Geophysical Journal International

Geophys. J. Int. (2013) **195,** 491–503
Advance Access publication 2013 July 4

doi: 10.1093/gji/ggt228

A framework for estimating stratospheric wind speeds from unknown sources and application to the 2010 December 25 bolide

Stephen J. Arrowsmith,[1] Omar Marcillo[1] and Douglas P. Drob[2]

[1] *Geophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, USA. E-mail: sarrowsmith@gmail.com*
[2] *Space Sciences Division, US Naval Research Laboratory, Washington, DC, USA*

Accepted 2013 June 4. Received 2013 June 3; in original form 2012 July 26

SUMMARY

We present a methodology for infrasonic remote sensing of winds in the stratosphere that does not require discrete ground-truth events. Our method uses measured time delays between arrays of sensors to provide group velocities (referred to here as celerities) and then minimizes the difference between observed and predicted celerities by perturbing an initial atmospheric specification. Because we focus on interarray propagation effects, it is not necessary to simulate the full propagation path from source to receiver. This feature allows us to use a relatively simple forward model that is applicable over short-regional distances. By focusing on stratospheric returns, we show that our non-linear inversion scheme converges much better if the starting model contains a strong stratospheric duct. Using the Horizontal Wind Model (HWM)/Mass Spectrometer Incoherent Scatter (MSISE) empirical climatology as a starting model, we demonstrate that the inversion scheme is robust to large uncertainties in backazimuth, but that uncertainties in the measured trace velocity and celerity require the use of prior constraints to ensure suitable convergence. The inversion of synthetic data, using realistic estimates of measurement error, shows that our scheme will nevertheless improve upon a starting model under most scenarios. The inversion scheme is applied to infrasound data recorded from a large event on 2010 December 25, which is presumed to be a bolide, using data from a nine-element infrasound network in Utah. We show that our recorded data require a stronger zonal wind speed in the stratosphere than is present in the HWM profile, and are more consistent with the Ground-to-Space (G2S) profile.

Key words: Inverse theory; Seismic monitoring and test-ban treaty verification; Wave propagation.

1 INTRODUCTION

Low-frequency sound, or infrasound, can propagate long distances through the atmosphere by refracting in the troposphere, stratosphere or thermosphere. Compared with the seismic and hydroacoustic communities, relatively little work has been done on using these waves to invert for medium properties along the path. State-of-the-art 4-D models for the atmosphere that are used in simulating infrasound propagation, such as the Ground-to-Space (G2S) model (Drob *et al.* 2003) incorporate data from independent ground-based measurements (e.g. radiosondes) and satellite measurements through numerical weather prediction data analysis fields available from, for example, NOAA and NASA operational numerical weather prediction centres. Recently, Drob *et al.* (2010a) formulated an algorithm for remote sensing of atmospheric winds by observing discrete ground-truth events with regional infrasound arrays. Another framework that utilizes discrete ground-truth events to invert for perturbations to wind speed profiles was recently proposed by Lalande *et al.* (2012). While these frameworks, which have been developed and tested on noise-free synthetic data, represent an important advance, the ability to accomplish the same goal using either discrete unknown sources or coherent infrasound noise would have more widespread utility. This paper attempts to extend the algorithms of Drob *et al.* (2010a) and Lalande *et al.* (2012) for application to unknown sources recorded on a regional infrasound array. We further explore the effects of noise on the stability of the inversion technique. Our results demonstrate that it is possible to invert for winds in the stratosphere using unknown sources recorded on regional infrasound networks, if one makes some reasonable assumptions.

The propagation of infrasound is affected by temperatures and winds, which vary as a function of altitude, spatial location and time. As discussed in Drob *et al.* (2010a), the geophysical uncertainties in state-of-the-art profiles of adiabatic sound velocity, which is derived from temperature, are $\sim$1–3 m s^{-1} whereas uncertainties in the wind components are on the order of 25 m s^{-1} or more above 50 km.

Following Drob *et al.* (2010a) and Lalande *et al.* (2012) this paper assumes that the adiabatic sound speed is known and inverts for perturbations to the zonal wind profile. This assumption neglects the fact that the temperature and wind are coupled through the thermal wind equation that relates temperature and wind gradients (Beer 1974). Wind fields in the G2S model in the stratosphere are derived exclusively from the geophysical fluid dynamic balance of the global pressure fields, which are themselves derived from temperature soundings and are not measured directly. However, even if temperature soundings can lead to wind estimates through the balance equation, those estimates are not expected to be as accurate as estimates obtained from direct measurement of wind. This is especially true since numerical differentiation worsens the effect of noisy observations (Chartrand 2011). Since the thermal wind equation relates wind to the horizontal pressure gradient, this is a major cause of error in the estimation of the wind field as discussed by Kalnay *et al.* (1985). Based on this principle, there is a significant body of work on direct satellite measurement of winds in the upper atmosphere to complement the temperature soundings [discussed by Reitebuch (2012), Baker *et al.* (1995), Weissman & Cardinali (2007) and Rahnama *et al.* (2006) to name some examples]. In this paper we effectively assume that temperature and wind are uncoupled to directly infer perturbations to the wind profile, which is less well constrained. Future work should assess how changing the wind field implies changing the temperature field; preferably in the context of multi-instrument variational global data assimilation techniques [e.g. Rabier (2005) and references therein]. Whether a solution for both wind components and the temperature can be reached in the absence of *a priori* dynamical constraints and other observations is a subject that is left for future research.

On 2010 December 25 a large event was detected by 12 arrays on the global infrasound network operated by the Provisional Technical Secretariat (PTS) of the Comprehensive Nuclear Test Ban Organization (CTBTO). The event is listed in the Reviewed Event Bulletin (REB) that is published by the PTS. Such a well-recorded event, which was located by PTS at 40.024411 156.99662 in the North Pacific, is thought to be from a bolide as there are no other likely causes of such signals from that location. Probably due to the remote location of the event, we have not found any eyewitness accounts of this event. We apply our inversion scheme to this event to illustrate the application of our methodology to real data.

The technique described in this paper exploits the fact that delay times between an unknown source recorded at spatially separated arrays can be extracted from cross-correlation. Haney (2009) successfully applied this technique to microbarom noise recorded over local distances. The scope of this paper is to develop and test the mathematical framework for inverting such delay times once they have been obtained, to assess the effect of noise on the inversion scheme and to apply the method to real data. The potential application to microbarom signals is described in the Discussion section of this paper.

2 A METHOD FOR ESTIMATING STRATOSPHERIC WIND SPEEDS FROM UNKNOWN SOURCES

Our method uses celerity, estimated using distances and time delays between sensor arrays, and backazimuth and elevation angles, estimated from array processing, to invert for wind in the stratosphere. We begin by describing the data processing steps that provide the inputs for our inversion method. Next, we introduce the parametrization scheme used to represent the 1-D winds; we outline the forward model used in this study; and finally we discuss the inverse approach itself.

2.1 Data processing

The methodology developed in this paper requires inputs that are estimated in two distinct processing steps: (1) estimation of the signal direction-of-arrival (DOA) at each array, and (2) estimation of the interarray time delay for each pair of arrays. The DOA is a vector that comprises both the backazimuth (θ) and elevation angle (ϕ). The time delay can be estimated from cross-correlation where the signals are consistent across the network. Since both estimates involve uncertainty, which can be quite significant in some cases, it is important to quantify the upper bound of uncertainty under which the algorithm can work. By exploring the effect of noise on the measurements, we further extend the work of Drob *et al.* (2010a) and Lalande *et al.* (2012) who assumed noise-free measurements.

The method illustrated in this paper assumes that the network is small enough such that a 1-D atmospheric model is a reasonable approximation, and that the signal is correlated on the individual arrays. Thus the cross-correlation derived delay-time between each pair of arrays can be modelled according to the measured backazimuth and inclination angle at the more distant of the two arrays—see Figs 1 and 2, which are discussed in more detail in Section 2.4. Measuring θ reliably is trivial; but determination of ϕ is more difficult (Szuberla & Olson 2004). The elevation angle, ϕ, is typically derived from the measured trace velocity of the signal

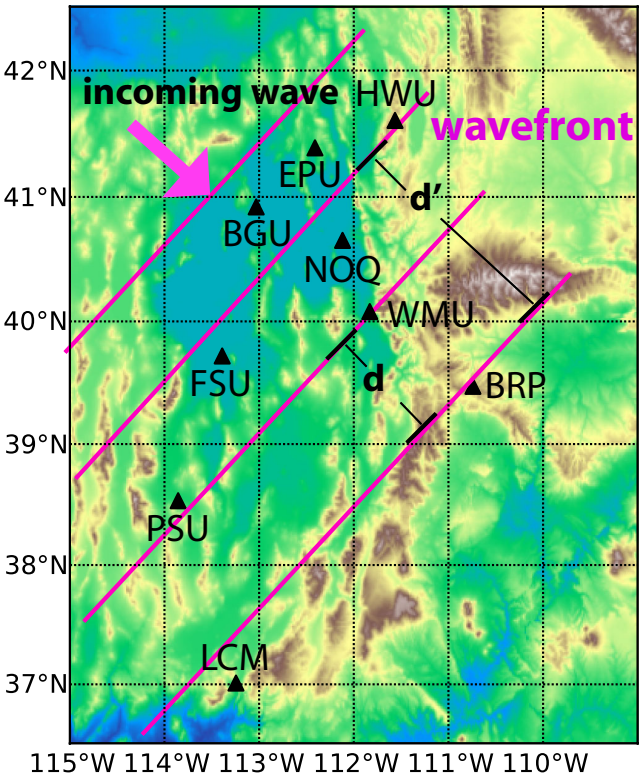

Figure 1. Map of the Utah network arrays (black triangles) with a schematic illustrating an incoming signal from a distant event, depicted here as a plane wave. Assuming that the signals are correlated at the different arrays, we can calculate a celerity given the estimated distance between the arrays in the direction of propagation (d and d' denote interarray distances between BRP-WMU and BRP-HWU).

