Unidata Community Equipment Awards

Enhanced Accessibility of Climate Data for Research and Teaching through a THREDDS Data Server

March 7, 2018

Principal Investigator:	Natalie Umphlett
Title:	Interim Director and Regional Climatologist
Institution:	High Plains Regional Climate Center School of Natural Resources, University of Nebraska-Lincoln
Telephone Number:	402-472-6764
FAX Number:	402-472-2946
Street Address:	711 Hardin Hall 3310 Holdrege Street Lincoln, NE 68583-0997
Email Address:	numphlett2@unl.edu

[redacted]

Co-PI:	Warren Pettee
Title:	Applications Programmer
Institution:	High Plains Regional Climate Center School of Natural Resources, University of Nebraska-Lincoln

Co-PI:	William Sorensen
Title:	Senior Applications/Systems Programmer
Institution:	High Plains Regional Climate Center School of Natural Resources, University of Nebraska-Lincoln

University Official:	Jeanne Wicks
Title:	Director, Office of Sponsored Programs
Street Address	2200 Vine Street, 151 Prem S. Paul Research Center at Whittier School Lincoln, NE 68583-0861
Telephone Number:	402-472-3171
FAX Number:	402-472-9323
Email Address:	jwicks2@unl.edu

[redacted]

Enhanced Accessibility of Climate Data for Research and Teaching through a THREDDS Data Server

Project Summary

The High Plains Regional Climate Center (HPRCC) is seeking funding for a THREDDS Data Server to enhance data delivery for research and teaching efforts across the region and throughout the nation. The HPRCC is a climate service provider that delivers value-added climate data and products to a diverse community of users. Although the Center's products are provided in a number of ways, the lack of access for those requiring netCDF file formats is potentially leading to an underutilization of these products, particularly among researchers and students. Implementation of a new THREDDS Data Server would allow the HPRCC to deliver netCDF versions of its products, such as the Standardized Precipitation Index (SPI), which could expand the community to a more diverse set of classrooms and researchers.

Project Description

The HPRCC is a NOAA/NCEI supported center, which is located at the University of Nebraska-Lincoln (UNL). As part of the Regional Climate Center (RCC) program, the mission of the HPRCC is simple – to increase the use and availability of climate data and information in the High Plains Region, which includes the states of Colorado, Kansas, Nebraska, North Dakota, South Dakota, and Wyoming. Center staff achieve this mission by providing climate services, developing climate data and information products, conducting applied climate research projects, and engaging stakeholders. The user base for the Center is large and varied, including researchers, educators, engineers, utilities, legal, media, and the private sector, in addition to local, state, federal, and tribal governments, among others. Although the Center's efforts are primarily focused on the High Plains states, in the past 5 years alone, the HPRCC has served the applied climate data needs of every state in the nation and over 20 countries and U.S. territories.

Approximately 28 percent of all the HPRCC's direct user contacts originate from research, teaching, and extension personnel at universities. These inquiries range from simple data retrievals to more complex data analyses. In a recent analysis of peer-reviewed publications citing HPRCC as a source of data, hundreds of manuscripts have been published from a broad range of disciplines, including agriculture, ecology, and engineering. Interestingly, only 10 percent of the publications were from the fields of meteorology and climatology (Figure 1).

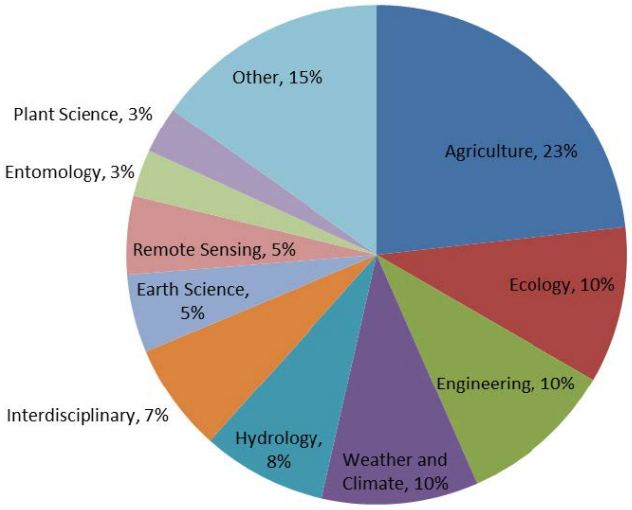


Figure 1: Research areas served by the HPRCC.

