# Development of Climate Grid to facilitate the construction of gridded data-sets of observed climate for Pakistan

Timothy D. Mitchell, Dan Hollis and Joseph Daron

August 2021

ARRCC
Asia Regional Resilience to a Changing Climate

Met Office

UKaid
from the British people

# Contents

# Executive Summary

**Gridded data-sets of observed climate are essential for understanding the past and predicting the future. The Met Office has developed Climate Grid, its software for building such data-sets, so that it may be shared with other organisations. Staff at the Pakistan Meteorological Department have been trained to run the software and are building a gridded data-set for Pakistan. This will be used internally in research and operations, and could eventually be supplied externally as a climate service.**

The Pakistan Meteorological Department (PMD) has made it a priority to build a gridded data-set of observed climate for Pakistan for use in a range of applications, including improving their understanding of Pakistan's past climate and the validation and verification of climate models. This priority became apparent in the course of a dialogue between the Met Office and PMD to explore the possible scope for collaboration within a Met Office Partnership programme, Asia Regional Resilience to a Changing Climate (ARRCC), which is funded by the UK's Foreign, Commonwealth and Development Office. In response to this need, the Met Office has shared its Climate Grid software, to enable PMD to create gridded data-sets of observed climate for Pakistan.

This activity required close collaboration between the Met Office and various teams at PMD that maintain and use climate observations. The Climate Grid code has been developed so that it may be used by organisations outside the Met Office. It is now available as a self-contained package that can be installed and run on non-Met Office hardware. It includes the capability to assess the accuracy of any grid that is built, which is essential for data-set development in a new context. Guidance materials have been developed to support its use, including comprehensive documentation, worked examples, sets of sample data, and template code for analysis. Consequently Climate Grid may now be run and supported for use in any country.

A workshop was held in June 2021 for 15 members of PMD staff working in Islamabad and Karachi. The workshop was run remotely due to Covid-related travel restrictions. While this created challenges, the workshop was very successful. In the workshop evaluation, all respondents stated that they were either likely (1/15), very likely (8/15) or guaranteed (6/15) to use what they had learned. PMD staff now have the skills and knowledge to use Climate Grid, with ongoing support being provided by the Met Office.

There is now potential for other organisations – in South Asia and elsewhere – to benefit similarly. The essential requirements for the development of gridded data-sets are people (experienced climatologists), information (station data), and computing resources. Where these are available, the climatologists may be trained to run Climate Grid in order to build gridded data-sets of observed climate for their region.

This activity has improved capacity at PMD to use climate data, and at the Met Office to share its Climate Grid software with other organisations. This capacity may be used to create downstream climate services and products that can better equip a range of stakeholders to understand past climate and adapt to changing climate conditions.

# 1. Introduction

## 1.1 Background and motivation

Reliable data-sets of observed climate are essential for understanding the recent past, for validating models used in weather and climate predictions, and for establishing a baseline for future projections. Observations from weather stations are useful, but only provide measurements for particular locations. Therefore it is common to interpolate these point-based data onto a two-dimensional grid (a 'climate grid'). This provides estimates of climate at regular intervals in space and time. A gridded data-set of observed climate may contain information for various climate variables (e.g. temperature and precipitation) for various time scales, from long-term averages to day-to-day variations.

While some gridded data-sets are produced for the globe, others are built by national meteorological and hydrological services (NHMSs) for their own country. For example, the Met Office has built HadUK-Grid for the UK (Hollis et al., 2019). However, no equivalent data-set has been built specifically for Pakistan. The Pakistan Meteorological Department (PMD) has made it a priority to fill that gap.

This priority became apparent in the course of a dialogue between the Met Office and PMD to explore the possible scope for collaboration. That dialogue formed part of a Met Office Partnership programme, Asia Regional Resilience to a Changing Climate (ARRCC), which is funded by the UK's Foreign, Commonwealth and Development Office. The Met Office saw an opportunity, through ARRCC, to help address that priority by sharing its software for gridding climate observations – 'Climate Grid' – with PMD. Once trained in its use, PMD staff could use Climate Grid to build gridded data-sets of observed climate for Pakistan. These might be first used internally, in both research and operations, and could eventually be supplied externally as a climate service.

