COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS

Grant agreement N°: 313085

Deliverable D5.52

Reanalyses and user needs with respect to Climate Change Services

31 March 2014

Theme: (SPA.2012.1.3-02) GMES Climate Change – Coordination of Earth observation data validation for re-analysis

Project start date (duration): 2013-01-17 (30 months)
Project web site: http://www.coreclimax.eu/
WP5: Intercomparing reanalysis results
Lead partner for deliverable: Finnish Meteorological Institute (FMI)
WP leader: Hilppa Gregow (FMI)
Task 5.1 Summarizing uncertainties related to reanalyses based on a survey on user requirements supported by peer reviewed literature

Authors by institute

Author	Institute	Role(s)
Hanna Mäkelä	FMI	Analysis of the questionnaire and writing
Kirsti Jylhä	FMI	Planning the questionnaire, analysis and writing
Hilppa Gregow	FMI	Coordination, planning the questionnaire, writing and editing this document
Andrea Kaiser-Weiss	DWD	Literature review, planning the questionnaire, writing this document
Andre Obregon	DWD	Literature review, planning the questionnaire, contributing to this document
Frank Kaspar	DWD	Reviewing this document according to user's view
Paul Poli	ECMWF	Collaboration with FMI and DWD
David Tan	ECMWF	Collaboration with FMI and DWD

Acknowledgments for contributing by institute

Contributor	Institute	Role(s)
Satu Kekki	FMI	Implementation of the questionnaire
Ali-Nadir Arslan	FMI	Commenting on survey
Several staff members	FMI	Testing the first versions of the questionnaire
Communication department	FMI	Distributing invitations and emails
	DWD	
	ECMWF	
Karolin Eichler	WMO	Sending the questionnaire to regional offices

Foreword

This report summarizes desired reanalysis characteristics as discussed in the scientific literature and as reported by a dedicated user survey (online from November 2013 until the end of February 2014).

The literature review provides an appraisal of reanalysis characteristics required from the perspective of climate service applications. The literature review points to the value of guidance to current products and the value of user access to feedback statistics. User-tailored post-processing could increase the user base and would enhance the benefits gained from reanalyses. Various future technical improvements for enhanced “climate quality” reanalysis, such as higher resolution and temporal stability could boost the usefulness and applicability of reanalysis for climate services.

The online questionnaire which was answered by over 2500 users of climate information focused on the awareness, skills and requirements of the respondents (mostly based in the scientific community) regarding utility and uncertainties in reanalyses. Implications for delivering future climate services were also assessed. In addition to a summary analysis of all responses, two pairs of contrasting sub-groups have been identified: “ERA-Interim users” versus “not ERA-users” and ”best-informed” users versus ”least-informed” users, in order to assist in identifying more specialized user needs. User survey respondents left almost 750 free comments and suggestions regarding their needs concerning the use of reanalysis data, and 320 free specifications, comments or suggestions regarding future climate services. Analysis of these will be included in Task 5.3. of the project CORE-CLIMAX and will contribute to the assessments of how reanalysis data could bring wider benefit for climate services and research.

Contents

List of abbreviations

CDR[1,2]: A Climate Data Record (CDR) is a series of observations over time that measures variables believed to be associated with climate variation and change. These changes may be small and occur over long time periods (seasonal, interannual, and decadal to centennial) compared to the short-term changes that are monitored for weather forecasting. Thus a CDR is a time series of a climate variable that tries to account for systematic errors and noise in the measurements.

COPERNICUS: The European Earth observation programme Copernicus, previously known as GMES (Global Monitoring for Environment and Security).

DWD: Deutscher Wetterdienst

ECMWF: The European Centre for Medium-Range Weather Forecasts

ECV[1,3]: An Essential Climate Variable (ECV) is a geophysical (atmospheric, terrestrial or oceanic) variable that is associated with climate variation and change as well as the impact of climate change onto Earth.

FMI: Finnish Meteorological Institute

GCOS: Global Climate Observing System

GFCS: Global Framework for Climate Services (see *WMO, 2013*)

GMES: Global Monitoring for Environment and Security

[1] http://ecv-inventory.com/ecv2/terminology/

[2] NRC, 2004: Climate Data Records from Environmental Satellites: Interim Report (http://www.nap.edu/catalog/10944.html)

[3] GCOS-82, WMO/TD 1143: The Second Report on the Adequacy of the Global Observing Systems for Climate in Support of the UNFCCC, April 2003. (available from https://www.wmo.int/pages/prog/gcos/Publications/gcos-82_2AR.pdf).

1. Introduction

The EU FP7 CORE-CLIMAX project coordinates efforts to form long time series of atmospheric, terrestrial and oceanic variables that are essential for monitoring and understanding of climate variations, trends and climate change impacts on our planet. The project consists of six work packages, each with specific objectives. The work package 5 (Intercomparing reanalysis results) will eventually propose a process for comparing outcomes from different reanalysis systems. It will specifically highlight the uncertainties and gaps in reanalysis systems that need to be addressed in order to deliver better climate services. That work is supported and informed by the current document, which summarizes desirable reanalysis characteristics as reported by a dedicated user survey (online from November 2013 until the end of February 2014) and as discussed in the scientific literature.

National, regional and global climate services disseminate climate information to a large variety of end users. The Global Framework for Climate Services (GFCS), established in the World Climate Conference-3 in 2009, aims to strengthen the provision and use of climate predictions, products and information worldwide. Coordinated by the European Commission, the European Earth Observation Programme Copernicus, previously known as GMES (Global Monitoring for Environment and Security), aims to ensure comprehensive and sustainable supply of reliable and up-to-date information on how our planet and its climate are changing.

While many atmospheric science researchers conventionally utilize weather station data or gridded data sets constructed from in-situ observations, reanalysis products are becoming increasingly widely-used. Reanalysis merges past in-situ and satellite-based observations through data assimilation in numerical weather prediction models (e.g., *Bosilovich et al. 2013*). A comprehensive global reanalysis provides a means to obtain time series of geophysical variables anywhere within the three-dimensional space under consideration, extending to remote districts with few measurements; regional reanalysis products with finer spatial resolution are also becoming available. An important advantage in reanalysis products is the fact that time series of different variables, including many Essential Climate Variables (ECVs; GCOS 2010), are consistent with each other. Even variables not directly measured can be provided. However, biases and deficiencies in input observations, numerical models and assimilations techniques all affect the quality of the reanalysis products. Communication of the strengths, limitations and uncertainties of the reanalysis data is therefore crucial and needs to be part of a continuous dialogue between reanalysis developers, observation developers and the research and climate service communities.

This report gives an analysis of the requirements of the reanalysis-user community on reanalysis products and uncertainties in them. The work was carried out by planning and implementing a targeted web portal enquiry and conducting a literature review. In particular, the awareness of the reanalysis-user community on targeted uncertainties in the reanalyses, and the effects of these to deliver climate services were examined. The survey also collected information on requirements from the reanalysis-user community regarding the improvement of reanalyses for better future climate services.

2. Background

This section provides a short background on climate observations in general, long-term Climate Data Records and the WMO Global Framework for Climate Services, highlighting the role of reanalyses.

2.1 Connection between climate services, reanalysis and climate data records

The aim of the Copernicus Climate Service is to provide products of 'climate quality', meeting the first three goals of GCOS [*Uppala et al., 2011*]: (i) monitoring, (ii) detecting and attributing climate change, (iii) assessing impacts and support of adaptation.

Definition of *"Climate Service": "Climate services are climate information prepared and defined to meet user's needs." [WMO, 2011]*

Climate services encompass datasets and statistical analysis, provided together with information and support for users to select the appropriate product for a specific application. This includes tailored information products, scientific studies, expert advice (e.g., on how to factor in the uncertainties) and ongoing support and user engagement. One of the fundamental pillars of a climate service is the availability of climate-quality datasets. These are largely based on observations (including ground-based and satellite observations), but other datasets support and contribute to this pillar: for example reanalysis datasets in which observations are combined, through the methods of data assimilation, with a numerical weather prediction model. Through such methods, reanalysis techniques enable the synthesis of information from a diverse range of observations (various parameters, different platforms and instruments, short- and long-term data sets).

For deriving climate information, long-term stable data records are required. Observations alone are not sufficient, as there is too little coverage, the inhomogeneities are often unknown or too large, and inevitable changes in the observing system over decadal timescales render their application difficult.

Reanalysis systems can provide remedy to a certain extent [*Bengtsson et al., 2007*]. In the 1980s, suggestions for the production of reanalyses for climate studies [*Bengtsson and Shukla, 1988*] and [*Trenberth and Olsen, 1988*] were followed by the NCEP/NCAR reanalysis [*Kalnay et al., 1996*] and ECMWF 15-year reanalysis [Gibson et al., 1997]. Since then, continuous efforts are focused on improving reanalysis for climate applications (see, e.g., *Bengtsson et al. 2007*, and references therein).

Definition of "reanalysis": "Consistent reprocessing of archived weather observations using a modern forecast system". "The aim is to derive a comprehensive description of the observed atmospheric circulation by using as much information as possible." [Dee et al. 2014].

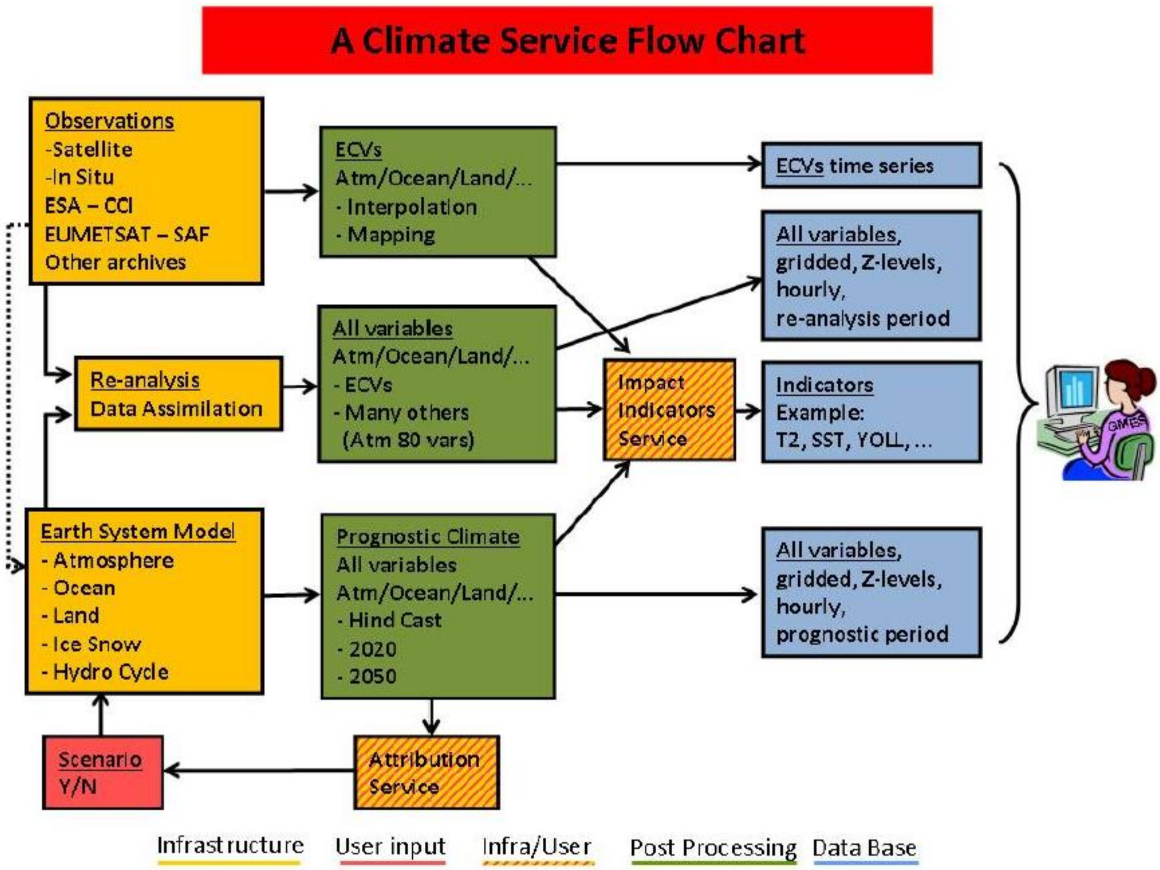


Figure 2.1 [Figure from *Uppala et al, 2011*] shows how reanalyses interfaces to other components in the flow of the envisaged Copernicus Climate

The advantage of the reanalysis method is that it can take into account shifts in a time series (via bias estimation), and is automatically recording the changes of a single observational data record with respect to the synthesis of all assimilated observations. Hence, reanalysis feedback statistics can give information on the climate quality of individual observational data records, with respect to all input observations and model physics. Though a reanalysis can be seen as best effort to arrive at climate information, more developments are required to arrive at what users wish, expect, and need, for instance, reliable information on longer time scale variability, including extreme events, which are often of particular interest to users. Figure 2.2 illustrates the requirements for a climate-quality reanalysis with increasing complexity. The success of different reanalysis data sets to meet these requirements are discussed and disputed in the scientific literature (e.g., *Trenberth 2007, Trenberth 2008, Trenberth et al., 2008, Thorne and Vose, 2010, Dee et al., 2011c, Thorne and Vose, 2011, Bosilovich 2007, van den Hurk, 2012, Trenberth et al., 2013).*

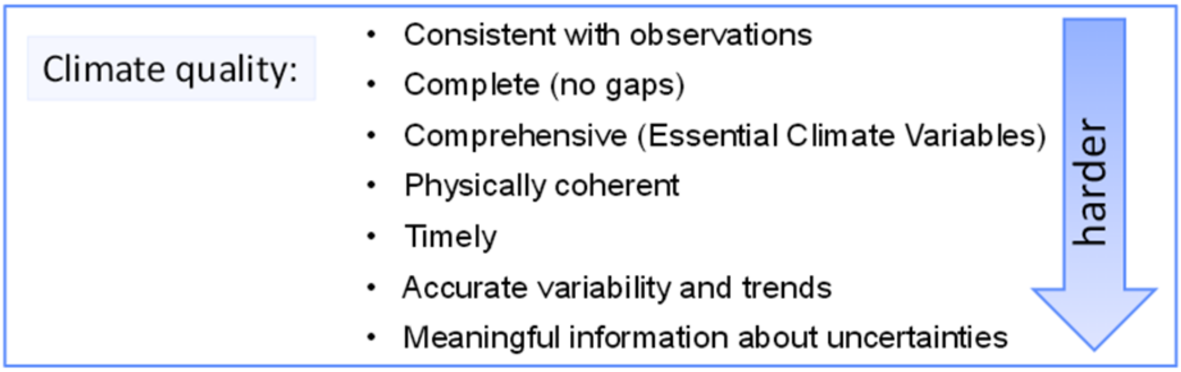


Figure 2.2 [Figure from the presentation of Dick Dee, EMS conference, Reading, 2013] Climate quality stands for a set of user requirements, where the consistency of reanalysis fields with observations, and completeness is easier to achieve than accurate variability and trend characterization. Hardest to achieve is meaningful information about reanalysis uncertainties which can be easily factored in during reanalysis application.

The link between observational records and reanalyses is two-fold:

1. Observations provide input for reanalyses, where the "reprocessing and intercalibration of observed records are critical to improve the quality and consistency of reanalyses" [Bosilovich et al., 2013].
2. Reanalyses provide input for quality control of the observational records, where congruence has to be interpreted in the light of (potential) observation influence on reanalysis.

Table 2.1. State-of-the-art global atmospheric reanalyses as summarized in the chapter 2 of the AR5 IPCC 2013 [Hartmann et al., 2013] in Box2.3, Table1 at p.185:

Institution	Reanalysis	Period	Approximate Resolution at Equator	Reference
Cooperative Institute for Research in Environmental Sciences (CIRES), National Oceanic and Atmospheric Administration (NOAA), USA	20th Century Reanalysis, Vers. 2 (20CR)	1871–2010	320 km	Compo et al. (2011)
National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR), USA	NCEP/NCAR R1 (NNR)	1948–	320 km	Kistler et al. (2001)
European Centre for Medium-Range Weather Forecasts (ECMWF)	ERA-40	1957–2002	125 km	Uppala et al. (2005)
Japan Meteorological Agency (JMA)	JRA-55	1958–	60 km	Ebita et al. (2011)
National Centers for Environmental Prediction (NCEP), US Department of Energy, USA	NCEP/DOE R2	1979–	320 km	Kanamitsu et al. (2002)
Japan Meteorological Agency (JMA)	JRA-25	1979–	190 km	Onogi et al. (2007)
National Aeronautics and Space Administration (NASA), USA	MERRA	1979–	75 km	Rienecker et al. (2011)
European Centre for Medium-Range Weather Forecasts (ECMWF)	ERA-Interim	1979–	80 km	Dee et al. (2011b)
National Centers for Environmental Prediction (NCEP), USA	CFSR	1979–	50 km	Saha et al. (2010)

2.2 The potential of global and regional reanalyses

The first decision any potential user of reanalysis is facing is: "should I pick reanalysis data for my application or not?". There are many users of traditional meteorological or climatological data, who have to decide this without having a profound background education in the field of reanalysis production and evaluation. For them, guidance is needed, specifically, they need to know for which parameters, and at which scales, reanalysis data might be superior to the (possibly scarce, locally influenced, inhomogeneous) observational records which they have at hand. In some occasions and over certain areas no observations are available. Is it then advisable to use reanalysis data or not? Guidance on reanalysis uncertainty would help to answer these questions.

The field of reanalysis has been, and is, rapidly evolving; see, e.g., 4th WCRP International Conference on Reanalysis *[Bosilovich et al., 2013]* with vigorous developments in the characterization of fitness for climate services. Comprehensive overviews on respective reanalysis activities are given by *Dee et al. 2013a*, and *Bosilovich et al., 2013.*
Today, global NWP reanalyses (e.g., as listed in Table 2.1.) capture a dynamically consistent state of the atmosphere, with the potential to provide climate quality global gridded products *[Trenberth et al., 2013].*

Principally, the information content captured in the reanalysis depends on:
1. input observations,
2. assimilation method,
3. the underlying forecast model.

The data assimilation method has to be developed together with the forecast model to make best use of the information content of the input observations. The historic evolution of reanalyses (conveniently described as 1st, 2nd, and 3rd generation) illustrates this combined development of model and assimilation systems, and the progress through more sophisticated data assimilation systems.

The skill increase of the reanalysis has been attributed 15 % to the observing system evolution and 85% to advances in modelling and data assimilation *[Dee et al., 2014*]. It is of paramount importance that the amount and quality of input observations is sufficient to constrain the model to climate quality results. But the ability to develop the required sophistication in forecast models and data assimilation systems remains strongly tied to the availability of high-quality observations.

Aiming for long-term stability in reanalysis datasets, unwanted effects arising from changing the forecast model and/or the assimilation system can be avoided (by fixing these components of the system), but the observing system cannot be changed for the past. This requires consideration of how to deal with the issue of a changing observing system. To estimate the severity of the effect, comparisons have been performed with reanalysis where also the observing system has been fixed (for instance, one such experiment is the 20th century reanalysis only using surface pressure and SST (Compo et al., 2011).

2.3 State-of-the-art global atmospheric reanalyses.

As of the time of writing, global reanalysis datasets are typically provided as gridded datasets, with sub-daily temporal sampling (say from 1 to 6 hours) and resolving horizontal scales on the order of 100 km. An overview summarizing the characteristics of the different products is maintained at reanalysis.org, a recent copy of which is included in Appendix A. The global reanalysis have been further processed to generate datasets at finer temporal and spatial scales by (1) statistical downscaling, (2) dynamical downscaling (i.e., nesting of a higher resolution NWP model) within the global reanalysis which provides the boundary condition, and (3) regional reanalysis.

Regional reanalysis have the potential benefits which come with the higher resolved model used for downscaling the background, plus the benefits from assimilating more data (which are more locally representative, and were thus rejected by the global reanalysis. This way, additional physical processes on the regional scale, e.g., mesoscale processes linked to coastline, topography, frontal systems, convection and interaction with the ground with variations at the range of km to several hundred kilometres are taken into account more explicitly. As the regional reanalysis is nested into the global reanalysis, it inherits the long-term (climate) stability from the latter at the synoptic scales. It is in the range of meso-to-microscale processes where the regional reanalyses can add value.

The value of both global and regional reanalyses depend on the spatio-temporal scale of interest. Although in theory one could increase the output resolution, there is a limitation to ever-increasing resolution because of the limit to which small-scale processes are modelled and to which data are available to constrain the model. Ensembles are one way to address the issue of uncertainty (one source of which is the lack of observational data). The computing effort strongly increases with resolution, and naturally with ensemble size.

2.4 Coupled atmospheric-oceanic reanalyses

In comparison to atmospheric reanalysis, oceanic reanalysis has substantially fewer observations, to constrain processes, which are generally on a smaller spatial scale than in the atmosphere. Therefore, it is much harder to reconstruct the state of the ocean.

Coupled reanalysis take into account both ocean and atmospheric models (also: ice and land surface models) together with respective observations, and their exchange of heat and momentum at the ocean/land surface. The NCEP Climate Forecast system Reanalysis (CFSR) is an example for this (see Saha et al., 2010). In theory, a coupled reanalysis should be better posed in describing the climate system, than decoupled atmospheric and oceanic reanalyses, but significant challenges remain. Some of these challenges relate to potential inconsistencies between the atmosphere- and ocean-components of the coupled models, for example in their representations of physical processes, especially at the ocean-atmosphere interface; and others relate to error-covariance estimation for the coupled system (a critical question for assimilation of sea-surface temperature information). Given the scientific and computational complexity of the coupled reanalysis problem, a pragmatic intermediate development is to use some form of “weakly-coupled” system in which the ocean surface provides boundary conditions for the atmospheric and oceanic reanalyses, and to consider the changes in the boundary condition. Similarly, the land surface can be treated as an atmospheric boundary condition.

2.5 Uncertainty characterization and validation

Evaluation of validity is part of the reanalysis production, see Fig. 2.3. The reanalysis producer generally monitors (i) the quality of fit to observations, (ii) the ability of the assimilation/forecast model to predict observations, and (iii) adjustments to the predictors of systematic observation error by the DA procedure (bias correction) [*Dee et al, 2011b*]).

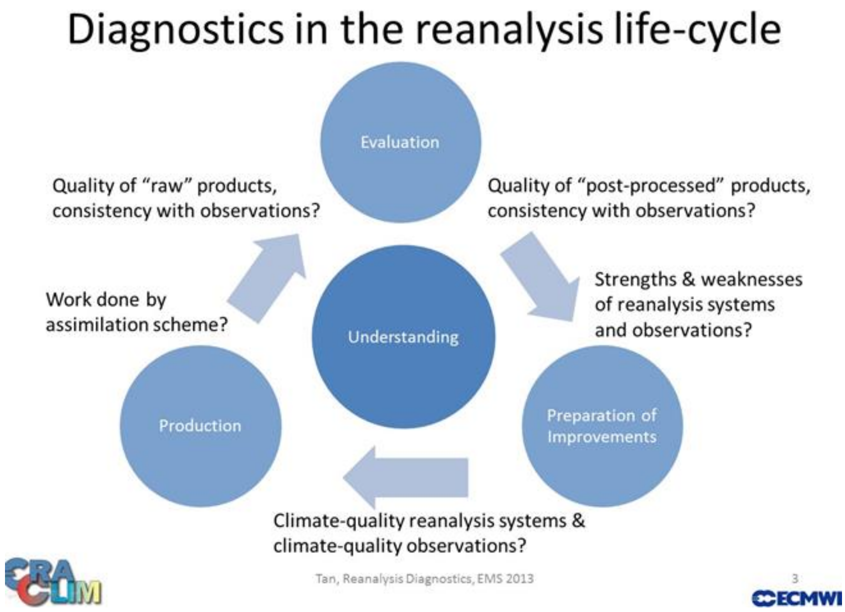


Figure 2.3. [Figure adopted from David Tan's presentation in Core-Climax Workshop in Helsinki 19 March 2014] Diagnostics of reanalysis performance come from production, evaluation and developing improvements.

Reanalysis users also have an important role in assessing the validity of the reanalysis for their chosen applications. The WMO Guide *[WMO, 2011]* recommends: The relative skill of the RA technique in representing the observed features should be assessed before using the data for further climatological studies.

Users need to acquire knowledge about, and contribute to the knowledge base on, the uncertainty of reanalysis in several aspects. Knowledge about uncertainty is sometimes defined with respect to raw data, or with respect to statistics, see Fig. 2.4 for the illustration of the many aspects users need information on.

Uncertainty comes from insufficient observation coverage, insufficient data quality, unknown observation uncertainties and model (and assimilation) deficiencies. Experiences on different aspects of uncertainty and data quality are often reported in the many publications of the users.

Assessing the skill of capturing certain observed physical processes, and comparison with independent data can show the validity of the reanalysis results. Here a so-called "post-processing" has to be applied to the reanalysis output, to ensure that a direct comparison with the specific type of observation is meaningful.

