Tropical Cyclone Intensity (TCI) 2015
High Density Dropsonde Quality Assurance Summary
Sponsored by the Office of Naval Research

By Michael M. Bell[1], James D. Doyle[2], Mark Beaubein[3], Todd Allen[3], Bonnie R. Brown[1], Jason P. Dunion[4,5], Patrick Duran[4], Joel W. Feldmeier[7], Lee C. Harrison[3,4], Eric A. Hendricks[6], Will Jeffries[3], William A. Komaromi[2], Jonathan Martinez[1], John Molinari[4], Jonathan R. Moskaitis[2], Daniel P. Stern[2], David Vollaro[4]

[1]University of Hawaii, [2]Naval Research Laboratory, [3]Yankee Environmental Systems, [4]State University of New York at Albany, [5]NOAA/AOML Hurricane Research Division and University of Miami/CIMAS, [6]Naval Postgraduate School, [7]Office of Naval Research

Data Quality Contact: Michael M. Bell <mmbell@hawaii.edu>
HDSS Contact: Mark Beaubein <mcb@yesinc.com>

Document Version Control

Dataset Version	Notes
1.0	Initial Release

I. Overview

This dataset consists of 786 quality-controlled eXpendable Digital Dropsondes (XDD) from the High Definition Sounding System (HDSS) released during the Tropical Cyclone Intensity (TCI) 2015 field experiment sponsored by the Office of Naval Research (ONR). Data were collected over 11 flights in four tropical cyclones in the eastern North Pacific and North Atlantic basins from the WB-57 high-altitude aircraft. The date, number of soundings per flight, and names of each tropical cyclone included in the dataset are listed in Table 1. This document describes the dataset and the quality control (QC) processing steps used to produce it. The quality assurance team would like to acknowledge the broader TCI science team whose support was critical for collecting this dataset. We would also like to acknowledge operational, technical and scientific support provided by the Office of Naval Research, the NASA WB-57 High Altitude Research Program, Yankee Environmental Systems, and NCAR's Earth Observing Laboratory, sponsored by the National Science Foundation.

Date	# of Sondes	Storm (Flight #)
30-Aug-15	59	Erika #1
27-Sep-15	57	Marty #1
28-Sep-15	84	Marty #2
2-Oct-15	84	Joaquin #1
3-Oct-15	78	Joaquin #2
4-Oct-15	84	Joaquin #3
5-Oct-15	83	Joaquin #4
20-Oct-15	13	Patricia #1
21-Oct-15	77	Patricia #2
22-Oct-15	83	Patricia #3
23-Oct-15	84	Patricia #4

Table 1. Number of quality-controlled XDDs per flight

II. Dataset Description

The HDSS data initially records the raw, binary logs from the XDD radio transmissions that include telemetry and atmospheric data for each receiver. After the field phase the raw binary data from the multiple receivers was merged and decoded to arrive at a best estimate of the transmission-error-corrected data. Additional details on the HDSS processing are provided in Chapter 6 of the TCI Operations Plan. Further processing of the data yields user accessible files for scientific use. Here we describe the standard levels of data processing for the HDSS data.

The **Level-0** product is the calibrated output in engineering units of every sensed measurement from the sonde. Users should note that the calibrations are stored within the sonde, and the data are sent calibrated by the sonde.

The **Level-1A** product involves modifications to the data to reconstruct the position data and to remove noise and gross errors from the measurements. This step is intended to produce a high quality dataset as close to the original data as possible. The Level-1A processing reconstructs a latitude and longitude and altitude for every time-step from the Global Positioning System (GPS) high-resolution 4 Hz motion data and low-resolution 0.1 Hz position data. The position integration is filtered using a G-H (aka Alpha-Beta) complementary filter, which is a simplified relative of a Kalman Filter. The filtered integration reduces the impact of single-sample variances in the latitude, longitude, and altitude data. For the TCI2015 sondes, a G-H filter coefficient of 0.05 is used for altitude, and a coefficient of 0.01 was used for latitude and longitude. Some derived quantities such as potential temperature, equivalent potential temperature, wind speed, and direction are also recorded in the Level-1A output. Level-1A data are available upon request, and can be identified by the prefix “ZZ” at the beginning of the filename.

