Received: 22 July 2022 | Revised: 17 November 2022 | Accepted: 14 December 2022
DOI: 10.1002/met.2111

RESEARCH ARTICLE

Meteorological Applications Open Access
Science and Technology for Weather and Climate
RMetS

The MÉRA Data Extraction toolkit

Levent Görgü[1] | **Daniel Hawtree**[2] | **Michael J. O'Grady**[1] | **Conor Muldoon**[1] | **Bartholomew Masterson**[3] | **Wim G. Meijer**[3] | **John J. O'Sullivan**[2] | **Gregory M. P. O'Hare**[4]

[1]UCD School of Computer Science, University College Dublin, Dublin, Republic of Ireland
[2]Dooge Center for Water Resources Research, UCD School of Civil Engineering, University College Dublin, Dublin, Republic of Ireland
[3]UCD School of Biomolecular & Biomedical Science, University College Dublin, Dublin, Republic of Ireland
[4]School of Computer Science and Statistics, Trinity College Dublin, Dublin, Republic of Ireland

Correspondence
Levent Görgü, UCD School of Computer Science, University College Dublin, Dublin, Republic of Ireland.
Email: levent.gorgu@ucd.ie

Funding information
EU INTERREG VA; European Regional Development Fund

Abstract
Historical meteorological datasets are indispensable for forming climatic models and the generation of weather forecasts. Such data are core to the training phase of prediction models and may also be harnessed for hydrological and environmental models. GRIB is the most common data format used in meteorology and represents the de facto standard for storing historical weather data. However, GRIB datasets are complex and do not constitute analysis-ready data without additional preprocessing. The MÉRA dataset of the Met Eireann, the Irish meteorological service, is archetypical of such datasets, emerging from a high-resolution climatic reanalysis of Irish weather data between 1981 and 2019. This article describes the MÉRA Data Extractor toolkit. This toolkit enables the intuitive, fast extraction and preprocessing of data from this extensive dataset. The toolkit is available as open source and will be of interest to those researching climate modelling in Europe.

KEYWORDS
big data, climate data, forecasting data, forecasting systems, model development, remote sensing

1 | INTRODUCTION

Fulfilling the potential of Earth Observation (EO) datasets demands that any barriers preventing their adoption be removed and that the data be made available for speedy analysis. Such analysis-ready data (ARD) (Huber et al., 2021), or tidy data (Wickham, 2014), allow scientists immediately focus on the task at hand. As ARD is closely coupled to the research question under investigation, it cannot be anticipated and produced by the data providers. For remotely sensed data and satellite imagery, data cubes represent one well-known effort for the uniform production of ARD (Giuliani et al., 2017). Raw data are often made available in GRIB format for climatic and weather datasets. However, additional preprocessing is needed to produce meteorological ARD for the analysis being conducted.

An extensive range of tools has been developed for working with GRIB data. These may be GUI-based and perhaps only useful for visualization. Alternatively, tools may be command-line driven but demand significant user expertise to use them effectively. Visualization tools lack critical functionality, such as extracting only the needed data and supporting subsequent manipulation or processing to produce ARD. This article describes a toolkit to produce ARD for GRIB datasets publicly available from Met Éireann, the Irish National Meteorological Service.

This article is structured as follows. Section 2 presents an overview of the MÉRA and the GRIB data format. Section 4 describes the design and the technical dimension of the MÉRA Data Extractor toolkit. Some use cases illustrating the toolkit's potential are documented in Section 5. Section 5 concludes this article.

2 | BACKGROUND

Met Eireann has recently completed a regional climate reanalysis for Ireland encompassing data from 1981 to 2019 (Gleeson et al., 2017; Whelan et al., 2018). The resultant dataset, MÉRA, covers all of Ireland and the United Kingdom (see Figure 1) on a 2.5-km^2 grid (Whelan et al., 2017). This dataset includes an extensive range of variables, 220 in total, and the entire dataset is more than seven terabytes. Its spatial and temporal resolution far exceeds alternative historical datasets. Thus, MÉRA has significant potential outside of core climatic and meteorological modelling and could inform diverse applications in environmental monitoring and services for smart agriculture (Hawtree et al., 2020). MÉRA is publicly available in the GRIB format and under a Creative Commons Attribution 4.0 International (CC BY 4.0) license (MetÉireann, 2022).

GRIdded Binary or General Regularly distributed Information in Binary (GRIB) is a data format commonly used in meteorology to store historical and forecast weather data. The GRIB data format is standardized by the World Meteorological Organization's Commission for Basic Systems. Data within GRIB files are in highly compressed binary format, rendering the data complex and challenging to use, and limiting its potential for archiving data (Caron & Oxelson, 2013). For researchers and service providers outside of the meteorological domain, using GRIB data incurs a steep learning curve. Thus, the potential of meteorological datasets, such as the MÉRA dataset, is compromised.

Currently, there are three versions of GRIB. Version 0 is no longer in operational use. The first version (GRIB1) is used operationally worldwide by most meteorological facilities. The MÉRA dataset conforms to GRIB version 1 (GRIB1). A second generation, GRIB2, has been introduced and is slowly being adopted. The data in GRIB1 are typically converted to integers using scale and offset, and then bit-packed. GRIB version 2 (GRIB2) also has possibilities for compression.

