Correlation Scales of Digital Elevation Models in Developed Coastal Environments

Christopher Small[a] and Robert Sohn[b]

[a] *Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964*
[b] *Woods Hole Oceanographic Institution, Woods Hole, MA 02543*

Abstract

Accuracy of digital elevation models (DEMs) often depends on how features of different spatial scales are represented. Scale dependence is particularly important in low gradient coastal environments where small vertical errors can affect large areas and where representation of fine scale topographic features can influence how DEMs are used for modeling inundation. It is commonly observed that different types of DEMs represent larger, coarse-scale topographic features similarly, but differ in how they represent smaller, finer-scale features. Spatial scale dependence of DEM accuracy can be quantified in terms of the correlation scale (λ_c); the spatial wavelength above which models agree with spectral coherency > 0.5 and below which they differ. We compare cross spectral analyses of the GDEM2 and SRTM global DEMs with 14,572 LiDAR-derived elevations along transects in diverse coastal environments of New York City. Both global DEMs have positive bias relative to LiDAR ground elevations, but bias (μ) and uncertainty (σ) of GDEM2 (μ: 8.1 m; σ: 7.6 m) are significantly greater than those of SRTM (μ: 1.9 m; σ: 3.6 m). Cross-spectral coherency between GDEM2 and the LiDAR DEM begins to roll-off at scales of $\lambda <$ ~3 km, while coherency between SRTM and the LiDAR DEM begins to roll-off at scales of $\lambda <$ ~1 km. The correlation scale below which coherency with LiDAR attains a signal to noise ratio of 1 is ~1 km for GDEM2 and ~ 0.5 km for SRTM; closely matching the divergence scales where the surface roughness of the land cover exceeds the roughness of the underlying terrain.

Introduction

Hazard assessments and inundation modeling of coastal areas rely heavily on both the accuracy and resolution of digital elevation models (DEMs). In many coastal areas global DEMs offer the most complete representation of coastal elevations and morphology available. Two distinct classes of global DEM are currently in widespread use: passive source stereographic models derived from optical imagery like the ASTER GDEM2 (Abrams et al. 2010), and active source ranging models derived from synthetic aperture radar like the SRTM (Farr et al. 2007). The accuracy of each model depends on multiple factors related to the sensing modality, the procedure used to estimate elevations, and the characteristics of the land surface (Farr et al. 2007; Lang and Welch 1999). The recent release of full resolution 30 m SRTM data for areas outside the US (previously degraded to 90 m) prompts the question of how the accuracy and effective spatial resolution of SRTM and GDEM2 compare, particularly in developed coastal environments where they may be used for inundation modeling and hazard assessments.

The accuracy and resolution of DEMs in coastal environments, where there are relatively small differences in elevation over large areas, are of special interest. At low elevations and gradients the signal magnitude approaches the noise level of the measurements, which can lead to large errors in inundation extent forecasts. This issue is particularly important for developed coastal environments where the spatial extent of inundation can have disproportionate consequences in terms of loss of life and property. There have been several comparative analyses of global DEM vertical accuracy (e.g. (Gesch et al. 2012; Meyer et al. 2012), (Tachikawa et al. 2011; Tadono et al. 2012), (Smith and Sandwell 2003)). Some analyses have included coastal areas (e.g. (Gorokhovich and Voustianiouk 2006), (Hvidegaard et al. 2012)), and some have incorporated land cover/use information (e.g. (Gesch et al. 2012), (Hofton et al. 2006), (Carabajal and Harding 2006)), but we are not aware of any that specifically consider the accuracy and spatial resolution of global DEMs in developed coastal environments. As explained below, the scale and diversity of land cover in developed coastal areas is fundamentally different from most of the environments where previous studies have focused.

The objective of this analysis is to assess the accuracy and scale dependence of the GDEM2 and SRTM global DEMs in developed coastal environments. We address the issue by quantifying the scale dependence of the agreement between these global DEMs and high-accuracy, high-resolution, LiDAR-derived elevations for a diverse variety of coastal environments within New York City (NYC). We quantify the scale dependence by using cross-spectral analysis to estimate the correlation scale (the length scale below which two signals are uncorrelated) of each global DEM with a co-registered DEM and digital surface model (DSM) derived from LiDAR. The LiDAR DEM (LDEM) and DSM (LDSM) have been thoroughly validated throughout the study area, and thus provide high-quality benchmarks for the analysis. We focus on quantifying the lateral length-scale at which the agreement between two models becomes random. This is complementary to, but distinct from, previous studies that used point-to-point

comparisons (e.g., GPS or fiducial) to measure the absolute accuracy of the global DEMs. To our knowledge, the only scale-dependent analyses of global DEMs are those of (Smith and Sandwell 2003) and (Rodriguez et al. 2006) but neither focus on developed or coastal environments.

