

Интерфейс пользователя для загрузки данных реанализа ECMWF ERA5

О массиве данных

ERA5 содержит в себе ежечасные, суточные и месячные данные по 257 атмосферным, приземным и океаническим параметрам. Пространственный шаг модельной сетки составляет 31 км ($\sim 0,28^\circ$), атмосферные параметры представлены на 37 вертикальных уровнях. Данные по волнению рассчитаны для сетки с разрешением $0,36^\circ$. Распространяемые данные были проинтерполированы в узлы регулярной сетки с разрешением $0,25^\circ$.

ERA5 доступен для скачивания через API Climate Data Store (CDS). CDS хранит данные на жестких дисках, а не в архивах, как ECMWF. Это существенно влияет на скорость доступа к данным. Массив находится в свободном доступе и может использоваться как в исследовательских, так и в коммерческих проектах (URL: <https://climate.copernicus.eu/>). В настоящее время ERA5 покрывает период с 1979 по 2018 гг.

Ожидается, что к концу 2020 г. ERA5 будет полностью рассчитан, покрывая временной интервал 1950-2020 гг. Далее массив будет пополняться с задержкой в 3 месяца.

CDS предоставляет пользователям возможность загружать данные как через веб-интерфейс, так и с помощью скриптов в автоматическом режиме. Веб-интерфейс удобен для скачивания небольших объемов данных и простых запросов, тогда как возможности API значительно шире. Для автоматизации скачивания используется язык программирования Python и модуль `cdsap`.

Интерфейс для загрузки данных

Для взаимодействия пользователя с клиентом `cdsap` был разработан визуальный интерфейс пользователя. Интерфейс был написан на языке FreePascal (URL: www.freepascal.org) в среде разработки Lazarus 2.0.6 (URL: www.lazarus-ide.org). Среда Lazarus позволяет создавать кроссплатформенные приложения, значительно облегчая построение визуальных интерфейсов пользователя в десктопном режиме.

Интерфейс спроектирован таким образом, чтобы шаг за шагом провести пользователя через выбор настроек, при этом не загромождая экранное пространство. Внешний вид главного окна программы представлен на рисунке 1.

