SI2-SSI: Lidar Radar Open Software Environment (LROSE)

University of Hawaii / NCAR Proposal
in Response to
NSF Program Solicitation 14-520
Software Infrastructure for Sustained Innovation
June 27 2014

PI: Michael Bell, U Hawaii
Co-PI: Michael Dixon, EOL, NCAR
Co-PI: Gary Barnes, U Hawaii
Co-PI: Wen-Chau Lee, EOL, NCAR

<u>Overview</u>: Radars (including wind profilers) and lidars are widely used for the remote sensing of the atmosphere in 3 dimensions. These are complex instruments that produce copious quantities of data. Both the size of the data, and the complexities associated with real-time applications and post-analyses, pose a challenge for researchers, students, and instrument developers. New software tools are needed to facilitate research and education and to maximize returns on NSF investments in observing facilities for weather and climate research.

The goal of the proposed work is to apply collaborative open source methods to help to address the 'big data' problem faced by users in the radar and lidar research community. The plan involves a series of steps: (a) identifying suitable open data exchange formats; (b) developing a suitable infrastructure framework for handling the data, translating formats, running scientific algorithms and visualizing the results; (c) developing 'standard' algorithm modules for those typical processing steps that are well understood and documented in the available literature; and (d) involving the user community in the development of new research modules that address the specific needs of the latest scientific research. This project will build on existing prototypes and available software elements, while facilitating community development of new techniques and algorithms.

Developing software as open source in collaborative projects has proved to be both innovative and efficient – as the development of the LINUX operating system and Python language both clearly demonstrate. The tools and mechanisms to do this effectively are widely available, and often free. The scientific community is starting to make good use of this approach, as the rapid growth of scientific Python shows. This proposal centers on using this approach to develop a framework and toolbox for scientists using radars and lidars. Collaboration would occur with other US universities and groups performing similar work, such as the DOE Argonne National Labs and European institutions.

<u>Intellectual Merit</u>: The outcome of this project would be (a) a completed infrastructure framework and (b) a suite of documented software modules for performing actions within that framework. These modules would each implement accredited scientific methods, ideally referencing published papers. The infrastructure and modules would allow researchers run 'standard' procedures (without having to develop these from scratch, thereby improving the reproducibility of the analyses) and also to develop new scientific approaches. The resulting software framework will be a significant improvement over the current radar and lidar software environment, in which researchers and students are required to spend a large fraction of their time handling data and developing code for routine tasks that are common to many research projects. Researchers will benefit from the extra time they have to devote to advancing understanding of weather and climate rather than handling mundane data tasks, leading to a positive outcome in terms of advancement of scientific knowledge. The use of collaborative open source methods will lead to a suite of available algorithmic modules that will allow innovative scientists to explore radar and lidar data in ways that they currently cannot.

<u>Broader Impacts</u>: The difficulty that users face in handling large complex data sets restricts the pace of progress of scientific research. Improving the software will speed up research in a synergistic way, as innovative researchers and students will leverage the improvements in ways that are not necessarily anticipated by the developers. Modern radars and lidars are a diverse class of instruments, capable of detecting aerosols, birds, bats and insects, 3-D winds, moisture, refractive index gradients, cloud particles and precipitation particles. Scientists and engineers use them to perform research into air quality and pollution, dangerous biological plumes, cloud physics, cloud extent, climate models, numerical weather prediction, road weather, aviation safety, severe convective storms, tornadoes, hurricanes, floods, and movement patterns of birds, bats and insects. Many of these are important topics for society and the biosphere. This project will open new avenues of scientific investigation, including data assimilation, by providing good software tools to researchers, students, and educators. Improving the effectiveness of this research will undoubtedly enhance our understanding of these diverse topics.

1. Introduction

Atmospheric radars and lidars of all types, including scanning weather radars, cloud radars, airborne radars, wind profilers, and Doppler and aerosol lidars, are at the forefront of research for the three-dimensional remote sensing of the atmosphere. The data from these instruments are critical to our ability to understand and predict atmospheric processes that have a significant societal impact, such as climate change, severe weather, hydrology and agriculture, air quality, aviation, and ecosystem impacts. This technology continues to evolve, leading to more complex instruments with high data rates, such as multi-spectral lidars and dual-polarization and multi-frequency radars. These technological advances improve the usefulness of the instruments, but also make efficient use of the resulting large data sets more difficult. Handling the large quantities of data produced, and conducting science using reproducible techniques, are challenging tasks and require sophisticated software tools. This proposal presents a plan for developing an open software infrastructure and the accompanying tools to allow scientists and students to meet this challenge, to enable them to handle the data more effectively, to perform analyses in a well-documented and reproducible fashion, to integrate disparate data sets, and to visualize the results more readily. The proposed effort will have a synergistic effect to empower atmospheric researchers, make radar and lidar data more accessible to the broader scientific community, improve productivity, and ultimately reduce research costs in the geosciences.

The project has three phases. First, best practices in scientific software engineering will be applied to the development of a data system infrastructure that will properly support the data formats, algorithm modules, and displays used in radar and lidar analysis. Second, a combined engineering and scientific approach will be employed to develop modules that perform 'standard' well-understood procedures, so that these modules are available for the more routine steps in the research process. And finally, new modules will be developed to support the latest research and development. The intent is that the modules created in the final step will, over time, be included in the toolbox of standard techniques available to researchers, students, and educators. The ultimate goal is that this software toolbox will encourage users to contribute, enhance and share tools with others in an open collaboration.

Radars and lidars transmit pulses of electromagnetic energy, and subsequently measure the returned signal received by the instrument. Radars operate at radio frequencies while lidars use lasers at frequencies in or close to the visible spectrum. These instruments emit a pulse of finite length from a highly directional antenna (for a radar) or telescope (for a lidar) that illuminates a volume of atmosphere – the 'measurement volume' or 'range gate'. Some fraction of that energy is scattered by the molecules or particles in the measurement volume, or by gradients in refractive index, and some of the backscattered energy is returned to the receiver. The properties of the measurement volume and its contents may be inferred by analyzing the returned signal, and the location of the illuminated volume may be estimated by knowing the pointing direction of the antenna and the time taken for the pulse to travel to the target and back to the receiver.

These remote sensing instruments may be fixed on the ground, or on mobile platforms such as vehicles, ships, aircraft and spacecraft. Many of them have actively scanning antennas/telescopes to sample the space around the instrument, while others point in a fixed direction, allowing the atmosphere to move past them as they sense the changes over time. One thing they all have in common is that they produce data at known distances from the instrument in a direction governed by the pointing angle of the antenna or telescope. This common feature allows for shared software infrastructure, including data formats, processing techniques and visualization tools. The data collected by these instruments can be processed to provide information on the following topics:

- Radars: severe weather research (storm formation, tornadoes, hurricanes, severe thunderstorms, hail, heavy rain), road weather, quantitative precipitation estimation (QPE), quantitative precipitation forecasting (QPF), flood forecasting, precipitation climatology, 3D winds in precipitation, temperature and wind profiles, aviation safety (wind shear, icing, turbulence), climate studies, moisture distribution (refractivity), biological behavior of insects, bats, and birds.

- Lidars: aerosols, air quality, air pollution, plume tracking, biological agent releases, cloud physics, moisture profiling, cloud detection, climate model parameterization, moisture profiling, 3D winds in clear air, wind energy research and operations, aviation safety (wake vortices).
- Wind profilers: wind profiles, turbulence and wind shear, gravity wave and rotor characteristics, boundary layer evolution.

Improved community software tools are needed to fully realize the potential of radars and lidars in these research domains and enable assimilating quality-controlled radar and lidar data into numerical models in near real time; thereby to improve our understanding and forecast of weather and climate.

1.1   Radar and Lidar Data Processing

This proposal seeks to improve the software used to process and analyze the large quantities of data produced by radars and lidars. Modern instruments can produce vast amounts of data. For example, the NCAR airborne HIAPER Cloud Radar (HCR) transmits 10,000 pulses per second, and measures the returned echo at a resolution of 10 m to a maximum range of 15 km (i.e. 1500 samples per pulse), yielding a data rate of 120 MBytes per second, or 5.2 TBytes on a 12-hour flight. The Doppler-On-Wheels (DOW) radars transmit at 5000 pulses per second, but at two separate frequencies, therefore their data rate is similar to that of the HCR. The 160 NEXRAD radars operated by the US National Weather Service produce raw data at an aggregated total of around 1.5 TBytes per day. These data sets are massive and complex, ranging from raw radar time series to processed products, with geometric intricacies resulting from the nature of propagation through the atmosphere and of moving platforms. This is 'Big Data' in the modern usage of the term. Handling such data can be overwhelming for a researcher or student, and improved tools are required to facilitate efficient research.