The RCCs have been at the forefront of operational support system development for over 20 years, and the Applied Climate Information System (ACIS) is the backbone of RCC climate product development and delivery (http://www.rcc-acis.org/). It was developed specifically to meet the needs of end users and is continuously being updated and improved. The HPRCC has been a part of this system development and has developed a number of climate products and services derived from the system. ACIS incorporates historical data from a number of datasets as well as near real-time station data from Unidata's Internet Data Distribution (IDD) service, which provides the most up-to-date climate information available. Data flowing through the IDD are acquired, processed, and stored by the HPRCC not only for inclusion into ACIS, but also for inclusion into NCEI's Global Historical Climatology Network-Daily (GHCN-D) dataset.

Mapped versions of HPRCC products that utilize ACIS, like the ones shown below in Figure 2, are accessed over 8,500 times per day by users from around the world. The maps are offered for several variables over timescales spanning from one week to two years and there is an archive of nearly 250,000 map images from the past 16 years. These products have become well established in the water resource management community for monitoring climate conditions and assessing drought impacts, and are also regularly used for agricultural applications. Despite the popularity of these mapped products, the gridded datasets associated with their creation have yet to be offered. Short of making special requests, researchers have had to rely on using the raw data to create their own products or using the prepared, static images.

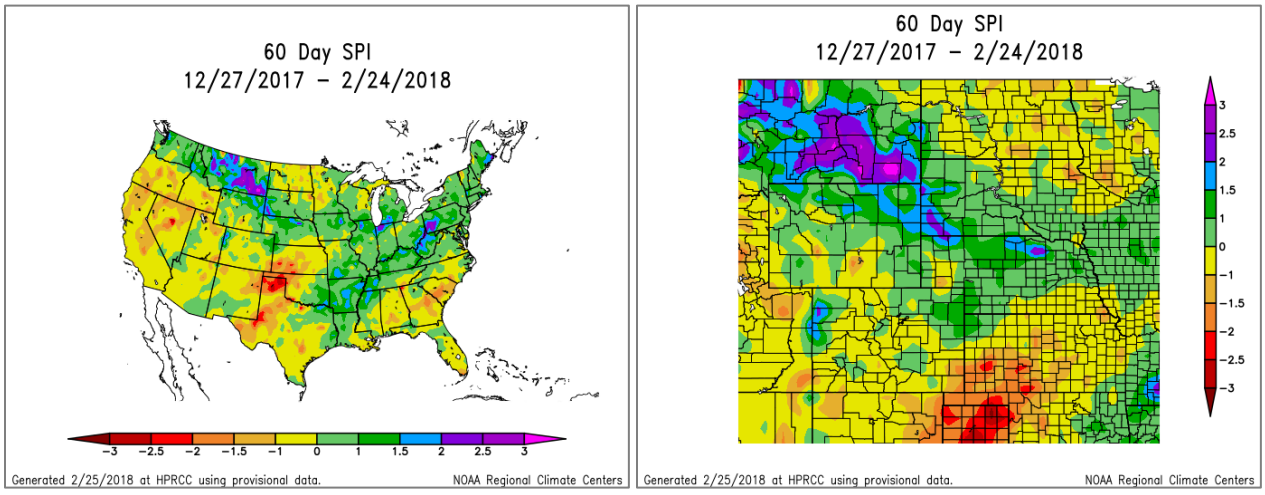


Figure 2: Researchers normally rely on raw data or pre-prepared images, such as these Standardized Precipitation Index (SPI) maps.

One way to fill this gap in services is to provide netCDF versions of these products. At this time, the HPRCC product generation suite has the capability of generating netCDF datasets, but inadequate infrastructure for deployment. As such, the HPRCC is seeking funding for a THREDDS Data Server in order to introduce new gridded products to the community. This server would be the preferred method of delivery because 1) the visualization tools could improve accessibility to data in the classroom, particularly due to the accompanying visualization tools (see Figure 3), and 2) researchers would gain direct access to datasets using the many tools available via Unidata for querying a THREDDS Data Server. This resource could

prove critical for programs centered around drought impacts and assessments, since HPRCC data are already used extensively by several organizations for this purpose, including the National Drought Mitigation Center and the U.S. Department of Agriculture. Additionally, for those in the HPRCC user-community who are unfamiliar with netCDF, the THREDDS server could encourage utilization through ease of use.

