

Earth and Space Science



RESEARCH ARTICLE

10.1029/2020EA001390

Key Points:

- A data assimilation method with synthetic ship-borne Global Navigation Satellite System data is implemented in the Cascadia region
- Ship elevation data improves tsunami forecasts over a targeted domain
- Sensitivity tests explore density of ships traveling in and around the source regions required for accurate forecast

Supporting Information:

- Supporting Information S1
- Movie S1

Correspondence to:

M. J. Hossen,
md.hossen@colorado.edu

Citation:

Hossen, M. J., Mulia, I. E., Mencin, D., & Sheehan, A. F. (2021). Data assimilation for tsunami forecast with ship-borne GNSS data in the Cascadia subduction zone. *Earth and Space Science*, 8, e2020EA001390. <https://doi.org/10.1029/2020EA001390>

Received 11 AUG 2020

Accepted 13 FEB 2021

Data Assimilation for Tsunami Forecast With Ship-Borne GNSS Data in the Cascadia Subduction Zone

M. J. Hossen¹ , Iyan E. Mulia^{2,3} , David Mencin^{1,4}, and Anne F. Sheehan^{1,5}

¹Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA,

²Earthquake Research Institute, University of Tokyo, Tokyo, Japan, ³Prediction Science Laboratory, RIKEN Cluster for Pioneering Research, Kobe, Japan, ⁴UNAVCO, Boulder, CO, USA, ⁵Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA

Abstract An efficient and cost-effective near-field tsunami warning system is crucial for coastal communities. The existing tsunami forecasting system is based on offshore Deep-Ocean Assessment and Reporting of Tsunamis and Global Navigation Satellite System (GNSS) buoys which are not affordable for many countries. A potential cost-effective solution is to utilize position data from ships traveling in coastal and offshore regions. In this study, we examine the feasibility of using ship-borne GNSS data in tsunami forecasting. We carry out synthetic experiments by applying a data assimilation (DA) method with ship position (elevation and velocity) data. Our findings show that the DA method can recover the reference model with high accuracy if a dense network of ship elevation data is used. However, the use of ship velocity data alone is unable to recover the reference model. In addition, we carried out sensitivity studies of the DA method to the ship spatial distribution. We find that a 20 km gap between the ships works well in terms of accuracy and computational time for the example source model that we explored. The highest accuracy is obtained when data from a sufficient number of ships traveling in and around the tsunami source area are available.

1. Introduction

Data assimilation (DA) is a technique that incorporates observations into a numerical model to improve forecasts. Several DA methods have been applied in operational weather forecasting, first with Optimal Interpolation (Daley, 1993), and later with the adjoint model (Courtier et al., 1993; Thepaut & Courtier, 1991). The DA method has also been used to study air quality (Sandu et al., 2005) and soil moisture (Bouttier et al., 1993). It can be used to predict physical parameters of the oceans, such as tidal currents, sea-surface elevation, and temperatures (Johannessen et al., 2000). Recently, it has been implemented in tsunami science to forecast tsunamis without the need to first determine the tsunami source or initial sea-surface displacement. Maeda et al. (2015) first showed the possibility of the DA method for tsunami forecasting and in numerical experiments successfully applied it to a retrospective forecast of the 2011 Tohoku Earthquake tsunami. Subsequently, numerical experiments with the DA method have been applied to tsunami events with NOAA Deep-Ocean Assessment and Reporting of Tsunamis (DART) sea-floor pressure data, Cascadia Initiative sea-floor pressure gauge data, ship-borne Global Navigation Satellite System (GNSS) data, and airborne radar observations (e.g., Gusman et al., 2016; Maeda et al., 2015; Mulia et al., 2017, 2020; Sheehan et al., 2019; Wang et al., 2017). The DA method does not require precomputed Green's functions (GFs), though some studies have shown that precomputed GFs can reduce computation time for fixed observation sites (Wang et al., 2017). The DA method shows much promise for incorporation into tsunami forecasting systems where appropriate real-time data are available.

Several other methods exist for tsunami forecasting. For example, Tsushima et al. (2009) developed a method for near-field tsunami forecasting, called tsunami Forecasting based on Inversion for initial sea-Surface Height (tFISH). The tFISH algorithm estimates an initial sea-surface displacement in a possible tsunami source area and then synthesizes coastal tsunami waveforms from the initial sea-surface deformation and precomputed tsunami Green's functions. A real-time tsunami inundation forecasting scheme, called Short-term Inundation Forecast for Tsunamis (SIFT) (Gica et al., 2008; Tang et al., 2009; Titov et al., 2005), is now fully operational in NOAA's Tsunami Warning Center. The NOAA tsunami forecasting system gives a real-time inversion of the tsunami source by forcing a combination of precomputed unit tsunami propagations

© 2021. The Authors. Earth and Space Science published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

to fit the DART measurements through an inversion process and also provides an offshore forecast (Gica et al., 2008; Titov et al., 2005; Wei et al., 2013). The French Polynesian tsunami warning center has developed a new tsunami forecast tool, named MERIT, for tsunami warning based on a tsunami propagation numerical forward model (Jamelot & Reymond, 2015).

Most of the tsunami forecast systems are based on high quality data from the DART buoy system (Gonzalez et al., 1998) and/or GNSS buoys (Falck et al., 2010; Kato et al., 2000). Several such systems are now running in demonstration mode in the near-shore region of Japan and show promise for providing short lead-time warnings for locally generated tsunamis (Kato et al., 2005). However, like the existing DART network, a deep-water GNSS-buoy system would be costly to build and maintain. Many countries with high tsunami risk cannot afford such systems. Foster et al. (2012) proposed that the commercial shipping fleet represents a vast infrastructure of potential open ocean GNSS platforms that could provide good spatial coverage around most tsunamigenic source regions. GNSS-based ship height-positioning data has been further investigated for offshore tsunami detection and forecasting by Inazu et al. (2016).

Ship positions are continuously tracked globally through the Automatic Identification System (AIS) under International Maritime Organization regulations (International Maritime Organisation, 2002). The AIS system provides information including date, time, MMSI (Maritime Mobile Service Identity), ship type, ship name, ship position (latitude/longitude), speed over ground (SOG), course over ground (COG), and ship heading (HDG). The AIS data from near-shore vessels (<~100 km from coast) are received by coastal stations, and those from ships farther offshore can be received by low-Earth-orbit satellites. The HDG and COG data from AIS can be used to estimate tsunami current observation if there is a significant deviation between them due to tsunamis (Inazu et al., 2018).

Previous studies have shown that a high-precision GNSS attached to a seagoing ship would have approximately 10 cm vertical precision and thus could detect a tsunami height of 10 cm or greater (Foster et al., 2012; Inazu et al., 2016). For this reason, ordinary ship GNSS equipment with meter-order precision would need to be replaced with high-precision GNSS equipment with advanced processing in order to detect tsunamis (Roggenbuck & Reinking, 2019; Roggenbuck et al., 2014). With the affordability of geodetic GNSS systems, and the availability of modern satellite communication systems, it should be possible to build a cost-effective tsunami monitoring network by equipping the commercial shipping fleet with real-time GNSS systems. This would ideally use the existing AIS communication system that is already mandated to use (and paid for). As a consequence, many countries with high tsunami risk could build a low-cost and efficient tsunami forecasting system based on ship data.

Here, we examine the feasibility of using ship-borne GNSS data in DA modeling to forecast tsunamis in the Cascadia subduction zone region. The ship positions (latitude/longitude) that we consider are real, obtained from AIS broadcasts, but the tsunami signals (height and velocity) are all based on models. Ship AIS broadcasts do not currently include ship elevation information, and we find that such information, if available at sufficient accuracy, would be of great utility for tsunami forecasting.

