# Data Quality Assessment Report:

# Ontario Radar DSM

**Water Resources Information Program (WRIP)**

**Geographic Information Branch**

**Ministry of Natural Resources**

[blank page for printing]

# Executive Summary

An accuracy assessment has been conducted on the Ontario Radar Digital Surface Model (DSM) to better understand the limitations and suitability of use for the product. This accuracy report seeks to create accuracy metrics for the DSM based on the United States National Digital Elevation Program (US NDEP) guidelines and discuss the suitability and limitations of the elevation data.

Two test sites were selected for the DSM accuracy assessment: the Hearst test site, located within the Hearst Forest Management Area (49.58° Lat.; -84.61° Long.), and the Belleville test site, extending over the City of Belleville and the surrounding rural area (44.01° Lat.; -77.25° Long.). Elevation accuracies were categorized by the various land cover types within the test sites in accordance with the NDEP guidelines. Relative elevation accuracy was also assessed.

The **Fundamental Vertical Accuracy (FVA)** is defined as the accuracy of the elevation data in open terrain conditions only, and is reported with respect to the 95% confidence level. The following results were calculated for this study:

- Tested 3.5 metres fundamental vertical accuracy at 95% confidence level in open terrain in the Heart test site;
- Tested 3.15 metres fundamental vertical accuracy at 95% confidence level in the open terrain in Belleville test site.

The **Supplemental Vertical Accuracy (SVA)** is used to describe all other land cover types and is reported at the 95% confidence level for normal error distribution patterns and at the 95$^{th}$ percentile for non-normal distributions. Results for this study were as follows:

- Tested 4.21 metres supplemental vertical accuracy at 95 % confidence level in treed wetland category in the Hearst test site;
- Tested 5.04 metres supplemental vertical accuracy at 95 % confidence level in sparse vegetation category in the Heart test site;
- Tested 6.14 metres supplemental vertical accuracy at 95 % confidence level in forest category in the Heart test site;
- Tested 4.84 metres supplemental vertical accuracy at 95$^{th}$ percentile confidence level in urban (built-up) category in the Belleville test site;
- Tested 10.12 metres supplemental vertical accuracy at 95$^{th}$ percentile confidence level in vegetated category in the Belleville test site.

The most significant limitation of the DSM data is its inability to penetrate surface features such as vegetation or anthropogenic features. In general, vegetation coverage has a negative effect on hydrological and geomorphological application of the DSM by not capturing the true elevation of the ground. The data product is well suited for landscape level applications where

local level elevation precision will have minimal impact.   Current uses of the data include land cover, landform and geology mapping, Far North watershed delineation and it has been incorporated into the provincial Integrated Hydrology Data product.

# Table of Contents

## List of Figures



## List of Tables

# 1 Introduction

The Ontario Radar Digital Surface Model (DSM) data product is a continuous raster elevation dataset covering the entire extent of the province of Ontario. A detailed description of the data product can be obtained from the product's metadata record, available through Land Information Ontario's metadata catalogue (MNR, 2012a) and within the technical specifications (MNR, 2012b).

The purpose of this document is to quantify the quality of the Ontario Radar Digital Surface Model (DSM) dataset using a series of vertical and relative accuracy measures as well as standard reporting methods. The document also identifies limitations and suitable uses of the data based on the results of the accuracy assessment.

The document is divided into three main sections. Section 2 (*Assessment Procedures*) describes the assessment methods and measures. Section 3 (*Results*) describes the accuracy assessment results and highlights key factors influencing elevation accuracy. Finally, Section 4 (Conclusions) describes potential applications for the data and known limitations based on the accuracy measures or other published information on the data product.

# 2 Assessment Procedures

## *2.1 Reference to NDEP Guidelines*

The accuracy assessment methods are based mainly on guidelines and recommendations developed by the National Digital Elevation Program (NDEP) for assessing digital elevation data (NDEP, 2004). The NDEP guidelines provide a set of flexible guidelines and recommendations for adoption by any organization for assessing the quality of elevation data products. At the present time, a similar guideline does not exist within Canada or the province of Ontario; therefore the NDEP guidelines are being referenced.

Section 1.5 of the NDEP guidelines is largely utilized for the accuracy assessment which focuses on data quality assessment and reporting mechanisms. The proceeding section of this report describes how Section 1.5 of the NDEP guidelines has been adopted for the accuracy assessment.

### 2.1.1 Vertical Accuracy

There is no universal acceptance of a single statistic as the best elevation data error metric for assessing vertical accuracy, although the NDEP guidelines suggest calculating accuracy as 1.96xRMSE. For the purposes of this assessment, the following statistics were identified as the most relevant error statistics that can be used in combination to obtain a more complete understanding of the error characteristics: Minimum, Maximum, Standard Deviation, Mean Magnitude, RMSE, and NSSDA (95% Confidence Level). For a description of each statistic, please refer to the List of Definitions (pg. 15).

### 2.1.2 Relative Accuracy

Point-to-point (or relative) vertical accuracy can be more important measure of accuracy depending on the application of the elevation data (NDEP, 2004, Maune, 2007)). For example, when elevation data is used for generating watersheds or delineating a flood plain, the morphology and point-to-point accuracy of the elevation surface will influence their accuracy.

Relative accuracy is specifically important for derivative products such as slope and aspect.

