– sPlot Letter No. 5 –

21 January 2016

Dear members of the sPlot Consortium,
Dear custodians and deputy custodians of sPlot member databases,

Happy New Year to all of you! Eleven months have passed since the 4th sPlot Letter, and we apologize for the silence since then, which was due to both the complex issues to be solved and the fact that two of us (Oliver Purschke and Jürgen Dengler) now have other main professional responsibilities. Today, however, we are happy to celebrate with you the release of sPlot 2.0, the most comprehensive vegetation-plot database ever in terms of geographic coverage and the first sPlot version available for scientific studies.

Table of contents:

sPlot 2.0 released!

Thanks to your support, **sPlot 2.0** (released today) contains **1,115,705 plots** from **110 source databases.** They originate from all **seven continents** and a total of **130 countries (Table 1).** This means an **increase by 39 databases, 199,179 plots and 40 countries** since the release of sPlot 1.1, presented in the sPlot Letter No. 4. **Nearly all sPlot data (99.9%) are geo-referenced** (see Figs. 1 and 2), albeit sometimes with low precision (but each plot contains a field for spatial precision). We are extremely grateful to Stephan Hennekens for having set up and for maintaining this unique database under **Turboveg 3**.

Table 1: Distribution of vegetation plots in sPlot 2.0 across continents. Note that some data bases contain data from more than one continent; thus, their numbers do not sum up to 110. The plot number is slightly less than the total because of few plots that are not yet assigned to a country.

Continent	Area (km²)	Data bases	Number of plots	Density (plots/100 km²)
Europe (incl. Turkey & Cyprus)	10,956,477	54	941,425	8.592
North America (ice-free part)	22,739,261	10	90,698	0.399
Australasia	9,008,500	5	23,452	0.260
Asia (excl. Turkey & Cyprus)	43,043,523	25	31,987	0.074
Africa	30,370,000	14	17,578	0.058
South America	17,840,000	11	10,235	0.057
Antarctica (ice-free part)	247,860	1	19	0.008
Total	**134,205,621**	**110**	**1,115,394**	**0.831**
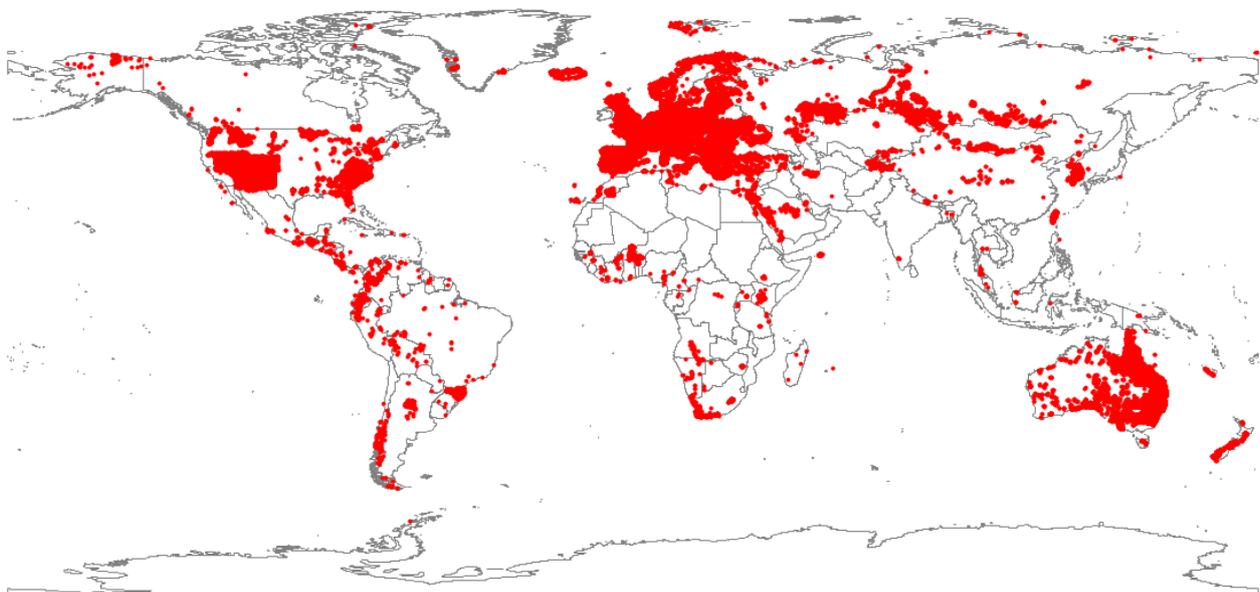
Fig. 1: Geographic distribution of vegetation plots in sPlot 2.0.

