

2007 User Survey Analysis

Introduction.

In September and October 2007 the BADC carried out an online survey to try and find out more about how the BADC user community found the present services as well as attempt to gather information about the user community's skills and requirements. The information obtained from the survey should provide a useful point of reference for the BADC to help target its development and to ensure that it remains focused on providing a quality service to the UK atmospheric research community. In short to meet the BADC mission statement:

The British Atmospheric Data Centre (BADC) is the Natural Environment Research Council's (NERC) Designated Data Centre for the Atmospheric Sciences. The role of the BADC is to assist UK atmospheric researchers to locate, access and interpret atmospheric data and to ensure the long-term integrity of atmospheric data produced by NERC projects.

Prior to the Survey only circumstantial evidence was available on user experience and thus the BADC felt there was a need to sound the user community in a more formal manner.

The aim of the 2007 BADC User survey was three fold:

- 1) To determine the skills base of the BADC user community.
- 2) To determine the present experience of the BADC user community of its datasets and supporting services.
- 3) To identify those areas where the BADC should improve, those areas where the BADC is doing well and those areas where the BADC could explore/develop.

Approach and Methodology

As the BADC is a web based service an online user survey was deemed to be the most appropriate format for the BADC user survey. After some time exploring various options for the survey an account was opened with Smart-survey as the user interface was reasonably intuitive and was known to work across the main internet browsers.

To encourage users to fill in the survey it was launched along side a competition to win a PDA, with the launch coinciding with the 2007 Royal Meteorological Society Student and Main Conferences. The survey was advertised via BADC beer mats, a link on the BADC website and a series of email shots to the BADC email list (one 24 hours prior to the launch, one 1 half way through and one at the start of the last week). In addition a reminder was attached to the bottom of Footprints emails during the period of the

survey (Footprints is the BADC Query handling system used by the BADC Helpdesk).

Only 100 people were expected to enter the survey, while 200 was felt to be desirable. In actuality 311 people went onto the survey with 285 filling in the survey, with the remaining 26 only filling in the competition fields. It was noticeable that as well as a steady trickle of completed surveys there were increases in completed responses in the 24-36 hours following the RMetS conferences and the reminder email shots to the BADC lists. The greatest uptake in a 36 hour period occurred during the last week after the final email shot was sent out.

It is hard to say what impact running the competition along side the survey had on encouraging people to fill in the survey.

Representativeness of Sample.

Before the results of the survey are examined it is worth first determining the representativeness of the group that filled in the survey as a sample of BADC user population.

Due to the BADC's archive containing data that can be access anonymously it is impossible to have a true idea of the BADC user population, thus the closest estimate to the population must be examined. Such an estimated population can be examined by use of data from the BADC User Data Base (hereafter the UserDB). However, as the UserDB contains information back to 1996 and therefore information that is no longer relevant further assumptions must be made as to what forms a representative "population" from the UserDB.

For the purposes of this report the BADC User population is defined as:

"Those users who are registered with the BADC and have access to one or more restricted datasets in the past 12 months."

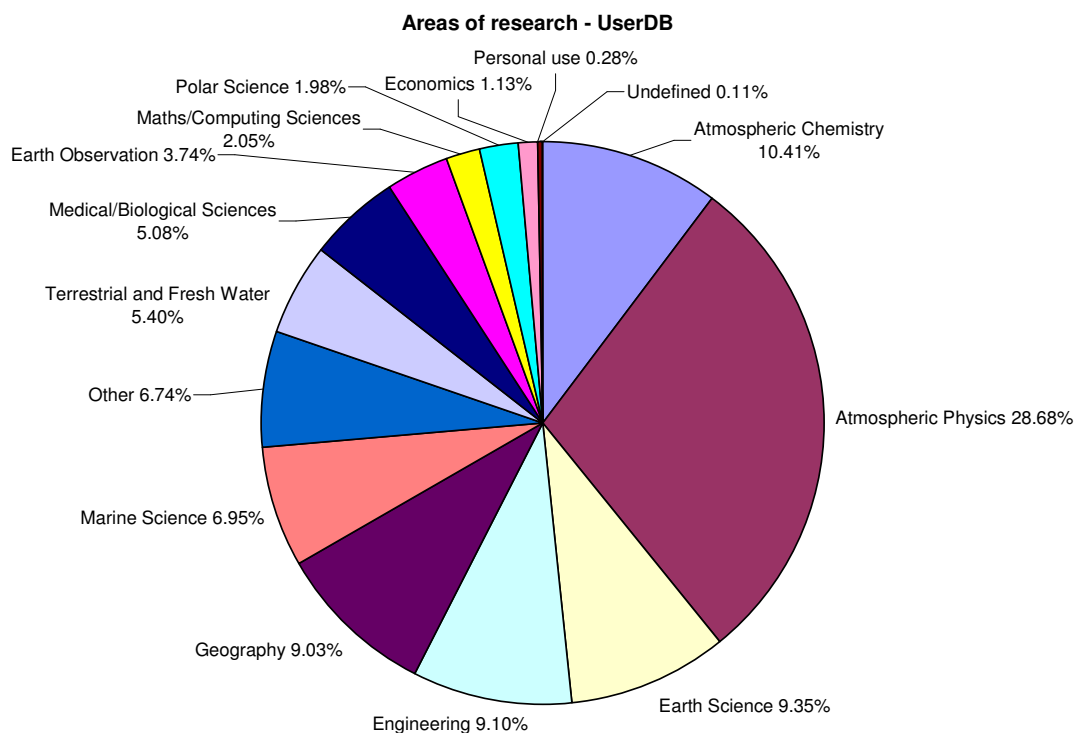
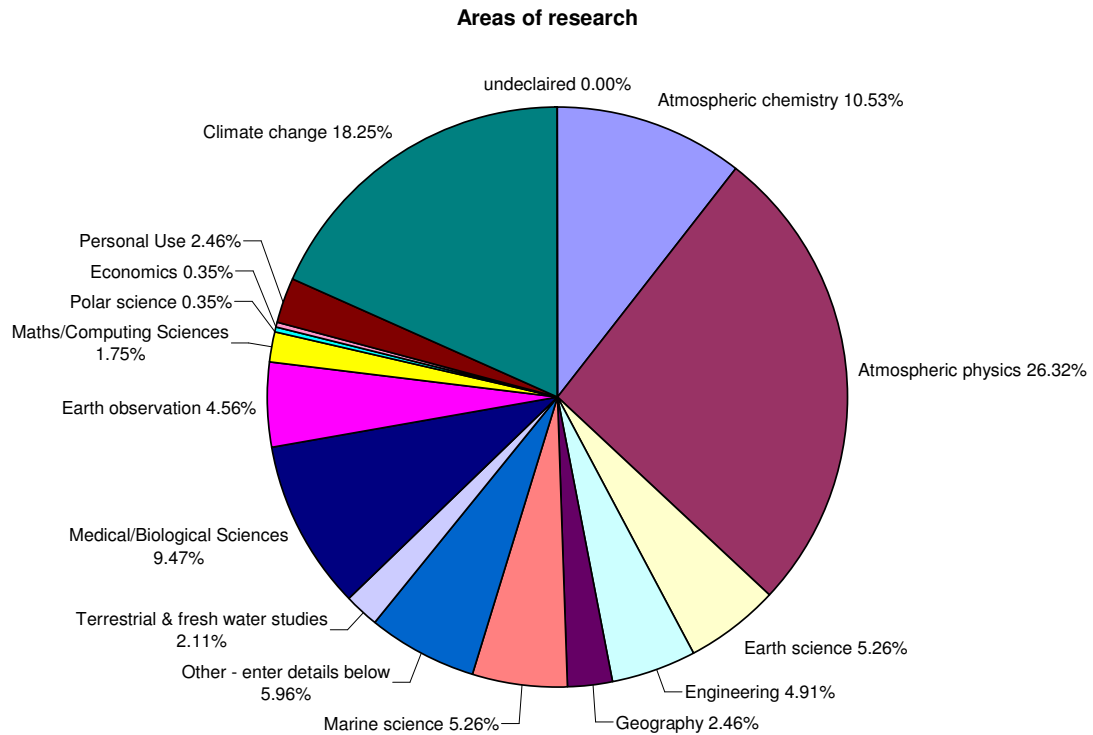
For the purposes of the survey results the sample will include all 285 responses that filled in the survey regardless of if they are registered with the BADC, not registered with the BADC or indicated that they didn't know if they were registered (87%, 7% and 6% respectively). For the survey's "target audience" subset, i.e. those to whom the service is provisioned, this will include those responses from :

"Those users who cite their research field as being atmospheric physics, atmospheric chemistry and climate change and are located in the UK"

Other subgroups will be defined below as and when appropriate.

Split of users by area of research

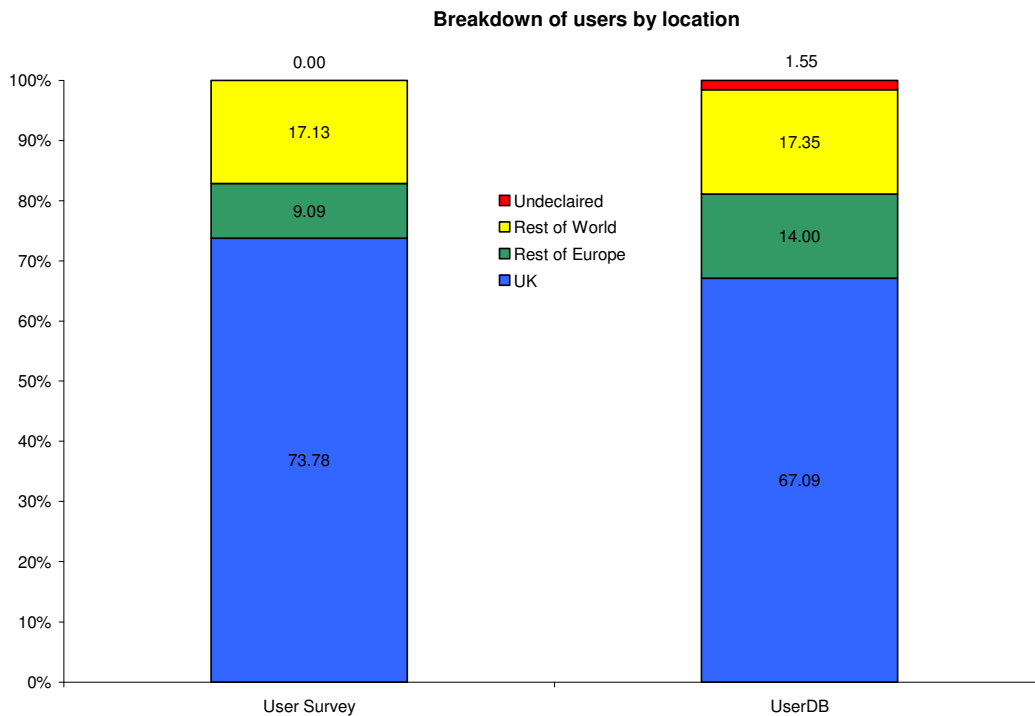
The two pie charts below show the breakdown by area of research for the survey respondents and the BADC user population. In the survey the options available for the area of research were based on those used in the UserDB, but expanded a little to split up the maths/computing science and marine/biological sciences and to add in a climate change option that was felt to be lacking from the UserDB. For the purposes of the comparison between the survey sample and the BADC user population the maths/computing science and marine/biological sciences groups have been re-merged. Climate change has been left in as this was an important group.



It is clear from this comparison that the atmospheric sciences were well captured in the survey and other fields were broadly represented in the sample. However, the large number of people who put their area of research as climate change is of importance and thus should be reviewed for inclusion on the UserDB list of options for people to choose in the future.

