University of Reading

Assessment of algorithms for computing moist available potential energy

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Harris, B. L. and Tailleux, R. ORCID: https://orcid.org/0000-0001-8998-9107 (2018) Assessment of algorithms for computing moist available potential energy. Quarterly Journal of the Royal Meteorological Society, 144 (714). pp. 1501-1510. ISSN 1477-870X doi: 10.1002/qj.3297 Available at https://centaur.reading.ac.uk/76186/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1002/qj.3297

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

University of Reading

Central Archive at the University of Reading

Reading's research outputs online

Received: 06 December 2017 | Revised: 13 March 2018 | Accepted: 20 March 2018 | Published on: 23 August 2018

DOI: 10.1002/qj.3297

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

Assessment of algorithms for computing moist available potential energy

Bethan L. Harris | Rémi Tailleux

Department of Meteorology, University of Reading, Reading, UK

Correspondence
Bethan L. Harris, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, RG6 6BB, UK.
Email: b.l.harris@pgr.reading.ac.uk

Funding information
NERC; NE/L002566/1.

Atmospheric moist available potential energy (MAPE) has traditionally been defined as the potential energy of a moist atmosphere relative to that of the adiabatically sorted reference state defining a global potential energy minimum. Although the Munkres algorithm can in principle find such a reference state exactly, its computational cost has prompted much interest in developing heuristic methods for computing MAPE in practice. Comparisons of the accuracy of such approximate algorithms have so far been limited to a small number of test cases; this work provides an assessment of the performance of the algorithms across a wide range of atmospheric soundings, in two different locations. We determine that the divide-and-conquer algorithm is the best suited to practical application, but suffers from the previously unexplored shortcoming that it can produce a reference state with higher potential energy than the actual state, resulting in a negative value of MAPE. Additionally, we show that it is possible to construct an algorithm exploiting a previously derived theoretical expression linking MAPE to Convective Available Potential Energy (CAPE). This approach has a similar accuracy to existing approximate sorting algorithms, whilst providing greater insight into the physical source of MAPE. In light of these results, we discuss possible ways to improve on the construction of Available Potential Energy (APE) theory for a moist atmosphere.

KEYWORDS
algorithms, available potential energy, convection, moist thermodynamics

1 | INTRODUCTION

Available Potential Energy (APE) theory, as originally outlined by Lorenz (1955), provides a framework to study the energy available to atmospheric motions. The theory is underpinned by the concept of an atmospheric background or *reference* state. Such a state has traditionally been envisioned as being obtained through an adiabatic mass rearrangement, such that the sum of the internal and potential energies of the atmosphere (total potential energy) is minimized. The APE is then found as the difference between the total potential energy of the atmosphere and the total potential energy of the reference state. In its reference state, the atmosphere is at rest and in hydrostatic equilibrium; its density stratification is therefore statically stable and horizontally uniform and no further conversion with kinetic energy can take place. The APE thus gives the total potential energy that is available for reversible conversions into kinetic energy. Assuming hydrostatic balance, minimization of the total potential energy is equivalent to minimization of the enthalpy H, so that

$$\mathrm{APE} = \int (h - h_{\mathrm{ref}})\, dm, \tag{1}$$

where h is the specific enthalpy and the integral is over all mass in the atmospheric domain considered.

For a moist atmosphere, the rearrangements are made via reversible adiabatic processes conserving total water content

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.
© 2018 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

(Lorenz, 1978). In this work, we refer to the APE of a moist atmosphere as the Moist Available Potential Energy (MAPE), following the terminology of Stansifer *et al.* (2017), and we focus only on the vertical component of MAPE. Unlike the dry case, for which reference pressure is uniquely determined by sorting potential temperature, there is no known analytical solution for obtaining the moist reference state from the distribution of entropy and specific humidity. As a result, previous methods of calculating MAPE have relied on heuristic approaches involving discretizing atmospheric domains into parcels of equal mass and sorting them according to density at differing pressure levels to obtain a reference state. From a computational viewpoint, the discretized approach to computing MAPE is equivalent to finding the permutation of the actual state with the lowest total potential energy. Tailleux and Grandpeix (2004) characterized such a problem as an asymmetric travelling salesman problem, but recently it was realized by Hieronymus and Nycander (2015) that the computation of such a reference state was in fact a linear assignment problem that can be solved by using the Munkres algorithm (Munkres, 1957). Whilst the Munkres algorithm is exact, it is also computationally expensive, and therefore it is still desirable to use approximate algorithms for speed. The time taken for the algorithms to compute MAPE is detailed in the Appendix. The Munkres algorithm can be used when the atmospheric domain considered comprises a small number of parcels n, but the runtime of the algorithm increases as n^3 (Stansifer *et al.*, 2017), so it quickly becomes infeasible for large domains.