The proliferation of 'legacy' data formats is also a significant challenge. Over the years many different formats have been developed by instrument developers, and frequently the user spends a large fraction of the available research time moving data around and trying (and often failing) to convert it into a form that is useful. In preparation for the proposed work, a new common data exchange format for lidar and radar was developed, along with the required format translation software for several commonly use data formats.

An important aspect of this work is to help facilitate research into the assimilation of radar and lidar data into numerical models. Accurate initial conditions are critically important for weather and climate research, and remote sensing instruments are among the few that can provide high-resolution, three-dimensional wind, aerosol, and precipitation fields. Current data assimilation techniques are limited in their ability to ingest large amounts of radar or lidar data, and significant data pre-processing such as quality control and data thinning is required. The optimal way to assimilate radar and lidar data into models is an ongoing area of research that is hindered by the lack of mature processing tools.

Figure 1 (below) shows a high-level view of the data flow for an end-to-end system incorporating radar and lidar instrumentation. Doppler radar data at its most raw is in the form of what is referred to as 'time series', with 2 values (so-called 'I' and 'Q') for each gate and each pulse. This is very voluminous data and in the past was usually discarded due to limited storage capacity. Some modern systems provide the option of saving this data for detailed analysis and display. From the I/Q time series, the radar 'moments' are computed, for example reflectivity (zero-th moment), velocity (first moment), and spectrum width (second moment). Lidars generally do not export time series data, instead treating moments as their 'raw' data.

Radars and lidars produce raw data that are inherently 'noisy' and contains artifacts and errors that must be handled before it is useful for either scientific or operational purposes. One can think of the processing steps as distilling the real *information* from the raw *data*. Generally these steps are applied to data still in the native coordinates. Spurious non-weather echoes from the ground, birds, wind-farms and the like must be identified and dealt with. Clutter mitigation, interference and artifact removal may be carried out. Velocity and range ambiguity problems (referred to as 'aliasing') are inherent to radars and lidars, which sometimes lead to uncertainty about where the echo is located or at what velocity the scatterers are

moving. Therefore, 'de-aliasing' algorithms may be applied as appropriate. Also at this stage, data quality metrics may be computed – these are necessary to remove bad data for analysis, or to supply estimates of uncertainty that are required by numerical model data assimilation systems. Once these steps have been applied, the data can be considered 'quality controlled'. Algorithms that operate in polar coordinates may be applied at this stage, and the data may be forwarded to models for data assimilation.
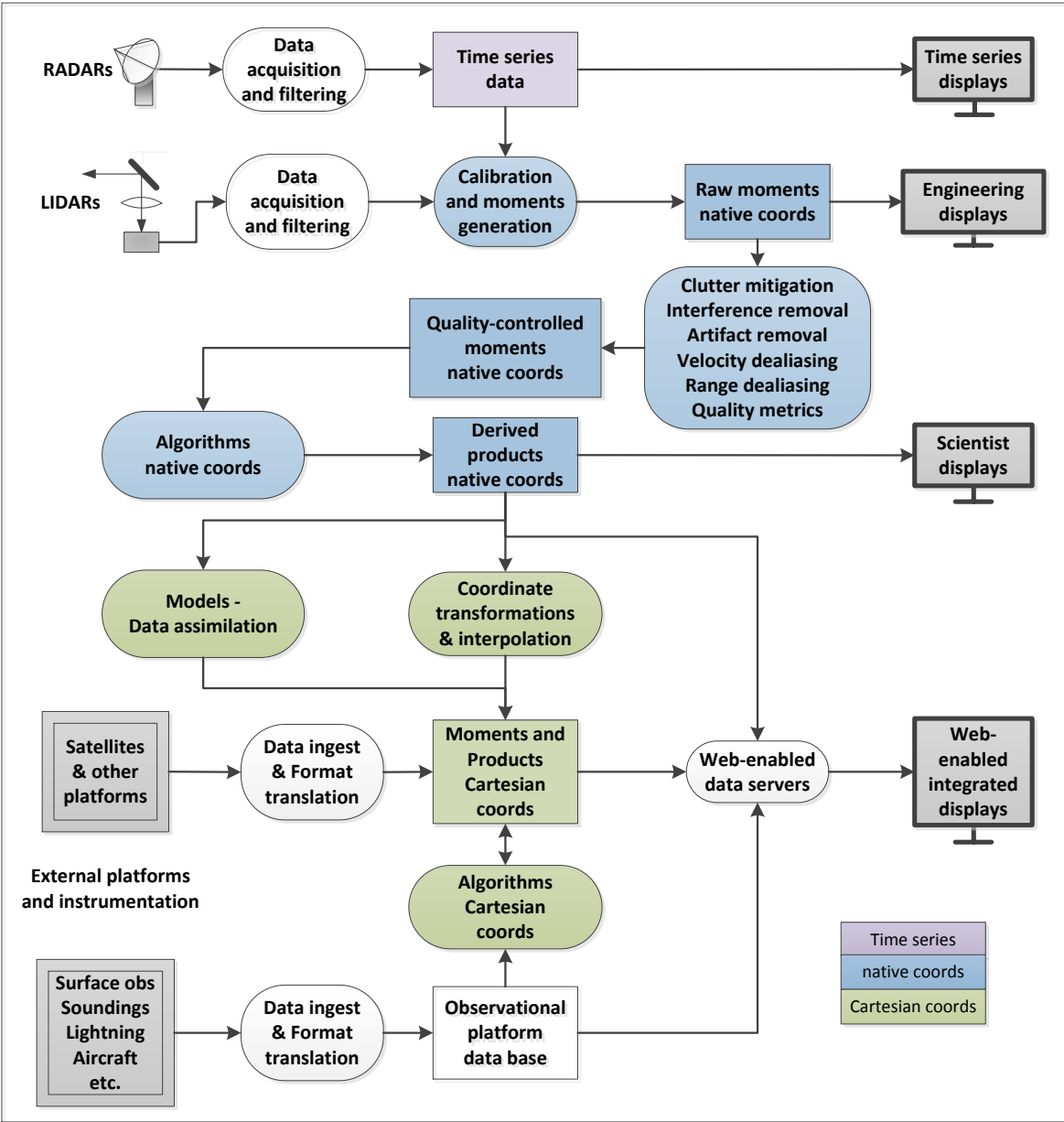

Figure 1: High-level data flow for modern end-to-end analysis of radar and lidar data.

For many applications the data must be transformed from polar to Cartesian coordinates. As a result of the complicated geometry of the propagation paths of electromagnetic pulses through the atmosphere, coupled with the moving coordinates for mobile instruments, sophisticated processing is required to correctly geo-reference the observations. There are a number of different approaches for interpolation and coordinate transformation, and a suitable method must be chosen. Once the data is in Cartesian coordinates, the results of numerical models may be imported, and algorithms appropriate to Cartesian data may be run.

As these systems become more complex and scientists wish to use the data in a wider context, for example in climate research, external environmental data sets (e.g. surface observations, soundings, satellite data) must be imported and integrated with the products from the radar/lidar analysis. Visualization of the results of the analyses at various stages of the processing chain requires that suitable displays be available for each stage and for different data types. While most people are familiar with the classic "radar scope", the variety and complexity of modern radar products requires more sophisticated software displays.

The goal of this project is to significantly improve the radar and lidar data processing capability for the community. Determining what functionality is most desired by the user community is a significant task due to many diverse research interests. A white paper developed in 2012 was presented to NSF and circulated widely to the radar and lidar user community at universities and government laboratories. This white paper was distributed prior to a workshop hosted by NSF from 27-29 November 2012 in Boulder, CO, titled *Community Workshop on Radar Technologies*. During one of the workshop sessions the attendees (users) were asked to consider and prioritize their urgent software needs, both for their personal research and what they thought would be best for the broader community. Table 1 summarizes the outcome of the survey, ranked according to highest need.

Radar Software Needs	Personal Research Needs	Community Needs	Total Score
NCAR-maintained centralized repository for radar software (esp. including wind synthesis) with data sets for software testing	55	41	96
Standardized software packages and toolkits (multi-platform, modular, menu-driven, easy for community to add to, ease of conversion among new and old radar data formats)	30	53	83
Training (workshops/online tutorials)	48	15	63
Ability to integrate radar and non-radar data sets	25	32	57
Open source tools and software	30	16	46
3D/4D Visualization Software (with publication quality output)	24	21	45
Next generation wind synthesis software to replace the legacy (REORDER/CEDRIC) algorithms, while maintaining current functionality	15	27	42
Common radar data format standard and a common metadata standard (e.g. CfRadial)	19	15	34
64-bit compatible real-time display software tool	19	11	30
Improved radar data quality control (solo) (Oye et al., 1995)	12	20	32
Automated quality control software	14	13	27
Detailed documentation for data products, tools, and code	18	7	25
Improved dual-polarization processing	10	12	22
Accessible variational Doppler radar assimilation and thermodynamic retrieval	7	4	11
Totals	326	287	613

Table 1. Radar software needs based on combined responses from workshop participants and an online survey of members of the radar community who were unable to attend the workshop. Similar topics among workshop group responses were consolidated. Scores represent the number of votes for each topic.

