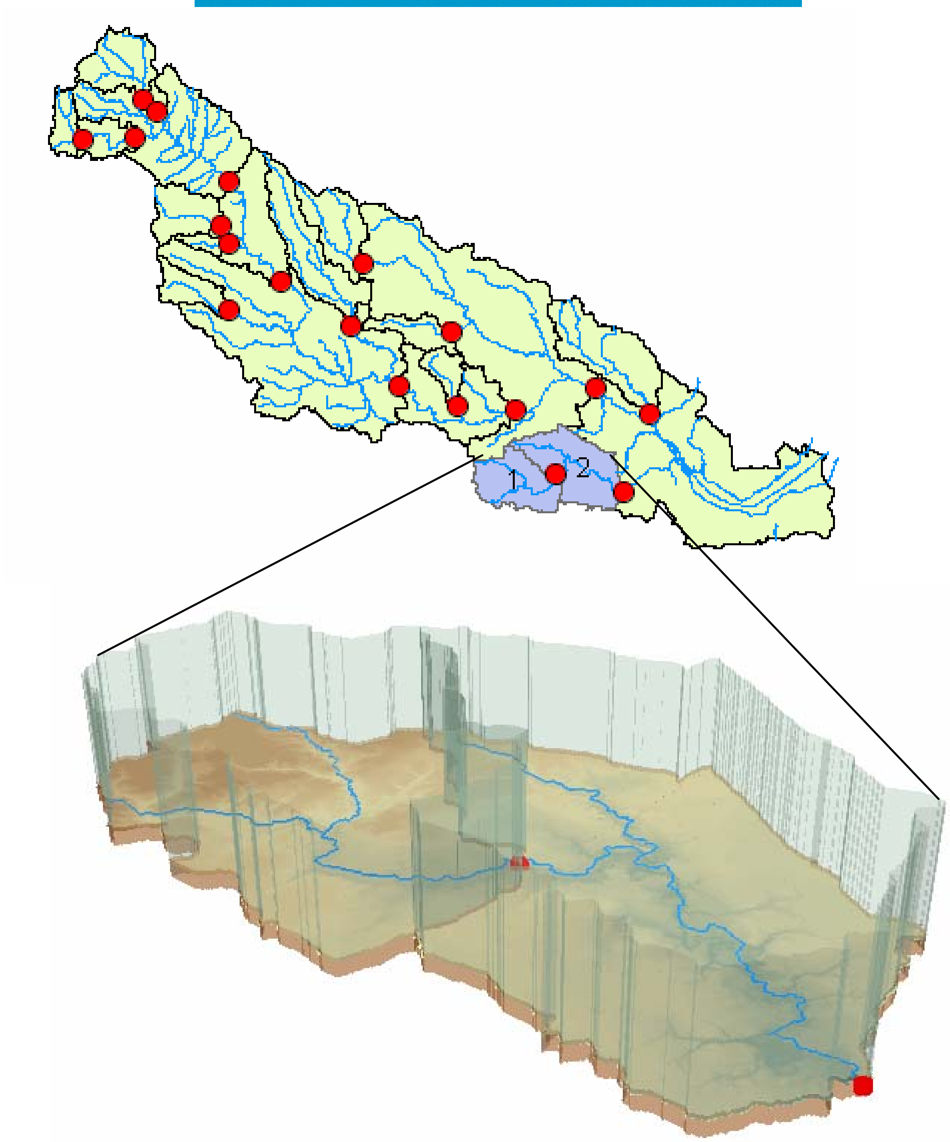
Hydrologic Information System
Status Report

CUAHSI
universities allied for water research



Version 1: September 15, 2005
Edited by David R. Maidment

Table of Contents

Chapter 4

User Needs Assessment

Christina J. Bandaragoda and David G. Tarboton,
Dept of Civil Engineering, Utah State University, Logan UT

David R. Maidment
Center for Research in Water Resources
University of Texas at Austin

Introduction

This chapter reports on a data collection effort targeting the Hydrologic Information System (HIS) User Community: who they are, what they do, and how they do it. Here we present the results of a web-based survey of the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) members and their affiliates which has clarified important HIS development issues, informed HIS project decision-making and will help create an effective, efficient, and functional HIS. CUAHSI is an organization representing more than 100 universities, sponsored by the National Science Foundation to develop infrastructure and services for the advancement of hydrologic science and education in the United States. The CUAHSI Hydrologic Information System (HIS) project is a component of CUAHSI's mission that is intended to improve infrastructure and services for hydrologic information acquisition and analysis. You can learn more about CUAHSI from the website www.cuahsi.org.

We have learned from the survey that there is a definitive, quantifiable need for a Hydrologic Information System. Most researchers spend a significant amount of time preprocessing data for their research and believe an information system such as the CUAHSI HIS would be helpful and relevant to their work. Data services are the most important services for the HIS to provide while addressing critical data use difficulties such as inconsistent data formats, the existence and consistency of metadata, and irregular timesteps.

The overall research objective of the HIS User Needs Assessment process was to assess how hydrologic information is used in research and to assess what functions are of greatest importance among the services that CUAHSI HIS may provide. The information collection focused on three main goals.

- Define the hydrologic data user community.
- Collect raw data on data use patterns, preferences, issues relevant to key decision points in HIS development.
- Prioritize future HIS developments.

The HIS User Assessment survey process included three main steps: preliminary information gathering, a pilot survey, and a web survey. The process began with collaborators to the HIS project team gathering preliminary information from their institutions. This preliminary information was presented at the HIS Symposium at

Austin, Texas, March 2005. At the symposium, a pilot paper survey was conducted using feedback from the preliminary information gathering efforts. The results of both the information gathering and pilot survey were then used to develop the web survey that was conducted in May 2005. This chapter primarily reports results from the web survey. Results from the information gathering and preliminary surveys are included as appendix 4.

Sampling Method

The web-based survey was developed to guide development of the CUAHSI HIS and to assess how hydrologic information is used in research and what functions are of greatest importance among the services that the HIS may provide. The hyperlink was emailed to approximately 100 CUAHSI contacts including representatives of CUAHSI member institutions or participants in CUAHSI sponsored projects or activities with the request that they take the survey as well as forward the request to others at their institution. The questionnaire used is given in Appendix 1. We received 76 responses from researchers at 39 different universities.

Results

Respondents

The HIS User Assessment respondents were approximately 40% hydrologists (surface and groundwater) with a notable majority of respondents identifying themselves as being from other disciplines including water resources, water quality/chemistry, environmental, ecology, atmosphere, GIS/spatial analysis, geomorphology, geology, statistics/mathematics, biology, social science and economics. (Figure 1). The current position of respondents was primarily University faculty (72%), but also included graduate students (20%), University professional/post-doc, working professionals, and others.

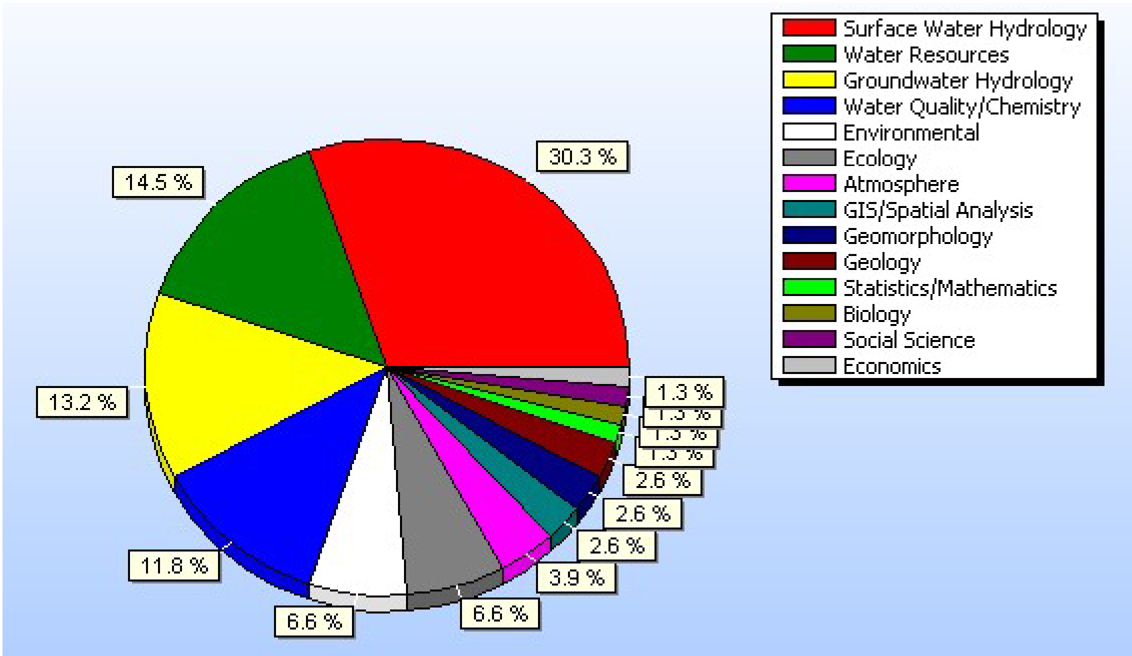


Figure 1. Distribution of respondent fields of research

Software Used for Hydrologic Research

The web survey began with questions related to software used in hydrologic research. Compatibility, inter operability and reliance on open source or professionally supported commercial systems are factors in the design of the HIS. Prior to this survey the distribution of operating system use in the HIS user community was unknown. The results show how many respondents use multiple operating systems (Figure 2). Findings show that 96% of respondents use the Microsoft Windows Operating systems for research, and 36% of respondents also use another operating system in addition to Windows. This shows that although nearly everyone uses Windows, a significant number of researchers also rely on other operating systems for their research.

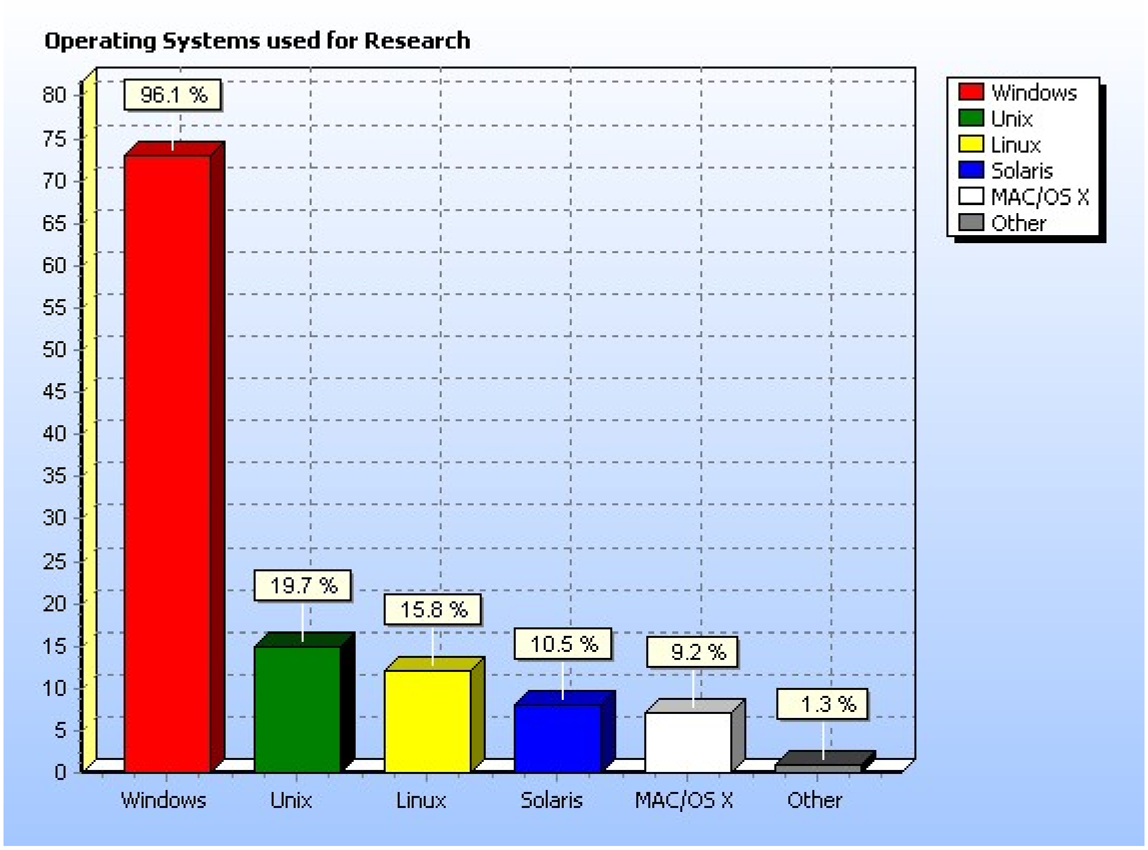


Figure 2. Distribution of operating systems used for hydrologic research. Respondents could indicate more than one operating system, resulting in percentages totaling more than 100%. 36% of respondents indicated one or more non Windows operating systems.