This report describes how the Met Office and PMD have collaborated to enhance the capacity of PMD to build gridded data-sets of observed climate for Pakistan. This activity formed workstream 7 of a project within ARRCC, the Climate Analysis for Risk Information and Services in South Asia (CARISSA), in which the Met Office is working in partnership with organisations in South Asia to enhance the uptake and use of climate information to inform climate change adaptation and build climate resilience.

## 1.2 Approach

The activity was co-developed with PMD from the beginning. A formal request was made to the Director General (DG) of PMD to engage in the collaboration. The DG agreed and nominated a focal person to engage with the Met Office team. Through a series of virtual meetings with the PMD team, and via shared documents, a workplan was jointly developed.

Under this workplan, the work has taken place in three phases:

1. During the early part of the activity, the majority of the work was done at the Met Office, in preparing the Climate Grid software for sharing, while PMD prepared their database of station observations.

2. In the middle part of the activity, the Met Office prepared and delivered a training workshop for PMD staff.

3. In the final part of the project the majority of the work is being done at PMD, who are using the software to develop climate grids for Pakistan.

The activities in this workstream combined the expertise available within the Met Office in the National Climate Information Centre (NCIC) and the International Climate Services (ICS) team. Dan Hollis (NCIC), who currently maintains the software, drove forward the required coding. Tim Mitchell (ICS) led the activity, liaised with PMD, provided feedback for the coding work, and wrote the user guide. The workshop was prepared and led by Tim Mitchell, with Dan Hollis as lead trainer on the technical materials. Joseph Daron (ICS) – lead for the CARISSA project – provided additional support.

This international engagement has been conducted entirely virtually, due to travel restrictions in response to the Covid-19 pandemic. This virtual engagement has presented challenges, and face-to-face interactions would have been preferable. However, good working relationships were developed between the Met Office and PMD that can be built on in future engagements.

This remainder of this report discusses the code development that was required to prepare the Climate Grid software for sharing (section 2), a training workshop held for PMD staff (section 3), and the prospects for building gridded data-sets of observed climate for Pakistan (section 4), before providing final conclusions (section 5).

# 2. Code development

## 2.1 Climate Grid software

The Met Office's Climate Grid software is complex and not straightforward to share, for a number of reasons.

Firstly, the Met Office has maintained code to grid observations for many years and the present Climate Grid software is the result of work done by a number of contributors over many years (Perry and Hollis 2005a, 2005b; Perry, Hollis, and Elms 2009; Hollis et al. 2019). The code's present state reflects its evolutionary development.

Secondly, the software has been developed for internal use in the Met Office, and to build observed grids of UK climate. It therefore included a substantial amount of code, parameters, file specifications and data that are specific to its intended use of building grids of UK observed climate.

Thirdly, the code has been migrated across platforms as supporting technologies have evolved. For a long time, the gridding method was coded as a customised version of ESRI ArcView GIS. Subsequently the code was ported to python and run on Linux, which is its current form. The code had not yet been fully migrated, so key components were not available in python.

Fourthly, the experimentation procedures and documentation that had been developed alongside the code were designed for internal use at the Met Office. The

code had not been designed for remote delivery or use. When Climate Grid had been shared previously, it was with internal users or with external users who visited the Met Office in Exeter (UK).

Therefore it was necessary to further develop Climate Grid for use by external and remote partners.

## 2.2 Redesign of Climate Grid

The most important constraint on the redesign of Climate Grid was that it should permit both internal use and external sharing, while avoiding the development of two independent versions of the core software. This was a challenge, as there are elements of the code that are specific to the UK and to the Met Office that are not appropriate for other contexts. Therefore Climate Grid has been restructured into a hybrid form. The foundation code is common to all versions. This supports, for the original Climate Grid, a UK-specific superstructure that will continue to be used within the Met Office. To support the use of Climate Grid in other contexts, a new generic superstructure has been developed.

