Foliar Trait and Spectroscopy Data, Greater Cape Floristic Region, South Africa, V2

Documentation Revision Date: 2026-09-17

Dataset Version: 2

Summary

This dataset holds foliar trait measurements, species-level life history traits, visible-near infrared (VNIR) spectroscopy, and vegetation survey data for plant communities within the Greater Cape Floristic Region (GCFR) of South Africa. Foliar traits measurements were made between 2010 and 2014. Releve surveys include plots from 1966 to 2014. Trait collection efforts extended over 5.8% of the GCFR's total area, where measurements were made for 1,328 species from 100 families. The trait dataset comprises 132,406 leaf and twig structural measurements, 7,467 canopy measurements and 11,949 elemental/isotopic measurements. The spectroscopy dataset contains 3,076 VNIR spectral measurements (450-949 nm) of 1,158 species from 97 families. These data contribute to research initiatives such as the Biodiversity Survey of the Cape (BioSCape) project and help understand this ecosystem from a trait-based perspective. The data are provided in comma separated values (CSV) format.

This dataset includes five files in comma separated values (CSV) format.

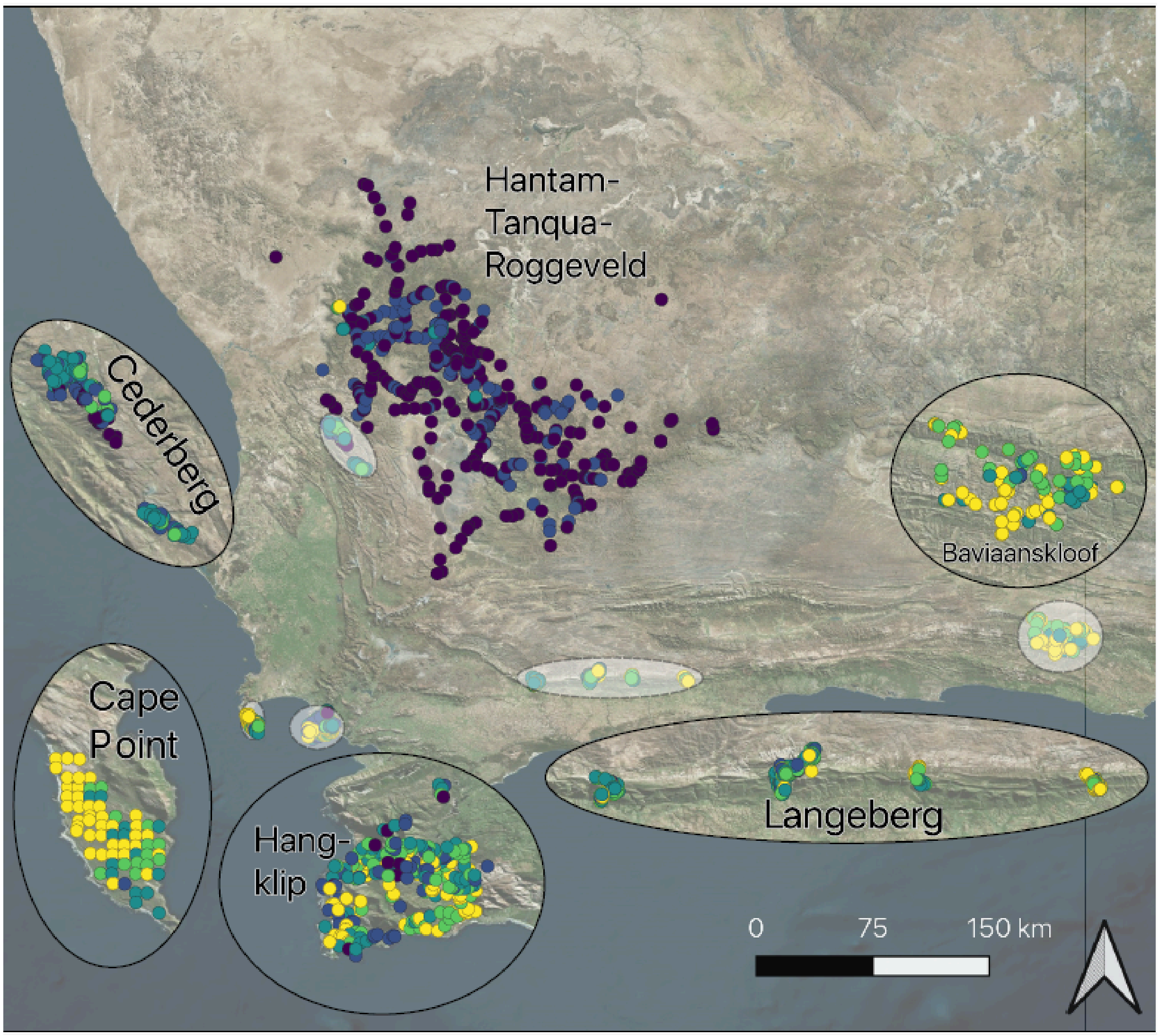

Figure 1. Map of releve plots from the most recent surveys of each subregion. Each releve plot is represented by a single point and is colored by its species richness. The Fynbos dominated subregions are shown with insets.