2. Methodology

In this work, we apply a sequential DA based on optimal interpolation outlined in Maeda et al. (2015). In our case, the method is seeded with synthetic tsunami height and velocity data at ship positions. The synthetic tsunami height and velocity is simulated using a linear tsunami propagation model that solves the 2-D linear long wave (LLW) equations in spherical coordinates using a finite difference scheme (Maeda et al., 2015). The linear model is valid for water depths greater than 50 m (Shuto, 1991).

The governing equations for tsunamis propagating in the deep ocean are given as

$$\begin{aligned}\frac{\partial \eta}{\partial t} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= 0 \\ \frac{\partial u}{\partial t} + gh \frac{\partial \eta}{\partial x} &= 0 \\ \frac{\partial v}{\partial t} + gh \frac{\partial \eta}{\partial y} &= 0\end{aligned}\tag{1}$$

where $\eta = \eta(x, y, t)$ is the tsunami water height (hereafter elevation), $(u, v) = (u(x, y, t), v(x, y, t))$ are the vertically integrated horizontal velocity components of the tsunami in the x and y directions (hereafter, referred to tsunami velocity), $h = h(x, y, t)$ is the water depth and $g = 9.8 \text{ m/s}^2$ is the gravitational constant. The state vector for the DA system is $\mathbf{x}_n = (\eta(x, y, n\Delta t), u(x, y, n\Delta t), v(x, y, n\Delta t))$ at a discretized time steps $t = n\Delta t$.

There are two steps in DA system as outlined by Maeda et al. (2015). In the first step (forecast step), the tsunami wavefield is calculated by a forward numerical simulation,

$$\mathbf{x}_n^f = \mathcal{M}(\mathbf{x}_{n-1}^a) \quad (2)$$

given an estimated wavefield \mathbf{x}_{n-1}^a at $(n-1)$ th time step. The forward model $\mathcal{M}(\mathbf{x})$ is lead by the time evolution of \mathbf{x} governed by Equation 1 after time discretization. In the second step (analysis step), the wavefield is updated by integrating observations with simulated waveforms as

$$\mathbf{x}_n^a = \mathbf{x}_n^f + W(\mathbf{y}_n - H\mathbf{x}_n^f). \quad (3)$$

where \mathbf{y} is observation, H is observational operator, and W is the weighted matrix. Equation 3 is the optimal solution \mathbf{x}^a of the cost function at n th time step

$$\mathcal{J}_n(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_n^f)^T B^{-1} (\mathbf{x} - \mathbf{x}_n^f) + (\mathbf{y}_n - H\mathbf{x}_n^f)^T R^{-1} (\mathbf{y}_n - H\mathbf{x}_n^f) \quad (4)$$

The weighted matrix W is calculated as

$$W = BH^T (R + HBH^T)^{-1}$$

where B is the background error covariance matrix and R is observation error covariance matrix. The entries of the matrix B are calculated by a normalized Gaussian function defined as $e^{-d_{ij}^2/L^2}$, where $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ is the distance between two grid points and L is the radius of influence. Since observation errors are assumed to be uncorrelated between stations, the matrix R is a diagonal matrix whose diagonal entries are ρ^2 where ρ is the relative error of the root-mean square of the observational error compared to the numerical forecast error. We set $L = 23 \text{ km}$ and $\rho = 0.9$ following Mulia et al. (2017). The inverse matrix $(R + HBH^T)^{-1}$ is calculated using LU decomposition.

The observation operator H projects the model space into observational space. The operator is linear since we assume that the matrix H has a value of 1 at numerical grids where ships are located and all other components are set to zero. The observed tsunami waveform \mathbf{y} at the stations can be extracted from the true wavefield \mathbf{x}_n^T by a sparse linear observation matrix H with a noise field ϵ using the relation

$$\mathbf{y}_n = H\mathbf{x}_n^T + \epsilon \quad (5)$$

The ship locations at which the synthetic observations \mathbf{y} are generated are real locations obtained from the AIS ship navigation system. We extract synthetic tsunami waveforms at the ship locations at 30 s time intervals using output snapshots of the JAGURS tsunami propagation model. JAGURS (Baba et al., 2015) solves the linear and nonlinear shallow water equations in spherical coordinates. The code is parallelized and we run it on multiple processors to reduce the computational time. We choose the linear model since we select ship locations where water depth is greater than 50 m. We used two nested grids-60 arc sec in the domain (132°W–122°W, 39°N–51°N) and 20 arc sec in (130°W–123°W, 40°N–50°N) as the model requires a ratio of 3 between nested grids. The bathymetry data are resampled from the 30 arc-sec bathymetry grid of GEBCO-08 digital atlas (Intergovernmental Oceanographic Commission et al., 2014). We set time step $\Delta t = 1 \text{ s}$ and save snapshots of the water height and velocity fields at every 30 s. To produce more realistic tsunami waveforms at the ship locations, we add noise ϵ from actual ship-based observations as suggested by Inazu et al. (2016). The noise levels are less than 10 cm at frequencies of 0.01–0.1 cpm.

Table 1
Performance of Ship Distribution at Different Hours of Operation Offshore
Cascadia Subduction Zone

Cases	Hours	No of ships	Elevation (%)	Velocity (%)	Elevcity (%)
Fixed	10–11	74	89	62	91
	11–12	58	92	67	84
	12–13	55	95	69	90
	13–14	42	93	68	83
	14–15	46	91	61	92
	15–16	40	92	64	93
Moving	10–11	74	89	62	90
	11–12	58	91	65	84
	12–13	55	95	74	87
	13–14	42	92	71	79
	14–15	46	91	63	90
	15–16	40	90	66	89

Notes. Accuracy calculated using Equation 6. Elevcity refers to models where both elevation and velocity data are used.

To evaluate the performance of the DA system, we used a formula proposed by Tsushima et al. (2009) to calculate accuracy by comparing maximum tsunami height of the reference model with the assimilated model. The formula is defined as

$$\text{Accuracy}(\%) = \left\{ 1 - \frac{\sum_{i=1}^N (w_i^s - w_i^r)^2}{\sum_{i=1}^N (w_i^r)^2} \right\} \times 100\% \quad (6)$$

where N represents the number of coastal points, and w_i^r and w_i^s represent the reference and assimilated maximum wave height at the i th coastal point, respectively.