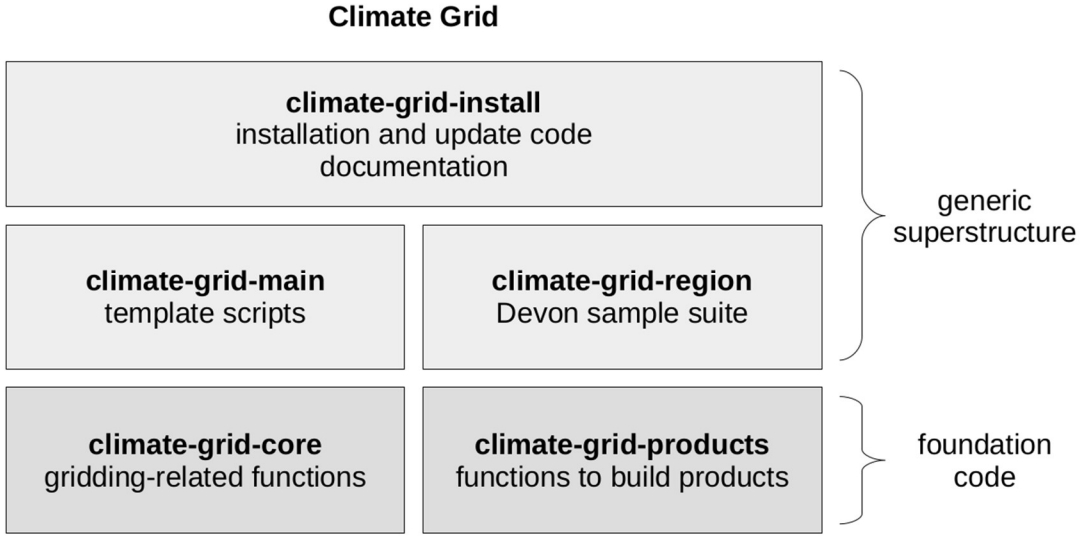


*Figure 1: The structure of Climate Grid shared with partners, as provided through five GitHub repositories.*

The version of Climate Grid that is shared with partners is shown schematically in Figure 1. The foundation code comprises the code to perform core gridding-related operations, such as interpolation onto a regular grid, and the code to build secondary products from the climate grids, such as maps or area averages. The generic superstructure provides a framework for a climatologist to carry out gridding-related tasks using station data for the region of interest. This new superstructure provides:

a) the ability to install Climate Grid locally as a self-contained package, and to run it on non-Met Office hardware;

b) a code architecture equivalent to the UK superstructure, so that Climate Grid may be run in a similar way in non-UK contexts;

c) user support, in the form of comprehensive documentation and worked examples, together with a set of sample data, resources and template scripts for Devon (UK).

## 2.3 Enhancements to Climate Grid

This redesign of Climate Grid was accompanied by a considerable number of enhancements, to make it readily shareable by partners.

Firstly, the code has been simplified by removing those elements of the software that are specific to the development of UK grids or to its use by NCIC. A framework has been provided to permit the user to add their own data, parameters, methods and functions without interfering with the core software. Core functionality has been reintroduced that had been omitted in earlier system migrations and that the new user will require, such as a leave-one-out method. This provides a useful approach to assessing grid accuracy, in which a station is omitted from the grid construction process, an estimate is made of the value at the station location, and the estimate is compared with the observation. The accuracy of the grid is assessed using an error statistic calculated after making this comparison for each station in turn.

Secondly, the code has been configured for installation by partners. The code has been migrated to a set of repositories on GitHub where it can be more easily accessed. A simplified installation procedure has been provided. A conda environment has been specified to ensure that the code runs in the same python environment as when run at the Met Office. A suite of test procedures checks that the core code produces the same outcomes on non-Met Office architectures.

Thirdly, the code documentation has been improved. A comprehensive user guide for the software has been written that can be accessed through a web browser. Template scripts and worked examples have been prepared. An API guide, built using Sphinx, documents the core functions for an advanced user.

# 3. Training workshop

## 3.1 Workshop overview

A virtual workshop was held in four 3-hour sessions on 9–10 and 15–16 June 2021 (group photo in Figure 2). The objective of the workshop was to train PMD staff in the use of Climate Grid, to equip them to develop climate grids for Pakistan. The four sessions developed the participants' capabilities step by step:

1. **Introduction to Climate Grid**: an overview of how Climate Grid works, using the set of sample data, resources and template scripts for Devon (UK).
2. **Configuration of Climate Grid for Pakistan**: an introduction to the configuration and script-editing aspects of using Climate Grid, in which the software was adapted to use sample datasets for Pakistan that PMD had prepared in advance with support from the Met Office.
3. **Product development cycle with Climate Grid**: an introduction to the development and evaluation of the methodological choices required when building gridded data-sets of observed climate, and the application of a procedure to test out various methodological choices that Climate Grid requires.