Citation

Frye, H., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, R. Turner, A.M. Wilson, and J.A. Silander Jr. 2026. Foliar Trait and Spectroscopy Data, Greater Cape Floristic Region, South Africa, V2. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2528

Table of Contents

1. Dataset Overview

This dataset holds foliar trait measurements, species-level life history traits, visible-near infrared (VNIR) spectroscopy, and vegetation survey data for plant communities within the Greater Cape Floristic Region (GCFR) of South Africa. Foliar traits measurements were made between 2010 and 2014. Relevé surveys include plots from 1966 to 2014. Trait collection efforts extended over 5.8% of the GCFR's total area, where measurements were made for 1,328 species from 100 families. The trait dataset comprises 132,406 leaf and twig structural measurements, 7,467 canopy measurements and 11,949 elemental/isotopic measurements. The spectroscopy dataset contains 3,076 VNIR spectral measurements (450-949 nm) of 1,158 species from 97 families. These data contribute to research initiatives such as the Biodiversity Survey of the Cape (BioSCape) project and help understand this ecosystem from a trait-based perspective.

This is Version 2 of this dataset and supercedes Version 1. The data has been updated to remove measurement outliers and data recording errors.

Project: Biodiversity Survey of the Cape (BioSCape)

The Biodiversity Survey of the Cape (BioSCape) is a research project that uses remote sensing and field data to understand the distribution, function, and importance of biodiversity on land and in the water around the Greater Cape Floristic Region (GCFR). Located in southwestern South Africa along the Atlantic coast, the GFCR is one of the world's 36 biodiversity hotspots. BioSCape is a collaboration between the US and South Africa. The team conducted an integrated field and airborne campaign in the hyperdiverse region in 2023. Researchers collected laser altimetry LiDAR data and UV/visible to short wavelength infrared (UVSWIR) and thermal imaging spectroscopy data over terrestrial and aquatic targets using four airborne instruments: AVIRIS-NG, PRISM, LVIS, and HyTES. This airborne dataset is unique in its size and scope and unprecedented in its instrument combination and level of detail. These airborne data are accompanied by a range of biodiversity-related field observations.

Related Publications:

Aiello-Lammens, M.E., J.A. Slingsby, C. Merow, H.K. Mollmann, D. Euston-Brown, C.S. Jones, and J.A. Silander. 2016. Processes of community assembly in an environmentally heterogeneous, high biodiversity region. Ecography 40:561–576. https://doi.org/10.1111/ecog.01945

Frye, H.A., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, A.M. Wilson, and J.A. Silander. 2021. Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. Global Ecology and Biogeography 30:1403–1417. https://doi.org/10.1111/geb.13306

Frye, H.A., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, R. Turner, A.M. Wilson, and J.A. Silander Jr. 2026. Foliar trait and spectroscopy data from the megadiverse Greater Cape Floristic Region. In review.

Schliep, E.M., A.E. Gelfand, R.M. Mitchell, M.E. Aiello-Lammens, and J.A. Silander. 2017. Assessing the joint behaviour of species traits as filtered by environment. Methods in Ecology and Evolution 9:716–727. https://doi.org/10.1111/2041-210X.12901

Slingsby, J.A., C. Merow, M. Aiello-Lammens, N. Allsopp, S. Hall, H. Kilroy Mollmann, R. Turner, A.M. Wilson, and J.A. Silander. 2017. Intensifying postfire weather and biological invasion drive species loss in a Mediterranean-type biodiversity hotspot. Proceedings of the National Academy of Sciences 114:4697–4702. https://doi.org/10.1073/pnas.1619014114

van Leeuwen, M., H.A. Frye, and A.M. Wilson. 2021. Understanding limits of species identification using simulated imaging spectroscopy. Remote Sensing of Environment 259:112405. https://doi.org/10.1016/j.rse.2021.112405

White, P.A., H. Frye, M.F. Christensen, A.E. Gelfand, and J.A. Silander. 2022. Spatial functional data modeling of plant reflectances. The Annals of Applied Statistics 16. https://doi.org/10.1214/21-AOAS1576

White, P.A., M.F. Christensen, H. Frye, A.E. Gelfand, and J.A. Silander. 2023. Joint multivariate and functional modeling for plant traits and reflectances. Environmental and Ecological Statistics 30:501–528. https://doi.org/10.1007/s10651-023-00574-7

White, P.A., H.A. Frye, J.A. Slingsby, J.A. Silander, and A.E. Gelfand. 2023. Generative spatial generalized dissimilarity mixed modelling (spGDMM): An enhanced approach to modelling beta diversity. Methods in Ecology and Evolution 15:214–226. https://doi.org/10.1111/2041-210X.14259

White, P.A., A.E. Gelfand, H. Frye, and J.A. Silander. 2024. Good modelling practice in ecology, the hierarchical Bayesian perspective. Ecological Modelling 496:110847. https://doi.org/10.1016/j.ecolmodel.2024.110847

Related Dataset:

Frye, H., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, R. Turner, A.M. Wilson,and J.A. Silander Jr. 2026. Foliar Trait and Spectroscopy Data Greater Cape Floristic Region, South Africa (Version 1). ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2482

- Version 1 of this dataset now superceded by this Version 2.

Acknowledgements:

The BioSCape Project is a multi-agency, NASA-led research project that integrates airborne imaging spectroscopy (AVIRIS NG) and lidar (LVIS) with a suite of surface measures of biodiversity. Financial support was provided by the National Science Foundation (NSF) (grant DEB-1046328), with a

Research Experiences for Undergraduates Broadening Participation (REU-BP) Supplement to J. A. Silander and C. S. Jones, and the South African Environmental Observation Network (SAEON). Compilation of data was supported by NASA grant 80NSSC22K1383 and National Research Foundation of South Africa (grant 142438). H.A. Frye was supported by funding from NSF (grant DBI-2021898) and NASA (grant 80NSSC22K0905). Additional funding was provided by the Critical Ecosystem Partnership Fund (CEPF) via the SKEP (Succulent Karoo Ecosystem Plan/Program) initiative.

