01 Dec 2008

Hydro-NEXRAD: An Updated Overview And Metadata Analysis

Witold F. Krajewski

Anton Kruger

Piotr Domaszczynski

Charles Gunyon

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/civarc_enveng_facwork/2665

Hydro-NEXRAD: An Updated Overview and Metadata Analysis

Witold F. Krajewski[1], Anton Kruger[1], Piotr Domaszczynski[1], Charles Gunyon[1], Bong Chul Seo[1], Radosław Goska[1], James A. Smith[2], Mary Lynn Beack[2] and Matthias Steiner[3]

[1]IIHR-Hydroscience and Engineering, University of Iowa, Iowa City, Iowa, USA
[2]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA
[3]University Corporation for Atmospheric Research, Boulder, CO, USA

Abstract

Hydro-NEXRAD is a prototype system that allows hydrologists to obtain user-specified rainfall data for their research. These data are based on observations collected by the national network of WSR-88D radars, known as NEXRAD. Users interact with Hydro-NEXRAD through a web-based interface that has map-based components for spatial navigation, calendar- and time series plot components for temporal navigation and a menu-based component for selection of processing options. Through the interface, users browse the Hydro-NEXRAD metadata and select data of interest. As the system is approaching the point of being fully operational, the authors and a group of test users have evaluated several aspects of the system.

Metadata remains very important for the system functionality. Radar-based, basin-based and point (for selected set of rain gauge locations) metadata serve multiple purposes: 1) enable users to efficiently search for subsets of data (SQL query, visual inspection), 2) provide information on quality of the collected data archive (missing or corrupt data), 3) and have a scientific value (basin-based metadata has a potential to be used as a precipitation input to hydrologic models).

The authors provide an updated overview of the Hydro-NEXRAD system. Additionally, the authors present the complete set of Hydro-NEXRAD metadata and discuss their possible applications.

Introduction

As a web-based system, Hydro-NEXRAD can be accessed by any researcher interested in radar rainfall data and derived rainfall products. The system gives access to radar data collected by the national network of WSR-88D radars (NEXRAD), and allows users to order radar-based products. System's graphical user interface (GUI) has been designed to make data selection easy and intuitive. Users interact with GUI in order to specify a number of parameters defining an order (Figure 1).