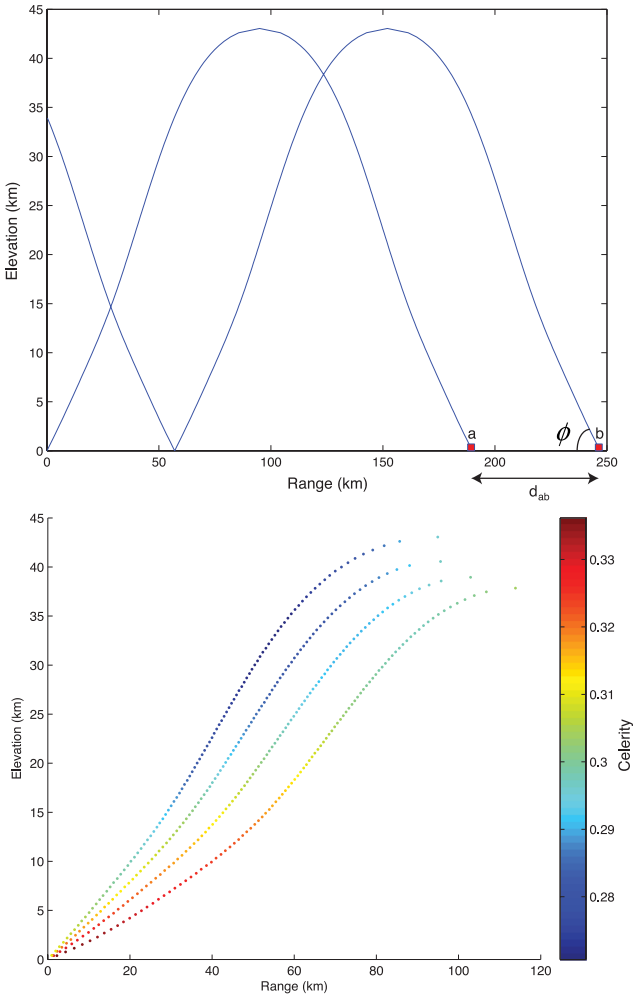

Figure 2. Top panel illustrates the network configuration from Fig. 1 in cross-section using two arrays. Because the inversion scheme uses the predicted celerity from the ground to the turning point, the minimum value of d_{ab} is estimated for the 2010 December 25 bolide by finding the range at which the celerity is 0.016 km s^{-1} less than the one bounce celerity (where 0.016 is based on the data uncertainties in Section 4.2). The second panel illustrates how the measured celerity would depend on the interarray distance for the four example rays shown.

across the array and the local average ambient temperature at the ground, but can also be measured directly using 3-D array processing (Edwards & Green 2012). Wind and temperature changes in the boundary layer can distort the arrival elevation angle; the absence of direct measurement of these meteorological parameters can result in a model error that compounds the measurement uncertainty. The required accuracy of the elevation angle measurements, in addition to the interarray delay time measurements, is tested numerically in Section 3.

2.2 Parametrization of the atmosphere using basis functions

Following Drob *et al.* (2010a), to reduce the number of unknowns, the wind profiles are represented by a set of basis functions (referred to as Empirical Orthogonal Functions, EOFs). The EOFs are calculated by taking the singular value decomposition of the profiles of the background winds from a multiyear time-series of Horizontal Wind Model (HWM)/Mass Spectrometer Incoherent Scatter (MSIS) or G2S profiles for a given location. The limitation of this approach is that it assumes that the EOFs capture the full variability of the wind profiles for a given location. The $m \times n$ matrix of zonal wind profiles $\mathbf{u}$ (where m is the number of time periods in the multiyear time-series with profiles every 6 hr covering the timespan from 2003 January 01 to 2011 February 28, and n is the number of discrete points in elevation with a sampling interval of 200 m and covering the range from 0 to 160 km) can be partitioned into the form $\mathbf{u} = \mathbf{U}\Sigma\mathbf{V}^*$ where $\mathbf{U}$ is an $m \times n$ matrix of amplitudes, $\sum$ is an $n \times n$ diagonal matrix of positive eigenvalues, and $\mathbf{V}^*$ is an $n \times n$ matrix of eigenvectors. The EOFs, $\psi_j(z)$ in eq. (1), are equivalent to the j-th row of $\mathbf{V}^*$ multiplied by the j-th eigenvalue. The linear scaling factors β_{kj} correspond to the value of $\mathbf{U}$ at the k-th row, which corresponds to the given time, and the j-th column. The zonal wind profile for a given time can now be represented by

$$\tilde{u}(z) = \sum_{j=1}^{n} \beta_{kj}\psi_j(z) + u_0(z), \tag{1}$$

where k represents the time and $u_0(z)$ is the multiyear average background state. The meridional wind component can be similarly parametrized.

Using the HWM/MSIS empirical climatologies (Hedin *et al.* 1996; Picone *et al.* 2002; Drob *et al.* 2008; Arrowsmith *et al.* 2010), each profile is suitably parametrized using only 10 EOFs (Fig. 3). EOFs derived from the empirical climatolgies were chosen over a set derived from the G2S time-series to reduce the number of degrees of freedom when applying the inverse methodology to the present set of observations. However, because the propagation of infrasound signals analysed in this study is approximately oriented in the zonal direction, we show that it is appropriate to invert only for perturbations to the zonal wind profile. Thus the total number of unknowns for a typical inversion is 10.

2.3 Forward model: infrasound propagation modelling

The forward model used in this study is based on the Tau-P equations (Garces *et al.* 1998; Drob *et al.* 2010b), which assume that the local effective sound speed velocity profile controls the first-order characteristics of infrasound propagation. The Tau-P equations, reproduced here for conciseness, are formulated in terms of the along track range (R), traveltime (T) and transverse offset (Q). Using the notation of Drob *et al.* (2010b), the along track range is

$$R(z, p) = 2\int_{z_0}^{z(p)} \psi(z, p)\left[\frac{p}{1 - u(z)p} + u(z)\zeta(z)\right]\mathrm{d}z, \tag{2}$$

where the ray parameter (p) is

$$p = \frac{k_z}{c_0}\left(1 + \frac{k_z u_0}{c_0}\right)^{-1}. \tag{3}$$

In eqs (2) and (3), k_z is the vertical wavenumber, $k_z = \sin(\phi)$, z_0 is the surface altitude and $z(p)$ is the upper limit, which is the first root above z_0 of the characteristic function

$$\psi(z, p) = \left(\zeta(z) - \frac{p^2}{(1 - u(z)p)^2}\right)^{-1/2}, \tag{4}$$