Software Used by category

The HIS may include the capability to interact with other software. Respondents were asked which software packages they use for hydrologic analysis in areas including programming, data management, programming for database client software, GIS (Geographic Information Systems), Mathematics/Statistics, and Hydrologic Models. Our question allowed respondents to select a first and second choice from lists of software programs specific to each category in order to determine which software is the most important among the numerous software programs available. The complete results and weighted average ranking of software programs is provided in Tables 1 through Table 6. Respondents were offered two drop down lists from which to select their first and second choices. Weighted averages were calculated simply with the first choice having twice the weight of the second choice.

Programming

FORTRAN is the most popular programming language used for research, followed by C/C++ and Visual Basic (VB). 85% of respondents indicated a programming language used in their research with 15% indicating programming is 'not applicable to my research'. Considering the first choice in programming languages, FORTRAN is twice as popular as C/C++ or VB. However, for the weighted average of first and second choices, the percentage of users selecting FORTRAN, C/C++ and VB is much closer. Java, Python, AWK and PERL are used by relatively few of the respondents.

Table 1. Programming languages for hydrologic research

Rank	Software	1st Choice	2nd Choice	Weighted Average
1	FORTRAN	42.1%	18.6%	34.3%
2	C/C++	19.7%	27.1%	22.2%
3	Visual Basic	18.4%	23.7%	20.2%
4	not applicable to my research	15.8%	15.3%	15.6%
5	Java	2.6%	5.1%	3.4%
6	Python	1.3%	5.1%	2.6%
7	AWK	0.0%	3.4%	1.1%
8	PERL	0.0%	1.7%	0.6%

Data Management

Microsoft Excel is the most popular software for managing data, followed by Microsoft Access. At least 93% of respondents indicated data management software that they use for their research with between 5% and 7% indicating 'not applicable to my research'. Almost 70% of respondents use Microsoft Excel as their first choice for managing data. This could be due to the simplicity of using Excel for the relatively small datasets common in hydrology. Less than half of the respondents program database client software to access data, but when they do, Visual Basic is primarily used (Table 3).

Table 2. Data Management software for hydrologic research

Rank	Software	1st Choice	2nd Choice	Weighted Average
1	Excel	69.3%	18.2%	52.3%
2	MS Access	10.7%	58.2%	26.5%
3	SQL/Server	12.0%	16.4%	13.5%
4	not applicable to my research	5.3%	7.3%	6.0%
5	PostgreSQL	2.7%	0.0%	1.8%

Table 3. Programming languages used to access database client software. Only a single selection was permitted for this question.

Rank	Software	1st Choice
1	Not applicable to my research	50.8%
2	Visual Basic	23.1%
3	FORTRAN	9.2%
4	C/C++	9.2%
5	Other	3.1%
6	Java	1.5%
7	Perl	1.5%
8	Python	1.5%
9	Awk	0.0%

GIS

ArcGIS (ESRI ArcMap, ArcInfo, ArcView) dominates GIS software use (Table 4). 92% of respondents selected ArcGIS as their first choice with the highest ranking second choice receiving only 30% of second choice selections apart from the 43% who indicated that a second choice was not applicable to their research. Apparently, most respondents rely only on ArcGIS and reliance on other GIS software is rare.

Table 4. GIS software for hydrologic research

Rank	Software	1st Choice	2nd Choice
1	ArcGIS (ESRI ArcInfo, ArcView, etc)	92.1%	0.0%
2	not applicable to my research	6.6%	43.3%
3	IDRISI (Clark Labs)	1.3%	13.3%
4	MapInfo	0.0%	30.0%
5	GRASS	0.0%	13.3%
6	TAS	0.0%	0.0%

Mathematics/Statistics

Matlab is the most popular software for mathematics and statistics, followed by Microsoft Excel and SAS, but there is a wide variability in software used (Table 5). Mathematics/Statistics software programs are used by at least 97% of respondents (less than 3% reported that use of a software program in this category was not applicable to their research). Matlab is the first choice of 42% of respondents, Excel is the first choice of only 24% of respondents. However, the difference in the weighted average between Matlab and Excel is only 9%. If use of mathematics/statistics software were to be incorporated into the HIS, both Matlab and Excel would need to be accommodated.

Table 5. Mathematics/Statistics software for hydrologic research

Rank	Software	1st Choice	2nd Choice	Weighted Average
1	Matlab	41.3%	19.0%	33.9%
2	Excel	24.0%	25.4%	24.5%
3	SAS	10.7%	11.1%	10.8%
4	SPSS	5.3%	11.1%	7.2%
5	R (Open Source Splus)	2.7%	14.3%	6.6%
6	Mathematica	5.3%	6.3%	5.6%
7	Minitab	2.7%	4.8%	3.4%
8	IDL	2.7%	3.2%	2.9%
9	Splus	1.3%	3.2%	1.9%
10	not applicable to my research	2.7%	1.6%	2.3%
11	Scilab (Open Source Matlab)	1.3%	0.0%	0.9%

Hydrologic Models

80% of respondents indicated that they use hydrologic models in their research, however the models used vary widely. The most important result reported in Table 6 may be that 'not applicable to my research' was the highest ranking response for choice in hydrologic model. Modflow is the most popular groundwater model but there is no predominant surface water model. A general, simple, standard, and open interface that could connect with many systems would be the only way to accommodate all of the models used.

Table 6. Hydrologic Models used in hydrologic research

Rank	Software	1st Choice	2nd Choice	Weighted Average
1	not applicable to my research	21.9%	15.3%	19.7%
2	Modflow/Visual Modflow	19.2%	16.9%	18.4%
3	U.S. Army Corps HEC models	11.0%	10.2%	10.7%
4	GMS Groundwater Modeling System	8.2%	11.9%	9.4%
5	TOPMODEL	11.0%	8.5%	10.2%
6	Sacramento/NWS/HSPF	5.5%	8.5%	6.5%
7	SMS Surface Water Modeling System	2.7%	6.8%	4.1%
8	SHE System Hydrologique European/Mike-SHE	0.0%	8.5%	2.8%
9	Groundwater Vistas	4.1%	5.1%	4.4%
10	TIN-based real time Integrated Basin Simulator (tRIBS)	4.1%	1.7%	3.3%
11	EPA Basins	2.7%	5.1%	3.5%
12	WMS Watershed Modeling System	2.7%	0.0%	1.8%
13	SWAT	4.1%	0.0%	2.7%
14	MMS/PRMS	2.7%	1.7%	2.4%

The questionnaire included space for respondents to list other software packages that should be considered for interfacing with the CUAHSI HIS. The complete list of responses received is given in Appendix 2. These responses mention a total of 35 additional software packages.

An interesting result from our preliminary survey at the CUAHSI HIS Symposium at Austin, Texas, March 2005, came from the comparison of all software programs without restriction to categories. Respondents were asked to rate each software program between 1 and 5, where 1 is "never use or do not find useful" and 5 is "use frequently and find indispensable". Figure 3 presents these results which show that Excel and ArcGIS scored highest as the two most popular software programs for hydrologic research among the Symposium participants.

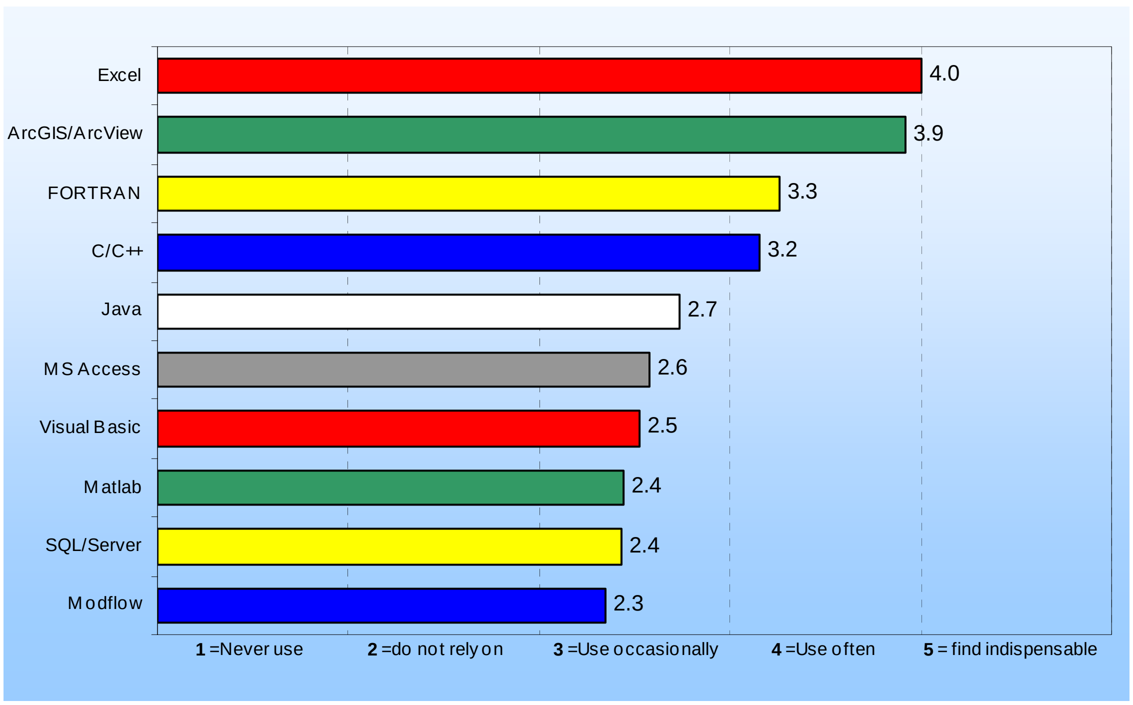


Figure 3. Rating of software packages and programming languages with respect to how important they are for hydrologic analysis. Results taken from preliminary survey at the CUAHSI HIS Symposium held in Austin, Texas, March 2005. Sample size n=39.