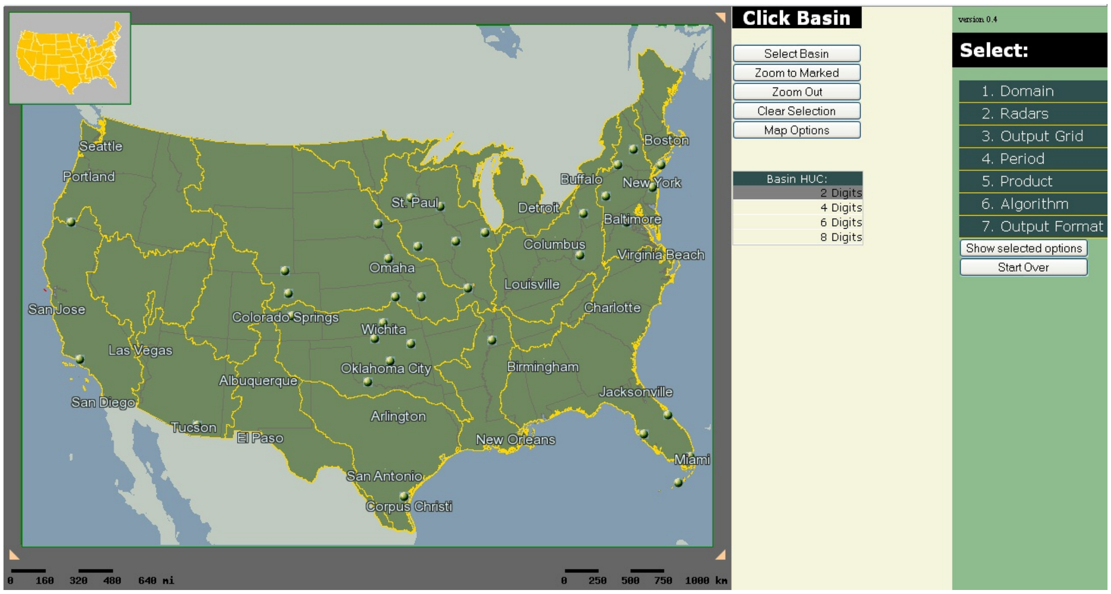

Figure 1: Hydro-NEXRAD graphical user interface

In the process of specifying an order, a user can select product's spatial domain to be an USGS hydrologic unit on any of four levels (region, sub-region, accounting unit or cataloging unit), a radar, a point (for selected set of rain gauge locations), or an arbitrary selected rectangular region. In case of hydrologic units and arbitrary domains, rainfall products can be computed from single radar data, or data from multiple radars can be merged, to provide full domain coverage. Merging can be done on data level or product level. Data level merging requires multiple radar data to be combined before any data processing is performed, whereas in product level merging, individual radar data is first processed separately and in the last step combined with other radars over the selected domain (Krajewski et al. 2008, this volume).

Temporal resolution of system's products for reflectivity and rainfall maps is approximately 5 minutes and is equal to NEXRAD system scanning intervals. Accumulation maps in the Hydro-NEXRAD system accrue rainfall in 15 minutes, 1 hour and 24 hour intervals. Time range selection is limited by data availability. The Hydro-NEXRAD system collects data from approximately 40 NEXRAD radars and data amounts differ between radars. The complete prototype system will have five years of data available for each location and for some selected location there will be a complete record of basic data (~10-14 years). Searching for data that meets user-specified criteria in case of such vast amounts of high resolution data is not trivial. In the Hydro-NEXRAD system, searching is strongly supported by pre-computed set of metadata, describing available data from hydrologic point of view.

Hydro-NEXRAD system provides a selection of grids for data reference. These grids have different characteristics and provide different product spatial resolution. NASA LDAS grid is a 1/8th degree latitude/longitude resolution grid commonly used for global modeling using satellite data. HRAP grid is a grid coordinate system used within the National Weather Service and hydrologic community for use with NEXRAD products. HRAP's spatial resolution is approximately 4 km by 4 km (Reed and Maidment 1999). Super-HRAP (S-HRAP) is a modification of HRAP grid increasing its original resolution to be approximately 1 km by 1 km. "LatLon" grid allows for data reference on the 1' by 1' resolution, latitude/longitude grid. In addition to mentioned grids, radar data can be presented on the original polar, radar-centered grid. Implementation of user specified grid is planned.

Computation of rain-rate maps and accumulation maps from radar reflectivity values requires additional parameters to be specified (*Z-R*, hailcap, etc.). Users with less experience can choose pre-defined algorithms with different level of data quality control (High-Fidelity, Quick-Look). Advanced users can control parameters for *Z-R* relationship, no-rain threshold, hailcap, and other options. Custom algorithms also allow choosing data source to be a single radar elevation scan or a CAPPI construct with user specified height. Additional data correction includes anomalous propagation (AP) removal, vertical profile of reflectivity (VPR) range effect correction, and advection correction.

Final products can be packaged as *ASCII* or *ArcASCII* formatted files. *ASCII* format allows for easy data interpretation, and its analysis is software independent. The *ArcASCII* format is an ESRI format suitable for easy data display using commercial software (for example *ArcMap*).

When Hydro-NEXRAD system completes an order, products become available at a remote *FTP* server, and user is notified by e-mail. Time required for completing an order is strongly related to the amount of quality control steps selected by a user. The variety of Hydro-NEXRAD products make the system suitable as a rainfall data source for many hydrologic applications including hydrologic models, radar versus rain gauge data comparison, individual storm investigations and more.

Metadata Computation Methodology

Searching vast archives of high spatial and temporal resolution NEXRAD data can be tedious and time consuming. For many scientific initiatives, radar-based precipitation is the main source of rainfall information. Well defined set of metadata can clearly identify periods of time when radars recorded high reflectivity values over significant coverage areas. Such information can significantly improve process of data subset selection. Example Hydro-NEXRAD metadata use is shown in Figure 2.