Data

The geological and geomorphic diversity of NYC includes a wide range of developed and natural coastal environments and land cover/use types. Four of NYC's five boroughs are islands. Included in NYC's 837 km of coastline are beaches, barrier islands, tidal wetlands, estuarine marshes, rivers, channels and a variety of reinforced embankments on the shores of New York Harbor, Long Island Sound, the Atlantic Ocean and the Hudson, Harlem and East Rivers (Fig. 1).

The LiDAR DEM and DSM were produced from a 17-day airborne campaign conducted by Sanborn Inc. in April 2010. The ALS-50 near infrared LiDAR instrument was flown at an altitude of 1100 m with a scan angle of 28°. The 937 km^2 collection of 15 x 10^9 elevation measurements results in a point density of 8 to 12 points/m^2. Comparison of LiDAR elevations with 1722 survey elevations throughout NYC yields a Root Mean Square Error (RMSE) of 0.075 m in elevation. Comparison with 200 building corner points yields a RMSE of 0.33 m horizontal error ((Ahern and Ahn 2011)). Extraction of first and last returns from the full waveform LiDAR allowed for better discrimination of true ground elevations under tree canopies and semi-penetrable land cover types. The LDSM measures tops of buildings, trees and infrastructure, as well as ground elevations where sufficient skyview exists. Known ground elevations were used to extract the DEM from the LDSM and last returns. The LDEM ground elevations under buildings are estimates based on surrounding true ground elevations. Some residual building footprints remain in the LDEM (Fig. 1) but their amplitude is generally < 1 m.

The GDEM2 and SRTM (v2.1) DEMs were obtained from the USGS in the form of 1 arc second (1" = ~28 m at 40°N) resolution geographic grids in the WGS84 horizontal and EGM96 vertical reference systems. The LDEM and LDSM were spatially averaged with a 1" Gaussian kernel and sampled at points coincident with the global DEMs along 6 transects (Fig. 1). Transects were chosen to maximize length, relief and diversity of coastal environments. For each transect, two adjacent rows (or columns) of the global DEMs were compared to check consistency of spectral estimates for similar (but not identical) realizations of the terrain. Transect WE1 extends westward from the East River through midtown Manhattan where neither global DEM was able to register accurate elevations because of the considerable height variability resulting from tall buildings, as can be seen in the LiDAR DSM. The NUM files that accompany SRTM v3 indicate that infill elevations were used to fill coverage gaps in midtown Manhattan and one pol on Jamaica Bay but not any other part of the transects selected for analysis. We exclude midtown from our spectral analyses but include it in Figure 1 to illustrate the infill elevations used to fill the gaps in DEM coverage. Within the NYC study area, the

SRTM grid used between one and four (mode = 3) acquisitions and GDEM2 used between one and twelve (mode = 7) acquisitions.

Analysis and Results

We use scatterplots and moments of difference distributions to quantify the point-to-point correspondence between the global DEMs and the 14,572 LiDAR elevations. The results show closer agreement between LDEM and SRTM than between LDEM and GDEM2 (Fig. 2). In the LDEM and SRTM comparison, 25% of the SRTM elevations underestimate the LiDAR ground elevations while almost all of the GDEM2 elevations overestimate the LiDAR ground elevations. An important difference between global DEMs is observed at the lowest elevations where GDEM2 always overestimates LDEM but SRTM often underestimates LDEM (Fig. 2). This difference is also illustrated by the large, low elevation wetlands (e.g. Jamaica Bay and New Jersey Meadowlands) in Figure 1. The 1st and 2nd moments of the difference distributions give the bias (μ) and uncertainty (σ) of each global DEM relative to the corresponding LiDAR model. Both global DEMs have overall positive bias relative to the LiDAR DEM but GDEM2 bias and uncertainty are much greater than those of SRTM (Fig. 2).

We conduct cross-spectral analyses of the four elevation models along all 6 transects to assess the scale-dependence of correlation. Power spectral density, cross-spectral phase and spectral coherency are estimated using the multitaper method (Thomson 1982) with adaptive weighting (Percival and Walden 1993) and a time-bandwidth product of eight. Multitaper estimation reduces the bias resulting from spectral leakage while minimizing the information loss inherent in the use of conventional tapers and avoiding the need for prewhitening (Thomson, 1982). Adaptive weighting minimizes the mean square error of the spectral estimates by determining the weights for each taper using an iterative procedure that accounts for the (nonwhite) spectral content of the data (Percival and Walden 1993).