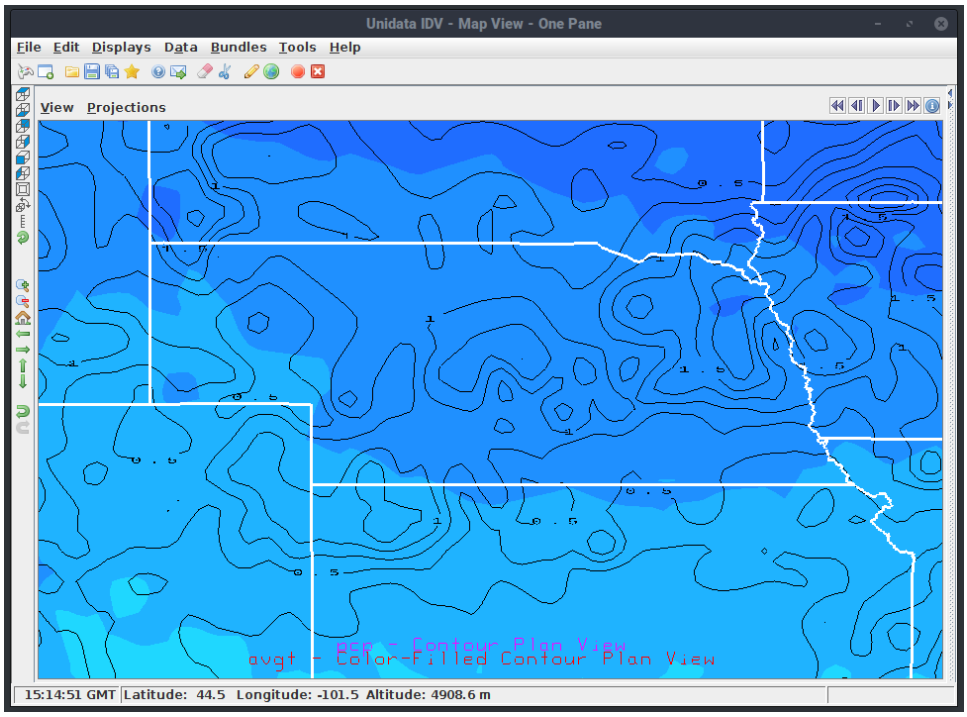


Figure 3: Academic programs can integrate gridded products into the curriculum using tools like Unidata's Integrated Data Viewer (IDV), seen here with a prototype of HPRCC data.

Budget

Funding in the amount of $6,219 is requested, which includes equipment only (Dell PowerEdge R630 server). At UNL, indirect costs are not charged on equipment grants.

Following Unidata's guidelines, the server will have a ZFS filesystem and be capable of supporting the I/O intensive operations required by THREDDS. The server offers 3.84TB of disk space across 8 hot pluggable drive bays, 32GB of RAM, and a 2.1GHz, 20M cache processor. After a RAID-Z2 filesystem is configured, the disk space will be adequate for maintaining growing daily datasets and an archive. HPRCC gridded products are two dimensional and offered daily for 14 different timescales. Products will likely be combined into a single file per timescale. Compared to model output, the disk space requirements are modest with a daily dataset (multiple timescales) requiring approximately 32MB. With additional products in development, these requirements will increase over time, but will remain far below thresholds set by many sub-daily datasets in the atmospheric sciences. A complete system description is provided at the end of the proposal.

The server will be housed on existing rack space already allocated to the HPRCC at the Scott Engineering Data Center at UNL (https://its.unl.edu/services/data-center-hosting). Server installation will be performed by HPRCC IT staff.

To ensure long-term support, this project has been built around existing resources and personnel. As one of six RCCs supported by NOAA, the HPRCC has professional IT staff who can maintain the server well into the future.

Equipment	Dell PowerEdge R630	$6,219
Indirect Costs		-
Total Budget		$6,219

Project Milestones

Phase I (1 Month): July 2018
Equipment will be purchased and installed immediately after grant funds are received. Delivery time for new equipment is typically one to two weeks followed by additional downtime while waiting for networking connections and ports. This initial process may take up to one month to complete.

Phase II (1 Month): August 2018
THREDDS will be installed on the new hardware and HPRCC daily gridded datasets will be directed to the new server. Internal testing will take place to ensure the server is operating properly. HPRCC will also identify ways to gather usage statistics through the preliminary period. This phase will take one month.

Phase III (6 Months): September 2018 – February 2019
The THREDDS server will be opened through the firewall for public access. The HPRCC will publicize the new service on social media and on their website. To further engage multidisciplinary researchers and students, a presentation will be offered at UNL's School of Natural Resources that provides information about the new service and the many ways the data can be accessed.

Phase IV: March 2019
A written article summarizing the use of grant funds will be prepared following Phase III. Usage data and any testimonies received will be included to help assess impacts made in the community.