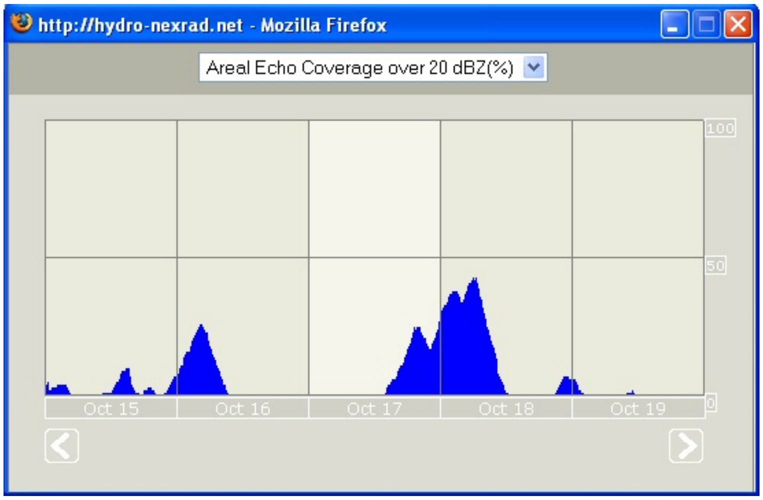

Figure 2: Example of metadata time series

The authors have specified requirements for metadata to be a good descriptor of available data from hydrologic point of view. At the same time computations of metadata should be fast not to slow down data ingest process. Metadata are computed for radars, USGS hydrologic units (basin metadata) and points separately. Each type is controlled by a different system module.

a. Radar Metadata

Depending on Volume Coverage Pattern (VCP) all NEXRAD radars generate Level II files with time resolution ranging from 4 to 10 minutes. Within the Hydro-NEXRAD system, each Level II file is converted into an RLE format file (Kruger and Krajewski 1997). RLE format allows for significant file size reduction and quality control of original data. Based on each RLE radar file a Constant Altitude Plan Position Indicator (CAPPI) scan is calculated. CAPPI scan is a fixed structure of 360 rays each consisting of 230, 1 *km* grid cells. Using a 1.5 *km* CAPPI scan as oppose to a base scan (0.5° elevation) for metadata calculation is dedicated to eliminate ground clutter echoes especially in radar proximity. From a CAPPI scan a set of radar based metadata is computed every 4-10 minutes and stored in the Hydro-NEXRAD database. We discuss members of this set below.

Percentage coverage with reflectivity above a threshold

Each cell from a CAPPI scan is checked for its reflectivity value and if the value exceeds any of the four threshold values (20 *dBZ*, 30 *dBZ*, 40 *dBZ*, and 50 *dBZ*) a variable representing specific metadata is increased accordingly. When all cells have been checked, each threshold metadata are normalized by the total radar coverage area.

Example formula for percentage coverage C of reflectivity above threshold Z_{thresh} is presented as equation:

$$\mathrm{C}(Z_{thresh}) = \frac{100\%}{A_R}\sum_{i=1}^{360}\sum_{j=1}^{230}\left(A_{ij}\cdot T_{ij}(Z_{thresh})\right) \qquad (1)$$

where A_R is the nominal radar coverage area based on 230 km range, A_{ij} is cell area at azimuth *i* and range *j*, and T_{ij} is an indicator function defined as:

$$T_{ij}(Z_{\text{thresh}}) = \begin{cases} 1 \; if \; Z_{ij} \geq Z_{\text{thresh}} \\ 0 \; if \; Z_{ij} < Z_{\text{thresh}} \end{cases} \qquad (2)$$

where Z_{ij} is radar reflectivity value at azimuth *i* and range *j*.

Selected metadata thresholds can be qualitatively related to different precipitation regimes: *20 dBZ* approximates rain/no rain threshold; *30 dBZ* corresponds to light rain; threshold of *40 dBZ* indicates heavy rain; and values *>50 dBZ* indicate heavy rain and possibly hail.