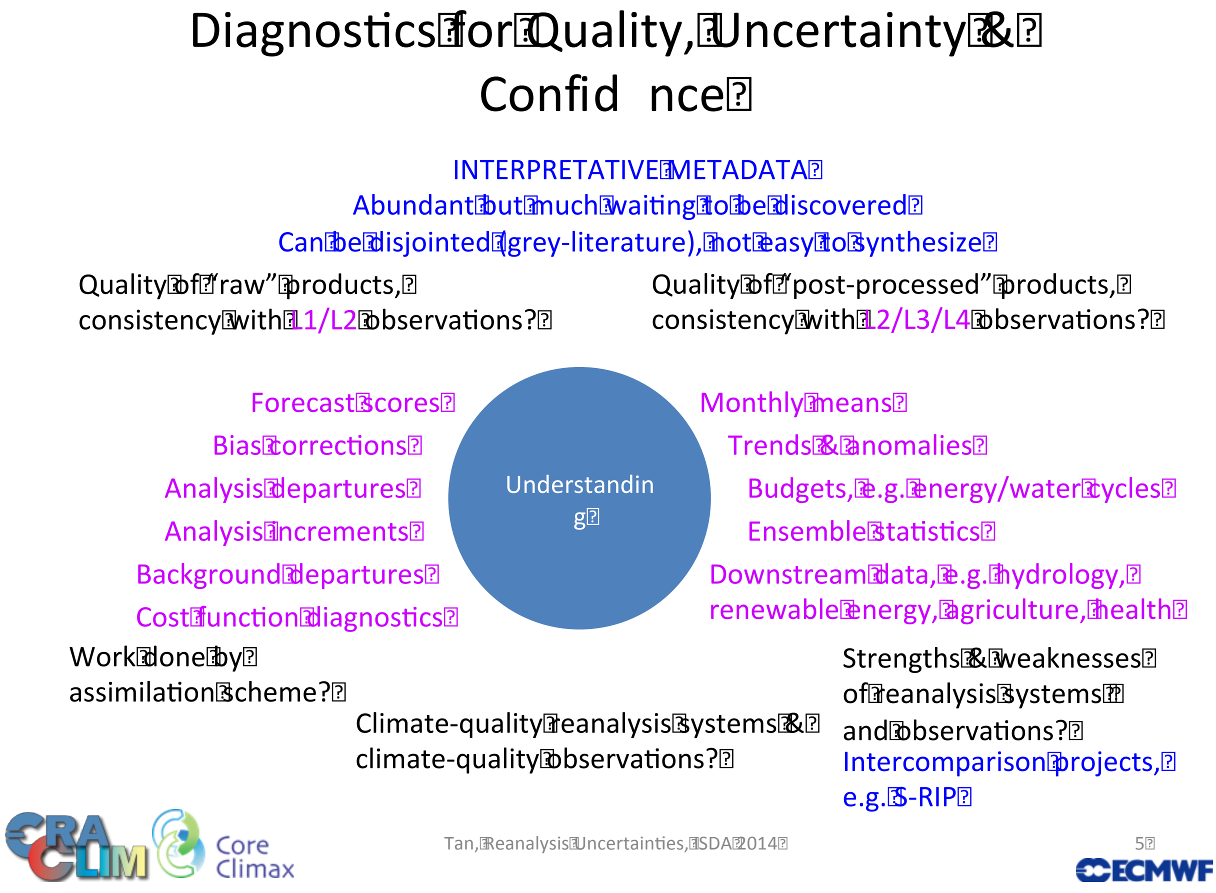


Figure 2.4. [Figure adopted from David Tan's presentation at the Int. Symposium on Data Assimilation, Munich 2014] Different kinds of uncertainty information are required for the users. This figure sorts these according to who is investigating uncertainty for which purpose.

However, independent observations are sparse and sample only a small part of the climate system. Thus, other estimates of uncertainties must be used to complement the picture. Typically, the forecast scores or skill measures employed in numerical weather prediction give a useful estimate of how well reality was captured. The underlying idea is that high-scoring forecasts indicate that the analysis (which the forecast was started from) was closely resembling the true atmospheric state.

Ascertaining the intrinsic uncertainty at smaller scales (where not enough data are available, or where certain physical processes cannot be resolved) is a formidable challenge. Current research is investigating whether this can be addressed with ensembles (of one reanalysis system).

It is worth noting that there is potential to use multiple-reanalysis ensembles to obtain information on model uncertainty, e.g., on trends, or other time-scales of interest, as well as for physical processes (like, e.g., frequency of blocking events).

2.6 Desired feedback statistics

It is potentially of high practical value for the users, to analyse the *feedback statistics*, which are routinely produced during the reanalysis.

Definition of *feedback statistics*: *relating any of observations, free forecasts, analysis results or analysis increments to each other.*

For instance, systematic changes in increments are due to biases in observations or model or both and indicate a deficiency in the system, whereas favourable statistics show that the frequency distribution and time series of observed and reanalysed parameters are matching. Thus, a portal where informed users can access feedback statistics would be desirable.

User-tailored post-processing based on current reanalysis products could provide derived products closer to the user needs. Standard products may include climatologies, global maps of mean and anomalies, comparisons among different reanalysis and independent data records. Downscaling, regional reanalysis and statistical post-processing all could provide derived products closer to the user needs.

2.7 Desired future technical improvements for enhanced "climate quality" reanalysis

Reanalysis will advance together with the numerical models, with successful coupling with the whole earth system, with further developments towards higher resolution, and with methodology developments in data assimilation, as well as with improved generation of climate data records which serve as observation input, and with the ever-increasing data-base resulting from data rescue activities.

"A true 'climate reanalysis' requires extra effort on selecting and preparing input data prior to data assimilation, with preference to observations that have been reprocessed, homogenized and otherwise prepared for climate applications" (citations from *Dee et al., 2014*).

3. Reanalysis user survey: background, objectives, structure and implementation

A previous user survey, conducted in 2004-2005 by ECMWF, focused on the ERA-40 reanalysis alone (Hollingsworth and Pfrang 2005). Most of the 127 respondents gave positive feedback about the quality and accessibility of the ERA-40 data. The survey revealed needs for increased resolution, longer time spans and more regular extensions of the time series to the present. These needs were addressed in the implementations of the ERA-Interim and ERA-20C reanalyses. It also indicated that there were relatively few users of the data in Africa and Latin America, and among researchers on ecosystems and biodiversity.

Since 2005 there has been an increase in the number of reanalyses and reanalysis products (Table 2.1, Appendix A). To collect information on the use of the various reanalysis products, on the knowledge of their limitations and on opinions about climate service activities, a web portal enquiry was conducted within work package 5 of the CORE-CLIMAX project. In this report, we outline the goals, structure and implementation of the questionnaire and present the basic analysis results (task 5.1). In a later phase of the project (task 5.3), a deeper analysis of the questionnaire will contribute to the assessments of how reanalysis data could bring wider benefit for climate services and research.

The reanalysis user and application questionnaire conducted in CORE-CLIMAX served as a tool to examine the use and usability of reanalysis product. The enquiry had the following two initial objectives:

- ✓ To survey the awareness of the user community on uncertainties in the reanalyses, and the effects of these on delivering climate services
- ✓ To collect information on requirements from the user community regarding the improvement of reanalyses for better climate services

The main target groups of the enquiry were

- ✓ existing users of reanalysis datasets
- ✓ users with a possible contribution to future climate services (e.g. governmental institutes, research institutes).

The different parties related to the envisaged use of reanalysis products in climate services, and the connections between them, are shown in Fig. 3.1. The three main components in this interaction scheme are i) the reanalysis, ii) the user, and iii) the climate services with their end-users (Fig. 3.1). The key questions to be addressed with the help of the survey results are indicated with question marks in Fig. 3.1 and can be formulated as follows

1) **How aware are the reanalysis data users of uncertainties and limitations in reanalyses?** What is their knowledge about strengths and weaknesses of the reanalysis products?
2) **What are the requirements of the reanalysis data users for the reanalyses?** What kind of improvement in reanalysis do the users need to deliver (better) climate services, and, further, to provide these to end-users?

3) **How are climate services/climate research affected by uncertainties/gaps in reanalyses?** Are the underlying limitations in reanalysis products influencing the quality and reliability of climate services?

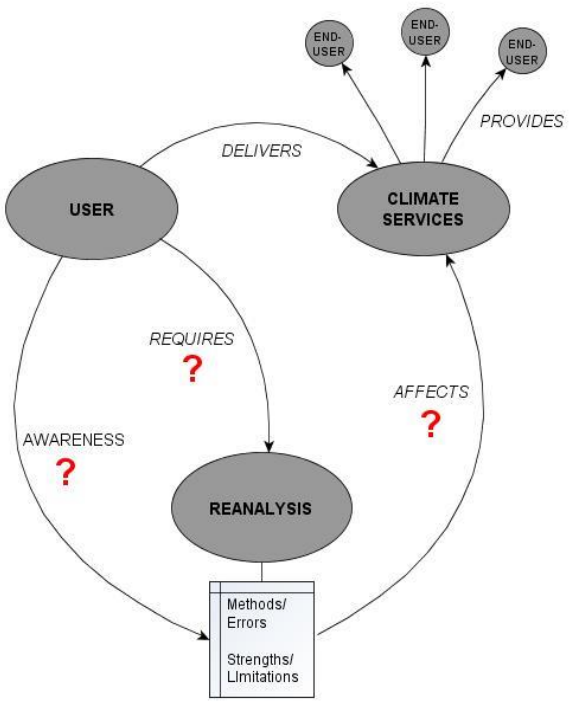


Figure 3.1. Schematic graph illustrating the purpose of the questionnaire.

In the initial planning phase of the survey it was realized that the first two questions could be directly made to the users via the questionnaire. The third question is less straightforward. Because many users of reanalysis data may have vague or contradicting definitions of or desires towards climate services, questions directly related to the functions of climate services were included in the enquiry. Implications of uncertainties and gaps in reanalyses for climate services and research will be discussed in a subsequent synthesis report after a deeper analysis of the outcome of the questionnaire. In that phase, the main issue to be examined is as follows:

- ✓ What are the most important aspects to know about reanalysis for better delivering climate services?

After a total of nine iterations, the final questionnaire comprised 11 main questions (Appendix B). These can be grouped into the following categories (Fig. 3.2):

- ✓ Respondent's background
- ✓ Reanalysis data
- ✓ Applications and methods
- ✓ User awareness and needs
- ✓ Climate services

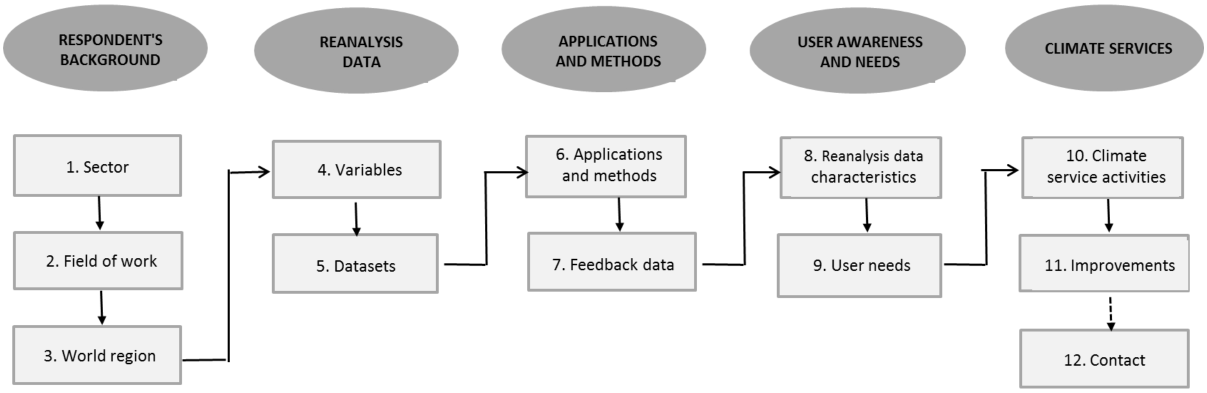


Figure 3.2. Flowchart of the questionnaire.

The web portal enquiry (deliverable D5.51) was linked to the CORE-CLIMAX webpage http://www.coreclimax.eu/ on around 15 November 2013. It was additionally linked to the web portal http://reanalysis.org, hosted by NOAA's Physical Sciences Division, to the DWD web site and to the FMI internal site. It was distributed to regional meteorological offices around the world with the help of the WMO and to universities, research institutes and COPERNICUS-Userforum members in Finland. Two very large emailing operations were conducted by ECMWF. Around 20,500 users of ECMWF reanalyses were contacted twice, on 22 January and 24 February 2014. The first email prompted 1,300 responses in the following few days, while the second email prompted an additional 800 responses in the following few days. By the end of February 2014 total number of respondents reached almost 2600 (2578 respondents in total; see Fig. 3.3). Because it was possible to choose not to answer each question, the number of respondents varied with questions.

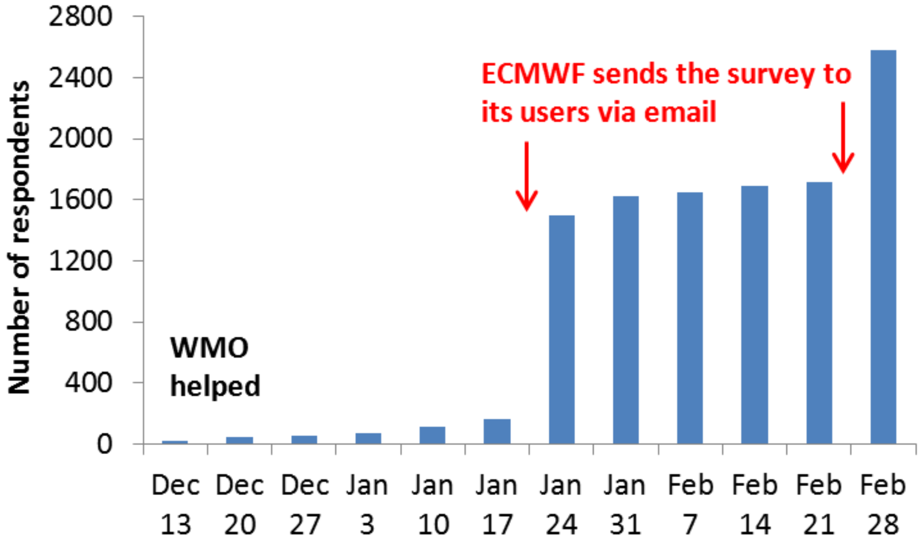


Figure 3.3. The number of the respondents as a function of time.

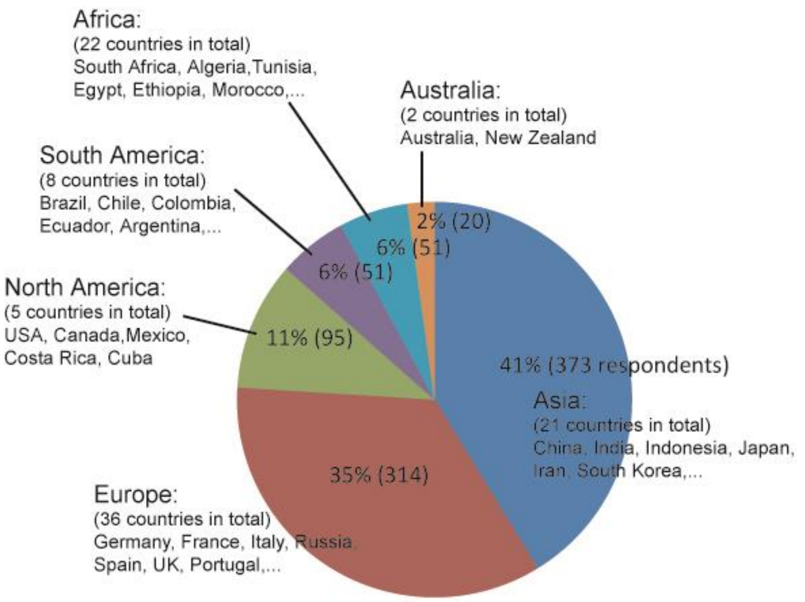
Three main analyses of the responses were carried out in this work. 1) An analysis of all respondents was first conducted (section 4 in this document). Two pairs of sub-groups were then identified: 2) “ERA-Interim users” versus “not ERA-users” and 3) ”best-informed” users versus ”least-informed” users in order to assist in identifying more specialized user needs. Main results of the two pairs of subgroups are shown in sections 5 and 6 in this document, and more detailed information about the background of those respondents can be found in Appendices C and D. Note that statistical significance tests of the differences between the samples were not included to this report.

Note: In general, propositions related to User awareness and needs -table (question 8 in Appendix B) were formulated in the way that by agreeing with the proposition the respondents indicated that they are satisfied with the issue in question. However, four propositions were originally formulated the other way around; meaning that agreement with the proposition indicated that the respondent is not satisfied with the issue in question. These propositions were (originally): “The data policy is too strict”, “The file sizes are too large to work with”, “The data tend to become too late for my needs”, and Time-varying biases make the data too instable for my needs”. Afterwards, when analyzing the results, these propositions were revised to be in line with the other statements. It is important to keep this in mind when interpreting the results of the User awareness and needs -table, because it is possible that some respondents have accidently picked the wrong option for these propositions.

4. Survey analysis: all respondents

4.1. Background of the respondents

Out of almost 2600 respondents 904 (35 %) left their contact information. The countries with most respondents were China (17% out of the 904 respondents), India (9%), the United States of America (8%), and Germany, France and Italy (4% each). About three quarters of the replies came from Asia or Europe (Fig. 4.1). In total, 94 different countries were mentioned and they covered all the continents.



TOP10 countries	
P.R. China	152
India	81
USA	71
Germany	34
France, Italy	32
Indonesia	30
Japan	29
Spain	27
Brazil	26
UK	21

Figure 4.1. Location of the respondents who gave contact information (904/2578) divided into continents. The list of countries with the highest number of respondents is also shown. Note that all countries for all the continents are not necessarily shown.

The sector of work, its field or subject and its regional focus were given by almost all respondents (Fig. 4.2). The public research & development sector was slightly more common than the education sector. The portion of the private sector was 7 %. The top three fields or subjects of work were climate, weather and oceans or seas. The number of alternatives was almost 30 (Appendix B), and on average 2.6 choices were made per respondent. “Energy” was selected by about 10 % of the respondents and “Fresh water resources and management” or “Ecosystems, biodiversity” by slightly fewer. In their work, the respondents most often focused on the whole globe, Europe or Asia. A respondent typically mentioned 1.8 regions. There was a large variety of regions, including polar regions and oceans where the traditional in-situ observation network is sparse.

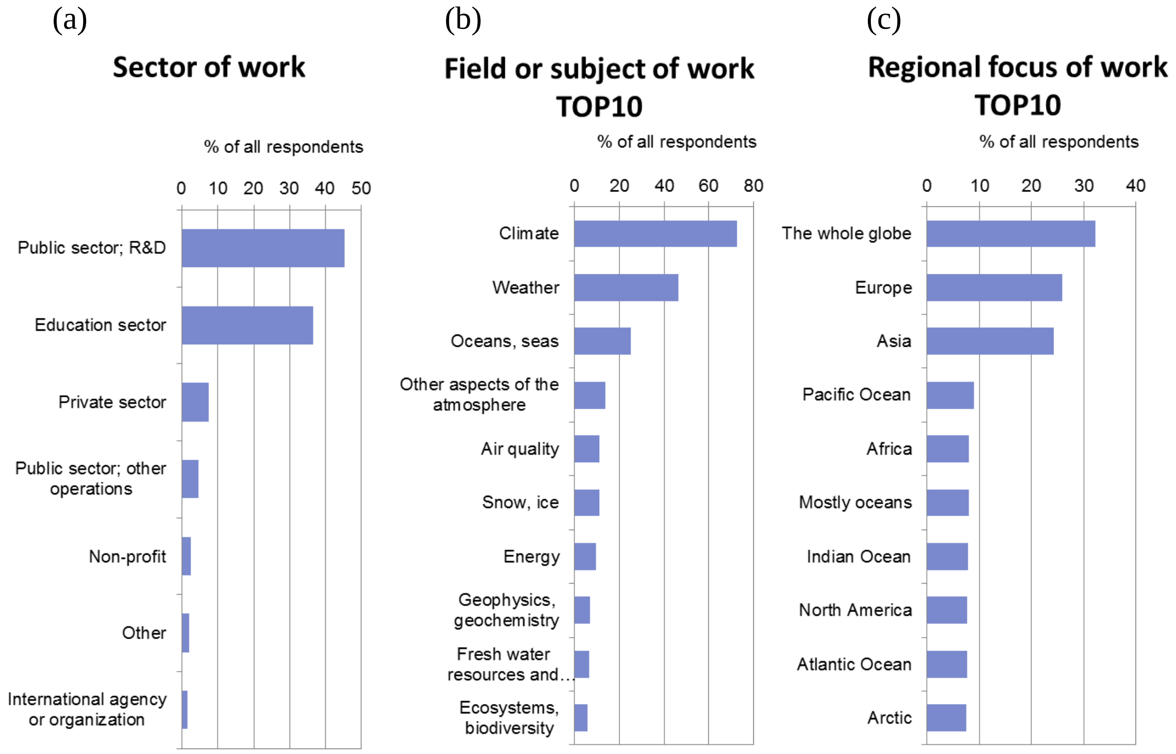
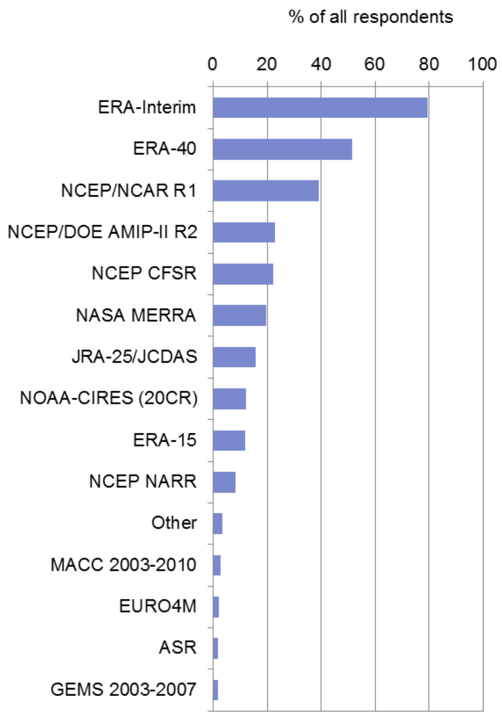


Figure 4.2. (a) Sectors of work of the respondents (2567). (b) Ten most popular topics that best describe the respondents' (2573) field or subject of work. It was asked to choose all that apply (see Appendix for the alternatives). (c) Ten most common regions of the world on which the respondents (2574) focused. It was asked to choose all that apply (see Appendix B for all the alternatives).

4.2 Reanalysis data and Essential Climate Variables

Among those respondents, the most widely-used reanalysis data sets were the Global ECMWF Interim Reanalysis (ERA-Interim), the Global ECMWF 40 year Reanalysis (ERA-40) and the Global NCEP/NCAR Reanalysis I (R1) (1948 to present) (Fig. 4.3a). On average 2.9 choices were made per respondent. Proportion of each dataset from all given responses is shown in Fig. 4.3b. The share of ECMWF's atmospheric reanalyses (ERA) was almost half (48%) of all given responses. Compared to the atmospheric reanalyses, the oceanic counterparts were less widely used by the respondents of this enquiry; on average 0.7 choices per respondent were recorded. The most widely-used oceanic reanalysis data sets were NCEP Global Ocean Data Assimilation System (GODAS), and ECMWF Ocean Reanalyses ORA S4 and ORA S3. Five to six per cent of the respondents indicated that they have used these oceanic reanalysis (Fig. 4.4a). Their share from all given responses was about one quarter (26%) (Fig. 4.4b).

A list of 50 Essential Climate Variables (ECVs) was given, and the respondents were asked to indicate the variables that they work with, stating whether they use reanalysis data or not. The number of respondents was 2569, and on average 13.7 choices were made per respondent. However, the distribution of the number of studied ECVs by the respondents revealed that the most common number of studied ECVs was five (Fig. 4.5). The distribution is broad; use of 2 to 13 different ECVs was common. There were also 107 respondents who indicated that they work with all given ECVs. Table 4.2 lists the ECVs for which it is more common to use reanalysis data than not to use (but to work with anyhow). The opposite is shown in Table 4.3, i.e. those variables are indicated that they are usually examined based on data sources other than reanalysis. Based on the tables, the respondents use reanalysis data especially for studying atmospheric upper air and surface temperature, pressure, wind speed and direction, and water vapour. By contrast, the reanalysis data is not commonly used to work with oceanic and terrestrial variables (with the exception of sea surface temperature).



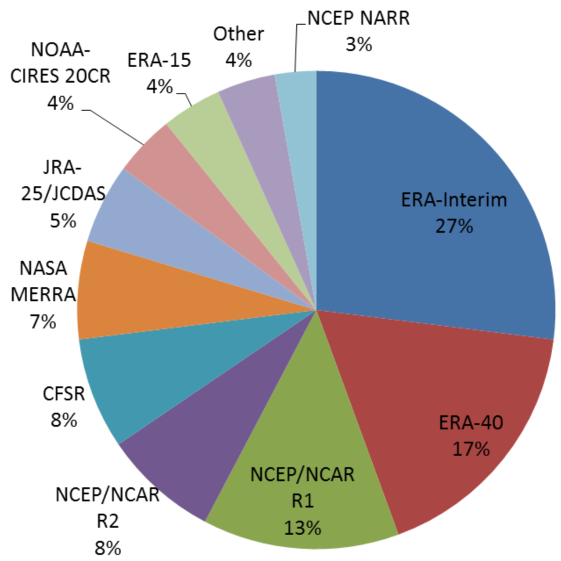


Figure 4.3. Atmospheric reanalysis data sets that the respondents (2502) most often used. (a) Percentages of all respondents (2502). (b) Percentages of all responses (7597) given by respondents. It was asked to choose all that apply.

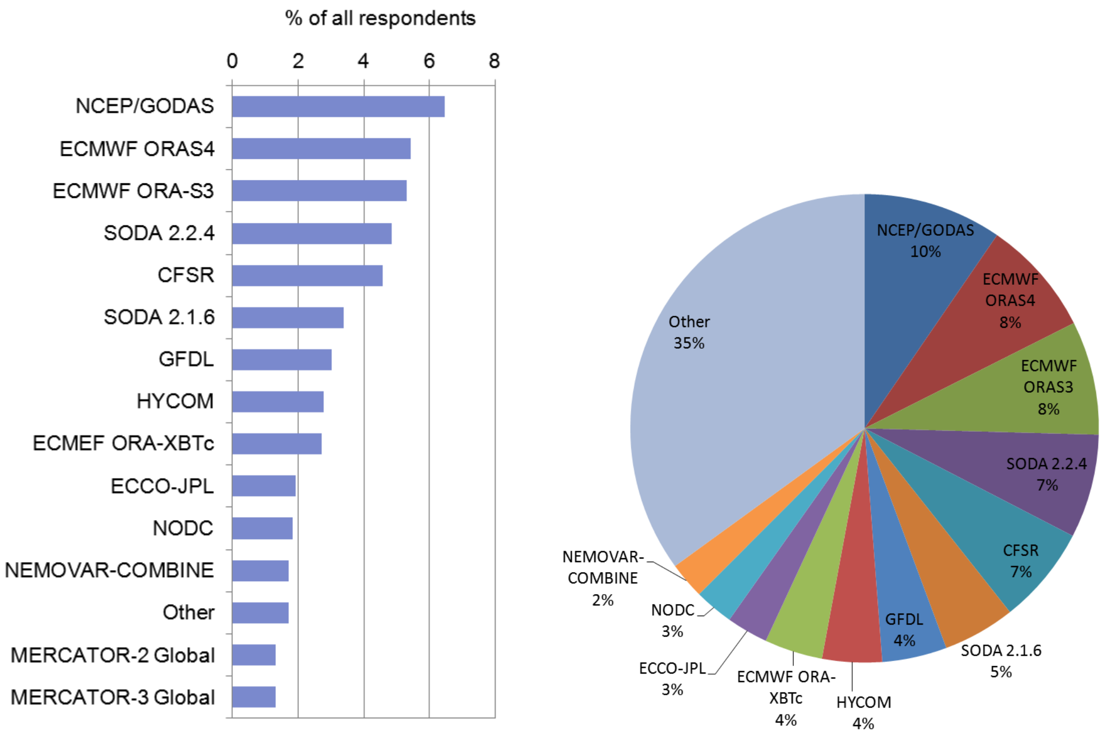


Figure 4.4. Oceanic reanalysis data sets that the respondents (2502) most often used. (a) Percentages of all respondents (2502). (b) Percentages of all responses (1746) given by respondents. The sector "other" consists of over 40 oceanic reanalysis data sets, whose share from all given votes is less than 2% each (see Appendix B for all the alternatives). It was asked to choose all that apply.

