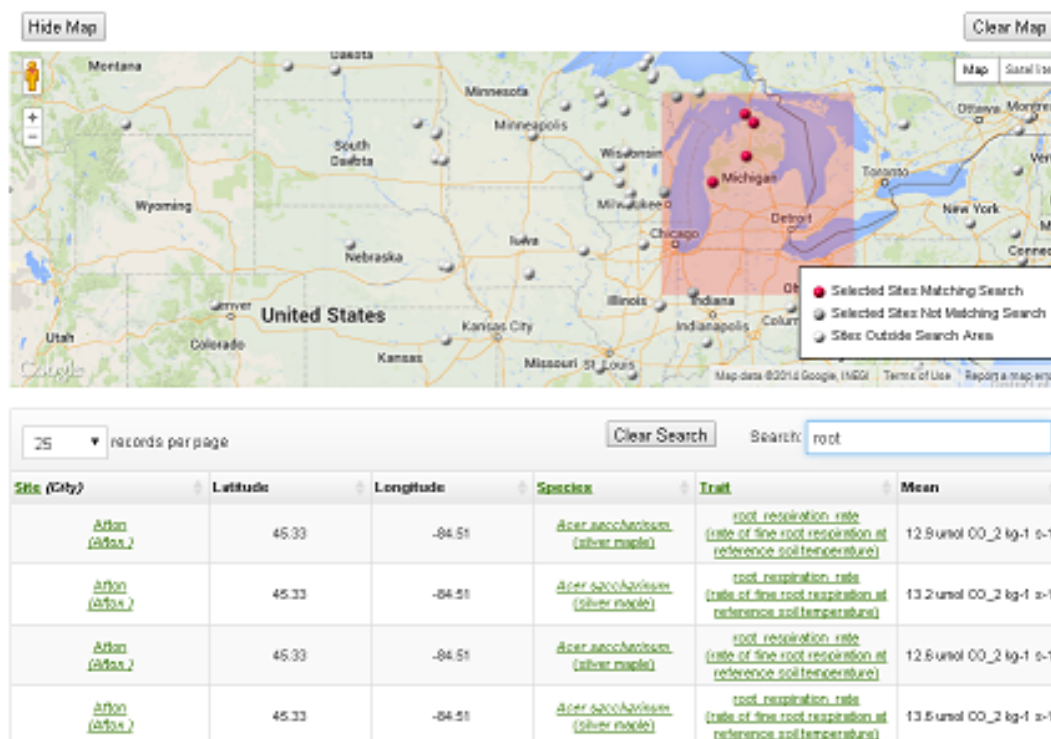


BETYdb Data Access

David LeBauer, Scott, Anjali Patel

Advanced Search

Download Results



1 Guide to accessing data from BETYdb

Data is made available for analysis after it is submitted and reviewed by a database administrator. These data are suitable for basic scientific research and modeling. All reviewed data are made publicly available after publication to users of BETY-db who are conducting primary research. Access to these raw data is provided to users based on affiliation and contribution of data.

All public data in BETYdb is made available under the Open Data Commons Attribution License (ODC-By) v1.0 (<http://opendatacommons.org/licenses/by/1-0/>). You are free to share, create, and adapt its contents. Data with an access_level field and value <= 2 is not covered by this license, but may be available for use with consent.

Please cite the source of data as:

LeBauer, David; Dietze, Michael; Kooper, Rob; Long, Steven; Mulrooney, Patrick; Rohde, Gareth Scott; Wang, Dan; (2010): Biofuel Ecophysiological Traits and Yields Database (BETYdb); Energy Biosciences Institute, University of Illinois at Urbana-Champaign. <http://dx.doi.org/10.13012/J8H41PB9> (<http://dx.doi.org/10.13012/J8H41PB9>)

1.1 Search Box

On the welcome page of BETYdb there is a search option for trait and yield data (Figure 1). This tool allows users to search the entire database for specific sites, species, and traits.

The results page provides a map interface (Figure 2) and the option to download the file in the upper right of the page.

The downloaded file is in csv format. This file provides meta-data and provenance information, including: the date and the SQL query used to access, as well as a citation for BETYdb itself (*to be updated on publication*) and the source of each record (Figure 3).

Instructions: The advanced search can be used by searching for trait and yield data through text. Doing so is very simple, type in the site (city), species or trait into the search bar and the results will show what is in the BetYdb database. The number of records per page can be changed to viewer preference and the search results can be downloaded in an excel sheet format.