This relation is valid for NEXRAD WSR-88D radars where conversion from reflectivity values [*dBZ*] to rainfall rate [*mm/h*] is based on *Z-R* power law with coefficients *a*=300, *b*=1.4:

$$Z = a \cdot R^b \qquad (3)$$

Maximum observed reflectivity value

From a CAPPI scan, a cell with the highest reflectivity [*dBZ*] value is selected.

$$Z_{MAX} = \max\left\{ \sum_{i=1}^{360} \sum_{j=1}^{230} \left(Z_{ij} \right) \right\} \qquad (4)$$

Detected anomalous propagation echo

Using algorithm developed by Steiner (Steiner and Smith 2002) each radar file in the Hydro-NEXRAD system undergoes anomalous propagation echo (AP) detection process. Radar cells are checked for possible AP echoes using vertical and horizontal echo structure. Detected Anomalous Propagation metadata store the total number of radar cells classified as AP and can be used as an indicator of weather conditions supporting AP occurrence.

$$N_{AP} = \sum_{i=1}^{360} \sum_{j=1}^{230} \left(AP_{ij} \right) \qquad (5)$$

where AP_{ij} is an indicator function defined as in (2) but operated on results of classification (“AP” or “no AP”) determined by the detection algorithm.

Echo advection vector

Using CAPPI scan information from two consecutive radar files, the Hydro-NEXRAD metadata system estimates advection vectors, showing magnitude and direction of horizontal echo movement. To speed up the calculations only one vector per radar field is computed.

Volume Coverage Pattern (VCP)

During data acquisition process, WSR-88D radar antenna is controlled by an automatic scanning program. VCP is selected to optimize data gathering for given meteorological situation. Since VCP information is not embedded in the RLE format, information on VCP is extracted from each radar file, based on number of radar elevation scans, used PRF sequence and scan time.

Volume scan duration

Depending on Volume Coverage Pattern (VCP) it can take from approximately 4 to 10 minutes to complete one volume scan. This does not hold for custom scanning strategies, which are not uncommon in situation of extreme weather events. To give more insight into radar scanning times, authors decided to extract this information and make it accessible to Hydro-NEXRAD users.

Missing radar data

Missing radar data is designed to provide information on completeness of radar information over time. Missing radar data can be caused by technical problems, routine radar calibration and maintenance, corrupt data due to hardware or software problems, and many more. When such vast data amounts are analyzed, it is very beneficial to provide users with information identifying when data is not available.

To enable metadata analysis on different time scales, daily metadata aggregates are computed. This allows for metadata display in a convenient calendar view (Figure 3).

Figure 3: Hydro-NEXRAD metadata calendar view

In the following we define a set of radar metadata daily aggregates:

Daily percentage coverage with reflectivity above a threshold

This group of daily aggregated metadata is based on percentage coverage with reflectivity above a threshold metadata. Metadata values from given radar, and for a given day are combined using the following expression:

$$DC(Z_{thresh}) = \frac{1}{N}\sum_{i=1}^{n}\left(C_i(Z_{thresh})\right) \quad (6)$$

where n is the number of $C(Z_{thresh})$ metadata occurrence during a given day, N is a number of $C(Z_{thresh})$ metadata with non-zero values and C_i is the i^{th} occurrence value. This procedure is repeated for all of the thresholds, i.e. 20 *dBZ*, 30 *dBZ*, 40 *dBZ*, and 50 *dBZ*.

Daily maximum reflectivity value

From time series of maximum observed reflectivity value (Z_{MAX}) metadata, the highest daily value is selected.

$$DZ_{MAX} = \max\left\{\sum_{i=1}^{n}\left(Z_{MAX_i}\right)\right\} \quad (7)$$

Daily rain fraction

Expressed in percent, fraction of a day when echoes classified as precipitation were recorded. To compute this value two previously computed metadata are used: percentage coverage above 20 *dBZ*, and volume scan duration.

$$DRF = 100\% \frac{1}{1440}\sum_{i=1}^{n}\left(VolumeDuration_i\right) \quad (8)$$

where $VolumeDuration_i$ is a metadata value representing volume scan duration time and n is a number of "*percentage coverage with reflectivity above 20 dBZ*" metadata occurrences within one day.