A **Level-1B** product was created in real-time during the field phase by processing the Level-1A data through the Atmospheric Sounding Processing Environment (ASPEN) software. This processing involved additional minimal quality control and file formatting for the production of real-time data files for in-field data quality assessment and mission planning. Further use of these files after the field phase is not recommended. These files can be identified by the prefix “rt” (for real-time) at the beginning of the filename.

The **Level-2** product is the primary data format recommended for publication quality analysis. These files have been subjected to a extensive quality control process described in the next section, and can be identified by the prefix “QC” at the beginning of the filename. The files are distributed in a text-based format containing a header with relevant metadata and seventeen columns of high-resolution data.

III. Level-2 Quality Control Procedures

Each sounding was quality controlled using a combined “subjective-objective” procedure. The objective process used the ASPEN software to apply a consistent set of QC checks and spatial filtering to retain the highest quality data. The subjective process involved visual inspection of every sounding by a minimum of two scientists and manual removal of poor quality or erroneous data. Each of these two steps is detailed below.

IIa. ASPEN objective processing

ASPEN is an interactive software package that has been used to quality control many operational and research soundings in hurricanes and other weather phenomena (Wang et al. 2015). It consists of a graphical user interface for data inspection and a series of QC checks to identify and remove bad data. Though the software has primarily been used with the Airborne Vertical Profiling System (AVAPS) and RD-93/94 dropsondes, the TCI quality assurance team worked with the ASPEN software developers to modify the software and processing configuration to use with the new HDSS data. A brief description of the specific configuration used for the QC is described here, with a more detailed description of all the processing steps available in the software User Manual.

The first few seconds of data after XDD release is unusable as the sensors adjust to the ambient environment. The ambient equilibration was set to a constant 5 seconds for pressure and temperature instead of allowing ASPEN to calculate the equilibration time due to hard-coded assumptions in the software that were not applicable to the XDD sensor response. The relative humidity was removed everywhere below -40 C since it was beyond the sensor range, such that the equilibration time is irrelevant for that variable. The dynamic correction for temperature described in Hock and Franklin (1999) was turned off due to similar hard-coded assumptions about the pressure-temperature-humidity (PTH) sensor suite, and was not needed due to the fast response of the thermistor. The final smoothing wavelength for the soundings was set to 3 seconds for PTH, which is faster than the default 5-second PTH filter used for the AVAPS. The 3-s filter is a "5-s AVAPS equivalent" since the XDDs are falling at ~1.5 times the rate of an RD-93/94. The 3-s filter leaves upper tropospheric features with scales approximately >240 m untouched and remove features < 80 m. The amplitude of features in between these scales would be damped. In the lower troposphere, the filter retains >120 m and remove features <40 m. This filter keeps the data very close to the original measurements, while removing very high frequency oscillations.

The U-blox-6 GPS sensor is similar to that used on recent AVAPS dropsondes, but the XDD sondes have a different aerodynamic response due to their size and lack of a parachute. A constant 10-second value for wind equilibration was chosen as the approximate length of time for the sondes to reach terminal velocity. The dynamic correction for winds was used for the XDDs, since this is a correction for the inertia of the sonde that is obtained directly from the measured acceleration and not from assumptions about the aerodynamics of the sonde. The correction is: $u_{new} = u_{obs} - (du/dt)*(dz/dt)/g$. Note that the dynamic wind correction is the only modification made directly to the HDSS data other than low-pass filtering, but is required in order to adjust the measured wind to the “true” wind. See Hock and Franklin (1999) for a detailed discussion and derivation of the wind dynamic correction. The smoothing filter used to calculate the dynamic correction was kept at the 10-second default to reduce noise in the measured acceleration. The final smoothing filter for winds was set to 3 seconds to match the PTH and keep the wind data close to the original measured frequency.

The “Vertical Velocity” check was modified from a threshold of 2.5 to 5 m s^{-1}. This check compares the vertical velocity calculated hydrostatically from the pressure and from the GPS descent, and then removes the winds if they differ by more than the specified threshold. Subjective inspection of the data suggests that this check is indeed valuable, as the winds tended to be noisy or erroneous when the GPS and pressure vertical velocity were different.

However, the default threshold seemed to be too restrictive due to the possibly realistic variations in the high fall speed of the sonde at upper-levels. The 5 m s^{-1} threshold removed the majority of erroneous winds, with additional removal possible during the subjective inspection.