4. **Further development with Climate Grid**: a concluding session that filled in gaps, reinforced learning, prepared PMD staff for the use of Climate Grid in practice, and included presentations from PMD staff that explored how Climate Grid might be used in Pakistan.

The workshop sessions achieved their stated objectives and positive feedback was received (see evaluation in section 3.5). By the end of session 2, the participants were successfully running Climate Grid tasks using Pakistan data. Between sessions 2 and 3 they succeeded in running new tasks. By the end of the workshop the participants were accustomed to configuring the software and editing scripts to carry out the tasks that they required.


*Figure 2: Workshop group photo, showing some of the participants, 9th June 2021.*

## 3.2 Workshop participants

The trainers were Tim Mitchell and Dan Hollis from the Met Office, with a welcome to the workshop from Joe Daron. From PMD, there was participation from four different divisions totalling 15 individuals:

1. National Weather Forecasting Centre (Islamabad: 3 participants).
2. Research and Development Division (Islamabad: 4 participants).
3. National Drought Monitoring Centre (Islamabad: 2 participants).
4. Climate Data Processing Centre (Karachi: 6 participants).

There were 9 pre-registered participants (7 male and 2 female) with a further 6 participants who joined various sessions (all male). There was a wide spread of roles, including one Director, three Deputy Directors, and specialists both in meteorology and in programming.

## 3.3 Workshop format

The workshop was carefully planned in view of the challenges of delivering remote learning, particularly of technical material in a context that had to surmount language and cultural barriers. The material was broken down into short presentations, interspersed with exercises that progressively built up the competence of the participants. The slides were written in clear, simple English and summarised succinctly in written form what was being explained orally.

The final session responded to the priorities of PMD that had emerged during the course of the workshop by answering questions that had been raised and by illustrating some capabilities of the software that the participants had requested seeing. This responsiveness was possible because of the level of participant contribution, which increased steadily throughout the workshop. In particular, the final session included presentations from the Islamabad and Karachi teams on how they expected to use Climate Grid, which gave the trainers the opportunity to make some useful comments.

## 3.4 Remote training

The provision of technical training remotely was a significant challenge. A brief test session (1 hour) a week in advance of the workshop was invaluable. It gave confidence in the practical arrangements, initiated the working relationships, and permitted a better sense of the level at which to deliver the training.

The videoconferencing software used was Microsoft Teams. Each Met Office participant connected from home. In Pakistan all participants were gathered around two workstations (desktop computers), one in Islamabad and one in Karachi. Microsoft Teams was run on both workstations, and also on some additional laptops. PMD chose this arrangement in the first instance because Climate Grid was installed on the workstations, and so the participants needed to see the workstation screen for the practical exercises.

In addition, this team-based approach to learning brought together people with the different skills that are required to make the best use of Climate Grid: coding, climatological expertise, user engagement. It permitted those with stronger English to assist those whose English is weaker. It encouraged discussion of difficulties and key issues, and the interactions were mutually respectful and inclusive. We would recommend this team-based approach to learning for future workshops.

The approach taken for the practical exercises evolved during the workshop. Initially the trainers displayed instructions on-screen and requested each team in Pakistan to follow these instructions and ask questions orally or using the chat function. This worked, but it was slow and did not allow the trainers to bring their expertise to bear. Therefore a different approach was tried and adopted. The trainers asked the Islamabad team to share their screen, and walked them through the steps necessary to carry out the practical task. The Karachi team were able to see what Islamabad were doing, and follow along. The trainers then checked whether Karachi had also succeeded in the task; if not Karachi shared their screen, enabling the trainers to spot the error and advise how to correct it. For the next exercise the two team roles were reversed.