It is clear from the NSF supported workshop and feedback from the white paper distribution that there is a strong need for improved software infrastructure in the radar and lidar community (Bluestein et al. 2014). The work in this proposal would meet that need by improving existing time-tested software and developing new open-source infrastructure, data exchange formats and displays, and algorithm implementations.

2 Software Design

From a software engineering point of view, the project has two major layers: (a) the enabling infrastructure, including the data layer and displays, and (b) the scientific software modules that extract the atmospheric information contained in the system.

We propose a relatively simple and well-understood infrastructure, based to a significant extent on existing code. This is an important layer, but it is not the intellectual heart of the system.

The intellectual merit will lie in the modules that contain the scientific algorithms. There are two categories of these modules: (a) standard processing steps that are already well understood, and often based on existing prototypes; and (b) new processing modules that will be developed by the project, and in the community, as a result of ongoing research and development in radar and lidar science.

The key to the development process is the concept of an *algorithmic module* that will take upstream data as input and produce some useful result. The choice of the module size and complexity is important – too many simple modules are unmanageable at the system level, and a large complex module may be difficult to debug and verify. The intent is that all modules should reference either accepted methods in the literature, or new methods under development that will soon be added to the literature. The scientific results will be tested as part of this project, and further reviewed by peers in the community to ensure quality.

The data and logic flow in Figure 1 suggests that a modular design is a natural fit for a software system designed to facilitate the required processing steps. A modular approach is nothing new in software engineering – rather, it is a well-proven technique that allows for flexibility, composability, and manageability of relatively simple components within an otherwise complex system.

The following sections introduce details of how such a modular system would be created to meet the needs of this project, with specific emphasis on the interoperability of the components.

2.1 Data exchange formats

A good definition of a *module* in the context of the proposed design is *an application that reads data in some form, probably from a file or queue, runs an algorithm or procedure on that data, and then writes the result, probably to another file or queue*. It is not necessary that the data exchange be file-based, but this is a useful paradigm when developing and testing components in a complex system. Once the modules are fully tested and verified, some of the file writing steps can be dispensed with.

One of the most challenging aspects for scientists and engineers dealing with radar and lidar data is the large number of data formats in use. A format suitable for data exchange should ideally be portable (i.e., computer-platform independent), maintainable, self-describing, easy to handle (to read/write/understand), properly documented and standardized for syntactic interoperability. UNIDATA NetCDF (http://www.unidata.ucar.edu/software/netcdf) provides a framework for such formats. For this project, the primary storage format will be NetCDF, using the Climate and Forecasting (CF) conventions (Eaton et al. 2011). The latest NetCDF 4 implementation is built on the NASA HDF5 layer (http://www.hdfgroup.org/HDF5), allowing for efficient compression. Both NetCDF and HDF5 are well documented and supported by open-source libraries, and are in wide use by the scientific community.

The CF conventions are an important requirement, because they are designed to help facilitate data exchange with the numerical modeling and climate communities by adopting standards for metadata. For Cartesian data, the CF conventions have been in common use for over 10 years (see http://cfconventions.org/). CF version 1.6 (or later) will be used for Cartesian data in this project, and upgrades will be incorporated as appropriate when new versions become available.

For radar and lidar data in native coordinates, the new CfRadial data format was developed at NCAR in 2010, and submitted to the CF Metadata process for review and approval (Dixon et al. 2013) (see

https://cf-pcmdi.llnl.gov/trac/ticket/59). Since then CfRadial has become one of the de-facto standards for radar data in polar coordinates. It has advantages over other formats in that it is CF compliant, self-describing, extensible, and non-lossy, i.e. it preserves the information received from the instrument (see https://www.eol.ucar.edu/content/standard-data-formats). This is an actively supported format with upgrades being made as required in response to feedback from the user community. A fully-featured C++ library (Radx) is available for handling this format, and for translating data to/from other formats (see https://www.eol.ucar.edu/software/radx). Furthermore, CfRadial is easy to read in any language that supports NetCDF, including Java, Python, Matlab and IDL. A Python module has been developed for it at DOE ANL (Heistermann et al. 2014).

NetCDF conventions are also available for auxiliary data as surface observations, profiles, atmospheric soundings and trajectories. These will be used as appropriate. For some data sets it is necessary to use a binary format for reasons of efficiency. This is true of radar time-series data, which is voluminous and is essentially a streaming format that is not suitable for NetCDF. The Integrated Weather Radar Facility (IWRF) format (see https://www.eol.ucar.edu/content/standard-data-formats) was developed as a joint project between NCAR and the Colorado State University CHILL National Radar Facility, and is used by both organizations for time series data.

2.2 Inter-module communication and module configuration

Within a single application, the Application Programming Interface (API) governs the communication between sub-modules. In a large system, communication is carried out between applications – referred to as inter-process communication or interoperability. The LROSE design is based on the latter approach, where at the macro level data will be passed from one module to the next via the file system, with the results being passed on to the next module in the chain. A suitable queue-based triggering mechanism will be provided for real-time operations. This modular design has the advantage of simplicity, and allows for easy communication between application modules that are written in different languages.

The user must be able to specify how each module should run. Some of the algorithms in LROSE will be complex and require a large number of configuration parameters. These will be supplied in files read as each application starts.

2.3 Implementation programming languages

The proposed system will be designed to make maximum use of existing software elements – see section 4 below. Therefore it will comprise a mixture of programming languages. The primary language for the core infrastructure will be C++, since a large number of existing modules are in C++, as is the bulk of the data infrastructure. With modern build techniques it is possible to create modules that are portable, particularly if these avoid GUI elements. However, C++ is generally a more expensive development environment than higher-level languages. Python is free and widely supported in the scientific community, and many excellent Python libraries are available. To the extent possible, new development will occur in high level languages such as Python, Matlab ® or Octave (open source Matlab), Julia, and IDL ®. Decisions concerning the design of the core suite will be made with maximum portability in mind.

Choosing the best language for display development is tricky since technology changes rapidly, especially in terms of web-capable displays. Java has proved to be a good portable display language (for example the JAZZ display developed at NCAR). Python and Julia also have excellent display support in the MatPlotLib library. Decisions on display technology will be deferred to the extent possible so that advantage can be taken of any new advances during the duration of the project.

3 Software elements

Previous work by the PIs and others at the University of Hawaii (UHM), NCAR and UNIDATA has laid the foundation upon which a significant fraction of the required modules may be developed. Many working prototypes are already available, defined here as existing software that performs the desired

functionality but without extensive testing and/or with a limited user interface. Building on prototypes will help to avoid duplication of effort and will save costs. (It is estimated that over 60 person-years of effort have been expended on this previous work at UHM and NCAR from various funding sources, primarily the NSF, the FAA and NASA.)

There are 4 categories relevant to this software: (a) the data system infrastructure, to provide a smooth interface with the NetCDF data layer; (b) the displays for visualization; (c) core algorithm modules and (d) community developed modules. The first three categories will be developed under support from this proposal.

3.1   Infrastructure

A well-tested core infrastructure, written in C++, has been developed at NCAR over the last 20 years. Much of this can be reused. A major upgrade in LROSE will be the integrated support for NetCDF.

The core infrastructure will handle the following tasks: (a) reading input data from sensors and other sources; (b) writing output data; (c) translating data into supported formats, primarily NetCDF; (d) moving data between hosts; (e) providing data to the algorithms; (f) serving data to the displays; (g) disseminating data to users.

A layer of bindings will be added to this C++ infrastructure for higher level languages such as Python, R etc., and, if applicable, native data interfaces in these higher level languages. A software distribution mechanism will be added, using GitHub. Table 2 lists the components to be developed as part of the proposed work.

Infrastructure Component	Prototype exists?
NetCDF data support layer	Yes
Data servers for display applications	Yes
Higher-level language bindings for C++ library classes	No
Higher-level language native data handling library	No
Portability layer – ensuring the system will compile on and/or port to the platforms of the intended users	Yes
Open source software distribution mechanism	Yes

Table 2. Infrastructure components.

3.2   Display applications

Visualization of data is a crucial aspect of any analysis package. Different scan modes require different types of displays, from vertically pointing data to the familiar polar scanning mode to more complex visualizations for time series data. For maximum scientific value, radar data needs to be combined with other atmospheric datasets for integrated analysis and visualization. Prototypes exist for several of these display modules, but additional development is required to improve their robustness and usability. Table 3 lists the display tools that will be included in the core suite.