3. Results

We carried out DA numerical experiments for tsunami forecasting with both fixed and moving ship locations and using three different types of ship position data: elevation, velocity, and elevation plus velocity (referred to as elevcity in Table 1). Previous studies (Inazu et al., 2016; Mulia et al., 2017) only considered fixed location because ship speed is much smaller than that of the tsunami, especially in deep water. For fixed ship location, we used the ship location directly without modification

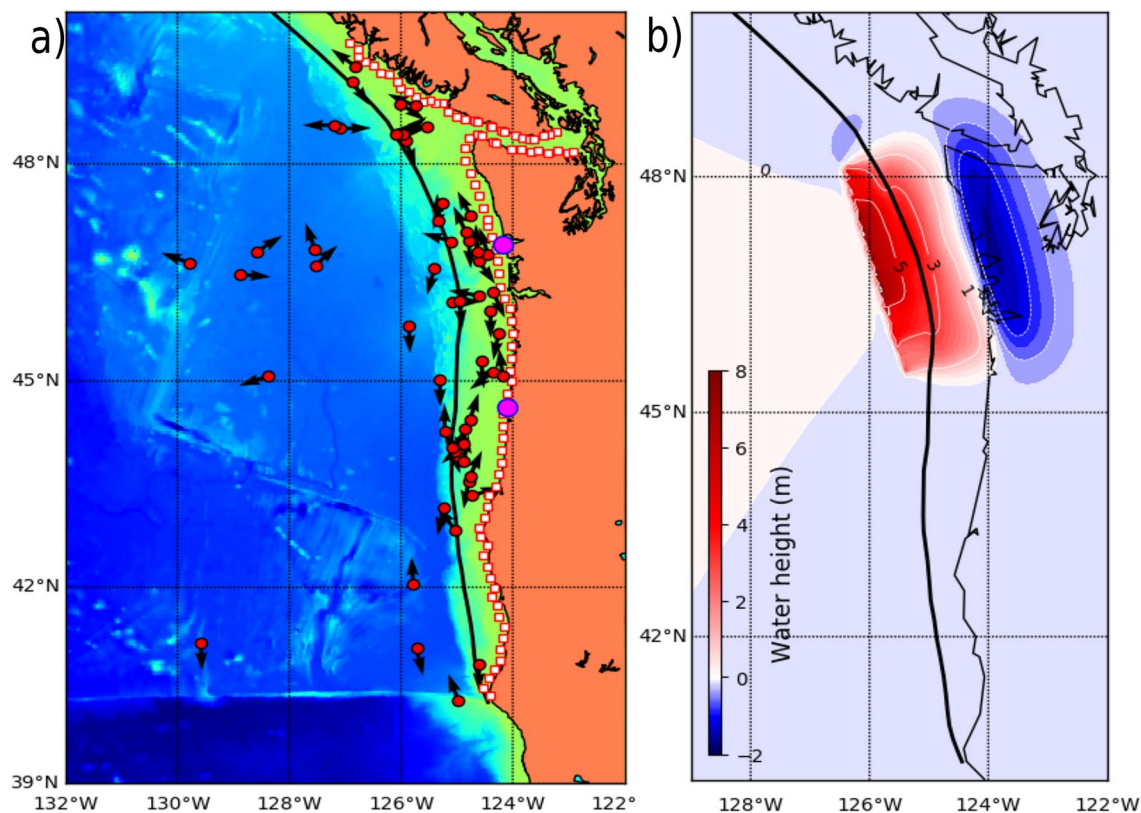


Figure 1. (a) Distribution of ships offshore Cascadia subduction zone 12–13 UTC, June 16, 2019, used in DA study. Red dots show ship locations and black arrows show ship heading. Magenta circles show Westport, Washington, and Newport, Oregon, the points at which we extract the waveform for comparison. White squares show coastal points at which we extract maximum water height from reference and assimilated models. Black solid line shows trench axis. Bathymetry from GEBCO-08 digital atlas is used in this plot. (b) Sea-surface displacement for Cascadia tsunami source scenario used in this study.

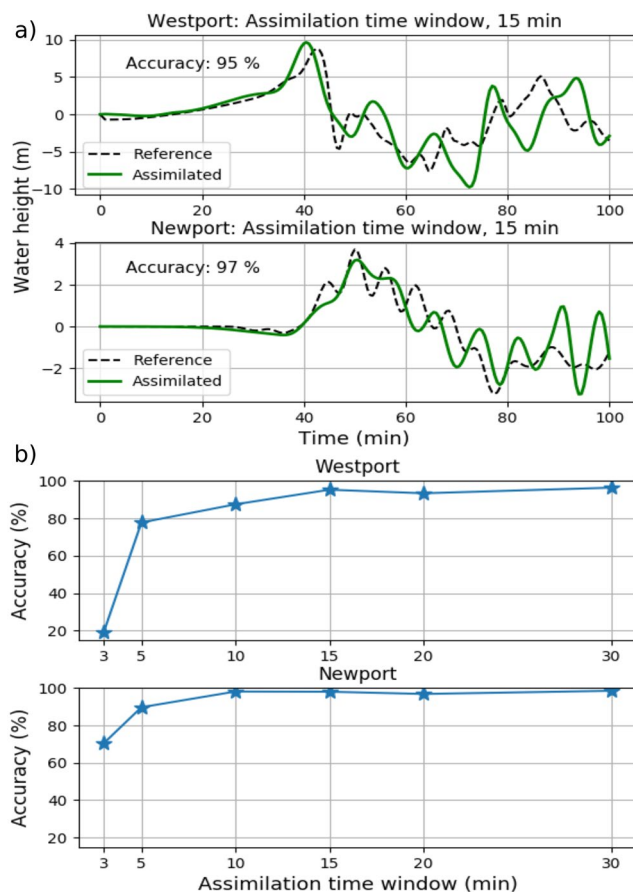


Figure 2. Upper panels (a) show the waveform fits (reference and DA model with elevation data) at Westport and Newport using 15 min assimilation window; lower panels (b) show the accuracy (%) obtained using different assimilation time window. Shorter assimilations result in more warning time, but lowered forecast accuracy. The initial value of each time window is 0 min.

Newport, Oregon (water depth 19 m), shown in Figure 1a. The waves corresponding to the Cascadia tsunami source scenario that we examine in this study (Figure 1b) reach Westport, WA in about 35 min and reach Newport, OR in about 40 min (Figure 2a). Thus, an assimilation time window of 30 min provides less than 10 min warning to these locations. We found that 15 min of data assimilated in the DA process provides 95% accuracy at Westport and 97% at Newport (Figure 2b). We have also included a movie (Movie S1) that shows the general pattern of a realistic tsunami wavefield begins to emerge at 15 min of DA. Hence, we choose 15 min time window for all computations. The results presented in Figure 3 show the maximum water height distribution at 10, 30, and 60 min for the reference model and for our fixed ship DA models constructed using ship elevation, ship velocity, and both elevation and velocity data. The assimilated wavefield at 10 min is obtained by assimilating 10 min data and the others at 30 or 60 min are obtained with 15 min data as our assimilation window is 15 min.

In Figures 4a–4c, we compare maximum water-height along the coast forecast by the DA models with that from the reference model. The accuracy, calculated with Equation 6, is 95% for the DA forecast using ship elevation data, 69% for the DA forecast using ship velocity data, and 90% for the DA model that uses both ship elevation and velocity data. We show the misfit between the reference and assimilated model in (Figures 4d–4f). When moving ship locations are considered, the performance of the DA method remains almost the same, Figures 5d–5f. The accuracy for elevation, velocity, and elevation + velocity are 95%, 74%,

(Figure 1a). For the case of moving ships, we used ship speed (assumed to be 6 m/s) and heading to create ship locations that move at each step of the simulation. To estimate moving ship locations, we used the formula, suggested by Inazu et al. (2018),

$$v = s \sin(\theta_c - \theta_h)$$

where s is speed over ground, θ_c is course over ground, the direction of true ship movement, and θ_h is heading, the direction of the ship heading. We use real ship locations from the Cascadia region obtained from satellite and terrestrial AIS broadcast data collected by Spire Global, Inc. (<https://spire.com>), however, the tsunami produced ship height and velocity variations are simulated.