The power spectral density estimates show the relative amount of variance over a range of spatial scales (wavelengths). Power spectra of elevation data are often used to quantify scale dependent variance as a proxy for topographic roughness (e.g. (Fox and Hayes 1985). The spectral shape and roll-off of each transect are functions of the elevation profile roughness (Fig. 3, top panels). For example, LDSM consistently has more power than LDEM at short wavelengths because the land cover upper surface (e.g. buildings and trees) is rougher than the underlying topography, especially in developed areas with multi-story buildings. GDEM2 consistently has greater power than SRTM at short wavelengths, perhaps because of noise in GDEM2 resulting from spectral heterogeneity of land cover imaged by ASTER. Spectral power levels for all four elevation models are similar for length-scales > ~1 km, but begin to diverge at length-scales of 0.5 – 1.0 km, where the relief associated with land cover (e.g. buildings and trees) has greater variance than the smoother underlying topography. We refer to this wavelength where the DSM and DEM separate (shown by arrows in Figure 3) as the surface-elevation divergence scale (λ_D).

The cross-spectral analysis (Figure 3, bottom panels) indicates that both global DEMs become incoherent at spatial scales < ~0.5 km. Coherency, which estimates correlation as a function of spatial wavelength, between GDEM2 and both LiDAR models begins to roll-off for wavelengths < ~3 km, while the coherency between SRTM and LDEM begins to roll-off for wavelengths < ~1 km. We define the correlation scale (λ_C) between two models as the length-scale at which the coherency drops below 0.5; a threshold value representing a signal-to-noise ratio of one (Bendat and Piersol 2010). We find that the correlation scale for GDEM2 relative to both LiDAR models is between 1.4 and 0.7 km, while the correlation scale for SRTM relative to LDEM is between 0.7 and 0.4 km. These correlation scales (λ_C) are similar to the surface-elevation divergence scales (λ_D) observed in the power spectra (Fig. 3). This indicates that the correlation scales where agreement among DEMs disappears occurs at divergence scales where the surface roughness of the land cover exceeds the roughness of the underlying terrain.

The cross-spectral analysis suggests that structure at scales < ~1 km is as likely to be noise as true elevation in both global DEMs. To illustrate this, we filter all four DEMs for profile NS2 to remove incoherent structure at wavelengths < 1 km (Fig. 4). The filtered profiles are in much closer agreement than the raw profiles, especially for the SRTM and LDEM, and the positive bias of GDEM2 is readily apparent.

Implications

Our results quantify the correlation scales of the GDEM2 and SRTM global elevation models for both developed and natural coastal environments in NYC. In general, we find that the global DEMs accurately resolve features with length scales > ~1 km, but at shorter length-scales noise overwhelms the elevation signal. The correlation scale for SRTM extends to shorter length-scales (~500 m) compared to GDEM2 (~1 km), and GDEM2 exhibits a systematic, ~8 m positive bias throughout the study area (Figs. 1 & 2). This may be due, in part, to vertical reference error in GDEM2.

The power spectra and correlation scales vary somewhat among the profiles, with most differences occurring at spatial scales finer than ~1 km where the land cover and the underlying topography signals begin to diverge. The accuracy of the global DEMs begins to deteriorate at about the same length-scale where heterogeneous land cover associated with developed environments becomes the dominant signal. The consistently higher variance (power) and lower coherencies we observe for GDEM2 compared to SRTM for length-scales of 0.5 – 3 km suggest that heterogeneous land cover in developed areas introduces more noise into GDEM2 elevation estimates compared to SRTM elevation estimates.

The scale and diversity of urban land cover poses challenges to stereography using decameter resolution imagery. Comparative multi-scale analyses of meter to decameter resolution optical imagery of urban environments reveal considerable intra-urban spectral diversity with characteristic spatial scales of 20 to 50 m (Small 2009). The spectral diversity and scale-dependent spectral mixing endemic to urban land cover violates the

assumptions of spectral homogeneity and Lambertian scattering that are implicit to stereography (Lang and Welch 1999). As a result, the 20-30 m spatial resolution of the ASTER sensor is not well suited to stereography in heterogeneous urban environments with abundant specular reflectors. However, this does not imply that GDEM2 is not well-suited to more spectrally homogeneous environments with greater topographic relief. GDEM2 is an important complement to SRTM because it provides coverage at higher latitudes and in areas of very steep terrain and sand dunes where SRTM often contains voids (Farr 2006).