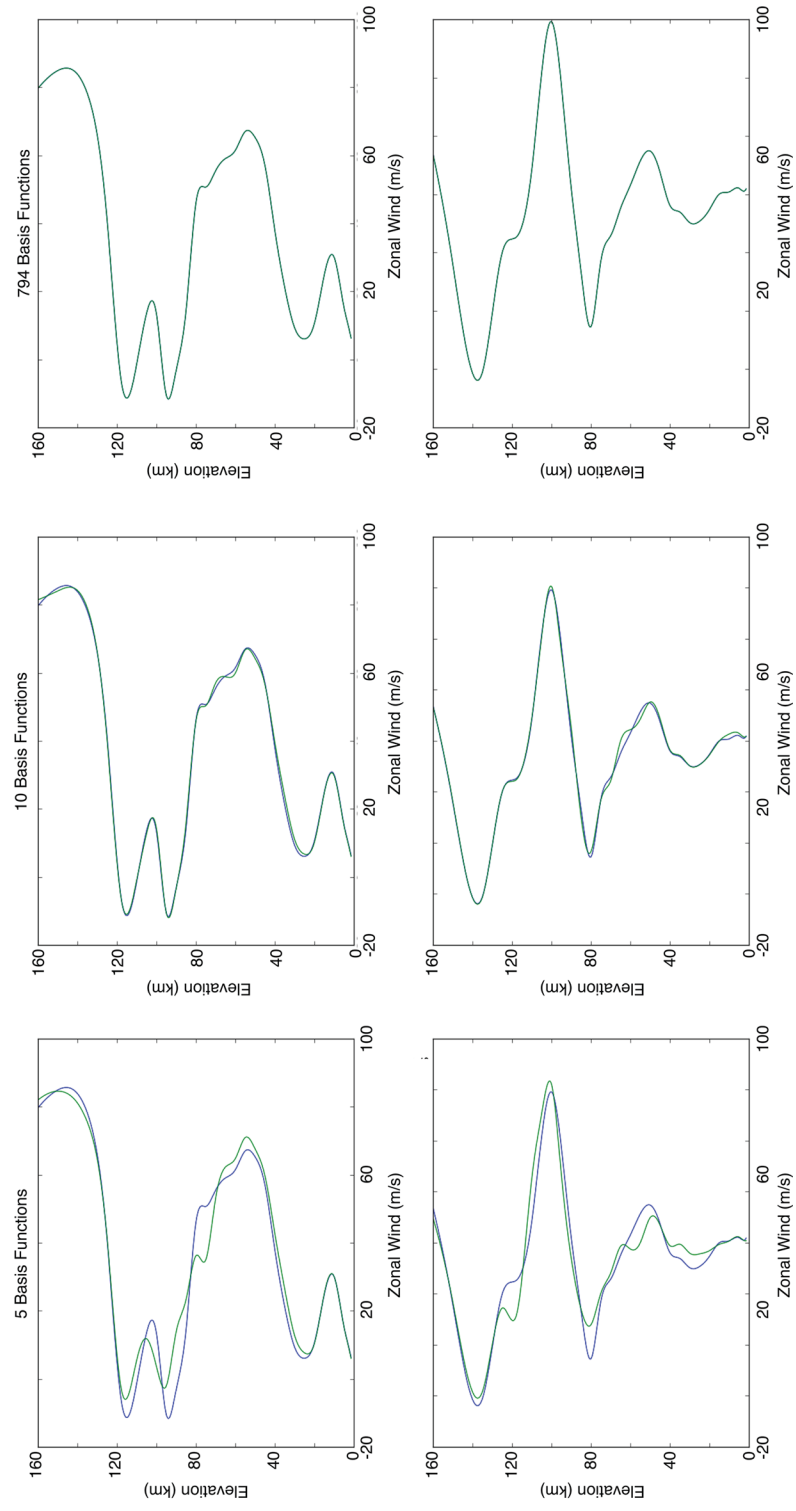

Figure 3. Comparisons between a starting HWM/MSIS model (blue curves) and the reconstructed profiles using different numbers of basis functions (green curves).

where $\zeta(z) = 1/c^2(z)$ and $u = u_{||}$ is the wind speed in the direction of propagation. The traveltime is

$$T(z, p) = 2 \int_{z_0}^{z(p)} \psi(z, p)\, \zeta(z)\, \mathrm{d}z, \tag{5}$$

and the transverse offset is

$$Q(z, p) = \int_{z_0}^{z(p)} \psi(z, p)\, \zeta(z) v(z)\, \mathrm{d}z, \tag{6}$$

where $v(z)$ is the horizontal wind component transverse to the direction of propagation.

Despite its simplicity, which makes it inappropriate for simulating long-range propagation, the Tau-P method is adequate for simulations over regional distances. For the inverse approach outlined below we only need to calculate celerities between pairs of arrays (from eqs 2 and 5). For the Utah network used in the simulations, the propagation ranges are short enough (the average interarray distance is 230.5 km) that the Tau-P method is suitable.

2.4 Inverse approach: estimation of zonal wind speeds

Drob *et al.* (2010a) and Lalande *et al.* (2012) used source location as a metric for refining a starting model. Here, the basic metric used to iteratively refine a starting model, which might be the HWM/MSIS climatology or G2S model for a given location and time, is the misfit between the measured and the predicted celerities. Using Figs 1 and 2 to illustrate this concept, and the assumptions therein, we can measure $v_{ab} = \frac{d_{ab}}{\Delta t}$ from the cross-correlation time lag obtained between arrays a and b. The use of cross-correlation assumes that the waveforms at a and b are sufficiently correlated. Our technique further assumes that the distance between arrays a and b is large enough such that the measured signal velocity represents the celerity of the signal. Our inversion algorithm in its current implementation minimizes the difference between the measured celerity between each pair of arrays and the predicted celerity that is extracted by ray tracing using the measured DOA at the more distant array from the source (defined as the last array receiving the signal). This means that the altitudes to which the inversion scheme is sensitive is strongly related to the interarray distances. The highest altitude to which the algorithm is sensitive is defined by the distance separating the most distant arrays. If the predicted celerities are calculated at the turning heights, we find that, as long as the arrays are greater than 60 km apart, the difference between the celerity measured by cross-correlation and model-based celerity calculated by ray-tracing from the ground to the turning point should be less than 0.016 km s^{-1} (equivalent to the estimated measurement uncertainty in celerity in Section 4). As the launch angle increases, the required minimum distance between the two arrays decreases as shown in Fig. 2. While a more rigorous study of the closest distance between each pair of arrays would be of value, the inversion algorithm could be readily modified such that the predicted celerity is only calculated up to a height based on the distance between the two arrays (and not always up to the turning point). This would enable the inversion algorithm to be applicable to scenarios where the arrays are closer together than 60 km, but is left for further study. Since array b is the more distant of the two arrays, this measured celerity should equal the celerity predicted by shooting a ray backwards from array b with a negative time step at the measured backazimuth (θ) and inclination angle (ϕ) at b. A further assumption inherent here is that the distance between arrays a and b must be small enough that range-dependence between the arrays is negligible; otherwise the measured apparent signal velocity between array pairs will not represent a celerity but rather will include velocity differences between different paths. For N arrays, this problem can be set up as a minimization problem, where we wish to minimize

$$\chi^2 = \sum_{i=1}^{M} [f(\mathbf{n}_i; \beta) - v_i]^2, \tag{7}$$

where $\mathbf{n}_i = [\theta_i, \phi_i]$ is the measured wave front normal vector for the i-th station pair (i.e. for the more distant of the two stations), $M = \frac{N!}{2!(N-2)!}$ is the number of pairs of arrays, and β are the 10 linear scaling factors described earlier. The function $f(\mathbf{n}_i; \beta)$ represents the predicted celerity given a set of basis function coefficients using eqs (2) and (5), where the term u is substituted by $\tilde{u}$ in eq. (1).

Based on synthetic tests discussed later, and similar to Drob *et al.* (2010a), we find that it is necessary to introduce prior constraints to reduce the non-uniqueness of the inversion by damping solutions with profiles that differ from the starting profile. This prevents non-physical variations of the profiles that may satisfy the data. With prior constraints, we wish to minimize the following function

$$\chi^2 = \sum_{i=1}^{M} [f(\mathbf{n}_i; \beta) - v_i]^2 + \lambda \sum_{j=1}^{P} \left[u_j(\psi; \beta) - u_j^s\right]^2, \tag{8}$$

where u_j^s is the zonal wind for the j-th altitude in the starting model, P is the number of altitudes over which damping is applied, $u_j(\psi; \beta)$ is the constructed zonal wind for the j-th altitude from the EOFs and linear scaling factors and λ is the damping parameter that controls the relative importance of each term in eq. (8). The second term is only summed over the bottom 35 km, where good radiosonde coverage exists and this is considered as reliable *a priori* information.

We search for the optimum values of β using the Levenberg–Marquardt algorithm (Seber & Wild 2003). The values of β are iteratively adjusted to obtain the reconstructed profile via eq. (1) that optimally minimizes χ^2 via eq. (8).

The azimuth of the incoming wave front affects the relative contributions of the zonal and meridional wind components. For example, if the incoming wave backazimuth is 270°, the transverse wind component is equivalent to the meridional wind component and thus the meridional wind has no effect on the traveltime (eq. 5). In this case, the meridional wind is completely unconstrained in the inversion and should not be allowed to change. Based on the results of a suite of synthetic tests for a regional infrasound network at mid-latitudes it was found that inverting for the zonal wind is robust over a wide range of azimuths, whereas the meridional wind speed is typically sufficiently low in the stratosphere that it is necessary to hold the relevant coefficients fixed in the inversion.

Another observation from numerous synthetic tests is that we must also constrain the upper atmospheric portion of the profiles constructed from the optimum values of β when computing the forward model (i.e. for each iteration of the Levenberg–Marquardt algorithm). Following a similar approach to Drob *et al.* (2010a), the upper atmosphere is constrained as follows:

$$u(z) = \pi(z) \cdot \tilde{u}(z) + (1 - \pi(z)) \cdot u_a(z), \tag{9}$$

where $u_a(z)$ is the starting profile in the upper atmosphere and the $\pi(z)$ is a sigmoid function defined by

$$\pi(z) = \left(1 + \mathrm{e}^{[0.2(z-70)]}\right)^{-1}. \tag{10}$$

3 SYNTHETIC TESTS USING REALISTIC NOISE (UNCERTAINTY) ESTIMATES

Following Rogers (2000) the purpose of performing synthetic tests is twofold: (1) to test the algorithm and code, and (2) to explore the limitations associated with the assumptions of (a) reliable measurements of the wave front normal vector, and (b) reliable measurements of the delay-times between each pair of arrays.

A suite of synthetic tests has been performed using the nine-array Utah infrasound network configuration (Fig. 1). For each test, an assumed source location is chosen and associated great-circle backazimuths are calculated for each array. Next, for a given atmospheric profile, represented by $c(z)$, $u(z)$ and $v(z)$, a distribution of launch angles is chosen randomly such that all the rays are ducted in the stratosphere when propagated backwards from each array at the given backazimuth and launch angle by using a negative time step in the Tau-P equations. This yields the synthetic data set, which comprises:

(1) Synthetic arrival parameters (i.e. the backazimuth and launch angle) at each array,
(2) Synthetic celerities associated with each arrival (calculated by dividing eq. 2 by eq. 5).