Approximate sorting algorithms have been employed to investigate the intensity of extratropical storm tracks (O'Gorman, 2010), using Lorenz's algorithm (Lorenz, 1979) to calculate MAPE, and the energetics of tropical cyclones (Wong *et al.*, 2016), using top-down and bottom-up algorithms. A review of existing approximate sorting algorithms is given by Stansifer *et al.* (2017), who discussed their accuracy compared with the exact Munkres algorithm. However, the comparison was made over only three test-case soundings. This showed that none of the approximate algorithms was able to compute the exact MAPE in every case, but clearly the small number of cases presented means that it is impossible to draw conclusions about the general relative performance of the algorithms and therefore it is difficult to know which is most useful to study atmospheric energetics.

It is also not certain that parcel-sorting algorithms calculate the most physically suitable form of MAPE. Finding the exact minimum enthalpy parcel rearrangement using the Munkres algorithm does not consider whether certain parcel movements may be restricted, for example, by the presence of Convective Inhibition (CIN). The bottom-up algorithm introduced by Wong *et al.* (2016) is designed to prevent the unrealistic release of Convective Available Potential Energy (CAPE) during sorting, but does not consider either CAPE or CIN directly in its computation.

In section 2, we briefly describe all the existing algorithms that have been designed to calculate MAPE. To investigate the possibility of using a more physically based approach to compute MAPE, we also develop an algorithm based on the relationship between CAPE and MAPE found by Emanuel (1994). As far as we are aware, this relationship has never been explored to investigate whether it can be used to obtain results similar to those of the parcel-sorting approaches.

We then apply all the MAPE algorithms to 3,130 soundings from the Atmospheric Radiation Measurement (ARM) station on Nauru and to 584 soundings from the ARM sites on the Southern Great Plains. This allows us to assess which of the approximate algorithms is likely to compute a MAPE close to the true value, and to investigate the variation in their accuracy over a large number of soundings. In section 3, we describe the data used for the assessment. Section 4 presents the results of the performance of the approximate algorithms against the Munkres algorithm, and compares their accuracy for the two locations. In section 5, we discuss how the results relate to what was previously known about the algorithms, and which algorithms are most suitable for practical application. We also discuss the implications of our results for the development of a satisfactory theory of APE for a moist atmosphere.

2 | ALGORITHMS FOR COMPUTING MAPE

In this section, we describe the algorithms that can be used to compute the reference state, and hence the MAPE, of an atmospheric sounding. We assume here that the sounding has been discretized into parcels of equal mass. To begin, we outline the Munkres algorithm, which finds the reference state corresponding to the exact minimum enthalpy rearrangement of the parcels. We then describe the parcel-sorting algorithms that have been designed to find approximations to the reference state. Due to their approximate nature, these methods are less computationally expensive than the Munkres algorithm, but their typical accuracy compared with the Munkres algorithm is unknown; this will be investigated in section 4. Finally we describe a method for calculating MAPE that does not rely on a sorting procedure, but instead makes use of the relationship between MAPE and CAPE, which was suggested by Emanuel (1994).

2.1 | Munkres algorithm

The Munkres algorithm (Munkres, 1957) may be used to obtain the exact minimum enthalpy rearrangement of a set of air parcels, by treating the computation of parcel reference pressures as a linear assignment problem (Hieronymus and Nycander, 2015; Stansifer *et al.*, 2017). This method first calculates a *cost matrix* $\mathbf{C}$, in which the entry c_{ij} is the enthalpy of the ith parcel at the jth pressure level. Using this cost matrix, the algorithm allocates parcels to the pressure levels resulting

in a minimized total enthalpy. This is done using the linear algebra procedure described by Munkres (1957), which tracks how difficult it is to find a low-enthalpy position for each parcel during the rearrangement process.

2.2 | Lorenz's algorithm

The first algorithm for approximating the minimized enthalpy reference state of a moist sounding was developed by Lorenz (1979). For a set of n parcels at pressures $p_1 < p_2 < \ldots < p_n$, this algorithm begins by calculating the virtual temperature that each parcel would have if it were lifted reversibly and adiabatically to p_1, denoted T_{v1}, and if it were similarly lifted to p_n, denoted T_{vn}. The algorithm first finds a parcel to assign to pressure level p_1, and then moves to progressively higher pressures. This assignment is determined as follows: at each level p_j, the unassigned parcels with the highest values of T_{v1} and T_{vn} are identified. If these values are maximized by the same parcel, this parcel is assigned to p_j. If the two identified parcels differ, then their virtual temperatures at $(p_j + p_{j+1})/2$ are calculated. The parcel with the higher T_v here is assigned to p_j. After n assignments are made in this way, all parcels will have been assigned a different reference pressure, thus determining the reference state. Equivalently, the specific volume may be maximized at each pressure rather than the virtual temperature, as has been done in our implementation.