We create synthetic observations (elevation, velocity) at the ship locations using output from JAGURS run with the tsunami source shown in Figure 1b. To create the source, we use a hypothetical earthquake slip model with parameters suggested by Mai and Beroza (2002) (see Supplementary Information for details of the slip model). Snapshots of JAGURS output are taken at 30 s intervals to produce synthetic observation data with 30 s sampling for the DA process. We apply the DA method by assimilating the synthetic data at every 30 s over a assimilation time window. We use the linear shallow water model given in Equation 1 over the computational domain shown in Figure 1a. The resolution of the domain we use is 60 arc-sec. For the fixed ship location case, we calculate the weight matrix W once at the beginning of assimilation and use it for the entire assimilation window. For the moving ship case, the matrix W is calculated at every assimilation time step since ship locations are changing in time. The assimilation process is very rapid. To complete the computation, it requires about 2.5 min for the fixed case and about 5.5 min for the moving case. However, the time to produce the tsunami forecast mostly depends on the assimilation time windows, that is, the time when observed data is available after the tsunami event.

To test the length of assimilation time window needed, we assimilated ship elevation data (fixed location) with different time windows from 3 to 30 min and calculated the accuracy by extracting tsunami signals at two points near the coast: Westport, Washington (water depth 11 m), and

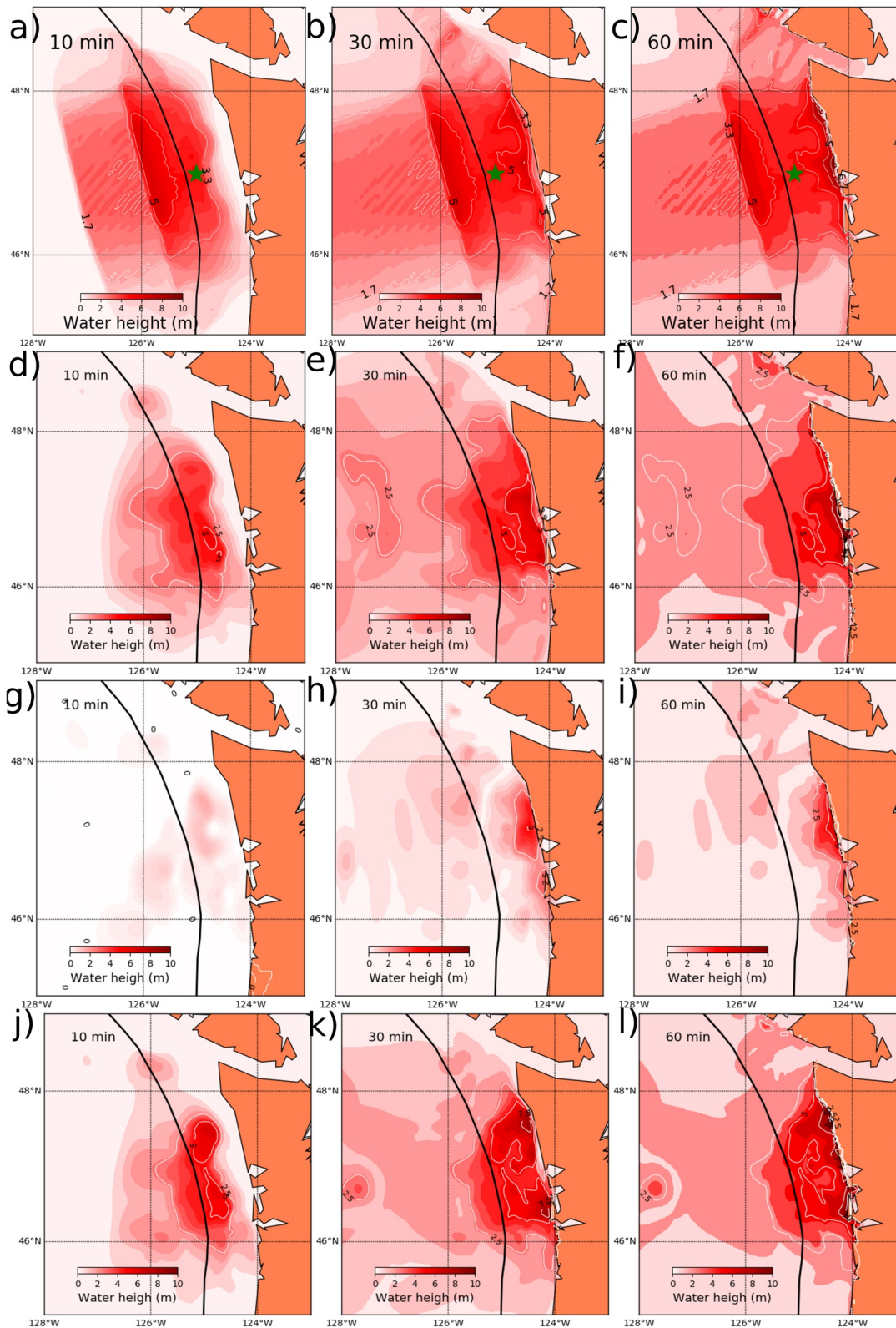


Figure 3. Maximum water height for the reference model and the data assimilated wavefield for three different scenarios. First row shows maximum water height for reference tsunami model (a, b, c), second row shows DA model constructed from ship elevation data (d, e, f), third row shows DA model from ship velocity data (g, h, i), and the fourth row shows the DA model constructed using both ship elevation and velocity (j, k, l). Left most columns show the wavefield at 10 min, middle columns at 30 min, and right most at 60 min.