The examples shown in this section use only the HWM model for brevity. The results using the G2S model are the same except that it is necessary to use 40 basis functions to represent the zonal wind profile, rather than 10. It should also be noted that, while the launch angle (ϕ) is not typically directly measured (the trace velocity is measured instead), we begin by assuming that this parameter, in addition to the backazimuth (θ), can be retrieved perfectly. Furthermore, celerities (v) are not measured directly in the real case, they are calculated from the cross-correlation delay times, and these are also assumed to be retrieved perfectly in the initial synthetic tests.

Next, the scaling factors (β) are perturbed by some amount to form the starting model for the synthetic test; this model differs from the true model, which is the one used to calculate the synthetic data set. We evaluate different levels of perturbations to assess the fidelity of the starting model that is required.

Although eq. (8) is used to control the inversion by minimizing the difference between the predicted and measured celerities with prior constraints, the actual properties we are interested in measuring are the zonal and meridional wind speeds in the stratosphere. Thus, to assess the accuracy of the inversion, we can calculate

$$\gamma = \frac{1}{N}\sqrt{\sum_i \Delta u_i^2}, \tag{11}$$

where γ is the mean difference between two zonal wind profiles over some altitude range. In the following discussion, γ_s refers to the rms difference between the starting profile and the true profile, while γ_e refers to the rms difference between the final estimated profile and the true profile. Since ground-based measurements from radiosondes cover the range from 0 to 35 km altitude, this paper focuses on improving wind specifications above this range. The range over which we estimate γ is dependent upon the turning points of the ray paths. For the synthetic tests below, the ray paths turn between 35 and 45 km, thus this is the range used. For summertime propagation conditions in the Northern Hemisphere, the stratospheric zonal wind jet peaks near 70 km altitude, about 10 km higher than the wintertime jet, which peaks between 45 and 55 km, so the height range for which the inversion is applicable will also vary to a degree with the actual geophysical conditions. We also note that the inverse methodology described here does not apply during Equinox where the stratospheric winds are more or less zero and stratospheric microbarom signals are not observed. In this paper we have chosen to focus on stratospheric returns but in principle the same methodology could be applied to thermospheric returns. In this case, seasonal variations and directionality effects would be less pronounced, and it may be possible to invert for both zonal and meridional perturbations.

3.1 Without prior constraints

The examples provided in this subsection illustrate the application of synthetic tests without the addition of prior constraints. For the first example, a suite of different source locations is chosen at different azimuths over which stratospheric returns are possible for a given wintertime atmospheric specification. The first five β coefficients are increased by 50 per cent (the effect of this is that the mean perturbation from 35 to 45 km is $\gamma_s = 25.6$ m s^{-1}). This example tests the dependence on successfully inverting for the zonal wind component while varying the azimuth of the source. The results (Fig. 4) show that all inversions have converged to a much-improved fit (i.e. the value of γ_e is much less than the value of γ_s). This result indicates that it is possible to reliably invert for the zonal wind component despite the source–receiver path having a significant N–S component. In contrast to the zonal wind results, the meridional wind solutions do not converge for each azimuth. This is because these profiles have such a minor influence on the synthetic celerities that they are essentially unconstrained.

For the next example, a hypothetical source location in the North Pacific is used (40°N, −170°E). The results from a suite of simulations with different levels of starting wind perturbations (γ_s) are shown in Fig. 5. It was found that, in some cases, the solution converged towards local minima—indicated by poor fits to the known celerities (large values of χ^2) and by poor fits to the wind profiles (large values of γ_e). The most common situation where this occurs is caused by rays that almost get refracted back to the ground in the stratosphere, but actually make it to the thermosphere where they are refracted. Such rays have celerities that are consistent with stratospheric returns, not thermospheric returns. An analysis of the cases where the inversion does not converge shows that these cases are associated with starting models that either have a weak stratospheric duct, or no duct. Of course, if we measure delay times between signals recorded at real spatially separated arrays that are consistent with stratospheric celerities, and we have additional supporting evidence that they are stratospheric signals, we 'require' a duct in the stratosphere to predict these arrivals. Thus, it makes sense for the starting model to have a strong stratospheric duct to avoid possible non-converging inversions. An analysis of different degrees of perturbations shows that even with a very strong rms increase in the stratospheric wind speed of ∼25 m s^{-1}, the solution converges satisfactorily (Fig. 6). In fact, as shown in Fig. 5, strong starting ducts improve the final solution.

3.2 With prior constraints

The addition of Gaussian-distributed pseudo-random noise to the measured backazimuths, launch angles and celerities requires damping to ensure that the resultant solution is consistently physically reasonable. However, based on a series of synthetic tests we have found that the addition of noise to the backazimuths does not

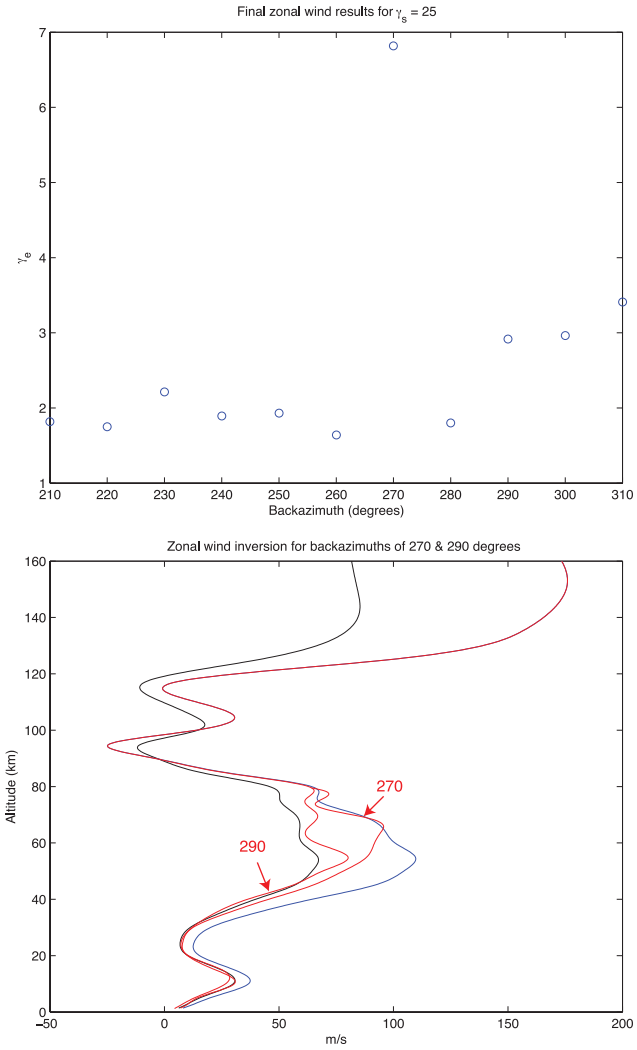

Figure 4. Top panel: Effect of azimuth. Bottom panel: Inversion results for azimuths of 270° and 290°. The black curve represents the true profile (used in generating synthetic data), the blue curve represents the starting (perturbed) profile and the red curves represent the final estimated profiles. The blue curve was obtained by perturbing the first five basis functions that make up the black curve by a constant factor. Above ∼80 km altitude the red and blue curves are identical as a consequence of implementing eqs (9) and (10).

require damping and does not significantly degrade the solution up to an unrealistically large level of added noise (Fig. 7). With respect to the addition of random noise in the synthetic data inversion experiments, there are actually two sources of noise, which are pertinent to the geophysical remote sensing problem at hand; the first being unresolved fluctuations of the geophysical variable being measured and not represented or resolvable by the forward model, and the second being the measurement or instrument noise resulting from the infrasound array measurement technique. In this paper we focus on the latter of these two components.

In contrast to the backazimuths, the addition of noise to the launch angles and measured celerities is somewhat more sensitive. Based on the estimated uncertainties in the measurements in Section 4, a suite of synthetic tests using HWM and G2S profiles for the centre of the Utah network, have been performed. Rather than perturbing the basis function coefficients directly, as in Section 3.1, here we forward model synthetic data using a true model ($\mathbf{m}_T$) that is identical

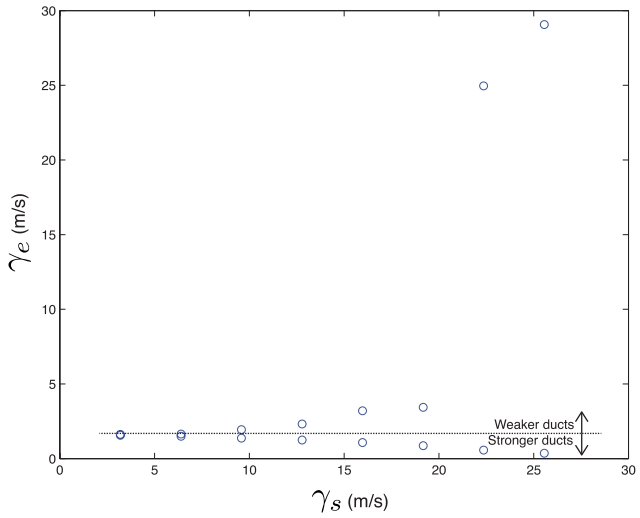

Figure 5. A comparison between γ_s (x-axis) and γ_e (y-axis) for multiple different synthetic simulations without using prior (damping) constraints. The dashed line separates inversion solutions with a starting profile that has a stronger duct from inversion solutions with a weaker duct. Stronger-duct starting models converge much better because they avoid the problem with a change in propagation regime associated with weak-duct models.