Split of users by location

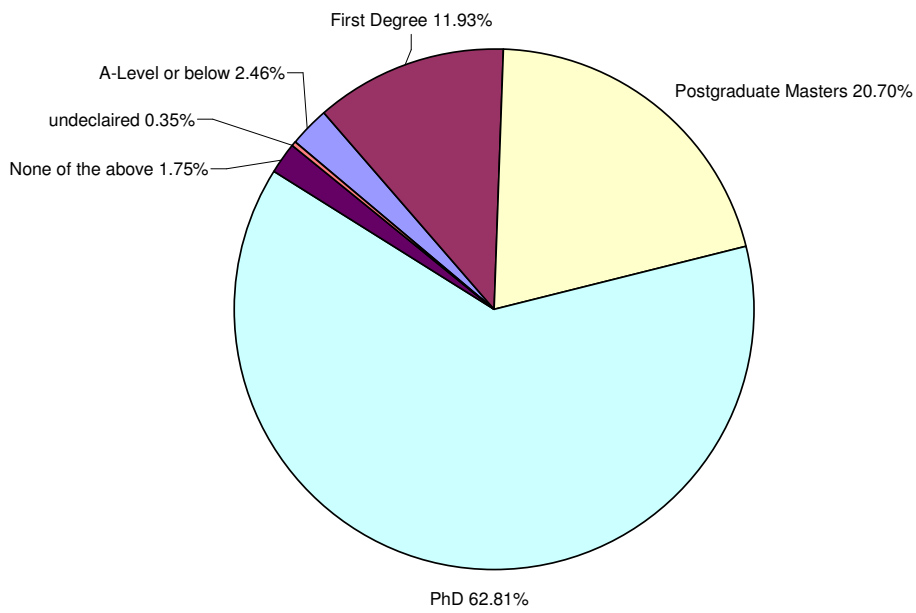
The following bar charts show that the split of geographical location of the survey sample agrees closely with that of the BADC user population.



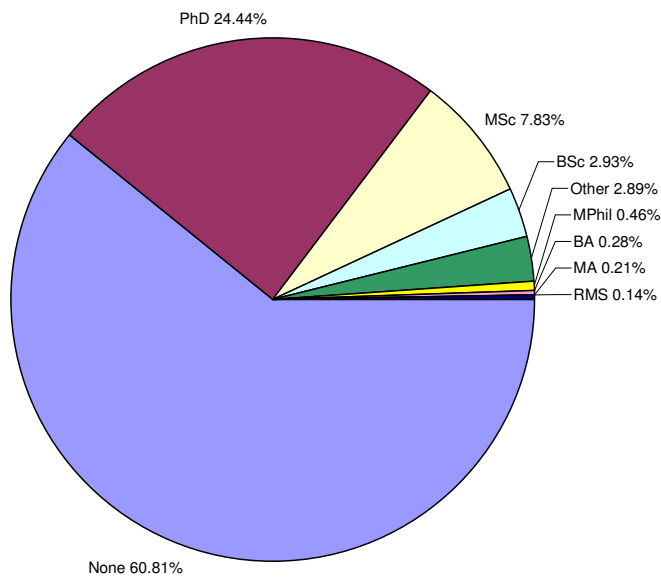
Split of users by level of education.

There does not exist a direct comparison between the survey results and the UserDB information due to the way the question was posed in the survey: in the survey users were asked for their level of current education while the userDB asks what study is being undertaken at present. Thus, a user may have a PhD and indicate so in the survey, but no longer actively studying therefore record themselves as not studying in the UserDB. However, comparing the two pie charts below it seems feasible to suggest that those in the UserDB as not studying are of PhD level or above and form the majority of the BADC's users. Second to this are those studying for a PhD at present.

Survey - level of education



UserDB - present study



Conclusion of representativeness.

From the above details it is clear that the sample returned by the User Survey is representative of the BADC user population and thus the analysis that follows may be taken as being applicable to the BADC users in general.

Survey Results.

The results shall be presented in the three broad sections of the survey:

- a) Users Skills
- b) User Experience of the BADC
- c) User desires for future development

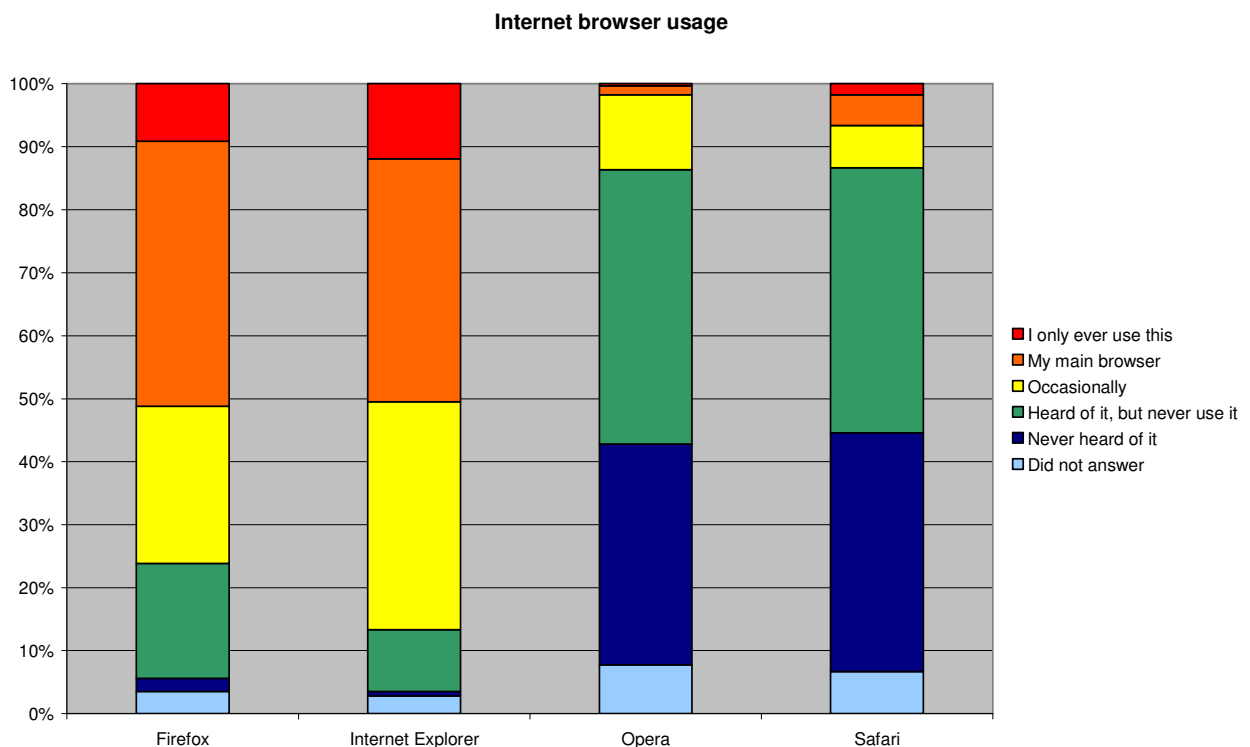
The User skills will examine the survey sample as a whole plus those of the following groups of users: the BADC “target” users, non BADC “target” users and those involved with projects whom the science support group liaise with on a regular basis. The other two sections will only consider the entire survey sample.

User skills

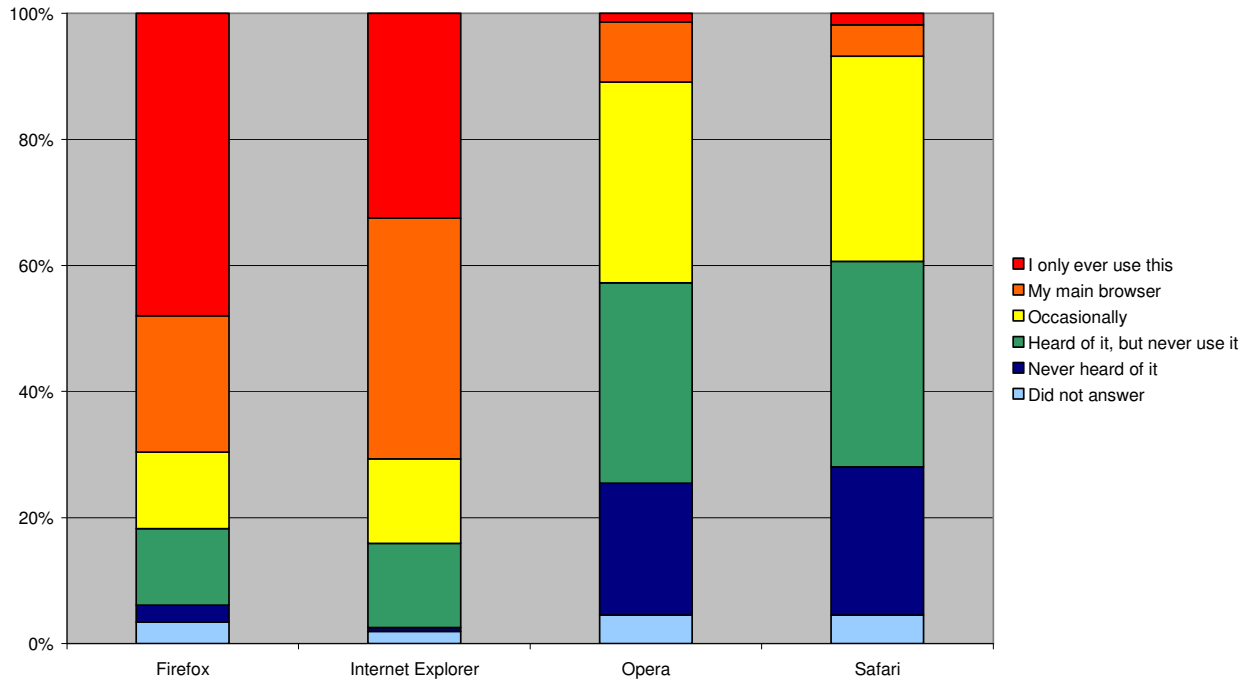
The main browsers used by the users and those that make up the various subgroups indicate that FireFox and IE are the principle browsers with their users showing a high degree of familiarity with them. The BADC target user group has a greater propensity towards choosing to use only FireFox or IE compared to the sample as a whole.

From those users who indicated that they used other browsers Konqueror, Lynx, Seamonkey and Maxthon were listed alongside Crasy Broswer, Camino, tt, Netscape and Mozilla.

While those listed under the “other” category were only small in numbers they none-the-less point to the need for the BADC services to be usable on a broad range of browser platforms.