Structurally, GRIB files are a collection of self-contained records of 2D data. Individual records stand as meaningful data, with no references to other records or an overall schema. For extensive datasets such as MÉRA, a bottleneck results as the size of the dataset renders it challenging to extract and process the data fully. Making such datasets analysis ready is thus a formidable task, exacerbated by a chronic lack of tools. For example, unprocessed MÉRA data provide wind as vectors (rather than speed and direction) and precipitation on a 33-h cumulative cycle. Thus, significant additional processing is needed to render the dataset effectively usable.

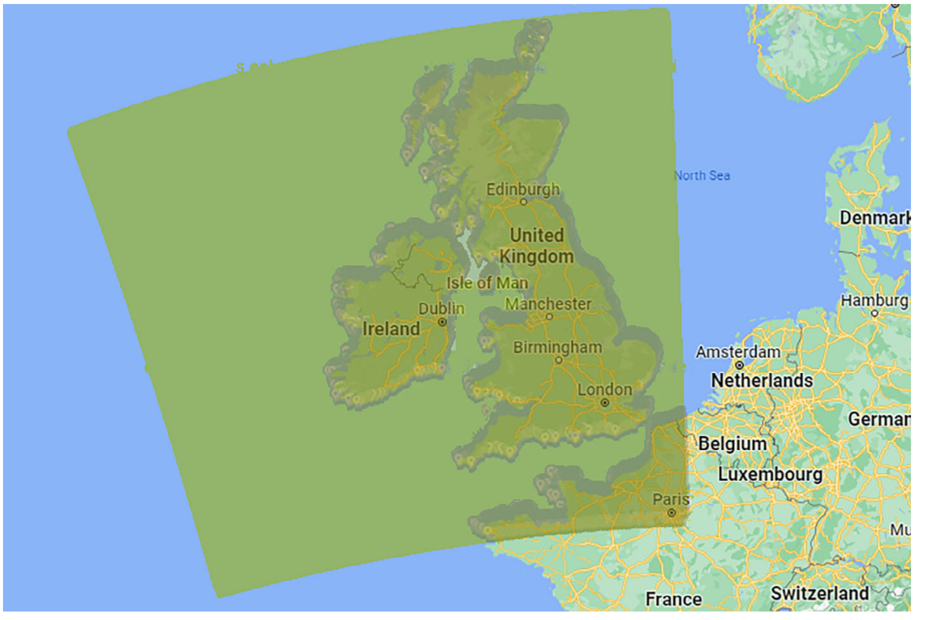
2.1 | GRIB toolkits

Various commercial, open source and freely available toolkits for processing GRIB data are widely documented. Such toolkits facilitate downloading GRIB files from various locations on the WWW and support general tools for manipulating and visualizing the data. It is in the generality of such tools that a fundamental difficulty arises. GRIB files are often different based on where they were initially generated. Moreover, there is no way to verify the origin of a GRIB file or check the consistency and quality of the data. Thus, to interact with GRIB files, only a set of common features can be extracted and visualized. Specific tools must be developed for GRIB files in which the internal structures are sufficiently known.

A detailed description of all the tools available is beyond the scope of this discussion. Appendix A lists some of the most common toolkits available to researchers. However, it is instructive to comment on some of them.

Amongst the most common GUI tools for GRIB files are Panoply (NASA, 2022) and MetView (Russell et al., 2020). Likewise, well-known commercial packages are Mathematica and ArcGIS. Most of these cannot extract the data except Panoply Data Viewer from NASA's Goddard Institute for Space Studies (GISS). GISS's Panoply can extract data for a limited time, often only an hour. Thus, even for data of 1 day, 24 separate extractions should be undertaken for all the available data locations, and the needed location(s) should be filtered from these extractions as a post-process. Panoply does not support mass data extraction based on location, which ends with an unusable extraction or an extraction with usage limited to another visualization tool. Wolfram's Mathematica and ArcGIS from ESRI are very sophisticated commercial packages that support various file formats, including GRIB. Mathematica supports the visualization of GRIB files, but its support for formats with complex internal structures is limited. ECMWF ecCodes (ECMWF, 2022) is another tool that supports extraction from GRIB files but only in JSON format, which is a standard text-based format for representing structured data based on JavaScript object syntax. Several computer tools, such as JSON, are briefly explained in

FIGURE 1 MÉRA model domain.



Appendix B for readers who are unfamiliar with these computer tools and associated computer science terminology. Extraction into other formats is possible only with some programming and utilizing the ecCodes' Fortran, C or Python libraries.

Ultimately, difficulties using GRIB coalesce around a need for metadata standards, controlled vocabularies and governance (Wright, 2014). The netCDF format is commonly used as it overcomes some of the problems associated with GRIB; however, it cannot help if the database was generated using GRIB format. Such issues have ramifications for using open science principles in climate research, limiting the application of the findable, accessible, interoperable, and reusable (FAIR) principles to meteorological datasets (de Vos et al., 2020). Efforts to resolve these problems are documented. Hales et al. (2021) demonstrate a platform for providing uniform access to multidimensional water data. The Copernicus C3S Climate Data Store (CDS) adopts a common data model to circumvent many problems resulting from heterogeneous file formats (Raoult et al., 2017).