To assess relative accuracy, elevation profiles can be used to evaluate the offset between the test and control dataset and influences on morphology and point-to-point accuracy (Maune, 2007).

### 2.1.3 Establishing Checkpoints

The following major steps are used to establish checkpoints for the accuracy assessment procedure in accordance with the NDEP guidelines:

1. A control data source and method is identified to ensure that the accuracy of the check points used to calculate elevation accuracy are at least three times higher than the elevation data being assessed.

2. Checkpoints are generated using randomly distributed points across the study area.

3. Separate checkpoints are established based on topographic variation (slope) and separate classes for open terrain and any other significant ground cover variation (e.g. forest, sparse forest, urban, etc.).

4. Collected checkpoints are edited to minimize errors:

    1.) Check points may be distributed more densely in the vicinity of important features and more sparsely in areas that are of little or no interest.
    2.) For a dataset covering a rectangular area with uniform positional accuracy, check points may be spaced at intervals of at least 10 percent of the diagonal distance across the dataset and at least 20 percent of the points are located in each quadrant of the dataset.
    3.) A minimum of 20 checkpoints (30 is preferred) are needed in each of the major applicable land cover categories.
    4.) Checkpoints should be selected on flat terrain, or on uniformly sloping terrain for x-metres in all directions from the checkpoint, where “x” is the DSM cell size.

5.) Whenever possible, terrain slope should not be steeper than a 20 percent grade.
6.) All ellipsoidal heights are converted to orthometric heights using the proper geoid model.

Because checkpoint surveys are not always error free, special care must be taken to ensure that all survey errors and blunders are identified during this step.

## *2.2 Test Sites and Control Data*

### 2.2.1 Test Sites

Two test sites were selected where LiDAR data, landcover data and high resolution imagery were available to be utilized as control datasets. The two sites also represent two distinct ecological areas in the province with distinctly different land cover, geology and terrain characteristics. The two sites are described below.

*Hearst Test Site*
The Hearst test site (Figure 1: 49.58$^{o}$ Lat.; -84.61$^{o}$ Long.) is 54,954 hectares in size and is located within the Hearst Forest Management Area situated on the very northern edge of the Ontario Shield Ecozone completely within the Lake Abitibi Ecoregion (Crins et al., 2009). The site is dominated mainly by forest cover (68%), which is mainly represented by jack pine, white birch, black spruce and poplar tree species. Forest harvesting is actively occurring in the area resulting in areas that are in stages of regeneration. A significant proportion of the test site is represented by treed wetlands (15%), with smaller percentages of open areas (7%), sparse vegetation (7%) and water (2%). A large portion of the test site (90.5%) has a gradient less than 3 degrees, mainly represented by sand, gravel and glacial deposits as well as glaciolacustrine deposits. The remaining 10.5% of the test site, where the gradient is greater than 3 degrees, is represented by an incised river valley and a large glacial moraine deposit.

*Belleville Test Site*
The Belleville test site (Figure 1: 44.01$^{o}$ Lat.; -77.25$^{o}$ Long.) is 108,872 hectares in size and is located in the Lake Simcoe-Rideau Ecoregion (6E) of the Mixedwood Plains Ecozone (Crins et al., 2009) extending over the City of Belleville and the surrounding rural area. The site is predominantly represented by agricultural and rural land (fields, forage crops, rural properties) at 44% of the area, with smaller percentages of open water (16%), swamp (13%), deciduous forest (11%), and marsh (7%). A large proportion (90%) of the test site has a gradient less than 3 degrees which is mainly underlain by sand, gravel and other glacial deposits. The remaining 10% of the test site exceeds a gradient of 3 degrees; typical to areas of shoreline features along Lake Ontario or the flanks of drumlins in the area.

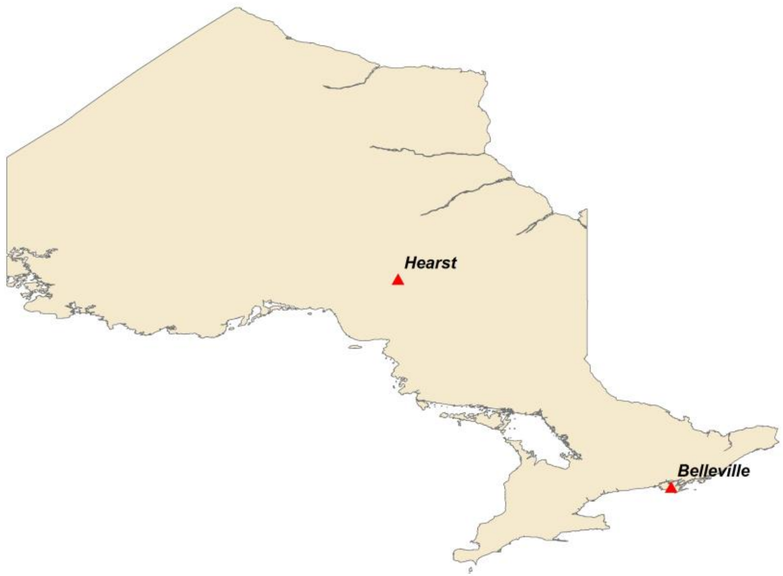



**Figure 1.** Location of the Hearst and Belleville test sites.

### 2.2.2 Elevation Control Data

The LiDAR datasets were previously tested using field GPS survey data. The accuracy of the LiDAR data in open terrain is 0.61 metres, and 0.75 metres under forest cover, both at 95% confidence level, which exceeds the specified accuracy requirement (three times higher) for check points (NDEP, 2004). Using the LiDAR dataset as a control dataset, it is possible to generate a larger, spatially distributed series of check points that can be used to assess the Ontario Radar DSM.