2.3 | Randall and Wang's algorithm

Randall and Wang (1992) noted that it was possible for Lorenz's algorithm to return a negative MAPE and designed a similar algorithm that eliminated this problem. For pressure levels $p_1 < p_2 < \ldots < p_n$ as before, the procedure begins by labelling $p_A = p_1, p_B = p_n$. Once again, the virtual temperatures for all parcels are calculated as if they were lifted to p_A and p_B, and those parcels with the highest values of T_{vA} and T_{vB} are identified. At this point, the two methods diverge. Randall and Wang next compute the total atmospheric enthalpy for two situations: if the parcel with the highest T_{vA} were lifted to p_A, with any intermediate parcels shifted down one pressure level; and if the parcel with the highest T_{vB} were lifted to p_A and the intermediate parcels shifted down. Whichever of these configurations results in the lowest total enthalpy is accepted as the new rearrangement and p_A is redefined as $p_A = p_2$. The method proceeds until $p_A = p_B$.

2.4 | Top-down algorithm

The top-down algorithm was used to compute reference states in the study of APE in tropical cyclones by Wong *et al.* (2016). The performance of the top-down algorithm was also analysed by Stansifer *et al.* (2017), who referred to it as the "greedy algorithm". The top-down algorithm for n parcels proceeds as follows: all n air parcels are moved reversibly adiabatically to p_1, the lowest pressure in the sounding. Their densities at this pressure are calculated and the parcel with the lowest density is assigned to have p_n as its reference pressure. This parcel is then eliminated from sorting. The remaining $n - 1$ parcels are moved to p_2, and again their densities are calculated and the least dense parcel assigned to p_2. The algorithm continues in this way until all parcels have been assigned to a reference pressure level.

2.5 | Bottom-up algorithm

Bottom-up sorting works similarly to top-down sorting, but the parcels are first moved to the highest pressure p_n, assigning the parcel with the highest density to this level, and proceeding to lower pressure levels $p_{n-1}, p_{n-2} \ldots$. Bottom-up sorting was suggested by Wong *et al.* (2016) to limit the inclusion of CAPE in the definition of MAPE. This may be desirable in practice, since not all the CAPE present in the atmosphere will be released, for example due to the presence of CIN or subsidence.

2.6 | Divide-and-conquer algorithm

The divide-and-conquer algorithm was introduced by Stansifer *et al.* (2017). It is similar to top-down or bottom-up sorting, but all the parcels are initially moved to the middle pressure level p_m, where $m = \lfloor (n+1)/2 \rfloor$. The m parcels with the lowest density at this pressure are assigned to the subdomain $[p_1, p_m]$ and the $n - m$ parcels with the highest density are assigned to $[p_{m+1}, p_n]$. The algorithm then acts recursively on the two subdomains. In the three test cases analysed by Stansifer *et al.* (2017), the divide-and-conquer algorithm was found to perform well. It computed the exact minimum enthalpy reference state in one test case and close to the minimum in the other two, even when other approximate algorithms failed to capture significant proportions of the MAPE. However, since the divide-and-conquer algorithm is not an exact enthalpy minimization procedure, the fact that it was tested only on a small sample of atmospheric soundings means that we do not have sufficient evidence to exclude the possibility of atmospheric conditions, yet to be identified, for which it might perform poorly.

2.7 | Estimation from Convective Available Potential Energy

Rather than using a parcel-sorting algorithm to compute the vertical component of MAPE, it is natural to consider its relation to CAPE, since both are measures of the energy available for vertical motion in a sounding. This link was noted by Randall and Wang (1992), who referred to the vertical component of MAPE as Generalised CAPE (GCAPE), but did not explore the link between CAPE and GCAPE. Tailleux and Grandpeix (2004) suggested the existence of a functional relationship between CAPE and MAPE, which could

permit the inexpensive computation of MAPE. However, it is still not known how CAPE-based measures of atmospheric energetics compare with the Lorenz MAPE of Equation 1. Here, we outline an algorithm for calculating MAPE using the CAPE-dependent equations of Emanuel (1994), which we will compare (in section 4) with the MAPE computed by the sorting algorithms described above.