3 | THE MÉRA DATA EXTRACTOR TOOLKIT

This section presents an overview of the design and architecture of a data extraction toolkit (see Figure 2). Conceptually, this may be regarded as a backend server, where MÉRA GRIB files are extracted and processed, and a frontend GUI where end users can select the necessary locations, the variable type and a time period before starting the data extraction process.

3.1 | Backend server

The original GRIB files for Met Eireann's MÉRA dataset must be downloaded prior to starting the extraction process. Downloading all MÉRA files was simply not feasible due to size issues. In practice, it is not necessary as most environmental models focus on historical data from the previous 20 years. In the MÉRA Data Extraction toolkit, we enabled data extraction from only the most recent 15 years of historical data (2004–2019). Moreover, for this particular use case, only seven variables were of interest out of 200 possible. Specifically, temperature (at 2 m), rainfall (total precipitation), soil moisture, wind (speed and direction separately at 10 m), sunlight (direct normal irradiance [DNI]) and pressure at mean sea level (PMSL) were all extracted to be utilized as predictive variables during the development of environmental models. Therefore, it is critical for our toolkit to provide these data in a format that is suitable as input without extensive additional processing. After excluding unnecessary years and unused data types, we managed to decrease the size of the MÉRA files needed to approximately 2.3 terabytes rather than the seven terabytes total size of the entire dataset mentioned in background Section 2. The implementation of the backend is built around an Apache server and a Python engine.

3.2 | The data extraction process

The grib_iterator implementation within the ecCodes Python Library is fundamental to the extraction process. While iterating through the GRIB file, customized Python scripts can manipulate the data to convert it into the

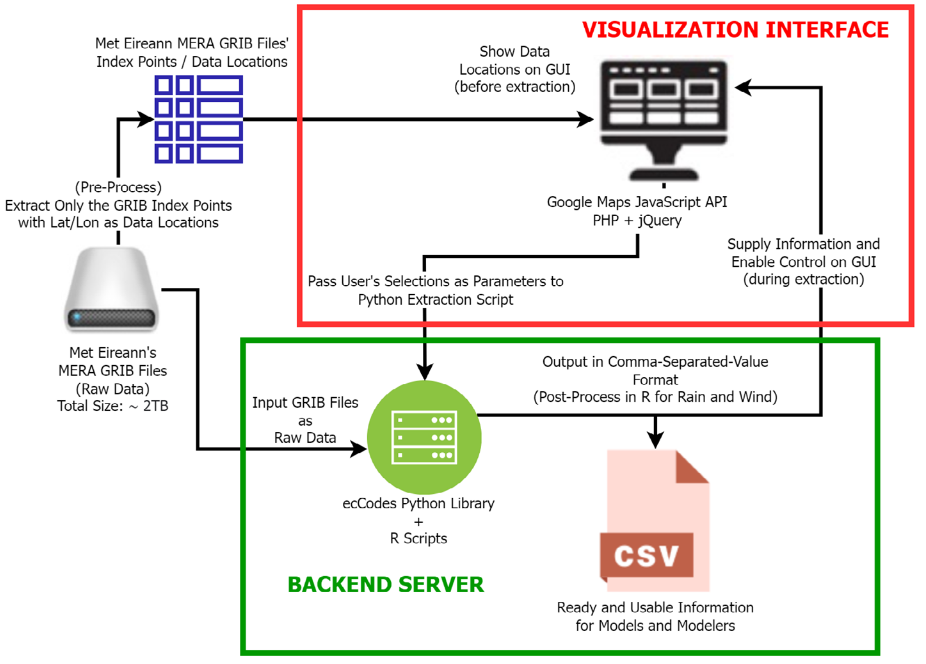


FIGURE 2 MÉRA Data Extractor backend and frontend architecture overview.

required format and output it into a comma-separated value (CSV) file. This format in CSV files has 11 columns, and the common header of the files includes the column names:

'index, lat, lon, value, dataDate, dataTime, validityDate, validityTime, name, shortname, units'.

Here, 'index' value is the identifier of the data point within the MÉRA file. This is a common index for all the files, which means the same index always points to the same lat/lon combination. The 'lat and lon' values are the latitude and longitude coordinates of the data point. The parameter 'value' is the related variable type's value for the data point. While 'dataDate and dataTime' are the date and time for the 'value'. From a temporal perspective, 'validityDate and validityTime' are the date and time that the 'value' is valid until, in case of a forecast. 'name and shortname' are the full name and the abbreviation of the name of the variable type, respectively. Finally, 'units' define in which unit the 'value' of the variable type is given.

3.3 | Data post-process

Simple processing, for example, converting Fahrenheit or Kelvin to Celsius for temperature data, is sometimes sufficient for producing ARD. However, different post-processes are needed for the rain and wind variables to get the data into an analysis-ready form. In MÉRA, rainfall data (total precipitation) are given in a 33-h cycle format called a 33-h forecast, starting with a 9-h warm-up/spin-up period (from 0:00 to 9:00). This is followed by 24 h from 10:00 to 9:00. There are two versions of 0–9 h in each cycle. The first (warm-up/spin-up) 9 h are used to overcome the problem of the starting conditions affecting model outputs. This format is necessary for producing consistent precipitation forecasts. Wind data in MÉRA are given in orthogonal velocity components (U-component and V-component). These variables are wind vectors for deriving wind speed and direction. U-component is the zonal velocity and is the component of the horizontal wind towards the east. V-component is the meridional velocity and is the component of the horizontal wind towards the north. Utilizing R scripting, the rain data are converted into the standard hourly and 24-h time series instead of a 33-h forecast cycle, and the wind data are converted into a regular direction and speed data format.