### 2.2.3 Landcover Control Data

In accordance with the NDEP guidelines, the accuracy of the Ontario Radar DSM accuracy was assessed across the dominant landcover types within each study area using existing provincial land cover datasets. Three land cover products were sourced to evaluate ground cover variation for each test site: Far North Land Cover, Provincial Land Cover 2000 and the Southern Ontario Land Resources Information System (SOLRIS).

# 3 Results

## *3.1 Vertical Accuracy*

In accordance with the NDEP guidelines, vertical elevation accuracy is estimated using the RMSE method and reported in ground distances at a 95 confidence level. The accuracies and major error statistics of the DSM for different land covers are shown in Table 1.

| Hearst Test Site | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Land Cover** | **n** | **Minimum** | **Maximum** | **Mean** | **Stdev** | **Skew** | **RMSE** | **Accuracy (m)** |
| Open Terrain | 10,000 | 4.00 | 5.87 | 0.91 | 1.53 | 0.36 | 1.79 | 3.50 |
| Sparse | 10,000 | -4.40 | 7.69 | 1.73 | 1.90 | 0.35 | 2.57 | 5.04 |
| Treed Wetland | 10,000 | -4.69 | 7.21 | 1.19 | 1.79 | 0.38 | 2.14 | 4.21 |
| Forest | 8,911 | 4.00 | 8.79 | 2.43 | 1.97 | 0.27 | 3.13 | 6.14 |
| **Belleville Test Site** | | | | | | | | |
| **Land Cover** | **n** | **Minimum** | **Maximum** | **Mean** | **Stdev** | **Skew** | **RMSE** | **Accuracy (m)** |
| Open Terrain | 3,628 | -5.17 | 4.33 | -0.55 | 1.61 | 0.45 | 1.58 | 3.10 |
| Vegetated Areas | 1,949 | -10.35 | 16.43 | 2.76 | 3.46 | 1.09 | -- | 10.12 |
| Urban (built-up) | 420 | -4.34 | 26.99 | 0.94 | 2.88 | 4.08 | -- | 4.84 |

***Note:*** Error statistics for urban (built-up) points for the Belleville test site resulted in a non-normal distribution with a long tail therefore mean and standard deviation values are not valid.

**Table 1.** Vertical accuracy statistics of the DSM at the Hearst and Belleville test sites.

The **Fundamental Vertical Accuracy (FVA)** is defined as the accuracy of the elevation data in open terrain condition only. It is reported with respect to the 95% confidence level as follows:

- "Tested **3.50 metres** fundamental vertical accuracy at 95 % confidence level in **open terrain**" in the Hearst test site.

- "Tested **3.15 metres** fundamental vertical accuracy at 95 % confidence level in **open terrain**" in the Belleville test site.

The **Supplemental Vertical Accuracy (SVA)** describes the error of the data in land cover types other than open terrain. For this data it is reported as:

- "Tested **4.21 metres** supplemental vertical accuracy at 95 % confidence level in **treed wetland** category" in the Hearst test site;

- "Tested **5.04 metres** supplemental vertical accuracy at 95 % confidence level in **sparse vegetation** category" in the Hearst test site;

- "Tested **6.14 metres** supplemental vertical accuracy at 95 % confidence level in **forest** category" in the Hearst test site.

- "Tested **4.77 metres** supplemental vertical accuracy at 95 percentile confidence level in **built-up** category" in the Belleville test site.

- “Tested **10.07 metres** supplemental vertical accuracy at 95 percentile confidence level in **vegetated** category” in the Belleville test site.

**Open terrain** has the highest overall accuracy ( NSSDA (95%) = 3.5 metres in the Hearst test site and 3.15 metres in the Belleville test site) and highest elevation precision (standard deviation (SD) = 1.53 metres).

**Forested areas** have the lowest overall accuracy: (NSSDA (95%) = 6.14 metres) and lowest elevation precision (Standard deviation (SD) = 1.97 metres) (showing in Figure 2).