Emanuel (1994) supposes that MAPE is due solely to the presence of CAPE in a thin boundary layer of depth Δp_b. In this case, an approximation to the MAPE is given by

$$\text{MAPE} \approx \frac{\Delta p_b}{g}\left(\text{CAPE}_b - \frac{1}{2}\Delta p_b \overline{p^{\kappa-1}} p_0^{-\kappa} R_d \Delta\theta_v\right), \qquad (2)$$

where CAPE_b is the mean CAPE in the boundary layer, $\kappa = R_d/c_{pd}$, and $\Delta\theta_v$ is the change in virtual potential temperature between the top of the boundary layer at $p_{b,top}$ and the boundary layer's level of neutral buoyancy, p_{LNB}. The overbar denotes a θ_v-weighted average from $p_{b,top}$ to p_{LNB}. The first term of Equation 2 corresponds to the release of CAPE when the boundary layer rises upwards to its LNB. The second term accounts for the energy change that occurs as a result of the remaining air parcels descending by Δp_b. We will henceforth refer to MAPE calculated using Equation 2 as the *Emanuel MAPE*.

We compute the Emanuel MAPE by calculating the value of Equation 2 for Δp_b depths ranging from 0 to 150 mb and selecting the maximum value of MAPE returned by any of these Δp_b values. We increment Δp_b simply by including the next lowest parcel in the sounding. Theoretically, it would be possible to use smaller increments in Δp_b and include fractions of parcels in the boundary layer. We have not done this because the sorting algorithms discussed earlier in this section are only able to rearrange whole parcels, so allowing this CAPE-based algorithm to lift only whole parcels provides a fairer comparison of the MAPE.

To compute the boundary-layer CAPE, CAPE_b, we use a parcel with a value of θ given by the pressure-weighted mean of θ in the boundary layer and q given by the mean q in the boundary layer. The CAPE is then

$$\text{CAPE}_b = \int_{p_{LNB}}^{p_i} \left(\alpha_p - \alpha_e\right)\, dp, \qquad (3)$$

where α_p is the specific volume of the parcel when it is lifted reversibly adiabatically, and α_e is the environmental specific volume. The parcel is lifted from its initial position p_i, which we take to be the bottom of the boundary layer (i.e. the surface), to its highest level of neutral buoyancy.

3 | DATA

To calculate the MAPE of a sounding, the sorting algorithms outlined in section 2 require the input of the temperature, pressure and total specific humidity profiles. The atmospheric profiles used to compare the algorithms are data obtained through soundings from the Atmospheric Radiation Measurement (ARM) Program (Stokes and Schwartz, 1994). We assume that the total specific humidity q_T in the soundings is equal to the specific humidity q, i.e. that no liquid water is present in the atmosphere. This widens our choice of data, since we do not require liquid water measurements, and is justified, since we do not expect large quantities of liquid water to be residing in the atmosphere for long periods of time.

We have used soundings from Nauru dating from April 1, 2001–August 16, 2006. These soundings contain data that have been interpolated on to 5 mb pressure levels and quality-controlled as described by Holloway and Neelin (2009). We take all soundings with at least 150 valid measurements of temperature and specific humidity, for which the valid measurements span at least the interval from 1,000–100 mb. Any missing temperature or humidity measurements are filled in by linear interpolation. This results in 3,130 soundings for which we can use the sorting algorithms to compute the MAPE in the 1,000–100 mb layer using 181 parcels of 5 mb depth.

To verify whether the performance of the algorithms is significantly affected if the soundings are from a different location, we have also used soundings from the ARM Southern Great Plains (SGP) sites during the Intensive Observation Period from June 4, 1997–July 7, 1997; this dataset is the one used by Tailleux and Grandpeix (2004). The pressure levels measured in the SGP soundings vary, so we select those soundings that have at least 2,000 valid measurements extending from 950 to 100 mb, and no more than 50 invalid measurements, resulting in a total of 584 suitable soundings. We interpolate the temperature and humidity data linearly on to 5 mb spaced pressure levels between 950 and 100 mb (resulting in 171 parcels per sounding), to match the parcel mass of the Nauru soundings. The results of section 4 were found to be insensitive to interpolating to a greater number of parcels.

The median profiles of temperature T and specific humidity q are shown for each location in Figure 1, along with the 25th–75th percentiles (dark shading) and 10th–90th percentiles (light shading). The profiles are similar in the two locations, with Nauru soundings exhibiting higher moisture at lower levels (this is reasonable because we have kept Nauru data at higher pressure levels, whereas there were insufficient measurements to do so for the SGP data). The Nauru soundings also show colder temperatures at high altitude. It is notable that there is very little variation about the median Nauru temperature profile and therefore differences in the ability of the algorithms to calculate MAPE accurately here will be due mostly to differences in humidity profiles between the soundings.