Display	Software Examples	Prototype exists?
ASCOPE for spectral radar I/Q time series	SpectraPlot	Yes
BSCAN for vertically pointing data	B-scan display for HCR	Yes
Low-level viewer and editor for polar data	solo3 (C++)	Yes
	emerald (MatLab),	Yes
	pysolo (python)	Yes
Platform-independent viewer for integrating radar and other data sets	Jazz	Yes
	IDV	Yes
Display and editor tool for profiler data	PyProf	Yes

Table 3. Displays to be included in the core suite.

3.3 Core suite algorithm modules

In order to facilitate scientific insight, the radar and lidar moments and time series spectra must be processed into atmospheric quantities such as wind velocity, hydrometeor size and shape, and precipitation rate. The conversion from the raw radar data first requires data quality control, and there is extensive literature on the removal of non-meteorological echoes and velocity de-aliasing (Steiner et al. 1995, James and Houze 2001, Bell et al. 2013), and on algorithms for processing wind profiler and lidar raw data. Once high data quality has been ensured, additional scientific algorithms can be applied to obtain relevant quantities. While many objective, citable algorithms to obtain winds (Ray et al. 1980, Bell et al. 2012), hydrometeor type (Vivekanandan et al. 1999) and rain rate (Ryzhkov et al. 2005) are documented in the literature, software to implement these algorithms is not widely available. Many of the algorithms are re-implemented over and over by graduate students due to lack of software availability, a task that is prone to error. While implementation of new algorithms by the community is expected, there are a number of core algorithms that are well established and existing implementations that are open-source. A core suite of algorithms that are expected to be most widely used will be included in the LROSE package. The availability of well-tested algorithms to perform common tasks will prevent "re-inventing the wheel". Table 4 lists the core algorithms to be included in LROSE:

Algorithm	Example	Prototype?
Clutter detection and mitigation	CMD (Hubbert et al. 2009)	Yes
Beam blockage analysis	RadxBeamBlock	Yes
Velocity de-aliasing	JamesD (James and Houze, 2001)	Yes
Quality metrics / error assessment	Bell et al. 2013	Yes
Attenuation correction in precipitation	Gu et al. 2011	Yes
Compositing data from multiple instruments	Henja and Michelson, 2012	Yes
Vertical profile of reflectivity – VPR	Kirstetter et al. 2013	No
Convective/stratiform partitioning	Steiner et al. 1995	Yes
Particle ID / Hydrometeor classification	Vivekanandan et al. 1999	Yes
Precipitation rate	Numerous	Yes
Quantitative Precipitation Estimation – QPE	Zhang et al. 2011	Yes
Storm tracking – convective	TITAN (Dixon and Wiener, 1993)	Yes
Storm tracking – stratiform	CTREC (Tuttle and Foote, 1990)	Yes
Vertically integrated liquid – VIL	Greene and Clarke, 1972	Yes
Single Doppler general wind retrieval	EVAD (Matejka and Srivastava,1991)	Yes
Single Doppler hurricane/tornado wind retrieval	GVTD (Lee et al. 1994, 1999)	Yes
Gridding / interpolation	SPRINT (Mohr et al. 1986)	Yes
Multiple-Doppler wind retrieval – geometric	CEDRIC (Miller et al. 1986)	Yes
Multiple-Doppler wind retrieval – 3D variational	SAMURAI (Bell et al. 2012)	Yes
Multiple-Doppler wind retrieval – 4D variational	VDRAS (Crook and Sun, 2002)	Yes
Thermodynamic retrieval	Roux et al. 1993	Yes
Wind shear detection	Albo and Kessinger 1996	No
Meso-cyclone detection	Stumpf et al. 1996	No
Wind profiler moments estimation	NIMA (Cornman et al. 1998)	No
Wind profiler clutter rejection	GABOR/Wavelet (Lehmann and Volker 2012)	No

Table 4. Algorithm elements to be included in the core suite

3.4 Community algorithm modules

We are proposing an open source approach that encourages community participation in algorithm and display development. Efforts to establish an open source weather radar software community have already begun (Heistermann et al. 2013). Work performed under this proposal will establish a core framework within which to develop, maintain, and upgrade commonly used functionality. The community will be able to use this framework for common processing tasks and to develop modules to meet more specialized

needs using efficient high level languages such as Matlab®, IDL®, NCAR Command Language (NCL), R, Julia, and Python. These modules will be shared with the full community in a manner similar to that adopted by the Weather Research and Forecasting model (WRF) community (http://www.wrf-model.org). A synergistic effect is expected, such that a multiplicative return on NSF research investment can be realized.

Communication with the user community will be continued through the life of the project. An annual workshop for interested users will be arranged at NCAR in Colorado, in addition to a town hall or another shorter forum at existing meetings, for example at the AGU annual meeting and AMS Conference on Radar Meteorology., These may be combined if it is deemed more effective for attendance and broader community feedback in a particular year.

4   Project plan

4.1   Software development plan

As listed above in Tables 2, 3 and 4, there are 3 main categories (or layers) in the proposed system: (a) the infrastructure for handling data ingest, converting formats and serving data to client applications, (b) displays for visualization and (c) algorithms for analysis and creation of advanced products. Previous experience has shown that it is unwise to try to develop these in a linear fashion. Rather, we propose an iterative approach, developing the layers in parallel. This allows for early prototype releases to the users, to obtain feedback on how well the applications work, how useful they are and what modifications are needed to improve them. The earlier this feedback is obtained in the development cycle, the more complete the components will be at the end of the project.

We are requesting funding for 4 years. Table 5 outlines the tasks planned for each of those years, showing the parallel development of the three software layers.

Time period	Proposed tasks		
	Infrastructure	**Displays**	**Algorithms**
Year 1	Enhance NetCDF data support layer. Enhance data server applications. Set up software distribution mechanism (GitHub). Enhance portability and compilation on target platforms.	Develop (or enhance) the most widely required displays (e.g. those that integrate radar/lidar data with other sources)	Develop/enhance tier 1 algorithms – i.e. those most in demand (refer to Tables 1 and 4)
Year 2	Enhance existing infrastructure. Begin development of bindings for higher level languages.	Enhance existing displays. Begin development on BSCAN and profiler editing display tools.	Develop/enhance tier 2 algorithms – i.e. those next most in demand
Year 3	Enhance existing infrastructure. Begin development of native APIs for higher level languages.	Enhance existing displays. Begin ASCOPE spectral display development.	Develop/enhance tier 3 algorithms – i.e. those next most in demand
Year 4	Enhance existing infrastructure.	Enhance existing displays.	Develop/enhance tier 4 algorithms – i.e. those next most in demand

Table 5. Planned software tasks

4.2   Team organization

The software development work will be shared between UHM and NCAR, the team comprising software engineers and scientists at both institutions. The work will be split roughly equally between the two

organizations. The PI will have overall management responsibility for the project, with the software engineering team led by the Co-PI (Dixon). The PI and 3 Co-PI's will form the *management team.*

Support is requested for UHM to hire one full-time software engineer for this work, and support two graduate students. NCAR will use a small team of software engineers and an associate scientist. We will refer to this group of staff as the *software engineering team.* The combined management and software engineering teams will be referred to as the *LROSE team.*

The software engineering team will collaborate using GitHub, which is a popular technology for open source development (Heistermann et al. 2014). For the most part meetings will be held as teleconferences, for cost and time efficiency. Provision is made in the budget to allow for some limited travel between the institutions for allow for face-to-face meetings between the teams. Further details on the team organization and project management are provided in the supplementary Management and Coordination Plan.

4.3   End-user feedback and performance metrics

Both UHM and NCAR have well-established groups of users in the scientific community, who depend on the software previously developed at both institutions. The development of LROSE will make considerably enhanced tools available to those user groups. As the new or upgraded display and algorithm applications become available in prototype form, they will be made available to the user community either via GitHub or via direct download from web servers.

NCAR will, on its main web site, maintain a JIRA ® (or equivalent) issue tracking web page, on which users can file bug reports, requests for enhancements and detailed feedback. JIRA has proved to be extremely effective for keeping track of outstanding issues and for managing change requests from users. Open communication with the user community will facilitate an active discussion between the users and the developers to ensure that the software requirements are fluid rather than fixed, and can be changed to meet the needs of actual users in the field.

In addition, each year we will approach the users by email to solicit feedback, and will hold workshop for the user community at NCAR in Colorado to facilitate direct feedback and a forum for active discussion. A town hall meeting or other shorter forum will also be utilized to obtain feedback at a widely attended scientific meeting. It is also possible that these might be combined if it is deemed more effective for broader community attendance and end-user feedback in a particular year.