The correlation scales of the SRTM and LDEM are consistent with, but larger than, previous studies that quantify scale-dependent resolution of SRTM. The continental-scale structure function analyses conducted by Rodriquez et al. (2006) find height error correlation functions dropping rapidly for scales < ~500m. The cross spectral analysis of LiDAR and SRTM for Mojave Desert terrain conducted by Smith and Sandwell (2003) found average spectral coherence of 0.5 at scales of ~200 m. The effective spatial resolution of the stereographic and synthetic aperture radar algorithms are generally coarser than the 30 m grid resolutions but finer than the 0.5 to 1 km correlation scales observed in this study. The larger correlation scales found in this study suggest that global DEMs may have lower effective spatial resolution (or that noise levels are correspondingly higher) in developed coastal environments compared to continental averages and mountainous, high desert, environments. We conjecture that the lower effective spatial resolutions of both global DEMs in developed coastal environments is a result (at least in part) of the heterogeneity of land cover with characteristic spatial scales comparable to the IFOV of the sensors used for the global models.

Overall, both global DEMs are well within their stated accuracy specifications and appear to have closer agreement to measured elevations than in some previous studies cited above. However, in addition to the apparently random errors discussed above, we do observe some systematic errors when comparing full 2D DEMs at 1" resolution. The largest and most obvious errors in both models occur in areas of high relief at pixel scales; (e.g. buildings in Manhattan). Both global DEMs overestimate ground elevation and underestimate building top heights by > 10 m in most areas where buildings exceed tree height. Consistent with the findings of (Hofton et al. 2006), SRTM overestimates elevation in areas of dense, closed canopy forest where the phase center lies within the canopy. GDEM2 also overestimates surface elevations in some wetlands.

Acknowledgements

This research was supported by the Office of Naval Research (grant N000-14-11-1-0653 to C.S.). We thank Tom Farr and Bob Crippen for helpful comments and suggestions. ASTER GDEM is a product of METI and NASA.

References

Abrams, M., Bailey, B., Tsu, H., & Hato, M. (2010). The ASTER global DEM. *Photogrammetric Engineering and Remote Sensing, 76,* 344-348

Ahern, S.C., & Ahn, H.J. (2011). Quality assurance and potential applications of a high density LiDAR data set for the City of New York, Proceedings of ASPRS, (p. 9).

Bendat, J.S., & Piersol, A.G. (2010). *Random Data: Analysis and Measurement Procedures, 4th Edition*: Wiley

Carabajal, C.C., & Harding, D.J. (2006). SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief. *Photogrammetric Engineering and Remote Sensing, 72,* 287-298.

Farr, T.G. (2006). Voids in SRTM data caused by sand dunes. In, *EUSAR 2006 - 6th European Conference on Synthetic Aperture Radar*. Dresden, Germany

Farr, T.G., Rosen, P., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics, 45,* 1-33

Fox, C.G., & Hayes, D.E. (1985). Quantitative methods for analyzing the roughness of the seafloor. *Reviews of Geophysics, 23,* 1-48

Gesch, D., Oimoen, M., Zhang, Z., Meyer, D., & Danielson, J. (2012). VALIDATION OF THE ASTER GLOBAL DIGITAL ELEVATION MODEL VERSION 2 OVER THE CONTERMINOUS UNITED STATES. In, *2012 XXII ISPRS Congress* (pp. 281-286). Melbourne, Australia: ISPR

Gorokhovich, Y., & Voustianiouk, A. (2006). Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics. *Remote Sensing of Environment, 104,* 409-415

Hofton, M., Dubayah, R., Blair, J.B., & Rabine, D. (2006). Validation of SRTM Elevations Over Vegetated and Non-vegetated Terrain Using Medium Footprint Lidar. *PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING, 72,* 279-285

Hvidegaard, S.M., Sorensen, L.S., & Forsberg, R. (2012). ASTER GDEM validation using LiDAR data over coastal regions of Greenland. *Remote Sensing Letters, 3*

Lang, H.R., & Welch, R. (1999). ATBD-AST-08 ALGORITHM THEORETICAL BASIS DOCUMENT FOR ASTER DIGITAL ELEVATION MODELS (STANDARD PRODUCT AST14). In: NASA