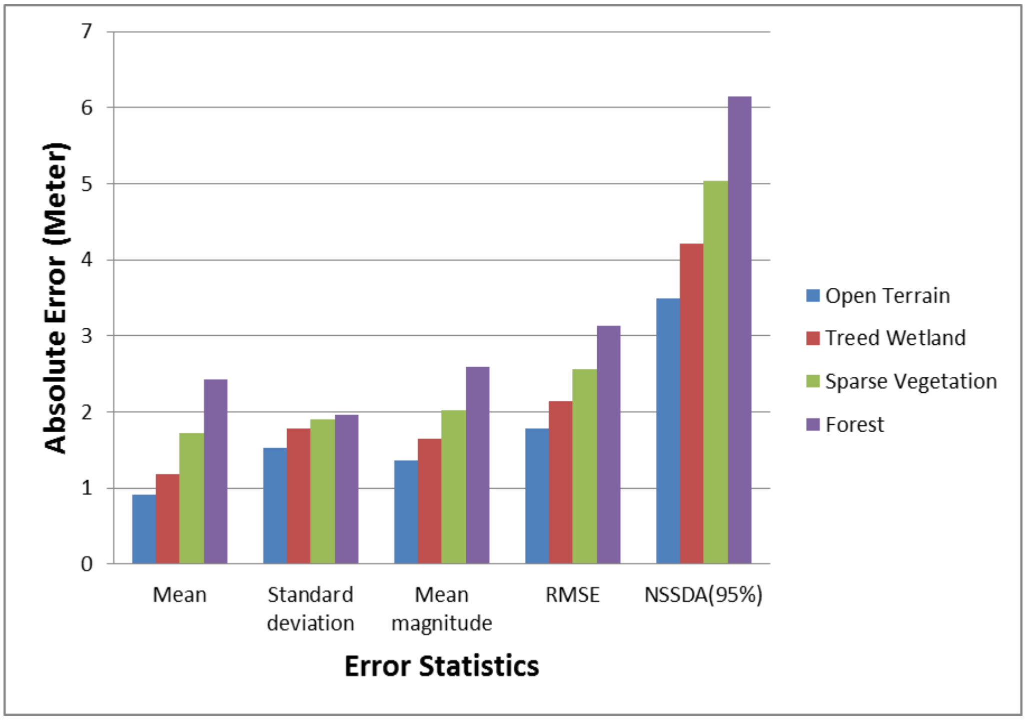



**Figure 2.** Error trends of different land cover types within the northern Hearst test site.

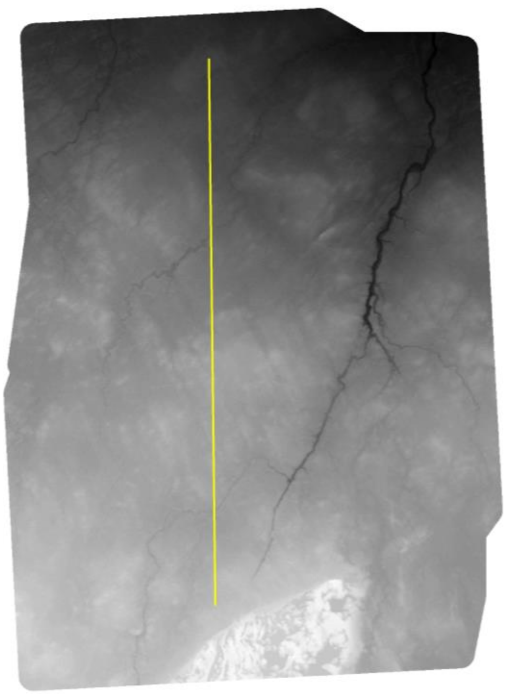
## *3.2 Relative Accuracy*

So far, this report has focused solely on absolute accuracy measures which statistically demonstrate how the DSM values compare to actual ground elevations. For many applications users are more interested in how well the DSM represents the topography. This is referred to as relative accuracy.

A relative accuracy assessment ignores systematic errors and focuses solely on how the random errors in the DSM impact the representation of topographic surface features. For example, a dataset may have a systematic error that causes all values to be shifted 5 metres vertically. The absolute accuracy of this data would be poor but the relative accuracy may still be quite high if surface morphology is well represented. The relative accuracy is particularly important to derivative products such as slope, curvature and feature mapping.

For the purposes of this report relative accuracy is tested using cross sectional profiles to compare the DSM elevations to bare earth LiDAR elevations. Figures 3 A and 4 illustrate a 38 km cross section in the Hearst test site. The resulting elevation profile of the DSM closely conforms to the overall morphology of the LiDAR but there is considerable influence from the vegetation. In most places, the DSM elevation is slightly higher than the LiDAR elevation due to vegetation influence while some specific locations are up to 8m higher due to tall, dense vegetation. A close examination of a cross section created for a river valley (Figures 3 B and 5) shows the impact of the riparian vegetation on the river banks.

A)



B)