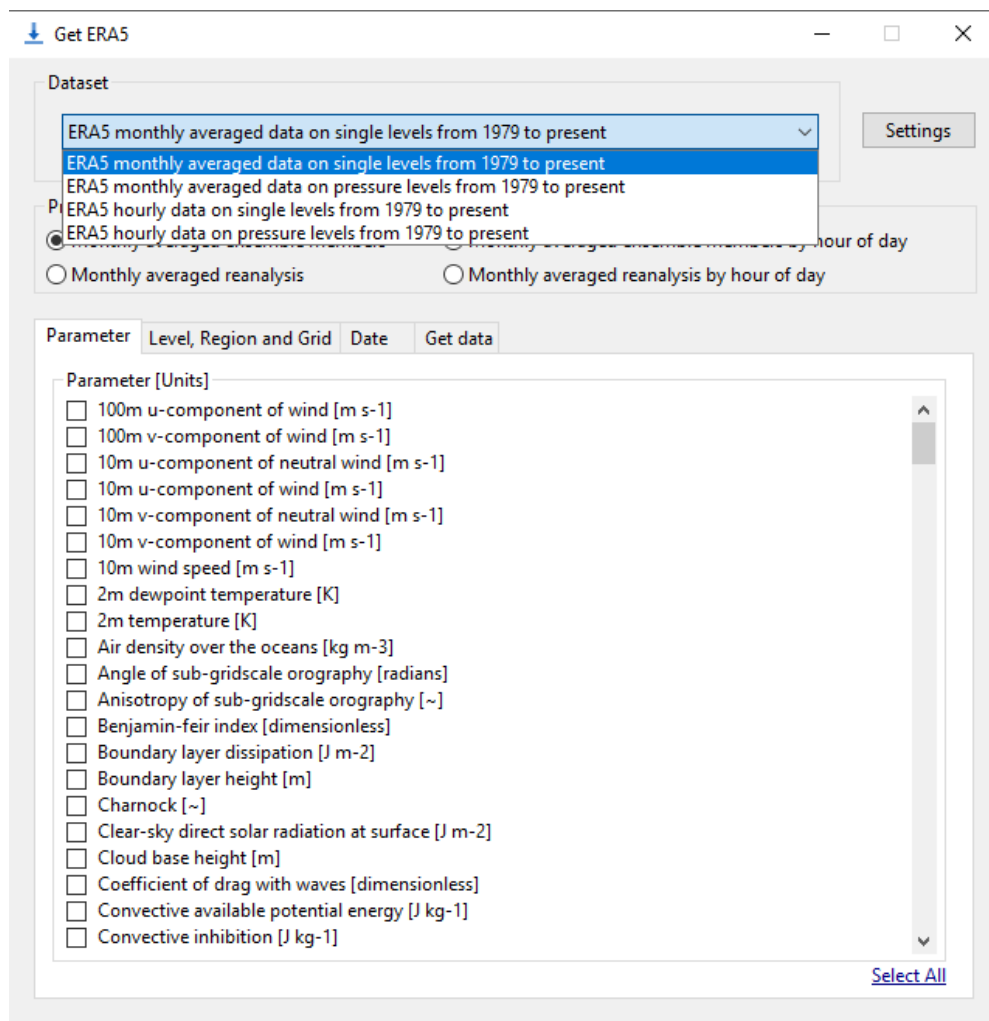


Рисунок 1 – Выбор массива данных

Перед началом работы пользователю необходимо установить (либо выбрать существующий) интерпретатор языка Python. Это можно сделать, перейдя в настройки программы (рисунок 2).

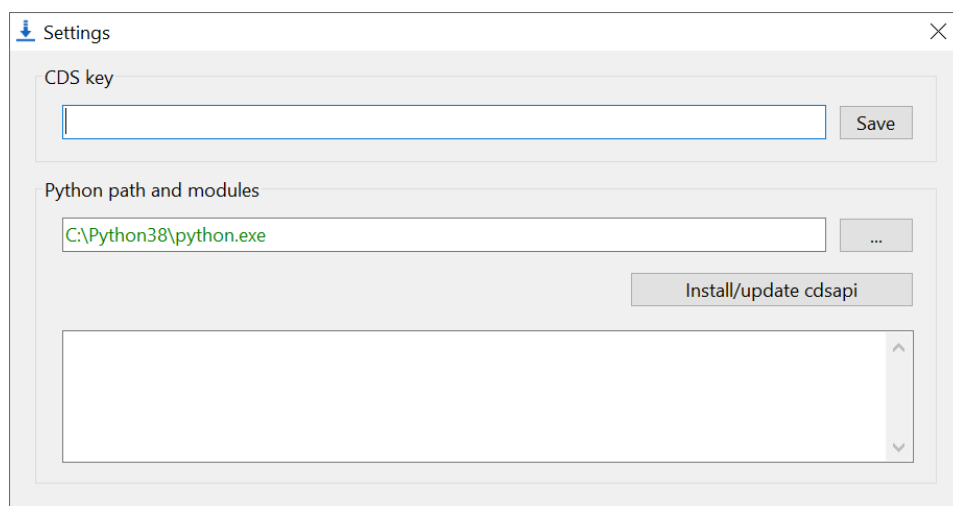


Рисунок 2 – Настройки программы

В зависимости от операционной системы путь может быть абсолютным, либо относительным. Последний зачастую используется в UNIX системах. Для работы программы необходимо установить интерпретатор языка Python 3. Следует отметить, что программа будет работать и с Anaconda, если возникнет такая потребность. Скачать интерпретатор Python 3 можно с официального веб-сайта проекта (URL: <https://www.python.org/>) совершенно бесплатно. После установки Python 3 необходимо указать путь к файлу python.exe, который расположен в корневой папке интерпретатора. После этого нужно установить модуль cdsapi. Сделать это можно нажатием кнопки «Install/update cdsapi».