Kirstenbosch Gardens, South African National Parks, Cape Nature, and Northern Cape Department of Environment and Nature Conservation gave permission to collect leaf spectra and traits and herbarium specimens. Logistical support was provided by Eastern Cape Parks and Tourism Agency and the Gamtoos Irrigation Board.

C. Carlson, E. Gray, S. Hall, J. Leonard, S. Mollmann, T. Moore, N. Oettle, N. Silander, N. Stevens, E. Timpe, R. de Gouvenain, E. O'Rourke, and B. Ramos assisted in field collections. K. Martinod assisted in the compilation of species-level traits.

2. Data Characteristics

Spatial Coverage: Greater Cape Floristic Region (GCFR), Northern Cape and Western Cape provinces, South Africa

Spatial Resolution: plots with average spacing of 1.1 km

Temporal Coverage: 2010-07-21 to 2014-08-26 for foliar traits; relevé data include plots from 1966 to 2014

Temporal Resolution: One time measurements

Site Boundaries: Latitude and longitude are given in decimal degrees.

Site	Westernmost Longitude	Easternmost Longitude	Northernmost Latitude	Southernmost Latitude
Greater Cape Floristic Region, South Africa	18.3744	24.4833	-30.3925	-34.3860

Data File Information

This dataset includes five files in comma separated values (CSV) format (Table 1).

Table 1. Files in this dataset.

File name	Content	Rows	Columns
BioSCape_foliar_canopy_leaf_chemistry_V2.csv	11,949 foliar chemistry and canopy measurements of 1,328 species	2509	24
BioSCape_foliar_leaf_struct_water_traits_V2.csv	132,406 foliar and twig structural measurements for 1,328 species	9547	30
BioSCape_foliar_species_traits_V2.csv	56 variables related to life history, growth form, and phenology for 2,213 species	2225	54
BioSCape_foliar_vnir_spectra_V2.csv	3,076 spectral measurements for 1,158 species	3076	511
BioSCape_foliar_releve_V2.csv	1,828 relevé surveys across a time span from 1966–2014	44,160	17

Table 2. Variables in *BioSCape_foliar_canopy_leaf_chemistry_V2.csv*.

Variable	Units	Description
sample_ID	-	A unique value that identifies the collection of a particular species, at a particular location, on a particular day. The value is based on Date_Sample_Subregion_Collector (e.g., "619_1_hangklip_CM", "826_9_htr_BHR")
scientific_name_original	-	Latin binomial assigned to plant based on most recent taxonomy at the time of data collection (e.g., Abutilon sonneratianum, Zygophyllum retrofractum)
family_MG	-	Family of plant based on the Manning and Goldblatt (2012) taxonomy (e.g., Acanthaceae, Zygophyllaceae)
scientific_name_WFO	-	Latin binomial of plant based on the World Flora Online v.20263.612 static backbone (e.g., Abutilon sonneratianum, Zygophyllum retrofractum)
scientific_name_authorship_WFO	-	Taxonomic authorship assigned to the World Flora Online taxa(e.g., "(A.DC.) F.White", "Willd.")
family_WFO	-	Plant family based on the World Flora Online v.20263.612 static backbone(e.g., Acanthaceae, Zygophyllaceae)

latitude	degrees north	Latitude in decimal degrees; range: -34.36736 to -31.24257
longitude	degrees east	Longitude in decimal degrees; range: 18.38403 to 24.34838
subregion	-	Subregion category
date	YYYY-MM-DD	Date of trait collection
collector	-	Collector initials, retained for provenance of unique identifier
sample	-	Sample number of individual collected
height_cm	cm	Vegetation height; range: 0 - 2500
canopy_axis_1_cm	cm	Length of canopy axis 1; range: 1 - 1400
canopy_axis_2_cm	cm	Length of canopy axis 2; range: 0.5 - 1300
canopy_cover_cm2	cm^2	Canopy cover calculation; range: 0.392 - 1429424.657
branch_order	-	Ordinal measure (0 to 4) of branch order, a measure of complexity of branching, with higher values representing greater complexity
pubescence	-	Categorical value indicating presence of pubescence (hair) on leaves (l), stems (s), both leaves and stems (b), or neither (n). NA indicates lack of measurement, not necessarily lack of pubescence
percent_N	percent	Percent of nitrogen content of leaf or primary photosynthetic surface; range: 0.200 - 4.837
percent_C	percent	Percent of carbon content of leaf or primary photosynthetic surface; range: 14.388 - 59.959
C_to_N_ratio	1	Carbon to nitrogen ratio; range: 6.654625191 - 234.89
d_15N_14N	‰	Ratio of δ15N to δ14N isotopes (in per thousand); range: -10.2654 to 17.38414
d_13C_12C	‰	Ratio of δ13C to δ12C isotopes (in per thousand); range: -33.2334 to -11.5092
quality_flag	-	Quality flags based on outlier assessment (e.g., "canopy_magnitude", "n_lma_covariation")

Table 3. Variables in *BioSCape_foliar_leaf_struct_water_traits_V2.csv*.