Meyer, D.J., Tachikawa, T., Abrams, M., Crippen, R., Krieger, T., Gesch, D., & Carabajal, C. (2012). SUMMARY OF THE VALIDATION OF THE SECOND VERSION OF THE ASTER GDEM. In, *2012 XXII ISPRS Congress* (pp. 291-293). Melbourne, Australia: ISPRS

Percival, D.B., & Walden, A.T. (1993). *Spectral Analysis for Physical Applications*. Cambridge, England: Cambridge University Press

Rodriguez, E., Morris, C.S., & Belz, J.E. (2006). A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing, 72,* 249–260

303 Small, C. (2009). The Color of Cities: An Overview of Urban Spectral Diversity. In
304 M. Herold & P. Gamba (Eds.), *Global Mapping of Human Settlements* (pp. 59-
305 106): Taylor and Francis
306 Smith, B., & Sandwell, D. (2003). Accuracy and resolution of shuttle radar
307 topography mission data. *Geophysical Research Letters, 30,* 20-21-24
308 Tachikawa, T., Hato, M., Kaku, M., & Iwasaki, A. (2011). Characteristics of ASTER
309 GDEM version 2. In, *International Geoscience and Remote Sensing*
310 *Symposium (IGARSS)* (pp. 3657-3660.): IEEE
311 Tadono, T., Takaku, J., & Shimada, M. (2012). VALIDATION STUDY ON ALOS
312 PRISM DSM MOSAIC AND ASTER GDEM 2. In, *2012 XXII ISPRS Congress*
313 (pp. 193-198). Melbourne, Australia: ISPRS
314 Thomson, D. (1982). Spectrum estimation and harmonic analysis. *Proceedings of*
315 *the IEEE, 70,* 1055–1096.
316
317
318