In building the CUAHSI HIS choices need to be made with respect to reliance on the capability of existing software, both proprietary and open source. Reliance on other software takes advantage of existing technology, avoids the need to repeat existing capability and may be more reliable and have professional support and maintenance. Respondents were asked opinions regarding the selection of open source or commercial software platforms for the CUAHSI HIS. Respondents are predominantly in favor of HIS client software being open source, but at the same time would like to leverage commercial software and have the capability to work on all operating systems (Table 7).

Table 7. Opinions on software development.

	Strongly Disagree	Disagree	Agree	Strongly Agree	No Opinion
HIS Client software should work on all computer operating systems	2.7%	10.7%	49.3%	30.7%	6.7%
HIS Software should leverage commercial software systems	2.7%	9.3%	40.0%	22.7%	25.3%
HIS Software should be open source	1.3%	8.0%	29.3%	42.7%	18.7%

In addition to opinions regarding the software platform issue, we were interested to know which issues researchers were concerned about when considering the use of open source or commercial software platforms. If the community has strong preferences for open source or proprietary software, it is useful to know why, or which concerns need to be addressed when decisions are made by the HIS development team. We developed the following list of common issues related to choice of operating platform and asked respondents to rank the three most important to them:

- Cost of commercial software required by the HIS user to exploit full HIS capability.
- Long term stability of commercial software and continuation of support by provider
- Existence of support and upgrade options for open source solutions
- Flexibility to scrutinize and modify source code
- The professional support provided by commercial software
- The functionality available in commercial software

The results show that the cost to the user of commercial software required to use the HIS is the greatest concern. This is closely followed by concern that the HIS have the stability, long-term support, and functionality available in commercial software (Figure 4).

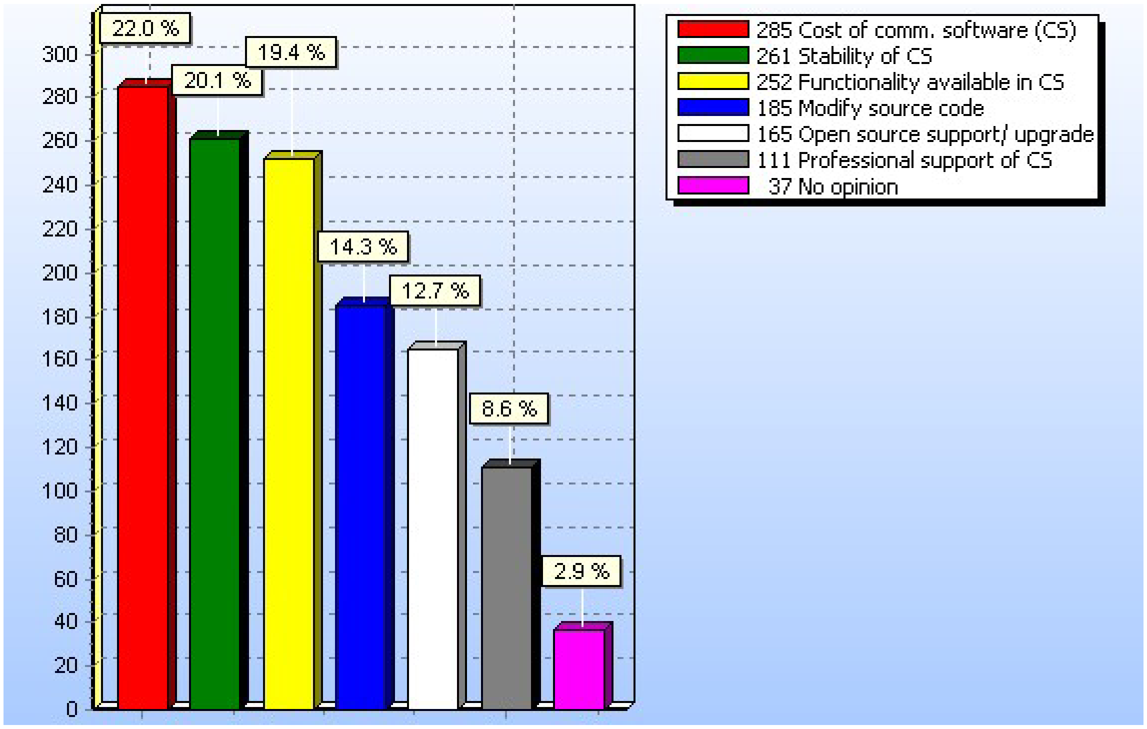


Figure 4. Importance of issues related to use of commercial and open source software (using a value score where first choice has a score of 3 points, second choice has a score of 2 points, and third choice has a score of 1 point).

Hydrologic Data Acquisition and Preparation

To understand the current patterns in hydrologic data acquisition and preparation, we asked what proportion of research time is spent preparing or preprocessing data into appropriate forms needed for research purposes. A significant fraction of research time is spent preparing and preprocessing data (Figure 5).

- More than 80% of respondents spend more than10% of research time preparing data.
- More than 35% of respondents spend more than 25% of research time preparing data.
- More than 12% of respondents spend more than 50% of research time preparing data.

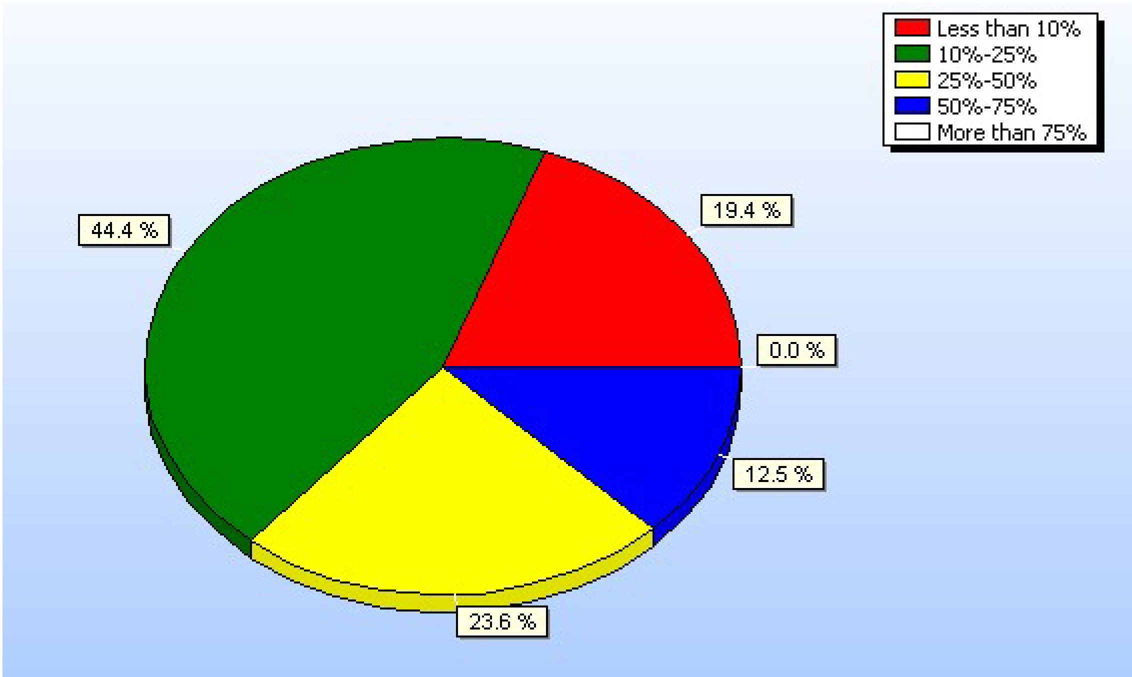


Figure 5. Proportion of research time spent preprocessing or preparing data.

A matrix of datasets that the CUAHSI HIS may incorporate was presented with four choices for rating the priority of each dataset for inclusion in the HIS. For each dataset, the respondent could choose 1) Essential to my research, 2) Am likely to use in my research, 3) I am aware of this, but not likely to use it, and 4) I have not heard of this dataset. The following datasets ranked highest for incorporating into the CUAHSI HIS[2].

1. USGS Streamflow
2. NCDC Precipitation
3. Remote Sensing data (e.g. LANDSAT, GOES, AVHRR)
4. National Elevation Dataset and derivatives (EDNA)
5. Other NCDC Weather and Climate Data
6. USGS Groundwater levels
7. National Land Cover dataset (NLCD)
8. Soils Data (STATSGO/SSURGO)
9. National Hydrography Dataset (NHD)
10. NCDC Pan Evaporation

The tabulated results (Table A3.1) and responses to the question about additional datasets to consider for inclusion in the HIS are listed in Appendix 3.[3] Isotope data,

---

[2] Rank was determined using a weighted average, (“Essential”*2 + “Likely to Use”)/3

[3] Other datasets respondents were asked to rank, but which scored lower, include in the following order: USGS Water Chemistry (NASQAN, HBN, Cooperative data), SNOTEL, EPA STORET Water Quality, NEXRAD Radar Precipitation, National Water Quality Assessment (NAWQA), Biological Data, USGS National Geology data, USGS Hydrologic Landscape Regions, PRISM Precipitation data, Climate Model Reanalysis data (e.g. NARR), Aquatic Ecoregions (AQUAECO), and Acidic Surface Waters (A_WATER).

water use and management data, listed among the additional dataset responses, are important datasets not previously identified.

In addition to ranking the above datasets for inclusion in the HIS, we also asked respondents which of these datasets are the most difficult to access and use (Figure 6). The HIS can provide a service to researchers by facilitating the dissemination of important data that is currently challenging to utilize. The top four datasets respondents believed would most benefit from increased ease of access through a Hydrologic Information System are:

1. EPA STORET Water Quality
2. USGS Streamflow
3. Remote Sensing data (e.g. LANDSAT, GOES, AVHRR)
4. NEXRAD Radar Precipitation

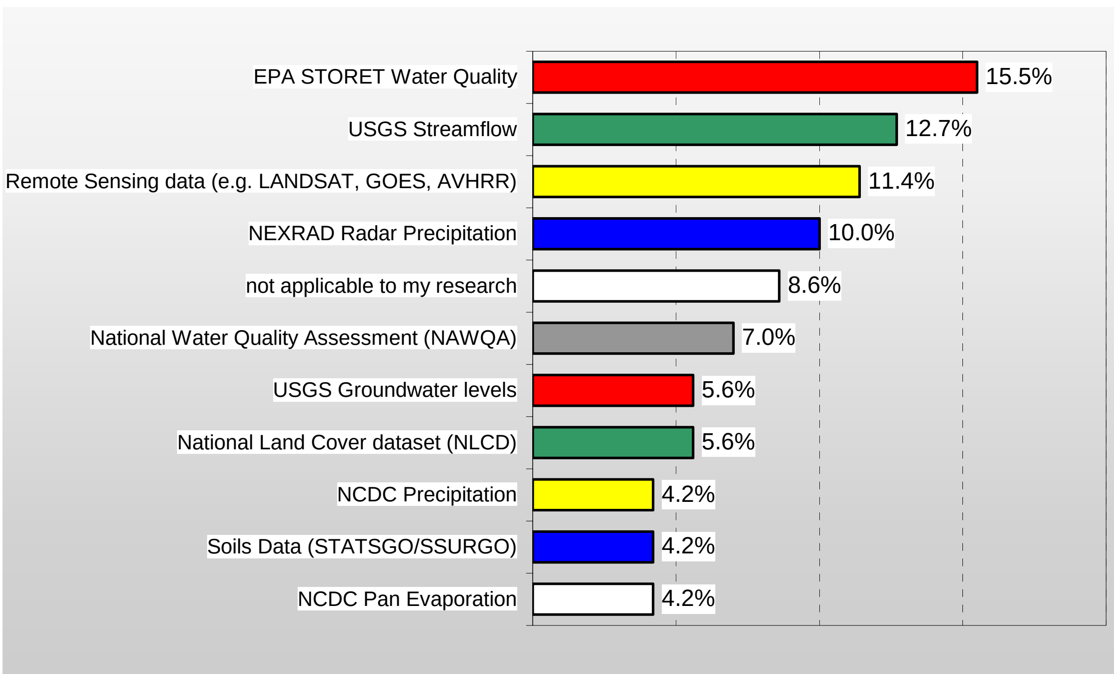


Figure 6. Datasets that are difficult to access and use which would most benefit from increased ease of access through a HIS. All datasets that appear in Table A3.1 were available for respondents to select.