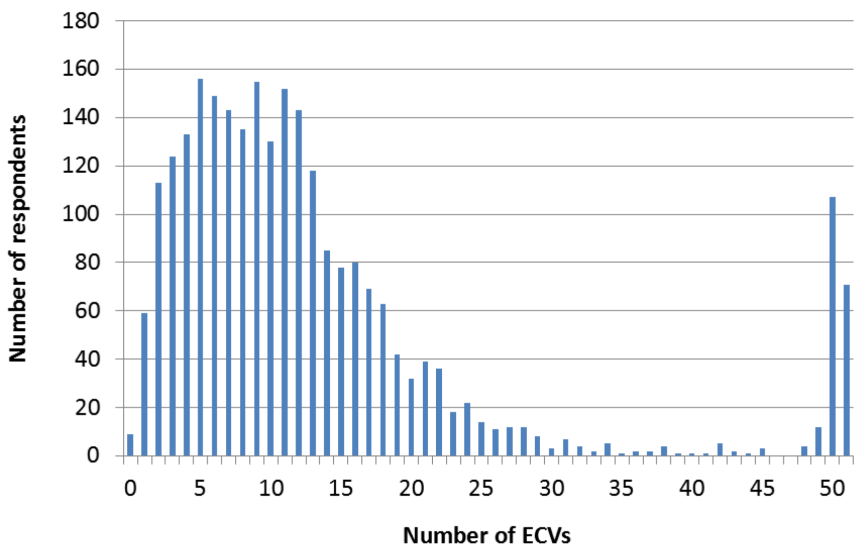


Figure 4.5. Distribution of the number of studied ECVs by the respondents.

Table 4.2. The ECVs for which the number of respondents USING reanalysis products IS LARGER than those NOT USING reanalysis products to work with. When the ratio A/B > 1, the respondents use more reanalyses than other data sets to work with this variable.

A = I work with this variable and use reanalyses for this;

B = I work with this variable but do not use reanalyses for this:

Essential climate Variable	A = use reanalyses to work with this	B = do not use reanalyses to work with this	A/B
AU: Temperature	1494	116	12.9
AS: Pressure	1792	150	11.9
AU: Wind speed and direction	1483	135	11.0
AS: Wind speed and direction	1894	177	10.7
AS: Air temperature	1910	202	9.5
AU: Water vapour	1094	161	6.8
AS: Water vapour	1265	208	6.1
AS: Precipitation	1453	415	3.5
OS: Sea-surface temperature	1153	343	3.4
AS: Surface radiation budget	914	299	3.1
AU: Earth radiation budget (including solar irradiance)	633	258	2.5
AU: Cloud properties	658	289	2.3
OS: Sea level	437	288	1.5
OSS: Temperature	419	282	1.5
OS: Sea state (waves)	302	215	1.4
T: Soil moisture	378	270	1.4
T: Snow cover	369	266	1.4
OS: Sea ice	381	277	1.4
T: Albedo	360	283	1.3
OS: Sea-surface salinity	331	298	1.1

AS: Atmospheric surface; AU: Atmospheric upper air; AC: Atmospheric composition; OS: Oceanic surface; OSS: Oceanic sub-surface; T: Terrestrial

Table 4.3. The Essential Climate Variables for which the number of respondents NOT USING reanalysis products IS LARGER than those USING reanalysis products to work with. When the ratio A/B < 1, the respondents use more other data sets than reanalyses to work with this variable.

A = I work with this variable and use reanalyses for this

B = I work with this variable but do not use reanalyses for this

	A = use reanalysis to work with this	B = do not use reanalysis to work with this	A/B
OS: Surface current	269	279	1.0
T: Land cover (including vegetation type)	338	355	1.0
OSS: Salinity	267	289	0.9
AC: Ozone and precursors	250	275	0.9
OSS: Current	252	287	0.9
T: Other variable, please specify	73	88	0.8
AC: Aerosols and precursors	221	278	0.8
T: Ice sheets	153	202	0.8
T: Leaf area index (LAI)	211	286	0.7
T: Glaciers and ice caps	154	211	0.7
T: Groundwater	155	223	0.7
AC: Carbon dioxide	190	278	0.7
T: River discharge	222	337	0.7
T: Water use	145	233	0.6
T: Lakes	139	232	0.6
T: Fraction of absorbed photosynthetically active radiation (FAPAR)	130	218	0.6
AC: Methane	136	245	0.6
AC: Other long-lived greenhouse gases	130	243	0.5
T: Above-ground biomass	111	211	0.5
T: Permafrost	97	186	0.5
OS: Phytoplankton	101	218	0.5
OS: Ocean colour	106	234	0.5
OSS: Nutrients	86	213	0.4
T: Soil carbon	82	209	0.4
OSS: Oxygen	78	210	0.4
T: Fire disturbance	76	209	0.4
OS: Ocean acidity	67	189	0.4
OSS: Tracers	72	204	0.4
OS: Carbon dioxide partial pressure	70	201	0.3
OSS: Carbon dioxide partial pressure	66	193	0.3
OSS: Ocean acidity	60	183	0.3

AS: Atmospheric surface; AU: Atmospheric upper air; AC: Atmospheric composition;
OS: Oceanic surface; OSS: Oceanic sub-surface; T: Terrestrial

4.3 Applications and methods

The respondents (2502) had an average 13.2 applications to work with or methods to employ. They were asked to indicate their source of data, stating whether they use 1) reanalysis products; 2) weather station, radiosonde or other in-situ observations; 3) satellite-based remote sensing data; and/or 4) weather radar based remote sensing data. It was found that in almost all given options of applications or methods (total of 28), reanalysis was the most frequently used data source. The only exception was “Production of in-situ-based data sets” where in-situ station data had the largest share (in-situ station data 223 answers, followed by reanalysis data 146 answers). It is nonetheless noteworthy that those involved in producing in-situ-based datasets make use of reanalysis data - more details on how would be valuable and could usefully be investigated in follow-up work.

Table 4.4. The TOP-five applications or methods for each type of the data source. It was asked to choose all that apply.

Applications or methods	number of respondents	% of all respondents
TOP5 Reanalysis data		
Studies of atmospheric dynamics	1335	52
Climate modelling	1051	41
Atmospheric modelling	1038	40
Time series analyses	933	36
Studies of atmospheric physics	815	32
TOP5 In-situ station data		
Time series analyses	701	27
Studies of atmospheric dynamics	591	23
Meteorological case studies	528	21
Atmospheric modelling	472	18
Studies of atmospheric physics	447	17
TOP5 In-situ gridded data		
Time series analyses	414	16
Climate modelling	398	15
Studies of atmospheric dynamics	393	15
Atmospheric modelling	375	15
Evaluation of climate models	311	12
TOP5 Satellite data		
Studies of atmospheric dynamics	589	23
Studies of atmospheric physics	444	17
Atmospheric modelling	434	17
Time series analyses	410	16
Meteorological case studies	394	15
TOP5 Weather radar data		
Meteorological case studies	229	9
Studies of atmospheric dynamics	213	8
Studies of atmospheric physics	163	6
Atmospheric modelling	163	6
Short-term forecasting	148	6

Studies of atmospheric dynamics, climate modelling and atmospheric modelling were the three most common applications to which reanalyses data were needed (Table 4.4).

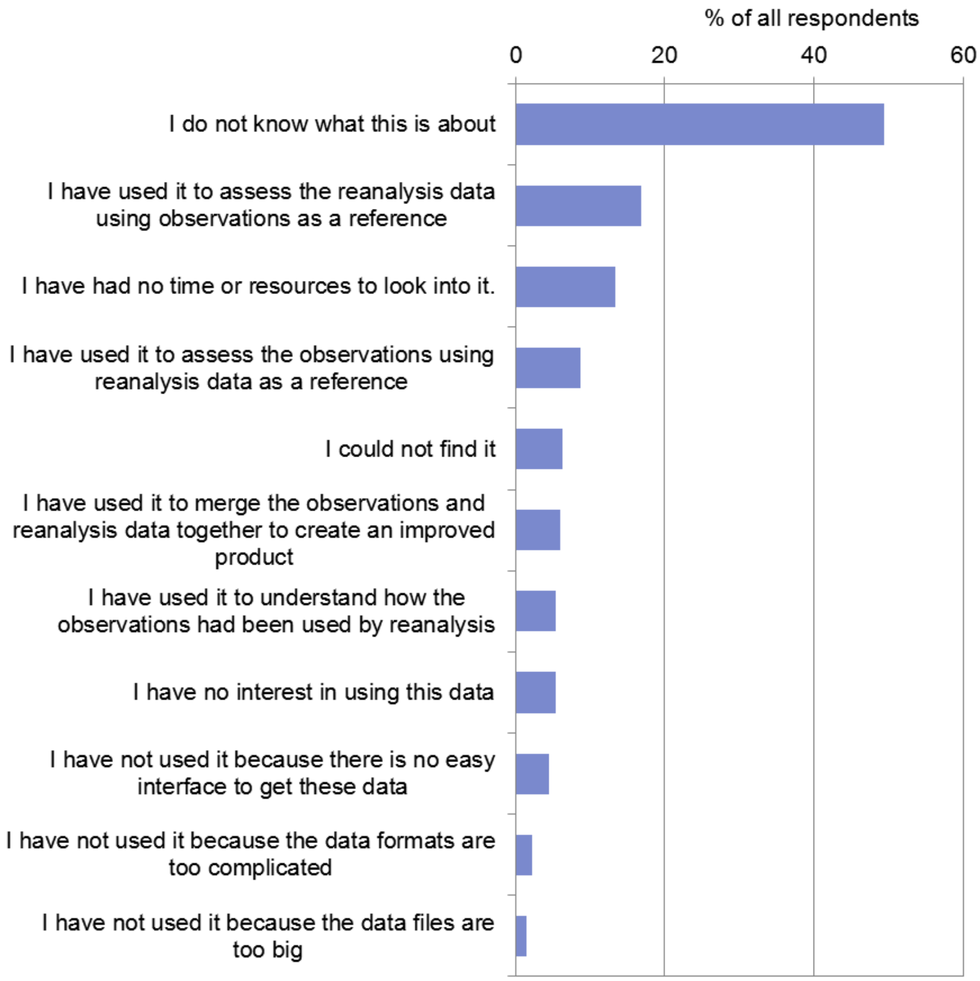


Figure 4.6. Responses (2473) to the use of reanalysis input observations and feedback data. It was asked to choose all that apply.

Almost half of the respondents seemed not to know what is meant by reanalysis input observations and feedback data (Fig. 4.6). The concepts and terminology are relatively well-known within the data assimilation and reanalysis-production communities - they relate to quantitative measures of the agreement between model fields (either before or after assimilation) and observations. Analysis of feedback information is a routine part of assessing the quality of a reanalysis and, to a lesser but growing extent, of assessing the quality of observations. (This is consistent with the survey responses which show that those having experience of feedback data more frequently assessed reanalysis data using observations as references rather than the other way round. The average number of choices per respondent was 1.2.)

Results of feedback analysis are communicated through a number of public channels including the scientific literature and presentations at scientific/technical meetings, as well as through bi-lateral communication with data providers. The feedback data for a particular observational dataset are themselves often shared with the data providers to facilitate improvements in the observational data via reprocessing. The reanalysis-production community has for some time realized the potential for downstream users to benefit from access to the feedback data, and are taking steps to make the feedback easily accessible. The survey suggests that uptake is limited primarily by user awareness, which could be remedied by further capacity-building in this area.

4.4 User awareness and needs

Regarding the characteristics of reanalysis data, a number of propositions were given with which the respondent could agree or disagree (1= fully agree, 5= fully disagree) or skip to the next proposition. In Fig. 4.7, the responses have been divided into three categories (fully or somewhat agree, in-between or did not answer, and fully or somewhat disagree). Also the given propositions are divided into four sub-groups: Data access and availability, Data resolution in space and time, Data quality and representation of uncertainties, and Background information.

The red bars in Fig. 4.7 give the share of those respondents who fully or somewhat disagreed with the proposition. These can be taken as an indication of the areas where there is need for improvements with the actual issue. As for the yellow bars in Fig 4.7, they show the amount of respondents being “in between” or who did not answer at all. In addition to including the respondents who think that the issue in question is unimportant for them, the yellow bars most probably include also those who are not aware of those issues. So these can be taken as an indication of the areas where there is need for increasing the awareness of the user community.

Propositions with large proportion (approximately 50% of the respondents) of those in-between were concerning i) awareness of the differences between the true and nominal spatial and temporal resolution, ii) awareness of the biases and uncertainties in the reanalysis data, iii) awareness of the observation input to reanalysis products, and iv) available training material on the web.

On average, the respondents were most satisfied with propositions “The data is easy to access”, “The time period covers my interest”, and “The data are consistent between the variables”. In the other end of the list, i.e., proposition that were least agreed, were “The uncertainties are well characterized”, “Plentiful training material is available on the web”, and “The observation input to reanalysis are clearly explained”.

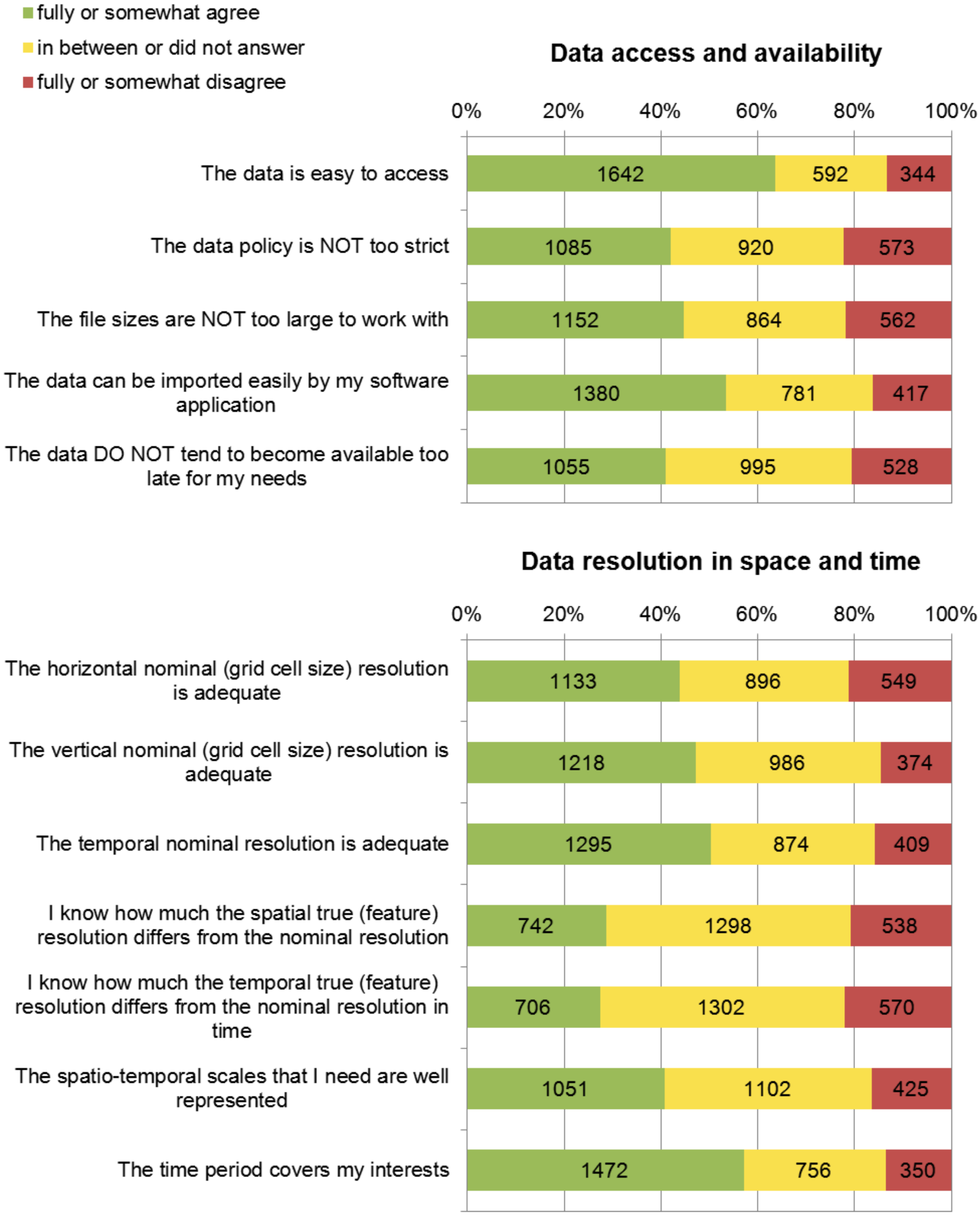


Figure 4.7. Responses (2486) to propositions about the characteristics of reanalysis data (continues).

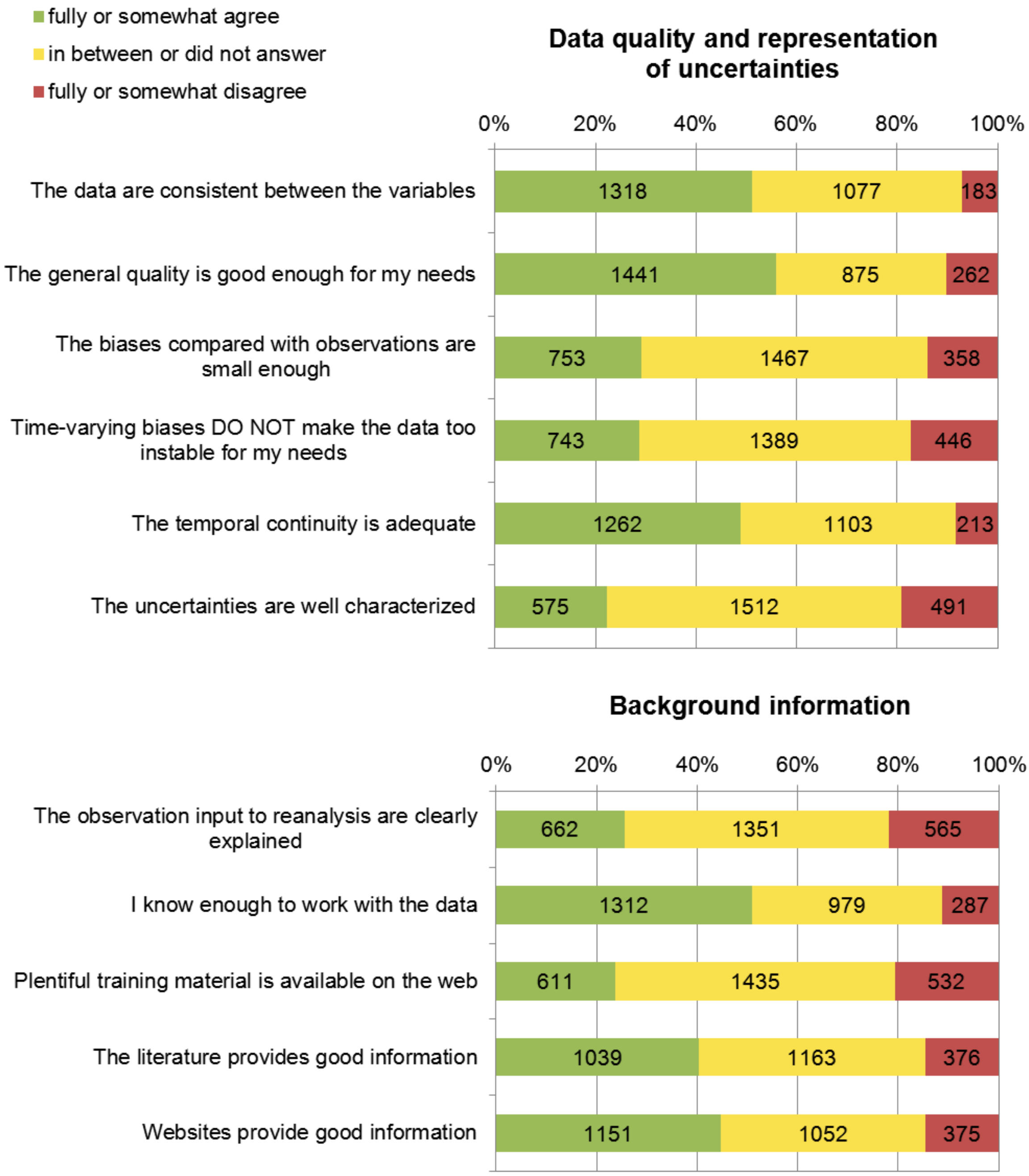


Figure 4.7. (continues) Responses (2486) to propositions about the characteristics of reanalysis data.

By calculating the average of all the responses for each proposition, and then sorting the propositions according to their score, a preliminary order of agreement can be assessed. The lower the score (average of the responses) the more the respondents agreed with the proposition:

1.0: Fully agree

2.1	The data is easy to access
2.2	The time period covers my interests
2.2	The data are consistent between the variables
2.3	The general quality is good enough for my needs
2.3	The temporal continuity is adequate
2.3	The data can be imported easily by my software application
2.4	I know enough to work with the data
2.4	The temporal nominal resolution is adequate
2.4	The vertical nominal (grid cell size) resolution is adequate
2.5	Websites provide good information
2.6	The spatio-temporal scales that I need are well represented
2.6	The file sizes are NOT too large to work with
2.6	The literature provides good information
2.6	The horizontal nominal (grid cell size) resolution is adequate
2.6	The data DO NOT tend to become available too late for my needs
2.7	The data policy is NOT too strict
2.8	The biases compared with observations are small enough
2.8	Time-varying biases DO NOT make the data too instable for my needs
2.9	For the climate variables I need, I know how much their spatial true (feature) resolution differs from the nominal resolution
2.9	I know how much the temporal true (feature) resolution differs from the nominal resolution in time
2.9	The observation input to reanalysis are clearly explained
2.9	Plentiful training material is available on the web
2.9	The uncertainties are well characterized

5.0: Fully disagree

Plentiful free comments and suggestions (from almost 750 respondents) related to user needs were given; analysis of them is underway.

4.5 Climate services

Opinions of the respondents were asked regarding tasks or activities (to be) included in future climate services (Fig. 4.8). The respondent could agree or disagree (1= fully agree, 5= fully disagree) or skip to the next proposition. In Fig. 4.8, the responses have been divided into three categories (fully or somewhat agree, in between or did not answer, and fully or somewhat disagree).

Regarding future climate services and reanalyses, free specifications, comments or suggestions from 320 respondents were given. Analysis of these is under way.

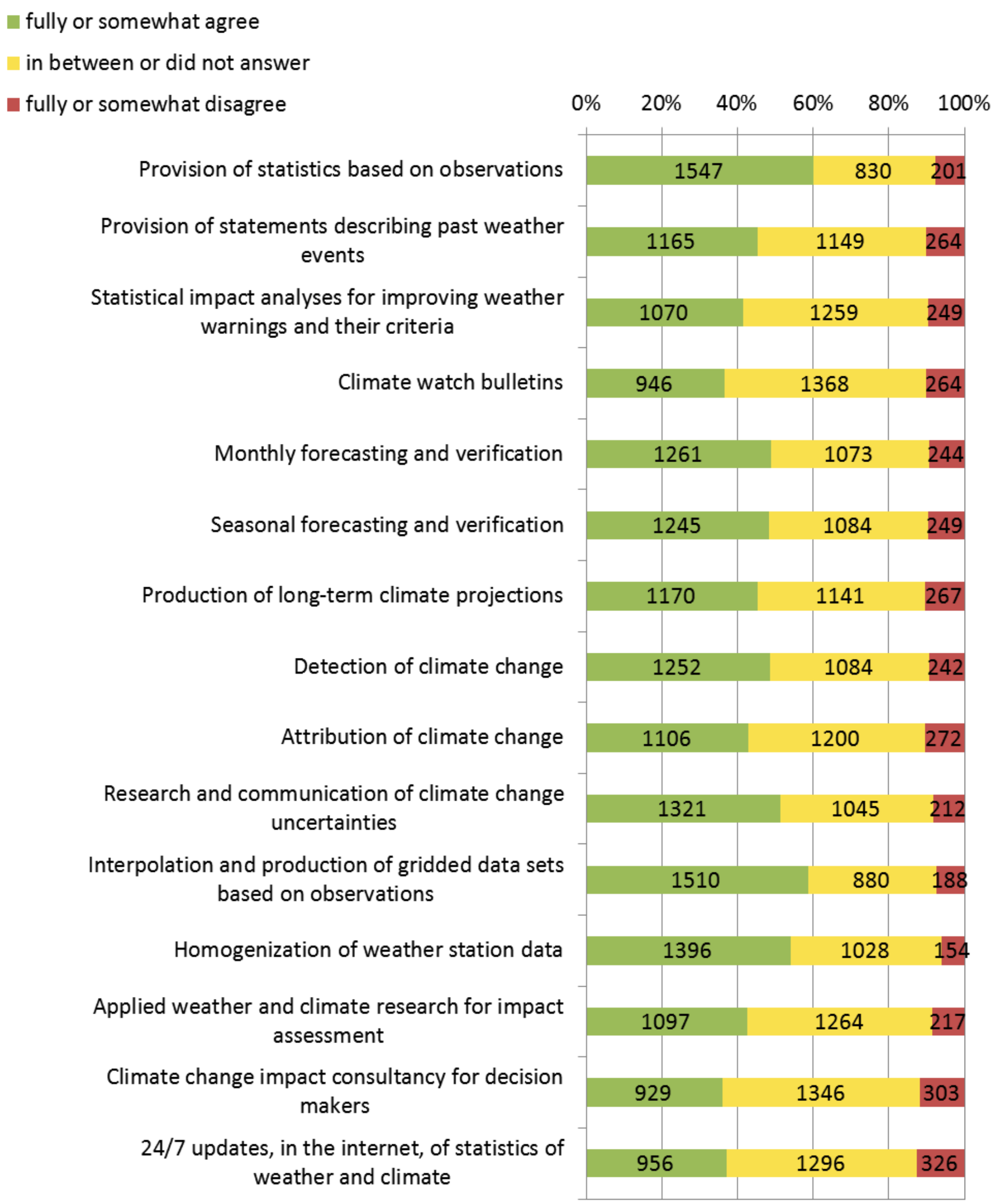


Figure 4.8. Responses (2192) to propositions about tasks or activities in future climate services.