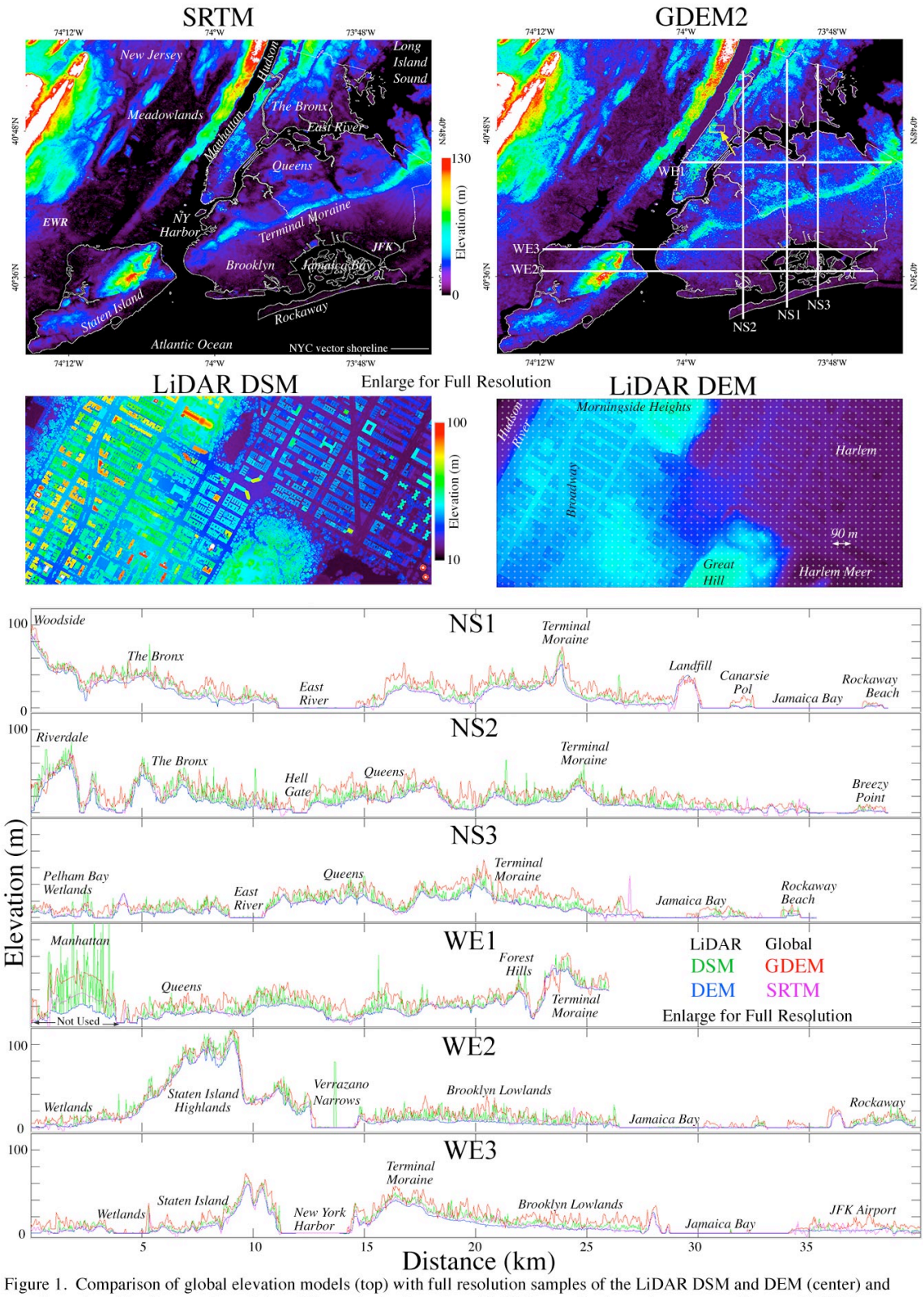

Figure 1. Comparison of global elevation models (top) with full resolution samples of the LiDAR DSM and DEM (center) and coregistered profiles from each model used for analysis. Location of LiDAR sample shown by arrow and box on GDEM2 map.