Variable	Units	Description
sample_ID	-	A unique value that identifies the collection of a particular species, at a particular location, on a particular day. The value is based on Date_Sample_Subregion_Collector.
scientific_name_original	-	Latin binomial assigned to plant based on most recent taxonomy at the time of data collection
family_MG	-	Family of plant based on the Manning and Goldblatt (2012) taxonomy
scientific_name_WFO	-	Latin binomial of plant based on the World Flora Online v.20263.612 static backbone
scientific_name_authorship_WFO	-	Taxonomic authorship assigned to the World Flora Online taxa
family_WFO	-	Plant family based on the World Flora Online v.20263.612 static backbone
latitude	degrees north	Latitude in decimal degrees; range: -34.36736 to -31.24257
longitude	degrees east	Longitude in decimal degrees; range: 18.38403 to 24.34838
subregion	-	Subregion category
replicate	1	Replicate number of leaf within individual sample plant
date	YYYY-MM-DD	Date of trait collection
collector	-	Collector initial code, retained for provenance of unique identifier
sample	-	Sample number of individual collected
num_leaves	-	Number of leaves used to make area and weight measurements

leaf_area_cm2	cm^2	Leaf area as measured with LI-COR Leaf Area Meter (LAM)
leaf_length_cm	cm	Leaf length of a single leaf, as measured either by LI-COR LAM or with calipers; range: 0.011 - 210.5
avg_leaf_width_cm	cm	Average leaf width, as measured with LI-COR LAM or with calipers; range: 0.002 - 23.5
max_leaf_width_cm	cm	Max leaf width, as measured with LI-COR LAM or with calipers; range: 0.002 - 47.0
leaf_thickness_mm	mm	Thickness of leaf lamina, as measured with micrometer or calipers; range: 0.002 - 34.5
leaf_fresh_wgt_g	g	Leaf fresh (wet) weight; range: 0.007 - 976.036
leaf_dry_wgt_g	g	Leaf dry weight, measured after drying in oven; range: 0.004 - 70.451
twig_fresh_g	g	Stem fresh (wet) weight; range: 0.002 - 103.784
twig_dry_g	g	Stem dry weight, measured after drying in oven; range: 7e-04 - 10.673
lma	$g\ m^{-2}$	Leaf mass area - calculated as leaf dry mass / leaf area; range: 13.461 - 3134.8410
lwc	$g\ g^{-1}$	Leaf water content - calculated as (leaf wet mass - leaf dry mass) / leaf dry mass; range: 0.067 - 63.405
succulence	$g\ m^{-2}$	Succulence - calculated as (leaf wet mass - leaf dry mass) / leaf area; range: 72.549 - 30501.125
ldmc	$g\ g^{-1}$	Leaf dry matter content - calculated as leaf dry mass / leaf wet mass; range: 0.016 - 0.9375
lwr	$cm\ cm^{-1}$	Leaf length-to-width ratio - calculated as leaf length / average leaf width; range: 0.211 - 3086.667
twig_fwc	$g\ g^{-1}$	Stem fresh water content - calculates as stem dry mass / stem wet mass; range: 0.005 - 46.50
quality_flag	-	Quality flags based on outlier assessment

Table 4. Variables in *BioSCape_foliar_species_traits_V2.csv*.

Variable	Units	Description
scientific_name_original	-	Latin binomial assigned to plant based on most recent taxonomy at the time of data collection
scientific_name_MG	-	Latin binomial of plant based on the Manning and Goldblatt (2012) taxonomy
family_MG	-	Family of plant based on the Manning and Goldblatt (2012) taxonomy
scientific_name_WFO	-	Latin binomial of plant based on the World Flora Online v.20263.612 static backbone
scientific_name_authorship_WFO	-	Taxonomic authorship assigned to the World Flora Online taxa
family_WFO	-	Plant family based on the World Flora Online v.20263.612 static backbone
perennial	-	Binary trait extracted from regional floras and expert opinion (1 = perennial, 0 = not perennial)
annual	-	Binary trait extracted from regional floras and expert opinion (1 = annual, 0 = not annual)
monoecious	-	Binary trait extracted from regional floras and expert opinion (1 = monoecious, 0 = not monoecious)
dioecious	-	Binary trait extracted from regional floras and expert opinion (1 = dioecious, 0 = not dioecious)
evergreen	-	Binary trait extracted from regional floras and expert opinion (1 = evergreen, 0 = not evergreen)
deciduous	-	Binary trait extracted from regional floras and expert opinion (1 = deciduous, 0 = not deciduous)