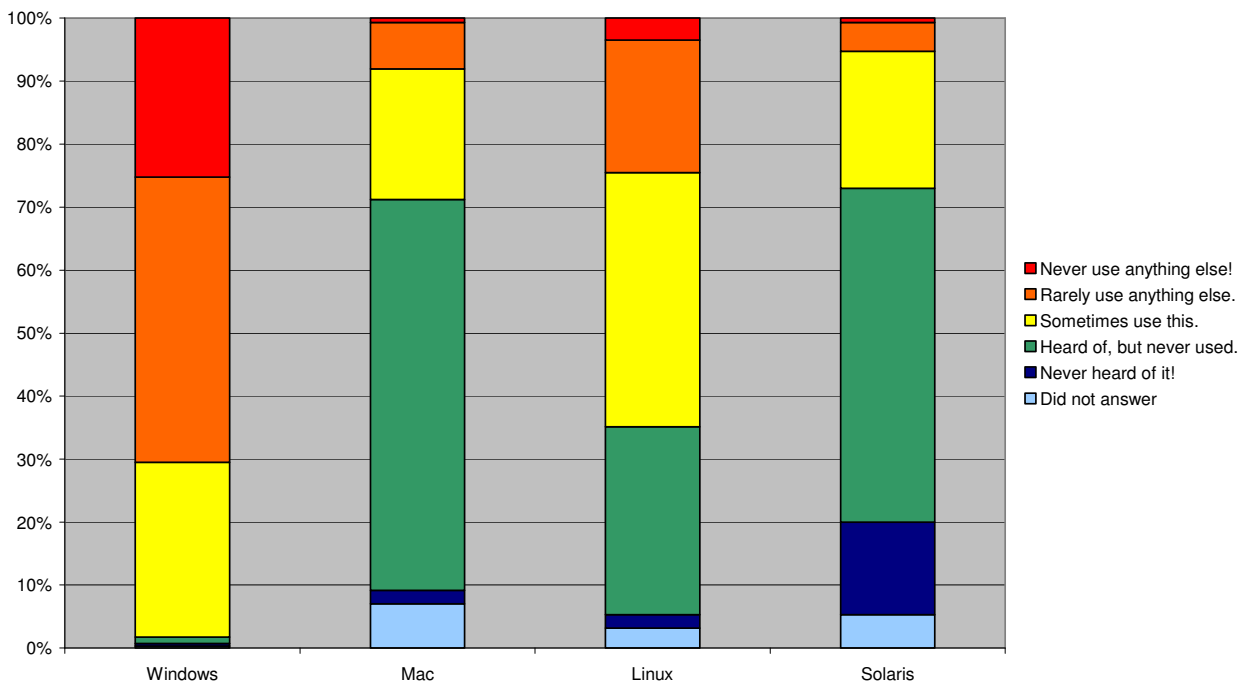


Internet browser usage - target users



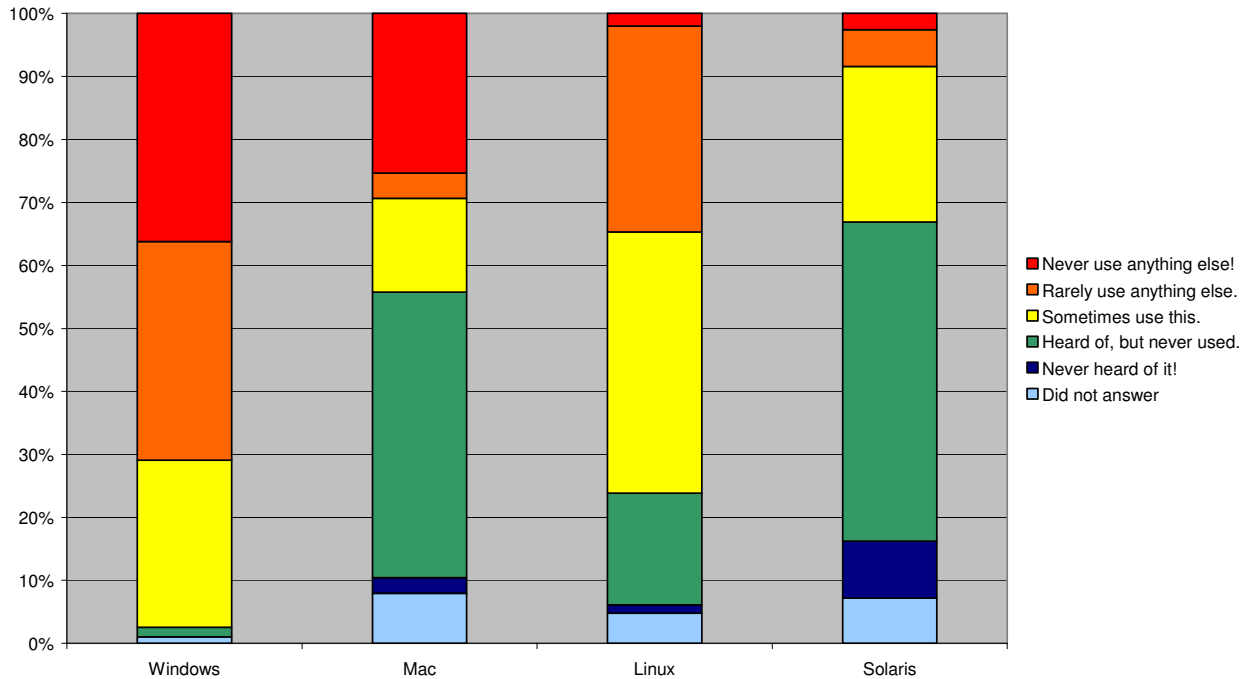
The lower numbers using Safari is also reflected by the lower numbers using a Mac (less than 30% used a Mac for at least sometime), with a greater uptake being seen for both Mac and Safari in the target BADC users group. By far the greatest operating system used by the users was Windows, although how this breaks down to the various versions is unknown. Linux was used to some degree by the majority of the BADC users with a slight increase of Linux use being seen by the target user group, but again there are no data on how this spans across the various flavours of Linux available.

Operating System

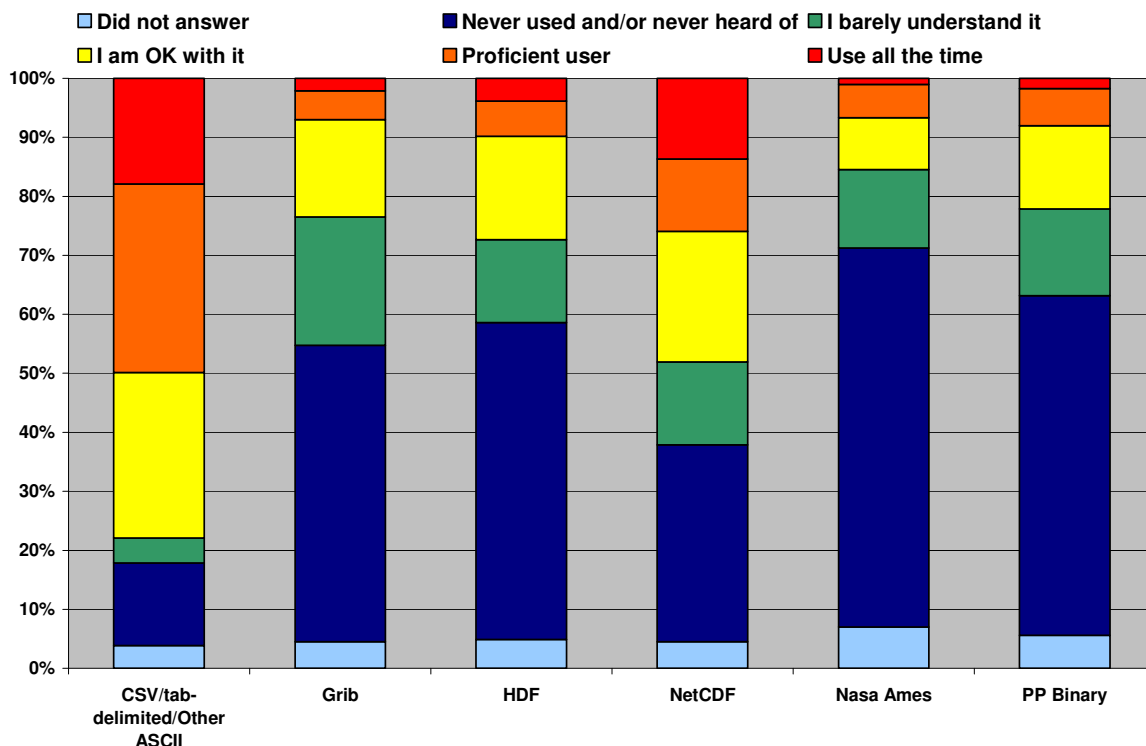


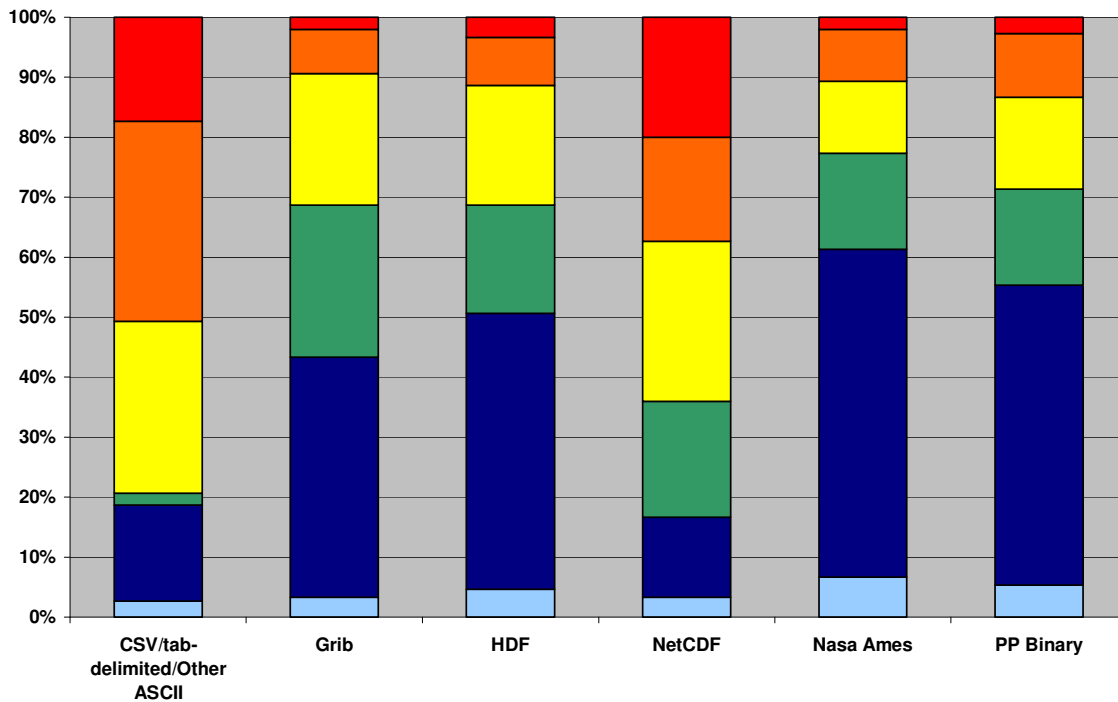
Only a handful of people (6) made a note of other operating systems they use which included UNIX and VMS. Other UNIX users would be included by the Solaris category and thus around a quarter of BADC users appear to use a UNIX system to some degree, moving to greater than 30% for the target users.

Operating System - target users



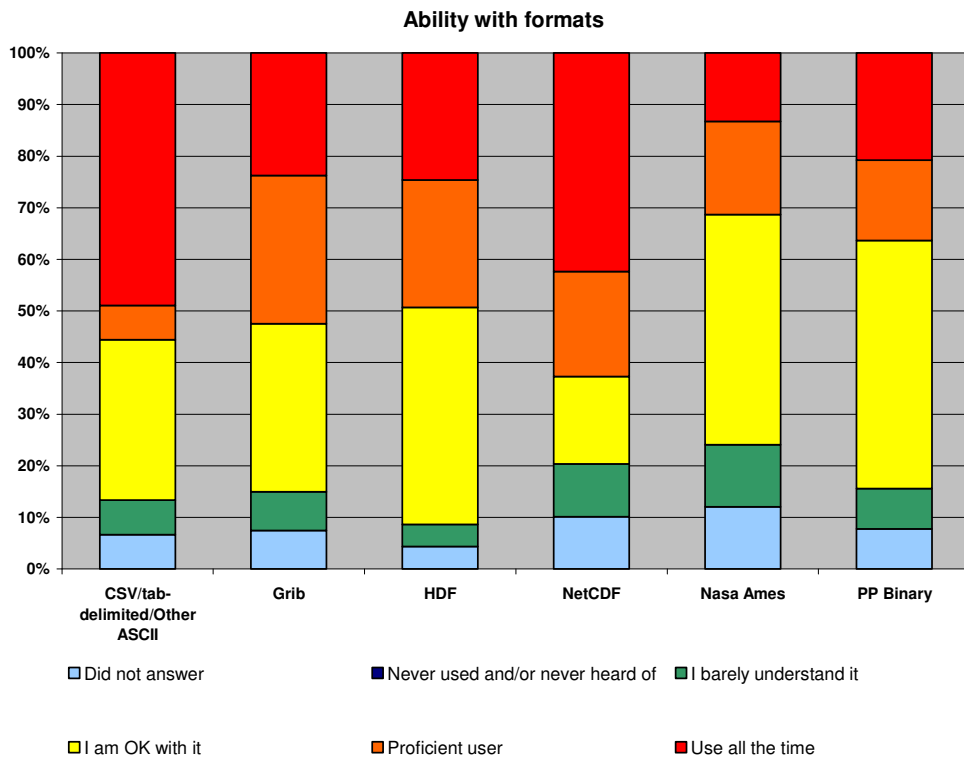
Key to the BADCs operations are the data formats used and whether or not the users are able to use such formats. When the users were asked about their ability with various file formats the results indicated that there is work to be done to educate the users in how to utilise the formats that the BADC supports.