4.4   Tangible metrics

Feedback from the users will be used to score the prototype applications using the following metrics:

- How complete is the application? Does it perform the functions you require? Is it fast enough? How can it be improved?
- Is the application easy to use?
- Is the application well documented?
- Is the application portable? Are you able to easily download and install it?
- Does the application fit well into the overall suite of available software?

In addition, the users will be asked to provide general feedback on topics such as:

- In general, are you able to understand the scope and direction of the LROSE project?
- Are you able to get information on new applications as they become available?
- How well does the LROSE project team respond to your requests for information, or changes to software?

4.5   Scientific Applications and Use Cases

The ultimate goal of LROSE is to directly enable scientists to meet the NSF mission *"to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national*

defense." Dedicated testing and application of the software elements will be conducted by the PIs on scientific applications and 'use cases' that are relevant for the broader radar and lidar communities.

a. Airborne Radar and Lidar Hurricane Application

Airborne radars have been used to probe the structure of many mesoscale weather phenomena such as bow echoes, convective boundary layers, fronts, hurricanes, mesoscale convective systems, tornadoes, and winter storms (Lee et al. (2003) and references therein). Hurricane reconnaissance flights using airborne Doppler radars have been conducted for several decades, and form a vast archive of existing radar data. With recent investments to the NOAA P3 Hurricane Hunter aircraft, upgrades to their radar systems, and new airborne phased-array radar technologies (Vivekanandan et al. 2014) this archive is expected to grow in the coming years. Despite the open availability of this unique dataset from both NOAA and NSF deployments, the lack of robust software and difficulty working with the data has hindered the use of this resource.

An example use case for this dataset is shown in Figure 2, illustrating the steps from raw *data* to scientific *information*. Raw airborne radar data contains both weather and non-weather echoes, but the non-weather echoes must be removed prior to performing an analysis. Fig. 2a highlights the variety of non-weather echoes present in the original data, which are not readily apparent to a non-trained user. A significant limitation in earlier airborne Doppler radar studies has been the lengthy time and effort required to quality control the radar data. An automatic technique to remove the non-weather echoes has been developed by the PIs (Bell et al. 2013) to reduce the editing time from several weeks down to a few days or hours, depending on the intended use of the radar data. The new editing technique is a powerful way to process large volumes of data, but the implementation depends on older, interactive software that is not ideal for batch mode processing. New displays and streamlined editing tools are required to improve the data processing capability.
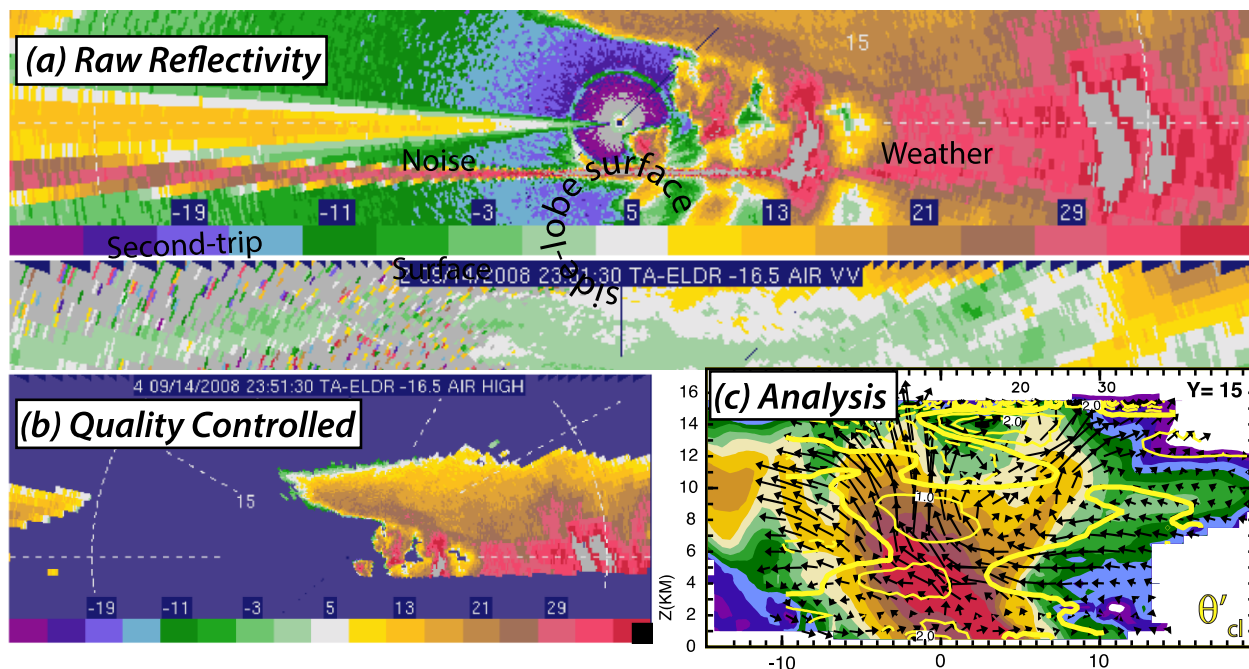

Figure 2. Sample processing of airborne Doppler radar from raw data to scientific information. (a) Raw airborne radar reflectivity from Typhoon Hagupit (2008) with weather and non-weather echoes highlighted and (b) automatic quality controlled data (from Bell et al. 2013). (c) Wind and buoyancy analysis from quality controlled data (from Bell and Montgomery 2010).

Once the data quality has been assured and the non-weather echoes removed (Fig. 2b), the data must be further processed into meteorological variables such as wind velocities. Fig. 2c shows the result of

processing the data using a 3D variational wind retrieval (Reasor et al. 2009) and thermodynamic retrieval (Roux et al. 1993). The resulting analysis reveals a strong updraft driven by a warm temperature anomaly in an intense thunderstorm. Both of the analysis techniques used in this example were performed using older software written in FORTRAN that is not extensible or easily modified. The input for the wind retrieval uses an old radar format, and outputs in a custom binary format. The thermodynamic retrieval reads in a different binary format, and then outputs into yet another different custom binary format. The PI has developed new analysis techniques and prototype software in C++ to perform these retrievals using newer data formats including netCDF (Bell et al. 2012; Bell and Foerster 2013). Further software development and testing is required to make these tools robust and available for the broader radar community.

The PIs have extensive experience working with airborne radar, and a graduate student under the advisement of the PI will extensively test the LROSE software developed under this proposal for hurricane applications using existing airborne radar and lidar datasets. The expectation is that this application and software use case will apply broadly to other university researchers interested in airborne radar data. In addition, the new software will provide a valuable set of tools to allow the numerical weather prediction community better access to radar data to improve hurricane forecasts. Hurricanes are high impact weather events that can profoundly affect lives, property, and our national economy, and preliminary results assimilating radar data into numerical forecast models are promising (Zhang et al. 2013). The PI is funded to conduct integrated hurricane research and radar education through an NSF CAREER award, and the software testing will be synergistic with those research and education efforts in radar meteorology.

b. **Ground-based WSR-88D and Mobile Radar Heavy Rain Application**

The Next-Generation Radar (NEXRAD) network consists of 160 Weather Surveillance Radars (WSR-88D's) that cover much of the contiguous United States, Hawaii, parts of Alaska and some American territories (Fig. 3). This network of 10-cm wavelength radars is a fundamental tool for short term forecasting (less than ~6 hours) of precipitation and all forms of severe weather (tornadoes, hail, strong straight line winds as well as flash floods). These radars collect a daunting amount of information (see section 1.1) but advancements in data storage and transmission have allowed full-resolution data to be made available in real-time and in archive mode (Crum et al. 1993). A ~$50M Congressional investment was recently made to upgrade the radars to dual-polarization, providing new avenues for microphysical research that make this a key dataset for the community. Polarimetric measurements provide the ability to distinguish among rain, snow, hail, and non-meteorological targets such as birds, insects and ground clutter, and can significantly improve rain rate estimates. The NEXRAD network is likely to be the primary source of remotely sensed observations with high spatial (1 km x 1 degree) and temporal resolution (sweeps as frequent as 360 degrees every minute) for the next decade or more. The development of LROSE will enable research scientists to more efficiently exploit this complex and vast data source.
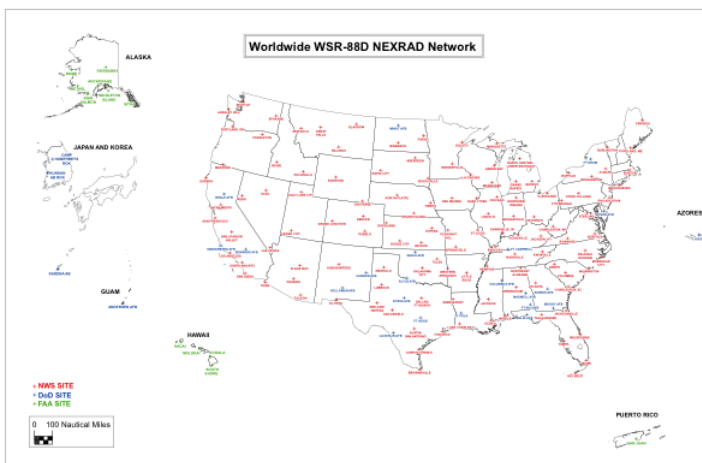

Figure 3. Worldwide WSR-88D NEXRAD Network. Locations indicate radar installations managed by the National Weather Service (red), Department of Defense (blue), and Federal Aviation Administration (green). From the National Weather Service.