seasonally_apparent	-	Binary trait extracted from regional floras and expert opinion (1 = species is seasonally apparent, 0 = species is apparent year round)
seasonally_identifiable	-	Binary trait extracted from regional floras and expert opinion (1 = species is identifiable only during part of the season, 0 = identifiable year round)
flower_postfire	-	Binary trait extracted from regional floras and expert opinion (1 = species does flower after a fire, 0 = species does not flower after fire)
resprout_postfire	-	Binary trait extracted from regional floras and expert opinion (1 = species does resprout after a fire, 0 = species does not resprout after fire)
serotinous	-	Binary trait extracted from regional floras and expert opinion (1 = species is serotinous, 0 = species is not serotinous)
flower_begin	month	Abbreviation of month flowering begins (e.g., "All year" - "Sep")
flower_end	month	Abbreviation of month flowering ends (e.g., "Apr" to "Sep")
flower_begin_alt	month	Ambiguous flowering begin month; there are several possible reasons a record may have included alternative end flowering, e.g., regional differences in range (e.g., "Apr" to "Sep")
flower_end_alt	month	Ambiguous flowering end month; there are several possible reasons a record may have included alternative end flowering, e.g., regional differences in range (e.g., "Apr" to "Nov")
max_height_cm	cm	Maximum height based on regional floras and expert opinion
herb	-	Binary trait extracted from regional floras and expert opinion (1 = species is herb, 0 = species is not an herb)
geophyte	-	Binary trait extracted from regional floras and expert opinion (1 = species is a geophyte, 0 = species is not a geophyte)
graminoid	-	Binary trait extracted from regional floras and expert opinion (1 = species is a graminoid, 0 = species is not a graminoid)
low_shrub	-	Binary trait extracted from regional floras and expert opinion (1 = species is a shrub less than 1 m tall, 0 = species is not a shrub less than 1 m tall)
mid_shrub	-	Binary trait extracted from regional floras and expert opinion (1 = species is a shrub between than 1 to 2 m tall, 0 = species is not a shrub between 1 to 2 m tall)
tall_shrub	-	Binary trait extracted from regional floras and expert opinion (1 = species is a shrub taller than 2 m, 0 = species is not a shrub taller than 2 m)
tree	-	Binary trait extracted from regional floras and expert opinion (1 = species is a tree, 0 = species is not a tree)
scrambler	-	Binary trait extracted from regional floras and expert opinion (1 = species is a scrambler, 0 = species is not a scrambler)
liana	-	Binary trait extracted from regional floras and expert opinion (1 = species is a liana, 0 - =species is not a liana)
hemiparasite	-	Binary trait extracted from regional floras and expert opinion (1 = species is a hemiparasite, 0 = species is not a hemiparasite)
parasite	-	Binary trait extracted from regional floras and expert opinion (1 = species is a parasite, 0 = species is not a parasite)
missing_growthform_class	-	Binary trait extracted from regional floras and expert opinion (1 = species has an unknown growth form class, 0 = species has a known growth form class)
succulent	-	Binary trait extracted from regional floras and expert opinion (1 = species is a succulent, 0 = species is not a succulent)
leaf_succulent	-	Binary trait extracted from regional floras and expert opinion (1 = species is a leaf succulent, 0 = species is not a leaf succulent)
stem_succulent	-	Binary trait extracted from regional floras and expert opinion (1 = species is a stem succulent, 0 = species is not a stem succulent)
other_succulent	-	Binary trait extracted from regional floras and expert opinion (1 = species is a succulent based on a different part of the plant, 0 = species is not a succulent based on a different part of the plant)

lvs_linear	-	Binary trait extracted from regional floras and expert opinion (1 = species has linear leaves and include shapes such as linear, filiform, needle-like, terete, or sword-shaped, 0 = species does not have linear leaves)
lvs_shape_other	-	All shapes not classified as linear or oval during regional flora extraction. (1 = species have leaves that include shapes such as lanceolate, oblanceolate, linear- lanceolate, linear- oblanceolate, falcate, linear-oblong, oblong-lanceolate, sword-shaped. 0 = species have either linear or rounded/orbicular shapes.)
lvs_oval	-	Binary trait extracted from regional floras and expert opinion (1 = species has shapes including ovate, oblong, cordate, ovate-lanceolate, rhomboid, subround, 0 = species does not have oval leaves)
lvs_compound	-	Binary trait extracted from regional floras and expert opinion (1 = species has compound leaves including descriptions such as leaflets, 3-folialate, trifoliate, pinnate, digitate, 0 = species does not have compound leaves)
lvs_lobed	-	Binary trait extracted from regional floras and expert opinion (1 = species has lobed leaves, 0 = species does not have lobed leaves)
lvs_dissected	-	Binary trait extracted from regional floras and expert opinion (1 = species has dissected leaves including specific terms such as pinnatisect, pinnatifid, pinnae and bipinnatisect, 0 = species does not have dissected leaves)
leaf_length	cm	Leaf length as measured by herbarium specimens; range: 0.10 - 160.84
virtually_no_leaves	-	Binary trait extracted from regional floras and expert opinion (1 = species has virtually no leaves, 0 = species has leaves)
stems_photosynthetic	-	Binary trait extracted from regional floras and expert opinion (1= species has photosynthetic stems, 0 - species does has photosynthetic stems)
functional_leaf	-	Categorical trait describing main photosynthetic surface of the plant: "leaf", "mixed", "culm_stem", and "none"
plant_longevity	-	Ordinal variable: 0 = <2 y; 1 = 2-8 y; 2 = 8-25 y; 3 = >25 y
flammability	-	Ordinal variable: low (l), med (m), high (h)
pollination	-	Categorical variable from regional floras and expert opinion describing the main pollinator of the plant: i = insect; b = bird; m = mammal; w = wind; and u = unknown
dispersal	-	Categorical variable from regional floras and expert opinion describing the main dispersal mechanism of the plant: i = insect; m = mammal; w = wind; p = passive; and u = unknown
lvs_frond	-	Binary trait extracted from regional floras and expert opinion (1 = species has fronds, 0 = species does not)
lvs_cladode	-	Binary trait extracted from regional floras and expert opinion (1 = species has cladodes, 0 = species does not)
lvs_phyllode	-	Binary trait extracted from regional floras and expert opinion (1 = species has phyllodes, 0 = species does not)
lvs_microphyll	-	Binary trait extracted from regional floras and expert opinion (1 = species has microphylls, 0 = species does not)

Table 5. Variables in *BioSCape_foliar_vnir_spectra_V2.csv*.