to the HWM model ($\mathbf{m}_H$) except between altitudes of 40 and 80 km. Starting our inversions with $\mathbf{m}_H$, we then assess the difference between the final estimated model ($\mathbf{m}_E$) and $\mathbf{m}_T$. Because the rays typically turn below 60 km altitude, our synthetic data provide no resolution on altitudes above this height (the only information comes indirectly from the EOFs, which provide some constraints on the atmosphere immediately above this range). For this reason, and due to eqs (9) and (10), profiles above ∼70 km are unconstrained by data and are controlled by the starting profile. The results, which have been calculated for a range of different damping parameters (where $0 \leq \lambda \leq 1$) with $\sigma(\phi) = 12^\circ$, $\sigma(v) = 0.016\,\mathrm{km\,s^{-1}}$, suggest that a damping parameter of $\lambda = 1\mathrm{e}-3$ should provide optimal results. The choice of $\sigma(\phi) = 12^\circ$ and $\sigma(v) = 0.016\,\mathrm{km\,s^{-1}}$ is based on the estimated uncertainties in the data observed from the 2010 December 25 bolide (see Section 4). An example from these tests (where the interpretation is independent of realization) is provided in Fig. 8. In the example provided in Fig. 8, for the damping parameter (λ) set to $1\mathrm{e}-6$, the recovered profile oscillates in the lowermost part of the atmosphere. In contrast, by incorporating more damping (setting $\lambda = 1\mathrm{e}-3$), the recovered profile is a much better realization of the true profile than the starting profile. By over-damping (e.g. $\lambda = 1$), the final solution is too heavily restricted from deviating from the starting model and thus cannot improve upon the starting model. We emphasize that the units of wind speed ($\mathrm{m\,s^{-1}}$) and celerity ($\mathrm{km\,s^{-1}}$) are not the same in eq. (8), which contributes to the low optimal values of λ. However, forcing them to have the same units does not allow for a more direct comparison of weights due to the different relative speeds of sound and wind and the way in which these parameters are incorporated in the inversion. Further work is required to investigate these uncertainties and their subsequent effect on the inversion scheme, especially to consider how using 3-D array processing (Edwards & Green 2012) might help reduce these uncertainties. In addition, the incorporation of additional arrays should help to better constrain the inversion by averaging out the effects of measurement errors.

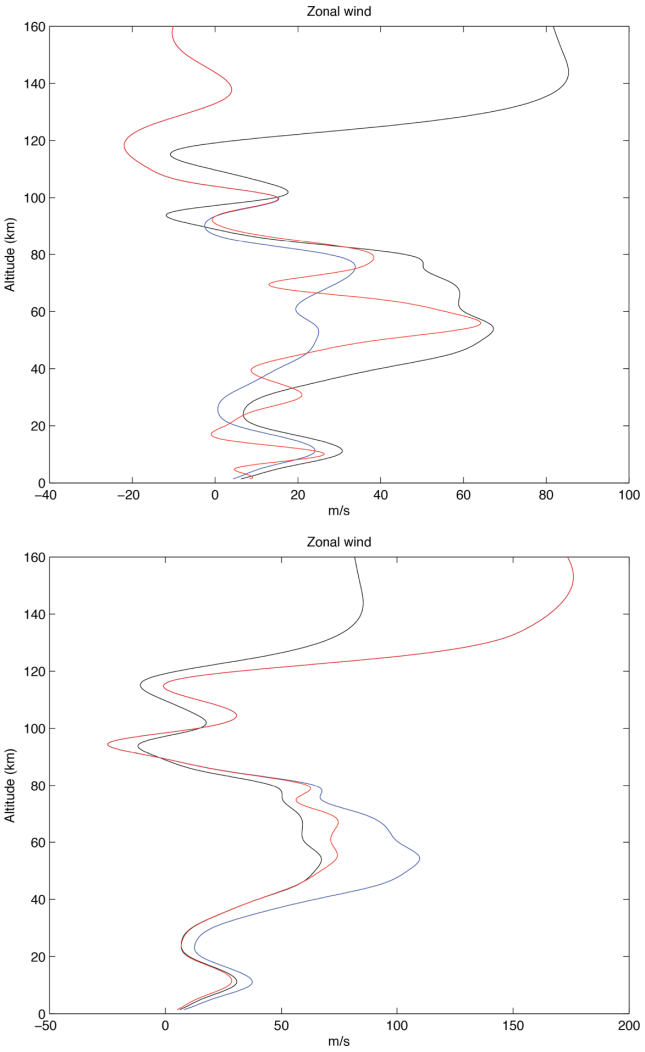

Figure 6. A comparison between a weak-duct starting model (top) and strong-duct starting model (bottom) without using prior (damping) constraints. Both models have exactly the same initial rms perturbation from the true model. The black curve represents the true profile (used in generating synthetic data), the blue curve represents the starting (perturbed) profile (obtained by perturbing the first five basis functions of the black profile by a constant factor), and the red curve represents the final estimated profile.

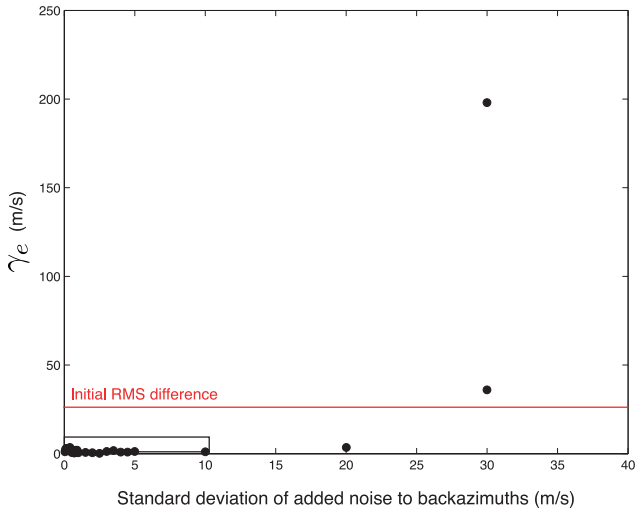

Figure 7. Results obtained after adding perturbations (a single realization for each case) to the synthetic backazimuths without using prior (damping) constraints. For tests with $\sigma = 20^\circ$ and below, the final profile is significantly closer to the true profile than the starting profile. At greater perturbations, the inversion breaks down as expected.

4 ANALYSIS OF INFRASOUND DATA FROM THE 2010 DECEMBER 25 BOLIDE

As described in Section 2.1, the inversion scheme developed in this study involves two distinct processing steps.

4.1 Measurements of the DOA

Using the adaptive F-detector (Arrowsmith *et al.* 2008), we detected signals from the event at nine operating infrasound arrays in Utah at distances between 7235 and 7486 km from the event location. Due to low signal-to-noise levels at two of the arrays, we focus in this paper on the signals recorded by seven arrays (Fig. 9). The signals occur at arrival times that are consistent with infrasound that has refracted multiple times in the stratosphere, based on the event location and origin time estimated by the CTBTO. While the CTBTO location is an estimate, the infrasound trace velocities (discussed in detail later) provide additional supporting constraints for the signals being refracted in the stratosphere.

The array analyses provide backazimuths and trace velocities at each array. Complicating the array analysis, the dominant frequency content of the bolide signal, which spans the frequency range of 0.03–2 Hz, overlaps significantly with the microbarom band (Fig. 9). In addition, the apertures of the Utah arrays (∼150 m) are designed for high-frequency infrasound signals. Thus, to obtain accurate estimates of the DOA for each array, we filter the data from 1–2 Hz with a from-pole Butterworth filter. This filter band removes most microbarom energy but retains energy from the bolide, and ensures the array response is not so broad that the estimates are poorly constrained.

The data are processed with a 40 s moving time window, overlapping by 50 per cent, with each window providing estimates of the F-statistic, correlation and DOA. We estimate the DOA by taking an average value of the azimuth and trace velocity in a window defined by the detection window (this region can be distinguished from the background at a statistically significant level). The DOA measurements for each array are provided in Table 1.