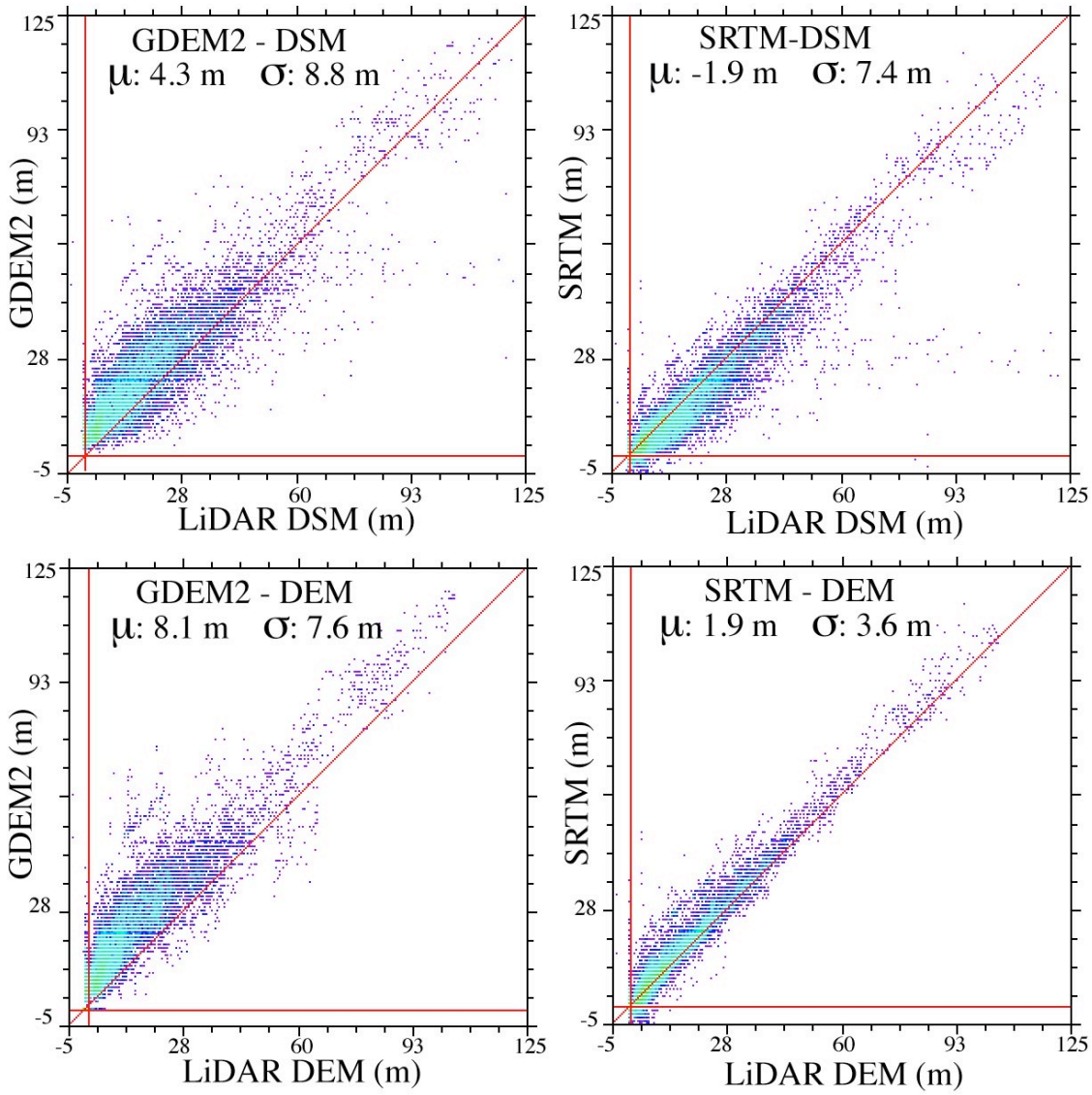

Figure 2. Density shaded scatterplots comparing global elevation models with LiDAR elevations. Distributions of 14,572 coregistered elevations are derived from 6 adjacent pairs of profiles sampled at 30 m postings along transects shown in Fig. 1. Moments of distributions of differences between global models and LiDAR quantify model bias (μ) and uncertainty (σ). Warmer colors indicate greater number of points.

