1953) does acknowledge that for purposes of his veld type map and descriptions, he relied on the work of Irvine (1941) for much of the latter area. Acocks had also been stationed in the province for more than two years and had travelled extensively with Irvine in the region and so was able develop a sound knowledge of the vegetation of these savanna areas himself. Relative to South Africa's biomes (Rutherford 1997), sampling intensity was relatively high in the Fynbos Biome and lowest in the extensive Savanna Biome (Table 2). Sampling appears high in the Forest and Desert Biomes but this is due to the very small areas of these units at biome scale.

Acocks' approximately 3 000 sample sites could not be expected to capture the often high proportion of rare species in the total flora of South Africa obtained from our available version of PRECIS. The estimate of the total can vary as different ways of calculating the total flora give somewhat different results (see Gibbs Russel 1985b). Nevertheless, ACKDAT contains around 42% of the total number of vascular plants (taken as only angiosperms and gymnosperms) in PRECIS for South Africa (Table 3). The Northern Cape is best represented. As with the sampling intensity, Limpopo Province is again lowest in representing the flora. Despite the relatively high sampling intensity in the Fynbos Biome,

Table 1: ACKDAT regional sampling intensity per province

Province	Number of sample sites	Sampling Intensity (sites/ 10 000km^2)
Eastern Cape	798	47.0
Western Cape	521	40.2
Free State	394	30.3
Kwazulu-Natal	247	26.8
Gauteng	40	22.1
Northern Cape	720	19.9
Mpumalanga	145	18.6
North-West	139	12.0
Limpopo	44	3.6
South Africa	3 048	25.0

Table 2: ACKDAT regional sampling intensity per biome (after Rutherford 1997)

Biome	Number of sample sites	Sampling Intensity (sites/ 10 000km^2)
Forest	3	44.0
Fynbos	301	35.3
Desert	2	31.8
Nama-Karoo	1 084	26.1
Grassland	925	23.3
Succulent Karoo	204	21.0
Savanna	521	11.6

the proportion of the total species flora in ACKDAT for this biome is not highly ranked relative to some other biomes (Table 4). This is primarily due to few sample sites in the many mountain fynbos areas of the 'spectacularly rich' flora of the Cape region (Cowling and Hilton-Taylor 1997), an undersampling acknowledged by Acocks (1953). The high proportion of total species incorporated in ACKDAT for the Nama Karoo Biome would be expected to decrease with further collections in this area that is undersampled by PRECIS (Gibbs Russel *et al*. 1984). One important area of plant endemism that Acocks did not sample was the Richtersveld section of the Gariep Centre of endemism (Van Wyk and Smith 2001).

For certain plant species, ACKDAT's geographical coverage exceeds that of PRECIS. Example comparisons are given in Figure 4 and expressed at the common quarter degree scale used by PRECIS. In all these example species, ACKDAT includes more than 80% of the total known localities (taken as union of PRECIS and ACKDAT occurrences) compared to mostly less than 30% for PRECIS. The striking differences in coverage may be due to PRECIS bias away from very common species (e.g. *Elytropappus rhinocerotis*), possibly less charismatic species (e.g. *Galenia africana*) and species where specimens are difficult to process (e.g. *Aloe ferox* and *Euphorbia mauritanica*).

Spatial precision

Acocks used sample units of varying area and shape. The vast majority of sites were along roads of the time. At his exhaustively sampled sites, he walked around until he no longer observed additional species. Time spent sampling sites varied from less than 20 minutes to 'three or four hours as a minimum' (Acocks 1953) depending on floristic diversity present. At many sites he appears to have walked several kilometres, often along a catena from bottom to the top of a slope, noting all the species he encountered along the way. There is no certainty of the precision of his placing the dots showing his sample sites on his 1:500 000 map but without the aid of more detailed maps or GPS, it would be challenging to locate sites to a precision of more than 1km at this scale. His descriptions of site localities are typically distances from towns along roads of his time to the site. Many modern roads have changed route. The resolution of these distances was to the nearest 1.6km (whole 1 mile) for about two thirds of his sites and 0.8km (2 mile) for almost all the remaining sites along roads. Coupled with the problem of not knowing whether the distance was measured from the centre or edge of the town (at the time), the spatial precision of his sites is at best to the nearest kilometre but probably more realistically to the nearest 2km or 3km. The precision of positioning sites can be improved by consulting addition-

Table 3: ACKDAT species per province as a proportion of vascular plant species found in PRECIS

Province	Total species (PRECIS)	Species in ACKDAT	AckDat as proportion of total (PRECIS) (%)
Northern Cape	4 916	2 751	56%
Eastern Cape	6 383	3 539	55%
Free State	3 001	1 440	48%
North-West	2 483	1 117	45%
Western Cape	9 489	4 042	43%
Kwazulu-Natal	5 515	2 007	36%
Gauteng	2 826	856	30%
Mpumalanga	4 593	1 170	25%
Limpopo	4 239	971	23%
RSA	17 761	7415	42%