3.4 | Visualization interface

Google Maps JavaScript API's abilities underpin the GUI for the extraction toolkit. Communication with the backend server is enabled through jQuery and PHP (see Appendix B).

GRIB Index points enable the extraction of the necessary positions (Lat/Lon coordinates). Extracting data points with only soil moisture values resulted in a dataset of 62 k data points covering the entire land area of Ireland, the United Kingdom and the northwest of France. Note that the number of data points in the original MÉRA dataset totals approximately 250 k. Restricting the dataset only to the land area of the island of Ireland resulted in a manageable dataset with a data point count of 14 k approximately (14108).

Visualization of MÉRA data locations (GRIB Index Points) on a map harnesses the Google Maps JavaScript API and its inbuilt support for marker clustering (see Figure 3). Casting the dataset an XML file (see Appendix B) and loading it directly into the browser was successful. Likewise, casting the dataset into a MySQL database (see Appendix B) and loading it directly from the database is a valid approach. However, both approaches take an inordinate amount of time due to the size of the dataset, rendering them effectively unusable. Thus, loading points directly in the page source code in JavaScript was the pragmatic solution adopted, though

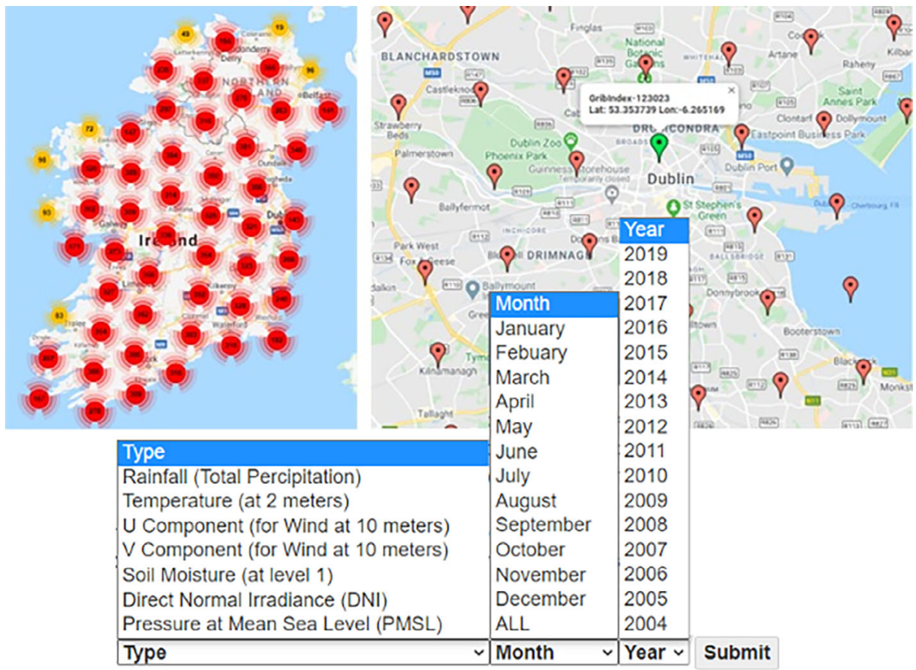
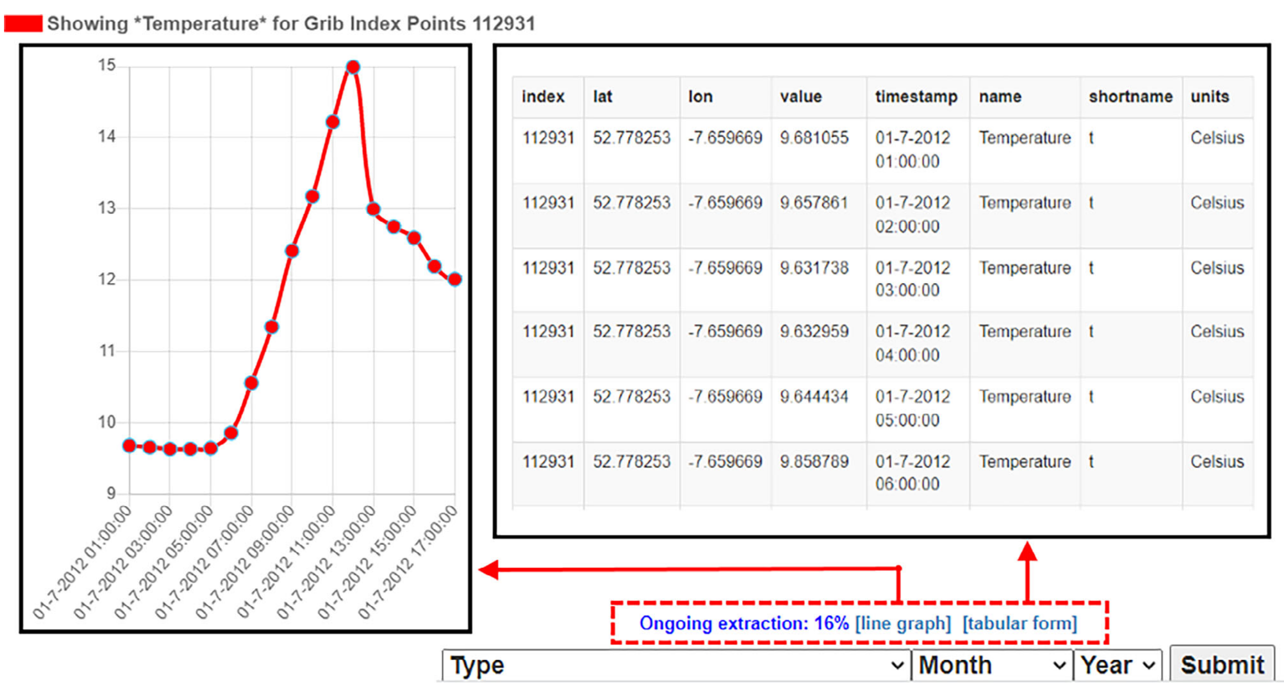