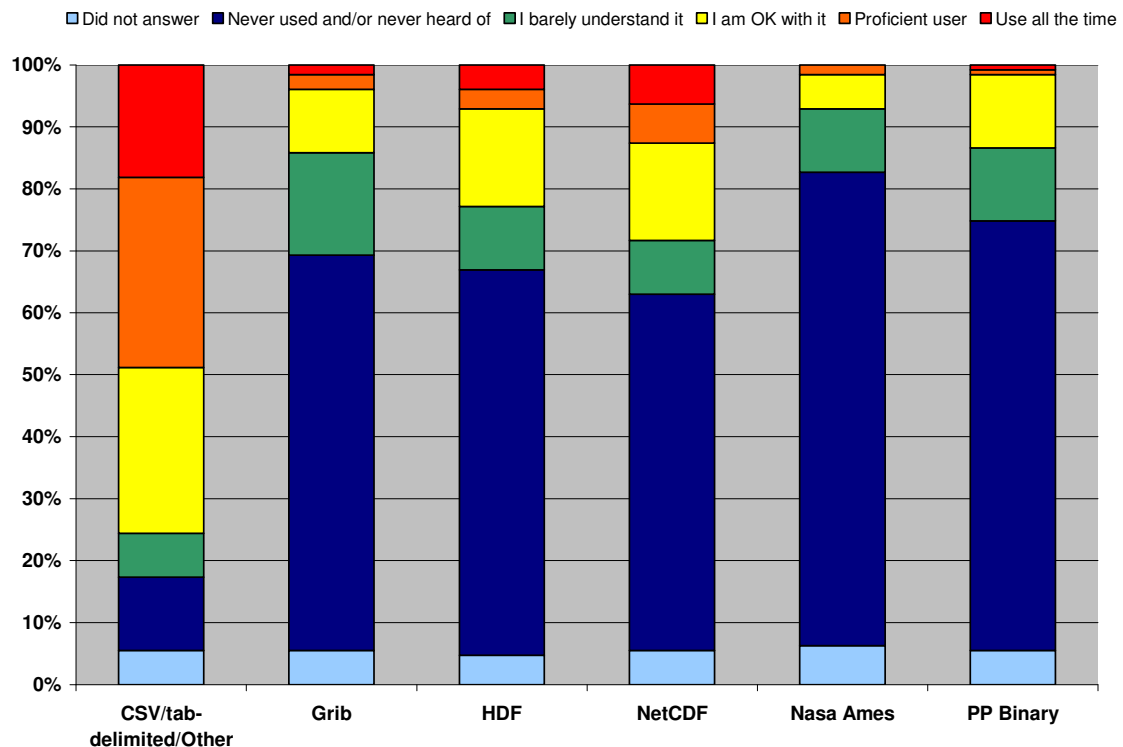
The above charts are for the survey sample as a whole and for the target user group. These show that there is some familiarity with the BADCs format of choice – NetCDF, but that this is exceeded by the generic CSV/tab-delimited/Other ASCII formats. The low numbers indicated familiarity with Nasa Ames could be as a result of people not knowing the name of this ASCII format and so selecting the CSV/etc. category in its place, but this can only be speculated at.

At first glance it would appear that the BADC has some work to do in educating and assisting the user community to better understand the various formats the BADC uses. However, taking into account that over 60% of users are at least ok with using NetCDF is encouraging. When the survey population is restricted to examine only those users using campaign data, chemistry data and remote sensing data (i.e. the group with whom the science support group liaises on a regular basis and hereafter referred to as program participants) the results are far more encouraging.



The impact of having scientists liaising closely with the research community to educate them how to use the formats found in the BADC archives becomes evident in the much stronger familiarity with all the data formats. In particular NetCDF has the greatest number of users who stated that they were proficient or used the format all the time.

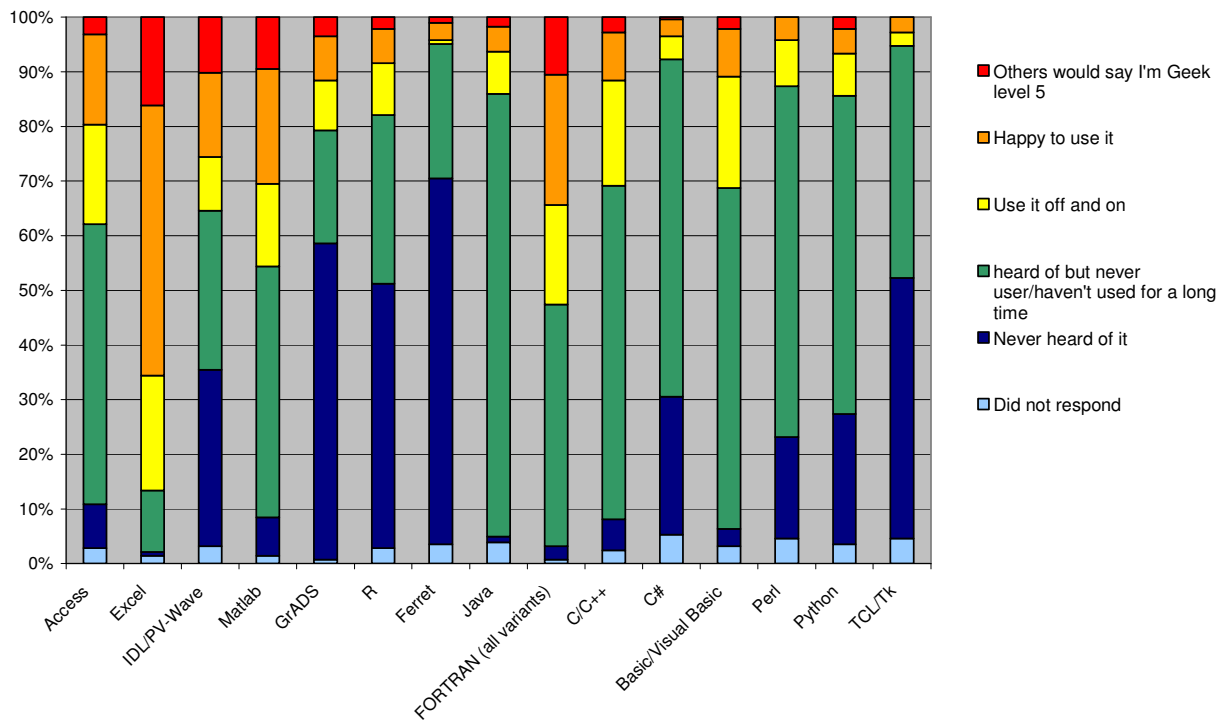
While this appears to show that the BADC is doing well in this regard it should be tempered when the statistics for the non-target user group are examined (see chart below). Here less than 30% of users are at least ok with NetCDF, with other formats fairing less well, except for the generic ASCII category.



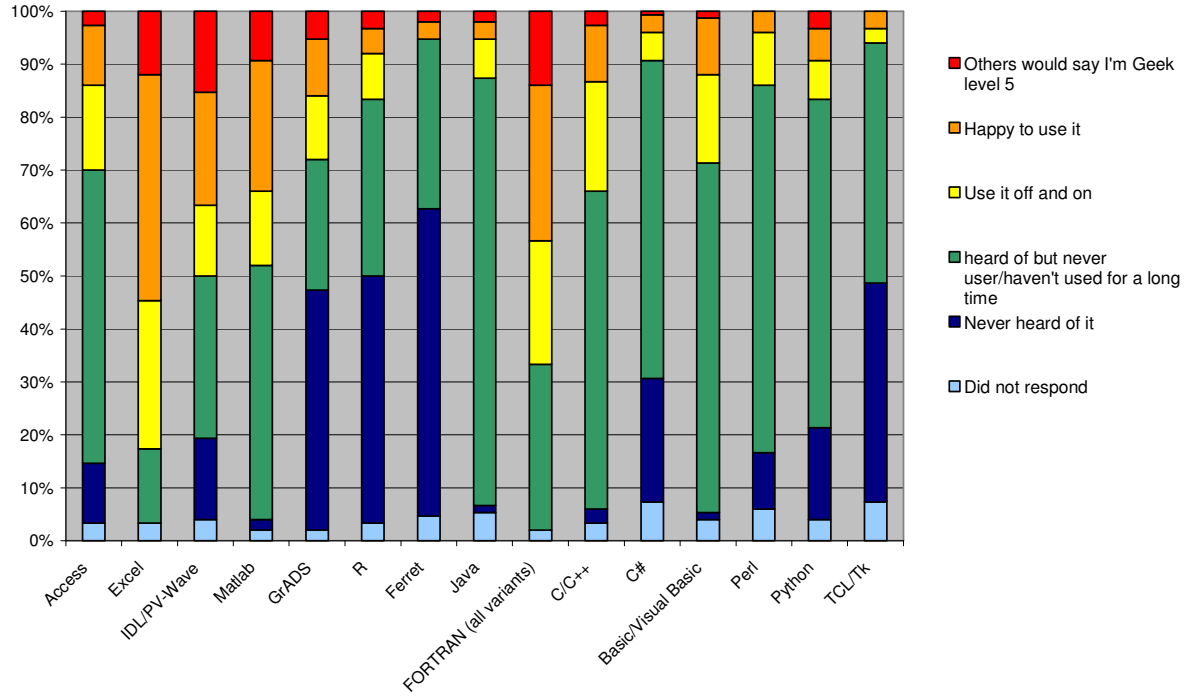
Various other additional formats were mentioned on the survey (see appendix for full list). These included GIS tools such as shapefiles, Arc and ESRI. IDL save and Matlab native files were also listed as were GrADS binary, Bufr and JCAMP. However, there were no format types that were noted by a large proportion of users that need to be considered by the BADC as a possible format type to be used.

When combined with how the above groups indicate their ability with various analysis tools and programming languages the ability of the non-user community appears to be orientated towards using Microsoft Access and Excel as opposed to the greater use of IDL/PV-wave, MatLab and traditional languages such as FORTRAN by the BADC target user group. These latter mechanisms lend themselves to using binary formats such as NetCDF and PP, while ASCII formats remain far more accessible to users using a much more interactive processing interface such as spreadsheets. The strong showing of FORTRAN can also be attributed to the large number of scientists who will have been educated to use this during their degrees and its use in the modelling community.

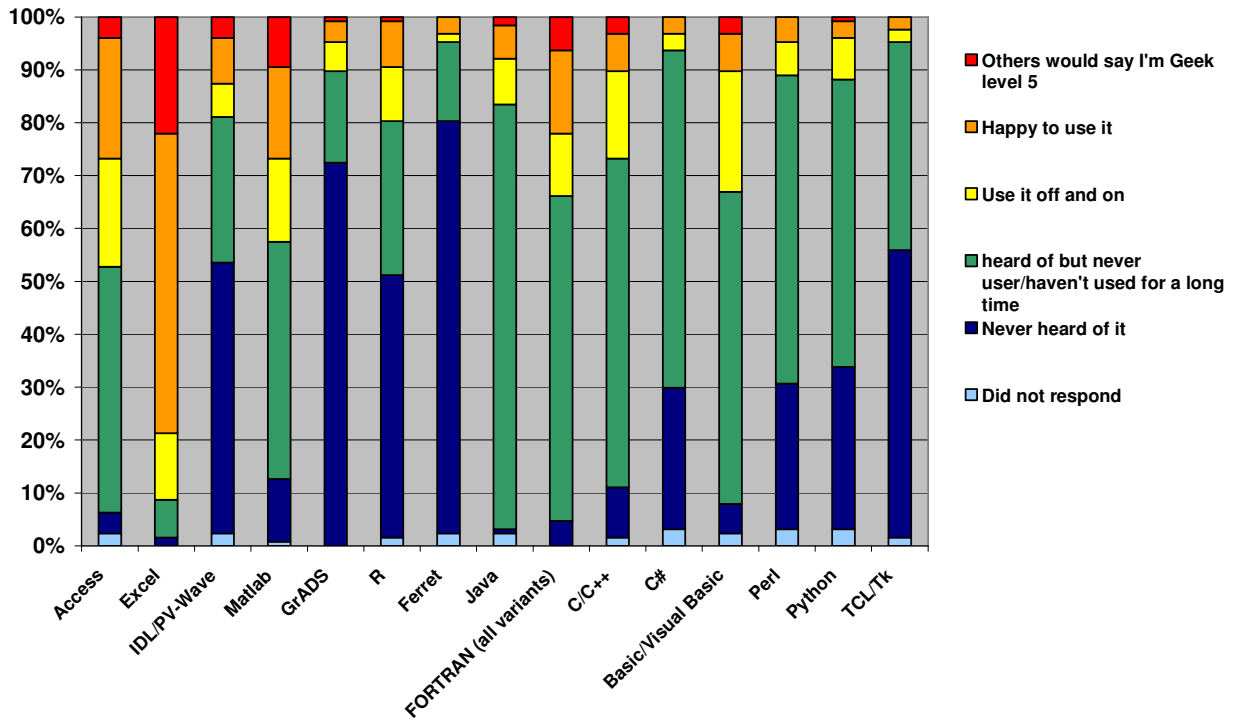
Proficiency of user by various analysis tools