(Figure 1)

Advanced search using maps is used by searching for a specific site(city), species and trait. The results of the search are then shown in Google Maps. Sites, species and traits can be searched further by clicking on the geographical location on the map. Additionally, there is a legend that shows selected sites that match search, don't match search, and sites that fall outside search area. All of these results can be downloaded by clicking the "download results" button on the upper right hand corner.

(Figure 2)

1.2 Details

1.2.1 Basic search:

1. Got to the home page. (This will be either: <http://pecandev.igb.illinois.edu/bety> (<http://pecandev.igb.illinois.edu/bety>), <https://ebi-forecast.igb.illinois.edu/beta> (<https://ebi-forecast.igb.illinois.edu/beta>), or <https://www.betydb.org>. (<https://www.betydb.org>))
2. Enter one or more terms in the search box. Each term must match one of (1) the scientific name of a species; (2) the common name of a species; (3) the name of a treatment. Thus, the more terms entered, the more restrictive the search. In addition to normal search terms, the keywords "trait" or "yield" may be entered to restrict the type of data returned.
3. Click the search icon next to the search box (or press the ENTER key). This will take one to the search results/advanced search page and show the result of the search.

1.2.2 Advanced search:

1. From the home page, you can get to the advanced search page simply by doing a basic search. (Or just click the search icon without entering any terms. Or enter the URL for the advanced search page directly by appending "/searches" to the URL given in

step one of the Basic search instructions.)

2. By default, the search will return results for both traits and yields. Use the radio buttons to restrict results to one or the other.
3. Enter one or more terms in the main search box. As with basic searches, each term must match one of (1) the scientific name of a species; (2) the common name of a species; (3) the name of a treatment. (Unlike basic searches, the “trait” and “yield” keywords are not recognized. Using them here will most likely eliminate all search results since they don’t match the name of a species or treatment. Use the radio buttons instead in order to restrict the search type.)
4. Press the Submit button to obtain search results.
5. Narrowing Search Results: Once you have obtained a search result, you may refine the result by applying further restrictions using the “Narrow your search” panel on the right side of the page. Enter terms in one or more of the three search boxes there to limit results by Site, by Species, or by Trait. Then either click one of the “Apply Limit” links or click the Submit button again.

Notes:

1. Search terms are case-insensitive and a search term will match a species name, treatment name, site name or trait name if it matches any part of the name. Thus, simply typing “misc” will match against “Miscanthus”, “Miscanthus sacchariflorus”, “Miscanthus sinensis”, or “Miscanthus X giganteus”.
2. The “species” search box in the “Narrow your search” panel only matches against scientific names. To match against common names, put the term in the main search box.
3. For sites having no formal site name, the city name is shown instead. Note however that when narrowing your search by site, the term entered will only be matched against site names.
4. The layout of the advanced search page looks best when your browser window is as wide as possible.

1.2.3 Example

1. Go to <http://pecandev.igb.illinois.edu> (<http://pecandev.igb.illinois.edu>)
2. Simple search: sugarcane yields -- Enter the search term "sugarcane" followed by the keyword "yields" in the "SEARCH FOR TRAITS AND YIELDS" box; then click the search icon.
3. You will see a new page with a table of results showing data about sugarcane yields.
4. Advanced search task 1: Switch to showing trait data -- Click the radio button next to the word "traits"; then click the Submit button.
5. Now you will see search results related to sugarcane trait data.
6. Advanced search task 2: Limit results to a specific site -- We just want to see the results for the 'Mackay' site. In the 'Narrow Your Search ...' area on the right side of the page, enter the word 'Mackay' into the 'By Site' box; then click the 'Apply Limit' link.
7. You will see only the search results for the 'Mackay' site.
8. Advanced search task 3: Limit results to a specific trait -- We only want to see the results concerning shoot density. In the 'Narrow Your Search ...' area on the right side of the page, enter the word 'density' into the 'By Trait' box; then click the 'Apply Limit' link.

9. The resulting table rows all concern total shoot density data for sugarcane at the Mackay site.

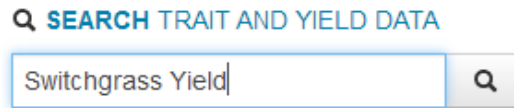


Fig. 1 Searching for switchgrass yield data using a textbox on the BETYdb home page (www.betydb.org (<http://www.betydb.org>)).