FIGURE 3 Clustered view of the data locations' markers (left) and zoomed-in view of the map with un-clustered markers and also one selected marker in green (right). Sensor type, month and year selection menu (bottom).



index	lat	lon	value	timestamp	name	shortname	units
112931	52.778253	-7.659669	9.681055	01-7-2012 01:00:00	Temperature	t	Celsius
112931	52.778253	-7.659669	9.657861	01-7-2012 02:00:00	Temperature	t	Celsius
112931	52.778253	-7.659669	9.631738	01-7-2012 03:00:00	Temperature	t	Celsius
112931	52.778253	-7.659669	9.632959	01-7-2012 04:00:00	Temperature	t	Celsius
112931	52.778253	-7.659669	9.644434	01-7-2012 05:00:00	Temperature	t	Celsius
112931	52.778253	-7.659669	9.858789	01-7-2012 06:00:00	Temperature	t	Celsius

FIGURE 4 Tabular and line-graph data visualizations and information window during extraction process.

not necessarily the most computationally efficient. In Figure 3, sensor type, month and year selection menu is shown as well. The user will need to select options from this menu and press the Submit button to start the data extraction.

Data visualization is supported in tabular form, using simple jQuery operations (see Appendix B), and in line-graph form, using Charts.js JavaScript Library (see Figure 4).

4 | THE MÉRA EXTRACTION TOOLKIT IN PRACTICE

4.1 | Case study: Catchments

The Environmental Protection Agency of Ireland has overall responsibility for managing Ireland's catchments, of which there are 583 sub-catchments. The coordinates of all catchments are freely available in geoJSON

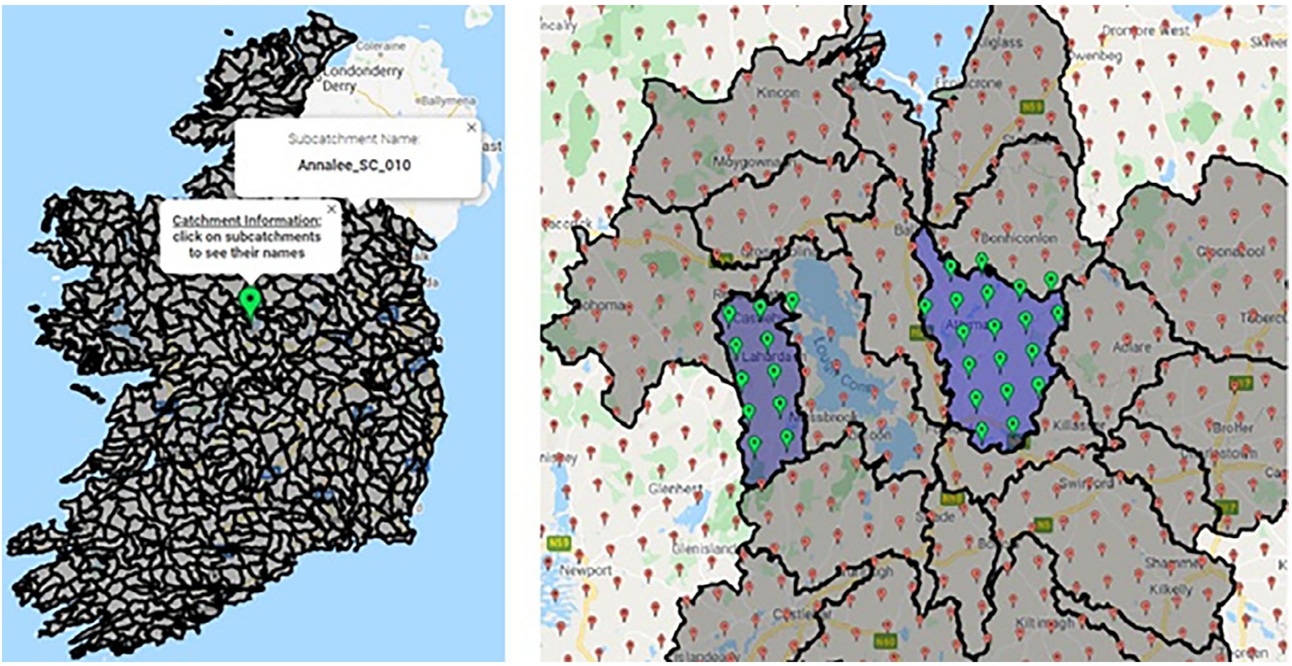


FIGURE 5 Sub-catchments view of Ireland (left) and zoomed-in view of the map with selected sub-catchments in blue and their selected markers in green (right).