Variable	**Units**	**Description**
sample_ID	-	A unique value that identifies the collection of a particular species, at a particular location, on a particular day. The value is based on Date_Sample_Subregion format (e.g., "619_1_HK", "825_9_HTR")
scientific_name_original	-	Latin binomial assigned to plant based on most recent taxonomy at the time of data collection (e.g., Abutilon sonneratianum, Zygophyllum procumbens)
family_MG	-	Family of plant based on the Manning and Goldblatt (2012) taxonomy (e.g., Acanthaceae, Zygophyllaceae)
scientific_name_WFO	-	Latin binomial of plant based on the World Flora Online v.20263.612 static backbone (e.g., Abutilon sonneratianum, Zygophyllum procumbens)
scientific_name_authorship_WFO	-	Taxonomic authorship assigned to the World Flora Online taxa

family_WFO	-	Plant family based on the World Flora Online v.20263.612 static backbone
latitude	degrees north	Latitude of sample location; range: -34.36736 to -31.27247
longitude	degrees east	Longitude in of sample location; range: 18.38403 to 24.34838
subregion	-	Subregion category
date	YYYY-MM-DD	Date of spectral collection
sample	1	Sample number of individual collected
X450 - X949 (500 columns)	1	Percent reflectance value by wavelength. Column names include wavelength in nm: *X<wavelength>*. Range: 0 - 90.24003.

Table 6. Variables in *BioSCape_foliar_releve_V2.csv*.

Variable	Units	Description
scientific_name_original	-	Latin binomial assigned to plant based on most recent taxonomy at the time of data collection (e.g., Abutilon sonneratianum, Zyrphelis microcephala)
scientific_name_WFO	-	Latin binomial of plant based on the World Flora Online v.20263.612 static backbone (e.g., Abutilon sonneratianum, Zyrphelis microcephala)
scientific_name_authorship_WFO	-	Taxonomic authorship assigned to the World Flora Online taxa
family_WFO	-	Plant family based on the World Flora Online v.20263.6 12 static backbone (e.g., Acanthaceae, Zygophyllaceae)
recovery_status	-	Taxonomic harmonizing match status ("flagged_non_taxonomic_entry ", "wfo_matched")
releve_ID	-	Unique identifier based on combination of plot and year (e.g., "BK_1_1992", "LB2_99_1993")
plot	-	Plot ID of relevé (e.g., "BK_1", "LB2_99")
year	-	The year in which the the survey was either conducted or published; range: 1966 - 2014
subregion	-	Subregion in which relevé was conducted (e.g., baviaanskloof, langeberg)
latitude	degrees north	Latitude of plot location
longitude	degrees east	Longitude of plot location
percent_cover	percent	Percent cover of observed plant; range: 0.01 - 87.5
abundance_class	-	Abundance class if abundance was originally recorded on an ordinal scale.Abundance class if abundance was originally recorded on a Braun-Blanquet scale. 0 corresponds to a single or sporadic individual (<1 % cover), 1 corresponds to insignificant cover (1-5% cover), 2 corresponds to low cover (5-25%), 3 corresponds to 25-50% cover, 4 corresponds to 50-75% cover, and 5 corresponds to 75-100% cover
cover_method	-	Method used to assess taxon cover ("condensed_Acocks_abundance_class_only", "percent_cover_continuous")
plot_percent_cover	percent	Sum total plot percent cover; range: 0.1 - 585
relative_percent_cover	1	The species percent cover divided by the total plot percent cover; range: 0 - 1
plot_size	-	When made available by survey documentation, size of relevé plot: "10x10m" or "20x20m"

3. Application and Derivation

The Greater Cape Floristic Region of South Africa is a megadiverse region threatened by climate change, human-driven habitat loss, and invasive species. Trait-based ecology offers a mechanistic approach to understand and predict how ecosystems function and respond to disturbance. Leveraging the interactions between plants and incoming solar radiation, remote sensing offers a way to measure foliar traits at landscape scales. In addition to greatly expanding the availability of foliar trait data for this region, these data are contributing to ongoing major research initiatives such as the Biodiversity Survey of the Cape (BioSCape) project and should prove invaluable for future efforts to understand this ecosystem from a trait-based perspective.

4. Quality Assessment

Data cleaning and taxon harmonization are documented in workflows available in Frye (2026). Unrealistic foliar chemistry and structural values (e.g., percent element values over 100% or negative water content values) were removed from the dataset or if possible corrected based on the original data entries. Where applicable, replicate measurements are fully retained to allow users to conduct appropriate assessments of uncertainty. Sampling constraints due to labor and availability of plant material may limit the amount of measurement replication.

Version 2 update

Foliar chemical and structural traits were compared to previously reported global ranges (e.g., Kattge et al., 2020) and flagged if they fell outside of the reported ranges. The spectroscopy library has received additional vetting such that reflectance values that fall below 30% at 800 nm are removed because those measures were likely cases where the leaf or leaf mat was not fully covering the accessory field of view.

In cases where replicate measures are present, each replicate was compared to the median value of the replicate set using the ratio of the value to its set median. Replicates with very high ratio values were removed due to possible measurement or data recording errors. An additional check was made using the coefficient of variation (CV) between replicates such that CV values in the 95th percentile were removed as potential measurement errors. Both the median ratio and CV criteria were used to flag and remove structural trait values from version 1: 60 measures of LMA, 61 measures of leaf water content, 178 measures of thickness, and 239 values of twig water content values. Canopy cover and height measurements did not have replication within samples and were flagged as outliers if the log of the ratio between the measure and the median of the plant family was greater than 2 (or 100 times greater than the median).

Leaf mass per area and nitrogen measures were joined based on sample ID and their log values were correlated to one another with standardized major axis regression, matching the leaf economic tradeoff of Wright et al. (2004). Samples were flagged as possible outliers if the absolute value of the scaled residuals was above 3.

