SANBI has been actively involved with locating localities of herbarium specimens since before 1976. It has been valuable to have quality species-locality information for bioclimatic modelling, for descriptions of vegetation types in South Africa, Lesotho and Swaziland, and assigning endemic plant species to a vegetation type. Of prime importance is for collectors to record accurate coordinates. With hundreds of thousands of herbarium and museum specimens in the National Herbarium (PRE) and Compton Herbarium (NBG) and other collections that had precision of centroid of a map sheet, it became necessary to develop effective methods for obtaining accurate coordinates for specimens. The South African Place Names was compiled in 1976 with 42,399 names in quarter degree grids. A far more comprehensive and precise gazetteer with over 530,000 named places in southern Africa has been compiled from many sources added to South African Place Names, and names digitised from historical maps. This has enabled the effective and efficient location of named places for georeferencing specimens. An ArcView 3.x extension that accessed this large gazetteer, together with historical maps, made it possible to georeference as many as ten times more specimens per day than using the gazetteer and other GIS applications. An increasing number of historical maps has been prepared for GIS in order to facilitate this work. With ever-increasing numbers of accurate records available, the number of applications requiring accurate georeferencing has grown to include National Biodiversity Assessment, Red-listing of plant and animal species, herbarium curation, and other work requiring accurate spatial data.

As submitted
L.W. Powrie, E-mail: L.Powrie@sanbi.org.za

Oral

Category: Biodiversity Information

Where exactly? Historical and modern geo-referencing challenges and solutions

L.W. Powrie

South African National Biodiversity Institute, Kirstenbosch Research Centre, P/Bag X7, Claremont, 7735, Cape Town, South Africa.

South African National Biodiversity Institute (SANBI) has been actively involved with locating localities of herbarium specimens since before 1976 and is currently involved with extracting species distribution information from field notes, surveys, specimens of all kinds of organisms, and other sources. Quality species-locality information has made possible bioclimatic modelling, species lists for descriptions and assigning endemic plant species to vegetation types in South Africa, Lesotho and Swaziland. With millions of herbarium and museum specimens in SANBI, Iziko Museums and many other collections, it became desirable to develop effective methods for obtaining accurate coordinates for geo-referencing records. The South African Place Names was compiled in 1976 with 42,400 names in quarter degree grids. A far more comprehensive and precise gazetteer with 520,000 named places in southern Africa compiled from many other sources, an ArcView 3.x extension that accesses this large gazetteer, using historical maps, and other tools have enabled increasingly effective and efficient location of named places for georeferencing.
specimens. With ever-increasing numbers of accurate records available, the number of applications benefiting from accurate georeferencing has grown to include National Biodiversity Assessment, Red-listing of plant and animal species, biodiversity science, policy and advice, among others. This talk addresses some challenges in pinpointing localities from historical records and guidance for recording locality information for future records and collections. By overcoming these challenges, the quality of information can be better used for many potential applications. It is presumably the dream of our colleagues, young and old, that information that they record will be valuable. Therefore, the data collector will surely wish to maximise the value of his or her contribution by recording accurate locational information.

As presented

Georeferencing – where was that dude?

Leslie W. Powrie

SANBI

7 Aug 1985: Tulbagh Waterfall Forest Reserve

Estimated locality
I was recently preparing some specimens to go into the herbarium. I estimated a point based on what the label information said. I guessed it would be west of Tulbagh, NE or E aspect. But no evidence of falls. I estimated 5 km confidence.


Then I decided to reconsider this. I searched for Tulbagh Waterfall Forest Reserve in the Gazetteer – not there. I searched for Tulbagh to see if there might be some likely match – no luck. Closest is ‘Tulbagh waterfall (Schlechter)’ I searched for Waterfall – 127 possibilities, filter on 33°S and 19°E, 2 matches not close to Tulbagh. I searched for Waterval – 560 possibilities, filter on 33°S and 19°E, 39 matches, one close to Tulbagh, ‘Waterval Forest Reserve’.

This looked much better in terms of NE and E aspects and there being a waterfall, plantations shown on 1:50 000 map, hill.