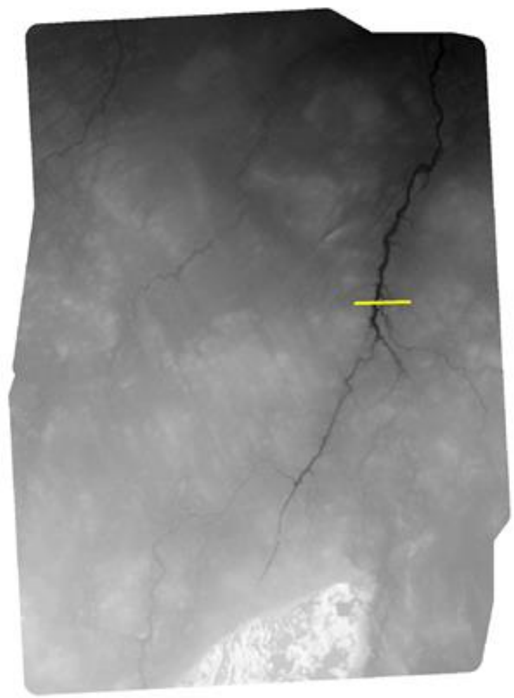


**Figure 3.** Locations of A) the NS cross section and B) the river valley cross section within the Hearst test site.

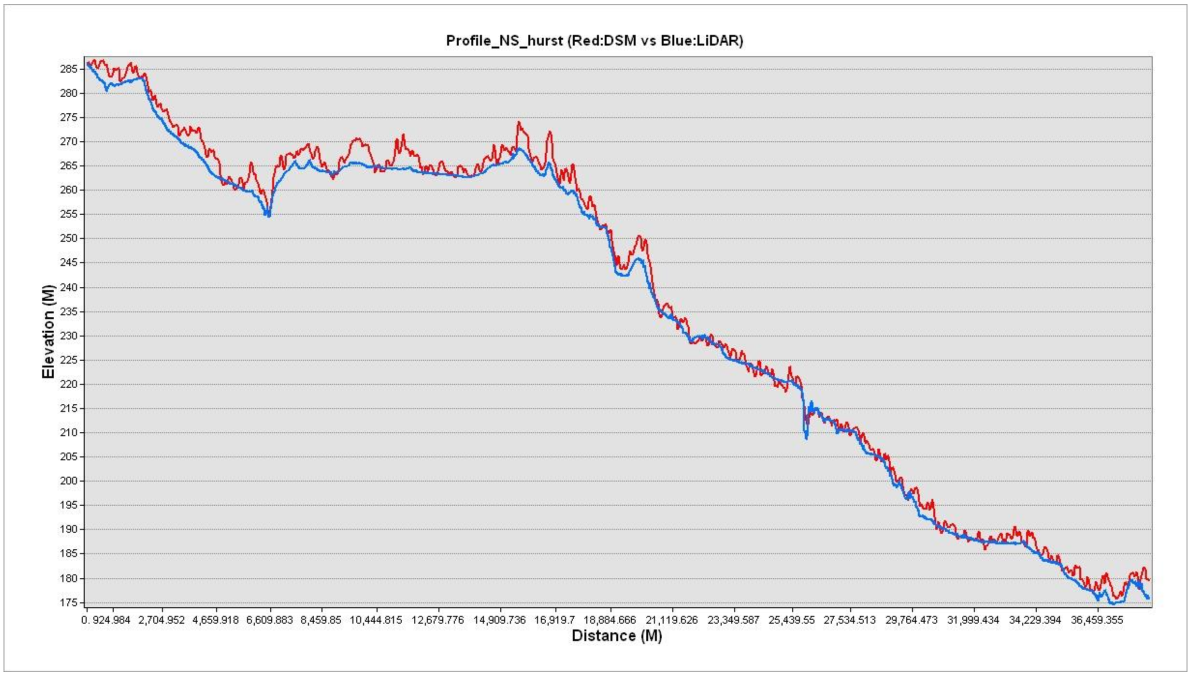



**Figure 4.** Elevation profile of the South-North (left to right) cross section within the Hearst test site (Red: DSM; Blue: LiDAR).

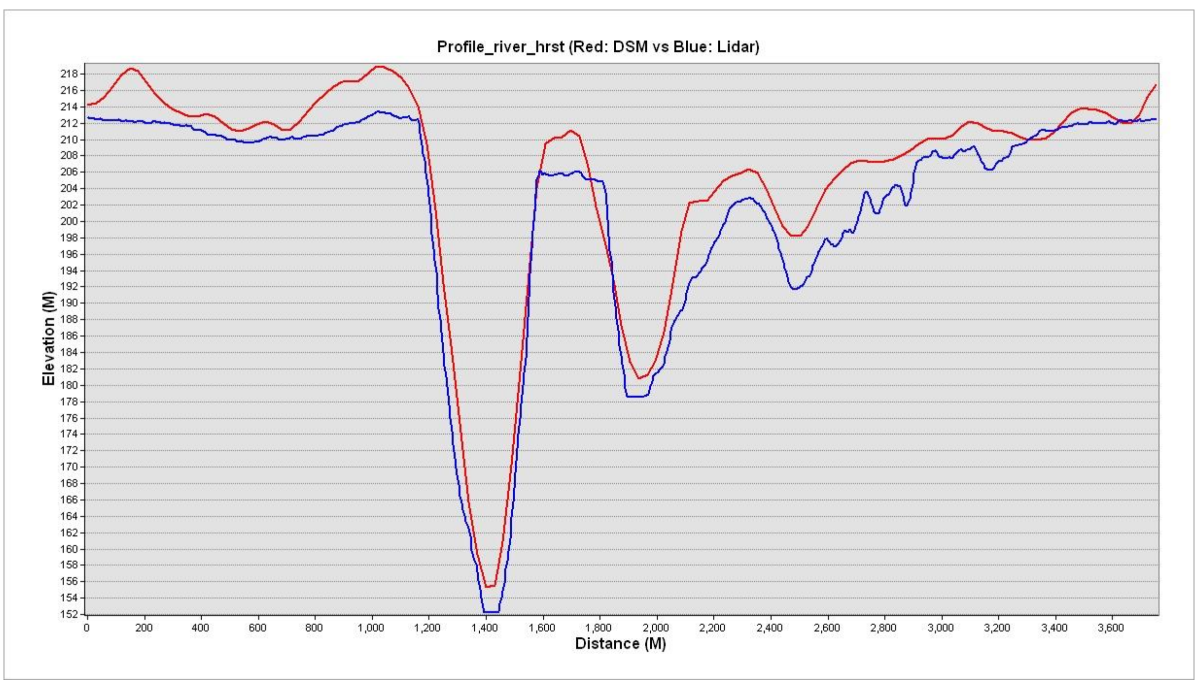



**Figure 5.** West to east profile of a river valley within the Hearst test site.

## 3.3 *Factors Influencing Elevation Accuracy*

The error of the DSM is positively associated to the amount of "openness" of the terrain. In the case of northern Hearst test site, the accuracy values decrease in order from open terrain, treed wetland, sparse vegetation, to forest. In the case of the Belleville test site, accuracy values decrease in order from open terrain, built-up area to vegetated areas.