Table 4: ACKDAT species per biome (after Rutherford 1997) as a proportion of vascular plant species found in PRECIS

Biome	Total species (PRECIS)	Species in ACKDAT	AckDat as proportion of total (PRECIS) (%)
Nama-Karoo	4 371	2 514	58%
Forest	1 122	632	56%
Grassland	7 058	3194	45%
Fynbos	8 965	3 681	41%
Succulent Karoo	5 232	2 107	40%
Savanna	9 242	3 653	40%
Desert	197		0%

Reduced to quarter degrees, hence some discrepancy with Table 2

Please supply better graphic

Please supply better graphic

a

b

Figure 5: Example of an Acocks sample site number 1278 photographed by Acocks on 9 September 1948 (a) and revisited by NBI staff on 20 November 1996 (b) at Kepskop near Carnarvon in the Nama Karoo. The comparison shows, apart from an apparent increased grassiness, very little change over 48 years. The visit also highlights the advisability of revisiting sites at the same date (and time of day) to better realise the potential value of using matched photographs

al descriptive information in Acocks' field notes and by using the mileages that Acocks often listed at every site along the road as it was routed at the time.

Site localities in the Version 1.0 of ACKDAT were rounded to the closest $^{1}/_{8}$ degree (7.5 minutes by 7.5 minutes). This represents approximately 13.8km in a north-south direction by 11.6km in an east-west direction. This approximation was to permit subsequent checking of localities against other sources of data held by the NBI, including photographs of sites. This process is ongoing.

Applications

The past decade has seen a strong increase in biodiversity related research worldwide. Spatial information on species underpins a new, digitally enhanced approach to conservation research and its application. The chief limitation to advances in this area lies in the lack of spatial data for analyses of species distributions and the associated spatial arrangement of biodiversity (e.g. Ferrier 2002). ACKDAT represents an unparalleled database at a large spatial scale that can be applied in answering a range of biodiversity and conservation questions. Examples where ACKDAT has been successfully employed include categorising functional plant response types (Rutherford *et al*. 1995), predicting survival in new environments (Rutherford *et al*. 1996), assessing the possible consequences of projected climate change on the floras of conservation areas (Rutherford *et al*. 1999), serving the needs for the plant biodiversity component of the South African Country Study on Climate Change that dealt with vulnerability and adaptation (Rutherford *et al.* 2000, Midgley *et al.* 2001) and testing relationships between a regional flora and fauna (Dean and Milton 2003).

It is tempting to think of using Acocks' accurately-dated recordings for detecting vegetation change over time. On the basis of independently captured data of the Acocks' records, Westfall and Greeff (1998) claim that Acocks' sites are 'extremely appropriate' for vegetation monitoring. Given the caveats set out above, especially in relation to precision of site location, it is difficult to see how much the perceived floristic changes can be attributed to actual change and how much to the possibility of partly or wholly 'resampling' a different area. The use of Acocks' data for monitoring has also, appropriately, been questioned by O'Callaghan (2000). Even where Acocks provided a photographic record at the site and the scene was reinspected (Figure 5), it is unclear exactly where the actual sampling area lies. For sensitive and reliable detection of change, monitoring techniques commonly use various forms of fixed points. Although for many other purposes ACKDAT would be regarded as a spatially high-resolution database, ACKDAT would be suitable for providing monitoring baselines only where the vegetation of an area was very uniform or where changes were major and widespread. O'Connor *et al.* (2003) also point to the problem of confidently relocating Acocks' sites for measuring change. However, they do demonstrate how even approximate relocation of sites can successfully be used to compare the effects of commercial versus communal grazing on compositional change over 50 years in KwaZulu-Natal.

Many new approaches to understanding the environmental determinants of species distributions have emerged in the past decade. Such knowledge is critical for predicting how species may respond to a major human-induced threat – climate change. Acocks himself was almost obsessed with understanding the root causes of desertification (Meadows 2003), and discussed climate change in his early correspondence. Linking ACKDAT through GIS applications to climate and other environmental surfaces allows advanced analytical techniques to be applied to plant species distributions. In this regard, Acocks' saturation sampling techniques are particularly powerful, in that a plant absence can be inferred with some degree of confidence where a presence is not recorded. This is the basis of a new generation of logistic regression-based models of plant species distribution, such as Generalised Additive Modelling (Hastie and Tibshirani 1990).

Clearly, ACKDAT could never have existed without Acocks' extraordinary dedication and unparalleled knowledge of one of the richest floras of the world. To have a working knowledge of over 7 000 species, often with plants in far from perfect condition in the field (e.g. heavily grazed), is simply unique.