About 7 km away from the first locality
The label mentions pine plantations. The 1:50 000 maps showed forest plantations. I looked at Google Earth, no obvious forest. I used the historical imagery and there was a suggestion of pine plantation from imagery of about 1990. Forest, overlooking private lands, this looked like a good match. What looks like forest is visible in imagery from 1984 to 1990.

It could be that a fire cleared the plantation, or it was harvested, and then not replanted. Either way, it was not evident in Google Earth.
The embarrassing thing is that I was that dude!

John Rourke and I were there on 7 Aug 1985, collecting Diastella for my MSc study. I believe Di Prosch was with us, possibly some other colleagues in the early morning climb to about 300 m where John said a particular species would start to appear.

I thought I was doing a reasonable job recording the locality. But I battled to find the locality 31 years later.

1985 and 2017

- No GPS, GPS on phone
- 5 minutes vs. several hours
- High precision vs. low confidence
- Introduce a georeferencer

GPS did not exist in 1985. Now almost anyone can have a GPS in a cell phone! It might have taken 5 minutes to record coordinates in 1985. From about 1987 I was carefully recording coordinates when I collected anything. It took me hours to determine where I had been 30 years earlier, and I am reasonably confident that I am within about 500 m or so. With GPS, 5 m precision is normal, convenient, simple.
Introduce a georeferencer, not with me in 1985, and not familiar with the area... Such a person lacks the advantage of even a vague recollection of the site, sometimes not even familiar with the geography of the area. Has nothing but a gazetteer and maps to help guess where the specimen was collected.

Why is precise locality important anyway?

Another example. On 20 Sep 2007 we were studying this fenceline contrast when we collected LWP 1446 Ornithoglossum sp. = Colchicaceae ? Ornithoglossum 20 Sep 2007 – the plant was about 6 cm high geophyte leaf, not identifiable. So, I took a GPS reading so that we could return on our last day at the site – this was day 3 of 11 that we were doing this survey. -32.99884, 19.89441

LWP 1446 - Sword Ledebouria

- 20 Sep 2007 – took GPS reading
  -32.99884, 19.89441

- 28 Sep 2007 returned and found it again within a few minutes, but still not identifiable.
  – Ornithoglossum sp. = Colchicaceae ? Ornithoglossum
28 Sep 2007 returned and found it again within a few minutes, but still not identifiable. But – we found that small plant! ID: Ornithoglossum sp. = Colchicaceae ? Ornithoglossum

Would I have been able to find it if I had collected it 7 Aug 1985 at Tulbagh Waterfall Forest Reserve? I am embarrassed to think how unlikely that is!

What has been done with georeferenced records?

Plant specimens in the National Herbarium, Pretoria, were first transcribed into a database in about the early 1980s, a leading initiative in the world.

Georeferencing, if done, was at quarter degree level, a precision with a radius of about 18 km at best. This represents the national topographic map series at 1:50 000 scale. It did enable plotting frequency of collecting by map sheet, species richness in provinces and biomes, and the like. Some grid cells have between 10 000 and 20 000 specimens collected.

First List of Species was 1984, so transcribing of specimen label information was well advanced by then.

If coordinates had been recorded, they were typed into the Notes field as there were not fields for Latitude and Longitude.

We saw the opportunity to relate species distribution to climate data to assess the potential impact of climate change on selected species.
Precise Georeferencing was done for selected taxa before 1999 for bioclimatic modelling.

The climate data were at about 1.5 km precision, so the effort was made to geo reference specimens for specific species to improve the 18 km radius to at least about 1.5 km precision.

For the descriptions of vegetation types on the national vegetation map, we checked species ranges relative to vegetation, particularly for endemic species.

The localities and habitats for endemics were carefully georeferenced and checked.

The species was only declared endemic to a vegetation type if it occurred primarily (90% or more of localities) in that vegetation type and did not occur in any other type (10% or less).
3308 species each occur in a single vegetation type.

16 species are listed in two vegetation types.

Limonium sp. nov. occurs in 4 vegetation types, classification still in preparation when the book was published.