Помимо Python на форме настроек необходимо указать ключ для работы с сервисами CDS. Для этого пользователь должен зарегистрироваться на портале CDS (URL: <https://cds.climate.copernicus.eu>), и в настройках аккаунта скопировать выданный системой ключ. При сохранении ключа он копируется в папку пользователя, где модуль cdsapi всегда может его найти. И интерпретатор Python, и ключ CDS указываются один раз, при самом первом запуске. Настройки сохраняются и далее используются программой по-умолчанию.

После выполнения первичных настроек можно переходить непосредственно к выбору данных.

Пользователь может выбрать любой из четырех поддерживаемых массивов данных:

- 1) Месячные данные на поверхности,
- 2) Месячные данные на вертикальных уровнях,
- 3) Ежечасные данные на поверхности,
- 4) Ежечасные данные на вертикальных уровнях.

Следует отметить, что всего 16 параметров представлены на вертикальных уровнях. Подавляющее же большинство переменных доступны только для поверхности Земли. Подробный список параметров приведен в документации к массиву ERA5 и не дублируется здесь ввиду большого размера.

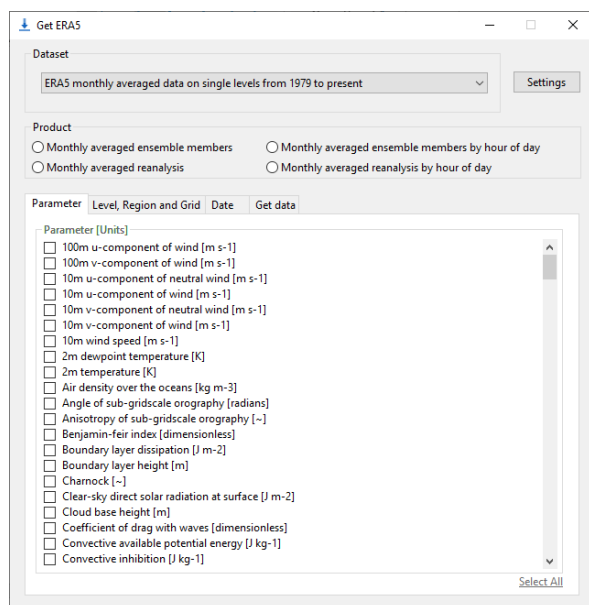
После выбора нужного массива данных пользователю предоставляется возможность выбора продукта. Продукты отличаются для месячных и ежечасных данных. Так для месячных данных (рисунок 3а) пользователь может выбрать:

- 1) Осредненные за месяц ансамбли данных,
- 2) Осредненный за месяц реанализ,
- 3) Осредненные за конкретный час ансамбли данных,
- 4) Осредненные за конкретный час данные реанализа.

Для ежечасных данных доступны следующие продукты (рисунок 3б):

- 1) Средние данные по ансамблю,
- 2) Спред по ансамблю,
- 3) Стандартное отклонение по ансамблю,
- 4) Реанализ.

а)



б)

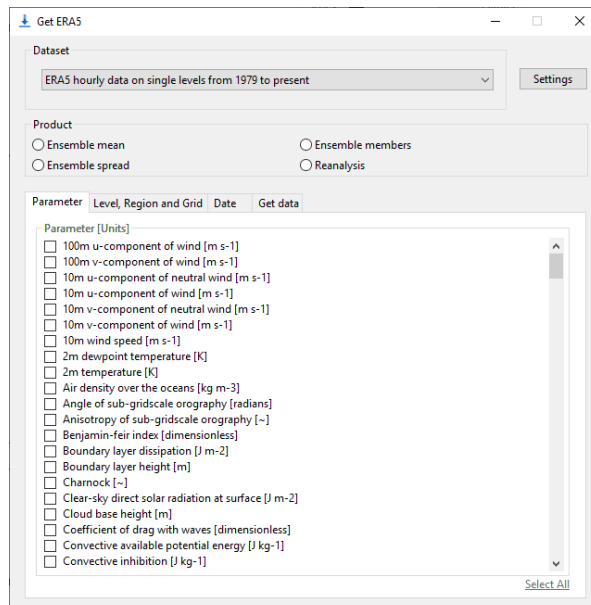


Рисунок 3 – Выбор продукта

После выбора продукта пользователь может выбрать из списка интересующие его параметры. Количество параметров не ограничено, но чем их больше, тем медленнее будет происходить выгрузка данных.