The priority/relevance of the suggested future climate service tasks and activities can be assessed by calculating the average of all the responses for each proposition, and then sorting the propositions according to their score. The lower the score (average of the responses) the more wanted the task or activity would be:

1.0: Fully agree

1.8	Interpolation and production of gridded data sets based on observations
1.9	Provision of statistics based on observations
1.9	Homogenization of weather station data
2.0	Research and communication of climate change uncertainties
2.1	Monthly forecasting and verification
2.1	Detection of climate change
2.2	Seasonal forecasting and verification
2.2	Production of long-term climate projections
2.2	Applied weather and climate research for impact assessment
2.2	Provision of statements describing past weather events
2.3	Attribution of climate change
2.3	Statistical impact analyses for improving weather warnings and their criteria
2.4	Climate watch bulletins
2.4	24/7 updates, in the internet, of statistics of weather and climate
2.4	Climate change impact consultancy for decision makers

5.0: Fully disagree

5. Survey analysis: "ERA-Interim users" and "not ERA-users"

Almost 80% of the respondents indicated that they have used ECMWF's ERA-Interim reanalysis. To learn from the possible differences between ERA-Interim-users and not ERA-Interim-users, a division into following subgroups was made:

- ✓ **"ERA-Interim users":** uses only ECMWF's ERA-Interim atmospheric reanalysis (391 respondents, 15% of all respondents).
- ✓ **"not ERA-users":** doesn't use any atmospheric reanalysis by ECMWF (241 respondents, 9% of all respondents)

Main findings and results concerning especially about the user awareness and needs, as well as future climate service activities are shown in the following pages. More detailed information and illustrations about the two sub-groups can be found in Appendix C.

5.1 Background of the respondents

Out of 391 ERA-Interim users 102 left their contact info, for the not ERA-users the number was 85 (out of 241). Half of respondents using only ERA-Interim were from Europe (50%), the second largest group was Asia (35%). For not ERA-users, Europe and Asia were almost equal size (38% and 33%, respectively). The share of the other continents (North and South America, Africa and Australia) was larger for the not ERA-users (almost 30%) than for the ERA-Interim users (15%). Those who are not using ERA reanalysis came more often from private sector and public sector's other operations (than R&D) compared to the ERA-Interim users, whose sector of work background was very similar to that of all respondents.

The not ERA-users were more interested to work with those subjects that were not that popular among all respondents and ERA-Interim users, such as "energy", "fresh water resources and management", "agriculture", "forests", and "ecosystems and biodiversity". Those who were not using any ERA reanalysis, were focusing more to Europe and less to the whole globe when compared with ERA-Interim users or all respondents. The proportion of those who were focusing on Africa or North America was also emphasized in not ERA-users.

Figures and tables related to the background of the respondents can be found in Appendix C: Figs. C.1-C.4 and Tables C.1-C.2.

5.2 Reanalysis data and Essential Climate Variables

Both ERA-Interim users and not ERA-users work on average with ten different Essential Climate Variables (9.9 for ERA-Interim users, 10.4 for not ERA-users). The distributions of the number of studied ECVs by the respondents are shown in Fig. 5.1 and 5.2. For the not ERA-users the distribution is somewhat wider than for the ERA-Interim users. The most common number of ECVs to work with was four (4) for the ERA-Interim users and five (5) for the not ERA-users.

The use of ECVs among the ERA-Interim and the not ERA-users are illustrated in Figs. C.5-C.8 shown in Appendix C. Those working with only ERA-Interim reanalysis use more reanalysis data to work with atmospheric surface and upper air variables compared to the not ERA-users, whereas for atmospheric composition, oceanic sub-surface and terrestrial variables the situation is vice versa. Similar to all respondents, the ECVs that the ERA-Interim users and the not ERA-users work with most using reanalysis data are related to atmospheric surface and upper air.

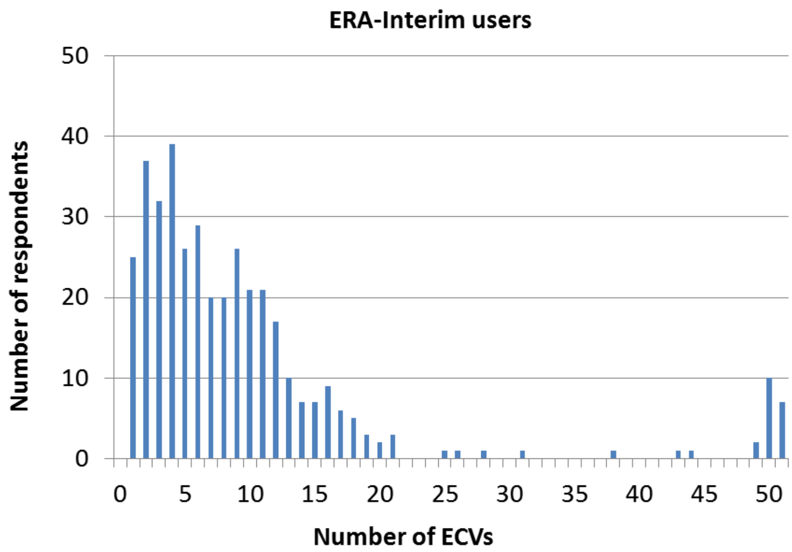


Figure 5.1. Distribution of the number of studied ECVs by the ERA-Interim users (391).

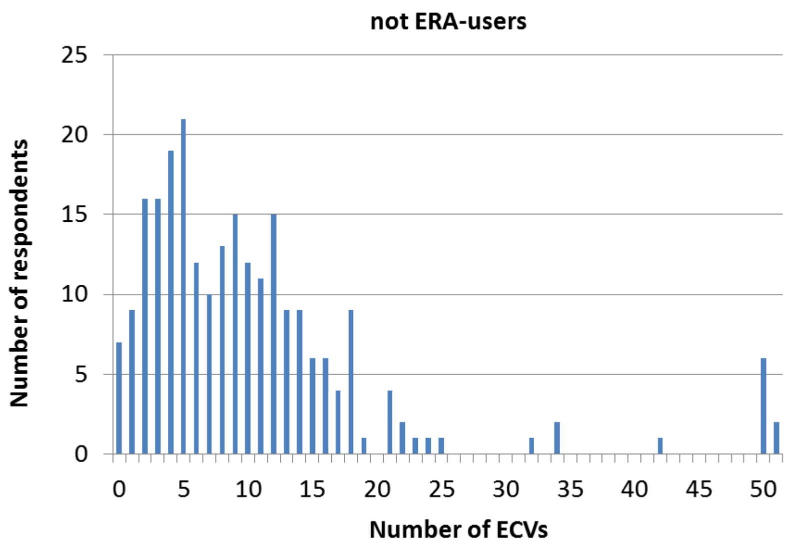


Figure 5.2. Distribution of the number of studied ECVs by the not ERA-users (234).

Those who are not using ECMWF's atmospheric reanalyses indicated that they use reanalysis data from on average 1.1 atmospheric reanalysis datasets (Table 5.1). The most often used datasets were NCEP reanalysis R1 (32% of all given responses) and R2 (13%) (Fig. 5.3).

Table 5.1. Used atmospheric reanalysis datasets by all respondents, the ERA-Interim users and the not ERA-users. It was asked to choose all that apply.

Atmospheric reanalyses:	All respondents	ERA-Interim users	not ERA-users
Global ECMWF Interim Reanalysis (ERA-Interim)	2049	391	0
Global ECMWF 40 year Reanalysis (ERA-40)	1325	0	0
Global ECMWF Reanalysis (ERA-15)	309	0	0
Global Japanese Reanalysis (JRA-25) or its continuation as JMA Climate Data Assimilation System (JCDAS)	409	0	14
Global NASA MERRA	505	0	27
Global NCEP Climate Forecast System Reanalysis (CFSR)	572	0	37
Global NCEP/NCAR Reanalysis I (R1) (1948 to present)	1010	0	79
Global NCEP/DOE Reanalysis AMIP-II (R2) (1979 to near present)	592	0	32
Global NOAA-CIRES 20th Century Reanalysis (20CR)	315	0	18
Arctic System Reanalysis (ASR)	43	0	4
NCEP North American Regional Reanalysis (NARR)	212	0	16
The European Reanalysis and Observations for Monitoring project (EURO4M)	53	0	0
Global MACC Reanalysis 2003-2010	75	0	3
Global GEMS Reanalysis 2003-2007	43	0	1
Other, please specify	85	0	26
Total	7597	391	257

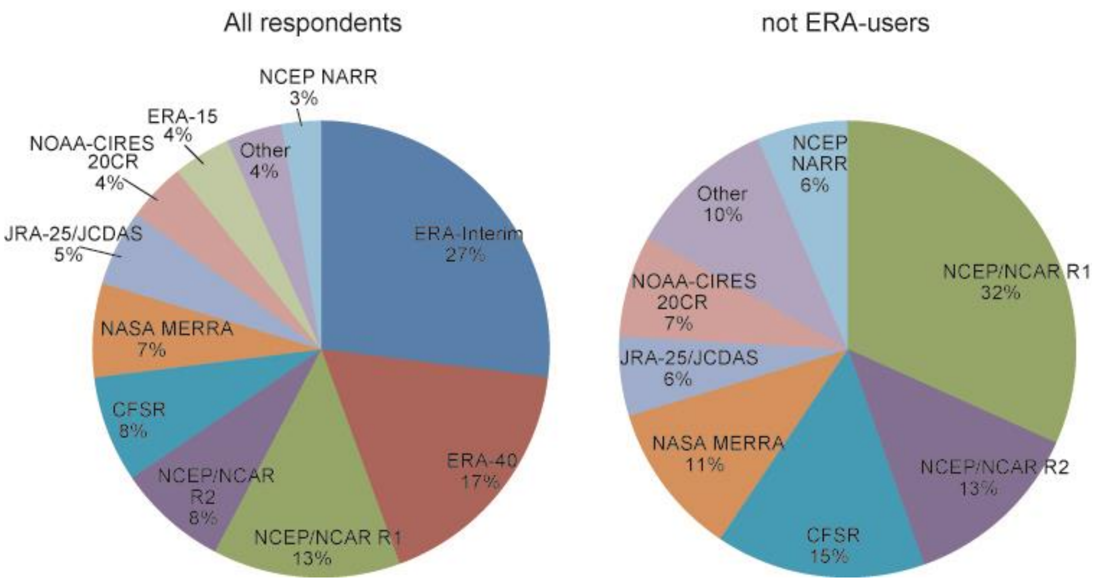


Figure 5.3. Proportion of atmospheric reanalysis datasets used by the respondents. Note that the percentage values are calculated from the total number of responses given for all the different reanalysis options (see Table 5.9), not from the number of respondents (all 2502, not ERA-users 241). It was asked to choose all that apply.

5.3 Applications and methods

Regarding the reanalysis input observations and feedback data, 62% of the ERA-Interim users indicated that they do not know what it is about. For the not ERA-users the proportion was 44%, which was similar to that of all respondents (49%). The proportion of those who think that the data formats of the reanalysis input observations and feedback data are too big or too complicated, or that there is no easy interface to the data, is more pronounced in the not ERA-users than in the ERA-Interim users (see Appendix C for illustration).

5.4 User awareness and needs

Questions about the characteristics of reanalysis data revealed that the ERA-Interim users were in general more satisfied with the reanalysis data than the not ERA-users (Fig. 5.4). This was evident especially for the Data access and availability aspects where more or less half of the ERA-Interim users fully or somewhat agreed on easy data access, not too strict data policy, not too large data files, and easy data import of the reanalysis data. For the not ERA-users 20-40% fully or somewhat agreed on these issues. Alternative interpretation could be that data access and availability aspects are factors in making ECMWF reanalyeses less attractive to some respondents.

For data quality and representation of uncertainties, 60% of the ERA-Interim users fully or somewhat agreed on data consistency between different variables. For the not ERA-users the proportion of those agreeing was 30%, and correspondingly the share of those in-between was clearly higher (roughly 60%) than for the ERA-Interim users (roughly 30%). General awareness of the characteristics of the reanalysis data sets was higher for the ERA-Interim users. The statement “I know enough to work with the data” was fully or somewhat agreed by 45% of the ERA-Interim users, whereas less than 25% of the not ERA-users fully or somewhat agreed on that.

In general, the share of those fully or somewhat disagreeing was in many cases roughly the same for both groups (red bars in Fig. 5.4). This means that the differences between the groups derive from those who either agree (fully or somewhat) or who are in between (or did not answer at all). The share of those who were in between or did not answer at all was for majority of the propositions clearly higher for the not ERA-users than for the ERA-Interim users. This could indicate that, in general, the share of less-informed respondents is higher in the not ERA-users than in the ERA-Interim users.

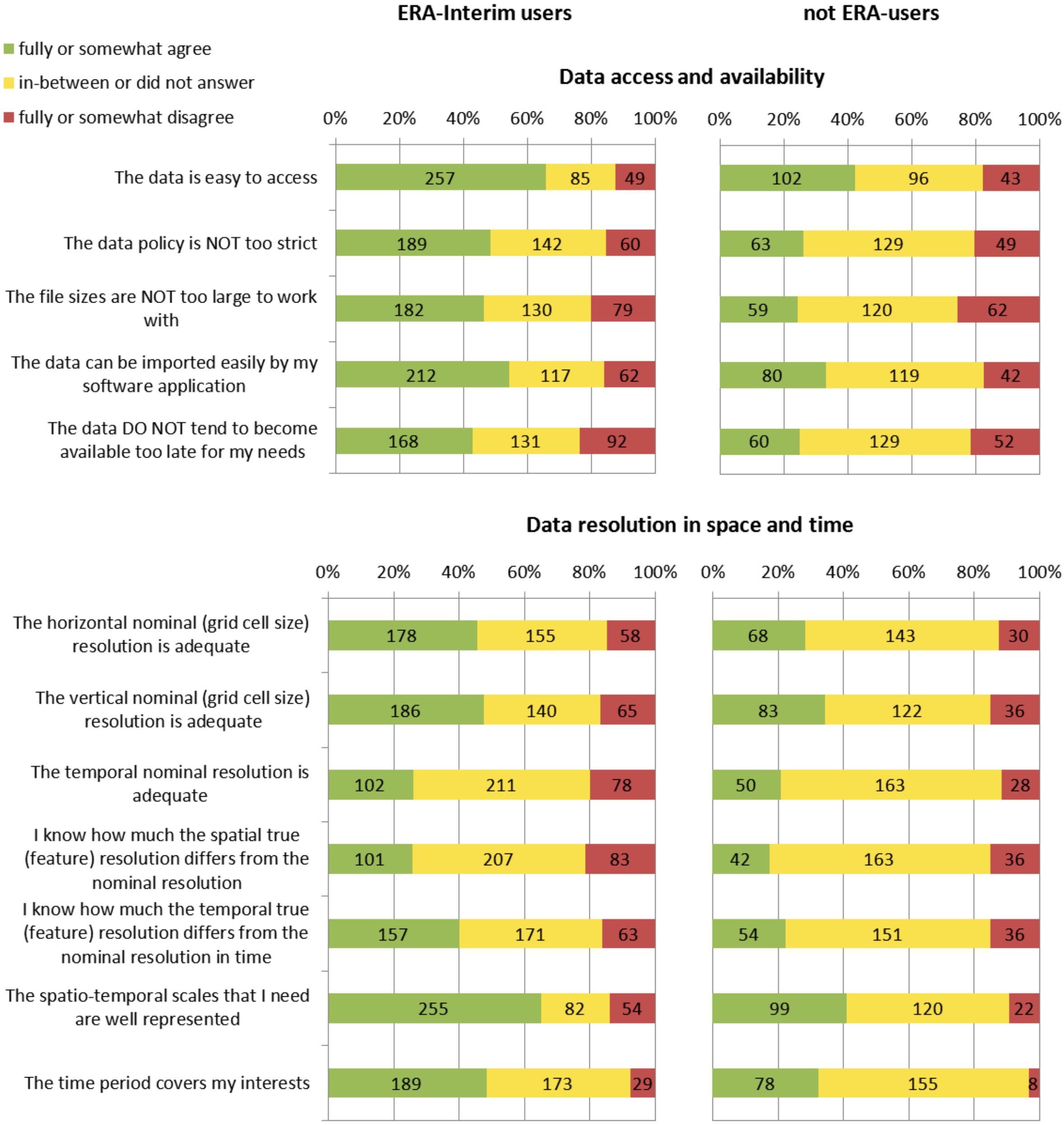


Figure 5.4. Responses (ERA-Interim users 381, not ERA-users 191) to propositions about the characteristics of reanalysis data. (continues)

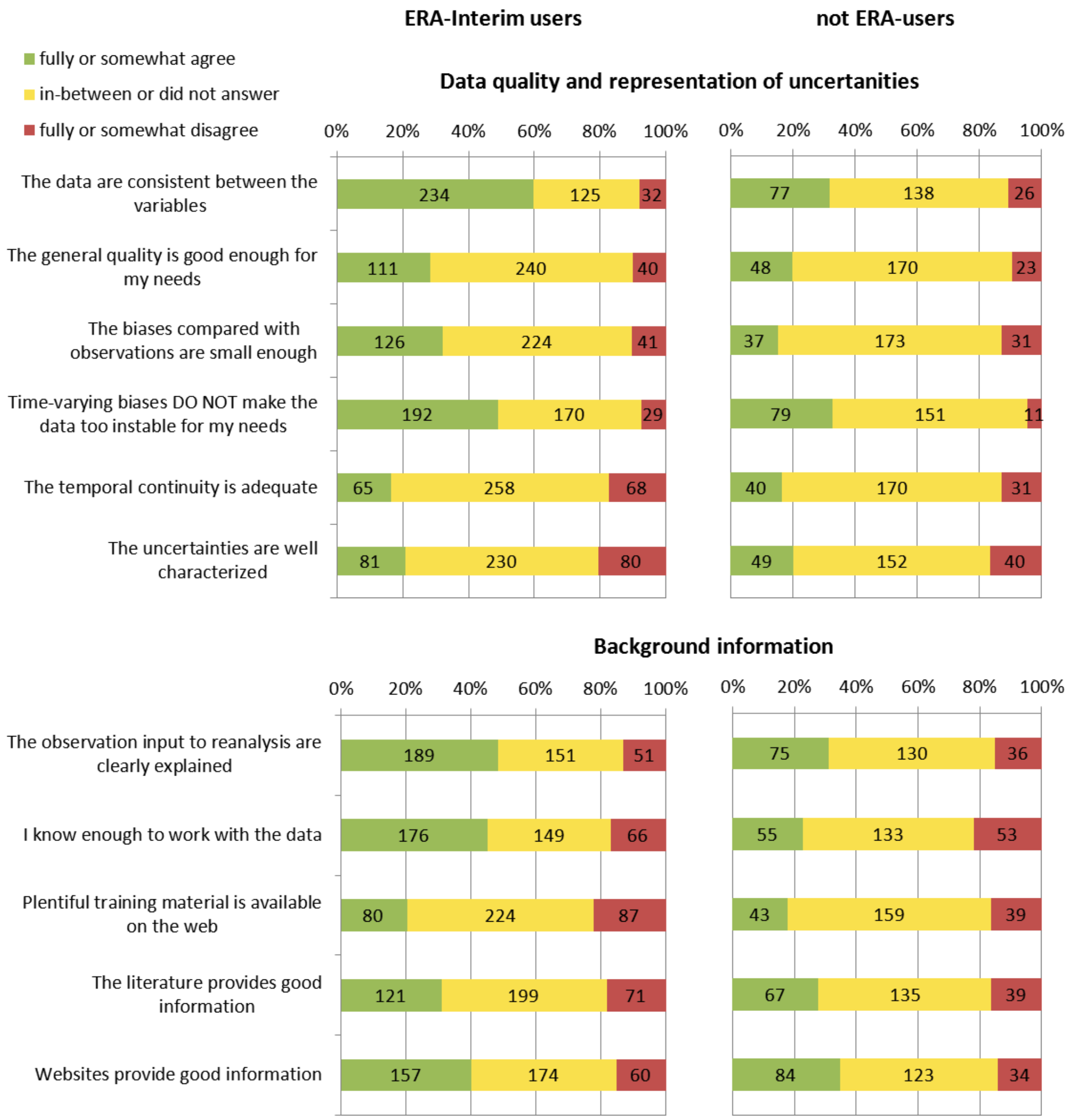


Figure 5.4. (continues) Responses (ERA-Interim users 381, not ERA-users 191) to propositions about the characteristics of reanalysis data.

5.5 Climate Services

When asking about wishes for the future climate service tasks or activities the replies from the ERA-Interim users and the not ERA-users were quite similar (Fig. 5.5). The most wanted tasks were "Interpolation and production of gridded data sets based on observations", "Provision of statistics based on observations" and "Homogenization of station data". These were the most wanted tasks also regarding all the respondents.

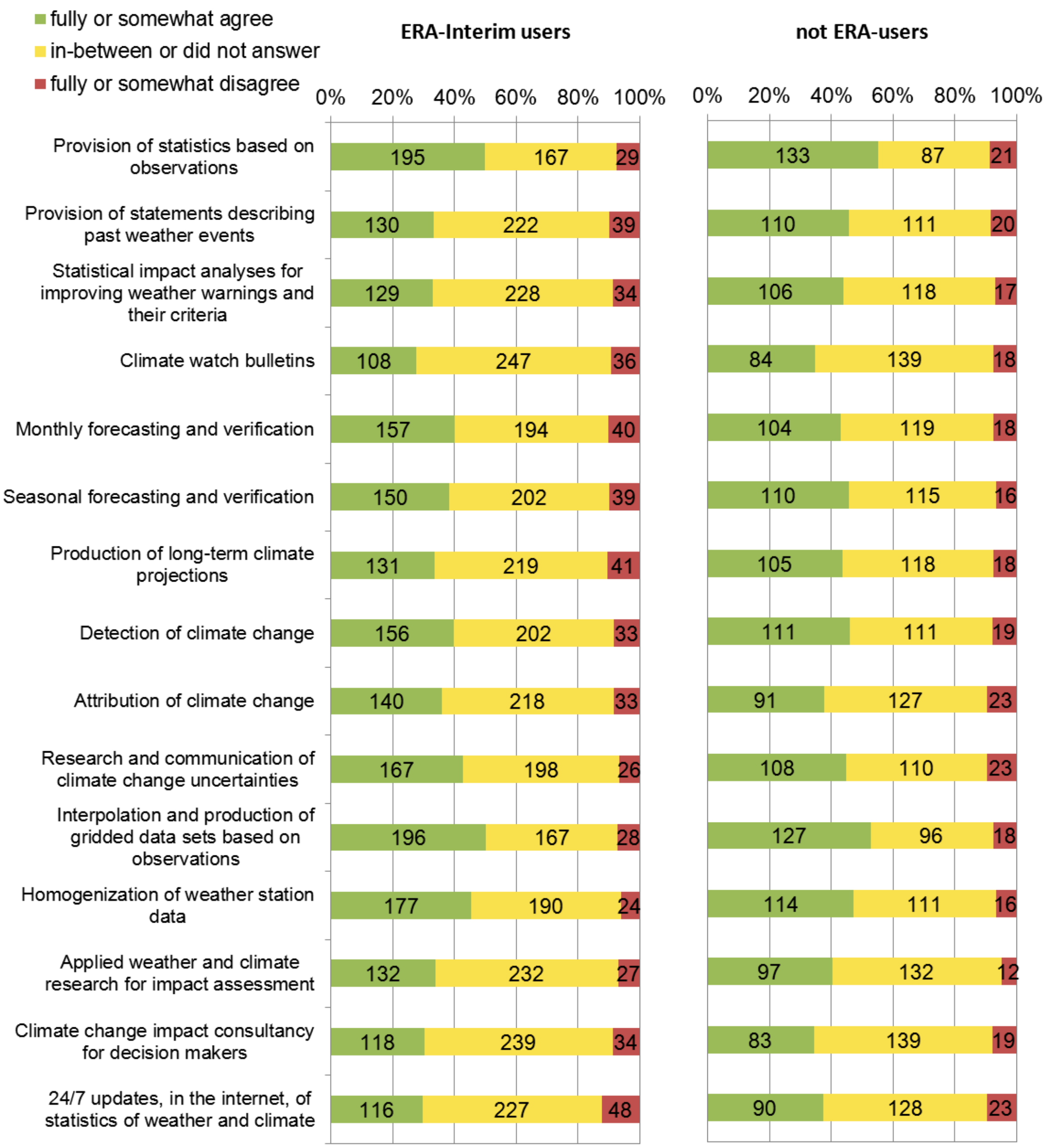


Figure 5.5. Responses (ERA-Interim users 315, not ERA-users 193) to propositions about tasks or activities in future climate services.

The priority/relevance of the suggested future climate service tasks and activities can be assessed by calculating the average of all the responses for each proposition, and then sorting the propositions according to their score. The lower the score (average of the responses) the more wanted the task or activity would be:

ERA-Interim users
1.0: Fully agree

2.0	Interpolation and production of gridded data sets based on observations
2.0	Provision of statistics based on observations
2.1	Homogenization of weather station data
2.1	Research and communication of climate change uncertainties
2.3	Monthly forecasting and verification
2.3	Detection of climate change
2.3	Seasonal forecasting and verification
2.4	Attribution of climate change
2.4	Applied weather and climate research for impact assessment
2.4	Statistical impact analyses for improving weather warnings and their criteria
2.5	Climate change impact consultancy for decision makers
2.5	Production of long-term climate projections
2.5	Provision of statements describing past weather events
2.5	24/7 updates, in the internet, of statistics of weather and climate
2.6	Climate watch bulletins

5.0: Fully disagree

Not ERA-users
1.0: Fully agree

1.8	Interpolation and production of gridded data sets based on observations
1.8	Provision of statistics based on observations
1.9	Homogenization of weather station data
2.0	Detection of climate change
2.0	Applied weather and climate research for impact assessment
2.0	Seasonal forecasting and verification
2.0	Statistical impact analyses for improving weather warnings and their criteria
2.0	Monthly forecasting and verification
2.0	Provision of statements describing past weather events
2.0	Research and communication of climate change uncertainties
2.1	Production of long-term climate projections
2.2	Climate change impact consultancy for decision makers
2.2	24/7 updates, in the internet, of statistics of weather and climate
2.2	Attribution of climate change
2.3	Climate watch bulletins

5.0: Fully disagree

6. Survey analysis: "best-informed" and "least-informed" users

In order to assist in identifying more specialized user needs a division into "best-informed users" and "least-informed users" was done. By using survey questions about user awareness and needs the following division into these sub-groups was done. As it turned out that the amount of respondents in these groups was fairly low compared to the whole respondent group (7% of all respondents fullfilling the definition of "best-informed users" and 10% fulfilling the definition of "least-informed users"), the results based on these groups can be regarded mainly approximate. Another ways of defining the best- and least-informed users could be worth considering in future.