One proposed application is the use of NEXRAD data to better understand and ultimately forecast heavy rains and flash

floods. In the contiguous United States there are over 3000 reports of 25 mm h^{-1} for an hour or greater per annum (Brooks and Stensrud 2000). Rain rates greater than 25 mm h^{-1} occur primarily during the warm season in the midlatitudes (Brooks and Stensrud 2000) and throughout the year in the tropics. Schumacher and Johnson (2006) cataloged 184 events over a 5-year period where the rainfall exceeded the 50 year recurrence for that location. Over 90% of these events produced a flash flood. Flash floods, which occur on the same temporal and spatial scales as the precipitation that caused them, are the leading cause of deaths in the United States due to weather phenomena (Wood 1994).

The research over the last 50 years has identified several fundamentals pertinent to heavy rains and flash floods on the U.S. mainland (Maddox et al. 1979, Chappell 1986, Doswell et al. 1996, Konrad 1997, Baeck and Smith 1998, Davis 2001, Schumacher and Johnson 2005, 2006) and Hawaii (Kodama and Barnes 1997, Lyman et al. 2005). Much of the results have been based on rain gages that have poor resolution in many regions. Application of the NEXRAD radar network to heavy rain and flash floods will allow scientists to better understand the convective and mesoscale processes that contribute to the event and supply information on spatial and temporal scales far more accurately than the relatively coarse rain gage network (e.g., Petersen et al 1999).

We will choose several heavy rain – flash flood events from contrasting parts of the country to test and validate the LROSE software. In addition to the NEXRAD network, these cases may also include data from the DOWs, the Front Range Observational Testbed (FRONT), or other ground-based radars. These case studies will lead to an identification of what modules in the initial wish list are most useful and what additions to the software should undergo development. The Co-PIs have extensive experience with radar datasets and the dynamics of heavy rain producing systems, especially those affecting Hawaii. A graduate student under the advisement of the Co-PI (Barnes) will extensively test the software developed under this proposal for heavy rain studies using existing NEXRAD datasets. The expectation is that this use case will apply broadly to other university researchers interested in ground-based dual-polarimetric radar data, and will provide a valuable set of tools to improve heavy rain forecasts.

c. Modern Multi-Sensor Field Projects

Modern atmospheric experiments supported by NSF can bring an impressive array of instrumentation into the field, including multiple radars, lidars, and profilers. A few example use cases of how the software would be used in recent and future experiments are detailed below.

HERO (Hawaiian Educational Radar Opportunity) was an NSF Educational Deployment led by the PI that was conducted on Oahu in October and November 2013 with the DOW radar (Bell 2014). 3-cm radars such as the mobile DOW are cheaper to develop and deploy than NEXRAD systems, but can attenuate in heavy rain. Mobile radars and profilers managed by NSF are popular requests by universities for education purposes, but many new potential users do not have much experience using radar software. Developing new tools for the Educational Deployment of NSF assets will help to expand this user base.

PECAN (Plains Elevated Convection at Night) will be conducted in Kansas and Oklahoma in summer of 2015. The project aims to improve our understanding of nocturnal convection over a stable atmospheric boundary layer, a phenomena which contributes significantly to U.S. rainfall totals. This large field project will include 13 ground-based lidars, 2 airborne lidars, 9 wind profilers, 7 ground-based radars, and 1 airborne radar. Multiple surface, tower, and sounding measurements will also be deployed. Performing the analyses for a large project such as PECAN will be very time consuming if good software is not available. The LROSE suite of applications can potentially streamline the analysis by enabling a common data format and sensor integration to reduce the analysis time and cost significantly. The PI is directly involved in PECAN, and would advertise LROSE and work with the PECAN Science Team to evaluate these new software tools using this extensive and diverse remote sensing dataset.

CSET (Cloud System Evolution in the Trades) will take place during the summer of 2015. Two pulsed instruments – a W-band radar (HCR) and a high spectral resolution lidar (HSRL), will be flown on an

aircraft transiting from California to Hawaii and back. The goals of CSET are to define the evolution of the cloud, precipitation and aerosol fields in stratocumulus clouds over the northern Pacific. The CSET PIs want to be able to overlay radar and lidar data on other data sets, such as numerical models and satellite observations, to evaluate model simulations of cloud system evolution. Handling the navigational information from the aircraft, and combining data with the radar and lidar observations, requires robust and well-tested software modules. The proposed LROSE core will provide that functionality. The Co-PI (Lee) will be involved with the field deployment of the HCR, and will consult with the CSET Science Team on LROSE applications.

5   Collaboration and synergy with other agencies

An advantage of open source projects is that they naturally enhance collaboration between groups working on similar projects, and are an effective means of avoiding duplication of effort (Von Krogh and Spaeth, 2007). In this case, there are at least three other groups working on similar open source radar software:

- US Department of Energy Atmospheric Radiation Measurement project (DOE/ARM), at Argonne National Labs: working on a very capable Python utility library for radar and lidar processing, with an emphasis on climate studies. See https://github.com/ARM-DOE/pyart
- University of Potsdam and University of Stuttgart, Germany: also working on a Python-based library for radar processing, concentrating on precipitation estimation and hydrology. (Heistermann et al. 2013). See http://wradlib.bitbucket.org
- BALTRAD – an advanced weather radar network for the Baltic Sea region: has developed a toolkit for radar data processing (Michelson et al., 2012). See http://git.baltrad.eu

6   Outreach and education plan

Outreach to the community will include: (a) organizing workshops for the user community to provide updates on the core suite and a forum to exchange community-developed algorithms; (b) providing training on the core suite; (c) developing documentation for the core suite and the algorithmic modules; (d) advertising the software to NSF radar and lidar facility users; (e) receiving feedback from the community

A strong community outreach effort has already begun by identifying community software priorities at the 2012 NSF Radar Workshop (Bluestein et al. 2014). Additional outreach will continue through the performance period, with a special emphasis on the AGU Annual Meeting, AMS Conference on Radar Meteorology, and relevant field experiments. The UHM and NCAR PI's are well established in the radar community, and can effectively communicate the availability of the new software to end users.

7   Sustainability plan

The key to sustainability of this system is the open source concept. Although is it not a panacea, open source has been shown to be every effective, both for developing new applications, and for maintaining existing ones. Because multiple people are invested in the system and make use of it, there is a tendency for bugs to be found and fixed, and information about these to be effectively disseminated.

The software source code will be hosted at the GitHub on-line version control repository for wide access to the user community.

The software developed under this proposal will be licensed as fully open source (St. Laurent, 2008). It will be distributed under a BSD-style license, using wording similar to the following:

Copyright (c) (relevant organization). If the software is modified to produce derivative works, such modified software should be clearly marked, so as not to confuse it with the version available from (organization)

Additionally, redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: (1) Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer. (2) Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other

materials provided with the distribution. (3) Neither the name of (organization) nor the names of its contributors, if any, may be used to endorse or promote products derived from this software without specific prior written permission.

DISCLAIMER - This software is provided by the copyright holders and contributors "as is" and any express or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall the copyright holder or contributors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

8   Results from Prior NSF Support

The PI was actively involved in the science and operations planning, execution, and recent post-analysis of the PREDICT field campaign. This research was supported under the "Multiscale Observational Analyses within the Marsupial Pouch of Pre-depression Tropical Disturbances" award (AGS-0851077, $615,890, 2010-2012). This field campaign to improve our understanding of how hurricanes form from tropical waves was summarized in a publication co-authored by the PI (Montgomery et al. 2012). After the very successful deployment, the PI has been involved in data post-processing and tropical cyclone research using the GV data and airborne Doppler radar data from the NOAA P-3 aircraft.