Respondents were asked which spatial scales are most relevant to the data resolution used in their research with the option to select multiple responses. The watershed scale was indicated as most relevant, followed closely by the Field and Sub-watershed resolutions (Figure 7). The prevalence of researchers investigating questions at more than one spatial scale indicates the importance for the HIS to integrate datasets so they may be used at multiple spatial scales. Respondents are predominantly interested in studies at a watershed scale or smaller.

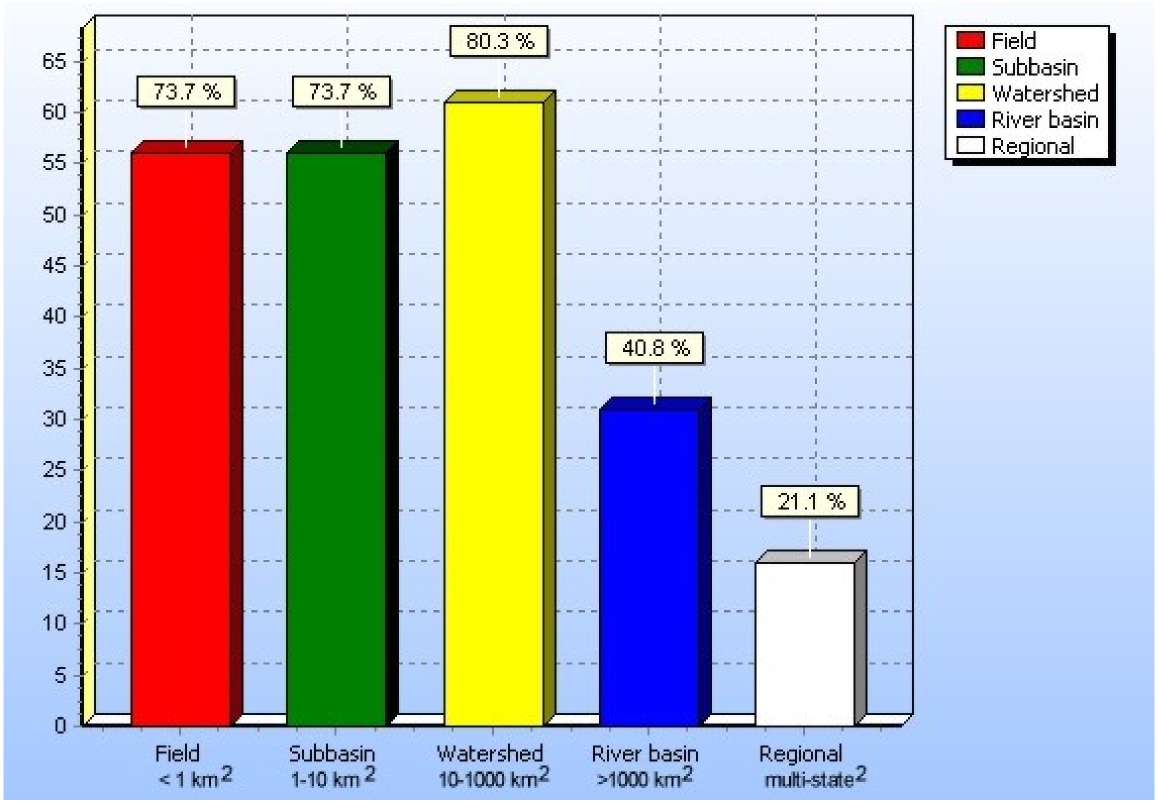


Figure 7. Spatial scales most relevant to data resolution used in research. Respondents could indicate more than one scale, resulting in percentages totaling more than 100%.

The HIS hopes to improve capability for integrating, analyzing, and synthesizing data from disparate sources. Respondents were asked to choose and rank three of the most critical difficulties for the HIS to address. Results show that the HIS under development must address the following common difficulties encountered when using hydrologic data for research (Figure 8):

1. Inconsistent data formats
2. Existence and consistency of metadata
3. Irregular and different timesteps

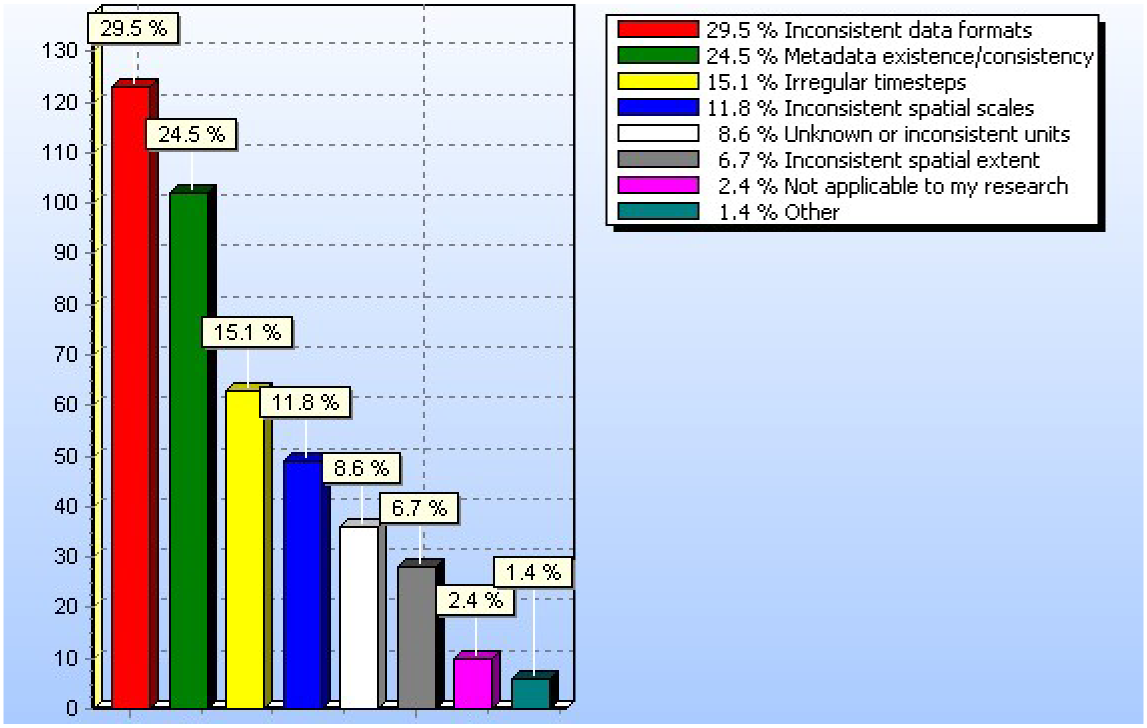


Figure 8. Difficulties in integrating, analyzing and synthesizing data that should be addressed by the HIS (using a value score where first choice has a score of 3 points, second choice has a score of 2 points, and third choice has a score of 1 point).

HIS Services

The CUAHSI HIS has under consideration four main categories of service to the hydrologic research community:

1. Hydrologic data services – these are services that can be used by a hydrologic researchers or students anywhere in the nation to obtain the hydrologic data they require quickly and easily, and in forms that they can readily use;

2. Hydrologic observatory services – these are information services that a CUAHSI Hydrologic Observatory will require to process, archive and display the data measured at the observatory;

3. Hydrologic science services– these are services needed to build the complex digital representations of hydrologic environments needed to support advanced hydrologic modeling, hypothesis testing, and constructing water, energy and mass balances of hydrologic systems;

4. Hydrologic education services – these are services needed to advance the use of hydrologic information in the classroom.

Respondents were asked to rank these four goals. Results in Figure 9 identify Data Services as the functionality that users perceive to be most important for the HIS to provide.

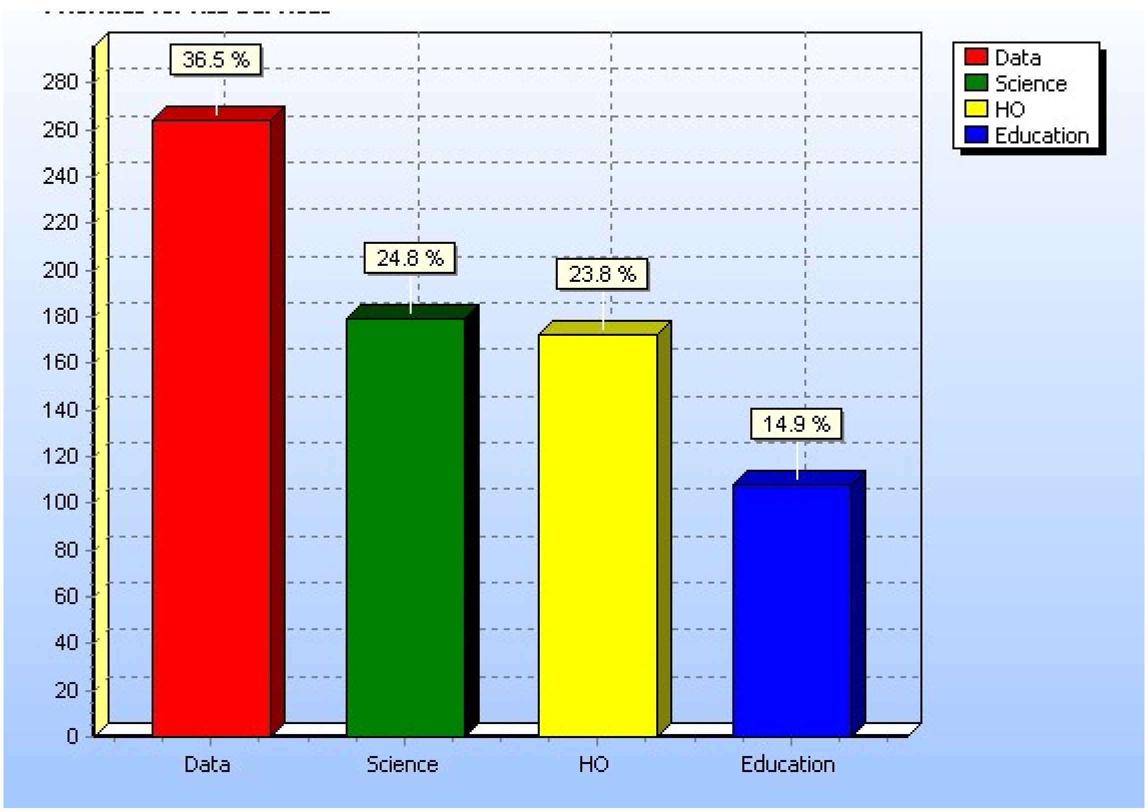


Figure 9. Priorities for general categories of HIS Services. (using a value score where first choice has a score of 4 points, second choice has a score of 3 points, third choice has a score of 2 points, fourth choice has a score of 1 point).