Dicrocaulon sp. nov. occurs in 5 vegetation types: Dicrocaulon sp. nov. (‘longifolium’) Dicrocaulon sp. nov. (‘pseudonodosum’) Dicrocaulon sp. nov. (‘nanum’) Dicrocaulon sp. nov. (‘neglectum’) Dicrocaulon sp. nov. (‘prostratum’)

In Mucina & Rutherford 2006:

7.7.4 Endemic Taxa

The concept of endemism is determined by the extent of the vegetation unit. This means that a plant taxon is listed as endemic in the description when it occurs exclusively within the unit concerned. We have relaxed the strict interpretation of the ‘unit-based endemism’ in the Fynbos Biome, where an endemic plant species may have less than 10% of localities outside the vegetation unit in question. In some other biomes we accepted a relatively narrow interpretation of the notion of ‘near-endemic’ where the clear concentration of the given taxon is within the vegetation unit and a few outliers occur nearby in one or more adjacent units. We are well aware of the fact that the current endemic status of many plants would change in future as our knowledge about distribution of the species becomes more detailed and the extent of our vegetation units becomes more precisely defined. As our vegetation units are naturally defined entities, we consider the use of the term ‘endemism’ in this sense more appropriate than using this term in a political context (defined by the boundaries of countries or tourist regions).

IUCN declarations for extinct, endangered, threatened, vulnerable species. Initially plants, then butterflies, reptiles, spiders, and other groups.
Species at risk

Essential to know where species occur, how many individuals, threats, health of populations.

Specimens of plants, crustaceans, insects and other organisms. There are various methods of collecting and preserving the specimens. There is debate about the ethics of collecting specimens, especially with rare or endangered species.

Specimens of plants, crustaceans and other organisms.
There are various levels of locality detail given, from latitude and longitude coordinates to province to nothing at all.

- 637 Cape of Good Hope Count Major Minor Country Loc Notes 600 Western Cape South Africa CAPE OF GOOD HOPE 24 Cape Province South Africa CAPE OF GOOD HOPE 6 Unknown Unknown CAPE OF GOOD HOPE 4 South Africa CAPE OF GOOD HOPE 1 Eastern Cape South Africa CAPE OF GOOD HOPE 1 KwaZulu-Natal South Africa CAPE OF GOOD HOPE 1 South Africa South Africa Cape of Good Hope

The legibility also varies from typewritten to difficult handwriting.

Type specimens for species

A new species of *Indigofera* L. (Fabaceae) from the eastern Cape

P.B. Phillipson

Department of Botany, Rhodes University, P.O. Box 94, Grahamstown, 6140 Republic of South Africa

*Indigofera elandsbergensis* P.B. Phillipson sp. nov., 1. cuncifoliae affine, a qua imprimit differt caulibus glabris, foliis sub-glabriss, foliolis plus parvis obovatis vel sub-ovatis, racemis minus densis glabris, stipulis et bracteis minus acuminatis glabris et floribus plus parvis.


Shrub to about 1.5 m. Entire plant except flowers somewhat glaucous. *Stems* erect, densely branched distally, glabrous.

Type specimens are the permanent reference for a species description. There must be at least one type specimen for a name.
Plants are sessile, rooted in the ground. Can measure anthers, leaves, petals, take photographs, describe texture, collect specimens.

Nuisance is strong wind, spines, succulent leaves. But they stay put.

Most butterflies, beetles, lizards, birds, fish and other animals move quickly and far.

This spider in my office moved the equivalent of my hiking from where I lived in 1983-1985 to the top of Lion's Head and back, and it did that in far less time than my hike would have taken.

Just try macro photography of butterflies, wasps, birds, any fast-moving thing. They just are not cooperative!

So, I think my argument has been made - georeferencing is needed:
Locate the plant;

- Species richness, diversity, ranges and distributions;
- Model the climate;
- Endemism;
- Red listing;
- Environmental impact assessments;
- Etc.

But rather than an area of 23 x 27 km covering Cape Town and Robben Eiland, we really need to get to about 2 X 2 m, the scale a single plant.
Studying animals that wander in a range of meters or kilometres require less precise georeferencing than plants, but it is amusing that precise coordinates were more often recorded for animals than for plants.

The definition of precision and accuracy.

If we had delegates wanting to get to this SAAB conference but insisted on using 0 decimal places (degree precision) they would be sitting in a boat out at sea 46 km away.

1 decimal place – in the military base 1.8 km away, dodging aircraft and risking arrest for trespassing.
2 decimal places – paddling in the waves 376 m away.