To analyze radar and aircraft observations, a variational analysis technique called SAMURAI (Spline Analysis at Mesoscale Utilizing Radar and Aircraft Instrumentation, Bell et al. 2012) was developed by the PI based primarily on the work of Ooyama (1987, 2002) and Gao et al. (2004). The SAMURAI analysis yields a maximum likelihood estimate of the 3D atmospheric state for a given set of observations and error estimates by minimizing a variational cost function. The technique has several advantages over traditional objective analysis techniques including: (i) observational error specifications for different instrumentation; (ii) use of more complex observations such as radar and lidar data; (iii) the addition of balance or physical constraints such as mass continuity; and (iv) a priori background estimates of the atmospheric state such as global model fields.

As described in section 4.5a, interactive airborne radar editing has been a hindrance due to the time and training required to properly identify non-weather radar echoes. A new algorithm has been developed by the PI in collaboration with NCAR colleagues that can decrease the editing time from weeks to minutes. The algorithm has been verified using published radar datasets edited by different radar meteorologists. The technical details of the algorithm, its skill at distinguishing weather from non-weather echoes in both Electra Doppler Radar (ELDORA) and NOAA P-3 radar data, and its impacts on dual-Doppler wind retrieval were published in a peer-reviewed manuscript authored by the PI (Bell et al. 2013). The SAMURAI software and airborne radar quality control package are open source and available at the PI's GitHub page (see https://github.com/mmbell).

The PI will be supported by a new CAREER award starting in July 2014 to analyze radar data in hurricanes and improve radar meteorology education at UHM ("Impacts of Convective and Stratiform Processes on Tropical Cyclone Intensity Change, AGS-1349881, $783,328, 2014-2019). The software tools developed in the current proposal will have a direct impact on the CAREER research and education plan, and a strong synergy is between the science, education, and technical work is expected.

The UHM Co-PI (Barnes) is currently supported under "Tropical Storms: A Bridge Between Formation and Intensification" (AGS-1042680, $474,455, 2011-2015). The research is aimed at developing more comprehensive descriptions of reflectivity, kinematic and thermodynamic structures within developing tropical storms using observations, including Doppler radar. Publications resulting from this award include Dolling and Barnes (2012a, b; 2014), Barnes and Dolling (2012), and most recently Barnes and Barnes (2014) that used lower fuselage radar to determine hurricane eye and eyewall characteristics. Prof. Barnes has extensive experience using radar data for numerous publications.

9 References

Albo, D., and Kessinger, C. 1996: A New Windshear Detection Algorithm for Doppler Radars, Workshop on Wind Shear and Wind Shear Alert Systems, Oklahoma City, Oklahoma, 13-15 November 1996, *Americ. Meteor. Soc.,* 125-132.

Baeck, Mary Lynn, and J. A. Smith, 1998: Rainfall Estimation by the WSR-88D for Heavy Rainfall Events. *Wea. Forecasting*, **13**, 416–436.

Barnes, C. E., and G. M. Barnes, 2014: Eye and eyewall traits as determined with the NOAA WP-3D lower fuselage radar. *Mon. Wea. Rev., in press.*

Barnes, G. M. and K. P. Dolling, 2013: The Inflow to Tropical Cyclone Humberto (2001) as Viewed with Azimuth–Height Surfaces over Three Days. *Mon. Wea. Rev.*, **141**, 1324–1336.

Bell, M. M., and M. T. Montgomery, 2010: Sheared deep vortical convection in pre-depression Hagupit during TCS08. *Geophys. Res. Lett.,* **37**, *L06802*, doi:10.1029/2009GL042313.

Bell, M. M., M. T. Montgomery, and K. E. Emanuel, 2012: Air-sea enthalpy and momentum exchange at major hurricane wind speeds observed during CBLAST. *J. Atmos. Sci., **69**, 3197-3122.*

Bell, M. M., W.-C. Lee, C. Wolff, and H. Cai, 2013: A Solo-based automated quality control algorithm for airborne Doppler tail radar data. *J. Appl. Meteor. and Clim., **52**, 2509-2528.*

Bell, M. M., and A. M. Foerster, 2013: A new methodology for thermodynamic retrievals in tropical cyclones from multi-Doppler radar analyses. *36th Conference on Radar Meteorology, Breckenridge, CO, 16 – 20 September.*

Bell, M. M., 2014: The Hawaiian Educational Radar Opportunity (HERO). National Science Foundation report, http://www.soest.hawaii.edu/MET/Faculty/mmbell/HERO/UHM_HERO_final_report.pdf

Bluestein, H. B., and co-authors, 2014: RADAR IN ATMOSPHERIC SCIENCES AND RELATED RESEARCH: Current Systems, Emerging Technology, and Future Needs. *Bull. Amer. Meteor. Soc., in press.*

Brooks, H. E., and Stensrud, D. J., 2000: Climatology of Heavy Rain Events in the United States from Hourly Precipitation Observations. *Mon. Wea. Rev.*, **128**, 1194–1201.Chappell, C. F., 1986: Quasi-stationary convective events. *Mesoscale Meteorology and Forecasting,* P.S. Ray, Ed., Amer. Meteor. Soc., 289-310.

Cornman, L.B., Goodrich, R.K., Morse, C.S., and Ecklund, W.L., 1998: "A Fuzzy Logic Method for Improved Moment Estimation From Doppler Spectra", *J. Atmospheric and Oceanic Technology*, 15, p 1287-1305.

Crook, N. A., and J. Sun, 2002: Assimilating Radar, Surface, and Profiler Data for the Sydney 2000 Forecast Demonstration Project. *J. Atmos. Oceanic Technol.*, 19, 888–898.

Crum, T. D., R. L. Alberty, and D. W. Burgess, 1993: Recording, Archiving, and Using WSR-88D Data. *Bull. Amer. Meteor. Soc.*, **74**, 645–653.

Davis, R. S., 2001: Flash flood forecast and detection methods. *Severe Convective Storms*. C. A. Doswell, ed., Amer. Meteor. Soc., Meteor. Monogr., 28, 481-525.

Dixon, M., and G. Wiener, 1993: TITAN - Thunderstorm Identification, Tracking, Analysis, And Nowcasting - A Radar-Based Methodology. *J. Atmos. Ocean. Technol.*, **10**(6), 785-797.

Dixon, M., W.-C. Lee, B. Rilling, and C. Burghart, 2013: CfRadial Data File Format – Proposed CF-compliant netCDF Format for Moments Data for RADAR and LIDAR in Radial Coordinates. URL http://www.eol.ucar.edu/system/files/CfRadialDoc.v1.3.20130701.pdf

Dolling, K. P., and G. M. Barnes, 2012: Warm core formation in Tropical Storm Humberto (2001). *Monthly Weather Review*, **140**, 1177 - 1190.
_____, and _____, 2012: The creation of a high equivalent potential temperature reservoir in Tropical Storm Humberto (2001) and its possible role in storm deepening. *Monthly Weather Review*, **140**, . 492-505..

_____, and _____, 2014: The Evolution of Hurricane Humberto (2001). *J. Atmos. Sci.*, **71**, 1276–1291.

Doswell, C.A., H. E. Brooks, and R. A. Maddox, 1996: Flash flood forecasting: An ingredients-based methodology. *Wea. Forecasting*, **11**, 560-581.

Eaton, B., J. Gregory, B. Drach, K. Taylor, S. Hankin, J. Caron, R. Signell, P. Bentley, G. Rappa, H. Höck, A. Pamment, and M. Juckes, 2011: NetCDF Climate and Forecast (CF) Metadata Conventions Version 1.6. URL: http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.6/cf-conventions.pdf

Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. J. Atmos. Oceanic Technol., 21, 457–469.

Greene, D. R., and R. A. Clark, 1972: Vertically Integrated Liquid Water—A New Analysis Tool. *Mon. Wea. Rev.*, 100, 548–552.

Gu, J.-Y., A. Ryzhkov, P. Zhang, P. Neilley, M. Knight, B. Wolf, and D.-I. Lee, 2011: Polarimetric Attenuation Correction in Heavy Rain at C Band. *J. Appl. Meteor. Climatol.*, 50, 39-58, doi:10.1175/2010JAMC2258.1.

Heistermann, M., S. Collis, M. J. Dixon, S. Giangrande, J. J. Helmus, B. Kelley, J. Koistinen, D. B. Michelson, M. Peura, T. Pfaff, and D. B. Wolff, 2014: The emergence of Open Source Software for the Weather Radar Community. *Bull. Amer. Meteor. Soc.*, *in press*.

Heistermann, M., S. Jacobi, and T. Pfaff, 2013: Technical Note: An open source library for processing weather radar data (wradlib*). Hydrol. Earth Syst. Sci.,* 17, 863-871.

Henja A., and D. Michelson, 2012: Improving the quality of European weather radar composites with the BALTRAD toolbox. *ERAD 2012 - The Seventh European Conference On Radar In Meteorology And Hydrology, Toulouse, France, Meteo France.*

Hubbert, J. C., M. Dixon, and S. M. Ellis, 2009: Weather Radar Ground Clutter. Part II: Real-Time Identification and Filtering. *J. Atmos. Oceanic Technol.*, **26**, 1181–1197.