As shown in Section 3, the inversion scheme is robust for the measurement errors of $\sigma(\phi) = 12^\circ$ and $\sigma(v) = 0.016\,\mathrm{km\,s^{-1}}$, estimated later. We note that measurement error can be further reduced using a 3-D solution (Edwards & Green 2012), but this requires accurate measurements of individual array elevations. For the Utah arrays, altitudes were measured using a handheld GPS unit (Hale, personal communication, 2013), which can have uncertainties of the order of tens of metres. Since the array elevation differences for the 100 m Utah arrays are generally selected to be small, we choose to assume planar arrays rather than to rely upon these GPS measurements. What cannot be easily quantified is the model error associated with the trace velocity measurements. The model error reflects the difference between the unknown atmospheric conditions (which include boundary-layer effects poorly resolved in the models) and the atmospheric specifications used in the inversion. Since this error cannot

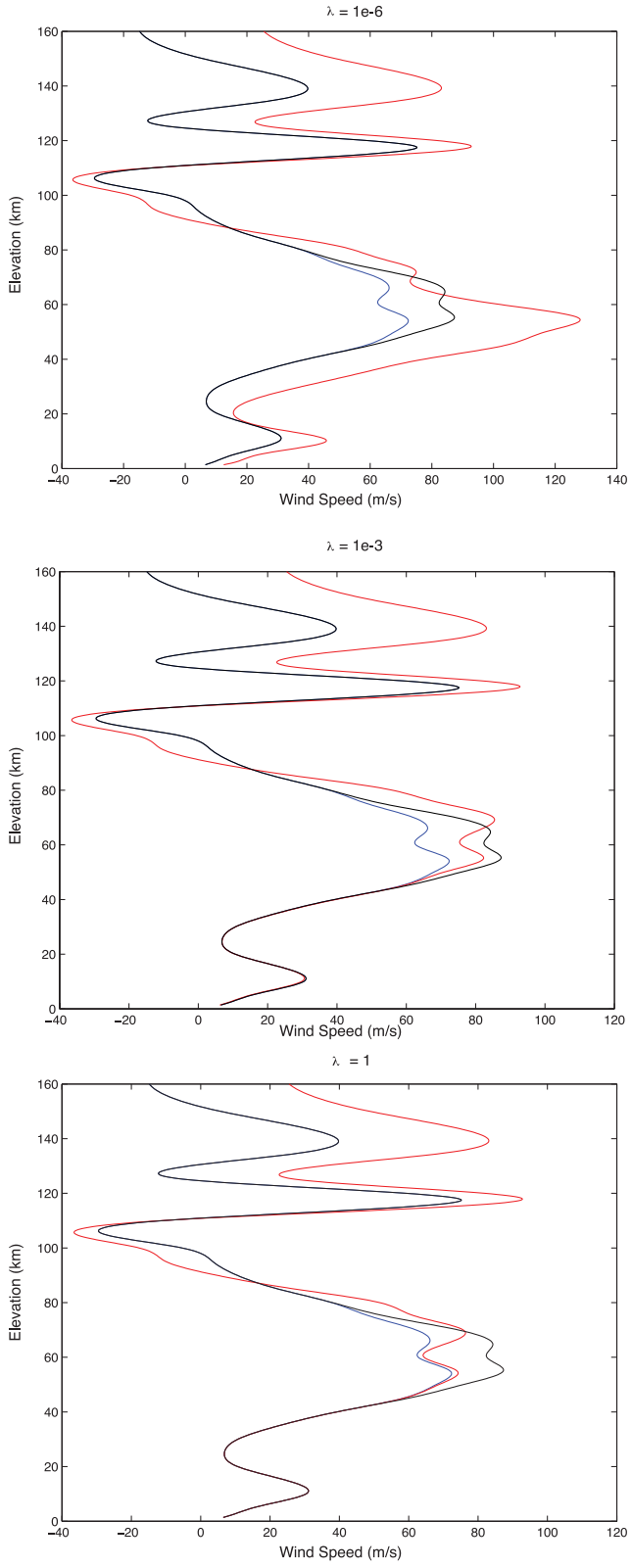

Figure 8. Results of inversions for purely synthetic data (red profiles) generated using the black profile and inverted using the blue profile as the starting profile. Gaussian noise with $\sigma(\phi) = 12^\circ$, $\sigma(v) = 0.016$ km s^{-1} have been added to the synthetic data.

be readily quantified, we choose to use an average of the trace velocities at all the arrays in the inversion scheme. Our rationale behind this is that this average should remove local boundary-layer effects that are unique to each array, thus minimizing the model error. This approach assumes the signals at each array come from the same duct. However, given the long-range propagation and small relative distances between arrays, we contend that this approximation is reasonable to first order. The *SD* of the measured trace velocities at all arrays, provided in Table 1, is 0.04 km s^{-1}; we use this as a crude estimate of the model error. Assuming a temperature of 20°C, this translates into an inclination angle uncertainty of $\sigma(\phi) = 12^\circ$.

4.2 Measurements of the interarray celerities

We require reliable estimates of the interarray celerities for accurate convergence of our inversion scheme. We use waveform cross-correlation to obtain interarray delay times, which can be converted to celerity given the measured backazimuths for the event. There are seven arrays that detect the bolide with clear signals (Fig. 9). For convenience in the notation below, each array is labelled with a number from 1 to 7, where 1 is the closest array (BGU) and 7 is the most distant array (BRP) from the estimated event location. We can calculate delay times between each pair of arrays using cross-correlation, noting that only array pairs with distances between arrays of >60 km in the direction of propagation are used.

The cross-correlations are evaluated for each array pair, using the traces between the red lines in Fig. 9 (the waveforms outside these limits are clipped), over a limited set of lag times associated with celerities that fall within the range of typical stratospheric values ($0.28 \leq v \leq 0.31$ km s^{-1}). In other words, we search for local maxima of the cross-correlation functions (which in most cases are the same as the global maxima). Since there is some uncertainty in the origin time of the event in the REB, we search over all stratospheric celerities to find the cross-correlation peaks in this range.

We denote each of the interarray delay times as t_{nm}, where n and m represent specific arrays. For example, t_{35} represents the delay time between arrays 3 and 5. For each of the interarray delay times, we wish to estimate the *SD* associated with that measurement. The *SD* can be estimated by utilizing closure relations defined by all pairs of arrays containing arrays 3 and 5 and a common array. For example, for t_{35} we have

$$\begin{aligned} t_{35} &= t_{34} + t_{45} \\ &= t_{36} - t_{56} \\ &= t_{37} - t_{57} \\ &= t_{15} - t_{13} \\ &= t_{25} - t_{23}, \end{aligned} \tag{12}$$

which effectively provides us with six separate estimates of t_{35}, or equivalently of v_{35} (using the distance between arrays 3 and 5 in the direction of propagation). The *SD* of these six estimates of v_{35}, which we denote $\sigma(v_{35})$, gives us an estimate of the measurement error associated with arrays 3 and 5. The mean of the *SD*s for all interarray pairs, $\sigma(v)$, is 0.016 km s^{-1}.

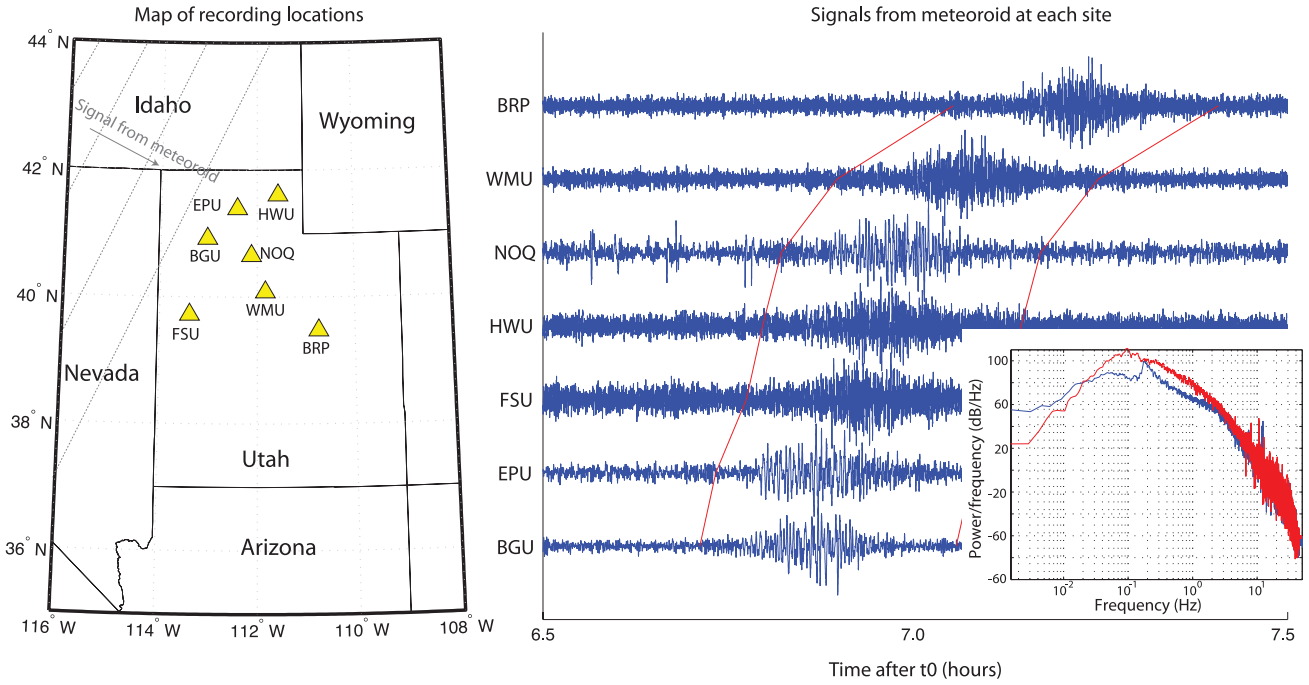

Figure 9. Left-hand panel: Map of the recording network (yellow triangles represent infrasound arrays, grey lines represent wave fronts coming from the direction of the event). Right-hand panel: Bandpass filtered data using a Butterworth filter with three poles and filter band of 0.03–2 Hz. The inset shows the frequency content at EPU (red = signal spectrum, blue = pre-event noise spectrum).