5. Data Acquisition, Materials, and Methods

Between 2010–2014 plant trait data was collected across the Greater Cape Floristic Region in South Africa. Sampling efforts were concentrated in six subregions where either historical or concurrently measured relevé surveys were conducted. While some of these floristic inventories took place several decades before the time of foliar collection and the local vegetation has likely changed due to the influence of fire and other drivers, these surveys still represent the dominant vegetation of the broader subregion. Foliar collections took place in the vicinity (typically < 1 km) of the relevé plots, but not within the original plots to avoid disturbance of vegetation for long-term monitoring efforts.

For foliar traits measurements, sunlit leaves from the top of the canopy were collected from a median of five individuals per species across all subregions. Leaves were placed in plastic bags with moist paper towels and kept in a cooler box or refrigerator and processed within 24 hours. Twenty-four morphological and chemical traits on leaves, stems, and canopies were measured following the standardized protocols of Cornelissen et al. (2003) and Pérez-Harguindeguy et al. (2013). These traits include: leaf area (cm^2), leaf length (cm), average leaf width (cm), maximum leaf width (cm), leaf thickness (mm), leaf fresh weight (g), leaf dry weight (g), twig fresh weight (g), twig dry weight (g), leaf mass per area ($g\ cm^{-2}$), leaf water content ($g\ g^{-1}$), leaf succulence ($g\ cm^{-2}$), leaf dry matter content ($g\ g^{-1}$), leaf length-to-width ratio ($cm\ cm^{-1}$), twig water content ($g\ g^{-1}$), canopy height (cm), canopy area (cm^2), branch order, pubescence, nitrogen (N; %), carbon, (C; %), carbon-nitrogen ratio (C:N), nitrogen isotope ratio ($\delta^{15}N$; ‰), and carbon isotope ratio ($\delta^{13}C$; ‰). In cases where the stem organ acted as the primary photosynthetic organ (e.g., culms of Restionaceae species or stems of succulent species) the stem material was treated as a leaf.

In addition to the foliar traits, leaf reflectance spectroscopy measurements were made on sunlit leaves using a USB-4000 spectrometer (Ocean Optics Inc., Largo, Florida, USA) with a leaf clip attachment. Note that mats of stems were measured in cases where the stem represented the main photosynthetic surface (e.g., some graminoids and stem succulents) covering the full field of view of the clip attachment. The spectrometer range was 450–950 nm, and the data were interpolated to 1-nm band centers.

As ancillary data, a compilation of species-level life history and phenological traits was created based on Germishuizen and Meyer (2003), Manning and Goldblatt (2012), and input from expert botanists. Both the historical and concurrent relevé survey data were collated for all subregions.

For further details regarding the description, measurement, and biogeographical context of these data, please see Frye et al. (2026).

To make ecologically informative inferences using abundance weighted data, researchers should combine the percent cover values from the relevé data and foliar trait data. Broad trait and environmental trends should always be interpreted with a degree of caution given the strong role that fire plays in the Fynbos ecosystems. Additionally, researchers are encouraged to link these data to regional abiotic data used in the Ecological Monitoring and Management Application project (EMMA; https://github.com/AdamWilsonLab/emma_envdata; Slingsby et al., 2020) and to explore NASA's related BioSCape Project. For the most recent regionally accepted taxonomic designations, check taxonomic names in these data with the *Botanical Database of Southern Africa* (South African National Biodiversity Institute, 2016).

Dataset Revisions for Version 2

An extensive effort to flag and remove possible outliers arising from measurement and data recording errors was implemented (described in the above in Section 3). The notable updates include:

- The relevé locations over the Hangklip subregion were updated. The original coordinates were estimated before GPS availability and on an older datum. When plotted, some of the original geolocations fell into the ocean or far away from their intended locations. To provide more accurate locations, the map figure reported in Boucher (1978) was rubber-sheeted into the modern UTM projection. and plot locations were hand-adjusted using the Georeferencer tool in QGIS v3.36. Ground control points were set on rocky features along the coastline using 25-cm resolution aerial imagery (Chief Directorate: National Geospatial Information (CD: NGI), 2025) as a base map. The Boucher figure was then georeferenced with a thin plate spline and bilinear resampling.

- The relevé dataset originally did conversions between abundance and percent cover scales. Due to the known challenges of this conversion (Damgaard, 2014; Fischer, 2015; van der Maarel, 2007), the data are reported as it was originally measured and clearly labeled to indicate which vegetation scale was used in each survey effort.
- A join error was discovered that dropped taxa from the Hantam-Tanqua-Roggeveld relevé surveys. This error is now fixed.
- In *BioSCape_foliar_species_traits_V2*, the leaf_type column was removed because of conflicts with other data sources in the table deemed more reliable. Non-problematic information from this column was used to create the lvs_frond, lvs_cladode, lvs_phyllode, and lvs_microphyll variables.

6. References

Aiello-Lammens, M.E., J.A. Slingsby, C. Merow, H.K. Mollmann, D. Euston-Brown, C.S. Jones, and J.A. Silander. 2016. Processes of community assembly in an environmentally heterogeneous, high biodiversity region. Ecography 40:561–576. https://doi.org/10.1111/ecog.01945

Boucher, C. 1978. Cape Hangklip area. II. The vegetation. Bothalia 12:455–497.https://doi.org/10.4102/abc.v12i3.1816

Cornelissen, J.H. C., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D.E. Gurvich, P.B. Reich, H. ter Steege, H.D. Morgan, M.G. A. van der Heijden, J.G. Pausas, and H. Poorter. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany 51:335-380. https://doi.org/10.1071/BT02124

Damgaard, C. 2014. Estimating mean plant cover from different types of cover data: A coherent statistical framework. Ecosphere 5:1-7. https://doi.org/10.1890/ES13-00300.1

Fischer, H.S. 2015. On the combination of species cover values from different vegetation layers. Applied Vegetation Science 18:169-170. https://doi.org/10.1111/avsc.12130

Frye, H. 2026. henryf6/GCFR_Traits: Code release (Version v2.0). Zenodo. https://doi.org/10.5281/zenodo.18189262

Frye, H., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, R. Turner, A.M. Wilson,and J.A. Silander Jr. 2026. Foliar Trait and Spectroscopy Data Greater Cape Floristic Region, South Africa (Version 1). ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2482

Frye, H.A., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, R. Turner, A.M. Wilson, and J.A. Silander Jr. 2026. Foliar trait and spectroscopy data from the megadiverse Greater Cape Floristic Region. In review.