The Need for HIS

Comments about the HIS project show that respondents are excited that the project is moving forward, think it is a great idea and will do much to advance hydrologic science. The HIS is expected to advance the quality of the hydrologic sciences and determine the nature of future research projects. When asked if a CUAHSI HIS was developed with the priorities the respondent had listed for a watershed where they conduct research, 77% responded ‘Yes, a user-friendly digital watershed for data access is what I need’. One would expect that the majority of respondent might answer ‘Yes’ since they are interested enough in the topic to take the time to answer the questionnaire. The important result shown in Figure 10 is that there are many who do not know enough about the CUAHSI HIS project to determine its expected utility. Almost 18% of respondents did not feel that had enough information to know if they would use the HIS or not. This speaks to a need for better communication and articulation to the community of the potential capabilities of the HIS.

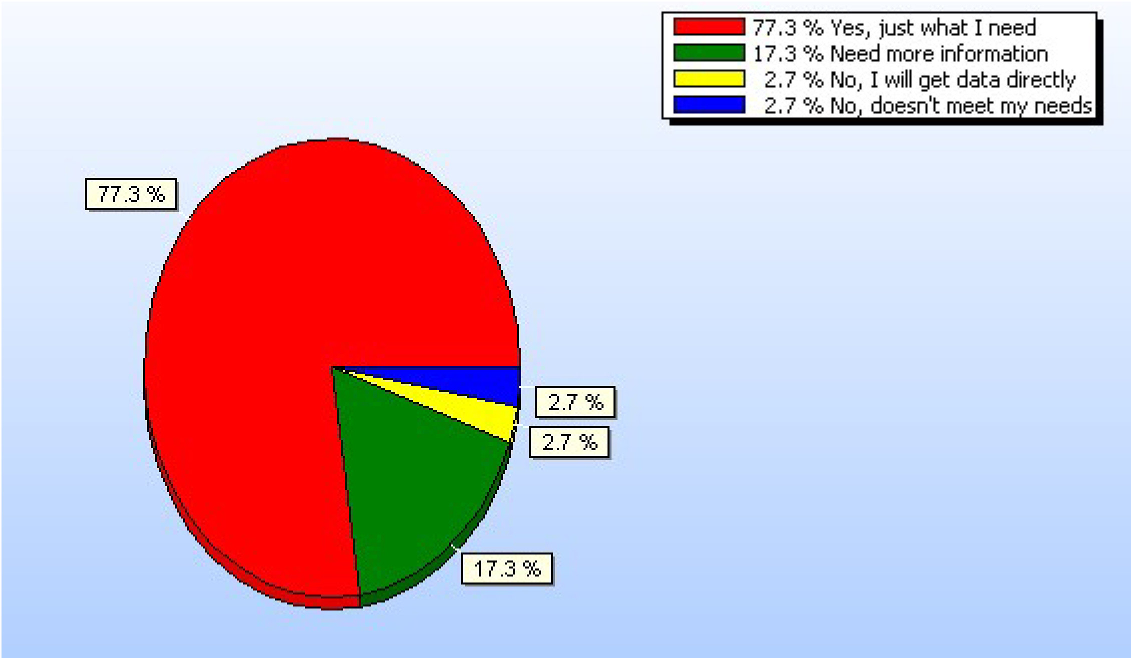


Figure 10. Anticipated use of CUAHSI HIS and the need for more information dissemination.

Conclusions and Recommendations

Our results verify the need for an HIS. Most researchers surveyed spend a significant amount of time preprocessing data for their research and expect that the HIS will be helpful and relevant to their work. Data services are the most important services for the HIS to provide while addressing critical data use difficulties such as inconsistent data formats, the existence and consistency of metadata, and irregular timesteps. Most agree that the HIS client software should work on all operating systems and that it should leverage commercial software as well as be open source. The cost related to the use of the HIS, and its long-term stability and functionality are critical issues to be addressed in the HIS development.

In hydrologic research, FORTRAN is the most popular programming language, Microsoft Excel most used for data management, Matlab is most used for mathematics and statistics work. Most respondents use GIS for their research (93%) and ESRI ArcGIS is the most used software. The survey shows that 80% of respondents use hydrologic modeling in their research. Modflow is the most popular for groundwater modeling, but there is no predominant surface water model of the more the 25 hydrologic models listed in the survey questions and mentioned by respondents. A general, simple, standard, and open interface that could connect with many systems would be the only way to accommodate all of the models used.

USGS Streamflow and NCDC Precipitation are viewed as the most essential datasets for inclusion in the HIS. Remote sensing data (e.g. LANDSAT, GOES, AVHRR), National Elevation Dataset and derivatives (EDNA), Other NCDC Weather and Climate Data, and USGS Groundwater levels also scored as high priorities. Datasets considered the most difficult to access and use that would most benefit from inclusion in the HIS include EPA STORET Water Quality Data, USGS Streamflow, Remote Sensing data, and NEXRAD Radar precipitation. Results show that most researchers work at multiple spatial scales, predominantly at the field, sub-basin, and watershed levels. The HIS should provide datasets in formats that can be easily utilized at a variety of spatial scales.

During the process of developing the HIS User Needs survey, three levels of HIS user information were outlined: 1) how people are currently conducting hydrologic analysis 2) how researchers believe an HIS could help them conduct hydrologic analysis with detailed HIS service prioritization, and 3) how data providers from other public and private institutions are currently assisting in HIS-like endeavors as well as how CUAHSI can coordinate HIS efforts with them. This paper focused on a web survey intended to collect the first level of information or a ‘client-based’ survey of software and data use patterns and current practices in hydrologic analysis by researchers at CUAHSI institutions. Future HIS user surveys should focus on prioritization of specific HIS services and how to best coordinate with other public and private providers of hydrologic data.

Acknowledgements

We would like to thank Richard Hooper, Jon Duncan, Xu Liang, Yao Liang, Chunmaio Zheng and Leroy Poff for their contribution and assistance with CUAHSI HIS User Needs preliminary surveys that led to the development of the web survey on which this paper is based.

Appendix 1. Web Questionairre used for HIS User Needs Assessment, May 2005. Values presented in drop down boxes are given with the presentation of results in Tables, Appendices and Figures.

CUAHSI

Hydrologic Information System (HIS) User Needs Assessment

The CUAHSI Hydrologic Information System (HIS) project is a component of CUAHSI's mission that is intended to improve infrastructure and services for hydrologic information acquisition and analysis. This user survey is designed to solicit input on ways to best focus efforts in developing the CUAHSI Hydrologic Information System. We want to understand how you use hydrologic information and how the technology, infrastructure and services that we are creating could be applied to help you accomplish your goals better. The questions in this survey are designed to quantify the relative priorities of various kinds of computing environments, software systems, information sources and services that the CUAHSI HIS could offer.

We appreciate you taking the time to fill out this survey. Results from this survey will be published in a CUAHSI HIS report and used for determining the direction of future HIS developments. To learn more about the CUAHSI HIS, visit our HIS homepage. If you have any other comments or guidance to offer please contact David Maidment at maidment@mail.utexas.edu.

We are requesting your name and institution in case we need to verify survey results. This information will remain strictly confidential, will not be used for any purpose beyond this survey, and you will not be identified in results reported from this survey.

Name

Institution

Which of the following most closely describes your field of research?

Which of the following most closely matches your current position?

Please indicate your affiliation. Select all that apply.

- [ ] CUAHSI Representative
- [ ] Participant in CUAHSI project or activity
- [ ] Selected at random to take this survey
- [ ] Other

Next Page (1 of 4)

CUAHSI

Hydrologic Information System (HIS) User Needs Assessment

Software Used for Hydrologic Research

Compatibility, inter operability and reliance on open source or professionally supported commercial systems are factors in the design of the HIS. These questions are intended to survey what computer systems and software are in use by the hydrology community and preferences and philosophies regarding software.

Which operating systems do you use for your research? If you use more than one operating system, select all that apply.

- [ ] Windows
- [ ] MAC/OS X
- [ ] Linux
- [ ] Solaris
- [ ] Unix

☐ Other (please specify)

If you selected other, please specify:

In the questions below we are asking that you rank selections picking a few that are highest priority to you. This form of question is designed to have you to tell us which is most important among the choices presented.

The HIS may include the capability to interact with other software. Please choose up to two software packages in each category that are most important for hydrologic analysis in your research.

Programming

1	
2	

Data Management

1	
2	

If you program database client software to access data, which programming language do you use?

If you selected other, please specify:

GIS (Geographic Information Systems)

1	
2	

Mathematics/Statistics

1	
2	

Hydrologic Models

1	
2	

Are there any other software packages we should consider for interfacing with the CUAHSI HIS?

In building the CUAHSI HIS choices need to be made with respect to reliance on the capability of existing software, both proprietary and open source. Reliance on other software takes advantage of existing technology, avoids the need to repeat existing capability and comes with professional reliability and support. Disadvantages include the costs to users (e.g. for a needed commercial system like Windows or GIS), lack of flexibility (e.g. to change something and access the code) and dependence on the business strategies of the provider.

Please indicate your opinion in the selection of commercial/open source software platforms for the CUAHSI HIS.

	Strongly Disagree	Disagree	Agree	Strongly Agree	No Opinion
HIS Client software should work on all computer operating systems					
HIS Software should leverage commercial software systems					
HIS Software should be open					

source					

In considering the choice between the open source and commercial software model for the HIS, the following considerations arise. Please indicate and rank the three issues that are most important to you.

- Cost of commercial software required by the HIS user to exploit full HIS capability.
- Long term stability of commercial software and continuation of support by provider
- Existence of support and upgrade options for open source solutions
- Flexibility to scrutinize and modify source code
- The professional support provided by commercial software
- The functionality available in commercial software

1)

2)

3)

Next Page (2 of 4)

CUAHSI

Hydrologic Information System (HIS) User Needs Assessment

Hydrologic Data Acquisition and Preparation

What proportion of your research time do you spend on preparing or preprocessing data into appropriate forms needed for research purposes?

- ☐ Less than 10%
- ☐ 10%-25%
- ☐ 25%-50%
- ☐ 50%-75%
- ☐ More than 75%

Consider the following datasets that the CUAHSI HIS may incorporate. Please rate each dataset for its priority for inclusion in the HIS.