All error measures, including mean, standard deviation, mean magnitude, RMSE, NSSDA (95%) follow the same increasing trend in the rank from open terrain, treed wetland, sparse vegetation, to forest.

Skewness tests on elevation error for open terrain (0.36), treed wetland (0.38), sparse vegetation (0.35), and forest (0.27) indicate that all land cover categories in the Hearst test site follow a normal distribution very well. This may, in part, be related to the presence of more uniform forest cover across the Hearst test site. However, skewness values increase for the Belleville test site with values of 1.09 in vegetated areas and 4.08 in built-up areas, therefore distributions were non-normal for these landcover classes.

According to the NASA JPL's product assessment, the source data, which is used to build the Ontario Radar DSM, has an absolute height error of 9.0 m at 90% confidence level (equivalent to 10.7 m at 95 % confidence level) for the continent of North America. The tested vertical accuracy of the Radar DSM is significantly better than the 9.0 metres expected for the continent of North America. The fundamental vertical accuracies for Open Terrain classes were 3.50 and 3.10 metres for the Hearst and Belleville sites, respectively, at the 95% confidence level.

Even in vegetated and urban land covers, the overall accuracy was typically between 5 and 6 metres. However, this error is heavily influenced by the height of the surface features. For example, in the Belleville test site where forests tend to be taller than in the Hearst test site, the calculated error was 10.10 metres with many of the error outliers occurring in the urban class situated around tall buildings.

The higher than expected accuracy may be attributed to the following factors:

1. Local Polynomial Interpolation used to create the Radar DSM

   Due to the nature of a SAR coherent imaging system, the source data contains a high level of speckle noise, which appears as a random, high-frequency, salt and pepper effect. Most imaging radar systems average many "looks" to reduce the speckling effect, however, the source data used to generate the DSM was optimized for a wide swath and thus acquired only 1 – 2 looks per sub-swath causing a relatively high speckle noise level (Farr et al., 2007).

Radar speckle is the main source of the noise and the most important limitation of the source data. Through research and testing, the Ontario Radar DSM was generated using local polynomial interpolation (LPI) with carefully chosen parameters. The LPI algorithm was deployed so that the local elevation variation trends were maintained while filtering the speckle noise as much as possible, which is supposed to increase the vertical accuracy of the DSM. A quantitative analysis on the effect of LPI on removing speckles will be carried out in a separate research project.

2. Implications of Land Cover

   Land cover type always plays an important role in the accuracy of elevation data (Hodgson and Bresnahan, 2004; Maune, 2007). For vegetated areas, the signal return was influenced by the vegetation height, structure, and density.

   Since the acquisition time of the source data was winter, the influence of vegetation and forest canopies on the elevation estimates (leaf-off condition) should be minimal for deciduous forests, though tree trunks and limbs produce a complex geometry of radar backscattering. The dense coniferous forests that dominate the Hearst test site will have a strong impact on the accuracy of the DSM, where the DSM accuracy level degrades from 3.50 metres in open terrain to 6.14 metres in forested terrain.

3. Implications to Slope and Aspect derivative surfaces

   Due to the side-looking character of radar imagery, geometric distortions vary with the orientation of topographic features. The slope and aspect of the terrain will vary the strength of reflected energy in SAR backscattering image and degrade the elevation accuracy.

   According to US NDEP, elevation accuracy should be assessed at the terrain surface with slope less than 20 degrees. Both the Hearst and Belleville test sites meet the NDEP standard with 90.5% and 90% having slopes less than 3 degrees respectively.

   For the Belleville test site, an initial test was run and the outliers were examined to determine the cause of extreme outliers. The majority of outliers (error > 3 sigma) could be attributed to areas with steep slopes (greater than 10 degrees). Therefore points located on slopes greater than 10 degrees were eliminated from this statistical analysis.

   Figures 6 and 7 highlight examples where high slopes produce errors greater than 10 metres in the Radar DSM.

**Figure 6.** Steep slopes caused by sand dunes that result in large errors in elevation (The green points on the upper-right of the image showing greater errors) .

The extreme positive and negative errors are caused by slope. Although check points located on slopes greater than 10 degrees were removed, points located at the edge of slopes (within 15 metres of the slope) yielded high errors as well. This can be attributed to the larger cell resolution of the Radar DSM (30 metres) which essentially averages the elevation throughout the cell.  Therefore, if a check point is located at the top of a slope, the Radar DSM will appear too low.  Conversely, if a check point is located at the bottom of a slope, the Radar DSM will appear too high. The following figure depicts this situation where the Radar DSM would be too low at the yellow point and too high at the green point.