"Best-informed users" are those who

- chose AT LEAST ONE of the following in Q7 (Have you used reanalysis input observations and feedback data?): "I have used it to assess the reanalysis data using observations as a reference", "I have used it to assess the observations using reanalysis data as a reference", "I have used it to merge the observations and reanalysis data together to create an improved product" and "I have used it to understand how the observations had been used by reanalysis"
- AND FULLY or SOMEWHAT AGREED on "For the climate variables I need, I know how much their spatial true (feature) resolution differs from the nominal resolution" in Q8
- AND FULLY or SOMEWHAT AGREED on "I know how much the temporal true (feature) resolution differs from the nominal resolution" in Q8.

There were 170 respondents (7% of all respondents) fulfilling the definition.

"Least-informed users" are those who

- CHOSE "I do not know what this is about" in Q7 (Have you used reanalysis input observations and feedback data?)
- AND FULLY or SOMEWHAT DISAGREED or were IN-BETWEEN on "For the climate variables I need, I know how much their spatial true (feature) resolution differs from the nominal resolution" in Q8
- AND FULLY or SOMEWHAT DISAGREED or were IN-BETWEEN "I know how much the temporal true (feature) resolution differs from the nominal resolution" in Q8
- AND FULLY or SOMEWHAT DISAGREED or were IN-BETWEEN "I know enough to work with the data" in Q8.

There were 245 respondents (10% of all respondents) fulfilling the definition.

Main findings and results concerning especially about the user awareness and needs, as well as future climate service activities are shown in the following pages. More detailed information and illustrations about the two sub-groups can be found in Appendix D.

6.1 Background of the respondents

Division into different sectors of work was surprisingly similar both among the best- and least-informed users and among all respondents. For the most popular topics to work with, the best-informed users chose more often "weather" whereas the least-informed were more to "climate" and "oceans, seas". The division of respondents' regional focus revealed that both best- and least-

informed users are focusing more on oceans, and specifically on Pacific Ocean, when compared to all respondents. Correspondingly, Europe and Asia are less studied by these two sub-groups.
Figures and tables related to the background of the respondents can be found in Appendix D: Figs. D.1-D.3 and Tables D.1-D.2.

6.2 Reanalysis data and Essential Climate Variables

The best-informed users work on average with 18.3 different Essential Climate Variables, for the least-informed users the number is 14.7. The distributions of the number of studied ECVs by the respondents are shown in Fig. 6.1 and 6.2. For the best-informed users the distribution was somewhat wider than for the least-informed users. Rather surprisingly, the most common number of ECVs to work with was for both groups 50, i.e., including all the variables that were given as an option (for the least-informed users including also the option “other, please specify”).

The use of ECVs among the best- and least-informed users is illustrated in Figs. D.4-C.4 shown in Appendix D. The share of the best-informed users compared to the least-informed users is larger for all ECVs that the respondents are working with USING reanalysis data, with the exception of surface air temperature, where the proportion of the least-informed users is slightly higher than the best-informed users (Fig. D.4). For the ECVs that the respondents are working with by NOT USING reanalysis data, the share of the least-informed users is in many cases larger than that of the best-informed users.

Compared to all respondents, the best-informed users indicated to work more with atmospheric composition, oceanic surface and sub-surface and terrestrial ECVs. The use of atmospheric surface ECVs was more consistent between these groups.

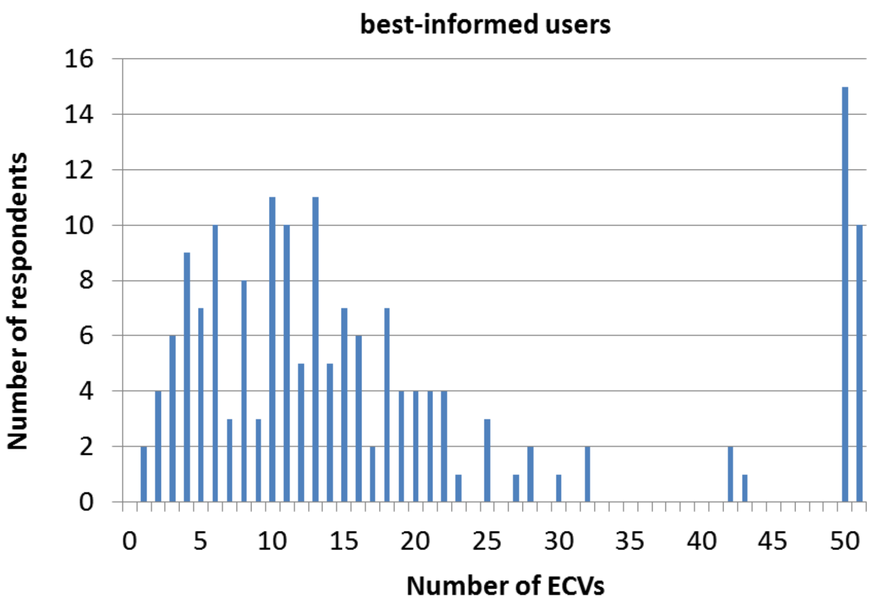


Figure 6.1. Distribution of the number of studied ECVs by the best-informed users (170).

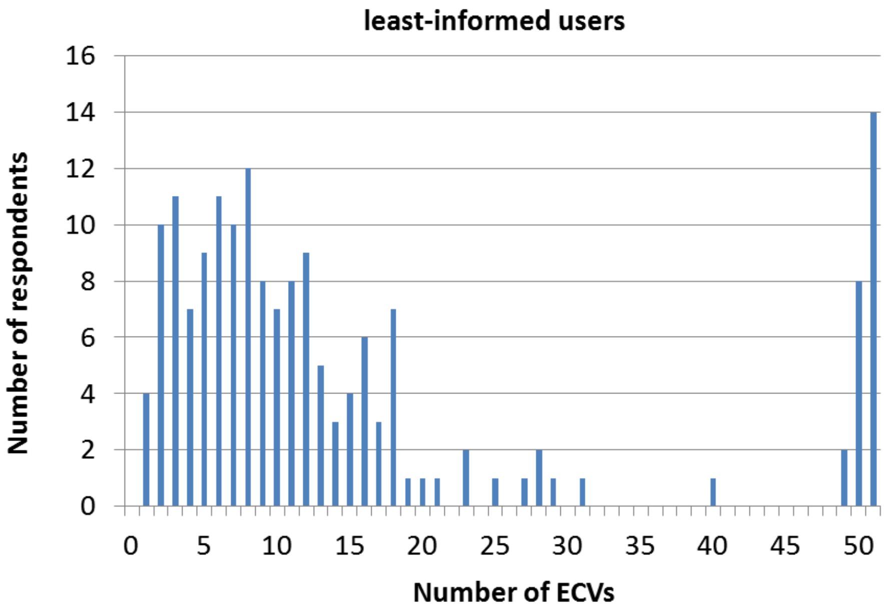


Figure 6.2. Distribution of the number of studied ECVs by the least-informed users (245).

6.3 User awareness and needs

Dealing with the best- and least-informed users, the responses to propositions about characteristics of reanalysis data are given in Fig. 6.10.

In general, the best-informed users are more satisfied with the different reanalysis data characteristics than the least-informed users. By definition, all the best-informed users fully or totally agreed to know what the difference is between the true and nominal spatial and temporal resolution. After this, the best-informed users were most satisfied with propositions concerning easy data access, consistency between different variables, and general awareness ("I know enough to work with the data"). The best-informed users were least satisfied with propositions concerning time-varying biases, the availability of the data in time, data file sizes, and data policy. However, these were the very propositions whose formulation in the original survey was inconsistent as compared to most propositions, and should be therefore regarded with some suspicion.

By definition, all the least-informed users fully or somewhat disagreed or were in between to know what the difference is between the true and nominal spatial and temporal resolution, and stating that they know enough to work with the data. Results show that the share of those who disagreed with these issues was higher than that of those being in between or not answering at all for the former propositions (about true and nominal resolution) and lower for the latter ("I know enough to work with the data"). After these, the least-informed users were least satisfied with information about the observation input to reanalysis, available online training material, and characterization of uncertainties. The least-informed users were most satisfied with easy data access, time period coverage, and data policy.

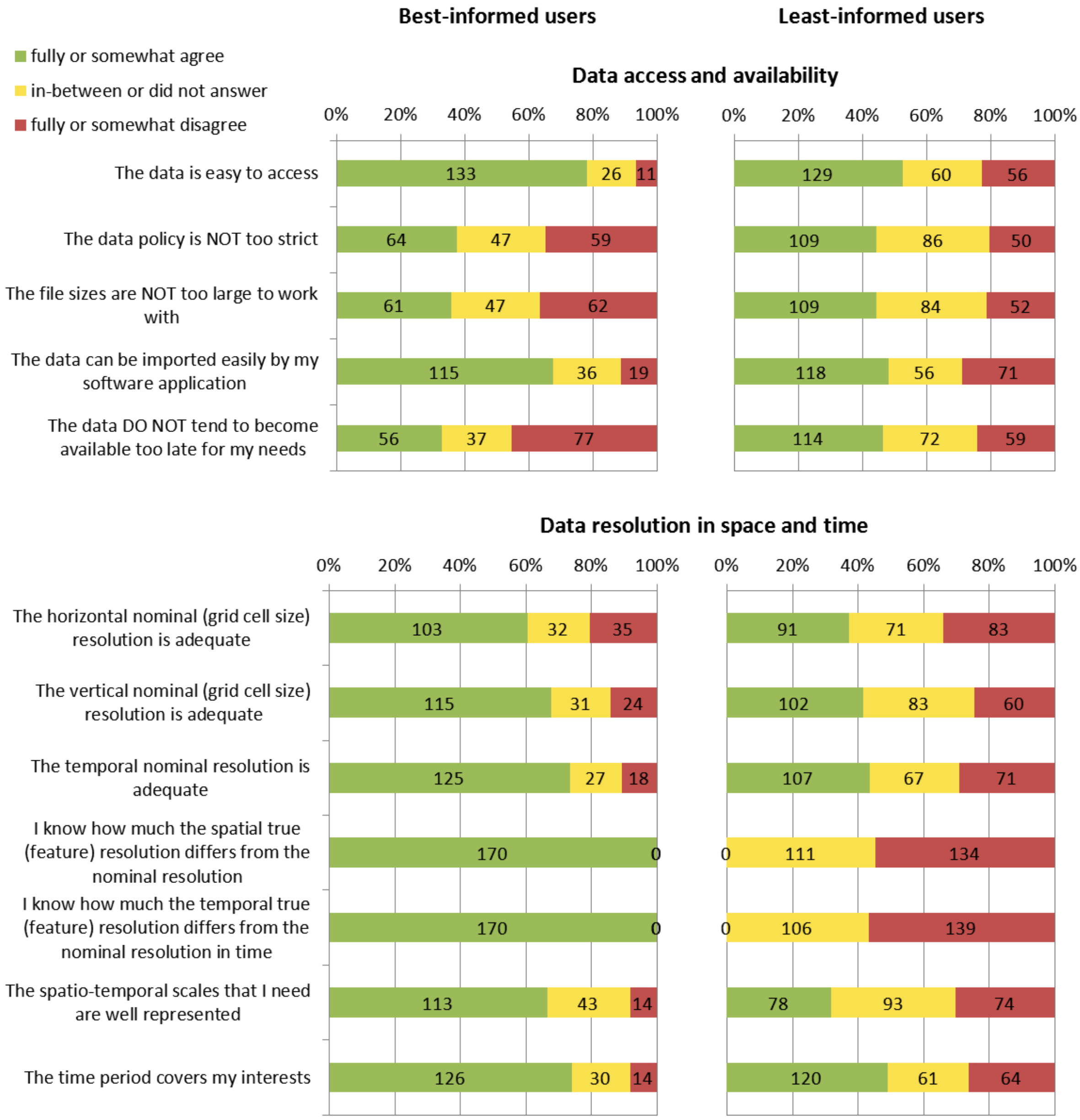


Figure 6.10. Responses (best-informed users 170, least-informed users 245) to propositions about the characteristics of reanalysis data (continues).

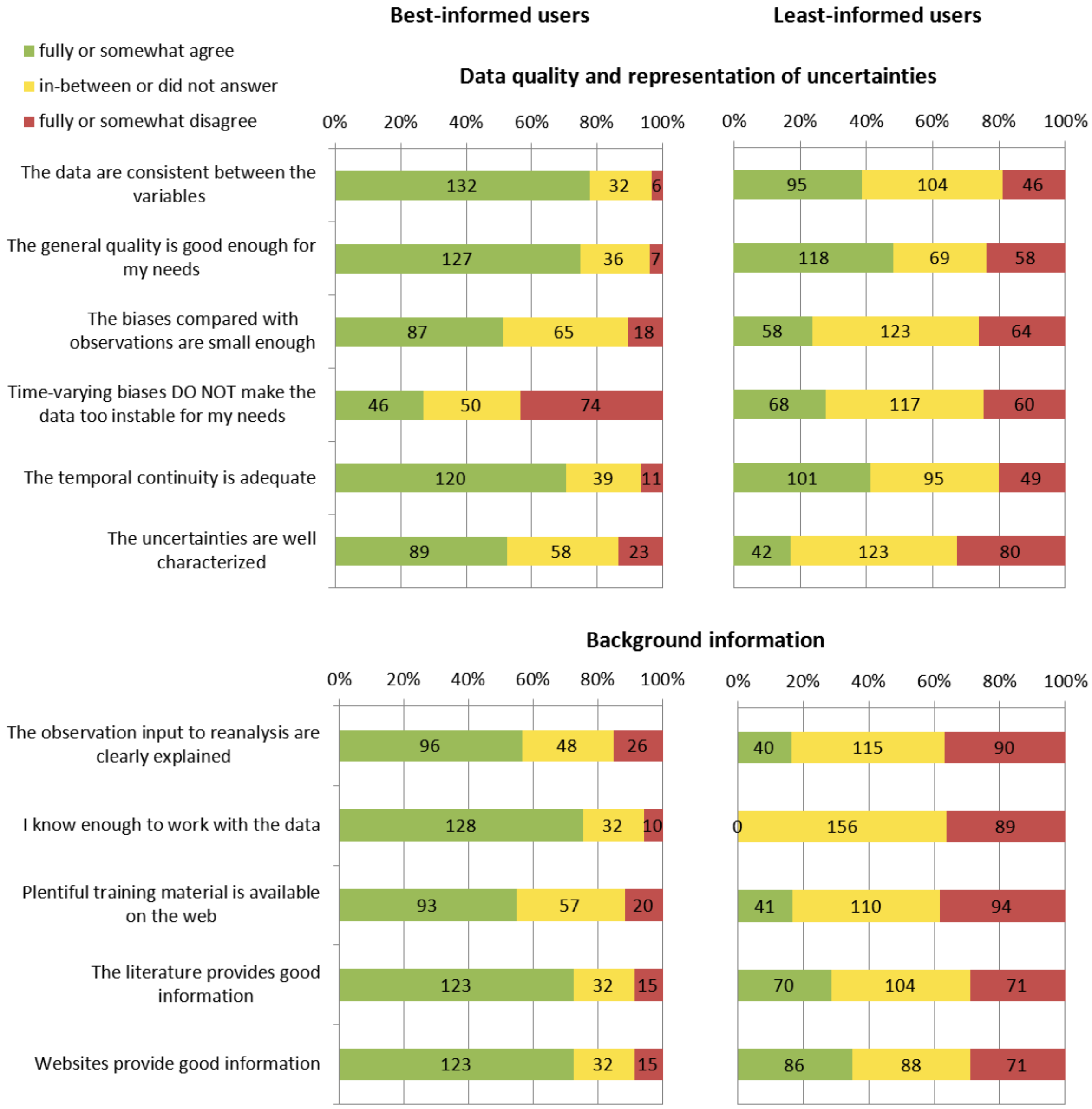


Figure 6.10. (continues) Responses (best-informed users 170, least-informed users 245) to propositions about the characteristics of reanalysis data.

6.4 Climate services

The most wanted tasks for future climate services were the same as for all respondents: "Provision of statistics based on observations", "Interpolation and production of gridded data sets based on observations", and "Homogenization of weather station data". However, the order of the top three was varying; best-informed users wished most for provision of statistics based on observations, whereas the least-informed wished most for homogenizations of the weather station data. For all users, the most wanted task was interpolation and production of gridded data sets.

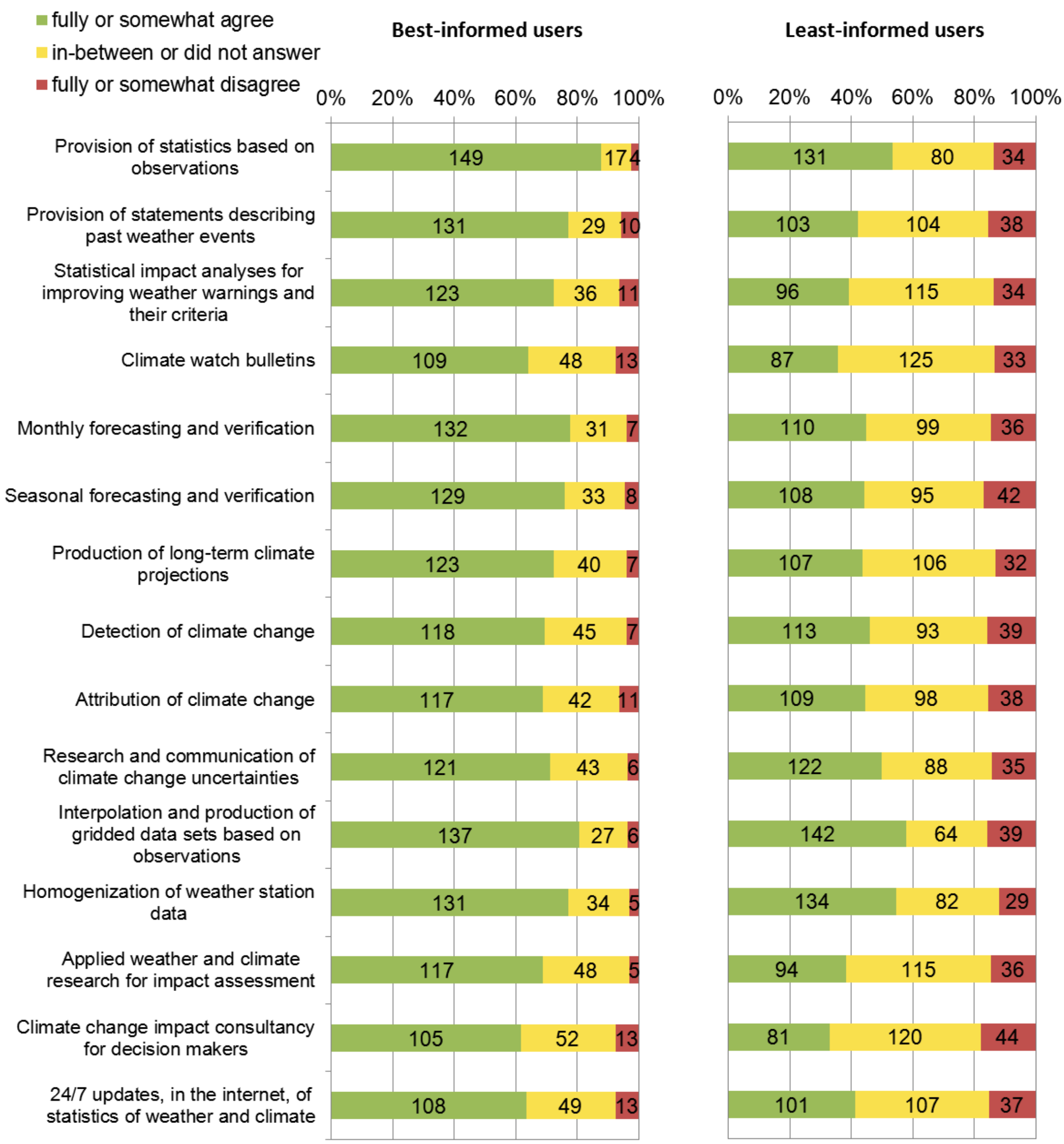


Figure 6.11. Responses (best-informed users 160, least-informed users 220) to propositions about tasks or activities in future climate services.

The priority/relevance of the suggested future climate service tasks and activities can be assessed by calculating the average of all the responses for each proposition, and then sorting the propositions according to their score. The lower the score (average of the responses) the more wanted the task or activity would be:

Best-informed users
1.0: Fully agree

1.4	Provision of statistics based on observations
1.5	Interpolation and production of gridded data sets based on observations
1.6	Homogenization of weather station data
1.6	Provision of statements describing past weather events
1.7	Monthly forecasting and verification
1.7	Seasonal forecasting and verification
1.7	Research and communication of climate change uncertainties
1.7	Production of long-term climate projections
1.7	Detection of climate change
1.8	Applied weather and climate research for impact assessment
1.8	Statistical impact analyses for improving weather warnings and their criteria
1.8	Attribution of climate change
1.9	24/7 updates, in the internet, of statistics of weather and climate
1.9	Climate change impact consultancy for decision makers
1.9	Climate watch bulletins

5.0: Fully disagree

Least-informed users
1.0: Fully agree

2.2	Homogenization of weather station data
2.2	Interpolation and production of gridded data sets based on observations
2.3	Provision of statistics based on observations
2.3	Research and communication of climate change uncertainties
2.4	Detection of climate change
2.4	Production of long-term climate projections
2.4	Attribution of climate change
2.5	Monthly forecasting and verification
2.5	24/7 updates, in the internet, of statistics of weather and climate
2.5	Provision of statements describing past weather events
2.5	Seasonal forecasting and verification
2.5	Applied weather and climate research for impact assessment
2.5	Statistical impact analyses for improving weather warnings and their criteria
2.6	Climate watch bulletins
2.7	Climate change impact consultancy for decision makers

5.0: Fully disagree

7. Summary and conclusions of the core components of climate services and the identified gaps in reanalyses based on the survey and literature review

To learn how widely and well used the reanalysis products are currently compared to other sources of data, the respondents were asked to indicate their source of data, stating whether they use 1) reanalysis products; 2) weather station, radiosonde or other in-situ observations; 3) satellite-based remote sensing data; and/or 4) weather radar based remote sensing data. It was found that reanalysis was the most frequently used data source in the applications of the respondents. The reanalysed data products were used especially when studying atmospheric dynamics, doing atmospheric and climate modelling and well as making time series analyses and studying atmospheric physics.

When asking all respondents what they think are part of the future climate services, it appeared that the core components of climate services should be 1) provision of statistics based on observations, 2) interpolation and production of gridded data sets based on observations, and 3) homogenization of weather station data. Reanalyses fit well in these categories as they are produced based on all available observations by employing sophisticated data assimilation schemes when calculating the final products. Therefore it is perhaps no surprise that the reanalyzed data sets, according to the survey results, were more commonly used even in time series analyses than were the in situ observations and in-situ gridded datasets. However, in interpreting the survey results, readers should be mindful of potential sampling bias: the timeline of responses suggests that the bulk of respondents quite probably came from those registered for reanalysis-use at ECMWF. There remain open questions about whether the same profile of responses would occur in a wider survey, and whether such reanalysis-users are fully representative of the wider climate-service community.

Considering the users' awareness of the 50 Essential Climate Variables (ECVs), the respondents were asked to indicate those variables that they work with, stating whether they use reanalysis data or not. The distribution was broad; use of 2 to 13 different ECVs was common. There were also 107 respondents who indicated that they work with all given ECVs. Most used variables were the surface temperature, wind and pressure, then the upper air temperature and wind.

Nearly 80 % of the respondents indicated that they have used ECMWF's ERA-Interim reanalysis. To learn from the possible differences between ERA-Interim-users and not ERA-Interim-users, a division into following subgroups was made:

- ✓ **"ERA-Interim users":** uses only ECMWF's ERA-Interim atmospheric reanalysis (391 respondents, 15% of all respondents).
- ✓ **"not ERA-users":** doesn't use any atmospheric reanalysis by ECMWF (241 respondents, 9% of all respondents)

This division indicated that the not ERA-users were more interested to work with those subjects that were not that popular among all respondents and ERA-Interim users, such as "energy", "fresh water resources and management", "agriculture", "forests", and "ecosystems and biodiversity". Those who were not using any ERA reanalysis, were focusing more on Europe and less on the whole globe when compared with ERA-Interim users or all respondents. The proportion of those who were focusing on Africa or North America was also emphasized in not ERA-users. Additionally, this division revealed that among the not-ERA-users the data formats of the reanalysis input observations and feedback data were found too big or too complicated, or that there was no easy interface to the data.

Considering the potential to improve reanalysis products and reanalysis based services, the users of reanalysis products would benefit from better presentation of the spatio-temporal scales employed in data assimilation and of somewhat smaller file sizes of data. For some users the release of the datasets could be quicker in order to be used e.g., in providing better daily climate services. Biases in data that exist in time and space are identified to possibly cause problems in climate service and research. Issues such as nominal resolution and feature resolution are not trivial for all users and could need to be explained and presented better. Training material and the underlying uncertainties are not recognized by all users and this should be paid attention to.

Overview documents covering comparisons between the different reanalyses, known strengths and weaknesses are desired. Guidance could include more information on uncertainties, criteria for product selection and scientific advice on limitations of usage and on the feature resolution which can be expected, as well as a description of successful use cases. Thus, a more comprehensive set of diagnostics as, e.g., called for in *Bengtsson et al., 2007*, would help the users to factor in the uncertainties inherent in the reanalysis data.

In conclusion following needs of users of reanalyses were identified:

1. Users need guidance to decide for which parameters, and at which scales, reanalysis data might be a superior alternative to the possibly scarce, locally influenced, or inhomogeneous observational records

2. Users would like to have support for choosing the most appropriate reanalysis (or the proper ensemble of reanalysis) for their application

3. Users would benefit of having more training material available on the web accompanied by a more comprehensive set of diagnostics

4. Users would like to have the reanalysis products processed and released real time to be able to use these in provision of daily climate services

5. Users would like to have the observation input to reanalyses and the nominal and feature resolution of the climate variables better explained

6. Tools for users would be desirable such that they can access information on uncertainties at their specific spatio-temporal scale of interest

7. Users need access to feedback statistics, to be able to compare the reanalysis fields against chosen observations

8. Increase in temporal stability as well as temporal and spatial resolution of reanalyses is desired

9. User-tailored post-processing could be developed such as to produce frequency distributions and other statistics comparable to traditional observations. This would open an easy access for applications based on traditional observations to use reanalysis data.