После выбора параметров пользователь может перейти к выбору региона и разрешения грида (рисунок 4). Несмотря на то, что разрешение массива составляет $0,25^\circ$, CDS предоставляет возможность задать несколько различных вариантов сетки. Реинтерполяция производится на серверах CDS, пользователь получает данные в нужном ему пространственном разрешении. На этой вкладке так же можно выбрать вертикальные уровни для атмосферных параметров.

На следующей вкладке пользователь переходит к выбору даты и времени (рисунок 6). В зависимости от массива данных и продукта доступны различные опции. Выбор года и месяца одинаков для всех продуктов. День можно выбрать только для ежечасных данных. Указать конкретные часы можно только для конкретных продуктов, которые поддерживают такой выбор. Более подробная информация приведена в описании ERA5 на сайте ECMWF.

Get ERA5

Dataset: ERA5 monthly averaged data on pressure levels from 1979 to present

Product:

☐ Monthly averaged ensemble members
 ☐ Monthly averaged ensemble members by hour of day
 ☒ Monthly averaged reanalysis
 ☐ Monthly averaged reanalysis by hour of day

Parameter | Level, Region and Grid | Date | Get data

Region:

Grid:

☒ 0.25
 ☐ 1.0
 ☐ 2.5
 ☐ 0.5
 ☐ 1.5
 ☐ 3.0
 ☐ 0.75
 ☐ 2.0
 ☐ 5.0

Level, hPa:

<input type="checkbox"/> 1000	<input type="checkbox"/> 875	<input type="checkbox"/> 750	<input type="checkbox"/> 500	<input type="checkbox"/> 250	<input type="checkbox"/> 125	<input type="checkbox"/> 20	<input type="checkbox"/> 2
<input type="checkbox"/> 975	<input type="checkbox"/> 850	<input type="checkbox"/> 700	<input type="checkbox"/> 450	<input type="checkbox"/> 225	<input type="checkbox"/> 100	<input type="checkbox"/> 10	<input type="checkbox"/> 1
<input type="checkbox"/> 950	<input type="checkbox"/> 825	<input type="checkbox"/> 650	<input type="checkbox"/> 400	<input type="checkbox"/> 200	<input type="checkbox"/> 70	<input type="checkbox"/> 7	
<input type="checkbox"/> 925	<input type="checkbox"/> 800	<input type="checkbox"/> 600	<input type="checkbox"/> 350	<input type="checkbox"/> 175	<input type="checkbox"/> 50	<input type="checkbox"/> 5	
<input type="checkbox"/> 900	<input type="checkbox"/> 775	<input type="checkbox"/> 550	<input type="checkbox"/> 300	<input type="checkbox"/> 150	<input type="checkbox"/> 30	<input type="checkbox"/> 3	

[Select All](#)

Рисунок 4 – Выбор региона и разрешения грида

Get ERA5

Dataset: ERA5 monthly averaged data on pressure levels from 1979 to present

Product:

☒ Monthly averaged ensemble members
 ☐ Monthly averaged ensemble members by hour of day
 ☐ Monthly averaged reanalysis
 ☐ Monthly averaged reanalysis by hour of day

Parameter | Level, Region and Grid | Date | Get data

Year:

<input checked="" type="checkbox"/> 1979	<input checked="" type="checkbox"/> 1984	<input checked="" type="checkbox"/> 1989	<input checked="" type="checkbox"/> 1994	<input checked="" type="checkbox"/> 1999	<input checked="" type="checkbox"/> 2004	<input checked="" type="checkbox"/> 2009	<input checked="" type="checkbox"/> 2014	<input checked="" type="checkbox"/> 2019
<input checked="" type="checkbox"/> 1980	<input checked="" type="checkbox"/> 1985	<input checked="" type="checkbox"/> 1990	<input checked="" type="checkbox"/> 1995	<input checked="" type="checkbox"/> 2000	<input checked="" type="checkbox"/> 2005	<input checked="" type="checkbox"/> 2010	<input checked="" type="checkbox"/> 2015	
<input checked="" type="checkbox"/> 1981	<input checked="" type="checkbox"/> 1986	<input checked="" type="checkbox"/> 1991	<input checked="" type="checkbox"/> 1996	<input checked="" type="checkbox"/> 2001	<input checked="" type="checkbox"/> 2006	<input checked="" type="checkbox"/> 2011	<input checked="" type="checkbox"/> 2016	
<input checked="" type="checkbox"/> 1982	<input checked="" type="checkbox"/> 1987	<input checked="" type="checkbox"/> 1992	<input checked="" type="checkbox"/> 1997	<input checked="" type="checkbox"/> 2002	<input checked="" type="checkbox"/> 2007	<input checked="" type="checkbox"/> 2012	<input checked="" type="checkbox"/> 2017	
<input checked="" type="checkbox"/> 1983	<input checked="" type="checkbox"/> 1988	<input checked="" type="checkbox"/> 1993	<input checked="" type="checkbox"/> 1998	<input checked="" type="checkbox"/> 2003	<input checked="" type="checkbox"/> 2008	<input checked="" type="checkbox"/> 2013	<input checked="" type="checkbox"/> 2018	

[Select All](#)

Month:

<input checked="" type="checkbox"/> January	<input checked="" type="checkbox"/> March	<input checked="" type="checkbox"/> May	<input checked="" type="checkbox"/> July	<input checked="" type="checkbox"/> September	<input checked="" type="checkbox"/> November
<input checked="" type="checkbox"/> February	<input checked="" type="checkbox"/> April	<input checked="" type="checkbox"/> Jun	<input checked="" type="checkbox"/> August	<input checked="" type="checkbox"/> October	<input checked="" type="checkbox"/> December

[Select All](#)

Day:

<input checked="" type="checkbox"/> 1	<input checked="" type="checkbox"/> 4	<input checked="" type="checkbox"/> 7	<input checked="" type="checkbox"/> 10	<input checked="" type="checkbox"/> 13	<input checked="" type="checkbox"/> 16	<input checked="" type="checkbox"/> 19	<input checked="" type="checkbox"/> 22	<input checked="" type="checkbox"/> 25	<input checked="" type="checkbox"/> 28	<input checked="" type="checkbox"/> 31
<input checked="" type="checkbox"/> 2	<input checked="" type="checkbox"/> 5	<input checked="" type="checkbox"/> 8	<input checked="" type="checkbox"/> 11	<input checked="" type="checkbox"/> 14	<input checked="" type="checkbox"/> 17	<input checked="" type="checkbox"/> 20	<input checked="" type="checkbox"/> 23	<input checked="" type="checkbox"/> 26	<input checked="" type="checkbox"/> 29	
<input checked="" type="checkbox"/> 3	<input checked="" type="checkbox"/> 6	<input checked="" type="checkbox"/> 9	<input checked="" type="checkbox"/> 12	<input checked="" type="checkbox"/> 15	<input checked="" type="checkbox"/> 18	<input checked="" type="checkbox"/> 21	<input checked="" type="checkbox"/> 24	<input checked="" type="checkbox"/> 27	<input checked="" type="checkbox"/> 30	

[Select All](#)

Time:

<input checked="" type="checkbox"/> 00:00	<input type="checkbox"/> 03:00	<input type="checkbox"/> 06:00	<input type="checkbox"/> 09:00	<input type="checkbox"/> 12:00	<input type="checkbox"/> 15:00	<input type="checkbox"/> 18:00	<input type="checkbox"/> 21:00
<input type="checkbox"/> 01:00	<input type="checkbox"/> 04:00	<input type="checkbox"/> 07:00	<input type="checkbox"/> 10:00	<input type="checkbox"/> 13:00	<input type="checkbox"/> 16:00	<input type="checkbox"/> 19:00	<input type="checkbox"/> 22:00
<input type="checkbox"/> 02:00	<input type="checkbox"/> 05:00	<input type="checkbox"/> 08:00	<input type="checkbox"/> 11:00	<input type="checkbox"/> 14:00	<input type="checkbox"/> 17:00	<input type="checkbox"/> 20:00	<input type="checkbox"/> 23:00