Table 1. Array measurements of the 2010 December 25 bolide recorded on arrays in the Utah network.

Array name	Sensor type	Great-circle distance from event location (km)	Great-circle backazimuth (°)	Measured backazimuth (°)	Measured trace velocity (km s^{-1})
BGU	Chaparral	7235.4	302.4	309.9	0.365
EPU	Chaparral	7251.2	302.5	294.0	0.290
FSU	IML	7281.5	302.7	297.5	0.389
HWU	IML	7297.5	302.8	302.7	0.385
NOQ	Chaparral	7316.1	302.9	294.1	0.290
WMU	IML	7371.3	303.3	305.0	0.359
BRP	IML	7486.2	304.0	300.2	0.359

5 APPLICATION OF THE INVERSE METHOD TO THE 2010 DECEMBER 25 BOLIDE

In applying the methodology described to the bolide signals recorded on the Utah network we consider two scenarios. In each scenario, we start with 1-D temperature and wind profiles from the location of the Utah Test and Training Range and a time of 12/26/2010 06:00:00 UTC (the signals arrive at the arrays between 06:25 and 06:55). Since the propagation paths are largely west to east, we invert for only changes to the zonal wind speed profile. In the first scenario, we use the HWM profile as our starting profile (i.e. starting values of β) and HWM basis functions to provide ψ in the equations mentioned above. In the second scenario, we use the G2S model to provide β and ψ. If each profile is completely self-consistent with the observed celerities and DOAs, the final profiles from the inversion will be identical to the starting profiles. The amount that each profile is required to change to improve the fit to the data gives us a measure of the consistency between that profile and the observed data.

Using the real data, and both HWM/MSISE and G2S profiles, we perform sensitivity analysis for the damping parameter (see results in Figs 10 and 11). These tests complement the tests on synthetic data in Section 3. Inversions where the damping parameter is too high prevent the model from deviating from the starting model. As we relax the damping parameter, enabling the models to fit the data better, we can characterize how the model changes. Using damping parameters that are too low enable the model to vary in the troposphere. For damping parameters between high and low extremes, the final models obtained using HWM as the starting model (and HWM basis functions) require a stronger duct in the stratosphere than is present in the HWM model (Fig. 10). In fact, comparison between the final estimated model and the G2S model show that the zonal wind speed in the final model is closer to the G2S model than the HWM model. In contrast, for the G2S model, the final model does not deviate significantly from the starting model in the stratosphere, indicating that the G2S profile is consistent with the recorded data in the stratosphere (Fig. 11).

Overall, our results are very encouraging as they suggest that our observations are consistent with the G2S profile. Further, the observations require a stronger duct in the stratosphere than is present in

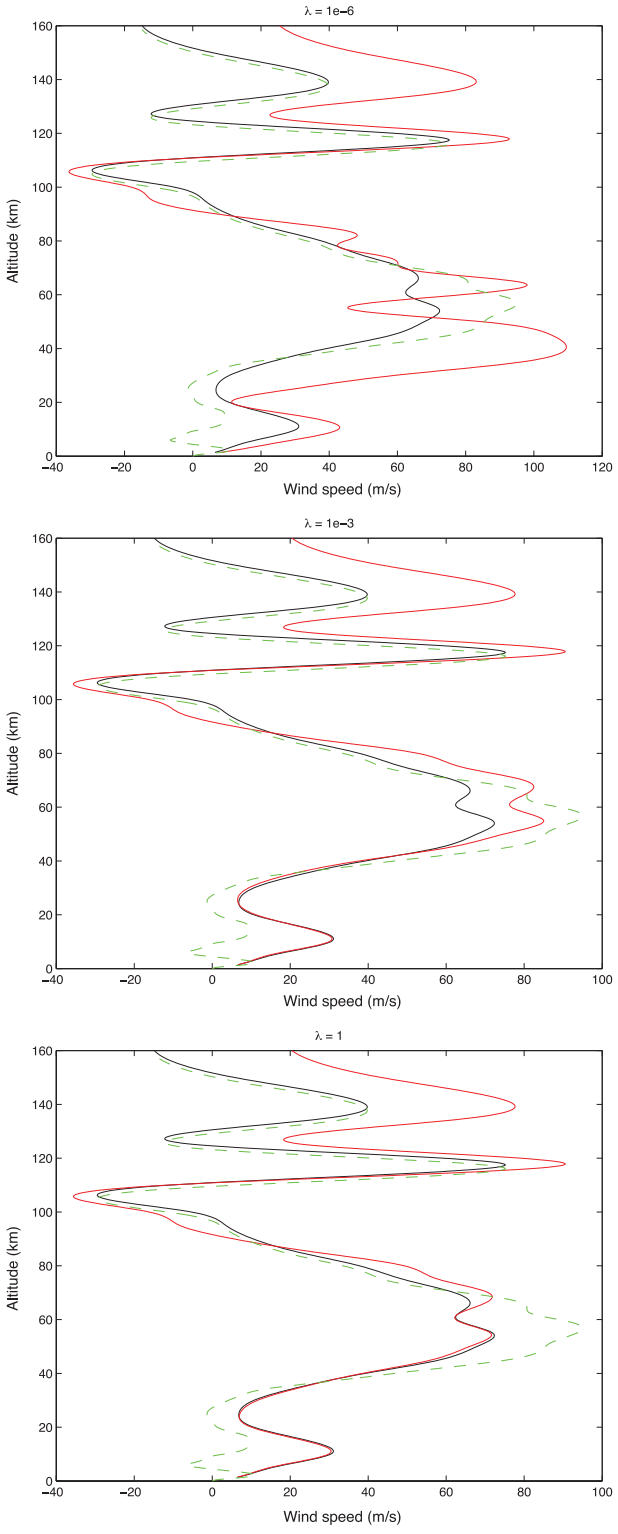

Figure 10. Sensitivity of the result for HWM for different damping parameters. In each panel, the black profile represents the HWM zonal wind profile for the location of the Utah Test and Training Range and a time of 12/26/2010 06:00:00 UTC, the red profile represents the estimated profile from the inversion, and the green dashed profile represents the G2S zonal wind profile for reference. Note that we increase the wind speed in the HWM starting model to ensure a stratospheric duct (the initial DOAs produce a thermospheric duct using the HWM model, which can result in a local minima). For this event, it is not necessary to do this for the G2S profile. Thus, above elevations of 70 km, the starting and final models are not the same.

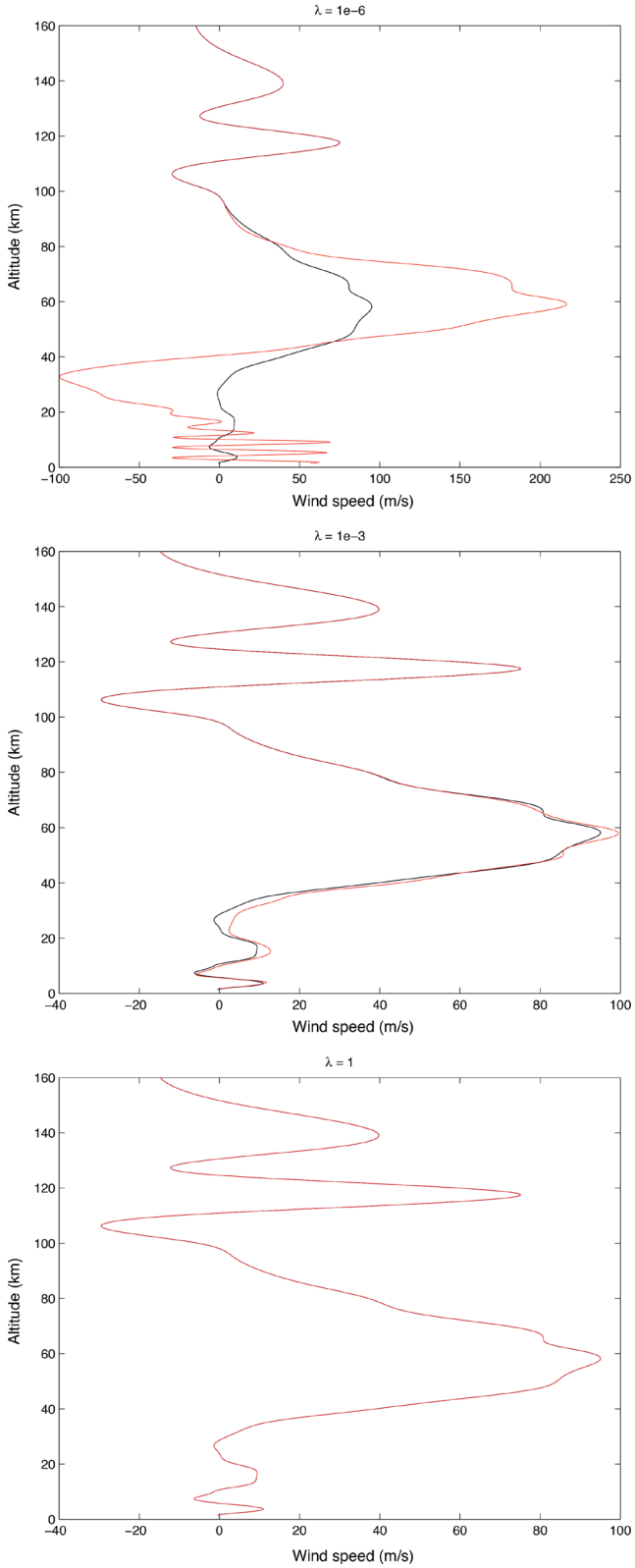

Figure 11. Sensitivity of the result for G2S for different damping parameters. In each panel, the black profile represents the G2S zonal wind profile for the location of the Utah Test and Training Range and a time of 12/26/2010 06:00:00 UTC, the red profile represents the estimated profile from the inversion.

the HWM profile. Of course, both inversions assume that the bottom 35 km is known, and we infer changes to the stratosphere given this assumption for both models. Changes to the bottom 35 km may indeed give different optimum solutions.