	I have not heard of this dataset	I am aware of this, but not likely to use it	Am likely to use in my research	Essential to my research
USGS Streamflow	☐	☐	☐	☐
USGS Groundwater levels	☐	☐	☐	☐
USGS Water Chemistry (NASQAN, HBN, Cooperative data)	☐	☐	☐	☐
National Water Quality Assessment (NAWQA), Biological Data	☐	☐	☐	☐
EPA STORET Water Quality	☐	☐	☐	☐

NCDC Precipitation	☐	☐	☐	☐
NCDC Pan Evaporation	☐	☐	☐	☐
Other NCDC Weather and Climate Data	☐	☐	☐	☐
SNOTEL	☐	☐	☐	☐
National Elevation Dataset and derivatives (EDNA)	☐	☐	☐	☐
National Hydrography Dataset (NHD)	☐	☐	☐	☐
Soils Data (STATSGO/SSURGO)	☐	☐	☐	☐
USGS National Geology data	☐	☐	☐	☐
National Land Cover dataset (NLCD)	☐	☐	☐	☐
USGS Hydrologic Landscape Regions	☐	☐	☐	☐
NEXRAD Radar Precipitation	☐	☐	☐	☐
PRISM Precipitation data	☐	☐	☐	☐
Climate Model Reanalysis data (e.g. NARR)	☐	☐	☐	☐
Remote Sensing data (e.g. LANDSAT, GOES, AVHRR)	☐	☐	☐	☐
Aquatic Ecoregions (AQUAECO)	☐	☐	☐	☐
Acidic Surface Waters (A_WATER)	☐	☐	☐	☐

Are there any other datasets you think should be included in a HIS?

What spatial scales are most relevant to data resolution used in your research? Check all that apply.

- [ ] Field/Site/Point/Project (< 1 square kilometer)
- [ ] Sub-watershed (1-10 square kilometer)
- [ ] Watershed (10-1000 square kilometer)
- [ ] River basin(> 1000 square kilometer)
- [ ] Regional (multi-state)

Some datasets used in your research may be difficult to access and use. Please indicate one dataset that you believe would most benefit from increased ease of access through a Hydrologic Information System (HIS).

The HIS hopes to improve capability for integrating, analyzing, and synthesizing data from disparate sources. In your experience, which of the following difficulties are most important for HIS to address?

- Existence and consistency of meta data
- Inconsistent data formats
- Inconsistent spatial scales
- Inconsistent spatial extent
- Unknown or inconsistent units
- Irregular and different timesteps

Please indicate the three most critical for the CUAHSI HIS to address.

1)
2)
3)

Next Page (3 of 4)

Hydrologic Information System (HIS) User Needs Assessment

HIS Services

There are many potential services that HIS can fulfill. These include:

1. Hydrologic data services – these are services that can be used by a hydrologic researchers or students anywhere in the nation to obtain the hydrologic data they require quickly and easily, and in forms that they can readily use;

2. Hydrologic observatory services – these are information services that a CUAHSI Hydrologic Observatory will require to process, archive and display the data measured at the observatory;

3. Hydrologic science services– these are services needed to build the complex digital representations of hydrologic environments needed to support advanced hydrologic modeling, hypothesis testing, and constructing water, energy and mass balances of hydrologic systems;

4. Hydrologic education services – these are services needed to advance the use of hydrologic information in the classroom.

Please rank these four HIS service categories for helping you.

1)
2)
3)
4)

If a CUAHSI HIS was developed with the priorities you have listed for a watershed where you conduct research, would you use it?

- ☐ Yes, a user-friendly digital watershed for data access is what I need.
- ☐ No, it doesn't meet my current research needs
- ☐ No, I would rather get data directly from various data providers.
- ☐ I don't have enough information to know.

Please provide any additional comments or suggestions regarding the HIS project.

Thank you for the information you have provided!

Submit Survey (4 of 4)

<u>Appendix 2</u>: Responses to question "Are there any other software packages we should consider for interfacing with the CUAHSI HIS?" (raw data)

- Maple
- Rivertools
- I use JMP for data management and statistics
- PHREEQC,NETPATH
- Rockware Rockworks Geochemists Workbench Surfer ModFlow
- AGWA
- Math: Mathcad, Scientific Notebook Curve fitting: TableCurve Parameter Estimation: Pest, UCODE Hydrologic Modeling: TOUGH and related
- FEFLOW
- Kineros model
- ENVI
- U.S. EPA Stormwater Management Model (SWMM5)
- Chemical equilibrium and speciation software such as PHREEQC
- HSPF
- CUENCAS being developed by our group at the univerrsity of colorado, boulder
- IPW (Frew/UCSB) ISNOBAL many others -- be flexible
- AutoCAD
- Highest priority is an efficient map based web interface. Second priority is a programmable application interface so that data access can be programmed and scripted.
- User developed data analysis, data visualization, or modeling applications.
- GAMS (for solving dynamic systems of equations) LIMDEP (statistical software) STATA (statistical software) GAUSS (statistical software)
- ERDAS Imagine a raster based image processing and GIS software with modeling capabilities.
- Imagine ERDAS for GIS image processing
- Penn State Integrated Hydrology Model
- For my research, I use the statistical software package Statistica (Statsoft, Tulsa, OK) extensively, Argus One
- MLAEM, Split, GFlow, TwoDAN, 3DFlow, Tim

Appendix 3. Important Datasets for inclusion in the CUAHSI HIS

Table A3.1 Datasets CUAHSI may incorporate ranked by priority ratings.

Rank		Essential to my research	Am likely to use in research	I am aware of this, but not likely to use	I have not heard of dataset	Weighted[1] Score
1	USGS Streamflow	60.8%	27.0%	9.5%	2.7%	49.5%
2	NCDC Precipitation	35.1%	44.6%	12.2%	8.1%	38.3%
3	Remote Sensing data (e.g. LANDSAT, GOES, AVHRR)	30.7%	34.7%	32.0%	2.7%	32.0%
4	National Elevation Dataset and derivatives (EDNA)	32.4%	31.1%	24.3%	12.2%	32.0%
5	Other NCDC Weather and Climate Data	22.5%	47.9%	21.1%	8.5%	31.0%
6	USGS Groundwater levels	25.7%	39.2%	28.4%	6.8%	30.2%
7	National Land Cover dataset (NLCD)	21.3%	46.7%	28.0%	4.0%	29.8%
8	Soils Data (STATSGO/SSURGO)	20.0%	44.0%	26.7%	9.3%	28.0%
9	National Hydrography Dataset (NHD)	25.3%	32.0%	22.7%	20.0%	27.5%
10	USGS Water Chemistry (NASQAN, HBN, Cooperative data)	20.0%	38.7%	32.0%	9.3%	26.2%
11	NCDC Pan Evaporation	18.7%	41.3%	29.3%	10.7%	26.2%
12	SNOTEL	26.0%	17.8%	35.6%	20.5%	23.3%
13	EPA STORET Water Quality	18.7%	32.0%	30.7%	18.7%	23.1%
14	National Water Quality Assessment (NAWQA), Biological Data	16.9%	35.2%	36.6%	11.3%	23.0%
15	NEXRAD Radar Precipitation	18.7%	30.7%	38.7%	12.0%	22.7%
16	USGS National Geology data	13.5%	33.8%	36.5%	16.2%	20.3%
17	PRISM Precipitation data	12.0%	29.3%	33.3%	25.3%	17.8%
18	USGS Hydrologic Landscape Regions	9.3%	33.3%	38.7%	18.7%	17.3%
19	Climate Model Reanalysis data (e.g. NARR)	6.7%	18.7%	38.7%	36.0%	10.7%
20	Aquatic Ecoregions (AQUAECO)	1.4%	16.4%	32.9%	49.3%	6.4%
21	Acidic Surface Waters (A_WATER)	1.3%	4.0%	29.3%	65.3%	2.2%

[1] Note: The weighted score is ("Essential"*2 + "Likely to Use")/3

Responses to question "Are there any other datasets you think should be included in a HIS?" (raw data)

- Isotope data
- A lot of USGS stations are being farmed to state or local resource agencies. This likely will continue. Is there any way to include these stations?
- SRTM! Calif Coop Snow Survey
- Other elevation data sets (state)
- USGS DRG Maps (Terraserver topo and aerial images)
- The many regional datasets that, for a given area, are far superior to the national datasets. The HIS needs the best coverage, not just national coverages, which will be what folks use only if they cannot get something better.
- Quantification of diversions, water use, irrigation
- Local agency data. Diversion flows in managed systems. Reservoir levels and volumes. GIS base layers (administrative boundaries, watershed boundaries, etc.) There are other existing sources of climate information. Many of these are publicly available, and may provide different data than NCDC.
- DEM digital elevation model data

- State data sets on water/use, socio-economic data
- snowcourse (whenever feasible), oceanography data (i.e SEASAT, circulation, etc)

Appendix 4: Preliminary Information Gathering and Pilot Survey

Preliminary Information Gathering

In order to develop a thorough and relevant web-based survey to collect specific information on HIS User Needs, HIS project collaborators were requested to conduct pre-surveys at their home institutions. Four separate surveys were conducted at Utah State University, University of California at Berkley, Virginia Tech University and the University of Alabama. Surveys were independently designed by HIS project collaborators and results were presented at the March 20, 2005 HIS Symposium in Austin, Texas by HIS Project collaborators: David Tarboton (USU), Xu Liang (UCBerkley), Yao Liang (Virginia Tech), Chunmaio Zheng (University of Alabama). LeRoy Poff (Colorado State University) presented perspectives representing Biology and Ecology HIS users.

The pre-surveys served as an initial information gathering effort. The questions developed by collaborators focused on fulfilling project goals, but how the questions were presented was left to their discretion. This created broad information on HIS User Needs, and helped focus the surveys that followed. There were three main goals for the preliminary information gathering:

- Set clear objectives for the web-based information collection; decide what information was necessary to collect from users
- Create a definition of the 'HIS User Community' in order to target the correct population or sample for the web-based survey
- Design clear questions in unambiguous formats

With the use of preliminary data collected from different institutions, we obtained feedback on which open-ended questions are relevant and began development of a survey that probes deeper into how data is used by different groups. In this appendix we present results from the Utah State University and University of California at Berkeley pre surveys as well as the pilot survey conducted at the March 2005 HIS Symposium in Austin.

Utah State University Preliminary Survey

At Utah State University, the mode of data collection was a questionnaire distributed by email to members of the campus water sciences community through the Water Initiative. The Water Initiative is a framework for all the Water Sciences at USU and includes a group of physical, biological, and social scientists, and engineers located in six colleges, multiple departments, and academic units whose careers focus on water-related science, engineering, and policy problems. Surveys were received back by eighteen respondents representing research fields in Hydrology, Watershed Science, Geomorphology, Fisheries, Biogeochemistry, Aquatic Ecology, Water resources engineer (includes irrigation), Environmental Engineering, Meteorology, Ecology, Natural Resources Sociology, and Remote Sensing/GIS. Twelve of the respondents were faculty, five graduate students, and one university professional.

Results

At USU, Windows is the most popular operating system used for research (80%) with some respondents also using Unix, Linux, and MAC/OS. Microsoft Excel and ESRI ARCGIS/Arcview are the most used software programs used by USU researchers; C++, Visual Basic, Matlab, MS Access and Fortran are also highly utilized. USU respondents were presented with a set of 10 software functions that could potentially be included in an HIS and asked to score each on a scale between 1 and 5 (where 1=Never use or do not find useful; 2=Have used but do not rely on this; 3=Use occasionally and am comfortable with its use; 4=Use often; 5=Use frequently and find indispensable). The following software functionality was the most important (score > 4.0) for including in an HIS: data storage and retrieval, visualization of spatial data, visualization of time series data, and building relational links. The software functionality USU respondents felt was less important (score <4.0) included: efficient coupling with 3rd party analysis software, presentation to non-technical audiences, development of publication quality figures, numerical analysis, multivariate statistical analysis, and univariate statistical analysis. On a scale of 1 to 5 (where 1= not important and 5 = essential), USU respondents believe that *CUAHSI HIS software should work on all computer systems (Windows, Linux, Mac, Unix)* (4.5) and that *CUAHSI HIS software should work independently from any 3rd party software (e.g. Matlab, ArcGIS)* (3.8).