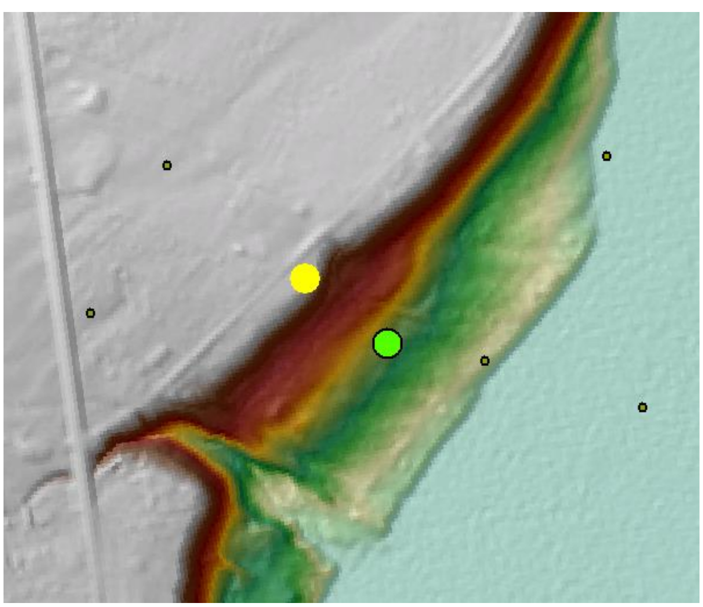


**Figure 7.** Extreme errors caused by slope on the DSM (The yellow point at the top of a slop showing lower elevation value, while the green point at the bottom of a slop showing higher elevation value)

# 4 Conclusions

## *4.1 Limitations of the Data*

*Limitations of a Surface Elevation Model*

The DSM elevation represents the reflective surface of the terrain including bare earth, vegetation cover and man-made features.

In general, vegetation cover has a negative effect on hydrological and geomorphological applications of the DSM by generally overestimating the true elevation of the ground. In addition, the elevation of the DSM does not represent the top of the forest canopy either. The radar signal penetrates the canopy to a certain distance. As shown in literature for canopy height and bio-mass estimation, a mathematic model is needed to establish the relationship between the DSM elevation and canopy elevation. It may be possible to identify the vegetation cover and build an algorithm to remove the impact of vegetation to get the bare earth value of the elevation.

Vector-format, stream datasets could be used to condition or enforce the DSM when creating the flow direction and watersheds (Maune, 2007) similar to the common practice with other digital elevation models.

## *4.2 Advantages of the Data*

The Ontario Radar DSM data provides the following major advantages:

*Consistency in Temporal Scale*

The source data used to construct the Radar DSM is from a near global coverage acquisition (from 60 degrees latitude north to 56 degrees latitude south) in a snapshot time period (11 days), which makes surface elevation data consistent and compatible provincially, nationally and globally.

*Data Acquisition Not Impacted by Cloud Cover*

The source data was acquired using space borne Interferometric Synthetic Aperture Radar (IFSAR). Unlike many other remote sensors which suffer from the coverage of clouds, IFSAR effectively limits the effect of cloud or atmospheric influences.

*High Accuracy in Open Areas at a Medium Scale*

The absolute vertical accuracy at 95% confidence level in open areas (referred to as the Fundamental Vertical Accuracy by NDEP) in this accuracy study is 3.5 metres in Hearst test site and 3.15 metres in the Belleville test site.

*Floating Point Elevation Values*

A continuous floating point elevation surface was interpolated using integer source data, providing a smoother elevation surface which is better suited for DSM analysis and derivative development. For example, a floating point DSM will allow better representation of hydrological trends at regional and medium-scales, which facilitates producing more realistic hydrologic derivatives, such as flow direction and watershed boundaries.

*Medium Spatial Resolution*

The Radar DSM is a surface elevation dataset with a 30-meter horizontal spatial resolution. It is best suited for regional planning purposes and not recommended for any large-scale or engineering level purposes. The relatively larger cell size of the raster elevation data will degrade the accuracy of derivatives, such as slope.

*Data Integration*

The horizontal geo-reference and elevation vertical datum of the DSM has been transformed to conform to existing provincial standards, which allows for integration with other provincial spatial datasets.

## *4.3 Suitable Uses for the Data*

The Ontario Radar DSM data is suitable for the following types of applications:

*A Reference Dataset for Generating Provincial Base Data*

The Radar DSM is the only elevation dataset developed for Ontario that provides high consistency in spatial and temporal scale. The data could be used as a “base” for generating other provincial base datasets, including applications such as image ortho-rectification or geo-referencing.

*Landscape Level Applications*

The data product is best suited for landscape level applications at a regional or provincial scale. It has been used, to-date, in several provincial and federal projects: the Ministry of Northern Development and Mines to assist in landform and geological mapping; IMA/MNR to support land cover classification; WRIP/MNR for generating hydro-conditioned DEMs and community based land use planning; Environment Canada on a project for the Rainy and Namakan Lakes in association with the International Joint Commission (IJC).

# 5 References

Crins, W.J., Gray, P.A, Uhlig, W.C., and Webster, M.C. 2009. The Ecosystems of Ontario, Part 1: Ecozones and Ecoregions, Technical Report SIB TER IMA TR-01, Ministry of Natural Resources, Science and Information Branch, Inventory, Monitoring and Assessment Section, 77p.

Farr, T.G., et al., 2007. The Shuttle Radar Topography Mission. Reviews of Geophysics, Vol: 45 (2).

Federal Geographic Data Committee (FGDC). 1998. Geospatial Positioning Accuracy Standards, Part 3: National Standard for Spatial Data Accuracy, FGDC-STD-007.3-1998.