6 DISCUSSION AND CONCLUSIONS

The purpose of this paper is to investigate the potential for using infrasound from unknown point or continuous-wave sources to probe the stratosphere. This paper builds on the work of Drob *et al.* (2010a) and Lalande *et al.* (2012), who outlined different frameworks for inverting for infrasound signals from discrete ground-truth events. Our goals in this paper are: (1) to develop the mathematical framework for inverting unknown signals for stratospheric winds, (2) to validate the framework with synthetic tests and (3) to test the framework on a real event. A key advantage of this framework, compared with the methods of Drob *et al.* (2010a) and Lalande *et al.* (2012), is that it does not require a ground-truth location or the simulation of the full propagation paths from source to receiver.

The synthetic tests presented in this paper demonstrate that our methodology is practically feasible and can account for realistic uncertainties in measurement error for a nine-array regional network in Utah. Incorporating additional arrays would enable the technique to handle larger values of $\sigma(\phi)$ and $\sigma(v)$, assuming random errors, since these uncertainties would be more effectively averaged out by the larger number of observations. Further research is needed to better quantify model error on $\sigma(\phi)$ in particular (i.e. what are the possible perturbations in measured trace velocities due to near-surface effects that are not captured in the starting profiles).

We have applied the framework to a large event (presumed to be a bolide) that was clearly detected by seven arrays in Utah. Our results demonstrate that the observed data clearly favour the G2S model over the HWM model for the time of arrival of infrasound at the Utah network. In particular, the results suggest that the zonal wind speed in the HWM model is too weak, but that the zonal wind speed in the G2S model is consistent with our observations.

The benefit of an approach like the one presented in this paper is clear: it does not require the use of a ground-truth event, one does not need to simulate the full propagation path from source to receiver, and the methodology can be potentially applied to continuous sources. Our ongoing research effort is currently exploring the use of cross-correlation techniques to extract time delays from microbarom signals recorded at pairs of arrays that are indicative of upper-atmospheric returns. Microbaroms are a well-known class of infrasound signal (Donn & Rind 1972; Garces *et al.* 2004; Le Pichon *et al.* 2006). Unlike signals from a discrete explosive event, microbaroms have a source region that is distributed over a geographical region, which can move slowly over time. Fortunately, microbarom azimuths at any given array are predominantly driven by the direction of stratospheric winds (Garces *et al.* 2004; Le Pichon *et al.* 2006). This implies that, for a small regional network, microbarom energy should predominantly come from the same general source region due to wind effects. This is important because our method exploits signal correlations, which require some consistency in source location. However, in this paper we have shown that by exploiting differences between traveltimes within a network, it is not necessary to know the source location or to assume a point source. The results from attempts to apply this inverse technique to microbarom signals will be the subject of forthcoming papers.

ACKNOWLEDGEMENTS

We are grateful to the editor, J. Wassermann, and to two anonymous reviewers for their thoughtful reviews that helped to improve this manuscript. We thank Rod Whitaker and Roger Waxler for their comments on aspects of this work, which helped to strengthen this paper. This research was funded by the Los Alamos Laboratory Directed Research and Development program. The MERRA/GEOS-5 data utilized in the G2S atmospheric specifications were provided by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center through the online data portal in the NASA Center for Climate Simulation. The NOAA GFS, also utilized in the G2S specifications, was obtained from NOAA's National Operational Model Archive and Distribution System (NOMADS), which is maintained at NOAA's National Climatic Data Center (NCDC). DPD acknowledges support from Office of Naval Research.

REFERENCES

Arrowsmith, S.J. *et al.*, 2008. Regional monitoring of infrasound events using multiple arrays: application to Utah and Washington State, *Geophys. J. Int.,* **175,** 291–300.

Arrowsmith, S.J., Johnson, J.B., Drob, D.P. & Hedlin, M.A.H., 2010. The Seismo-Acoustic Wavefield: a new paradigm in studying geophysical phenomena, *Rev. Geophys.,* **48,** doi:10.1029/2010RG000335.

Baker, W.E. *et al.*, 1995. Lidar-measured winds from space: a key component for weather and climate prediction, *Bull. Am. Meteorol. Soc.,* **76**(6), 869–888.

Beer, T., 1974. *Atmospheric Waves*, John Wiley & Sons, New York.

Chartrand, R., 2011. Numerical differentiation of noisy, nonsmooth data, *ISRN Appl. Math.*, **2011,** doi:10.5402/2011/164564.

Donn, W.L. & Rind, D., 1972. Microbaroms and the temperature and wind of the upper atmosphere, *J. Atmos. Sci.,* **29,** 156–172.

Drob, D.P., Picone, J.M. & Garces, M., 2003. Global morphology of infrasound propagation, *J. geophys. Res.,* **108,** doi:10.1029/2002JD003307.

Drob, D.P. *et al.*, 2008. An empirical model of the earth's horizontal wind fields: HWM07, *J. geophys. Res.,* **113,** doi:10.1029/2008JA013668.

Drob, D.P., Meier, R.R., Picone, J.M. & Garces, M., 2010a. Inversion of infrasound signals for passive atmospheric remote sensing, in *Infrasound Monitoring for Atmospheric Studies*, pp. 701–731, eds Le Pichon, A., Blanc, E. & Hauchecorne, A., Springer, Dordrecht.

Drob, D.P., Garces, M., Hedlin, M.A.H. & Brachet, N., 2010b. The temporal morphology of infrasound propagation, *Pure appl. Geophys.,* **167,** 437–453.

Edwards, W. & Green, D.N., 2012. Effect of interarray elevation differences on infrasound beamforming, *Geophys. J. Int.,* **190,** 335–346.

Garces, M., Hansen, R.A. & Lindquist, K.G., 1998. Traveltimes for infrasonic waves propagating in a stratified atmosphere, *Geophys. J. Int.,* **135,** 255–263.

Garces, M., Willis, M., Hetzer, C., Le Pichon, A. & Drob, D.P., 2004. On using ocean swells for continuous infrasonic measurements of winds and temperature in the lower, middle, and upper atmosphere, *Geophys. Res. Lett.,* **31,** doi:10.1029/2004GL020696.

Haney, M.M., 2009. Infrasonic ambient noise interferometry from correlations of microbaroms, *Geophys. Res. Lett.,* **36,** doi:10.1029/2009GL040179.

Hedin, A.E. *et al.*, 1996. Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys.,* **58,** 1421–1447.

Kalnay, E., Jusem, J.C. & Pfaendtner, J., 1985. The relative importance of mass and wind data in the present observing system, Paper Presented at *Proceedings NASA Symposium on Global Wind Measurements*, NASA, Columbia, MD.

Lalande, J.-M., Sebe, O., Landes, M., Blanc-Benon, P., Matoza, R., Le Pichon, A. & Blanc, E., 2012. Infrasound data inversion for atmospheric sounding, *Geophys. J. Int.,* **190,** 687–701.

Le Pichon, A., Ceranna, L., Garces, M., Drob, D.P. & Millet, C., 2006. On using infrasound from interacting ocean swells for global continuous measurements of winds and temperature in the stratosphere, *J. geophys. Res.,* **111,** doi:10.1029/2005JD006690.

Picone, J.M., Hedin, A.E., Drob, D.P. & Aikin, A.C., 2002. NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues, *J. geophys. Res.,* **107,** doi:10.1029/2002JA009430.

Rabier, F., 2005. Overview of global data assimilation developments in numerical weather-prediction centres, *Q.J.R. Meteorol. Soc.,* **131,** 3215–3233.

Rahnama, P., Rochon, Y.J., McDade, I.C., Shepherd, G.G., Gault, W.A. & Scott, A., 2006. Satellite measurement of stratospheric winds and ozone using Doppler Michelson interferometry. Part I: instrument model and measurement simulation, *J. Atmos. Oceanic Technol.,* **23,** 753–769.

Reitebuch, O., 2012. The spaceborne wind lidar mission ADM-Aeolus, in *Atmospheric Physics: Background - Methods - Trends*, pp. 815–827, ed. Schumann, U., Springer, New York.

Rogers, C.D., 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Pub. Co. Inc., Singapore, 240 pp.

Seber, G.A.F. & Wild, C.J., 2003. *Nonlinear Regression*, Wiley-Interscience, Hoboken, NJ.

Szuberla, C.A.L. & Olson, J.V., 2004. Uncertainties associated with parameter estimation in atmospheric infrasound arrays, *J. Acoust. Soc. Am.,* **115**(1), 253–258.

Weissmann, M. & Cardinali, C., 2007. Impact of airborne Doppler lidar observations on ECMWF forecasts, *Quart. J. R. Meteorol. Soc.,* **133,** 107–116.