The sonde weight was set to its measured value of 58 grams. Though the XDD has no parachute, the parachute size was set to a heuristic value of 55.5 cm^3 that yields an approximately zero vertical velocity for a sample of clear air sondes. The blend length of 4 seconds was changed to 2 seconds to retain the scale equivalent filtering in the lowest part of the sounding. All other parameters were left as the default values used in the "Research Dropsonde" configuration.

IIb. Scientist subjective processing

Each sounding was visually inspected and independently analyzed by a minimum of two scientists. While the objective checks from ASPEN did an excellent job of removing poor quality data and correcting the winds for sonde inertia, some instrumentation errors and unphysical features required manual inspection. Each scientist on the quality assurance team processed two flights, with different pairs of scientists for each flight to allow for diversity of experience and scientific viewpoints. Using ASPEN, scientists could interactively remove temperature, humidity, or winds that were subjectively determined to be instrument error or have too poor quality for scientific analysis. After the independent assessment, the majority of soundings had similar if not exact QC choices applied. Any differences in the data removal were then reconciled during a merger process that involved detailed discussion among the quality assurance team. Due to this extensive QC procedure, there is generally high confidence in the quality of the data in the Level-2 products. The subjective data inspection consisted of identifying and correcting 3 main types of instrument errors: a) sensor wetting errors, b) GPS wind errors, and c) altitude errors. Each of these is discussed below.

Three types of sensor wetting errors were identified in the HDSS data: i) anomalous cooling of the temperature due to ice accumulation or evaporation, ii) a greater than moist adiabatic lapse rate while saturated below cloud layer, and iii) erratic behavior in the humidity sensor after being struck by a drop or ice crystal.

The thermistor is coated with a hydrophobic layer, so the number of evaporative cooling problems for the sensor were small. However, some instances of strong cooling or unphysical isothermal layers near 0°C were observed. This "type I" wetting error was sometimes identified by a severe superadiabatic layer beneath a 0°C isothermal layer, indicating that ice had accumulated on the thermistor which was then shed leading to a rapid temperature recovery. However, most instances of isothermal layers observed near the melting level are believed to be real features, so these superadiabatic layers were only removed in limited circumstances. To determine whether a wetting error occurred or a real air temperature measurement, the thermistor temperature was checked against the thermometer on the humidity sensor. The alternate thermometer is slower and less accurate than the thermistor, but provided an independent reference. In most cases the alternate thermometer was colder than the thermistor since it lagged the faster sensor, but wetting errors could be identified in some cases where the alternate thermometer began warming faster than the thermistor. The majority of "type I" wetting errors are believed to have been identified through this procedure, but without further verification it is possible that some limited cases still exist in the dataset.

If the profile is saturated but has a lapse rate greater than moist adiabatic then this can either be a Moist Absolutely Unstable Layer (MAUL) or a wetting error. Since a MAUL must be mechanically forced, they are uncommon and only found in cloud. A wetting error is more common, especially right below cloud base where the humidity sensor does not recover quickly enough to dry conditions after saturation. A subjective examination of whether the saturated dry adiabatic layer was estimated to be within or below cloud base was used to identify these “type II” wetting errors. Since the temperature measurement was still correct in these cases, the procedure to correct this error was to remove the humidity in the suspect layer.

The most common problem in the dataset was the failure of the humidity sensor after being struck by a raindrop or ice crystal. This “type III” wetting error produced variable behavior in the sensor, ranging from complete sensor failure to an apparent dry bias in the lower troposphere. Once the sensor was hit with liquid it rarely recovered to a usable value, such that the humidity data were manually removed in the lower troposphere in many soundings. While the remainder of the humidity data retained in the Level-2 dataset are believed to be valid, the humidity data has the lowest confidence of the measurements due to the prevalence of this error throughout the project and the relatively slow response of the humidity sensor compared to the thermistor.

GPS-derived altitude is known to have significant errors in some cases, such that the preferred altitude for dropsondes is calculated from an integration of the hydrostatic equation. Comparisons with RD-94s during the test flights suggest that good accuracy of the heights was obtained when integrating the XDD pressure and virtual temperature measurements upward from the surface. However, tests integrating downward from the recorded WB-57 aircraft launch altitude suggest the launch altitude was not accurate enough to obtain the dropsonde altitude with sufficient accuracy. If a dropsonde was determined to have not reached the surface then the geopotential altitudes were set to missing, but the GPS altitudes were kept intact. A subjective procedure involving visual inspection and comparison with neighboring surface pressures was used to determine whether the sondes reached the surface. Geopotential heights could also be missing even when the sonde did hit the surface if a significant fraction of temperature data were not available for the hydrostatic calculation. It is also noted that integration errors accumulate upward, such that the geopotential heights in the upper troposphere have greater uncertainty than those in the lower troposphere.