Daily heavy rain fraction

Using 40 dBZ threshold as an indicator of significant rain, the same procedure as in computation of daily rain fraction is applied to estimate what part of daily radar records represent heavy rain.

$$DHRF = 100\% \frac{1}{1440}\sum_{i=1}^{n}\left(VolumeDuration_i\right) \quad (9)$$

Completeness of daily radar data

This metadata represent aggregation of missing radar data periods to a daily scale.

$$\text{MissingData} = \frac{100\%}{1440}\sum_{i=1}^{n}\left(\text{MissingRadarData}_i\right) \quad (10)$$

where $MissingRadarData_i$ represents duration of i^{th} missing data period.

b. Basin Metadata

Hydrologic unit codes developed by USGS are a way of identifying all of the drainage basins in the United States in a nested arrangement from largest (regions) to smallest (cataloging units). For each of those hydrologic units the Hydro-NEXRAD system computes a set of metadata with a fixed 10-minute temporal resolution. Since every basin has its own irregular shape and size and multiple radars can cover the same area, a multistep procedure has been adopted to allow for efficient and accurate computation of metadata. First, reflectivity mosaic is generated. Hydro-NEXRAD mosaic is a 1' longitude by 1' latitude grid, covering continental United States. Each grid cell has its index and is designed to hold reflectivity values. CAPPI products from radar files are used to fill the mosaic with reflectivity values. WSR-88D NEXRAD radars are not synchronized in time with each other, so we have chosen 10-minute time window for calculation of basin metadata. Thus, every 10 minutes there is a new value that characterizes rainfall (or no rainfall) for each basin. Positioning of radars, and by that, polar CAPPI products with respect to the mosaic does not change over time, so to make the process of generating a mosaic faster and more efficient, authors pre-computed a set of look-up tables. Radar look-up tables define relationship between CAPPI product in polar coordinates and mosaic cells. Each basin can be identified on a mosaic as a set of individual cells (Figure 4).

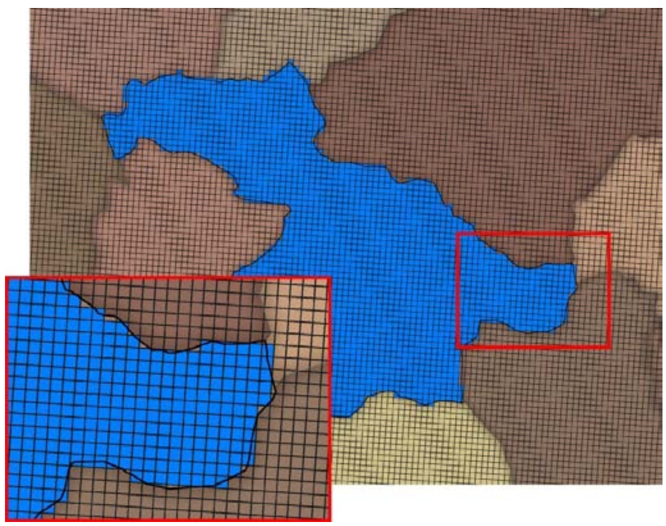

Figure 4: Identification of a basin on Hydro-NEXRAD mosaic.

Once all the cells constituting a basin on the mosaic are known, computation of basin metadata can start. Currently in the Hydro-NEXRAD system there are seven types of basin metadata defined:

Mean areal precipitation

From all the cells constituting a given basin, cells with reflectivity values greater than 20 *dBZ* (rain/no rain threshold) are selected, and converted to rain rate using *Z-R* power law with coefficient a=300 and b=1.4. Estimated value is than normalized by total number of cells constituting a basin N_B.

$$R_{MAP} = \frac{1}{N_B}\sum_{i=1}^{n}\left(RR_i \cdot T_i(Z_{thresh})\right) \qquad (11)$$

where RR_i represents rain rate value of the i^{th} basin cell and T_i is an indicator function defined by (2).