Hardware Quote



You have saved an eQuote 1020312712075

An eQuote is now saved in your Dell Online Store.
This will be held for 60 days and will expire on 04/20/2018

Your eQuote has been sent to:

Emailed to: ttran13@unl.edu

To retrieve this eQuote

Login to Premier
Sign in to University of Nebraska
Click on "Quotes" in the top menu bar and search for eQuote number 1020312712075

eQuote Name	HPRCC Server 2018
Saved By	ttran13@unl.edu
eQuote Description	
Authorized Buyer	
Notes/Comments	
Account Name	University of Nebraska
Contract Code	93ACC

eQuote Summary

Description	Quantity	Unit Price	Subtotal
PowerEdge R630	**1**	**$6,218.05**	**$6,218.05**

eQuote Subtotal	$6,218.05
Shipping*	$0.00
Shipping Discount*	$0.00
Tax*	$0.00
Environmental Disposal Fee*	$0.00
eQuote Total*	**$6,218.05**

*The eQuote total, including applicable taxes and additional fees, may be viewable online.

Hardware Details

eQuote Details

Description	Quantity	Price
pe_r630_1337 | PowerEdge R630	**1**	**$6,218.05**

Module	Description	Product Code	Sku	ID
PowerEdge R630	PowerEdge R630 Server	R630X	[210-ACXS] [329-BCZI]	1
Trusted Platform Module (TPM)	No Trusted Platform Module	NTPM	[461-AADZ]	1574
Chassis Configuration	Chassis with up to 8, 2.5" Hard Drives, Software RAID, up to 2 PCIe Slots (With Optional Riser)	8S2O	[321-BBKI]	1530
Shipping	PowerEdge R630 Shipping- 8 Drive Chassis	D8SH	[340-AKPS]	1500
Processor	Intel® Xeon® E5-2620 v4 2.1GHz,20M Cache,8.0GT/s QPI,Turbo,HT,8C/16T (85W) Max Mem 2133MHz	26200	[338-BJCZ]	1550
Additional Processor	No Additional Processor	1P	[374-BBBX]	1551
Processor Thermal Configuration	1 CPU up to 120W	HS1201	[370-ABWN] [412-AAEE]	1697
Memory DIMM Type and Speed	2400MT/s RDIMMs	2400MT	[370-ACPH]	1561
Memory Configuration Type	Performance Optimized	PEOPT	[370-AAIP]	1562
Memory Capacity	32GB RDIMM, 2400MT/s, Dual Rank, x4 Data Width	32GBMM	[370-ACNS]	1560
RAID Configuration	No RAID with Embedded SATA (1-8 SATA HDDs or SATA SSDs)	NRES	[780-BBJG]	1540
RAID Controller	Embedded SATA	NCTRLR	[405-AACD]	1541
Hard Drives	(8) 480GB Solid State Drive SATA Read Intensive MLC 6Gbps 2.5in Hot-plug Drive, S3520	SSD25	[400-APCX]	1570
Network Daughter Card	Broadcom 5720 QP 1Gb Network Daughter Card	5720QP	[540-BBBW]	1518
Embedded Systems Management	iDRAC8 Express, integrated Dell Remote Access Controller, Express	I8EXP	[385-BBHN]	1520
Internal Optical Drive	No Internal Optical Drive for 8 HDD Chassis	NODVD	[429-AAQN]	1600
Bezel	Dell EMC 1U Standard Bezel	BZL630	[325-BCJU]	1532
Rack Rails	ReadyRails™ Sliding Rails Without Cable Management Arm	RDYRL	[770-BBBC]	1610

Power Management BIOS Settings	Performance BIOS Setting	HPBIOS	[384-BBBL]	1533
Power Supply	Dual, Hot-plug, Redundant Power Supply (1+1), 495W	495R	[450-ADWQ]	1620
Power Cords	(2) NEMA 5-15P to C13 Wall Plug, 125 Volt, 15 AMP, 10 Feet (3m), Power Cord, North America	125V10	[450-AALV]	1621
System Documentation	No Systems Documentation, No OpenManage DVD Kit	NODOCS	[631-AACK]	1590
Operating System	No Operating System	NOOS	[619-ABVR]	1650
OS Media Kits	No Media Required	NOMED	[421-5736]	1652
Shipping Information	US No Canada Ship Charge	USNONE	[332-1286]	111
Service	5 Year ProSupport and NBD On-site Service	U5IP	[976-7650] [976-7659] [976-7728] [989-3439]	29
Deployment Services	No Installation	NOINSTL	[900-9997]	714
Remote Consulting Services	Declined Remote Consulting Service	NORCS	[973-2426]	735

eQuote Subtotal	$6,218.05
Shipping*	$0.00
Shipping Discount*	$0.00
Tax*	$0.00
Environmental Disposal Fee*	$0.00

eQuote Total* **$6,218.05**

*The eQuote total, including applicable taxes and additional fees, may be viewable online.