References

Bengtsson, L. and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. Bull Amer. Meteor. Soc., 69, 1130-1143.

Bengtsson, L., Arkin, P., Berrisford, P., Bougeault, P., Folland, C. K., Gordon, C., Haines, K., Hodges, K. I., Jones, P., Kallberg, P., Rayner, N., Simmonds, A. J., Stammer, D., Thorne, P. W., Uppalöa, S., Vose, R. S. 2007: The need for a dynamical climate reanalysis. BAMS, DOI: 101.1175/BAMS-88-4-495.

Bosilovich, M., 2012: After Two Decades of Retrospective-Analysis: What Next? GEWEX NEWS: 22, No. 1.

Bosilovich, M. G., J. Kennedy, D. Dee, R. Allan and A. O'Neill, 2013: On the Reprocessing and Reanalysis of Observations for Climate. *Climate Science for Serving Society: Research, Modelling and Prediction Priorities*, Asrar, G., Hurrell, J. W. eds., Springer, 51-71.

Compo, G. P., J. S. Whitaker and P.D. Sardeshmukh, 2011: The Twentieth Century Reanalysis Project. Quart. J. Roy. Meteor. Soc., 137, 1-28.

Dee, D., Berrisford, P., Bosilovich, M., Chelliah, M., Compo, G., Ebita, A., Jones, P., Kobayashi, S., Kumar, A., Rutledge, G., Saha, S., Sato, H., Simmons, A., Smith, C. & Vose, R., 2011a: The use of reanalysis data for monitoring the state of the climate. State of the Climate in 2010, Bull. Amer.Meteor.Soc., 92, S34-S35.

Dee et al., 2011b: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quart. J. R. Meteorol. Soc., 137, 553-597.

Dee, D. P., E. Källén, A. J. Simmons, L. Haimberger, 2011c: Comments on "Reanalyses Suitable for Characterizing Long-Term Trends". Bull. Amer. Meteor. Soc., 92, 65–70.

Dee, D. P., M. Balmaseda, G. Balsamo, R. Engelen, A. J. Simmons and J.-N. Thépaut, 2014: Toward a consistent reanalysis of the climate system. BAMS early online release accepted 2013 *http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-D-13-00043.1*

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., ... & Ishimizu, T. (2011). The Japanese 55-year Reanalysis "JRA-55": an interim report. *Sola*, *7*, 149-152.

Gibson, J.K., P. Kallberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano, 1997: ERA description: ECMWF ERA-15 Project Report Series no1, 71 pp.

Hartmann, D.L., A.M.G. Klein Tank, M. Rusticucci, L.V. Alexander, S. Brönnimann, Y. Charabi, F.J. Dentener, E.J. Dlugokencky, D.R. Easterling, A. Kaplan, B.J. Soden, P.W. Thorne, M. Wild and P.M. Zhai, 2013: Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hollingsworth, A. and C. Pfrang, 2005: A preliminary survey of ERA-40 users developing applications of relevance to GEO (Group on Earth Observations*). ECMWF Newsletter*, 104, 5-9.

Kalnay et al., 1996: The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Meteor. Soc., 77, 437-470.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). NCEP–DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, *83*(11).

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bull. Amer. Meteor. Soc.*, 82, 247-268.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., ... & Taira, R. (2007). The JRA-25 reanalysis. *J. Meteorol. Soc. Jpn*, *85*(3), 369-432.

Manton, M.J.; Belward, A.; Harrison, D.E.; Kuhn, A.; Lefale, P.; Rosner, S.; Simmons, A.; Westermeyer, W.; Zillman, J., 2010: Observation Needs for Climate Services and Research. Procedia Environmental Sciences, 1, 184-191.

Rienecker, M.M., D. Dee, J. Woollen, G.P. Compo, K. Onogi, R. Gelaro, M.G. Bosilovich, A. da Silva, S. Pawson, S. Schubert, M. Suarez, D. Barker, H. Kamahori, R. Kistler, and S. Saha, 2012: Atmospheric Reanalyses—Recent Progress and Prospects for the Future. A Report from a Technical Workshop, April 2010. NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM–2012-104606, Vol. 29, 56 pp.

Saha et al., 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc., 91, 1015–1057.

Thorne, P. W., R. S. Vose, 2010: Reanalyses Suitable for Characterizing Long-Term Trends. Are they really achievable? Bull. Amer. Meteor. Soc., 91, 353–361.

Thorne, P. W., R. S. Vose, 2011: Reply. Bull. Amer. Meteor. Soc., 92, 70–72. http://dx.doi.org/10.1175/2010BAMS3145.1

Trenberth K.E. and J. G. Olson, 1988: An evaluation and intercomparison of global analyses from NMC and ECMWF. Bull Amer. Meteor. Soc. 69, 1047-1057.

Trenberth K.E., 2007: Atmospheric reanalyses: a major resource for climate services. SP-636 Envisat Symposium.

Trenberth, K.E., 2008: Observational needs for climate prediction and adaptation. WMO Bulletin, 57, 17-21.

Trenberth, K.E., Koike, T. Onogi, K., 2008: Progress and prospects in reanalysis. Eos, 89, 26, 24 June 2008, 234-235.

Trenberth, K. E.; Anthes, R. A.; Belward, A.; Brown, O.; Haberman, E.; Karl, T. R.; Running, S.; Ryan, B.; Tanner, M.; Wielicki, B., 2013: Challenges of a sustained climate observing system. *Climate Science* *for Serving Society: Research, Modelling and Prediction Priorities*, Asrar, G., Hurrell, J. W. eds., Springer, 13-50.

Uppala et al., 2005: The ERA-40 re-analysis. Quart. J. R. Meteorol. Soc., 131, 2961-3012.

Uppala, S., Broad, A., Delécluse, P. Köchy, M., Strahlendorff, M., Doherty, M., Schulz, J., Simmonds, A., Kattenberg, A., Noyes, S., Fuchs, T., Jol, A., Erhard, M., 2011: Draft report GMES Climate Service: Towards a European knowledge base in support of mitigation and adaption, Brussels, 10/06/2011 DG ENTR/H.4/CC .

van den Hurk, B, 2012: Reanalyses and their Applicability for Climate Research. GEWEX NEWS: 22, No. 1.

WMO, 2011: Climate Knowledge for Action: A Global Framework for Climate Services – Empowering the Most Vulnerable. The Report of the High-Level Taskforce for the Global Framework for Climate Services, WMO No. 1065, World Meteorological Organization, Geneva, Switzerland.

WMO 2013: GFCS - Intergovernmental Board on Climate Services. WMO Bulletin, 62 - Special issue.

APPENDIX A: Currently maintained collections and comparisons of reanalysis characteristics in the internet

Comparison table from reanalyses.org (March 2014):

Name	Source	Time Range	Assimilation	Model Resolution	Model Output Resolution	Publicly Available Dataset Resolution	Dataset Output Times and Time Averaging
Arctic System Reanalysis (ASR)	Byrd Polar Research Center Polar Meteorology Group	2000-2010 (30km) 2000-2011 (10km)	WRFDA-3DVAR	10 and 30km 71 sigma levels	10 and 30km	10 and 30km	3-hourly WRF outputs; selected variables for surface and upper-air fields. Monthly averages of selected fields.
ECMWF Interim Reanalysis (ERA Interim)	ECMWF	1979-present	4D-VAR	T_L255L60 and N128 reduced Gaussian	T_L255L60 and N128 reduced Gaussian (~79km globally)	User defined, down to 0.75x0.75	3-hourly for most surface fields; 6-hourly for upper-air fields Monthly averages of daily means, and of 6-hourly fields
ECMWF 40 year Reanalysis (ERA-40)	ECMWF	1958-2001	3D-VAR	T_L159L60 and N80 reduced Gaussian	T_L159L60 and N80 reduced Gaussian (~125km globally)	2.5x2.5 / 1.125x1.125	3-hourly for most surface fields; 6-hourly for upper-air fields Monthly averages of daily means, and of 6-hourly fields
Japanese Reanalysis (JRA-25) JMA Climate Data Assimilation System (JCDAS)	Japan Meteorological Agency	1979-2004 2005-present	3D-VAR	T106L40	T106L40 Gaussian	1.25x1.25 2.5x2.5	6-hourly daily monthly
NASA MERRA	NASA	1979-present	3D-VAR, with incremental update	2/3 lon x1/2 lat deg; 72 sigma levels	2/3 lon x1/2 lat deg 3d Analysis and 2d variables; 1.25 deg 3d Diagnostics; 72 model levels and 42 pressure levels	2/3 lon x1/2 lat deg 3d Analysis and 2d variables; 1.25 deg 3d Diagnostics; 72 model levels and 42 pressure levels	2d Diagnostics - 1 hourly avg, centered at half hour; 3d Diagnostics - 3 hourly avg, centered at 0130, 0430 ... 2230; 3d Analysis - Instantaneous 6 hourly; 2d Diagnostics, Monthly mean diurnal average; Monthly means for all collections; daily averages processed at servers on-the-fly
NCEP Climate Forecast System Reanalysis (CFSR)	NCEP	1979-?	3D-VAR	T382 L64	.5x.5 and 2.5x2.5	.5x.5 and 2.5x2.5	Hourly, 4 times daily
NCEP/DOE Reanalysis AMIP-II (R2)	NCEP/DOE	1979-present	3D-VAR	T62 L28	2.5x2.5	2.5x2.5	4 times daily/daily/monthly, also LTMs
NCEP/NCAR Reanalysis I (R1)	NCEP/NCAR	1948-present	3D-VAR	T62 L28	2.5x2.5 and 2x2 gaussian	2.5x2.5 and 2x2 gaussian	4 times daily/daily/monthly also LTMs
NCEP North American Regional Reanalysis (NARR)	NCEP	1979-present	RDAS	32km	32km	32km	4/8 times daily/daily/monthly also LTMs
NOAA-CIRES 20th Century Reanalysis (20CR)	NOAA/ESRL PSD	1871-2010	Ensemble Kalman Filter	T62 L28	2x2	2x2	4/8 times daily, daily,monthly, also LTMs

Comparison table from climatedataguide.ucar.edu (March 2014):

Summary of Atmospheric Reanalysis Data Sets discussed in Climate Data Guide

Name & Link to Page	Original Source	Scheme & model vintage	From	To	Model resolution	Domain	available timestep(s)	available format(s)	available resolution
Arctic System Reanalysis (ASR)	Byrd Polar Research Center / David Bromwich, Sheng-Hung Wang NCAR CIRES U Illinois	WRF-VAR	2000/01	2010/12	30 km (prototype)	Arctic	Sub-daily, Monthly	netCDF	30 km (360x360) ; ~0.27 deg of lat spacing
Climate Forecast System Reanalysis (CFSR)	NCEP	3DVAR 2009	1979/01	2010/12	T382 x 64 levels	Global	Sub-daily, Monthly	GRIB	.5°x.5° & 2.5°x2.5°, 0.266 hPA top
ERA-15	ECMWF		1979/01	1993/12	T106 (1.125)	Global	Sub-daily, Monthly	GRIB	T106, 2.5 x 2.5
ERA-Interim	ECMWF	4DVAR 2006	1979/01	2013/01	T255, 60 levels	Global	Sub-daily, Daily, Monthly	netCDF, GRIB	0.75°x0.75°x60 lev 0.1 hPA top
ERA40	ECMWF	3DVAR 2004	1957/01	2002/12	T159, 60 levels	Global	Sub-daily, Monthly	netCDF, GRIB	2.5°x2.5° / 1.125°x1.125°; 60 levels 0.1 hPA top
JRA-25	Japanese Meteorological Agency	3DVAR 2004	1979/01	2004/12	T106, 40 levels	Global	Sub-daily, Monthly	GRIB	1.125x1.125/2.5x2.5; 0.4 hPA top
JRA-55	Japanese Meteorological Agency	4DVAR 2009	1958/01	2012/12	T319 x 60 levels	Global	Sub-daily, Monthly	GRIB	T319 x 60 levels, 0.1 hPA top
NASA MERRA	NASA	GEOS IAU 2009	1979/01	2013/01	0.5° x 0.667° x 72	Global	Sub-daily, Monthly	netCDF, HDF	0.5° x 0.667° x 72 , 0.01 hPA top
NCEP NARR	NCEP	3DVAR 2003	1979/01	2012/09	32km x 45 eta	North America	Climatology, Sub-daily, Monthly	GRIB	32km
NCEP Reanalysis (R2)	NCEP DOE	3DVAR 2001	1979/01	2012/12	T62 28 levels	Global	Sub-daily, Daily, Monthly	netCDF, GRIB	2.5°x2.5° 28 levels 3 hPA top
NCEP-NCAR (R1): An Overview	NCEP NCAR	3DVAR 1995	1948/01	2013/01	T62 - 28 levels	Global	Sub-daily, Daily, Monthly	netCDF, GRIB	2.5°x2.5°; 3 hPA top
NOAA 20th-Century Reanalysis, Version 2	NOAA ESRL CIRES CDC	Ensemble Kalman Filter 2009	1871/01	2010/12	T62 28 levels	Global	Sub-daily, Daily, Monthly	netCDF, GRIB	2°x2°, 28 levels 10 hPA top

APPENDIX B: Webportal questionnaire

Core Climax

Reanalysis User and Application Survey

Dear recipient,

The aim of this questionnaire is to understand the use of reanalysis products and their limitations.

The results of the questionnaire will be used to assess how, and if, reanalysis data could bring wider benefit for research and climate services.

The questionnaire is conducted within the EU FP7 project Coordinating Earth Observation Data Validation for RE-analysis for Climate Services (2013-2015).

There are altogether 11 questions about reanalyses, applications and methods, user awareness and climate services. It takes approximately 15 minutes to answer the questionnaire.

Please feel free to skip questions that you find non-relevant. If needed, you may break and continue later.

Analyses of the responses will be available by the end of May 2014.

We appreciate your cooperation !

Respondent´s background

1. What is your sector of work? (please choose one)

- Private sector
- Education sector
- Public sector; research and development
- Public sector; other operations (local, regional, or national)
- International agency or organization
- Non-profit
- Other, please specify

2. What best describes your field or subject of work? (please choose all that apply)

- Weather
- Climate
- Air quality
- Other aspects of the atmosphere
- Oceans, seas
- Fresh water resources and management
- Snow, ice
- Agriculture, food production
- Forests
- Ecosystems, biodiversity
- Erosion / flooding of coastal areas
- Energy
- Industry
- Transportation (land, air, marine)
- Economics
- Insurance
- Architecture, urban or other spatial design/management
- Construction and municipal engineering
- Health, human well-being
- Tourism, recreation
- Safety and security issues, disasters
- Indigenous peoples' issues
- Other social aspects
- Communication
- Geophysics, geochemistry
- Geoinformatics
- Geology
- Geography
- Other, please specify

3. On which region of the world do you mainly focus your work?

- The whole globe
- Mostly continents
- Mostly oceans
- Asia
- Africa
- Europe
- North America
- Middle East
- Oceania / Australia / New-Zealand
- Central America / Carribean
- South America
- Arctic
- Antarctic
- Atlantic Ocean
- Pacific Ocean
- Indian Ocean
- Specific countries, please specify
- Other, please specify

Reanalysis data

4. Please indicate below the variables that you work with, stating whether you use reanalysis da not

	I work with this variable and use reanalysis data for this	I work with this varia do not use reanalysis this
Atmospheric surface:		
Air temperature	○	○
Wind speed and direction	○	○
Water vapour	○	○
Pressure	○	○
Precipitation	○	○
Surface radiation budget	○	○
Atmospheric upper air:		
Temperature	○	○
Wind speed and direction	○	○
Water vapour	○	○
Cloud properties	○	○
Earth radiation budget (including solar irradiance)	○	○
Atmospheric composition:		
Carbon dioxide	○	○
Methane	○	○
Other long-lived greenhouse gases	○	○
Ozone and precursors	○	○
Aerosols and precursors	○	○
Oceanic surface:		
Sea-surface temperature	○	○
Sea-surface salinity	○	○

Sea level	○	○
Sea state (waves)	○	○
Sea ice	○	○
Surface current	○	○
Ocean colour	○	○
Carbon dioxide partial pressure	○	○
Ocean acidity	○	○
Phytoplankton	○	○
Oceanic sub-surface:		
Temperature	○	○
Salinity	○	○
Current	○	○
Nutrients	○	○
Carbon dioxide partial pressure	○	○
Ocean acidity	○	○
Oxygen	○	○
Tracers	○	○
Terrestrial:		
River discharge	○	○
Water use	○	○
Groundwater	○	○
Lakes	○	○
Snow cover	○	○
Glaciers and ice caps	○	○
Ice sheets	○	○
Permafrost	○	○
Albedo	○	○
Land cover (including vegetation type)	○	○
Fraction of absorbed photosynthetically active radiation (FAPAR)	○	○
Leaf area index (LAI)	○	○
Above-ground biomass	○	○
Soil carbon	○	○
Fire disturbance	○	○
Soil moisture	○	○
Other variable, please specify ____	○	○

5. If you indicated that in some case(s) you use reanalysis data, from which reanalysis dataset(s) is that? (please choose all that apply)

Atmospheric reanalyses:

- ☐ Global ECMWF Interim Reanalysis (ERA-Interim)
- ☐ Global ECMWF 40 year Reanalysis (ERA-40)
- ☐ Global ECMWF Reanalysis (ERA-15)
- ☐ Global Japanese Reanalysis (JRA-25) or its continuation as JMA Climate Data Assimilation
- ☐ System (JCDAS)
- ☐ Global NASA MERRA
- ☐ Global NCEP Climate Forecast System Reanalysis (CFSR)
- ☐ Global NCEP/NCAR Reanalysis I (R1) (1948 to present)
- ☐ Global NCEP/DOE Reanalysis AMIP-II (R2) (1979 to near present)
- ☐ Global NOAA-CIRES 20th Century Reanalysis (20CR)
- ☐ Arctic System Reanalysis (ASR)
- ☐ NCEP North American Regional Reanalysis (NARR)
- ☐ The European Reanalysis and Observations for Monitoring project (EURO4M)
- ☐ Global MACC Reanalysis 2003-2010
- ☐ Global GEMS Reanalysis 2003-2007
- ☐ Other, please specify

Ocean reanalyses:

- ARMOR3D
- BLUElink
- CERFACS
- CFSR
- CMCC
- DePreSys
- ECMWF Ocean ReAnalysis ORA-S3
- ECMWF Ocean Reanalysis System 4 (ORAS4)
- ECMWF Ocean ReAnalysis ORA-XBTc
- ECMWF NEMOVAR COMBINE
- ECCO-GODAE MIT/AER
- ECCO-JPL
- ECCO-SIO
- ECCO2-cube78
- EN3
- FNMOC/GODAE
- FOAM Global
- FOAM North Atlantic
- FOAM Mediterranean Sea
- GECCO
- GECCO2
- GFDL
- GloSea
- GMAO
- G2V1
- HYCOM
- INGV
- K-7
- K7-ODA
- K7CDA
- LEGOS
- MERCATOR-2 North Atlantic
- MERCATOR-2 Global
- MERCATOR-3 Global

- ☐ MOVE/MRI.COM-C
- ☐ MOVE/MRI.COM-CORE
- ☐ NCEP/GODAS
- ☐ NODC
- ☐ PEODAS
- ☐ POAMA
- ☐ PSY3V3
- ☐ SODA 2.1.6
- ☐ SODA 2.2.4
- ☐ UR025.4
- ☐ Other, please specify

Applications and methods

6. Please indicate below the application(s) that you work with or the method(s) that you use. Please also indicate your source of data, stating whether you use
• reanalysis products;
• weather station, radiosonde or other in-situ observations;
• satellite-based remote sensing data; or
• weather radar based remote sensing data
(please choose all that apply).

	Reanalysis data	In-situ station data	In-situ gridded data	Satellite data	Weather radar data
Studies of atmospheric dynamics					
Studies of atmospheric physics					
Atmospheric modelling					
Climate modelling					
Production of remote sensing -based data sets					
Production of in-situ-based data sets					
Short-term forecasting					
Monthly to seasonal scale forecasting					
Verification of forecasts					
Meteorological case studies					
Climate monitoring					
Operational climate service					
Climate change detection					
Evaluation of climate models					
Climate change projections					
Climate engineering					
Climate change impacts					
Impacts of variability and extremes in weather					
Time series analyses					
Extreme value analyses					
Climatological means determination					
Climate variability characterization					
Visualization of climate data for the media					
Resource assessment for renewable energies					
Assessment of energy use/ demand					
Opening of new shipping/ trade routes					
Consultancy / decision making					
Other, please specify					

7. Have you used reanalysis input observations and feedback data? (please choose all that apply)

- I do not know what this is about
- I have used it to assess the reanalysis data using observations as a reference
- I have used it to assess the observations using reanalysis data as a reference
- I have used it to merge the observations and reanalysis data together to create an improved product
- I have used it to understand how the observations had been used by reanalysis
- I could not find it
- I have had no time or resources to look into it.
- I have not used it because the data files are too big
- I have not used it because the data formats are too complicated
- I have not used it because there is no easy interface to get these data
- I have no interest in using this data

User awareness and needs

8. Regarding the characteristics of reanalysis data, would you say that (1= fully agree, 5= fully disagree)

	1	2	3	4	5
The data is easy to access					
The data policy is too strict					
The file sizes are too large to work with					
The data can be imported easily by my software application					
The horizontal nominal (grid cell size) resolution is adequate					
The vertical nominal (grid cell size) resolution is adequate					
The temporal nominal resolution is adequate					
For the climate variables I need, I know how much their spatial true (feature) resolution differs from the nominal resolution					
I know how much the temporal true (feature) resolution differs from the nominal resolution in time					
The spatio-temporal scales that I need are well represented					
The time period covers my interests					
The data are consistent between the variables					
The general quality is good enough for my needs					
The biases compared with observations are small enough					
Time-varying biases make the data too instable for my needs					
The temporal continuity is adequate					
The uncertainties are well characterized					
The observation input to reanalysis are clearly explained					
I know enough to work with the data					
The data tend to become available too late for my needs					
Plentiful training material is available on the web					
The literature provides good information					
Websites provide good information					

Free comments

9. Regarding reanalysis data, do you have any free comments or suggestions related to your needs (length 100 marks and high 10 rows)

Future climate services

10. Regarding future climate services, do you think, that they should include the following tasks or activities (1= fully agree, 5= fully disagree)

The overall goal is to improve quality and quantity of future climate services.

	1	2	3	4	5
Provision of statistics based on observations	○	○	○	○	○
Provision of statements describing past weather events	○	○	○	○	○
Statistical impact analyses for improving weather warnings and their criteria	○	○	○	○	○
Climate watch bulletins	○	○	○	○	○
Monthly forecasting and verification	○	○	○	○	○
Seasonal forecasting and verification	○	○	○	○	○
Production of long-term climate projections	○	○	○	○	○
Detection of climate change	○	○	○	○	○
Attribution of climate change	○	○	○	○	○
Research and communication of climate change uncertainties	○	○	○	○	○
Interpolation and production of gridded data sets based on observations	○	○	○	○	○
Homogenization of weather station data	○	○	○	○	○
Applied weather and climate research for impact assessment	○	○	○	○	○
Climate change impact consultancy for decision makers	○	○	○	○	○
24/7 updates, in the internet, of statistics of weather and climate	○	○	○	○	○

11. Regarding future climate services and reanalyses, do you have any free specifications, comments or suggestions related to your needs (length 100 marks and high 10 rows)

12. Optional: You may wish to leave your contact information

Name

Email

Phone or Mobile

Company / Organization

Department

Address

ZIP / Post code

City

Country

APPENDIX C: Survey analysis: "ERA-Interim users" and "not ERA-users"

C.1 Background of the respondents

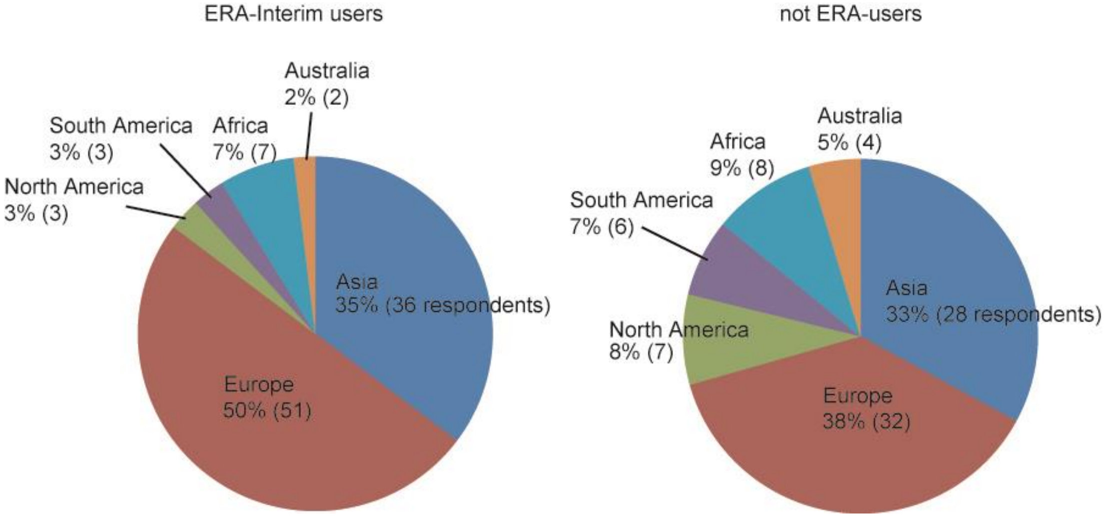


Figure C.1. Location of the respondents who left their contact info (102 ERA-Interim users, 85 not ERA-users) divided into continents.

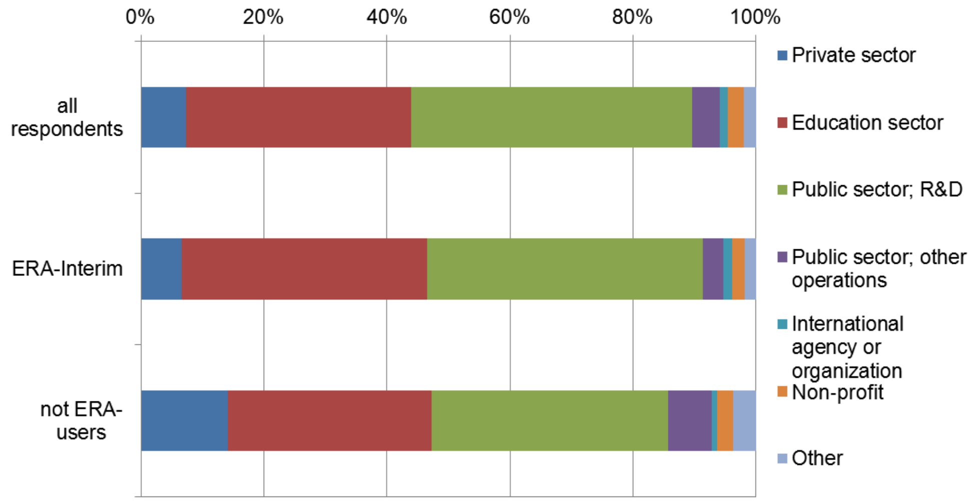


Figure C.2. Sectors of work of the respondents (all 2568, ERA-Interim users 391, not ERA-users 239).