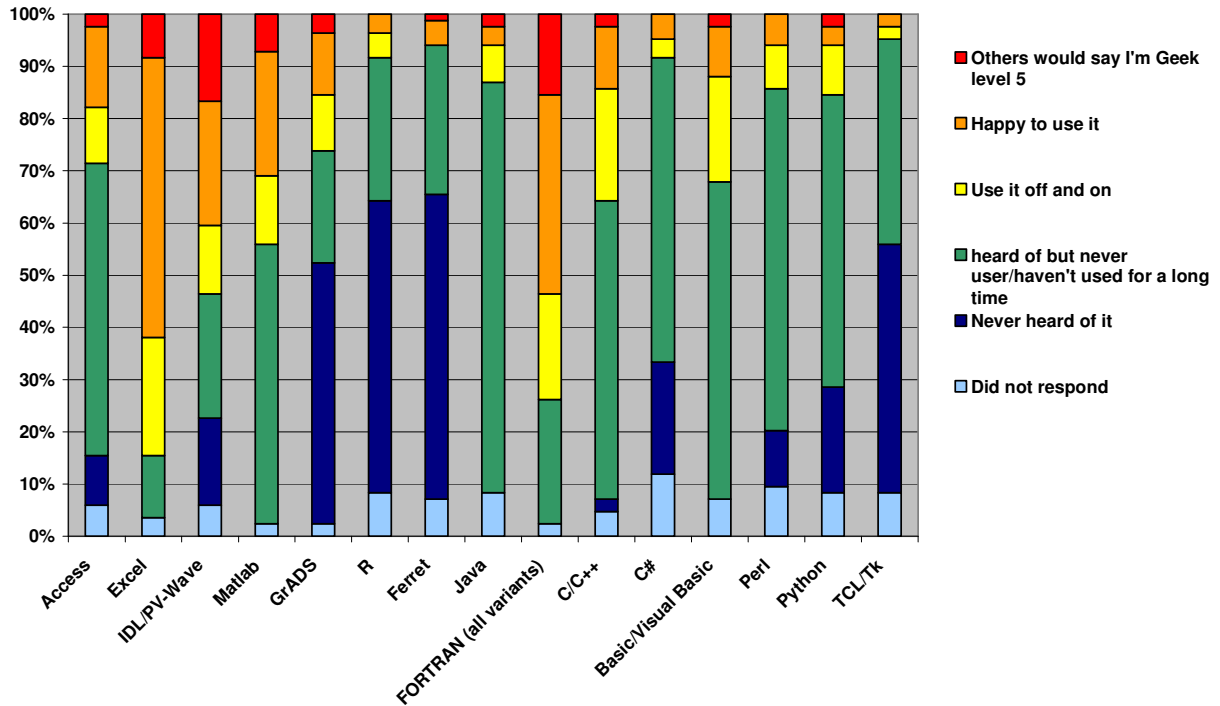
Proficiency of target users by analysis tools



Proficiency of non-target users by analysis tools



Comparison of the program participants group shows that they have a greater familiarity with IDL/PV-wave than the survey group as a whole. Meanwhile there is a slight shift away from using Excel with fewer people being happy to use it or indicating a strong familiarity with it.

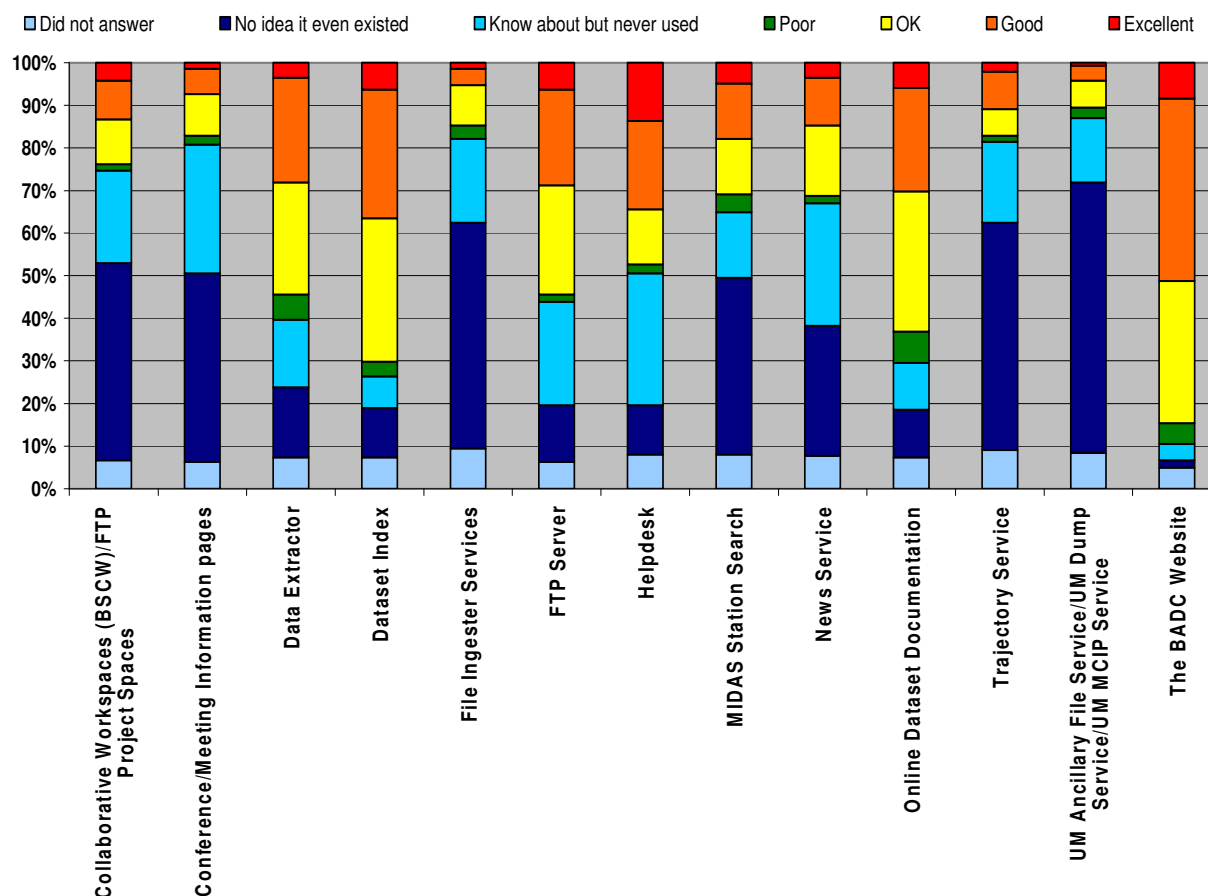


Other analysis tools mentioned (by 74 of the respondents) in the “other” section included ArcGIS, Genstat, MathCad, Mathematica, Minitab, SAS, SPSS, Stata and XMgrace. Some included IDL and Matlab here too, but were not included in the statistics for the above charts as they did not state their level of proficiency with the data.

Note regarding data providers: While it was hoped that a subgroup of data providers could be pulled from the survey responses by filtering the results by those who rated the file ingester services as poor to excellent (i.e. had an opinion as they had used the service) the resulting sub-set did not appear to constitute data-providers (names of these respondents were available from the competition entries). The only way such a group could effectively be surveyed would be to prepare a survey target at such people. Such a section was removed from the initial survey due to space limitations.

Experience of BADC.

This section examines the experiences of all those that filled in the survey to highlight those areas of weakness of the BADC as well as those where the BADC is doing well.

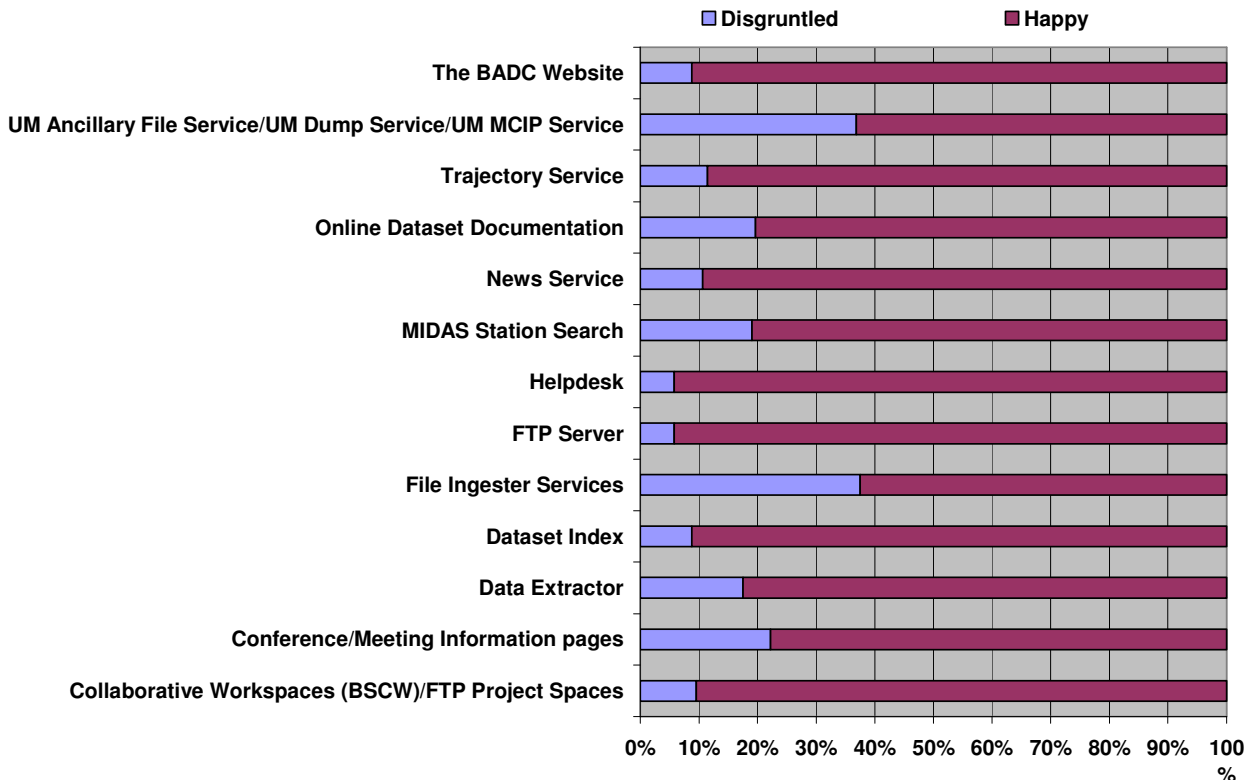


Overall the BADC appears to be doing an ok to good job with only a small number of respondents rating a service as poor. The largest dissatisfaction comes from the online documentation (7.4%), the Data Extractor (6.0%), the BADC website (4.9%) and the MIDAS station search (4.21 %). When examining the complaints left by the users the recurring themes were:

- The online documentation could be improved
- The data extractor can be slow and unreliable
- The website is difficult and slow to navigate

If a comparison is made between user who responded 'poor' (a "disgruntled" group) to those who responded 'good' or 'excellent' (a "happy" group), then it

is clear that the worst performing services are the UM services and the file ingester and UM services, while the helpdesk and ftp services are the best performing.