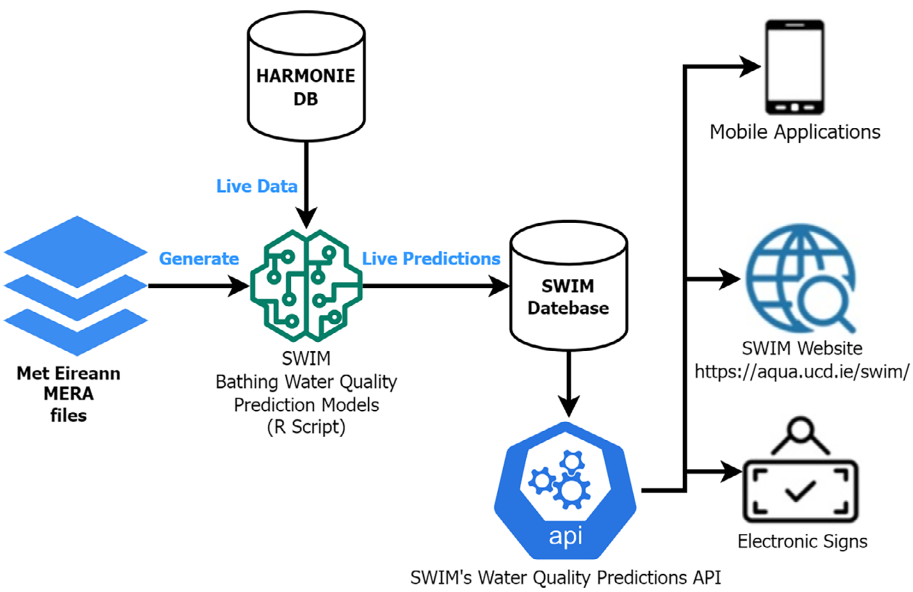


FIGURE 6 Model operations of SWIM software system.

format (see Appendix B) via a public API. The objective is to visualize corresponding MÉRA data points for each catchment as determined by the end user. The polygons representing each sub-catchment are rendered using standard Google Maps functionality on the toolkit GUI. End users then selected the required catchment to initiate the data extraction process (see Figure 5).

4.2 | Case study: Bathing water quality prediction

The SWIM Interreg project (https://aqua.ucd.ie/swim/) harnessed the MÉRA Extraction Tool for developing models for predicting bathing water quality. These models harnessed MÉRA historical data for generation, configuration and fine-tuning. However, real-time data were provided through HARMONIE, a weather forecasting model from Met Eireann. Currently, the SWIM (Hawtree et al., 2020) system (see Figure 6) delivers hourly predictions. Every hour, SWIM automatically runs the models, which are coded in R. These models communicate with the HARMONIE (Bessardon et al., 2020) and query for live weather information. Model predictions are recorded in the SWIM database and made publicly available via SWIM's public API. This API is currently used by SWIM's electronic information signs at beaches, its mobile Apps (both Android and iOS) and its official web page to publicize predictions.

The United Kingdom's Environment Agency (UK EA) compared the outputs of wind (both speed and direction) and sunlight variables from the MÉRA Data Extractor with their own datasets from the UK Met Office. As these two datasets closely matched after comparison, they conflated them together and doubled the size of the data used to build their environmental models, which will give more confidence in the results. This good correlation with the UK Met Office data served as confirmation of the reliability of the Extraction Toolkit.

Northern Ireland Water (NI Water) harnessed a wind dataset constructed using the MÉRA Data Extractor while building a coastal and river model for the Newcastle (Co. Down) bathing beach. During the coastal model build process, NI Water noticed that the currents are significantly affected by wind-induced flows and that applying a static wind source produced mediocre model results. MÉRA data enabled the application of a spatially varying wind field to the coastal model.

Irish Water requested temperature, wind and rainfall data extractions for Waterford harbor. These data were used for developing a Marine Model for Waterford harbor that guided their Wastewater Disinfection Programme. Feedback confirmed the usefulness of the data for validating internal datasets and improving local environmental models.

5 | CONCLUSION

Open source tools, namely Python, R and ECMWF's ecCodes, form the foundation of the MÉRA extraction toolkit. The toolkit remains a work in progress—it is envisaged that it will be adopted and customized in response to individual research needs. Significant advantages accrue for further integrating the post-processing functionality with the extraction process. Rather than two separate stages as currently implemented, future post-processing functionality could be seamlessly integrated into the processing chain in a similar way that temperature is currently converted, for example. In this way, the production of ARD becomes seamless and customizable.

Currently, MÉRA complies with the GRIB1 format. Going forward, it is anticipated that Met Eireann will migrate to GRIB2. In response, the toolkit should be upgraded.

This toolkit can work with any GRIB files with the help of ecCode's Python library. However, the fact that each GRIB file can show significant differences based on the originator of the file must be considered. To enable a level of data extraction from any GRIB file, first, the location information of GRIB file's data points ought to be examined and extracted beforehand in order to adapt the Python scripts enabling them to work on the targeted GRIB dataset. Gathering location information from the files necessitates only basic programming skills to undertake simple editions on toolkit's Python scripts and to use the toolkit via the command line without any user interface.