Further manual inspection of the data was performed by each scientist to assess data quality, and some additional data were manually removed due to noise, unphysical values, or other instrumentation errors. Any manually removed data in the final dataset was confirmed by at least two scientists. While it is very difficult to ensure that 100% of errors have been removed, the final Level-2 products consist of a rigorously inspected, subjectively and objectively quality controlled dataset suitable for publication quality analysis.

III. Level-2 Dataset specifications

The sensor specifications for the XDD data are listed below in Table 2. Further details on the HDSS system can be found in the Appendix of the TCI Operations Manual.

The sounding format consists of a header with launch, project, and platform information, followed by a space-delimited column data. The specific measurements and units are listed at the top of each column in every file. Missing data are denoted by the “-999.0” value. The naming convention for each file starts with the prefix "QC" followed by the date and time of the launch in YYYYMMDD-HHMMSS format and a unique 4-character hexadecimal identifier for each sonde. The suffix ".eol" denotes the NCAR/EOL ASCII format Version 1.0. All soundings were processed by ASPEN V3.2-243 using a custom HDSS configuration developed for this project. The eXtensible Markup Language (XML) file with the complete configuration information is available upon request.

Note that the descent rate and geopotential altitude are calculated from the hydrostatic fallspeed and not the GPS fallspeed. While it is possible to assess the air vertical velocity by subtracting the theoretical fall speed, the variable aerodynamic characteristics of the ballistic sondes make this difficult. The vertical velocity was therefore not calculated for inclusion in the Level-2 dataset.

It is also noted that the infrared (IR) sea surface temperature measurements were not included in this dataset. Preliminary assessment of the sensor suggested that only a subset of direct measurements were made of the sea surface due to extensive cloud cover in the hurricane environment. Further quality control of the IR sensor is required, and will be released as a separate dataset.

Measurement	Sensor Type	Range	Accuracy	Resolution	Sampling Rate
Pressure	MEMS Electronic	10 – 1150 hPa	±1.5 hPa @25°C ±2.5 hPa over full range	0.1 hPa	2 Hz
Temperature	Thermistor	-90 to 50°C	±0.14°C @25°C (dark) ±0.5°C over full range with sunlight	1/64°C	2 Hz
Humidity	Poly MEMS Electronic	0 – 100% above -40°C	±1.8% from 10-90% @25°C	0.1% full scale	2 Hz
Horizontal Winds	GPS	0-150 m s^{-1}	0.1 m s^{-1} velocity, 5° heading (under good GPS conditions)	0.1 m s^{-1}	4 Hz

Table 2. XDD Sensor Specifications

IV. Dataset Citation

Please cite this dataset in any scientific publications using the following citation:

Bell, M. M., J. D. Doyle, M. Beaubein, T. Allen, B. R. Brown, J. P. Dunion, P. Duran, J. W. Feldmeier, L. C. Harrison, E. A. Hendricks, W. Jeffries, W. A. Komaromi, J. Martinez, J. Molinari, J. R. Moskaitis, D. P. Stern, and D. Vollaro, 2016: ONR Tropical Cyclone Intensity 2015 NASA WB-57 HDSS Dropsonde Data, Version 1.0. doi:10.5065/D6KW5D8M

V. References

Hock, T. F., and J. L. Franklin, 1999: The NCAR GPS dropwindsonde. *Bull. Amer. Meteor. Soc.*, **80**, 407-420.

Wang, J-H., K. Young, T. Hock, D. Lauritsen, D. Behringer, M. Black, P.G. Black, J. Franklin, J. Halverson, J. Molinari, L. Nguyen, T. Reale, J. Smith, B. Sun, Q. Wang and J. A. Zhang, 2015: A long-term, high-quality, high vertical resolution GPS drosponde dataset for hurricane and other studies. *Bull. Amer. Meteorol. Soc.*, **96,** 961-973.