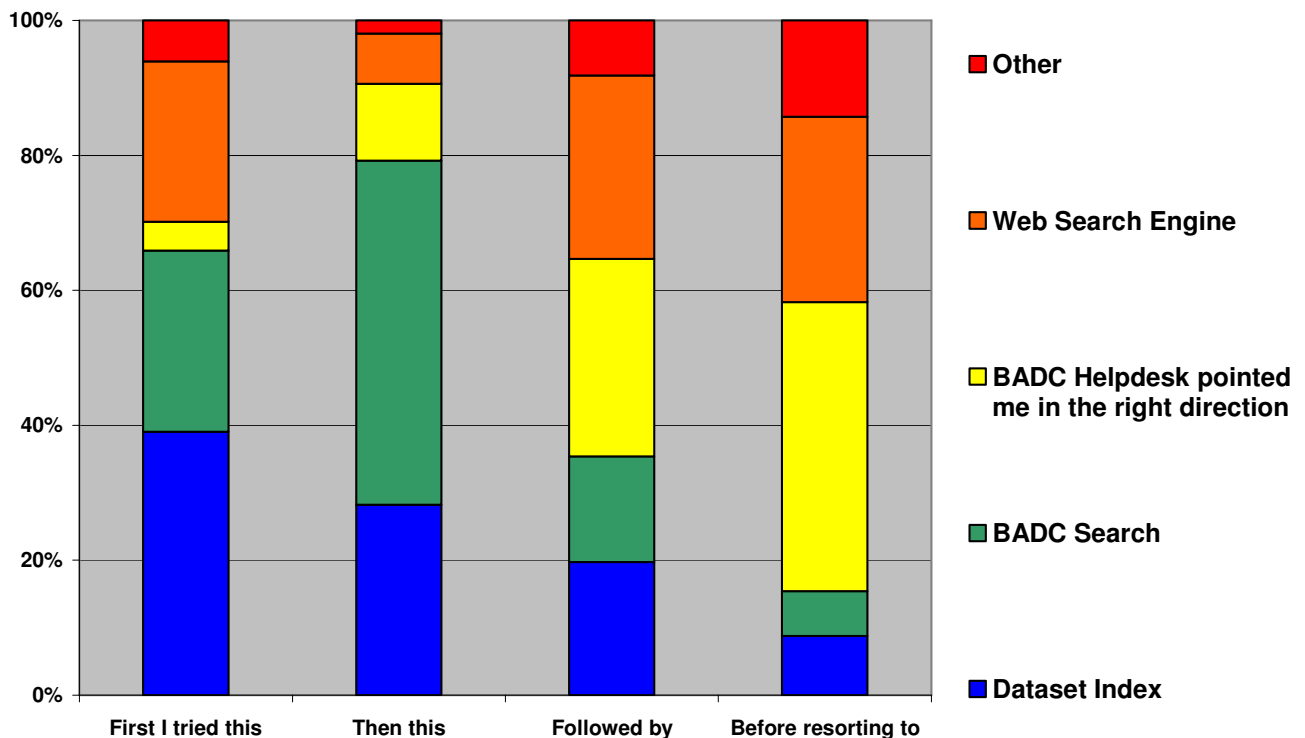
A survey carried out a few years ago indicated that the users rated the BADC website good to excellent, thus it appears that while the website may not have deteriorated over the intervening years its growth and the improvement of web standards means that the present site is looking dated and is becoming increasingly difficult to use. With the advent of the NDG services some changes will occur that should address some of the issues raised by the user responses.

Of perhaps slight concern is the number of people who indicated that they did not know a BADC service existed. Some of these are understandable as they are targeted at specific users, e.g. collaborative workspaces, but others such as the News service and MIDAS station search are for greater use. Experience from queries received by the Helpdesk show that such services are too hidden from the user and thus their use is diminished.

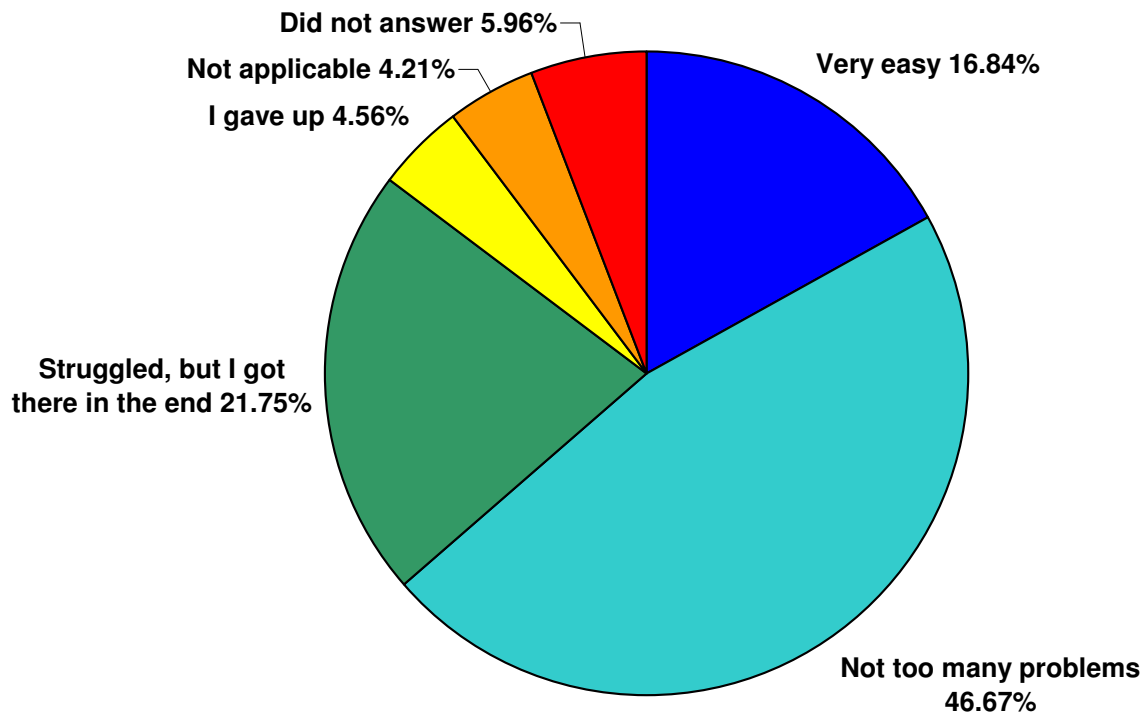
There were numerous general complements for the BADCs services as a whole, with a number of users making explicit references to how useful it has been/is to their research. However, two principle themes in the complements were that users are happy with the efficiency and efficacy of the helpdesk and that the range of datasets and services provided by the BADC are important and welcome resources to the research community. Interestingly there were also a number of complements for the website, contrasting the complaints about the website.

Obtaining data from the BADC

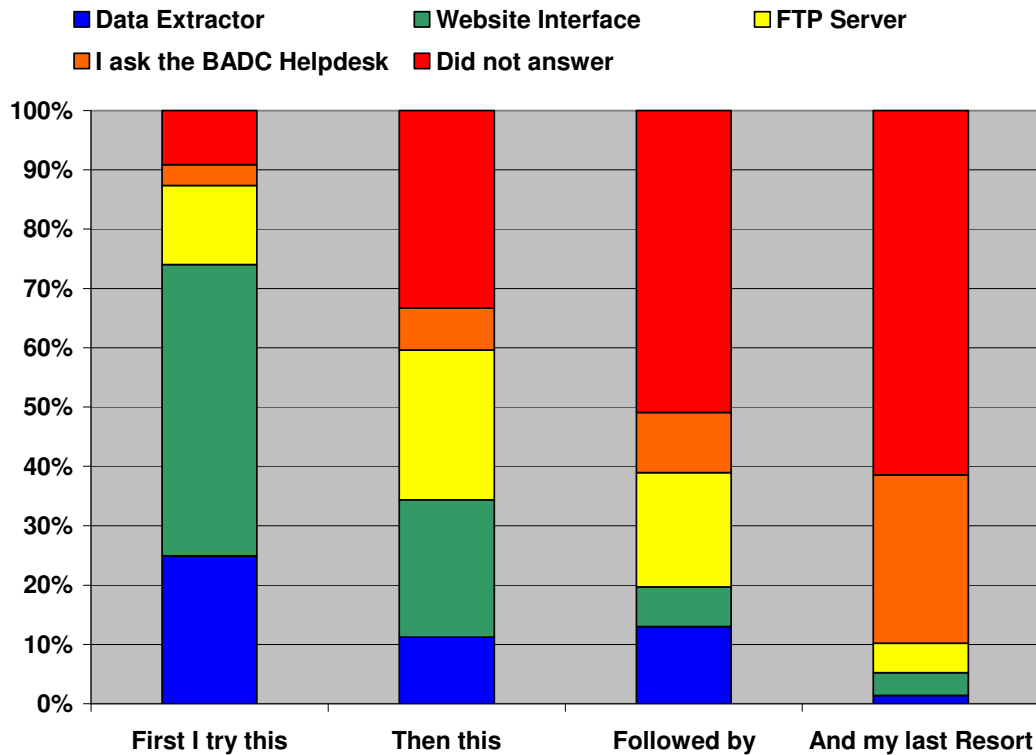
One of the aims of the survey was to establish the user experience of obtaining data from the BADC. To try and establish this, three questions were asked. Firstly, “how easy did you find it to locate the data you wanted?” - i.e. how do they discover the data. The breakdown of responses is in the chart below which shows that on the whole users first if all attempt to locate the data either by using knowledge of the dataset index or by using some search function, usually beginning by starting at the BADC website. After trying to locate the data for themselves the user then tends to resort to the helpdesk. The experience of the help desk is that while such enquiries can be with regards to all data types and from users of all backgrounds there is a greater tendency for such enquiries to come from non-target users with a large number enquiring about the MIDAS dataset.



As for the ease of the user to locate the data using the above routes, question two asked: “How easy did you find it to locate the data you wanted?” The following pie chart shows that over 2/3rds of users are able to locate the data they are after with little or not problem. Only a small fraction indicated that they struggled to the point that they gave up. While the figures look positive, the sample may be misleading as the results do not give any indication to the number of people who try to use the BADC to locate the data they require only to give up and never attempt to use the BADC again. Discussions with some users indicate that there are users who have resorted to other data centres for their data in the past due to the difficulties of locating data at the BADC.



Finally, the third question – “how do you normally obtain data from the BADC?” - focused on how the user then went on to get data from the BADC archive to their local machine. The results are shown in the following chart:



Clearly the website interface has been playing a key role in how users obtain their data, with the data extractor also playing an important role. Looking at these results and comparing them to the comments made it is clear that users desire a way to subset the data they are interested in before downloading it. This is particularly important for the larger datasets or at least for those with a large number of variables/points in them, e.g. the MIDAS data set. There were a number of calls for such a service to be extended and further supported.

Once again the helpdesk plays an important service being able to help when a user struggles to use other ways to obtain the data they require. Primarily such assistance is provided with data extractions as while the data extractor has a 66% success rate the large number of users of the service still means that there is a sizable number of users who approach the helpdesk for assistance with extractions.

Future Developments

No qualitative questions were posed to address the question of “what would the BADC users like to see in the future”. While it may have been possible to ask if users would have made use of such services as a data discovery porthole, it was felt more important to ask the users for their input, rather than suggest those services we are developing at present.

Thus, an open text field was provided which returned results that split along the following categories:

- a) Data Arranged/Formatted Differently
- b) New Service Suggestions
- c) Data Extraction/Upload
- d) Metadata/Documentation Improvements
- e) Additional Dataset Suggestions

The full list of answers is included in the appendix. A sizable number of requests will be answered by the NDG service once implemented. Other aspects (such as more surface data) are, to a degree, beyond the control of the BADC as we act as a data centre rather than a data collecting agency. Other requests could be examined to see if they are feasible or not, such as obtaining a high resolution radiosonde data from more stations than at present, more satellite data and MOZAIC chemistry data from commercial aircraft. One recurring theme was the desire for improved provision of NWP data and related products.

Conclusions and response from the BADC

The positive response from the BADC user community to fill in the BADC 2007 user survey has resulted in a useful insight into the user community's abilities and needs. At the start of this report the three aims of the survey were set out:

- 1) To determine the skills base of the BADC user community.
- 2) To determine the present experience of the BADC user community of its datasets and supporting services.
- 3) To identify those areas where the BADC should improve, those areas where the BADC is doing well and those areas where the BADC could explore/develop.

The above results show that all three aims have been met:

- 1) The user community consists of a number of identifiable subsets, each with varying skill levels. The group with the greatest skill base with regards to having the ability to make use of the various formats of the BADC using a number of tools are those involved with various research programs. However, these users only account for 30% of the users in total and so a significant skills gap still remains for the remainder of the users. This was most stark when results from non-target users (taking in geographers, oceanographers, biological systems scientists etc.) were examined (nb this included some program participants). This group tended to favour ASCII based formats which could then be used in various Microsoft (or similar) tools.
The figures for the BADC's target user group (i.e. those involved with the atmospheric sciences) are still not as encouraging as the program participants, but are encouraging nonetheless. Generally this group has a breadth of skills, being familiar with a variety of data formats, including the BADC's preferred format: NetCDF. They also show skills with various analysis tools beyond Microsoft products and also beyond just a Windows platform. Such tools require a degree of programming skills and lend themselves to data handling and manipulation. Skills with FORTRAN are noticeably strong across the board, with other languages fairing less well. Proprietary packages such as IDL/PV-wave and MatLab are also in reasonably wide spread use amongst the target user group.
- 2) From the experience of the user community it appears that the BADC is providing a **good** service that meets the expectations of the user community. The high level of praise received through the survey is encouraging. However, there still remains work to be done as highlighted by the users. These are principally needing to improve the online documentation (for both datasets and serviced), improve the layout of the website to provide efficient and ease of navigation and provide better methods for accessing and/or subsetting the data. A significant number of people were unaware of some BADC services; there is a need for better advertising and navigation for these.