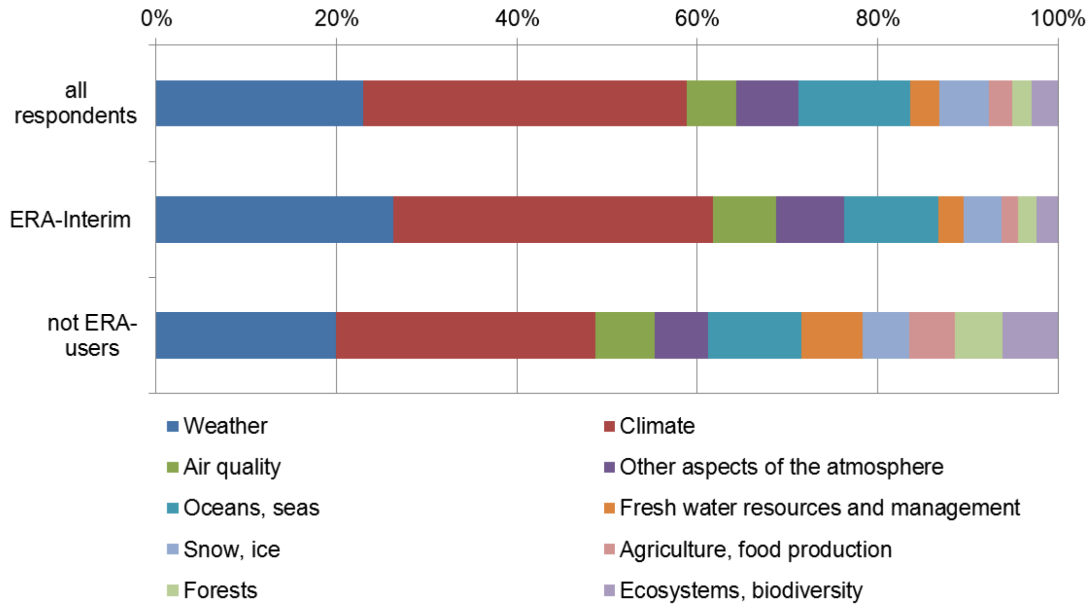


Figure C.3. Ten most popular topics (in respect to all respondents) that best describe the respondents' field or subject of work. Number of respondents: all 2573, ERA-Interim users 391, not ERA-users 239. It was asked to choose all that apply.

Table C.1. Division of respondents' field or subject of work.

	% of all respondents	% of ERA-Interim users	% of not ERA-users
Weather	47	45	38
Climate	73	61	54
Air quality	11	12	12
Other aspects of the atmosphere	14	13	11
Oceans, seas	25	18	20
Fresh water resources and management	7	5	13
Snow, ice	11	7	10
Agriculture, food production	5	3	10
Forests	4	4	10
Ecosystems, biodiversity	6	4	12
Erosion / flooding of coastal areas	4	3	5
Energy	10	7	15
Industry	2	2	1
Transportation (land, air, marine)	3	3	5
Economics	1	1	4
Insurance	1	1	1
Architecture, urban or other spatial design/management	1	0	3
Construction and municipal engineering	1	1	3
Health, human well-being	2	2	6
Tourism, recreation	1	0	2
Safety and security issues, disasters	4	2	4
Indigenous peoples' issues	0	0	1
Other social aspects	0	0	0
Communication	1	0	1
Geophysics, geochemistry	7	7	7
Geoinformatics	3	3	5
Geology	2	1	3
Geography	6	4	10
Other, please specify	5	8	10

Table C.2. Division of respondents' regional focus of work.

	% of all respondents	% of ERA-Interim users	% of not ERA-users
The whole globe	32	28	14
Mostly continents	6	4	5
Mostly oceans	8	7	5
Asia	24	19	18
Africa	8	10	12
Europe	26	27	35
North America	8	6	8
Middle East	4	3	5
Oceania / Australia / New-Zealand	4	4	2
Central America / Carribean	2	1	1
South America	7	7	8
Arctic	8	6	4
Antarctic	5	6	2
Atlantic Ocean	8	7	7
Pacific Ocean	9	5	7
Indian Ocean	8	6	6
Specific countries, please specify	7	7	13
Other, please specify	5	6	8

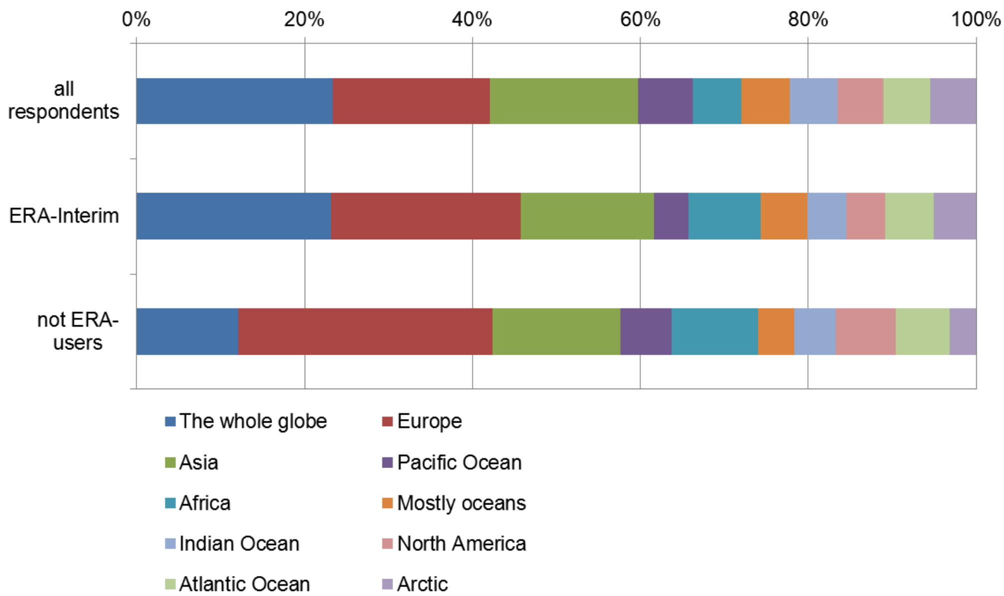


Figure C.4. Ten most common regions of the world (in respect to all respondents) on which the respondents focused. Number of respondents: all 2574, ERA-Interim users 391, not ERA-users 240. It was asked to choose all that apply.

C.2 Reanalysis data and Essential Climate Variables

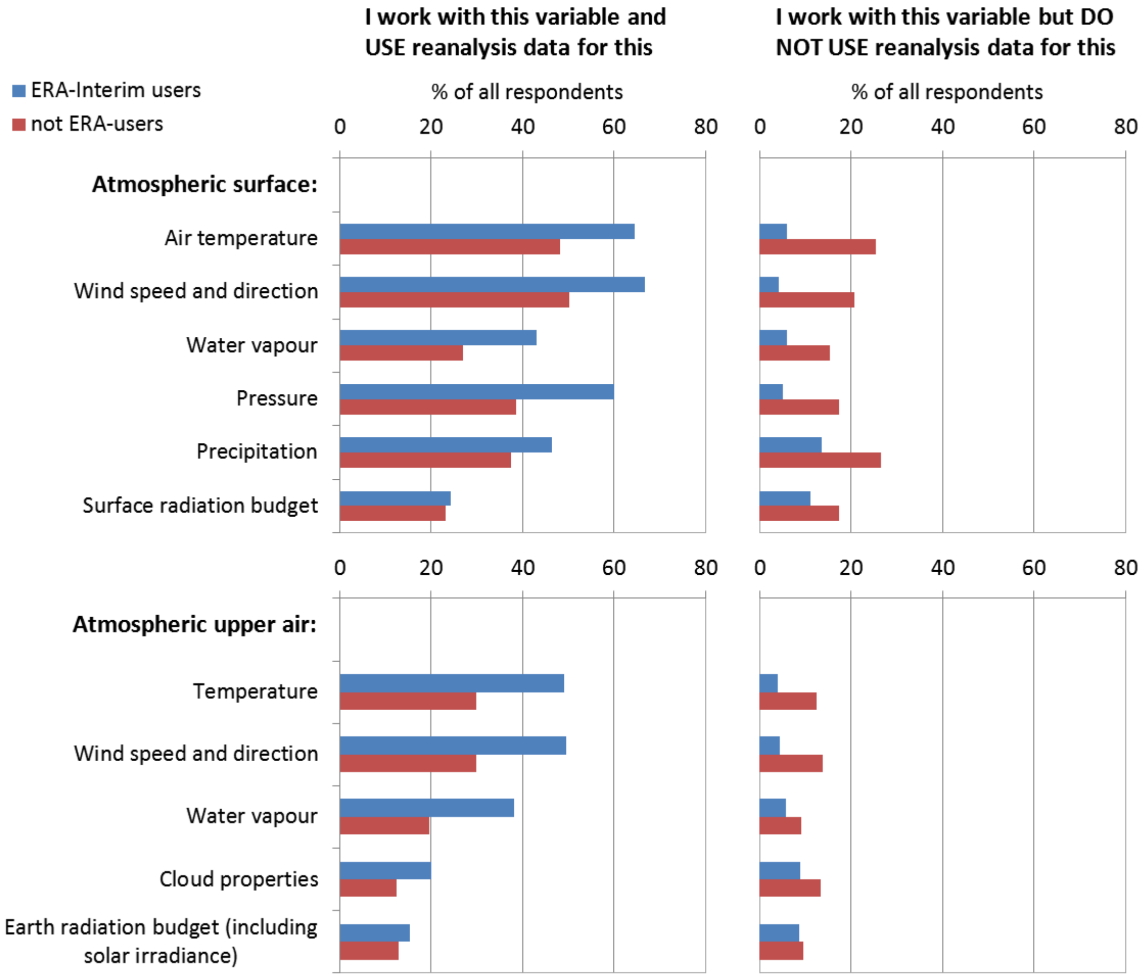


Figure C.5. Use of reanalysis data for working with different atmospheric surface and upper air variables within ERA-Interim users and not ERA-users.

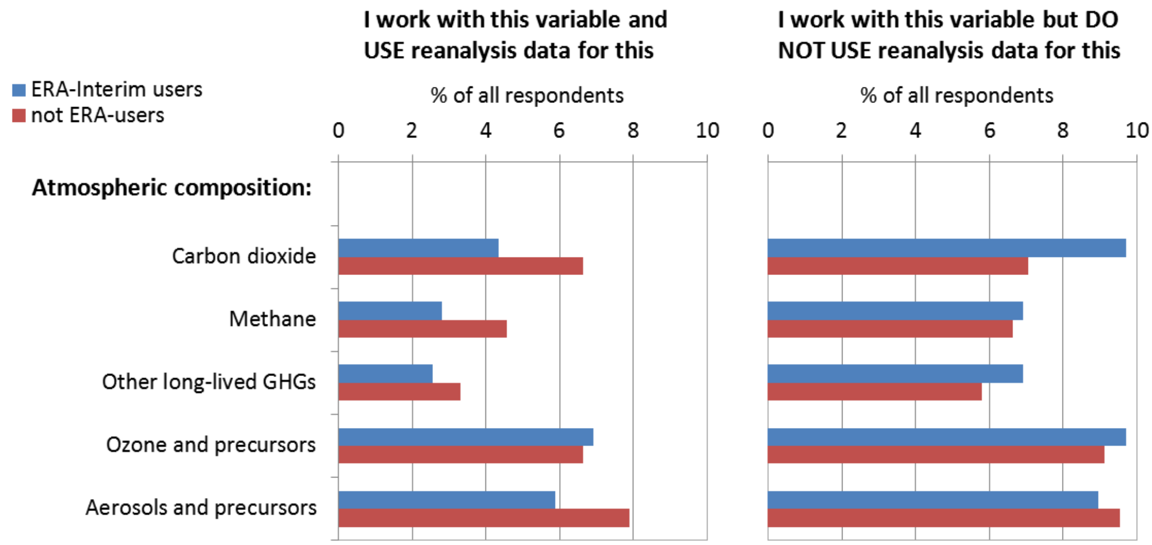


Figure C.6. Same as Fig. C.5 but for atmospheric composition variables.

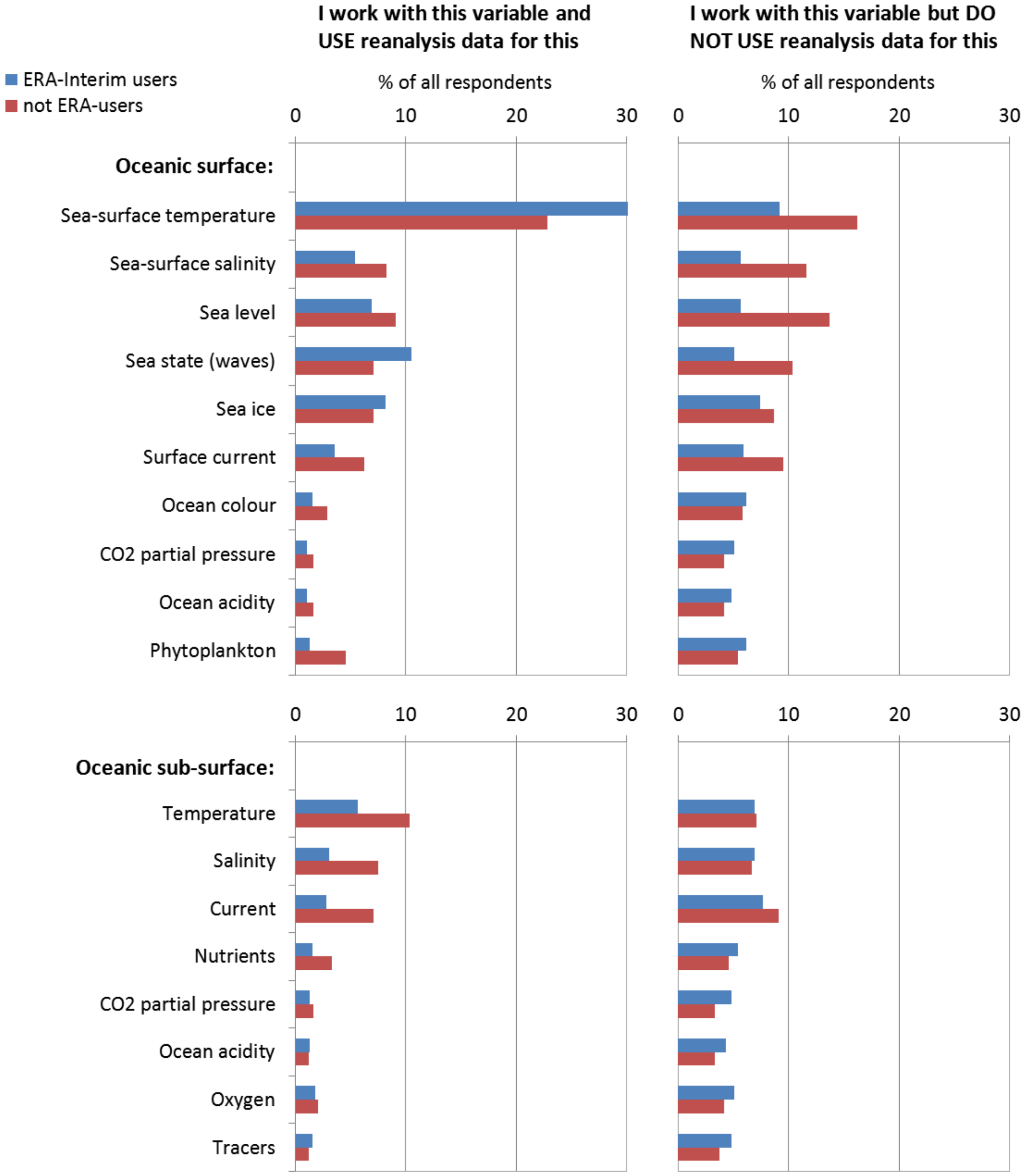


Figure C.7. Same as Fig. C.5 but for oceanic surface and sub-surface variables.

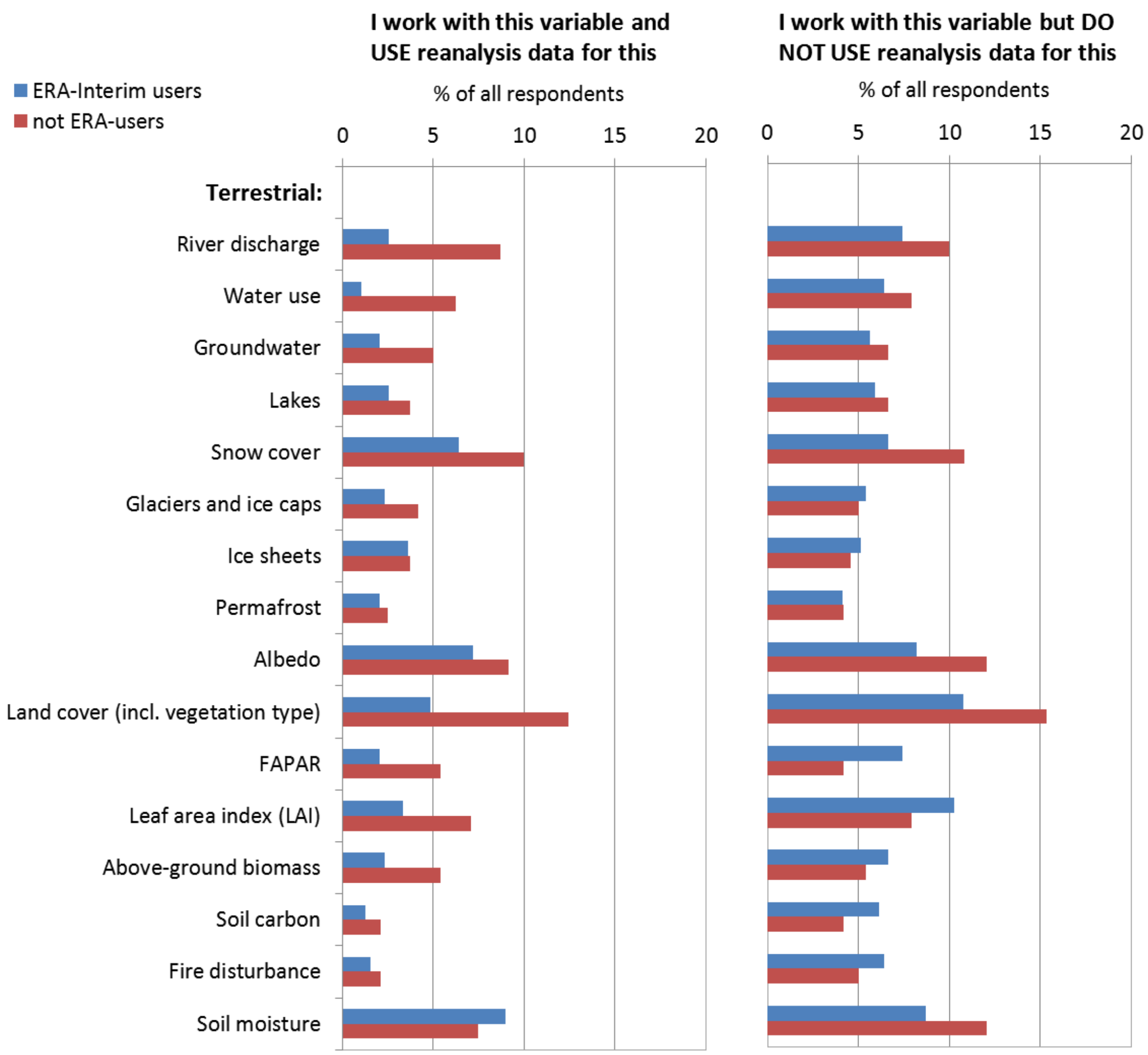


Figure C.8. Same as Fig. C.5 but for terrestrial variables.

C.3 Applications and methods

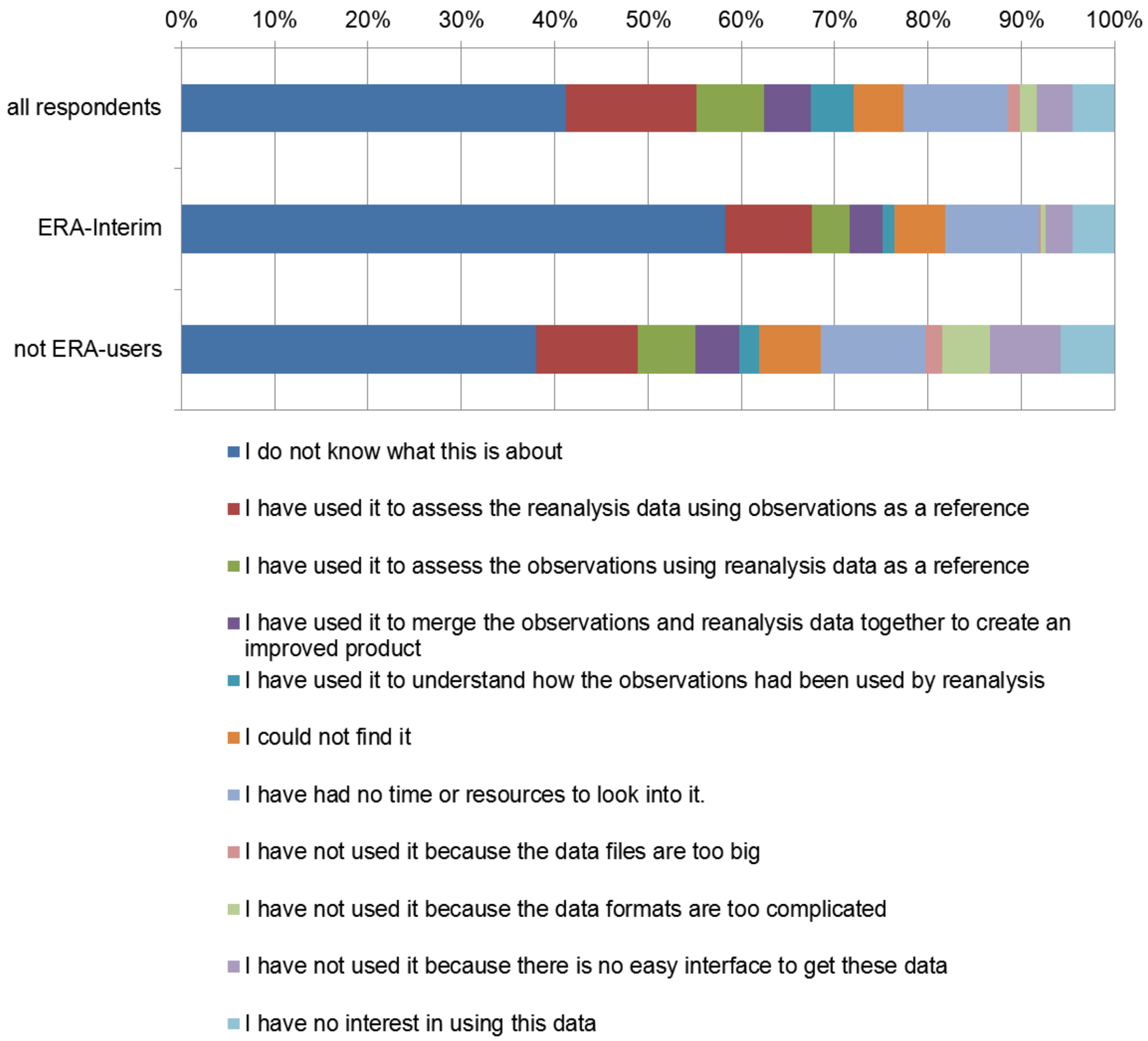


Figure C.9. Responses to the use of reanalysis input observations and feedback data. Number of respondents: all 2473, ERA-Interim users 372, not ERA-users 216. It was asked to choose all that apply.

APPENDIX D: Survey analysis: "best-informed users" and "least-informed users"

D.1 Background of the respondents

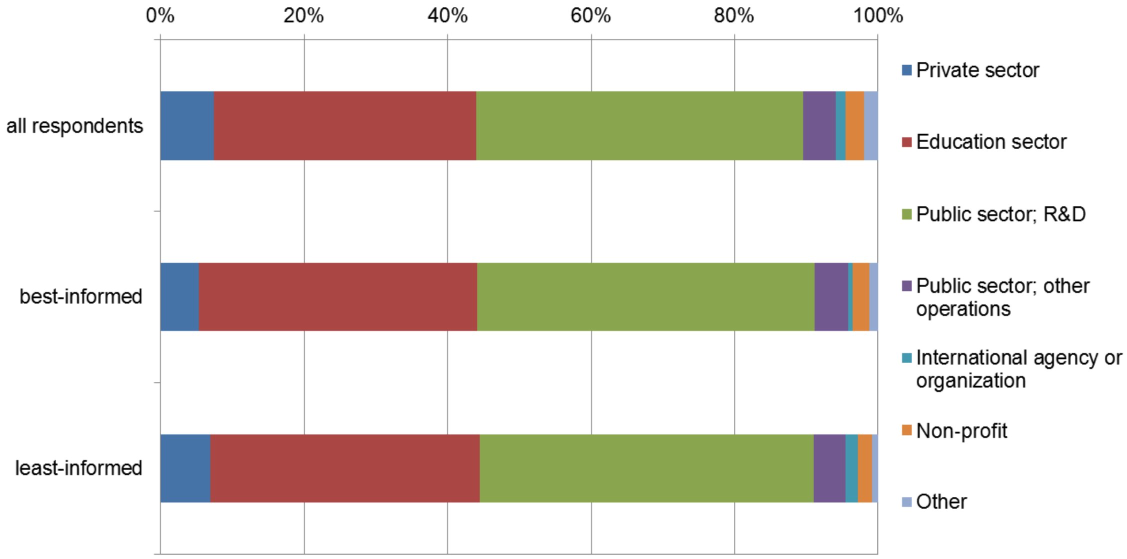


Figure D.1. Sectors of work of the respondents (all 2568, best-informed 170, least-informed 245).