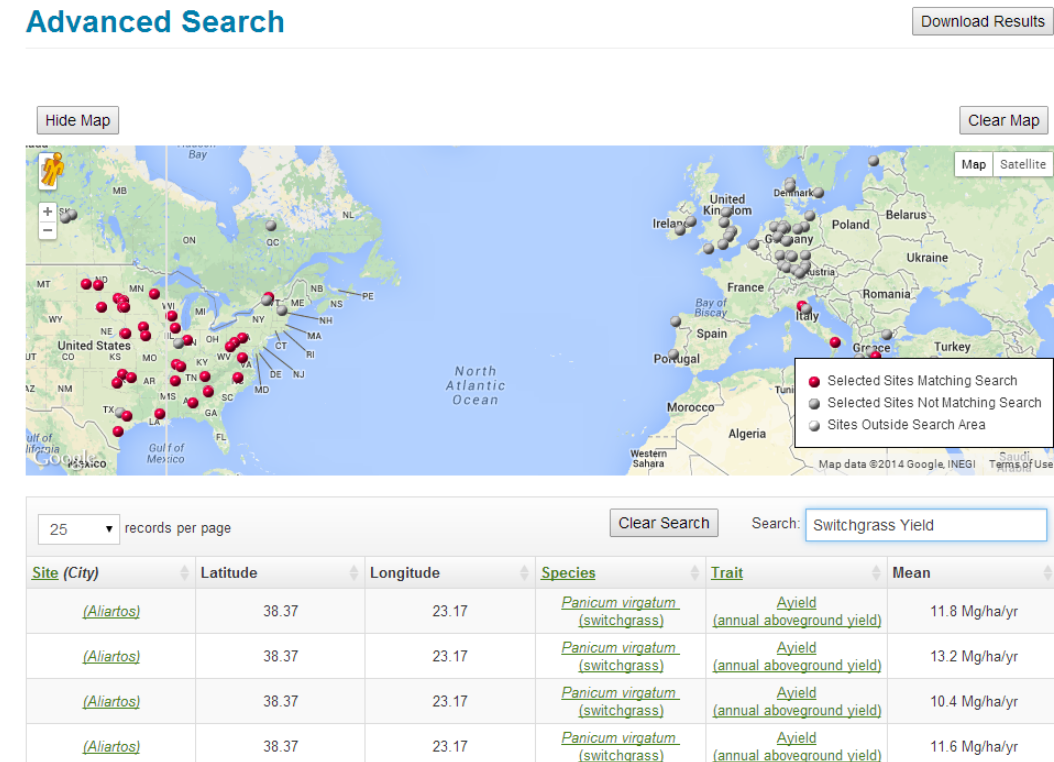


Fig. 2 Viewing results for a search for "switchgrass yield": a map of sites for which data is available, including the option to download.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
2	David LeBauer, Dan Wang, and Michael Dietze, 2010. Biofuel Ecophysiological Traits and Yields Database Version 1.0. Energy Biosciences Institute, Urbana, IL																					
3	SQL: diebauer@illinois.edu																					
4	SELECT "traits_and_yields_view".* FROM "traits_and_yields_view" WHERE (access_level >= 4) AND ((LOWER(traits_and_yields_view.scientificname) LIKE "%switchgrass%" OR LOWER(traits_and_yields_view																					
5	Date of #####																					
6																						
7																						
8	sitename	city	lat	lon	scientific	common	genus	author	citation	y	treatment	date	month	year	dateloc	trait	mean	units	n	statname	stat	notes
9	Alariatos		38.37	23.17	Panicum v switchgra:	Panicum	Alexopolc	2008				#####	11	1998	6	A yield	15.2 Mg/ha/yr				[missing]	
10	Alariatos		38.37	23.17	Panicum v switchgra:	Panicum	Alexopolc	2008				#####	11	1998	6	A yield	11.8 Mg/ha/yr				[missing]	
11	Alariatos		38.37	23.17	Panicum v switchgra:	Panicum	Alexopolc	2008				#####	11	1998	6	A yield	13.2 Mg/ha/yr				[missing]	
12	Alariatos		38.37	23.17	Panicum v switchgra:	Panicum	Alexopolc	2008				#####	11	1998	6	A yield	10.4 Ms/ha/yr				[missing]	

Fig. 3 The first few results (of almost 1388 total) for the query "switchgrass yield".