- 3) Areas for possible investigation and/or development were wide ranging from the suggestions supplied by those who completed the survey. There were a number of datasets that were requested (e.g. A-train satellite data, improved high resolution radiosonde data coverage) suggested as well as various additional services/amendments to services. A number of the service suggestions should be addressed by the implementation of the NDG.

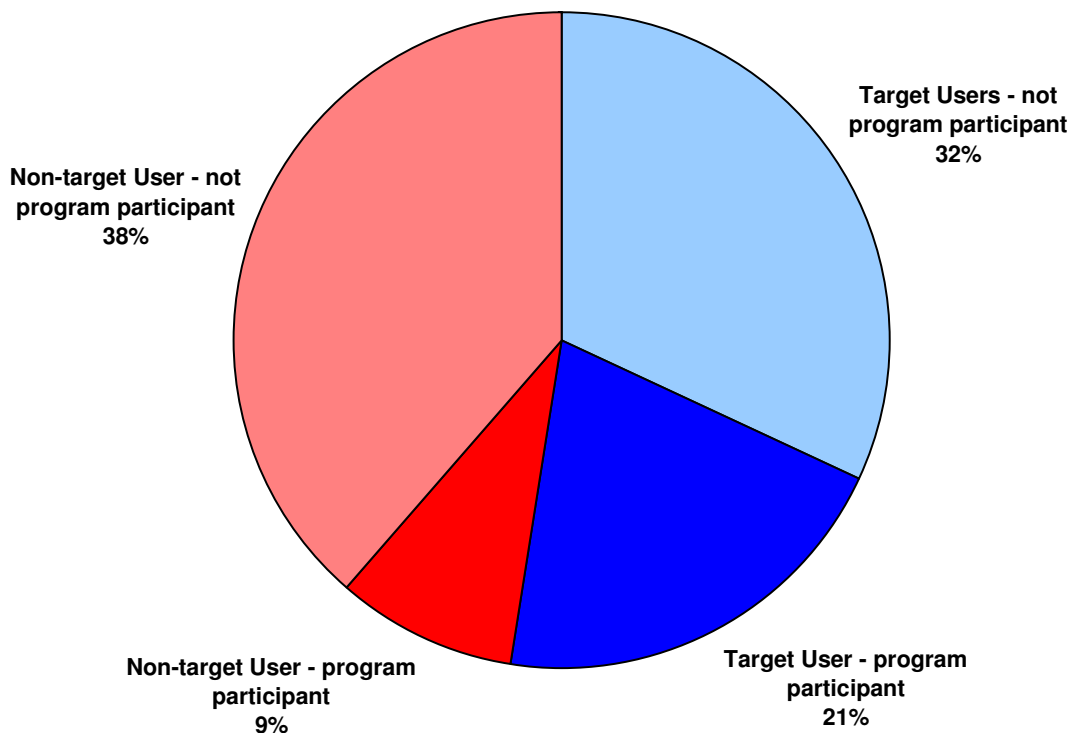
Finally, it is suggested that a separate survey be conducted to explore the skills, experiences and needs of the data supplier community, an aspect that was not covered in this survey.

The BADCs response.

A response to each point raised by the users in the survey is not possible, but the following how the BADC will attempt to respond to the points raised above:

- 1) Data Formats – NetCDF is in increasing use. This survey provides evidence that this is not just hearsay. NetCDF will be the format of choice for all BADC datasets. Plain ASCII formats are still the predominate formats in use and the BADC will continue to supply this data this way even if the underlying archive format is NetCDF.
- 2) The BADC Website. The response to the survey suggest that the website is generally useable, but significant number of navigation problems exist. Some of these problems will be address by the deployment of software developed in the NERC DataGrid project (<http://ndg.badc.rl.ac.uk>), however a more general review of the sites navigation and general presentation will also be undertaken.
- 3) Increasing datasets provided. We will look at the suggested new data sets and add those that have significant demand and are within the data centres scope.
- 4) Improved documentation. We recognise the inconsistencies in the level of documentation between datasets. We are in the initial stages of a project that will add a document repository for a more organised and systematic documentation procedure. This should identify where documentation is lacking.
- 5) Data discovery and other data services. The new NERC Data Discovery service (<http://ndg.nerc.ac.uk/discovery>) will improve the discovery of the BADC datasets as well as the other NERC data centres. The provision of other data services will be carefully considered as opportunities arise, but cost of maintaining this service is often considerable and needs to be evaluated.

- 6) Future issues to be addressed in the next survey. The survey had a good response and is a valuable tool in setting our future direction. A specific group of users that need further probing is data providers. Questioning of NERC funded data producers is routinely gathered, however third party suppliers, like the Met Office are not systematically surveyed. A single consistent survey of data supplier needs would provide a powerful tool for BADC policy setting.
- 7) User base. The survey has confirmed that our users are who we think they are. Statistics from our user database are consistent. This confirms that our designated community of atmospheric scientists is augmented by a high proportion of non-atmospheric science users. We will continue to consider the non-atmospheric user base as the BADC service evolves.



Appendix

Comments/Complaints/Compliments

Comments

- Though I am from Agricultural field, BADC provides me restricted data sets for our use.
- Sometimes it's difficult to identify the latest version of files when they appear in the Workspace and the main BADC. But otherwise, an excellent service. Thanks!
- I find I still have to put a lot of work in when collating data for the same site from different data types, and I'm trying to develop a work-around. The use of the annual files for the MIDAS data set is useful for downloading, but difficult to manage. The data extractor helps with this, but I still find the collation of data to be a big job.
- Site is good but difficult to navigate all the information
- These are some comments how I think the service could be improved: - pp-format for MetOffice has quite a steep learning (tool writing) curve and is too difficult to use; a simple tool to convert to ascii would be necessary to use this - ascii data should have default values in blank fields, the current format should needs to be converted each time to be used in Fortran, gnuplot, etc; then also separation by space rather than comma is possible, which is also easier to handle - data overview could be more structured as opposed to the current simple list - allocation of stations to a geographic location is tedious (eg MIDAS), lat/lon should be part of the data set and not just station code or number
- There are a few bugs in some of the automated data extraction methods, but the help desk are quite helpful in resolving problems.
- I would like to use the UM datasets but do not have the time /resources to devote to learning how to plot /retrieve info from them
- An easy guide to displaying the datasets might be useful.
- I am overall very happy with the system. I have come some problems with web extractor although I recognise it works. I would appreciate to access the data via an opendir server. The few times I contacted the helpdesk I got my question answered quickly.
- The other day I tried to get surface air pressure for somewhere around Eastbourne, Southeast England for 2001 to 2006 but somehow could not find it. Should be easy to find in a sort of google environment... It's the occasional user that probably struggles the most.
- BADC is a great service. I like the ftp access, but some would like it to be improved. I know some want back data sets of the um produced, but I would rather see a good set of data produced for the future. I would like to see a repository for visualisation /analysis codes / routines for matlab and idl for tephigram / power spectral analysis etc. set up in the badc system as a free ware facility
- A good service - very useful. I think that in the future you could look at providing the data a little more dynamically (rather than having the user download one file at a time you could have an interface that allowed the user to specify the timespan of data they wanted, and your interface could extract all those files and zip them into a single file for download, and perhaps also allow the user to select one of several formats for download like an Access database?).
- I rarely think of BADC for getting hold of data but probably I should. My first port of call is a google web search. If the datasets and their availability are advertised better and NetCDF files are provided I would certainly use BADC more.
- Generally I find it to be a fantastic resource. Occasionally I find the page navigation to be a little counter-intuitive, and the recent provision of the data extractor is a mixed blessing as I found it harder to find the data I was after.

Compliments

- The services and information provided by BADC have been excellent and are very useful.
- Thanking you all for your great help and services. Hope BADC will keep growing.
- The BADC is a very useful tool and provide access to a large range of data. People very helpful and do their best to answer to emails
- The helpdesk is really helpful and extremely friendly. This is greatly appreciated.

- Belinda is fantastic! Whenever I've emailed in a panic she is quick to reply and always provides a useful solution to my problems. I'm also very grateful to the Met office for making so much data freely accessible via the BADC.
- The high resolution radiosonde data has helped me a lot in my PhD research. Thanks!
- Sometimes it's difficult to identify the latest version of files when they appear in the Workspace and the main BADC. But otherwise, an excellent service. Thanks!
- Thanks, the BADC is a great resource!
- I was impressed by the data extraction/plotting software demonstrated by Stephen Pascoe on the BADC stand at the RMetS conference. It would be desirable to have the same flexibility for plotting/viewing plots from all BADC datasets.
- I've always found that the level of service of the help desk has been excellent
- Swift response from helpdesk whenever I've used it - very pleased!
- So far I did not have any problems with data downloading, everything is clear enough on the web (sufficient documentation, description and references).
- The BADC is a great service and keeps a unique set of data which I really appreciate.
- Responsive and nice people at the helpdesk ...
- I had problems with the Data Extractor, and Helpdesk was always very good and efficient in solving them. Thank you guys!
- On the whole quite a good service.
- The site is great, compliments on that!
- Hi have always found the helpdesk team very efficient and helpful. My feeling is that the BADC is a fantastic resource and should be fully supported in providing a high quality service.
- Service and data coverage extremely good.
- Thank you for your regular emails about new datasets etc. Thank you for the existing data, which have been most useful. Congratulations on your good work.
- Thank you for the nice website and for all the help.
- BADC is very good service to the young researchers and scientists all over the world. Really, it is very good and I used for my research purpose.
- I think both data server and help desk are very good. But I only accessed UKMO data which is very limited. Thanks very much!
- I found the use of the data from BADC to be very helpful. It was great to get old land temp data for remote areas of Scotland.
- whenever I have used the service it has been very good, thanks.
- BADC has been very useful for my research. I regularly use global radiosonde data, in particular, which resulted in a couple of research papers.
- i find web site easy to use
- I have always found BADC to be very helpful and I have always been able to obtain the data I need if it is available.
- The helpdesk was very efficient in solving my doubts.
- Great Service
- I find the staff at badc are very helpful. They respond swiftly to my email queries. And they keep responding with careful personal attention until the problems that I run into have been resolved.
- The services provided by BADC are an important resource for scientists researching links between climate and health
- Overall, the BADC provides a very important service in a very competent manner. I would be very disappointed to see a cut in services.
- Thanks for your work!
- Congratulation for your work. Thank you and carry on the good work
- Thanks to the team, that provides fast reply and advice to enquires
- BADC has been source of great scientific centre for research development and I quite applaud the centre for this wonderful service to the scientific world in which we belong.
- All in all thank you, good job :) Helpdesk staff are efficient and friendly as well
- Very pleased with the BADC - an excellent resource, but too much of a well-kept secret.
- No complaints, very happy with BADC so far. Instances in which I have needed help or more info have been dealt with very satisfactorily via your helpdesk. I suspect that it is impossible to get a web based service perfect so the option of well informed, quick, efficient helpdesk

support is very important. Also important to be able to talk to helpdesk by phone where necessary.

- I am very glad, that the BADC exists.
- Great Job, very useful service. Thanks a lot
- Friendly and helpful folk in general when met at conferences etc...
- Excellent service provided by BADC enhances my research
- Emailing BADC support has always led to very friendly, helpful and speedy responses.
- I have always received good support from help desk.
- I feel the BADC does an excellent job overall.
- A good service professionally run - please continue to do what you do so well
- Excellent that this data is accessible to ordinary students and independent researchers - well done! Thanks to helpdesk, too, for all your help.
- You do a great job!