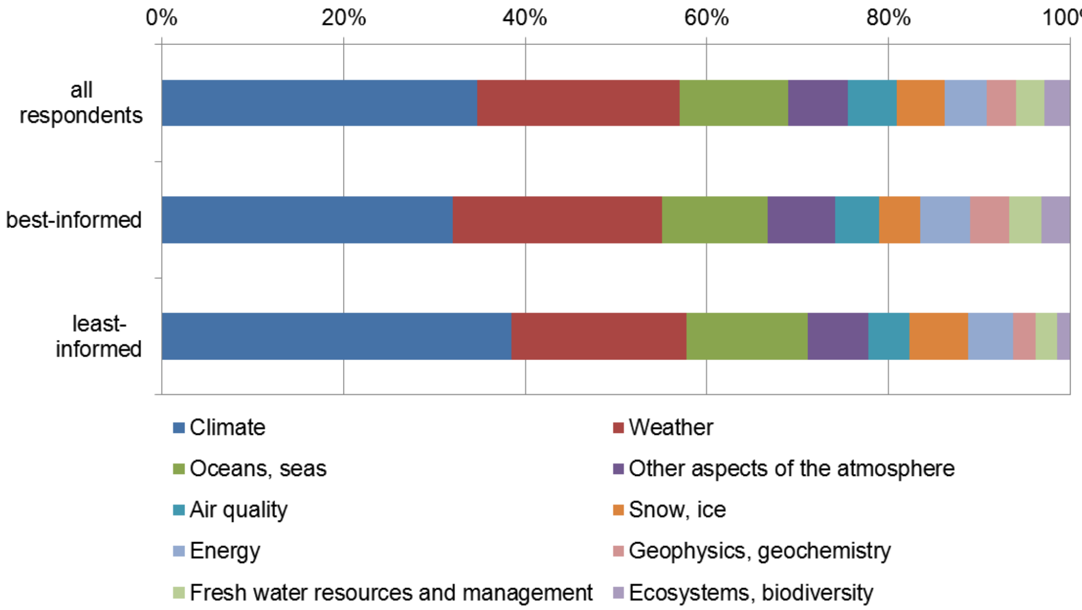


Figure D.2. Ten most popular topics (in respect to all respondents) that best describe the respondents' field or subject of work. Number of respondents: all 2573, best-informed 170, least-informed 245. It was asked to choose all that apply.

Table D.1. Division of respondents' field or subject of work.

	% of all respondents	% of best-informed users	% of least-informed users
Weather	47	56	38
Climate	73	79	76
Air quality	11	12	9
Other aspects of the atmosphere	14	18	13
Oceans, seas	25	29	26
Fresh water resources and management	7	9	4
Snow, ice	11	11	13
Agriculture, food production	5	9	4
Forests	4	8	3
Ecosystems, biodiversity	6	8	3
Erosion / flooding of coastal areas	4	6	3
Energy	10	14	10
Industry	2	4	1
Transportation (land, air, marine)	3	7	1
Economics	1	5	1
Insurance	1	4	2
Architecture, urban or other spatial design/management	1	2	1
Construction and municipal engineering	1	2	1
Health, human well-being	2	4	2
Tourism, recreation	1	3	0
Safety and security issues, disasters	4	5	3
Indigenous peoples' issues	0	1	0
Other social aspects	0	1	0
Communication	1	2	1
Geophysics, geochemistry	7	11	5
Geoinformatics	3	4	2
Geology	2	4	2
Geography	6	8	6
Other, please specify	5	4	5

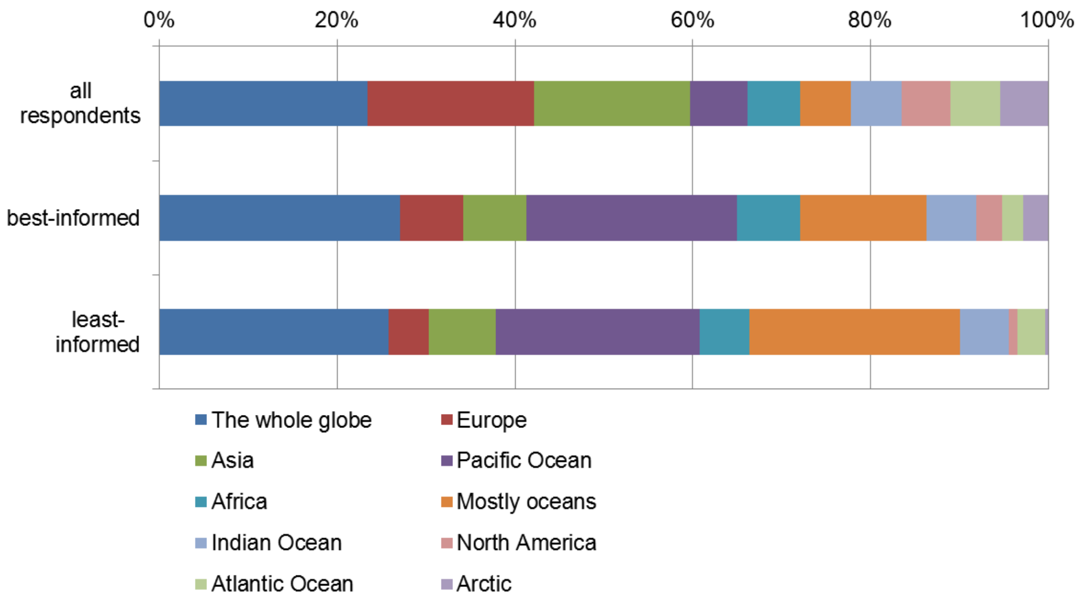


Figure D.3. Ten most common regions of the world (in respect to all respondents) on which the respondents focused. Number of respondents: all 2574, best-informed 170, least-informed 245. It was asked to choose all that apply.

Table D.2. Division of respondents' regional focus of work.

	% of all respondents	% of best-informed users	% of least-informed users
The whole globe	32	34	31
Mostly continents	6	3	7
Mostly oceans	8	18	28
Asia	24	9	9
Africa	8	9	7
Europe	26	9	5
North America	8	4	1
Middle East	4	13	7
Oceania / Australia / New-Zealand	4	11	7
Central America / Carribean	2	4	4
South America	7	7	9
Arctic	8	4	0
Antarctic	5	8	7
Atlantic Ocean	8	3	4
Pacific Ocean	9	29	27
Indian Ocean	8	7	7
Specific countries, please specify	7	11	7
Other, please specify	5	11	10

D.2 Reanalysis data and Essential Climate Variables

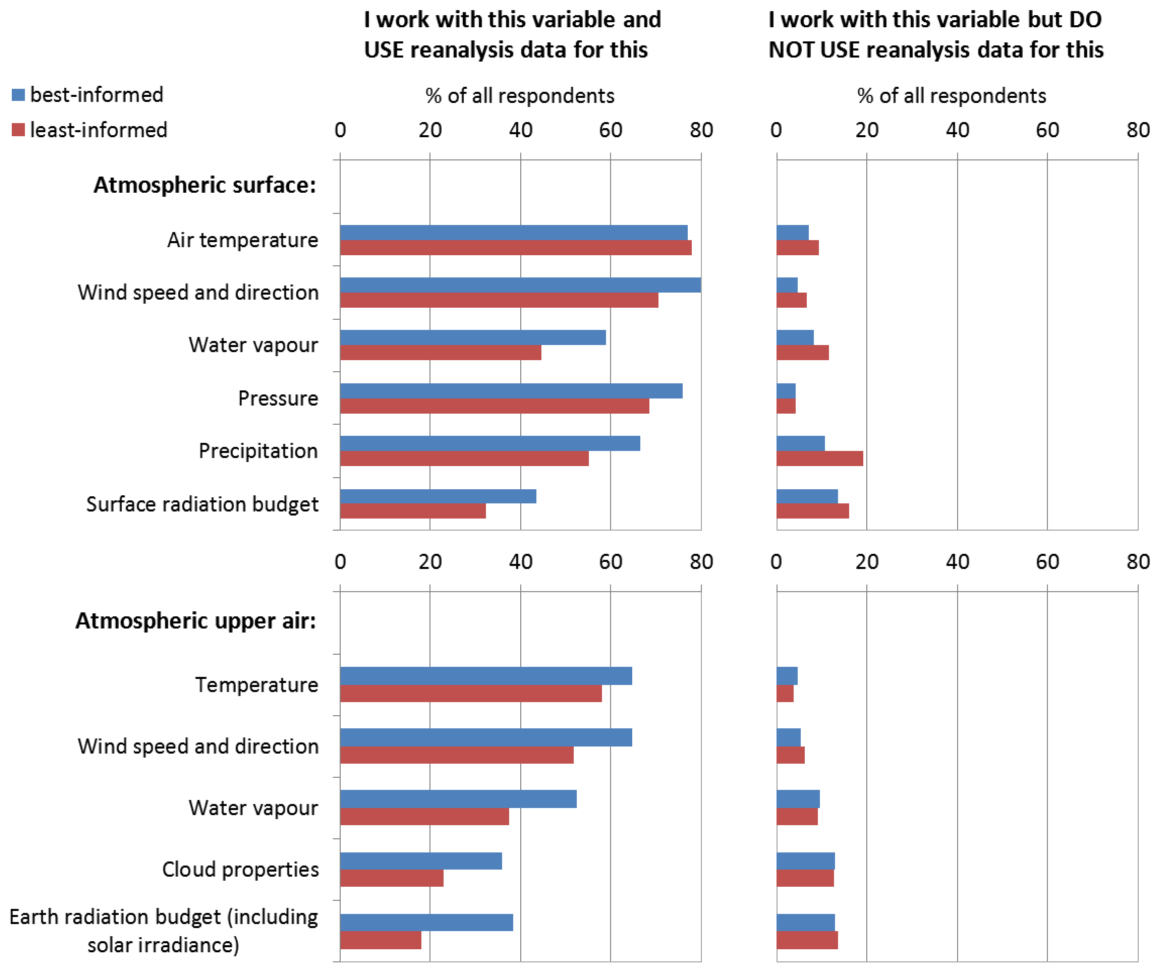


Figure D.4. Use of reanalysis data for working with different atmospheric surface and upper air variables within best- and least-informed users.

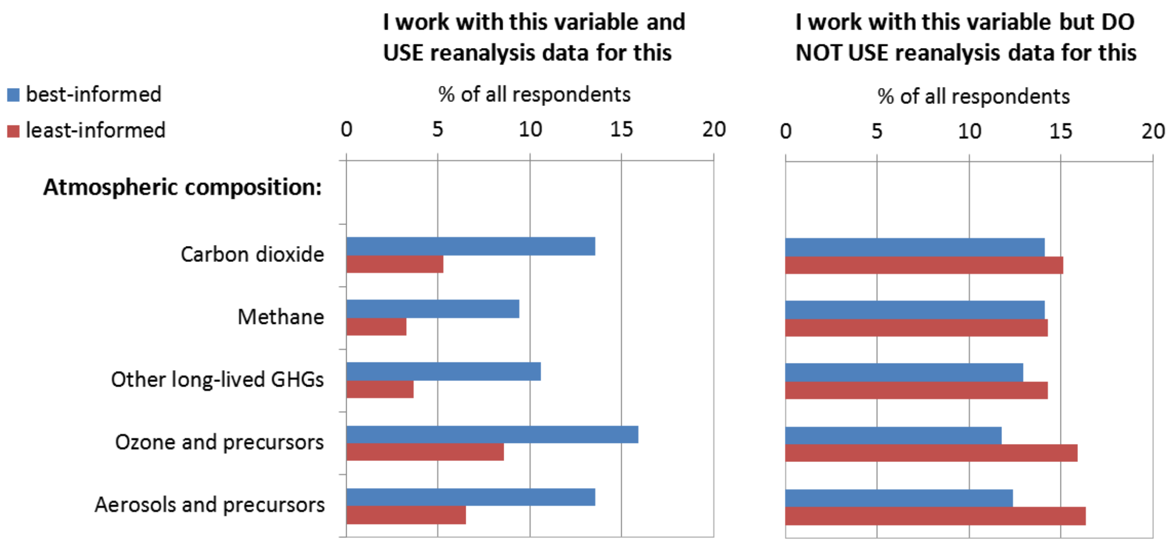


Figure D.5. Same as Fig. D.4 but for atmospheric composition variables.

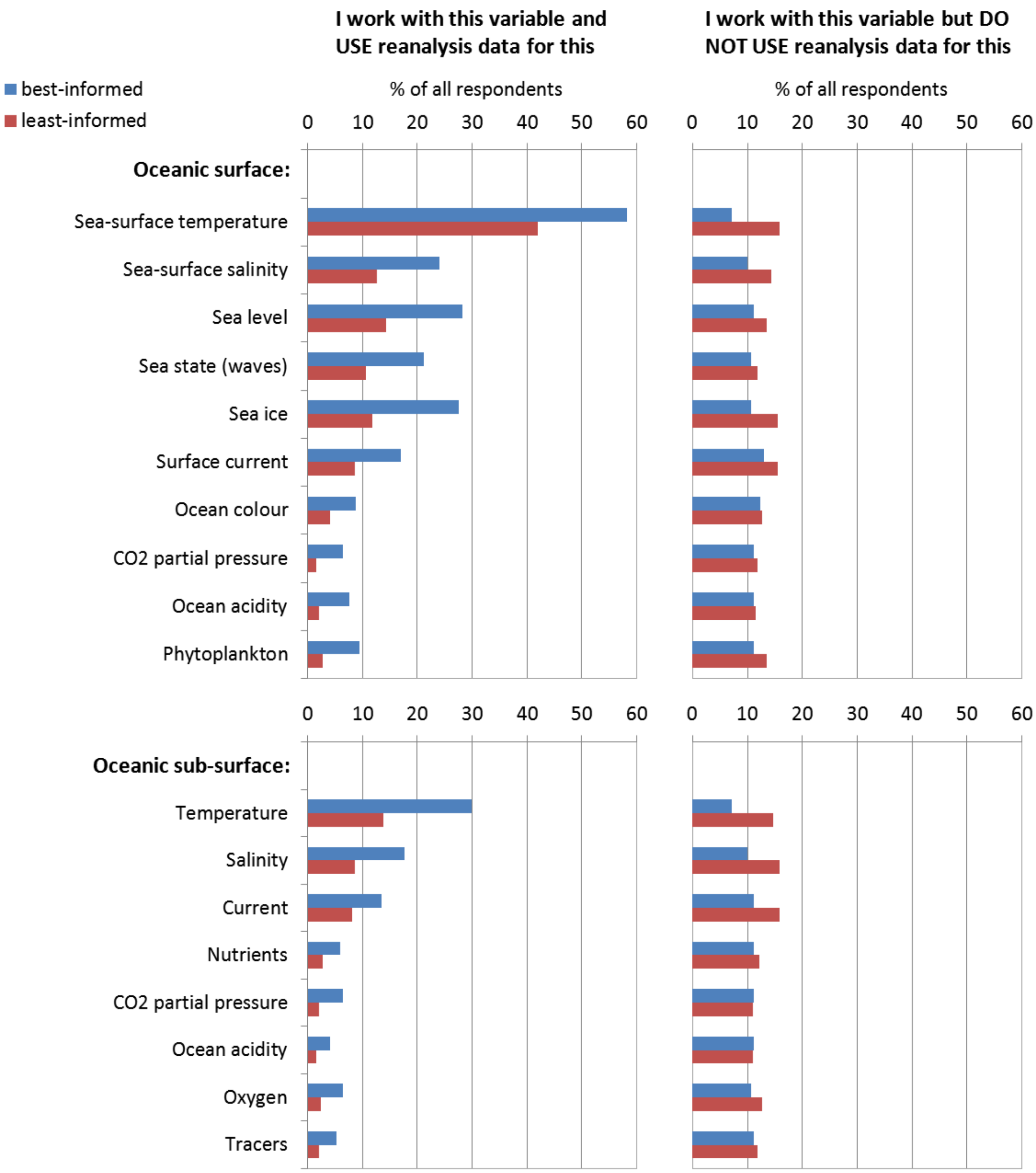


Figure D.6. Same as Fig. D.4 but for oceanic surface and sub-surface variables.

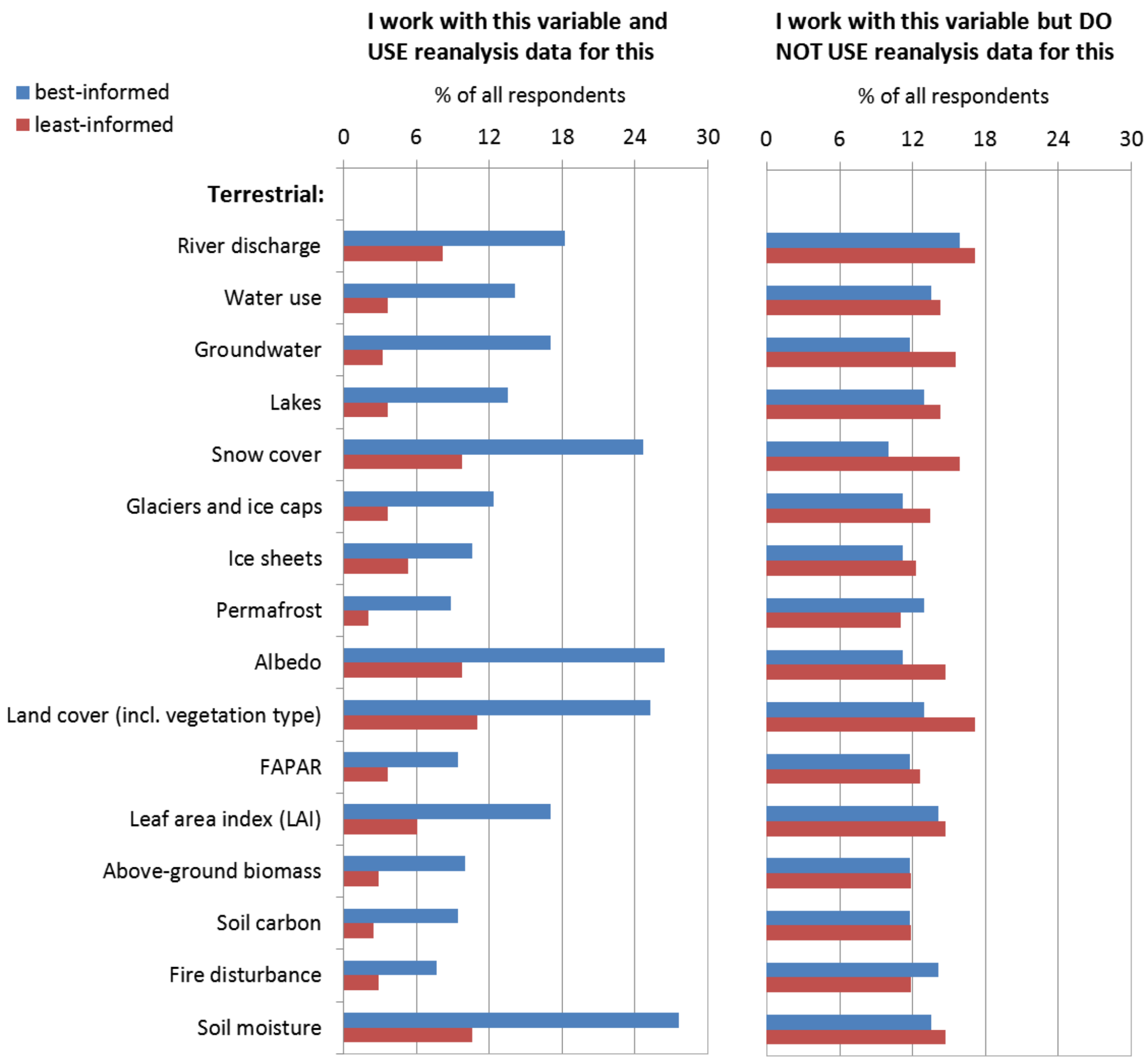


Figure D.7. Same as Fig. D.4 but for terrestrial variables.

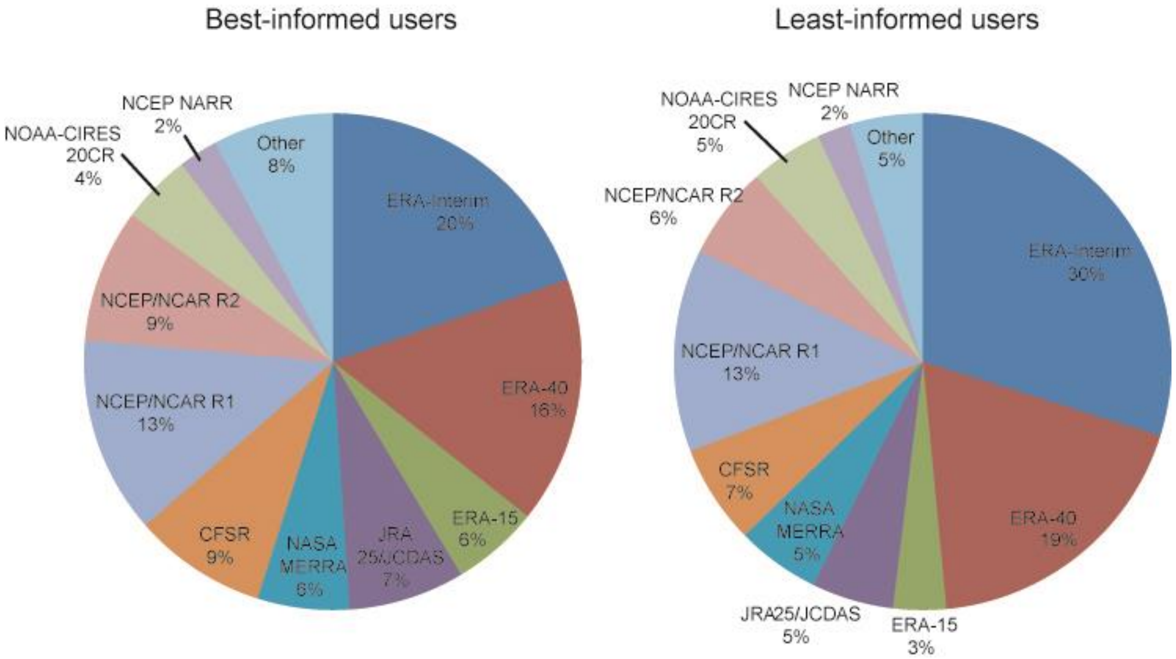


Figure D.8. Proportion of atmospheric reanalysis datasets used by the best- and least-informed users. Note that the percentage values are calculated from the total number of responses given for all the different reanalysis options, not from the number of respondents (best-informed users 170, least.-informed users 242). It was asked to choose all that apply.

D.3. Applications and methods

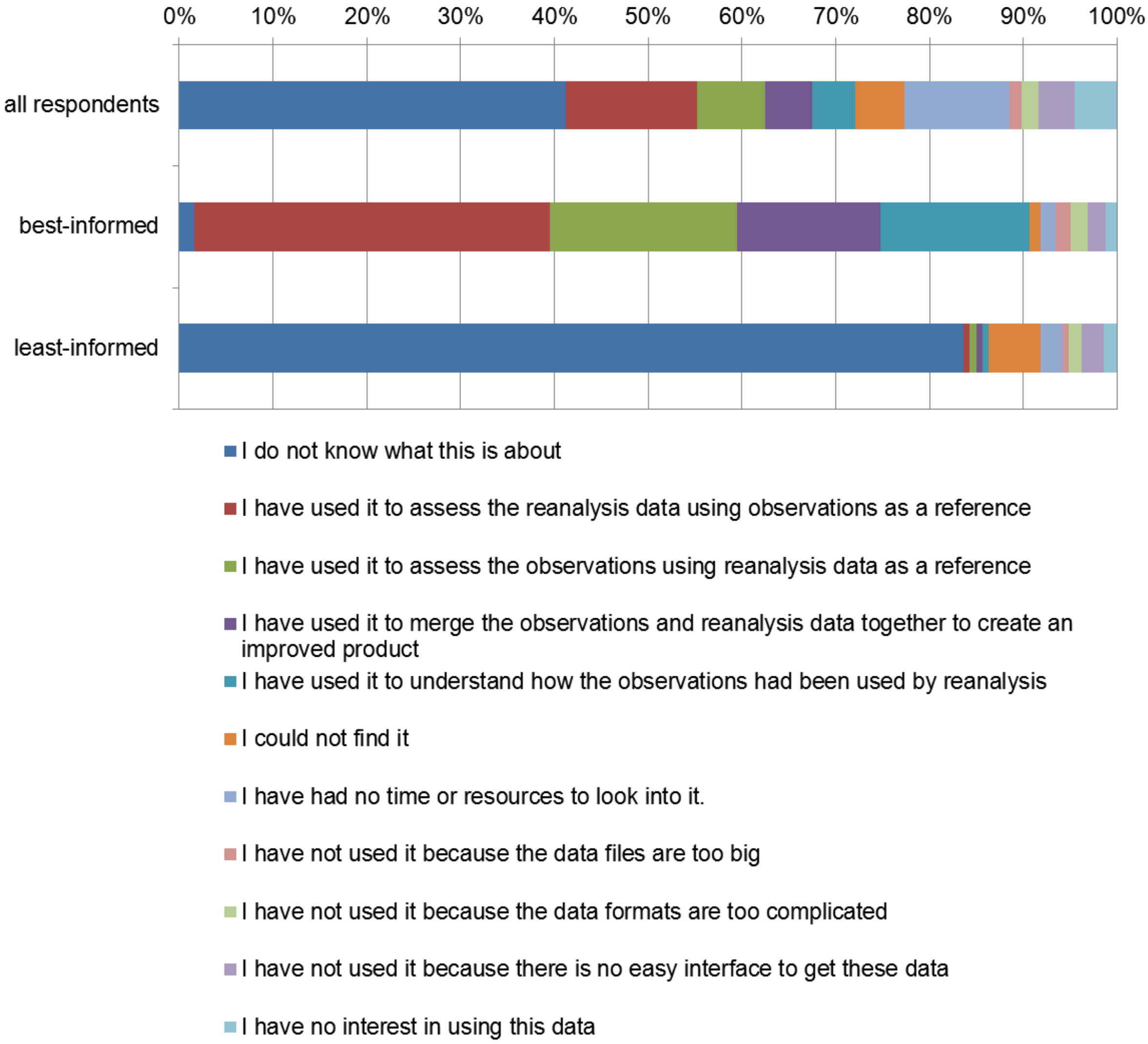


Figure D.9. Responses to the use of reanalysis input observations and feedback data. Number of respondents: all 2473, best-informed users 170, least-informed users 245. It was asked to choose all that apply.