When asked about the priority of specific datasets for inclusion in the HIS (where 1 =low and 5 = high), USU respondents believe that the National Elevation Dataset, USGS Historical Streamflow, NCDC Precipitation, and the National Hydrography Dataset are the most important (scores >4.5). When asked about the priority of specific roles for the HIS, 'Retrieval of relevant National, Community, and Hydrologic Observatory datasets' and 'Uploading, archival and sharing of hydrologic data with collaborators and the CUAHSI community' were the highest priority for USU respondents (scores > 4.4). To comment on the quality of the data collection, all datasets presented scored at 3.7 or higher, and all HIS roles presented scored 3.1 or higher, showing a bias towards everything presented as a priority. Information on local data and standards was collected in an open format. Results from these questions were used to develop some of the questions presented at the Texas Symposium. The USU survey took an average of 15 minutes (with a range of 5-30 minutes).

UC Berkeley Preliminary Survey

This survey was conducted by the faculty and graduate students at UC Berkeley and researchers at the Lawrence Berkeley National Laboratory (LBNL). A total of 29 individuals from five departments, who conduct research in the general areas of hydrology or some related research projects, participated in the survey. The five departments include: Civil and Environmental Engineering (CE), Environmental Science, Policy, and Management (ESPM), Earth and Planetary Science (EPS), Landscape Architecture and Environmental Planning (LAEP), and LBNL (see Figure A4.1 below). The contributions from researchers in related fields helped to understand the status of the use of hydrologic information and systems at Berkeley in a more interdisciplinary context.

The specialties of those who participated in the survey were divided into seven categories raging from hydrology to ecology (see Figure A4.2 below). The diversity of participants' specialties revealed that people from related fields use the same or similar data sources and systems and that we have common challenges to cope with for better scientific information use.

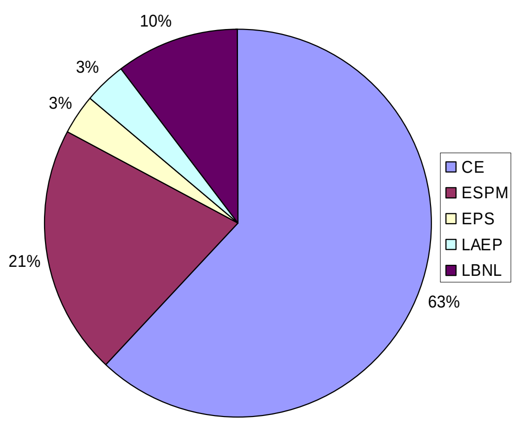


Figure A4.1. Participation by department

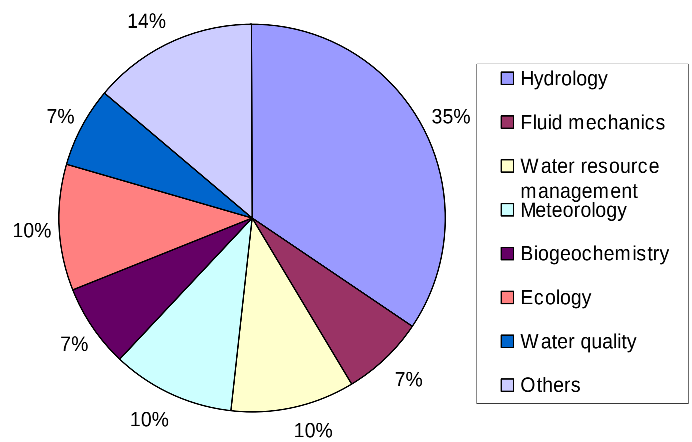


Figure A4.2. Participation by specialty

This survey was implemented using two approaches: Web-based approach and paper survey approach. 23 people used the Web-based version, and 6 people used paper questionnaire. The survey was divided into four sections: 1) systems and software; 2) data and sources; 3) needs of a data system for research, applications, and education; and 4) CUAHSI HIS. Each section included 4-7 questions. The survey was performed from January 20 to February 24, 2005.

Results

More than 70 percent of respondents at Berkeley use Windows (Figure A4.3). Those who work on modeling were found to use either Unix/Linux or a combination of two or three platforms. When asked about software that the participants prefer to use for data analysis, participants indicated that they preferred easy-to-use software with appropriate functions rather than programming languages such as C and Java (Figure A4.4). For this question, people were allowed to provide multiple answers.

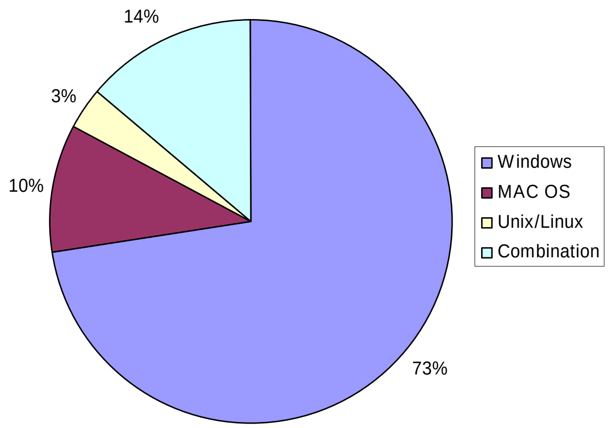


Figure A4.3. Operating systems in use at Berkeley.

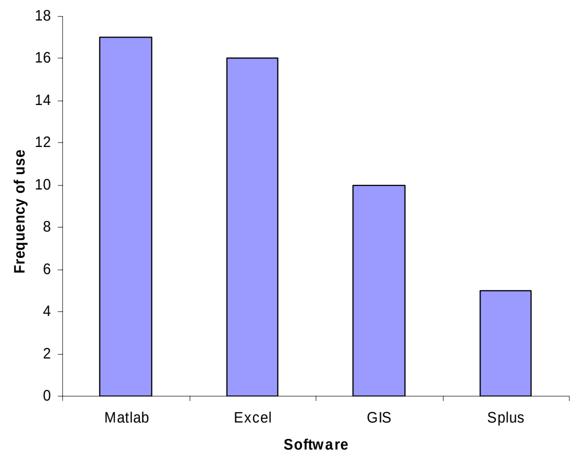


Figure A4.4. Preferred data analysis software.

In a question asking whether the participant developed any software for his/her research, 8 out of 29 answered "yes." The types of software development ranged from Matlab programs to a Web-based data analysis system. The participants described that most of their system administrators have multiple skills such as server management and hardware maintenance. Also, the participants indicated their preference of being provided by an effective and easy-to-use interface so that they can easily connect their own programs (e.g., Matlab or other software programs) to HIS data. Ten of the participants also indicated that they preferred to use GIS software and have a GIS type of components in the HIS system in a user-friend way.

In the section on data and sources, the participants indicated that they spent, on an average, about 30 % of their total research time on data processing (Figure A4.5(a)). When they were asked about questions of what data sources they often used and what kind of difficulties there were in using these data, the participants indicated the following main concerns and would like see they are to be addressed by HIS. These include:

- Necessity of assess to many different data sources with very different interfaces
- Lack of data visualization tools
- Large uncertainties associated with data
- Lack of basic functions to conduct data analysis (e.g., checking consistency, basic statistics, etc.) before downloading the data

When asked which data sources participants would like to use, participants replied that they preferred to use data that were from more established data providers such as UGSG and NCDC and expected to continue using them (Figure A4.5(b)). Thus, the participants suggested that CUAHIS HIS provide easier access to all of the existing popular data sources.

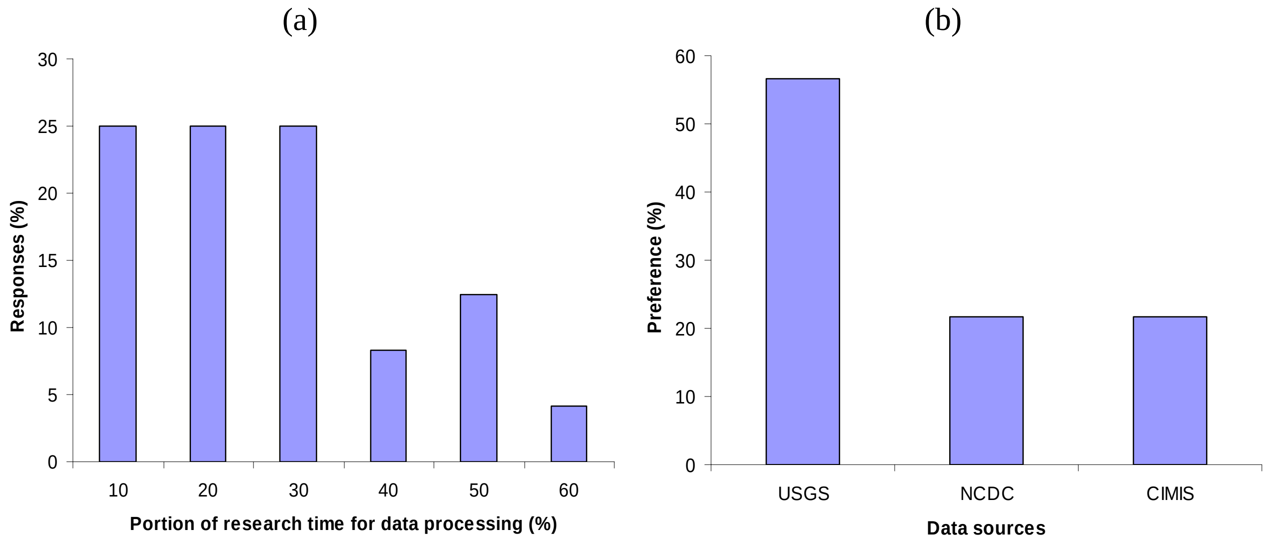


Figure A4.5. Results in percentage (%) of (a) the total research time for data processing and (b) preferred data sources.

In the section regarding needs of a data system for research, applications, and education, answers can be summarized through the identifications of the following needs:

- to address common problems that people encounter
- to provide quick and easy-to-use visualization and basic statistic functions to check the datasets before the user downloads the data
- to integrate various data sources in a single Web system
- to provide easy access to various data sources
- to provide assess to existing popular data providers
- to provide a user-friend connection to popular software programs for further in-depth data analysis

When asked about the common problems that they encounter, the participants indicated the lack of basic functionalities in most of the current data sources. Also, 100 % of the participants express their needs to have an ease of getting data, and 40% indicated that a complicated data system would be helpful, but not necessary (Figure A4.6). Most people wanted to see improvement in data visualization, in particular 3-D visualization and contour plots, followed by statistical analysis (Figure A4.7). Regarding the question of how people use hydrologic information in their research, main responses included:

- Modeling such as hydrological, atmospheric, groundwater, and water quality modeling
- Calibration and validation of numerical models
- Ecosystem modeling (e.g., climate/plant interactions, relationship of species meta-population with water management, wetland dynamics, etc.)
- Watershed and river restoration

From the survey results, it was clearly indicated that the hydrologic information be required by a broad range of research applications. Therefore, CUAHSI HIS needs to consider researchers in related fields.