2 The BETYdb API

TODO: What is an API? Why is it useful?

The API has two options: easy and hard.

2.1 Easy Way

The easy way provides most useful information in a single table in a csv format that is easy to use - in any spreadsheet software or scripting language (or json or xml, see below).

Data can be downloaded as a `.csv` file, and data from previously published syntheses can be downloaded without login. For example, to download all of the Switchgrass (*Panicum virgatum* L.) yield data,

```
https://www.betydb.org/traits.csv?genus=Switchgrass
```

2.2 Hard Way

The "hard way" downloads individual tables from the database. This approach allows more complex queries and faster programatic access. In the future, this feature will be used to access BETYdb by the `rotraits` package (<https://github.com/ropensci/rotraits/issues/3>) under development by rOpenSci. It would be straightforward to translate functions in the PEAdb package to use the API. Currently, PEAdb accesses the PostgreSQL database directly, via ssh tunnel.

However, most of the useful information, for example about sites and treatments, is provided on other tables. Thus, useful queries will should trouble is that this file will need to

2.2.1 Cross-table queries using CSV, JSON and XML APIs

BetyDB has the ability to return any object in these three formats. All the tables in BetyDB are RESTful, which allows you to GET, POST, PUT, or DELETE them without interacting with the web interface. Here are some examples. In all of these examples, you can use exchange `.json` and `.xml` depending on the format you want.

The format of your request will need to be:

2.2.2 Examples

Return all citations in json format (replace json with xml or csv for those formats)

```
https://www.betydb.org/citations.json
```

Return all yield data for the genus 'Miscanthus'

```
https://www.betydb.org/yields.json?genus=Miscanthus  
https://www.betydb.org/yields.json?genus=Miscanthus&species=giganteus
```

Return all yield data from the with author = Heaton and year = 2004 (can also be queried by title)

```
https://www.betydb.org/citations.xml?include[]=yields&author=Heaton&year=2004
```

Find species associated with the `biocro.salic` pft

```
https://www.betydb.org/species.xml?include[]=pfts_species&include[]=pfts&name=biocro.s  
alix
```

Return all citations with their associated sites (you use the singular version of the associated tables name - site - when the relationship is many to one, and the plural when many to many; hint: if the id of the table you are attempting to include is in the record - relatedtable_id - then it is the singular version.

```
https://www.betydb.org/citations.json?include[]=sites
```

Return all citations with their associated sites and yields (working on ability to nest this deeper)

```
https://www.betydb.org/citations.json?include[]=sites&include[]=yields
```

Return all citations with the field journal equal to 'Agronomy Journal' and author equal to 'Adler' with their associated sites and yields.

```
https://www.betydb.org/citations.json?journal=Agronomy%20Journal&author=Adler&include[]=sites&include[]=yields
```

Return citation 1 in json format, can also be achieved by adding '?id=1' to line 1

```
https://www.betydb.org/citations/1.json
```

Return citation 1 in json format with it's associated sites

```
https://www.betydb.org/citations/1.json?include[]=sites
```

Return citation 1 in json format with it's associated sites and yields

```
https://www.betydb.org/citations/1.json?include[]=sites&include[]=yields
```

2.3 API keys

Using an API key allows access to data without having to enter a login. To use an API key, simply append @?key=@ to the end of the URL. Each user must obtain a unique API key.

3 Installing a local version of BETYdb

The following scripts make it easy to install the latest version of BETYdb locally (on Linux or OSX), populated with all public data. This is the recommended way to use with PEcAn (<https://www.pecanproject.org>); in fact, these scripts are part of the PEcAn repository (<https://github.com/PecanProject/pecan>) and their use is described in the developer's guide to installing PEcAn (<https://github.com/PecanProject/pecan/wiki/Installing-PEcAn>). Users interested in this option, but unfamiliar with Linux should either a) start with the PEcAn Virtual Machine or b) find someone who is familiar with Linux.

- MySQL: <https://github.com/PecanProject/pecan/blob/master/scripts/update.mysql.sh>
(<https://github.com/PecanProject/pecan/blob/master/scripts/update.mysql.sh>)
- PostgreSQL:
<https://github.com/PecanProject/pecan/blob/master/scripts/update.psql.sh>
(<https://github.com/PecanProject/pecan/blob/master/scripts/update.psql.sh>)