0 < density <= 0.01
0.01 < density <= 0.1
0.1 < density <= 1
1 < density <= 10
10 < density <= 100
density > 100
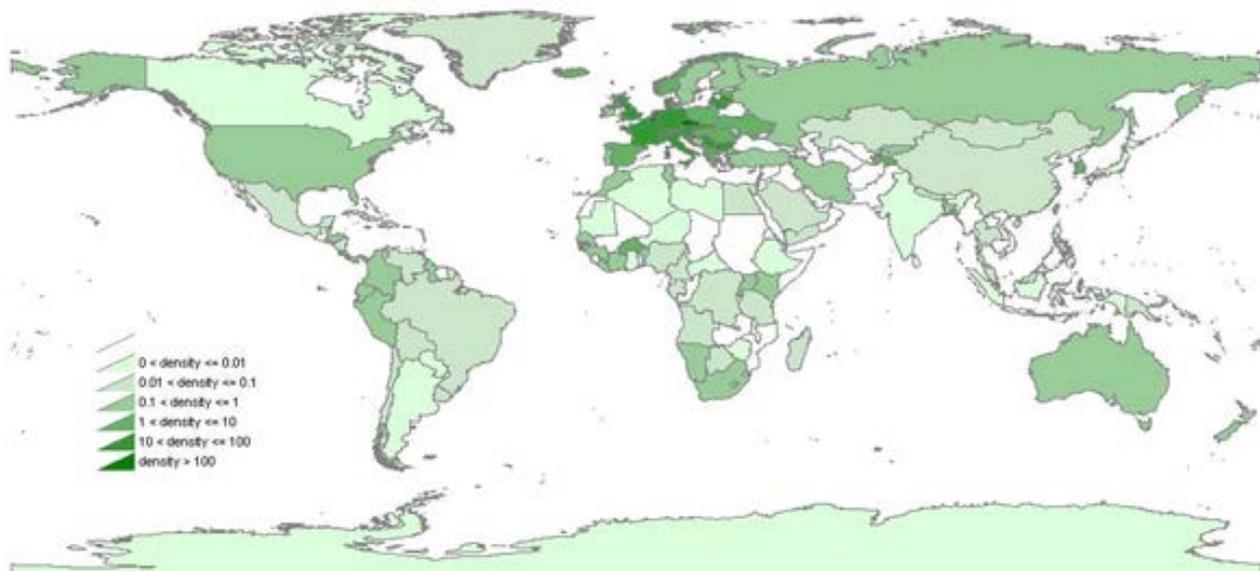
Fig. 2: Density of vegetation plots (number per 100 km²) per country as contained in sPlot 2.0.

Taxonomic backbone linking sPlot 2.0 to TRY 3.0

To link sPlot 2.0 to **TRY 3.0** in an as accurate as possible way, and to maximize the match between the two databases, there is now a unified taxonomic backbone (generated by Oliver Purschke in November 2015). The backbone unifies the in total **122,901** (**partly non-standardized**) **species names** in the two databases, of which 62,628 names were unique to sPlot 2.0, 35,477 unique to TRY 3.0 and 24,796 shared between the two databases. Our iterative name standardization procedure (based on string manipulation algorithms, *The Plant List*, *Taxonomic Name Resolution* Service and manual corrections) revealed that there were 95,485 accepted names, 20,952 synonyms, 4,772 unresolved and only 1,692 names (1.38%) for which no suitable matches were found even after manual search. A total of 105,818 names could be resolved down to the species level, 13,383 to the genus level and 1,880 to the family level. We finally ended up with **90,969 standardized unique names** (sPlot and TRY together), belonging to **665 families**, of which 95.4 % were vascular plant species.