The MÉRA Data Extraction toolkit is available as open source: https://github.com/jrgarga/MeraDataExtractor.

AUTHOR CONTRIBUTIONS

Levent Görgü: Resources (equal); software (lead); writing – original draft (equal); writing – review and editing (equal). **Daniel Hawtree:** Software (supporting); writing – review and editing (supporting). **Michael J. O'Grady:** Supervision (equal); writing – original draft (equal); writing – review and editing (lead). **Conor Muldoon:** Software (supporting). **Bartholomew Masterson:** Funding acquisition (supporting); supervision (supporting). **Wim G. Meijer:** Funding acquisition (supporting); supervision (supporting). **John J. O'Sullivan:** Funding acquisition (supporting); supervision (supporting). **Gregory M. P. O'Hare:** Funding acquisition (equal); supervision (supporting).

ACKNOWLEDGEMENTS

This research has been realized thanks to the contributions of both the European Union (EU) Acclimatize and SWIM projects. The Acclimatize project was funded by the European Regional Development Fund through the Ireland Wales Cooperation programme and the EU SWIM project was funded by the EU INTERREG VA programme, managed by the Special EU Programmes Body (SEUPB). We gratefully acknowledge the assistance and advice of Eoin Whelan from Met Éireann during this study.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in MÉRA (Met Éireann ReAnalysis) Climate Reanalysis Technical Report at https://data.gov.ie/dataset/m-ra-met-ireann-reanalysis-climate-reanalysis, referenced as MetÉireann, 2022 in the article.

ORCID

Levent Görgü https://orcid.org/0000-0001-7566-3421
Daniel Hawtree https://orcid.org/0000-0002-7649-3839
Michael J. O'Grady https://orcid.org/0000-0002-3024-3666
Conor Muldoon https://orcid.org/0000-0003-1381-2561
Bartholomew Masterson https://orcid.org/0000-0001-9016-1673
Wim G. Meijer https://orcid.org/0000-0003-1302-4338
John J. O'Sullivan https://orcid.org/0000-0002-9524-1274
Gregory M. P. O'Hare https://orcid.org/0000-0002-5124-1686

REFERENCES

Bessardon, G., Clancy, C., Daly, C., Darcy, R., Gleeson, E., Hally, A. et al. (2020) Met Éireann Updates. *ALADIN-HIRLAM Newsletter*, 10, 70–78. http://edepositireland.ie/handle/2262/91446

Caron, J. & Oxelson, J. (2013) BUFR and GRIB file formats considered harmful for data archiving. In: *EGU general assembly conference abstracts, EGU2013–1961*. Göttingen, Germany: Copernicus Publications. https://ui.adsabs.harvard.edu/abs/2013EGUGA.15.1961C/abstract

de Vos, G., Martine, W.H., Bari, D., Behrens, J., Bendoukha, S., Garcia-Marti, I. et al. (2020) Open weather and climate science in the digital era. *Geoscience Communication*, 3(2), 191–201. Available from: https://doi.org/10.5194/gc-3-191-2020

ECMWF. (2022) EcCodes Home - ECMWF. https://confluence.ecmwf.int/display/ECC/ecCodes+Home

Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J.-P., Allenbach, K. et al. (2017) Building an earth observations data cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data*, 1(1–2), 100–117. Available from: https://doi.org/10.1080/20964471.2017.1398903

Gleeson, E., Whelan, E. & Hanley, J. (2017) Met ÉIreann high resolution reanalysis for Ireland. *Advances in Science and Research*, 14(March), 49–61. Available from: https://doi.org/10.5194/asr-14-49-2017

Hales, C., Riley, E.J., Nelson, G.P., Williams, N.J., Ames, D.P. & Enoch Jones, J. (2021) The grids python tool for querying spatiotemporal multidimensional water data. *Water*, 13(15), 2066. Available from: https://doi.org/10.3390/w13152066

Hawtree, D., O'Sullivan, J., O'Hare, G., Görgü, L., Muldoon, C., Meijer, W.G. et al. (2020) The development of a water quality forecasting system for recreational coastal bathing waters in Ireland. In: *EGU general assembly conference abstracts, EGU2020*, p. 15223. Göttingen, Germany: Copernicus Publications. Available from: https://doi.org/10.5194/egusphere-egu2020-15223

Huber, R., DOnofrio, C., Devaraju, A., Klump, J., Loescher, H.W., Kindermann, S. et al. (2021) Integrating data and analysis technologies within leading environmental research infrastructures: challenges and approaches. *Ecological Informatics*, 61(March), 101245. Available from: https://doi.org/10.1016/j.ecoinf.2021.101245

MetÉireann. (2022) MÉRA (Met Éireann ReAnalysis) Climate Reanalysis - Data.gov.ie. https://data.gov.ie/dataset/m-ra-met-ireann-reanalysis-climate-reanalysis

NASA. (2022) NASA GISS: Panoply 5 NetCDF, HDF and GRIB Data Viewer. https://www.giss.nasa.gov/tools/panoply