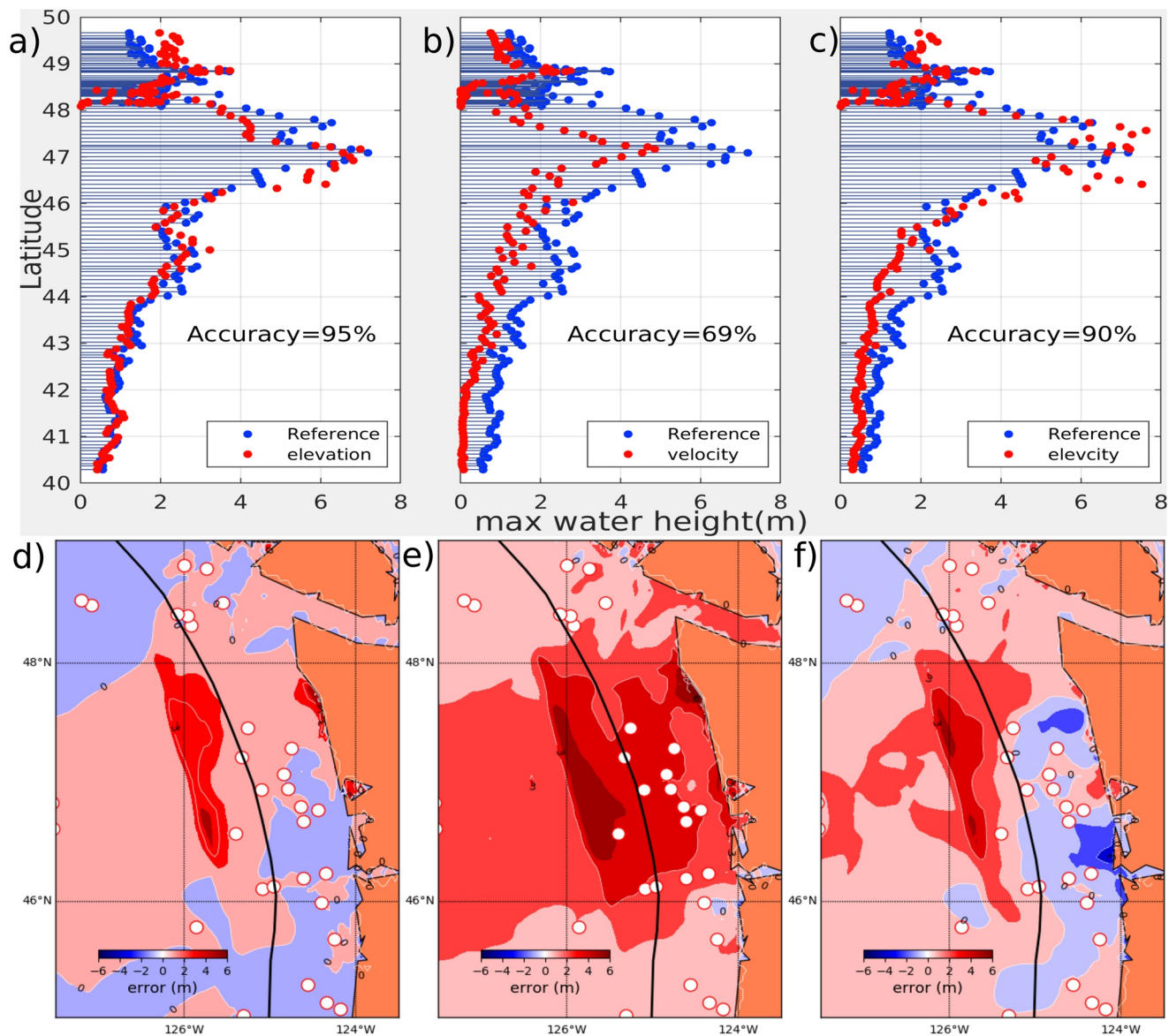


Figure 4. Comparison of tsunami height from reference tsunami model and from the DA models constructed using synthetic tsunami data at 55 fixed ship locations (as shown in Figure 1a). Top plots show reference (gray) and DA model (red) maximum tsunami height at points along the coast (coastal locations shown in Figure 1a). Bottom plots show map view of maximum tsunami height misfit between the reference tsunami model and the DA models. Left plots are for DA simulation using ship elevation data (a, d), middle plots for model using ship velocity (b, e), and right plots for model using combination of elevation and velocity data (c, f).

and 87%, respectively; the performance with elevation data remains the same; velocity data improves the performance by 5%; but the performance with elevation plus velocity data decreases by 3%.

Interestingly, in both fixed and moving cases, the use of only ship elevation data shows excellent performance (95% accuracy). In contrast, velocity data performs worst among the three; it underestimates the reference model, whereas the combination of velocity and elevation overestimates the height at some points. In both cases, the DA method with the three different datasets (elevation, velocity, and combination) cannot recover the maximum height of the tsunami source over the seaward side of the trench line; the error is very large for all cases (Figures 4 and 5). It could be due to lacking sufficient numbers of observations or the inability of the DA method to recover the sharp edge of the source model using few stations. However, the error is near zero in the region between the trench and coastline when elevation data is used (Figures 4a and 5a).

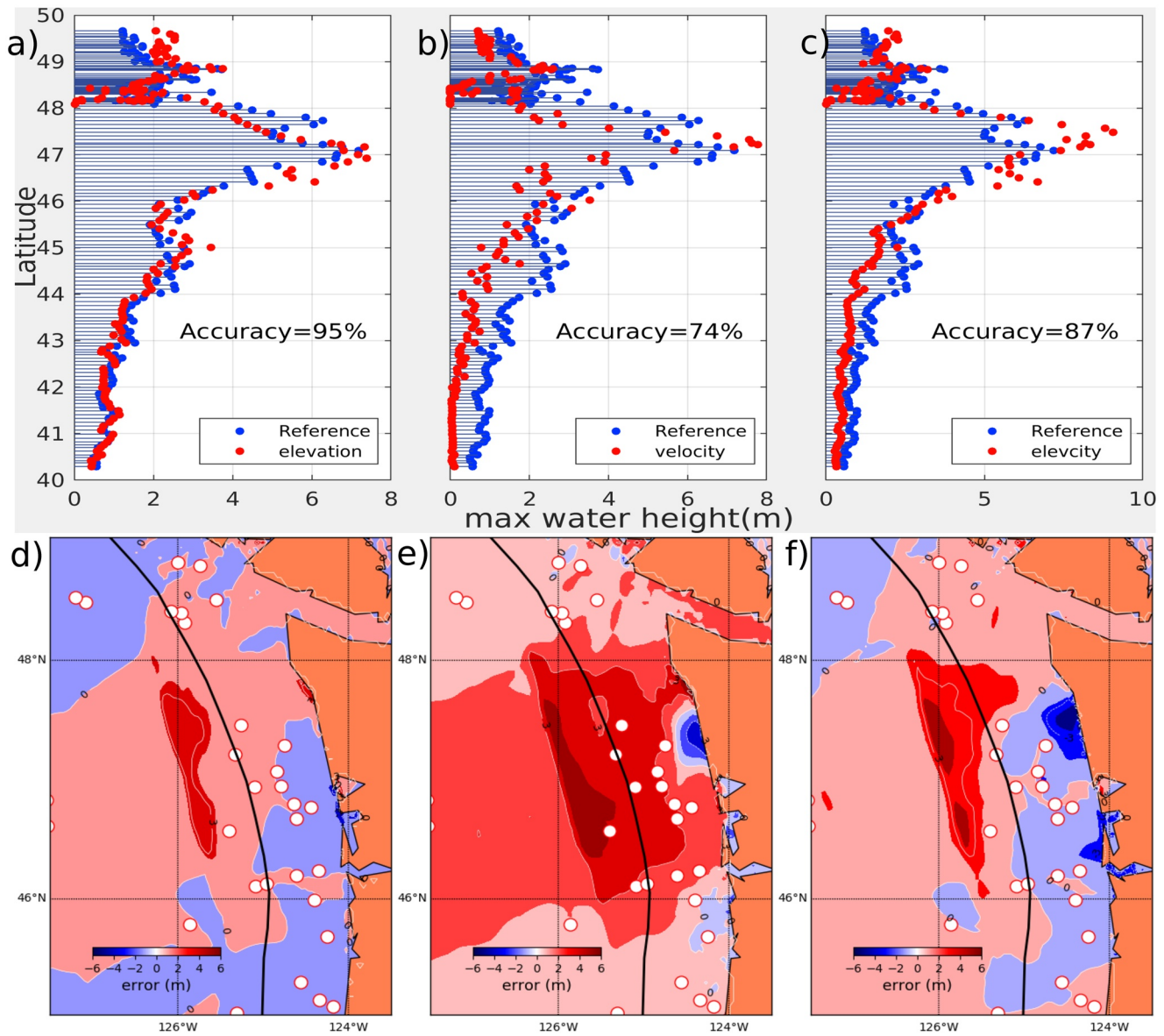


Figure 5. Same as Figure 4, but with moving ship locations. Left plots are for DA simulation using ship elevation data (a, d), middle plots for model using ship velocity (b, e), and right plots for model using combination of elevation and velocity data (c, f).

4. Discussion

Our findings show that DA simulations using fixed and moving ship data provide nearly the same results. Given the similar performance, we favor the use of fixed ship locations because of the reduced computational cost. We also find that ship elevation data provides superior results to ship velocity data in the DA simulations. Ship elevation data is currently not included in the AIS broadcasts since ships are assumed to be at sea level, but could be a valuable addition. While marine ship traffic tends to follow common routes, the number and spatial distribution of ships varies with time, thus we next examine examples of the effect of ship distributions on the DA tsunami forecast models. These examples and sensitivity tests are illustrative rather than exhaustive.