Hodgson, M.E. and P. Bresnahan. 2004. Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget, Photogrammetric Engineering and Remote Sensing, 70(3), 331-339.

Ministry of Natural Resources. 2009. Accuracy Assessment Report: Far North Land Cover (2000). Inventory Monitoring and Assessment Section, Version 1.0.

Maune, D.F., 2007. Digital Elevation Model Technologies and Applications: The DEM Users Manual, 2ed Edition, the American Society for Photogrammetry and Remote Sensing

Ministry of Natural Resources. 2012a. Ontario Radar DSM, Water Resources Information Program, http://geohub.lio.gov.on.ca/datasets/mnrf::ontario-radar-digital-surface-model

Ministry of Natural Resources. 2012b. Ontario Radar DSM, Water Resources Information Program, URL: https://www.sdc.gov.on.ca/sites/MNRF-PublicDocs/EN/CMID/Ontario%20Radar%20DSM%20-%20User%20Guide.docx

National Digital Elevation Program (NDEP). 2004. Guidelines for Digital Elevation Data, Version 1.0, United States of America.

Spectranalysis Inc., 2004. Introduction to the Ontario Land Cover Database, Second Edition.

## List of Definitions

**Checkpoint** - One of the points in the sample used to estimate the positional accuracy of the dataset against an independent source of higher accuracy.

**Confidence Level** - The probability that errors are within a range of given values.

**Independent source of higher accuracy** - Data acquired independently of procedures to generate the dataset that is used to test the positional accuracy of a dataset. The independent source of higher accuracy shall be of the highest accuracy feasible and practicable to evaluate the accuracy of the dataset.

**Fundamental Vertical Accuracy (FVA)** - The fundamental vertical accuracy of a dataset must be determined with check points located only in open terrain, where there is a very high probability that the sensor will have detected the ground surface. The fundamental accuracy is the value by which vertical accuracy can be equitably assessed and compared among datasets. Fundamental accuracy is calculated at the 95-percent confidence level as a function of vertical RMSE (NDEP, 2004).

**Light Detection and Ranging (LiDAR)** – An instrument that measures distance to a reflecting object by emitting timed pulses of light and measuring the time between emission and reception of reflected pulses. The measured time interval is converted to distance.

**Minimum** and **Maximum -** indicates the highest and lowest error values, covering the entire systematic and random error values.

**Mean Magnitude Error** - indicates the average of all errors regardless of whether they are positive or negative (absolute error values).

**National Standard for Spatial Data Accuracy (NSSDA) 95% Confidence Level** – An accuracy statistic represented as 1.96×RMSE, was developed by the U.S. Federal Geographic Data Committee (FGDC) when they published the National Standard for Spatial Data Accuracy (NSSDA) (FGDC-STD-007.3, 1998). This metric is also known as the Linear Error at 95% Confidence (LE95) which assumes that the RMSE is measured from a normal error distribution (with no or minimum bias).

**95$^{th}$ percentile** – A statistic indicating that 95 percent of the errors in the dataset will have absolute values of equal or lesser value and 5 percent of the errors will be of larger value. With this method, $\text{Accuracy}_z$ is directly equated to the 95th percentile, where 95 percent of the errors have absolute values that are equal to or smaller than the specified amount.

**Relative Vertical Accuracy** - A measure of the point-to-point vertical accuracy within a specific dataset. To determine relative vertical accuracy, the vertical difference between two points is measured. That difference is then compared to the difference in elevation for the same two points on the reference. The difference between the two measures represents the relative accuracy. The reference must have at least three times the accuracy of the intended product accuracy, insuring that all systematic errors and blunders have been removed. Relative vertical accuracy, an important characteristic of elevation data used for calculating slope, should be documented in the DEM metadata file.

**Root Mean Square Error (RMSE)** – A statistic representing the square root of the mean of squared errors for a sample. The RMSE includes both random and systematic error. If there is no systematic error component (i.e. the RMSE is measured from a normal error distribution with no or minimum bias) then the RMSE is the statistical equivalent of the standard deviation.

**Slope** - The measure of change in z-value over distance, expressed either in degrees or as a percent. For example, a rise of 4 metres over a distance of 100 metres describes a 2.3° or 4 percent slope.

**Standard Deviation** – A statistic indicating how spread-out the elevation vertical errors are around the (arithmetic) mean. It can be used as a measure of the elevation data precision.

**Supplemental Vertical Accuracy (SVA)** - The result of a test of the accuracy of z-values over areas with ground cover categories or combinations of categories other than open terrain. Obtained utilizing the 95th percentile method, a supplemental vertical accuracy is always accompanied by a fundamental vertical accuracy. See fundamental and consolidated vertical accuracies. (NDEP, 2004).

## List of Acronyms

DSM – Digital Surface Model

FVA – Fundamental Vertical Accuracy

FGDC – Federal Geographic Data Committee

JPL – Jet Propulsion Laboratory

MNR – Ministry of Natural Resources

NDEP – National Digital Elevation Program

NSSDA – National Standard for Spatial Data Accuracy

LiDAR – Light Detection and Ranging

RMSE – Root Mean Square Error

SOLRIS – Southern Ontario Land Resources Information System

SVA – Supplemental Vertical Accuracy

WRIP – Water Resources Information Program