Raoult, B., Bergeron, C., Alós, A.L., Thépaut, J.-N. & Dee, D. (2017) Climate service develops user-friendly data store. *ECMWF Newsletter*, 151, 22–27. https://www.ecmwf.int/en/newsletter/151/meteorology/climate-service-develops-user-friendly-data-store

Russell, I., Magnusson, L., Janousek, M. & Kertész, S. (2020) Metview's python Interface opens new possibilities. *ECMWF Newsletter*, 162, 36–39. https://www.ecmwf.int/en/newsletter/162/computing/metviews-python-interface-opens-new-possibilities

Whelan, E., Gleeson, E. & Hanley, J. (2018) An evaluation of MÉRA a high-resolution mesoscale regional reanalysis. *Journal of Applied Meteorology and Climatology*, 57(9), 2179–2196. Available from: https://doi.org/10.1175/jamc-d-17-0354.1

Whelan, E., Hanley, J. & Gleeson, E. (2017) "The MÉRA Data Archive". Met Éireann Technical Note No. 65. Met Éireann Technical Note No. 65. http://www.tara.tcd.ie/bitstream/handle/2262/81711/technicalnoteno.65.pdf?sequence=2

Wickham, H. (2014) Tidy Data. *Journal of Statistical Software*, 59 (10), 1–23. Available from: https://doi.org/10.18637/jss.v059.i10

Wright, B. (2014) The challenges of using NetCDF and GRIB for managing forecast data at the met office. Closing the GRIB/NetCDF Gap Workshop. https://www.ecmwf.int/en/elibrary/13709-challenges-using-netcdf-and-grib-managing-forecast-data-met-office

How to cite this article: Görgü, L., Hawtree, D., O'Grady, M. J., Muldoon, C., Masterson, B., Meijer, W. G., O'Sullivan, J. J., & O'Hare, G. M. P. (2023). The MÉRA Data Extraction toolkit. *Meteorological Applications*, *30*(1), e2111. https://doi.org/10.1002/met.2111

Meteorological Applications Science and Technology for Weather and Climate Open Access RMetS

APPENDIX A

Most of the existing tools should be accepted as only visualization tools that lack any kind of data extraction. These visualization tools are supported by different levels of GUI, and most of them lack the option to extract the data in any form. Table A1 below summarizes several tools in terms of some key features, specifically:

- *License*—whether the application is commercial or free.
- *Desktop or Mobile*—applications supporting mobile versions along with their desktop version are mentioned as 'Both' in Table A1.
- *Data Visualization*—whether the application supports any level of data visualization from GRIB files.
- *Data Processing*—refers if the application supports any level of data manipulation on GRIB files. We specifically mentioned (as 'Extract') if there is any *extraction* ability of the application and give details about the *extraction* type supported.
- *Interface*—whether the application supports any level of 'GUI' or if it is solely a command-line based tool or it is supporting 'Both'.

TABLE A1 Features comparison for standalone GRIB tools.

Name	Open source	Desktop or Mobile	Data visualization	Command line or GUI	Data processing
ATMOGRAPH ModelVis, http://atmograph.com/	No	Desktop	Yes	GUI	Open Extract: Image and Video
GrADS, http://cola.gmu.edu/grads/gadoc/grib.html	Yes	Desktop	Yes	Both	Open/Edit and Convert (other raster formats) Extraction: Image Format
Wgrib2, http://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/index.html	Yes	Desktop	No	Command	Open/Edit and Convert (other raster formats) Extraction: GRIB2 files
LuckGrib, https://luckgrib.com/	No	Both	Yes	GUI	No
OpenCPN, https://opencpn.org/	Yes	Desktop	Yes	GUI	NO
CDO, http://code.zmaw.de/projects/cdo	Yes	Desktop	Yes	Command	Open/Edit and Convert (other raster formats) Extraction: Image Format
GribAE, https://www.enviroware.com/portfolio/gribae/	No (Freeware available)	Desktop	Yes	GUI	Open/Edit and Convert (other raster formats) Extraction: GRIB2 files
Weather4D, https://www.weather4d.com/	No	Both	Yes	GUI	No
glgrib, https://github.com/pmarguinaud/glgrib	Yes	Desktop	Yes	Both	No

APPENDIX B

Several computer tools not related to GRIB were utilized during the development and implementation of the MÉRA Data Extraction toolkit. They are listed in Table B1 with brief information.

TABLE B1 Computer tools used in the MÉRA Data Extraction toolkit not related to GRIB.

Name	Description
JSON, https://www.json.org/	JavaScript Object Notation is a standard text-based format for representing structured data based on JavaScript object syntax.
jQuery, https://jquery.com/	jQuery is a small and feature-rich JavaScript Library that makes HTML document traversal, manipulation and event handling much simpler.
PHP, https://www.php.net/	PHP: Hypertext Preprocessor is a fast, flexible, pragmatic and general-purpose scripting language that is especially suited to web development.
XML, https://www.w3.org/XML/	Extensible Markup Language is a file format for storing, transmitting, and reconstructing arbitrary data with a set of rules for encoding documents in a format that is both human-readable and machine-readable.
MySQL, https://www.mysql.com/	MySQL is an open source relational database management system.
GeoJSON, https://geojson.org/	GeoJSON is an open standard format designed for representing simple geographical features, along with their non-spatial attributes.
Chart.js, https://www.chartjs.org/	Chart.js is a free, open source JavaScript Library for data visualization.