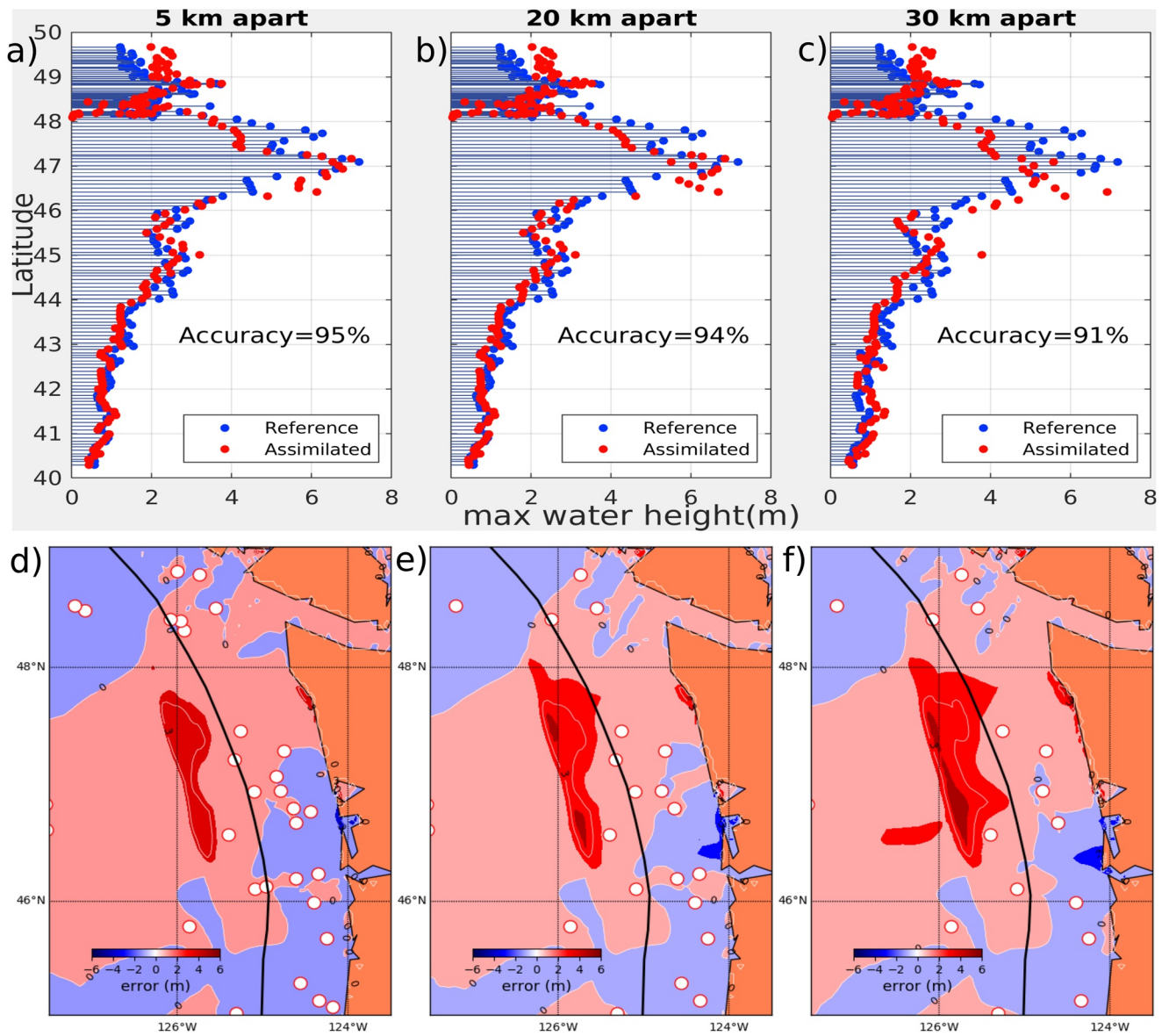


Figure 6. Accuracy of DA simulated model with different minimum spacing between ships. Ship elevation data, but not velocity, is used in these simulations. Left panels show the result with a 5 km gap between ships (a, d), middle panels (b, e) 20 km, and right panels (c, f) 30 km gap between ships.

4.1. Sensitivity of DA Method to Ship Distribution at Different Times

We first considered ships traveling at different hours of operation on a given day. We examined ship distributions at 10:00, 11:00, 12:00, 13:00, 14:00, and 15:00 UTC on June 16, 2019. The number of ships meeting our latitude-longitude and water depth criteria ranged from 40 to 74 during these hours. DA simulations were run with each of these ship distributions, and results are given in Table 1. The results for the different distributions are very similar, and are not necessarily improved when a larger number of ships are available. The elevation data provide the best overall accuracy, the velocity data the worst, and the combination of elevation plus velocity is similar to using elevation data alone.

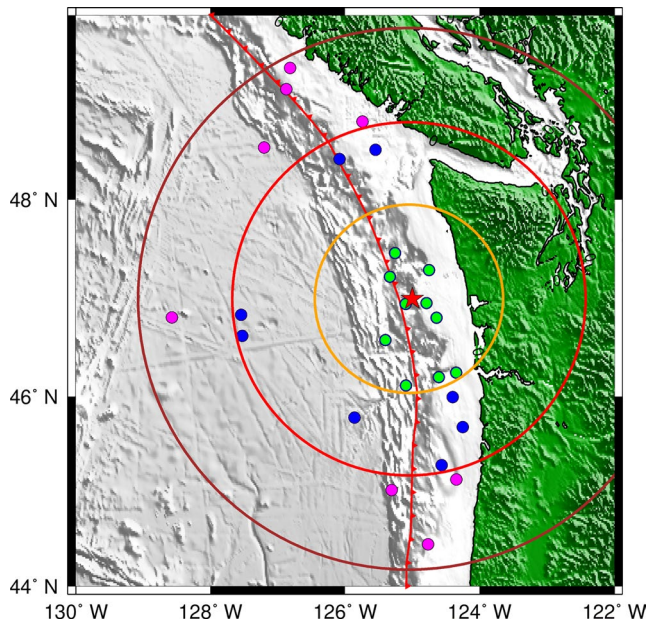


Figure 7. Ship distribution within different coverage radii centered on tsunami source region (red star). Radii of 100 km (orange circle), 200 km (red circle), and 300 km (brown circle) considered. Ships within 100 km shown as green dots, ships from 100 to 200 km by blue dots, and ships from 200 to 300 km by magenta dots.

height in the reference model (Figure 7). We first considered ships located within 100 km of the source region. There are only 10 ships within the radius, and they are well distributed in the region where maximum tsunami energy is propagated, 92% accuracy is attained (Figure 8a). Next, we considered radius of coverage of 200 (18 ships) and 300 km (26 ships), respectively; the performance is the same (94% for radii 200 and 300 km). The improvement in performance is likely due to number of ships and azimuthal coverage. Ships within 200 km provides 94% accuracy (Figures 8c and 8f) which is the same as the result with wider coverage (41 ships in the gap of 20 km) in Figure 6b. From these findings, we infer that the area of ship coverage should be wide enough to cover at least the initial source region.

5. Conclusion

We applied a DA method to synthetic ship elevation and velocity data to explore their utility for tsunami forecasts in the Cascadia region. We considered three types of ship data: elevation, velocity, and elevation plus velocity, all extracted from numerical simulations with an assumed tsunami source. We find that the DA method produces similar results for fixed and moving ship locations; fixed locations are preferred to lower computational cost. Among the three types of ship data considered, elevation data provides the best coastal tsunami forecast accuracy. We also considered different numbers of ship observations with varied spatial density. We found that elevation data always outperforms the velocity data irrespective of ship distribution. Our sensitivity study shows that a minimum of 20 km gap between ships in regions with high ship density is sufficient to recover the reference model, and that good coverage of observations in and around the tsunami source region is crucial.