James, C. N., and R. A. Houze, 2001: A Real-Time Four-Dimensional Doppler Dealiasing Scheme. *J. Atmos. Oceanic Technol.*, **18**, 1674-1683.

Kirstetter, P.-E., H. Andrieu, B. Boudevillain, and G. Delrieu, 2013: A Physically Based Identification of Vertical Profiles of Reflectivity from Volume Scan Radar Data. *J. Appl. Meteor. Climatol.*, 52, 1645–1663.

Kodama, K., and G. M. Barnes, 1997: Heavy rain events over the south-facing slopes of Hawaii: Attendant conditions. *Wea. Forecasting*, **12**, 347-367.

Konrad, C. E., 1997: synoptic-scale features associated with warm season heavy rainfall over the interior southeastern United States. *Wea. Forecasting*, **12**, 557-571.

Lee, W.-C., P. Dodge, F. D. Marks Jr. and P. Hildebrand, 1994: Mapping of Airborne Doppler Radar Data. *J. Atmos. Oceanic Technol.*, 11, 572 – 578.

Lee, W.-C., F. D. Marks, C. Walther, 2003: Airborne Doppler Radar Data Analysis Workshop. *Bull. Amer. Meteor. Soc.*, **84**, 1063–1075.

Lee, W.-C., F. D. Marks, and R. E. Carbone, 1994: Velocity Track Display—A Technique to Extract Real-Time Tropical Cyclone Circulations Using a Single Airborne Doppler Radar. *J. Atmos. Oceanic Technol.*, 11, 337–356.

Lee, W.-C., B. J.-D. Jou, P.-L. Chang, and S.-M. Deng, 1999: Tropical Cyclone Kinematic Structure Retrieved from Single-Doppler Radar Observations. Part I: Interpretation of Doppler Velocity Patterns and the GBVTD Technique. *Mon. Wea. Rev.*, 127, 2419–2439.Lim, E., and J. Sun, 2010: A Velocity Dealiasing Technique Using Rapidly Updated Analysis from a Four-Dimensional Variational Doppler Radar Data Assimilation System. *J. Atmos. Oceanic Technol.*, 27, 1140–1152.

Lehmann and Volker, 2012: Optimal Gabor-Frame-Expansion-Based Intermittent-Clutter-Filtering Method for Radar Wind Profiler. *J. Atmos. Oceanic Technol.*, **29**, 141–158.

Lin, J. W.-B., 2012: Why Python Is the Next Wave in Earth Sciences Computing. *Bull. Amer. Meteor. Soc.*, 93, 1823–1824.

Lyman, R. E., T. A. Schroeder and G. M. Barnes, 2005: The heavy rain event of 29 October 2000 in Hana, Maui. *Wea. Forecasting*, **20**, 397- 414.Maddox, R. A., Hoxit, L. R., C. F. Chappell, and F. Caracena, 1978: Comparison of Meteorological Aspects of the Big Thompson and Rapid City Flash Floods. *Mon. Wea. Rev.,* **106**, 375–389.

Maddox, R. A., C. F. Chappell and L. R. Hoxit, 1979: Synoptic and meso-alpha scale aspects of flash flood events. *Bull. Amer. Meteor. Soc.,* **60**, 115-123.

Matejka, T., and R. C. Srivastava, 1991: An Improved Version of the Extended Velocity-Azimuth Display Analysis of Single-Doppler Radar Data. *J. Atmos. Oceanic Technol.*, 8, 453–466.

Michelson D., J. Koistinen, T. Peltonen, J. Szturc, and M. R. Rasmussen, 2012: Advanced weather radar networking with BALTRAD+. ERAD 2012 - *The Seventh European Conference On Radar In Meteorology And Hydrology, Toulouse, France*. URL: http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/NET_073_ext_abs.pdf

Miller, L. J., C. G. Mohr, and A. J. Weihheimer, 1986: The simple rectification to Cartesian space of folded radial velocities from Doppler radar sampling. *J. Atmos. Oceanic Technol.*, 3, 162-174.

Mohr, C. G., L. J. Miller, R. L. Vaughan and H. W. Frank, 1986: The merger of mesoscale datasets into a common Cartesian format for efficient and systematic analysis. J. Atmos. Oceanic Technol., 3, 146-161.

Montgomery, M. T., C. Davis, T. Dunkerton, Z. Wang, C. Velden, R. Torn, S. Majumdar, F. Zhang, R. Smith, L. Bosart, M. Bell, J. Haase, M. A. Boothe, J. Jensen, and T. Campos, 2012: The Pre-Depression Investigation of Cloud Systems in the Tropics (PREDICT) Experiment: Scientific Basis, New Analysis Tools and Some First Results. *Bull. Amer. Meteor. Soc.,* **93**, *153-172.*

Ooyama, K. V., 1987: Scale controlled objective analysis. Mon. Wea. Rev., 115, 2479–2506.

Ooyama, K. V., 2002: The cubic-spline transform method: Basic definitions and tests in a 1d single domain. Mon. Wea. Rev., 130, 2392–2415.

Oye, R., C. Mueller and S. Smith, 1995: Software for the radar translation, visualization, editing and interpolation. *Preprints, 27th Conf. on Radar Meteorology, Vail, CO*, Amer. Meteor. Soc., 359-361.

Petersen, W. A., L. D. Carey, S. A. Rutledge, J. C. Knievel, R. H. Johnson, N. J. Doesken, T. B. McKee, T. Vonder Haar, and J. F. Weaver, 1999: Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull. Amer. Meteor. Soc.*, **80**, 191-216.

Reasor, P. D., M. D. Eastin, and J. F. Gamache, 2009: Rapidly intensifying Hurricane Guillermo (1997).

Part I: Low-wavenumber structure and evolution. Mon. Wea. Rev., 137, 603–631.

Roux, F., V. Marecal, and D. Hauser, 1993: The 12/13 january 1988 narrow cold-frontal rainband observed during MFDP/FRONTS 87. Part I: Kkinematics and thermodynamics. J. Atmos. Sci., 50, 951 – 974.

Ryzhkov, A., M. Diederich, P. Zhang, and C. Simmer, 2014: Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking. *J. Atmos. Oceanic Technol.*, 31, 599-619.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnic, 2005: The Joint Polarization Experiment: Polarimetric Rainfall Measurements and Hydrometeor Classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824.

Schumacher, R.S., and R. H. Johnson, 2005: Organization and environmental properties of extreme-rain-producing mesoscale convective systems. *Mon. Wea. Rev.*, **133**, 961-976.

_____, and _____, 2006: Characteristics of U.S. extreme rain events during 1999-2003. *Wea. Forecasting*, **21**, 69-85.

St. Laurent, A. M.,2008: Understanding Open Source and Free Software Licensing. O'Reilly Media, Sebastopol, CA, USA.

Steiner, M., R. A. Houze, and S. E. Yuter, 1995: Climatological Characterization of Three-Dimensional Storm Structure from Operational Radar and Rain Gauge Data. J. Appl. Meteor., 34, 1978–2007.

Stumpf, G. J., A. Witt, E. D. Mitchell, P. L. Spencer, J. T. Johnson, M. D. Eilts, K. W. Thomas, and D. W. Burgess, 1998: The National Severe Storms Laboratory Mesocyclone Detection Algorithm for the WSR-88D*. *Wea. Forecasting*, **13**, 304–326.

Tuttle, J. D., and G. B. Foote, 1990: Determination of the Boundary Layer Airflow from a Single Doppler Radar. J. Atmos. Oceanic Technol., 7, 218–232.

Vivekanandan, J., D. S. Zrnic, S. M. Ellis, R. Oye, A.V. Ryzhkov, and J. Straka, 1999: Cloud Microphysics Retrieval Using S-Band Dual-Polarization Radar Measurements. *Bull. Amer. Meteor. Soc.*, **80**, 381–388.

Vivekanandan, J., Lee, W.-C., Loew, E., Salazar, J. L., Grubišić, V., Moore, J., and Tsai, P.: The next generation airborne polarimetric Doppler weather radar, *Geosci. Instrum. Method. Data Syst. Discuss.*, 4, 1-42, doi:10.5194/gid-4-1-2014, 2014.

Von Krogh, G., and S. Spaeth, 2007: The open source software phenomenon: Characteristics that promote research. *J. Strat. Inf. Sys.*, **16**(3), 236-253.

Wood, R. A., 1994: *Storm Data 1993 with Annual Summaries*. National Climatic data center, 80-81.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, Results, and Future Plans. Bull. Amer. Meteor. Soc., 92, 1321–1338.

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, C.H. Bishop, 2009: Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **137**, 2105–2125.