3 decimal places – teetering on the edge of the building 25 m away.

4 decimal places – 5.7 m from 5 decimal place within 10 m. We are approaching the accuracy of a GPS.

5 decimal places – within about 2 m. I always recommend using 5 decimal places of a , 3 decimal places of a minute or 1 decimal place of a second.

More than 6 decimal places is giving false precision.

Imagine looking for a little bulbous plant using these coordinates. Well, maybe it was here before the hotel was built?
A GPS is precise to about 5 decimal places (7 m).

The GPS concept was born in 1973. As it developed for DoD purposes, concern started to arise regarding the enemy being able to aim at targets and so Selective Availability was introduced in 1990 to downgrade the accuracy for the general users. SA was turned off 2 May 2000. So, accuracy went from 50 m horizontally and 100 m vertically. Currently, civilian GPS fixes under a clear view of the sky are on average accurate to about 5 meters (16 ft) horizontally. [https://en.wikipedia.org/wiki/Error_analysis_for_the_Global_Positioning_System](https://en.wikipedia.org/wiki/Error_analysis_for_the_Global_Positioning_System)

In the process of this bioclimatic modelling and vegetation mapping I developed the SANBI Gazetteer – now being incorporated into Botanical Database of Southern Africa (BODATSA) To be available online A living, expanding resource.

In 2008 Dee Snijman asked me if I knew where Kommando Drift (near the Tanqua River and Ceres-Calvinia road) was where Acocks collected a specimen that she was studying. It was not on the maps that she was using. PRECIS showed only QDS.

I found it on the map that Acocks had used, and it is on two other older maps. But it is not shown on newer maps. It is now in the Gazetteer.

More than 445 000 unique place names with

- Locality
- Minor (district)
- Major (province)
- Country
- Type of locality (e.g. waterfall, dam, farm, kraal)
- Radius in metres indicates precision for coordinates (point, line or polygon)
- Altitude

There are some challenges with a gazetteer, but it is really valuable.
Palmietfontein occurs four times on the same map. One can be precise or vague depending on the details recorded by the collector.

Remhoogte Pass is not shown on the modern map, and the route has changed. So the map used by the collector must be the reference for routes rather than the newer map.

Historical maps

- Various ages
- Various scales
  - Different details
- Various ages
- Various scales
  - Different details

- Historical maps
- Various ages
- Various scales
- Different details

Cape Peninsula

Cape Peninsula, 1909
Cape of Good Hope Colony, 1895

1:50 000
Using specimens from and expedition yields much better results in a fraction of the time than by species cupboard. Georeferencing specimens one at a time, took about 40-50 hours spread over 35 years, resulting in 68 specimens being mapped to a precision of about 18 km, 9 with errors ranging from 320 km to 650 km.

Mapping using the itinerary took less than 3 hours, and 151 specimens were mapped, with 128 at 1-2 km, 21 at 2.5-5 km precision. The remaining specimens have no locality information. If collecting register or field notes are available, they might add information.

Of the 16 specimens collected on one day, 7 were not mapped, 7 were mapped to the Pella [Mission Station] map sheet, and 2 were mapped to another locality about 650 km away.
Of the 182 collections from the expedition that included the day near Pella [Mission Station]: 105 were not georeferenced. Of the 77 georeferenced:

- 9 are outliers
- 2 at another Pella (510 km away)
- 5 at another Concordia (320 km away)
- 2 at another Wolftoon (Namibia, 650 km away)

‘From’ presumably means ‘after leaving’ – but in what direction? It is much better to give a direction, e.g. ‘SW’ as in this case.

The 16 specimens collected on the day at Pella were sorted and distributed throughout the herbarium. Some of the individual species were encoded (specimen label information typed), and some of these were georeferenced.
The best is for the collector to georeference at the time. This is the most accurate and least time consuming -seconds vs. tens of minutes of guessing.

When georeferencing a batch of specimens it is worth spending time to do detective work, and the confidence in the results is much greater.

I propose that one focus on transcribing label information and then, when there are a good number of specimens, to focus on batch georeferencing unless urgent georeferencing is needed. Transcribe field notes and collecting registers if possible as these will give context for georeferencing specimens.