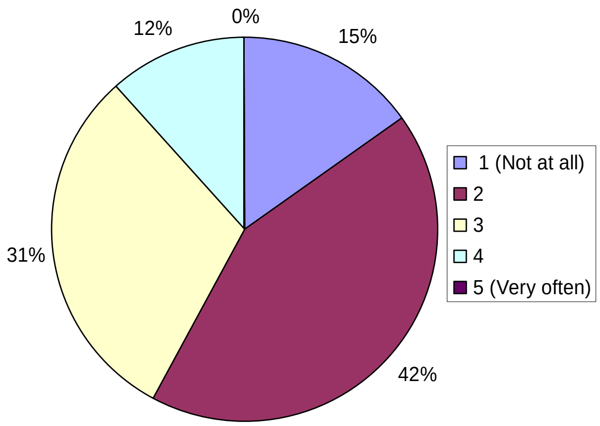


Figure A4.6. Survey participants' recognition of the need for a complicated hydrologic information system.

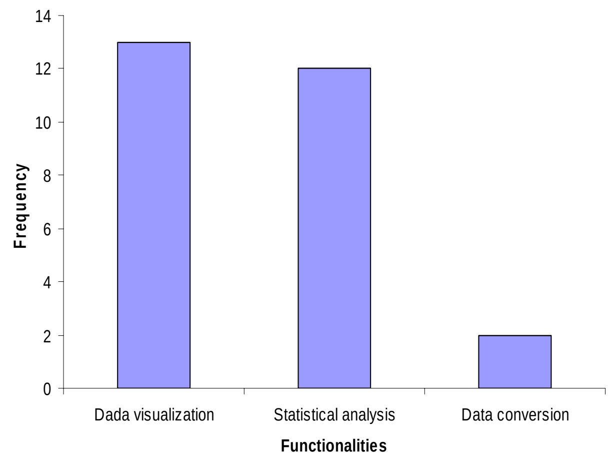


Figure A4.7. Most needed functionalities.

In the final section, we asked the participants about CUAHSI HIS to find its potential applicability at Berkeley. When asked whether they heard about CUAHSI HIS, one-third answered "yes." When asked about the expected infrastructure and services from CUAHSI HIS, main expectations from the participants were:

- Capability of data sharing (e.g., easy to ingest data into and to retrieve data from HIS)
- Standard data transferability (e.g., temporal and spatial resolution conversion)
- Support of various data formats (e.g., Ascii, Bin, HDF, etc.)
- Easy data configuration
- User-friendly cataloguing and indexing
- Service for people in other fields (e.g., ecology)
- Single interface, web-based data system
- Data visualization
- Basic statistical analysis functions
- Easy connection to other popular software programs (e.g., Matlab, Excel, GIS, Splus, etc.) for further in-depth analysis
- Open source approach
- Complicated data system is helpful but not necessary
- Prototype its integrated system, and receive feedbacks

From this list, the need for user-friendly cataloguing and indexing is notable which suggests that the user be able to know what is in the HIS when they access it.

A4.3 Pilot Survey at HIS Symposium 2005

The pilot survey was directed at those attending the HIS Symposium 2005 to focus on a subset of the CUAHSI membership most interested in the development of the HIS. The aim of the pilot survey was to refine questions based on internal surveys in order to improve the effectiveness of the web survey, which was later presented to the entire CUAHSI membership.

A paper survey questionnaire was distributed to participants in the conference information packet. Attendees were requested to fill out the survey during a break in the first day of the symposium.

Results

The pilot survey had 38 respondents from 23 different Universities and 3 different government institutions. A wide range of disciplines was represented at the Symposium (Figure A4.8). About half of the respondents represent disciplines outside hydrology or engineering, but use hydrology data for their research. The high number of computer scientists in attendance was due to interest in the development of the computer specifications of the HIS, a main thrust of the Symposium. The majority of the respondents were University faculty (57%), followed by graduate students (16%), university professionals (14%), working professionals (8%), and others (5%).

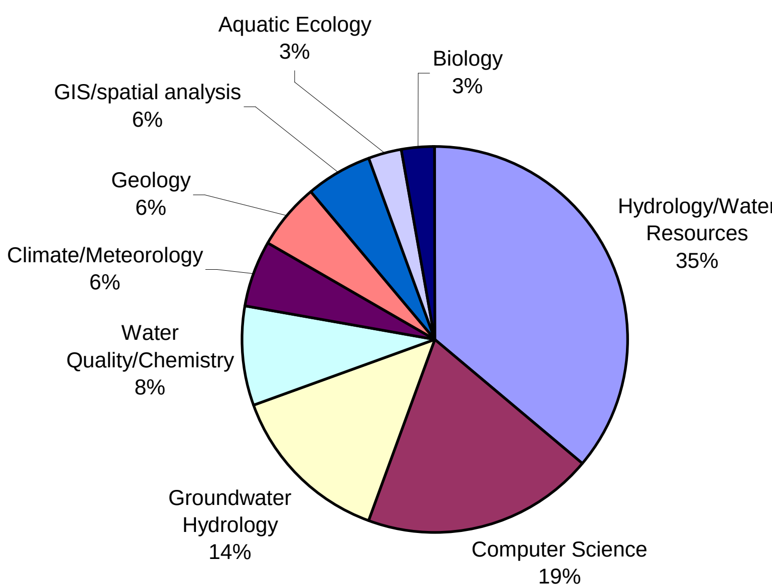


Figure A4.8. Specialties represented at the HIS Symposium 2005

Respondents were asked which operating systems they use for hydrologic research and what percentage of their time they spend using different operating systems.

Microsoft Windows was the preferred operating system: 36 of the 38 respondents use Windows an average of 75% of the time they are conducting their research. The remaining one quarter (25%) of research time is spent on using other operating systems (Table A4.1). Interestingly, only 11 of the respondents (30%) reported using Windows operating system 100% of the time.

Table A41. Operating system use by average percent of time used for research

Operating System	Number of Respondents	Average percent of research time
Windows	36	75%
Linux	15	25%
Unix	11	20%
Solaris	7	14%
MAC/OS X	5	54%
Other	2	18%

In order to prioritize software and programming languages utilized by the HIS, respondents were asked to rate the importance using a scale of 1 to 5 (where 1= Never use and/or do not find useful to your research, 2= Rarely use and/or do not rely on this, 3= Occasionally use and/or am comfortable with its use, 4= Often use and/or rely on for your research, 5= Frequently use and/or find indispensable). Microsoft Excel (4.0) and ESRI ArcGIS/Arcview (3.9) were the most used software and FORTRAN (3.3) and C/C++ (3.2) the most used programming languages. The other software and programming languages[4] we asked about averaged below the "occasionally use" range. The average results actually reflect the fact that most people "Never use" most of the software listed, while a few others find that same software "Indispensable".

Considering the datasets that the HIS could incorporate, respondents were asked the priority for including in the HIS on a scale from 1 to 5 where 5 is a high priority. Almost all of the datasets listed averaged above 4.0 as a priority for inclusion. Those that scored 4.5 or higher include: National Land Cover Data (4.7), Groundwater level (4.7), NCDC Precipitation (4.6), USGS Historical Streamflow (4.6), Water quality/Chemistry (4.6), National Hydrography Dataset (4.6), NEXRAD Radar precipitation (4.6), EPA STORET Water Quality Data (4.5), USGS Real Time Streamflow (4.5), SNOTEL data (4.5), and USGS National Geology data (4.5)[5].

For all the datasets listed, a score was given for ease of use. To help define the niche where the HIS can help researchers the most, we were interested to understand the intersection between the data priorities and the difficulty associated with using data. EPA

[4] Others listed in question: Java, MS Access, Visual Basic, Matlab, SQL/Server, Modflow, Adobe Illustrator, HEC models, GMS, WMS, SMS, R, SWAT, Sigma Plot, Surfer, SPSS, SAS, HSPF, S-Plus, GRASS, Visual Modflow, Mathematica, PostgreSQL, Tecplot, Groundwater Vistas, Kaleidagraph,

[5] Others listed included: Score 4.4 [National Elevation Dataset, Water use,Evapotranspiration, PRISM Precipitation Data, USGS Hydrologic Landscape Regions], Score 4.3 [SSURGO soils data, STATSGO soils data], Score 4.2 [LANDSAT Satellite Imagery], Score 4.0 [NCEP North American Regional Reanalysis (NARR) climate data, Real-time weather and Nexrad data from Unidata], Score 3.8 [University of Washington Gridded Meteorological Data]

STORET Water Quality Data (2.6) and SNOTEL data (2.6) were the most difficult to use of the high priority datasets. Those datasets scoring low on 'ease of use' (score of 2.5 or lower) included: Evapotranspiration (2.3), NEXRAD Radar precipitation (2.4), NCEP North American Regional Reanalysis (NARR) climate data (2.5), Water use (2.5), and Real-time weather and Nexrad data from Unidata (2.5), none of which were listed as the highest priorities for including in the HIS.

When asked to rate the priority of a list of HIS roles and system functionalities between 1 and 5 where 5 is a high priority, respondents did not select any of the options as a low priority. The highest priority roles for the HIS included (score above 3.5 average):

1. (4.7) Retrieval of relevant National, Community, and Hydrologic Observatory datasets
2. (4.6) Uploading, archival and sharing of hydrologic data with collaborators and the CUAHSI community
3. (3.8) Interfacing of hydrologic datasets in standard format with third party analysis software
4. (3.8) Development of community data models and standards for data representation

The highest priority HIS functions included (score above 4.0 average):

1. (4.5) Store and retrieve digital products from a hydrologic digital library
2. (4.4) Include GIS data on terrain, soils, land cover, geology, stream networks
3. (4.2) Allow connection to hydrologic models
4. (4.1) Include information from weather and climate models, and Nexrad
5. (4.1) Design metadata and develop tools for preparing it
6. (4.1) Support intelligent searching for hydrologic data, models, reports and papers
7. (4.1) Include remote sensing information
8. (4.0) Automatically harvest hydrologic observation data from agency websites

Interesting comments we received about the HIS project at the Texas Symposium included concern about data uncertainty, working with datasets at different scales, and including anthropogenic influences on the landscape. Others highlighted the need to create a system that is easy to use, has intuitive interfaces and is responsive to users. Comments also reflected that it is currently unclear to the community whether the HIS will be a data storage system, possibly replicating work of other agencies, or a data dissemination system that includes capabilities for data visualization, manipulation and analysis.

The pilot survey conducted in Austin was critical to identify problems with the question design and to focus the questions planned for inclusion in the web survey. One problem was the tendency for selection of 'everything is important' when respondents were asked which issues HIS should address and which datasets should be included. A change of format in the web version to a ranking question forced respondents to select the most important issues and datasets. The distribution of ranked responses gives much more information for HIS planning than a simple result that everything is important. Using an importance scale for software use (1=Never use or do not find useful; 2=Have used but do not rely on this; 3=Use occasionally and am comfortable with its use; 4=Use often; 5=Use frequently and find indispensable) was difficult to interpret. Averaged results did not represent the popular use of the software as well as a ranked choice of software most used by researchers. We also learned that not everyone was familiar with all of the long lists of datasets and software tools presented. Our attempt to simplify the questionnaire by presenting too wide a range of options resulted in off-putting some respondents. The information gathering and Symposium pilot surveys took an average of 15 minutes (with a range of 5-30 minutes), which was deemed too long for the web survey.