This approach requires a ‘directive’ style in which the trainer continually asks participants to start/stop sharing their screen. It also requires a team-based approach to learning, as individuals may be less keen to have their mistakes exposed to everyone else. The approach may become harder to implement as the number of participating teams rises, but worked well for this workshop.

## 3.5 Workshop evaluation

A short feedback form was completed by each of the 15 participants. Figure 3 displays the responses when asked about their learning, showing that all participants learned at least as much as they expected, and many (6/15) learned more or much more than expected. Figure 4 shows the participants' responses when asked about their likely use of what they learned. All responded that they were either likely (1/15), very likely (8/15) or guaranteed (6/15) to use what they had learned. The feedback form also included a space for comments. A selection of these comments are reproduced in Table 1. These comments broadly reflect and confirm the responses shown in Figures 3 and 4 and matched the oral feedback from participants during the workshop.

*Table 1: Representative selection of comments made by participants on the feedback forms, in response to the question: "Do you have any other comments on the training?"*

| |
|---|
| **LEARNING EXPERIENCE** |
| **1.** I feel really good by getting this training, very healthy and lively discussions. |
| **2.** I am really thankful to Met Office for conducting very smooth training. |
| **3.** The training was organized in an excellent way and the delivery of knowledge was very well executed. The flow of different modules was designed to take the participants along and explain each and every function in the software. |
| **4.** I appreciate the support extended by Met Office to provide the software and training for gridded dataset generation. The training was delivered in a very professional way keeping all the participants fully engaged throughout the sessions. That was not an easy task having limitations of virtual trainings. |
| **FUTURE USE OF CLIMATE GRID** |
| **5.** This training has opened a lot of opportunities to develop and use observed climate gridded data sets for Pakistan. We are going to formulate a plan based upon the suggestions provided by Met Office to successfully build a reliable gridded data set for desired parameters in the coming days/months. |
| **6.** Due to technical constraints and limitation of methodology, one has to stick with the best available approach. Lots can be done in future using multiple approaches especially in the context of Pakistan due to complex terrain and highly variable weathers. I highly acknowledged the efforts of Met Office for providing the technical assistance to PMD regarding this subject and hopeful to have a strong collaboration ahead. |
| **7.** I am looking forward the continued Met Office support to make this programme a success story. |

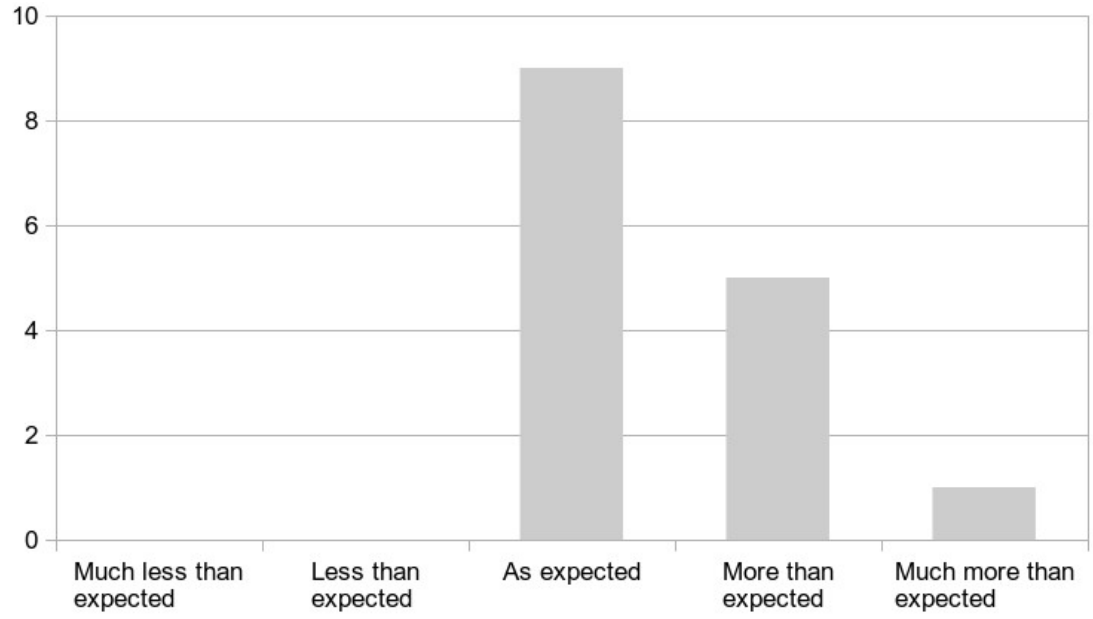


*Figure 3: Responses by the participants to the question: "How much did you learn from the training?"*