Our numerical simulations use real ship lateral (latitude and longitude, but not elevation) positions obtained from AIS broadcast data, but the ship elevation and velocity are modeled. Currently ship elevation (vertical position) is not provided as part of the AIS system. While this work demonstrates the great utility of such data should it become available, more work is needed to characterize and improve the precision of ship-based GNSS systems and determine ways to obtain such data in real time for it to be of use in tsunami forecasting. Once the issues are resolved, we will apply the DA process to forecast tsunami by extracting tsunami signals from GNSS data in real-time.

4.2. Sensitivity of DA Method to Ship Spatial Density

We next carried out experiments with ships at different minimum separations, imposing minimum separations between ships of 5 (55 ships), 10 (52 ships), 20 (41 ships), and 30 (32 ships) km. Use of fewer ships can result in faster computational time, but potentially at the expense of forecast accuracy. For these experiments, we used elevation data of the ships traveling between 12 and 13 UTC on June 16, 2019. The distances between adjacent ships are not uniform; in the ship position data that we obtained from Spire we find that ships are closely spaced between the trench and the coastline, and are more sparse seaward of the trench line. It is reasonable to assume that all ships in low ship density areas should be considered, and in areas with many ships, it is likely that we do not need to consider every ship. We find that 94% coastal forecast accuracy can be obtained with a gap up to 20 km between the ships (Figure 6b); the accuracy is decreased by 3% if ship spacing is increased to 30 km (Figure 6c). The error distribution in Figure 6e shows low misfit in the region between trench and coast line when using a gap of 20 km. Thus, for this example, a gap between ships of 20 km works well in terms of accuracy and computational time.

4.3. Sensitivity of DA Method to Radius of Coverage of Ships

In this section, we examine sensitivity of the DA method to radius of coverage of ships. We consider ship locations of 20 km apart based on the output of the previous section. We chose ships located within radii (100, 200, and 300 km) of coverage area from the location of maximum water

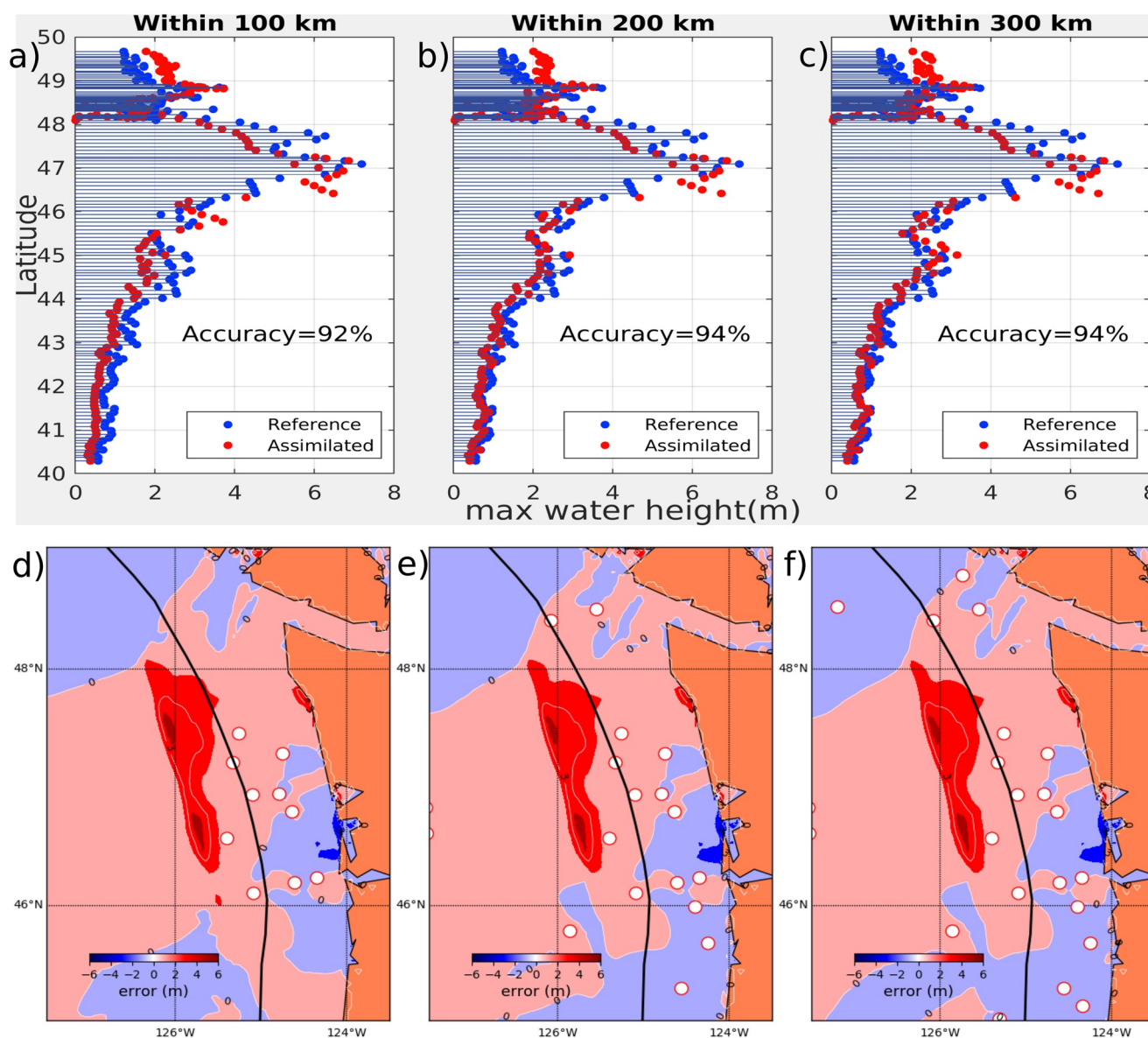


Figure 8. Coverage of the ship distribution within a distance. Upper panels (a, b, c) show comparison of reference and assimilated model extracted at the coastal point and lower panels (d, e, f) show difference between reference and assimilated wavefield. Left panels (a, d) show the result with ships within 100 km (10 ships) from the location of maximum water height, middle panels (b, e) show within 200 km (18 ships) and right panels (c, f) show within 300 km (26 ships) radius of coverage.

Acknowledgments

The first author thanks the Cooperative Institute for Research in Environmental Sciences (CIRES) for supporting this research through a postdoctoral visiting fellowship. This work was supported in part by NSF Award Number ICER 1855090. We thank the company Spire Global, Inc. (www.spire.com) for providing ship location data. This work utilized the RMACC Summit Supercomputer, which is a joint effort of the University of Colorado Boulder, and Colorado State University.

Data Availability Statement

The raw data is available in the link: <https://doi.org/10.6084/m9.figshare.12816488.v5>

References

- Baba, T., Takahashi, N., Kaneda, Y., Ando, K., Matsuoka, D., & Kato, T. (2015). Parallel implementation of dispersive tsunami wave modeling with a nesting algorithm for the 2011 Tohoku tsunami. *Pure and Applied Geophysics*, 172(12), 3455–3472.
- Bouttier, F., Mahfouf, J., & Noilhan, J. (1993). Sequential assimilation of soil moisture from atmospheric low-level parameters. Part II: Implementation in a mesoscale model. *Journal of Applied Meteorology*, 32(8), 1352–1364.
- Courtier, P., Derber, J., Errico, R., Louis, J.-F., & Vukićević, T. (1993). Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus A: Dynamic Meteorology and Oceanography*, 45(5), 342–357.
- Daley, R. (1993). *Atmospheric Data Analysis*: Cambridge Atmospheric and Space Science Series, Cambridge: Cambridge University Press.