Complaints

- I found the dataset listed on the BADC website rather old and not very often updated.
- There does not seem to be enough people to answer specific requests. As a result, some problems may take time to be solved. To me, a centre with more people available to users would save a large amount a time to atmospheric scientists. Thanks.
- UM Ancillary files are never complete for my area of interest (Antarctica)
- Data advertising ie 1950-1960 misleading, when there are gaps or sometimes only a few years within that timeframe. Any gaps could be shown as 1950-1960 (discontinuous). Otherwise hopes are raised, only to be dashed!!
- I have been working as an amateur scientist, and have been intimidated from using the system because I was not a registered student.
- Data Access to UKMO UM data allowed then denied. BADC should provide the details of who is responsible for decisions to deny access. BADC's refusal to supply this information demonstrates a lack of accountability.
- Your web site can be slow at times, and the direct ftp route is not well publicised or easy to find. This is frustrating when trying to access large amounts of data. Your web system works well for single files but is tedious for multiple files.
- netCDF is still fairly impenetrable for first-time / occasionally users
- I think some of the dataset webpages could be a bit better laid out with more obvious entries for their description, key references, what they contain, and how to download. In particular some model datasets seem very ill-described in terms of the model setup etc.
- Please provide BETTER DOCUMENTATION! I have yet to find any documentation on the MIDAS metadata available to me. In addition, many of the data fields and abbreviations are still a mystery to me. The 00_README files contain next to zero information, if any at all. The website should be re-organised and structured to assist a new user, with links to appropriate documentation from the dataset pages. It doesn't help that I have to be psychic in order to help myself. Given the poor website layout, I have been extremely reluctant to utilise the Help options available, as it is quite clear the entire effort was underfunded. Yet, and without any sense of exaggeration, I would not be able to accomplish my research without the BADC, it's resources are that important to me, my colleagues and my field of study. Better support for Linux and alternative web browsers, such as Mozilla, as most of us power-users avoid Windows and proprietary limitations.
- As a general rule I have found the database to be very helpful. At times I think more information could be provided within the notes.
- not very easy to see what is in each file type e.g. i have to keep clicking back and forth to see what .HWX means etc.
- I have the idea that the MIDAS weather station search engine has changed in the last year. It has lost one of the fine search options: to search via grid reference, that is unfortunate.
- The Met Office needs to release data to the BADC in a more timely manner, i.e. providing more recent land surface data for all stations.
- Data not being up to date. Why if data has been requested as it is missing is it then not posted on the site?
- ECMWF operational could be more up-to-date (currently a lag of 10 days seems to be the norm)

- I had a few problems with the quality checks of the data
- Would be nice to sort out the midas data so that there is comprehensive guide to what all the fields are in the data, all the headers match up, no duplicate records, an idea of why the metoffice has different version numbers... But their more metoffic communication problems than with the BADC.
- Generally I find this site quite useful. However I sometimes find it to be a bit of a maze to try and find everything that I need as there are so many pages relating to the data set that I am interested in. I think this could be cleared up by having all possible links for the FAAM data sets put onto a page and then I can see where everything is that I need.
- Transversing BADC web interface (up and down the directory structure) is SLOW! so slow that I end up resorting to ftp... Also the interface will not allow enough "depth" in terms of directory levels.
- It can be annoying that access to website pages is slow. Otherwise it's a great service
- Web site is sometimes slow to navigate.
- My only complaint is that the web interface is slow, but this is not very serious given the speed of the ftp server.
- The website always seems very slow to load, and to go between pages.
- The MIDAS data extractor is very slow and not user friendly.
- The web-based data extractor is useless and time consuming. Each time, the user can only extractor 100MB data, which is less than 5 day worth of data if one needs all the fields. I tried the extractor for a few times, and I had to give it up. I contacted the data centre and got the data by using FTP.

Section 3: About what you would like to see

Data Arranged/Formatted Differently

- Sets of atmospheric data organized by atmospheric species.
- compress the radiosonde data for fast transfer
- The UM ancillary file software only gives climatological values but sometimes more accurate values are needed when running the UM for a particular case.
- better preprocessing of data before it is given to the user. Some users want "raw" data as now but others want a clearer presentation of data in a format that they can understand.
- A mind map or diagram of the folder layout. I find it difficult to extract information I need or to click through folders to find the right dataset, especially for ECMWF. I often download the incorrect dataset by mistake.
- It would be helpful if the raw unprocessed data was available - this would help in understanding variations across datasets.
- rename the datasets to an easier format e.g. if its a wind file name it wind instead of .WNX etc
- Download in csv, because it can be readily loaded into lots of different programs
- WCS,WMS,WFS Want to use GIS
- altitude (km) data in addition to pressure data in trajectory files
- Improved software modules to read specific datasets to speed up my use of the data.
- Data available in variety of formats, preferably ascii and NetCDF
- Just keep archiving the Met Office coarse resolution "UARS corrolative" data in netcdf format. Thanks!

New Service Suggestions

- Extraction of projected climatological data for using in the hydrological prediction analysis.
- computation via web, ie, spectrum, seasonal and annual means, etc etc
- Data converters from one format to another
- Easier location data extraction
- Example uses of data by the user community - this is currently being updated for the EDINA Digimap website (useful for GIS admin boundary data & OS maps). It provides an interesting introduction to some datasets & inspiration for what work can be done.
- It would be good to be able to query data (in my case sea surface temperatures) for a specific longitude/latitude over a specified time period to obtain only the data required - this would save a lot of time.
- A preprints server for atmospheric and meteorological papers would be nice
- Maybe it has been implemented since I last used the database, but it would be nice to be able to enter a date or period and a location and be able to download different datasets: eg radar, radiosondes, ecmwf etc..
- Running user supplied tools to extract/pre-processes data to reduce data volume for transfer.
- quicklook facility
- Links to more work done with the data.
- R packages available for data analysis or at least a list of toolboxes
- Ability to manipulate data (e.g. run FORTRAN or IDL program) in the project space before FTPing to local system.
- A More update newsletter combining the climate change information
- I was surprised the data search found nothing under degreedays, especially as I arrived here due to a reference in CIBSE TM34. Degreedays are used to calculate expected heating and cooling performance in buildings, amongst other things.
- It would be very useful to be able to expand the BADC search facility to enable all data sets containing a specific data type to be identified. For example, a search for 'land surface temperature' or 'skin temperature' would lead you to the Met Office Cardington observation data set and the ECMWF operational and reanalysis data. This would allow people with a specific interest in one data type to find all data sets containing this data type very easily.

- Links to Facebook
- opendap server
- I would like an improvement in pp to anc / viz-versa and other data manipulation tools for the UM. This is not an easy task, and will take some effort. I wonder if there could be a consistent plan to make sure we keep all meso scale model results from the um. I suspect tyhios is out of your control though.
- Analysis tools for quick online viewing of data prior to acquisition.
- it is cumbersome searching for datasets using the data index. Problems: 1) they are ordered alphabetically rather than grouped into similar content 2) there is no way of sorting that index by other fields. Also, the search tool does not allow you to specify logical operators like AND - this would be very useful. I would also like to see a graphic added to each of those index entries that shows the time span of each dataset - something that you can glance at and know instantly the span of years the dataset covers. This is a key variable that determines what datasets I can use, but I have to hunt through several layers of pages to find it (and then it is often not correct/up to date).
- Better index to Met Office surface stations e.g. a clickable map which you can filter for different years or data types. Synchronising Met Office surface data across several years, so that format is the same. Holding several years as concatenated files in wider range of formats e.g. SAS. Linked metadata to data files i.e. so you can select a particular column and obtain details like variable definition and units, method of measurement.
- more space available for projects
- An overview of types of data available on a single web page with quick links. I always struggle to find the relevant datasets.
- Tools for analysis of ERA-40 data similar to those available for NCEP data at their Electronic Atlas

Data Extraction/Upload

- I've never managed to use the ECMWF data extractor, that I believe is connected to BADDC.
- I'm a great fan of your data extractor, and would like this to be developed further. You might like to improve the graphics package that comes with it as the graphics are not of the best quality.
- web data service never seems to be able to download from a teir of directories so often resort to downloading everything via ftp and then sort and delete locally.
- Easier FTP upload on the workspace. having to upload files individually is a bit annoying.
- To be able to select variables using the data extractor. At the moment the selection is compulsory for all variables.
- A more advanced data extractor with larger possible filesizes (I quite like the ECMWF attempt for ERA data). More efficient than downloading large files via ftp and pruning them myself.
- better upload services (this might have been changed recently) used in 2005/06
- download datasets directly
- I find the WCRP CMIP data site very good to use. i.e. good search engine - not too difficult to download from the web-site. but ftp direct to my unix disc as with BADDC is the preferred method.
- just make it easier to download the data straight over the web please!
- I have used your ssh access/workspace, which would benefit from being able to use xconv/convsh or the likes for data extraction. I also had to ask for a (working) FORTRAN compiler for use in that area - would be useful if such tools were automatically available.
- I would like to be able extract Era40 reanalysis and ECWFM operational data easily.
- Improved reliability of data extractors. E.g. I could never get ECMWF ERA-40 data from the data extractor for more than one day at a time.

Metadata/Documentation Improvements

- the metadata surrounding key datasets (such as MIDAS) are poor. The acronyms and lack of cross referencing are impossible to sort through.
- How to use hdf and netcdf format for new users

- I don't think I really make the most of the services, so maybe just a guide to the existing ones would help.
- To help interact with scientists concerned with observations so that any extra information not provided in the data set can be obtained
- Technical notes, illustrated to instal softwares and extract data
- More explanation of Met office terms for those not used to Met data?
- I think the src_id.xls file could be accompanied by the a file showing (for each site) the periods of sampling by data type, the frequency of measurements and the number of valid (version_num = 1) measurements within that period. It would save a lot of time tracking usable sites.
- Easy to use index to what the data mean! What are the actual columns and WHAT ARE THE CODINGS? I have obviously missed something, but so far I have been unable to find this out and it has prevented me from actually using BADC data.
- More info on datasets that someone without a meteorological background can easily understand. And, more comprehensive information on field variables within the datasets (this specifically applies to the MIDAS data, which is the only BADC dataset that I have used).With regards to MIDAS data it would also be helpful to know things like - are the daily average temperature measurements just an average of the hourly data (both available in separate tables). And, if so how many hourly measurements are required to report a daily average. Also there are differences where (example) wind speed my be unavailable in an hourly table but a daily average wind speed may be available for the same station, and the same day, in a daily average table. Why? ie some explanation on missing data within MIDAS datasets would be helpful.
- Better metadata
- I would appreciate some additional notes when using the MIDAS dataset to explain how some of the data may have been 'managed' from the original data.

Additional Dataset Suggestions

- some more model datasets would be very useful, e.g. 3hourly ECMWF data.
- Land temp data for North America
- I would like to see you store ncep /ecmwf data sets. This is also a huge task and may be beyond your resources. More and more people will use WRF, and ecmwf and ncep data sources are used to start this code.
- more satellite data
- seasonal climate prediction datasets
- What I would like to obtain are spectra of terrestrial radiation obtained by balloons at intervals of 100 metres from the surface of the Earth.
- More archive upper air (radiosonde) data sets
- Infra-red and other energy levels of radiation breakdowns
- Possibly high resolution Landsat / MODIS images of Britain. The University of Maryland has access to some images, but these are infrequent (e.g. a single tile) with no temporal overlap so prevents comparisons such as seasonal vegetation changes or land-use changes (e.g. comparing Landsat between years)
- High resolution radiosonde data from more stations.
- JAIVEx data from ARM facility, 'cause I can't get it from them.
- Mesoscale UM analyses!!!
- ndvi data
- Eumetsat EPS satellite lv1 and lv2 (GOME-2,IASI,AVHRR) (to avoid time intensive GUI frontend and order delays)
- A large array of regional precipitation datasets, eg covering Indian monsoon region
- Earth Observation data (A-train data)
- More data from HadCM - not only the control run.
- More surface radiation data is always good for me!
- Data from the EUMETSAT SAFs e.g. LST from the landsaf
- Archived climate model simulations, i.e. HadCM3 and HadRM3 for time periods other than the 1960-90 climate normal period.
- model forecasts
- global aerosol and dust data

- Physical ocean parameters
- upto date versions of NCEP / ECMWF reanalysis
- MOZAIC chemistry data set (collected from comercial aircraft) would be nice.
- data in SPARC data-centre?
- Data set over West Africa or Africa in general. Planetary Boundary layer data and upper air data with the support softwares for analysing them.
- sea surface global chlorophyll maps (say weekly means)
- Access to full UKMO-UM ensemble and ECMWF ensemble forecast archives.
- More observational ozone data, especially the zone vertical profile data.
- world-wide simultaneous wind speed data and solar irradiance data, all from the same set of years. About 10 should do it as hourly averages.
- More climateprediction.net data NCEP/NCAR reanalysis data
- operational radar data (rather than translated NIMROD data)
- Aerosol optical properties, modeled and/or measured, would be useful for my work
- Run of wind from the MIDAS stations so that we can use the other met data to calculate evoptranspiration and potential evapotranspiration.
- UKCIP08 scenarios