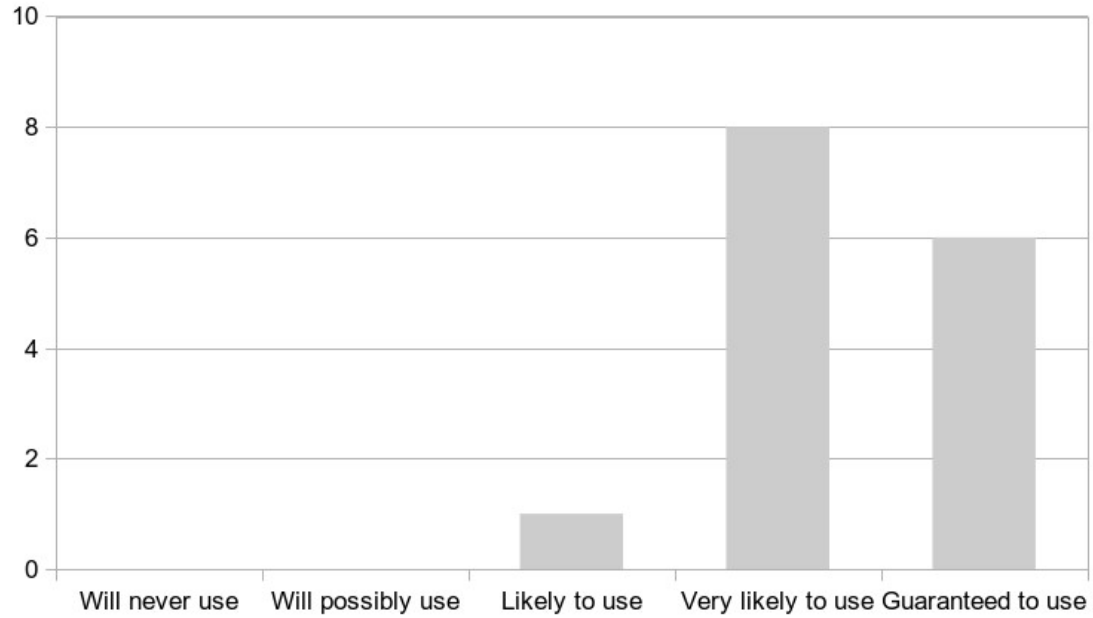


*Figure 4: Responses by the participants to the question: "How likely are you to use the information or knowledge you gained in the training in your work?"*

# 4. Gridded data-sets of observed climate for Pakistan

Currently the team in Karachi processes weather and climate data for ~100 stations and hopes to soon reach ~150 stations. PMD plan to use this station data to build gridded data-sets of observed climate for Pakistan. They intend to experiment with the various methodological choices required by the software to optimise it for various

climate variables. They recognise that gridded values will not be as accurate in regions with few stations and are concerned that in some cases their observation network may be too sparse to allow accurate grids to be built. As part of the ARRCC project, the Met Office team will continue to support PMD remotely as they work through these challenges.

Participants in the workshop highlighted some improvements to Climate Grid that they would find valuable in this task.

1. The leave-one-out functionality (section 2.3) might be expanded to report additional error statistics, report by station, and to provide maps of accuracy, to permit a better understanding of where the observing network is sufficiently dense to permit accurate grid-building, and where additional stations would be beneficial.
2. It would help in assessing grids if the gridding functionality could report the number of stations contributing to each grid point.
3. It would be valuable to port to python the gap-filling functionality that would enable climate normals (standard 30-year averages) to be calculated for many stations without a full record. This would enable station anomalies to be calculated relative to the station normals, rather than to adjacent grid points.

Many of these desirable improvements may be made by the Met Office as part of future projects, and can be shared with PMD as the code is developed. It is also conceivable that PMD staff might themselves make significant improvements to Climate Grid that others may share.