- Falck, C., Ramatschi, M., Subarya, C., Bartsch, M., Merx, A., Hoeberechts, J., & Schmidt, G. (2010). Near real-time GPS applications for tsunami early warning systems. *Natural Hazards and Earth System Sciences*, 10(2), 181–189.
- Foster, J. H., Brooks, B. A., Wang, D., Carter, G. S., & Merrifield, M. A. (2012). Improving tsunami warning using commercial ships. *Geophysical Research Letters*, 39(9). <https://doi.org/10.1029/2012GL051367>
- Gica, E., Spillane, M., Titov, V., Chamberlin, C., & Newman, J. (2008). *Development of the forecast propagation database for NOAA's Short-term Inundation Forecast for Tsunamis (SIFT)* (Tech. rep., NOAA Tech. Memo. OAR PMEL-139, 89 pp). Retrieved from <https://repository.library.noaa.gov/view/noaa/11079>
- Gonzalez, F., Milburn, H., Bernard, E., & Newman, J. (1998). Deep-ocean assessment and reporting of tsunamis (DART): Brief overview and status report (19). Proceedings of the International Workshop on Tsunami Disaster Mitigation. Retrieved from <https://www.ndbc.noaa.gov/dart/brief.shtml>
- Gusman, A. R., Sheehan, A. F., Satake, K., Heidarzadeh, M., Mulia, I. E., & Maeda, T. (2016). Tsunami data assimilation of Cascadia seafloor pressure gauge records from the 2012 Haida Gwaii earthquake. *Geophysical Research Letters*, 43(9), 4189–4196. <https://doi.org/10.1002/2016GL068368>
- Inazu, D., Ikeya, T., Waseda, T., Hibiya, T., & Shigihara, Y. (2018). Measuring offshore tsunami currents using ship navigation records. *Progress in Earth and Planetary Science*, 5(1), 38.
- Inazu, D., Waseda, T., Hibiya, T., & Ohta, Y. (2016). Assessment of GNSS-based height data of multiple ships for measuring and forecasting great tsunamis. *Geoscience Letters*, 3(1), 25.
- Intergovernmental Oceanographic Commission, IHO, & BODC. (2014). *Centenary edition of the GEBCO digital atlas, published on CD-ROM on behalf of the Intergovernmental Oceanographic Commission and the International Hydrographic Organization as part of the General Bathymetric Chart of the Oceans*. Liverpool, UK: British Oceanographic Data Centre. Retrieved from <https://www.gebco.net>
- International Maritime Organisation. (2002). *Guidelines for the onboard operational use of shipborne Automatic Identification Systems (AIS)*. Assembly 22nd session, resolution 917. Retrieved from http://rise.odessa.ua/texts/A917_22e.php3
- Jamelot, A., & Reymond, D. (2015). New tsunami forecast tools for the French Polynesia tsunami warning system Part II: Numerical modelling and tsunami height estimation. *Pure and Applied Geophysics*, 172(3–4), 805–819.
- Johannessen, O. M., Sandven, S., Jenkins, A. D., Durand, D., Pettersson, L. H., Espedal, H., et al. (2000). Satellite earth observation in operational oceanography. *Coastal Engineering*, 41(1–3), 155–176.
- Kato, T., Terada, Y., Ito, K., Hattori, R., Abe, T., Miyake, T., et al. (2005). Tsunami due to the 2004 September 5th off the Kii peninsula earthquake, Japan, recorded by a new GPS buoy. *Earth Planets and Space*, 57(4), 297–301.
- Kato, T., Terada, Y., Kinoshita, M., Kakimoto, H., Isshiki, H., Matsuishi, M., et al. (2000). Real-time observation of tsunami by RTK-GPS. *Earth, Planets and Space*, 52(10), 841–845.
- Maeda, T., Obara, K., Shinohara, M., Kanazawa, T., & Uehira, K. (2015). Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters*, 42(19), 7923–7932.
- Mai, P. M., & Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research*, 107(B11), ESE–10. <https://doi.org/10.1029/2001JB000588>
- Mulia, I. E., Hirobe, T., Inazu, D., Endoh, T., Niwa, Y., Gusman, A. R., et al. (2020). Advanced tsunami detection and forecasting by radar on unconventional airborne observing platforms. *Scientific Reports*, 10(1), 1–10.
- Mulia, I. E., Inazu, D., Waseda, T., & Gusman, A. R. (2017). Preparing for the future Nankai Trough tsunami: A data assimilation and inversion analysis from various observational systems. *Journal of Geophysical Research: Oceans*, 122(10), 7924–7937. <https://doi.org/10.1002/2017JC012695>
- Roggenbuck, O., & Reinking, J. (2019). Sea surface heights retrieval from ship-based measurements assisted by GNSS signal reflections. *Marine Geodesy*, 42(1), 1–24.
- Roggenbuck, O., Reinking, J., & Härting, A. (2014). Oceanwide precise determination of sea surface height from in-situ measurements on cargo ships. *Marine Geodesy*, 37(1), 77–96.
- Sandu, A., Daescu, D. N., Carmichael, G. R., & Chai, T. (2005). Adjoint sensitivity analysis of regional air quality models. *Journal of Computational Physics*, 204(1), 222–252.
- Sheehan, A. F., Gusman, A. R., & Satake, K. (2019). Improving forecast accuracy with tsunami data assimilation: The 2009 Dusky Sound, New Zealand, tsunami. *Journal of Geophysical Research: Solid Earth*, 124(1), 566–577. <https://doi.org/10.1029/2018JB016575>
- Shuto, N. (1991). Numerical simulation of tsunamis: Its present and near future. *Natural Hazards*, 4(2–3), 171–191.
- Tang, L., Titov, V., & Chamberlin, C. (2009). Development, testing, and applications of site-specific tsunami inundation models for real-time forecasting. *Journal of Geophysical Research*, 114(C12). <https://doi.org/10.1029/2009JC005476>
- Thepaut, J.-N., & Courtier, P. (1991). Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Quarterly Journal of the Royal Meteorological Society*, 117(502), 1225–1254.
- Titov, V. V., Gonzalez, F. I., Bernard, E., Eble, M. C., Mofjeld, H. O., Newman, J. C., & Venturato, A. J. (2005). Real-time tsunami forecasting: Challenges and solutions. *Natural Hazards*, 35(1), 35–41.
- Tsushima, H., Hino, R., Fujimoto, H., Tanioka, Y., & Imamura, F. (2009). Near-field tsunami forecasting from cabled ocean bottom pressure data. *Journal of Geophysical Research*, 114(B6). <https://doi.org/10.1029/2008JB005988>
- Wang, Y., Satake, K., Maeda, T., & Gusman, A. R. (2017). Green's function-based tsunami data assimilation: A fast data assimilation approach toward tsunami early warning. *Geophysical Research Letters*, 44(20), 10–282. <https://doi.org/10.1002/2017GL075307>
- Wei, Y., Chamberlin, C., Titov, V. V., Tang, L., & Bernard, E. N. (2013). Modeling of the 2011 Japan tsunami: Lessons for near-field forecast. *Pure and Applied Geophysics*, 170(6–8), 1309–1331.