Acknowledgements — the authors wish to acknowledge the input of

M. O'Callaghan, MT Hoffman and many others in earlier developments of ACKDAT and T. Arnold for the use of the PRECIS database.

References

Acocks JPH (1953) Veld types of South Africa. Memoirs of the Botanical Survey of South Africa **28**: 1–192

Acocks JPH (1975) Veld types of South Africa (2nd edn). Memoir of the Botanical Survey of South Africa **40**: 1–128

Acocks JPH (1988) Veld types of South Africa (3rd edn). Memoirs of the Botanical Survey of South Africa **57**: 1–146

Arnold TH, de Wet BC (1993) Plants of southern Africa: names and distributions. Memoirs of the Botanical Survey of South Africa **62**: 1–825

Cowling RM, Hilton Taylor C (1997) Phytogeography, flora and endemism. In: Cowling RM, Richardson DM, Pierce SM (eds) Vegetation of Southern Africa. Cambridge University Press, Cambridge, pp 43–61. ISBN 0–521–57142–1

Dean WRJ, Milton SJ (2003) Did the flora match the fauna? Acocks and historical changes in Karoo biota. South African Journal of Botany (in press)

Ferrier S (2002) Mapping spatial pattern in biodiversity from regional conservation planning: where to from here? Systematic Biology **51**: 331–363

Gibbs Russel GE (1985a) PRECIS, the National herbarium's computerised information system. South African Journal of Science **81**: 62–65

Gibbs Russel GE (1985b) Analysis of the size and composition of the southern African flora. Bothalia **15**: 613–629

Gibbs Russel GE, Retief E, Smook L (1984) Intensity of plant collecting in southern Africa. Bothalia **15**: 131–138

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. Chapman and Hall, London

Hoffman MT, Cowling RM (2003) 'Enough to be considered useful': John Acocks' contribution to South African botany. South African Journal of Botany (in press)

Irvine LOF (1941) The Major Veld-types of the Northern Transvaal and their Grazing Management in the Light of Research Conducted within the Area. DSc thesis, Faculty of Agriculture, University of Pretoria, Pretoria, South Africa

Killick DJB (1980) Obituaries: John Phillip Harison Acocks (1911–1979). Bothalia **13**: 239–244

Killick DJB (1982) Corrections to Acocks's Veld Types of South Africa (1975). Bothalia **14**: 143

Meadows ME (2003) John Acocks and the expanding Karoo hypothesis. South African Journal of Botany (in press)

Midgley GF, Rutherford MC, Bond WJ (2001) The heat is on: impacts of climate change on plant diversity in South Africa. National Botanical Institute, Claremont. ISBN 1–919684–32–8

O'Callaghan M (2000) National grid of vegetation sites: reality check. South African Journal of Science **96**: 231–233

O'Callaghan M, Powrie LW, Hurford JL, Rutherford MC (1994) ACKDAT: The national plant ecological database. Unpublished paper presented at the Fynbos Forum, 13–15 July 1994, Bien Donné, Stellenbosch, South Africa

O'Connor TG, Morris CD, Marriott DJ (2003) Change in land use and botanical composition of KwaZulu-Natal's grasslands over the past fifty years: Acocks's sites revisited. South African Journal of Botany (in press)

Rutherford MC (1997) Categorization of biomes In: Cowling RM, Richardson DM, Pierce SM (eds) Vegetation of Southern Africa. Cambridge University Press, Cambridge, pp 91–98. ISBN 0–521–57142–1

Rutherford MC, Midgley GF, Bond WJ, Powrie LW, Roberts R, Allsopp J (2000) Plant biodiversity. In: Kiker G (ed). Climate change impacts in southern Africa. Report to the National Climate Change Committee, Department of Environmental Affairs and Tourism, Pretoria, South Africa

Rutherford MC, O'Callaghan M, Hurford JL, Powrie LW, Schulze RE, Davis GW, Hoffman MT, Mack F (1995) Realized niche spaces and functional types: a framework for prediction of compositional change. Journal of Biogeography **22**: 523–531

Rutherford MC, O'Callaghan M, Powrie LW, Hurford JL, Schulze RE (1996) Predicting survival in new environments through analytical GIS application. Environmental Software **11**: 113–121

Rutherford MC, Powrie LW, Schulze RE (1999) Climate change in conservation areas of South Africa and its potential impact on floristic composition: a first assessment. Diversity and distributions **5**: 253–262

Van Wyk AE, Smith GF (2001) Regions of Floristic Endemism in Southern Africa: A Review with Emphasis on Succulents. Umdaus Press, Hatfield. ISBN 1–919766–18–9

Westfall RH, Greeff A (1998) A national grid of vegetation monitoring sites. South African Journal of Science **94**: 150–151

Edited by RM Cowling and MT Hoffman