Match between sPlot 2.0 and TRY 3.0

Based on the taxonomic backbone, we calculated the match between the resolved species names in sPlot and the set of species for which a fully gap-filled trait matrix (see Shan et al. 2012, *ICML*) or observed values for six of the most common traits in TRY 3.0 (see Díaz et al. 2015, *Nature*; and Fig. 3 below) are available. While there was an overall match of > 40% between all vascular plant species in sPlot 2.0 and the gap-filled species in TRY 3.0, the match increased to >60% when based on the most frequent species in sPlot (Fig. 3). This is very good news, since most of planned trait-based analysis will be based on the gap-filled trait data.
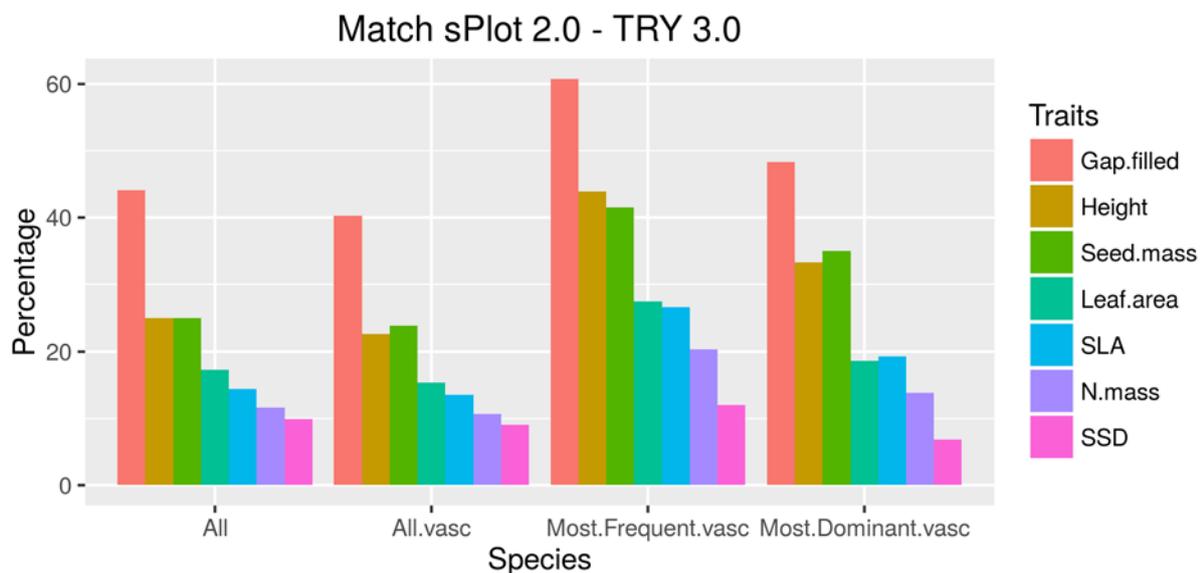

Fig. 3: Match (in %) between species names in sPlot 2.0 and TRY 3.0, based on (i) all species in sPlot (incl. non-vascular plants), (ii) all vascular plant species, (iii) the most frequent vascular species (upper 33.3 percentile of species overall occurrence frequencies in sPlot) and iv) the most dominant vascular species (upper 33.3 percentile of average species abundances in sPlot). Statistics are presented for the sets of species for which gap-filled trait information or trait measurements for the following six traits (also see Díaz et al. 2015, *Nature*) are available in TRY 3.0: plant height, seed mass, leaf area, specific leaf area, leaf N content per dry mass (N.mass) and stem specific density (SSD).

Metadata management via GIVD

As announced earlier, sPlot has made an agreement with the **Global Index of Vegetation-Plot Databases (GIVD)** to **manage our "metadata"** as it also does it for EVA. Most sPlot databases were already in GIVD and, for those that were not, many custodians have recently created an entry in GIVD. For those few remaining databases that did not have a GIVD entry, we created one that is currently invisible to the public and only serves the maintenance of your metadata. The custodians and deputy custodians of the respective databases have received an e-mail from Jürgen about how they can manage their metadata there and how they could make the information about their database in GIVD publically visible if they like.

The most **important metadata sPlot retrieves from GIVD** are the proper name of your database, the identity of custodian and deputy custodian, their e-mail and addresses and whether a database report has been published somewhere that should be cited when your data are used via sPlot. Therefore, please ensure that the names, e-mails and addresses of the custodian and deputy custodian in GIVD are correct and up-to-date. We will use this information to contact you and to decide whom to invite for co-authorship of sPlot papers. The GIVD metadata also allow you to list more than two persons who are originators of or contributed to your database, to give them proper credit.

More **benefits of having your database registered in GIVD**, you can find at http://www.givd.info/givd/faces/info_benefits.xhtml.

Preparation of sPlot publications is starting now

With today's release of version 2.0, sPlot becomes available for scientific studies aiming at publications. We remind both data contributors and potential lead authors of the regulations and procedures for data use and authorship as defined in the sPlot Rules No. 5 and 6.

The **first paper** is the **database paper** directly described in Rule 5b. This is an **opt-out paper** and you find the paper **proposal attached**, which has been approved by the Steering Committee. According to the sPlot Rules, the custodians of all 110 member databases (plus those of the two data aggregators EVA and TAVA) become co-authors automatically, unless they notify the **lead author (Jürgen Dengler)** that they wish to be unlisted. The sPlot Steering Committee has decided include additionally the deputy custodians of the most important databases (i.e. those with ≥ 5,000 plots within Europe or ≥ 2,000 plots outside Europe)

This first sPlot paper is planned to be submitted to *Global Ecology and Biogeography* and will be kind of a hybrid between the two following database papers, which you also find attached:

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J.H.J., Aćić, S., Agrillo, E., (…) & Yamalov, S. in press. European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science* 19: 173–180.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M.W., Reich, P.B., Wright, I.J. (…) Wirth, C. 2011. TRY – a global database of plant traits. *Global Change Biology* 17: 2905–2935.