Frye, H.A., M.E. Aiello-Lammens, D. Euston-Brown, C.S. Jones, H. Kilroy Mollmann, C. Merow, J.A. Slingsby, H. van der Merwe, A.M. Wilson, and J.A. Silander. 2021. Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. Global Ecology and Biogeography 30:1403-1417. https://doi.org/10.1111/geb.13306

Germishuizen, G. and N.L. Meyer (eds). 2003. Plants of southern Africa: An annotated checklist. South African National Biodiversity Institute (SANBI Publishing); Pretoria, South Africa.

Kattge, J., G. Bönisch, S. Díaz, S. Lavorel, I.C. Prentice, P. Leadley, S. Tautenhahn, G.D.A. Werner, et al. 2019. TRY plant trait database – enhanced coverage and open access. Global Change Biology 26:119–188. https://doi.org/10.1111/gcb.14904

Manning, J. and P. Goldblatt. 2012. Plants of the Greater Cape Floristic Region: Vol. 1: The Core Cape flora. South African National Biodiversity Institute (SANBI Publishing); Pretoria, South Africa. ISBN: 978-1-919976-74-7

Perez-Harguindeguy, N., S. Diaz, E. Garnier, S. Lavorel, H. Poorter, P. Jaureguiberry, M.S. Bret-Harte, W.K. Cornwell, J.M. Craine, D.E. Gurvich, C. Urcelay, E.J. Veneklaas, P.B. Reich, L. Poorter, I.J. Wright, P. Ray, L. Enrico, J.G. Pausas, A.C. de Vos, N. Buchmann, G. Funes, F. Quetier, J.G. Hodgson, K. Thompson, H.D. Morgan, H. ter Steege, M.G. A. van der Heijden, L. Sack, B. Blonder, P. Poschlod, M.V. Vaieretti, G. Conti, A.C. Staver, S. Aquino, and J.H. C. Cornelissen. 2013. New handbook for standardised measurement of plant functional traits worldwide. Australian Journal of Botany 61:167-234. https://doi.org/10.1071/BT12225

Schliep, E.M., A.E. Gelfand, R.M. Mitchell, M.E. Aiello-Lammens, and J.A. Silander. 2017. Assessing the joint behaviour of species traits as filtered by environment. Methods in Ecology and Evolution 9:716-727. https://doi.org/10.1111/2041-210X.12901

Slingsby, J.A., C. Merow, M. Aiello-Lammens, N. Allsopp, S. Hall, H. Kilroy Mollmann, R. Turner, A.M. Wilson, and J.A. Silander. 2017. Intensifying postfire weather and biological invasion drive species loss in a Mediterranean-type biodiversity hotspot. Proceedings of the National Academy of Sciences 114:4697-4702. https://doi.org/10.1073/pnas.1619014114

South African National Biodiversity Institute. 2016. Botanical Database of Southern Africa (BODATSA). https://posa.sanbi.org/

van der Maarel, E. 2007. Transformation of cover-abundance values for appropriate numerical treatment- Alternatives to the proposals by Podani. Journal of Vegetation Science 18:767–770. https://doi.org/10.1111/j.1654-1103.2007.tb02592.x

van Leeuwen, M., H.A. Frye, and A.M. Wilson. 2021. Understanding limits of species identification using simulated imaging spectroscopy. Remote Sensing of Environment 259:112405. https://doi.org/10.1016/j.rse.2021.112405

White, P.A., H. Frye, M.F. Christensen, A.E. Gelfand, and J.A. Silander. 2022. Spatial functional data modeling of plant reflectances. The Annals of Applied Statistics 16. https://doi.org/10.1214/21-AOAS1576

White, P.A., M.F. Christensen, H. Frye, A.E. Gelfand, and J.A. Silander. 2023. Joint multivariate and functional modeling for plant traits and reflectances. Environmental and Ecological Statistics 30:501–528. https://doi.org/10.1007/s10651-023-00574-7

White, P.A., H.A. Frye, J.A. Slingsby, J.A. Silander, and A.E. Gelfand. 2023. Generative spatial generalized dissimilarity mixed modelling (spGDMM): An enhanced approach to modelling beta diversity. Methods in Ecology and Evolution 15:214–226. https://doi.org/10.1111/2041-210X.14259

White, P.A., A.E. Gelfand, H. Frye, and J.A. Silander. 2024. Good modelling practice in ecology, the hierarchical Bayesian perspective. Ecological Modelling 496:110847. https://doi.org/10.1016/j.ecolmodel.2024.110847

7. Dataset Revisions

Version	Release Date	Revision Notes
2	2026-09-18	See details about quality assurance in Sections 4 and Dataset Revisions in Section 5.
1	2026-02-17	Original publication: https://doi.org/10.3334/ORNLDAAC/2482