It is likely that the gridded data-sets built in this project will be a first version, and that subsequent versions may improve upon them. It may even be the case that the process of developing version 1 itself provides a stimulus to improve the observational network, as it highlights the impact of a sparse network on the accuracy of gridded data-sets .

PMD envisage such data-sets being used for a wide range of purposes, including:

- assessing climate variability over time and space, including building climate normals and routine monitoring;
- monitoring climate and climate change, including the development of long-term trends in climate change indicators and in extremes for different parts of Pakistan;
- verification of long-term predictions, such as seasonal and sub-seasonal forecasts;
- assessment of climate models, satellite products and other gridded data-sets of observed climate;
- bias correction and statistical downscaling of global climate model outputs;
- as a source for related disciplines, including rainfall runoff modelling, streamflow modelling, hydropower forecasting, basin management, flood control, ecological modelling (species evolution), agricultural modelling (smallholder crop productivity), infrastructure resilience.

# 5. Conclusions

## 5.1 Extension to other partners

This activity has prepared Climate Grid, the Met Office's software for building gridded data-sets of observed climate, for use by partner organisations. It has successfully shared Climate Grid with the Pakistan Meteorological Department (PMD), enabling PMD to build gridded data-sets for Pakistan, rather than relying on third party products. Their knowledge of the climate of Pakistan will be invaluable in using the software to build new gridded datasets. Their relationships with potential users of this information in other organisations may facilitate the provision of new locally-sourced climate services.

This activity might be readily repeated with other organisations, including NMHSs in other countries. The majority of the work required has already been completed in this activity, with the software prepared, training materials developed, and with Met Office staff having gained practical experience in delivering suitable training. The main requirements for replicating this exercise are:

1. a willing and interested partner organisation, with some climatological and coding expertise, access to good quality station observations, and sufficient computing power and storage;
2. resources to enable the Met Office to engage with the partner, prepare sample data, prepare and deliver a training workshop, and provide subsequent support.

## 5.2 Open source connectivity

The other major opportunity for developing Climate Grid lies in connecting it to other software. At present Climate Grid stands alone; it is not directly connected to other software, other than through functionality to connect to the UK database of station observations. Users must therefore prepare their station data in particular file formats, and the gridded outputs are not designed to serve as inputs to other software. In this respect, Climate Grid is much like most other software for storing and manipulating observed climate data.

It is widely agreed that a Climate Data Management System (CDMS) is needed to implement common standards and promote inter-operability (WMO 2014, 2019). There is work ongoing to develop an open-source CDMS, called OpenCDMS. Climate Grid could potentially be used either within this framework or by taking outputs from OpenCDMS as its inputs. This would greatly facilitate the scope for researchers to build gridded datasets from climate observations.

Taking Climate Grid in such a direction would require placing it in the public domain. The work done in this activity to prepare the software for sharing with PMD has already taken significant steps to facilitate that. Making Climate Grid open source might have further benefits, including exposing the software to external scrutiny, expanding the number of people working on it, and extending its use.

# References

Hollis, Dan, Mark McCarthy, Michael Kendon, Tim Legg, and Ian Simpson. 2019. “HadUK-Grid - a New UK Dataset of Gridded Climate Observations.” *Geoscience Data Journal* 6 (2): 151–59. https://doi.org/10.1002/gdj3.78.

Perry, Matthew, Dan Hollis, and Margaret Elms. 2009. “The Generation of Daily Gridded Datasets of Temperature and Rainfall for the UK.” Climate Memorandum 24. Met Office.

Perry, Matthew, and Daniel Hollis. 2005a. “The Development of a New Set of Long-Term Climate Averages for the UK.” *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25 (8): 1023–39. https://doi.org/10.1002/joc.1160.

———. 2005b. “The Generation of Monthly Gridded Datasets for a Range of Climatic Variables over the UK.” *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25 (8): 1041–54. https://doi.org/10.1002/joc.1161.

WMO. 2014. “Climate Data Management System Specifications Version 1.0.” No. 1131. World Meteorological Organization.

———. 2019. “Manual on the High-Quality Global Data Management Framework for Climate.” No. 1238. World Meteorological Organization.