Evidently it is not possible to prepare and discuss such a paper with about 200 co-authors. Therefore a **smaller team of core authors** (listed in the proposal) will do the main work. If some of you other custodians would like to become actively involved, please let the lead author know as soon as possible and he will be happy to include you in the further mailing. All

listed authors will have to see the manuscript two weeks prior to submission, check their names and affiliations and of course provide some last input.

There are currently seven subsequent **analytical papers** envisaged based on sPlot 2.0. They are listed in sPlot Letter No. 3 with their prospective content and lead author. Note that these opt-in papers will only start officially when the lead author has handed in a proposal like the one for the opt-out paper, and this proposal has been approved by the Steering Committee and sent to all sPlot Consortium members. When you receive such a proposal, you have one month to declare your interest to become co-author (**opt-in-paper**). Lead authors have to accept one such offer from each database whose data are included in the analyses. However, this does not necessarily need to be the custodian; rather the custodian and deputy custodian should ensure that there is a fair benefit sharing among the contributing parties of their database. Each opt-in author is expected to provide at least some intellectual input to the respective paper.

You, the sPlot Consortium members, could propose further paper projects (see Rule 5c). However, you should be a little bit patient in this respect because the Steering Committee has to ensure that there are no conflicts between papers (Rule 5c). Though the seven existing proposals will have priority, papers that are topically unconnected to these can be proposed already now.

New sPlot homepage

The **new sPlot homepage** (www.idiv.de/splot), has been online for nearly half a year, but you might not have recognised it yet. It provides updated content and much more functionality, including an overview of all sPlot Consortium members and contributing databases, overview on the current sPlot content and materials for download, including conference presentations about sPlot.

Since some of you requested to have an **"attribution page"** available to which we could link from every paper using sPlot data, we will create this within the sPlot homepage. Each sPlot member database is invited to provide to Jürgen Dengler any text they wish to appear there. For databases that do not provide specific text to us, we will include a full citation of their latest database report in Phytocoenologia or Biodiversity & Ecology (you can also provide citations of additional such database reports), and where such a report does not exist, as a minimum, a link to the URI (unique resource identifier) of your GIVD entry.

sPlot on conferences

Since the last sPlot Letter, our project was presented at four conferences: In April 2015, Oliver Purschke presented us at the EGU conference in Vienna, Austria, with a poster (attached). In May 2015, Jürgen Dengler gave an oral presentation at the conference of the European Vegetation Survey (EVS) in May 2015 in Rennes, France. Then in August 2015 Jens Kattge presented us at the ESA conference in Baltimore, USA. Finally, Helge Bruelheide showed our state of art at the iDiv Workshop in December 2015 in Leipzig, Germany (attached).

Future additions to sPlot

While sPlot 2.0 as a fixed version will serve for the first set of publications, i.e. all will use exactly the same dataset, we are in parallel slowly starting the **preparation of sPlot 3.0**, which we intend to release as a next fixed version **in perhaps 12–18 months** from now.

Since we are not sure about funding and the availability of some of the key players in this period, we cannot promise how many new datasets we can add in release 3.0, but likely we will have to be more restrictive than for sPlot 2.0, where the smallest extra-European dataset contained only 24 plots. In any case you are highly welcome to **contact Jürgen already during the next months and inform him about data that you or someone else would like to contribute** to sPlot, be it a new version of a previously included dataset or a new dataset. The best way to describe the main properties of a not yet included dataset for consideration through us is to upload its metadata in GIVD (www.givd.info) and to send us its GIVD code. Note that European datasets (including Turkey and the Caucasus countries) can only be contributed via our partner **EVA (European Vegetation Archive)**, and tropical African datasets should preferably be contributed via our partner **TAVA (Tropical African Vegetation Archive)**.

With our best regards,

Jürgen Dengler (sPlot Coordinator: juergen.dengler@uni-bayreuth.de)
Oliver Purschke (sPlot Founder: oliver.purschke@idiv.de)
Milan Chytrý (sPlot Steering Committee: chytry@sci.muni.cz)
Valério De Pillar (sPlot Steering Committee: vpillar@ufrgs.br)
Jens Kattge (sPlot Steering Committee: jkattge@bgc-jena.mpg.de)
Brody Sandel (sPlot Steering Committee: brody.sandel@biology.au.dk) &
Helge Bruelheide (sPlot Founder & sPlot Steering Committee: helge.bruelheide@botanik.uni-halle.de)

----
sPlot URL: http://www.idiv.de/splot

German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig
Deutscher Platz 5e
04103 Leipzig
Germany
URL: http://www.idiv.de/