Convective mean areal precipitation

Mean basin coverage with strong, convective precipitation, where only cells with reflectivity values greater than 40 *dBZ* are considered. Symbolic definition is similar to mean areal precipitation definition, but convective precipitation threshold is selected to be 40 *dBZ*.

$$R_{CMAP} = \frac{1}{N_B}\sum_{i=1}^{n}\left(RR_i \cdot T_i(Z_{thresh})\right) \tag{12}$$

Fractional basin coverage

Metadata closely related to mean areal precipitation, represents percentage coverage of a basin with echoes of reflectivity greater than 20 *dBZ*.

$$FCB = \frac{1}{N_B}\sum_{i=1}^{n}\left(T_i(Z_{thresh})\right) \tag{13}$$

where T_i is an index function defined by (2).

Convective fractional basin coverage

This estimates how much of a basin area is covered with strong, convective precipitation. To estimate convective fractional basin coverage sum of all the cells with reflectivity values greater than 40 *dBZ* is normalized with the total number of cells constituting a basin.

$$CFCB = \frac{1}{N_B}\sum_{i=1}^{n}\left(T_i(Z_{thres})\right) \tag{14}$$

where T_i is an indicator function defined by (2).

Maximum observed rainfall rate

For a given basin, the highest observed rainfall rate is stored every 10 minutes.

Location of maximum observed rainfall rate

This metadata stores location of maximum observed rainfall rate in geographical coordinates [decimal degrees of latitude and longitude]. Estimated location corresponds to the center of the cell with the highest rainfall rate value.

Basin coverage with radar information

The Hydro-NEXRAD system continuously collects data from approximately 40 NEXRAD radars. Those radars provide precipitation information for some 800 basins (USGS cataloging units). Some of the basins are only partially covered by radar data. Additionally radars undergo routine maintenance during which no radar data are available. Initially, mosaic cells are set to a “no value” flag. After the process of filling mosaic cells with radar data is over, each basin is checked for a number of “missing cells”. This number is than normalized by the total number of cells constituting a basin to calculate what part of a basin area does not have radar information.

$$BC = \frac{1}{N_B}\sum_{i=1}^{n}(T_i) \quad (15)$$

where T_i is an indicator function defined by (2) an operated on existence of valid values.

c. Point metadata

In the Hydro-NEXRAD system, locations of approximately 1000 rain gauges are stored. For each of these locations, point metadata is computed with 10 minute temporal resolution. Geographical location of a rain gauge is converted to cell indexes of the pre-computed mosaic file. Point metadata for a single location has a form of a 3 by 3 matrix to provide information about rainfall directly over a point and its neighboring cells. Such formulation of point metadata allows users to obtain more information about spatial extend of rainfall experienced over a rain gauge.

Conclusions

With an increasing variety and quality of available atmospheric measurements, it is important for scientists and engineers to have a good access to the data for use in research. Hydro-NEXRAD project enables researchers interested in hydrologic data an easy and efficient way of accessing precipitation information based on vast archives of NEXRAD data.

Whenever a substantial set of data is available a problem with selecting a subset that meets specific user criteria occurs. Hydro-NEXRAD project addresses this problem making available an extensive set of hydrologically meaningful set of metadata. Similar set of metadata is now being considered at the National Climatic Data Center, the national archives of NEXRAD data. Stored in the Hydro-NEXRAD database pre-computed metadata enhances the possibility of making efficient selection of user-specified subset of NEXRAD-based precipitation data. Selected method includes necessary quality control and allowed for a fast and accurate computation. Rich set of metadata, available to Hydro-NEXRAD users in a form of time series plots and colored calendar view greatly improves search and selection of data subsets, provides data quality overview and quality control.

References

Krajewski W.F., B.C. Seo, A. Kruger, P. Domaszczynski, G. Villarini, and J.A. Smith, Multiple radar data merging in Hydro-NEXRAD, 2008 (this volume).

Kruger, A. and W.F. Krajewski, Efficient storage of weather radar data, *Software Practice and Experience*, 27(6), 623-635, 1997.

Reed S.M., and D.R. Maidment, Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling, *Journal of Hydrologic Engineering*, 4(2), 174-182, 1999.

Steiner, M., and J.A. Smith, Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data *Journal of Atmospheric and Oceanic Technology*, 19(5), 673–686, 2002.