[Select All](#)

Рисунок 5 – Выбор даты и времени

После выбора нужной конфигурации пользователь может приступить к загрузке данных на вкладке «Получить данные» (рисунок 6). При нажатии на кнопку «Download» интерфейс пользователя автоматически генерирует скрипт на языке Python, который запускается на исполнение. Результаты работы скрипта выводятся в окне, что позволяет пользователю отследить этапы загрузки. Также предусмотрена возможность вывести на экран скрипт, который передается в интерпретатор Python.

По завершении загрузки автоматически откроется папка с данными, а интерфейс проинформирует пользователя об успешном завершении операции. Выходной формат данных - netCDF4. Несмотря на то, что оригинальный реанализ хранится в формате GRIB2, перекодировка в netCDF происходит на серверах CDS за минимальное время.

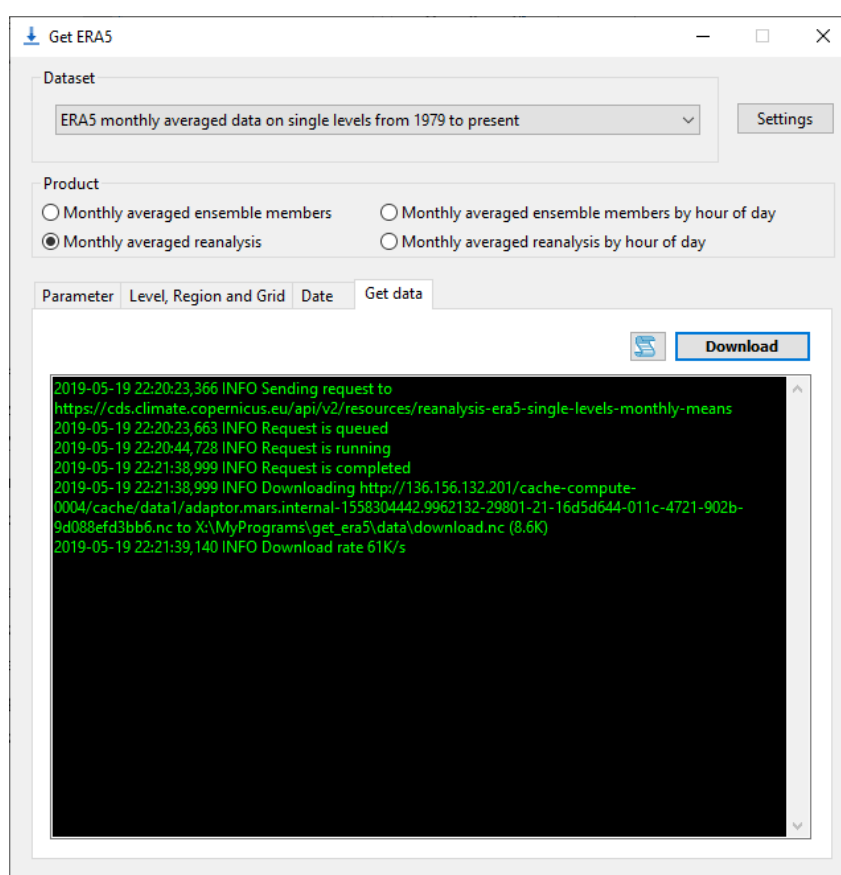


Рисунок 6 – Загрузка данных

Разработанный интерфейс был протестирован для различных массивов данных и продуктов ERA5. Он может распространяться в сборках для Microsoft Windows (XP и выше) и Linux. Поддерживаются как 64-битные, так и 32-битные системы.

С переводом реанализа из хранилищ ECMWF в CDS значительно возросла скорость и удобство доступа к данным. Используя модуль загрузки, даже неподготовленный пользователь сможет выбрать и закачать нужные ему данные, не ограничиваясь малофункциональным веб-интерфейсом.