321
322

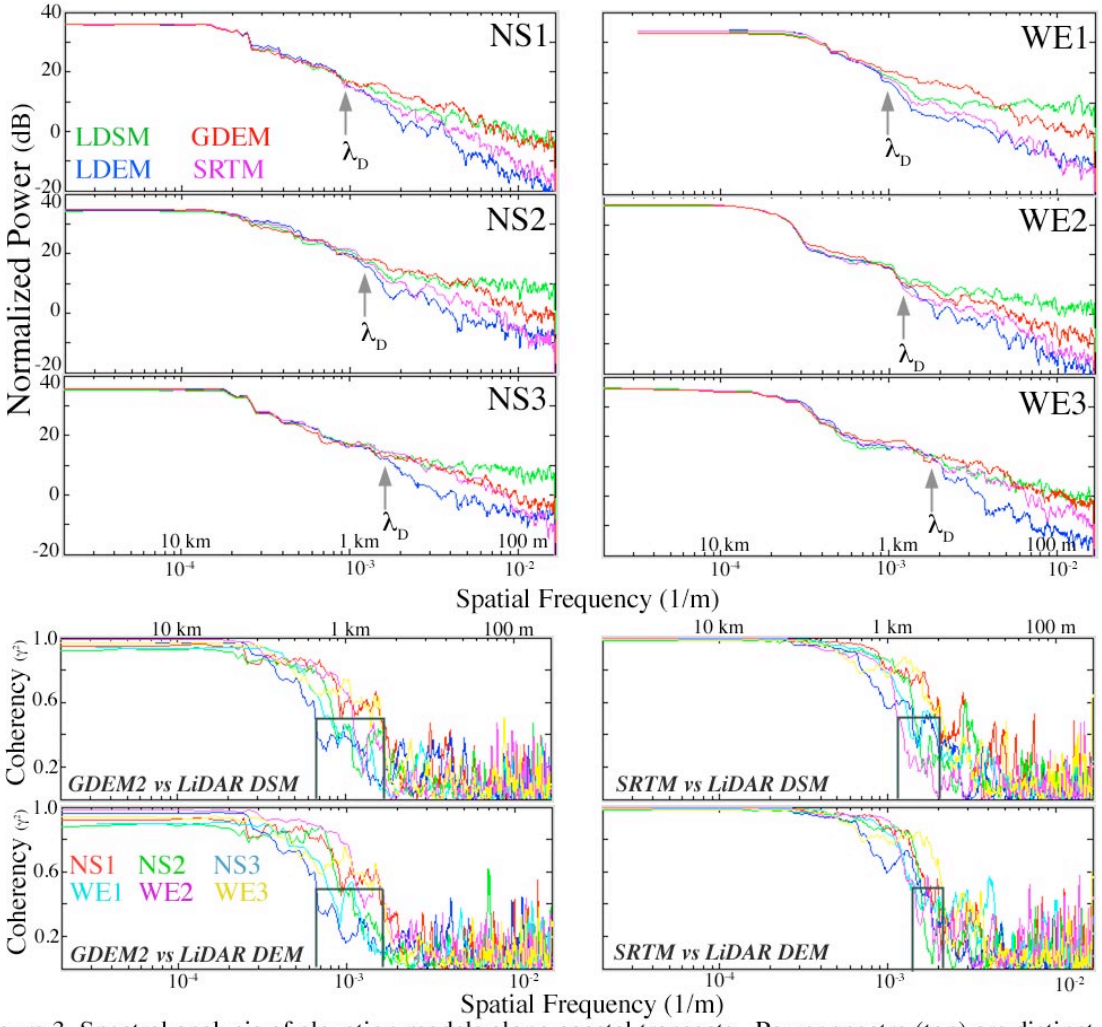

Figure 3 Spectral analysis of elevation models along coastal transects. Power spectra (top) are distinct for each of the 6 transects but show similar relationships among models for all transects. Spectral slopes at higher frequencies show rougher "tops" (DSM) and smoother "bottoms" (DEM). Spectra diverge for $\lambda < \sim 1$ km for all profiles but coherency begins to deteriorate at longer λ for GDEM2 than SRTM. The correlation scales λ_C where coherency reaches 0.5 (boxes) corresponds to the divergence scale λ_D where the roughness of land cover begins to exceed the roughness of the underlying terrain (arrows).

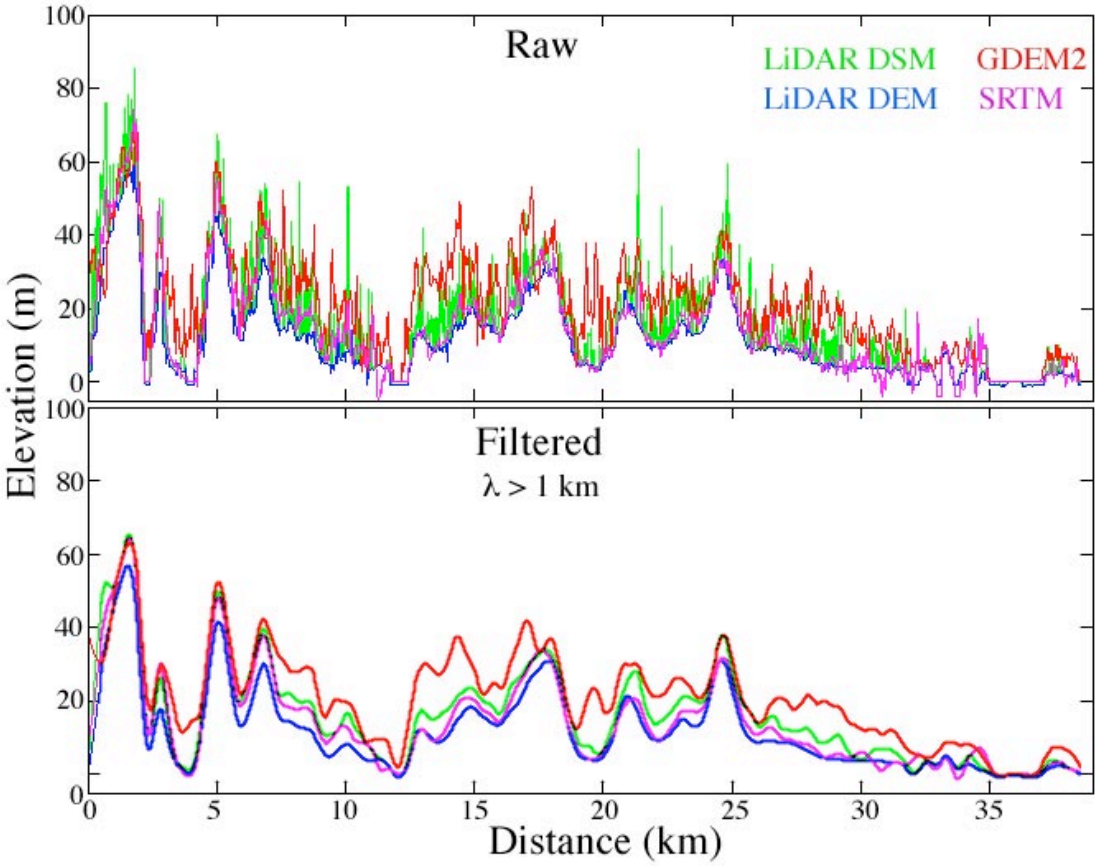

Figure 4 Comparison of raw and filtered elevations. Profile NS2 illustrates the effects of removal of incoherent short wavelength ($\lambda < 1$ km) variance from all 4 models. The LiDAR DEM and SRTM agree well in most areas. The GDEM2 model retains some short wavelength ($\lambda < \sim 3$ km) artifacts and positive bias relative to the LiDAR DSM. Note GDEM amplification of short λ features of DSM.