4 | COMPARISON OF ALGORITHMS

All the sorting algorithms discussed in section 2 were used to calculate the MAPE of each of the 3,714 ARM

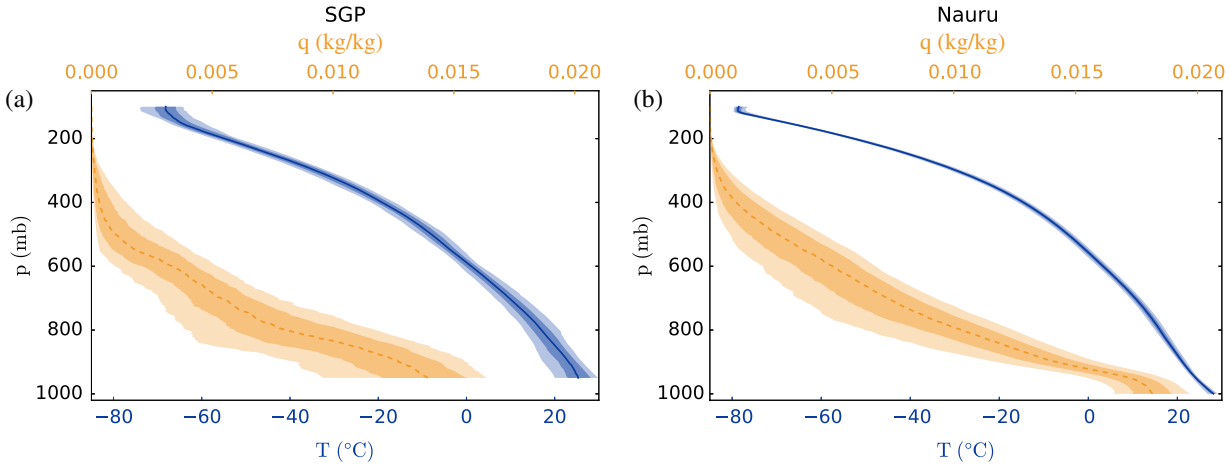

FIGURE 1 Median profiles of temperature T (°C, solid line) and specific humidity q (kg/kg, dashed line) for the ARM soundings from (a) Nauru and (b) the SGP used to test the MAPE sorting algorithms. The dark shading shows the 25th–75th percentile and the light shading the 10th–90th percentile [Colour figure can be viewed at wileyonlinelibrary.com]

soundings described in section 3. To summarize, these algorithms are as follows: Munkres, Lorenz, Randall and Wang, top-down, bottom-up, divide-and-conquer, and Emanuel. We use the implementation of the Munkres algorithm from Stansifer *et al.* (2017), modified to compute the cost matrix using a vectorized iterative method as described in the Appendix. For the other algorithms we use our own implementations, incorporating the iterative method approach. The MAPE found by the Munkres algorithm is the maximum MAPE computable by sorting; in the following section, we compare this with the MAPE computed by the approximate algorithms to assess their accuracy.

To quantify the accuracy of each algorithm, we define the percentage relative difference in MAPE as

$$D_{\mathrm{R}} = \begin{cases} \frac{|\mathrm{MAPE}_{\mathrm{Munk}} - \mathrm{MAPE}_{\mathrm{app}}|}{\mathrm{MAPE}_{\mathrm{Munk}} + \mathrm{MAPE}_{\mathrm{app}}} \times 100, & \text{if } \mathrm{MAPE}_{\mathrm{app}} \geq 0 \\ 100, & \text{otherwise,} \end{cases} \tag{4}$$

where $\mathrm{MAPE}_{\mathrm{app}}$ is the MAPE computed by the approximate algorithm and $\mathrm{MAPE}_{\mathrm{Munk}}$ is the MAPE computed by the Munkres algorithm. This provides a measure of the amount of MAPE that each approximate algorithm fails to capture. All the approximate algorithms that are based on sorting parcels must compute a MAPE lower than the value computed by the Munkres algorithm, while the Emanuel MAPE may exceed this value.

The distributions of D_{R} for each approximate algorithm across all the soundings are displayed in Figure 2. It is clear from these results that Randall and Wang's algorithm is the most accurate of the six approximate algorithms, with a median D_{R} of 0.0077% for the SGP soundings and 0.0015% for the Nauru soundings. However, there remain outlying cases in which even Randall and Wang's algorithm fails to capture a large proportion of the MAPE. Of the other sorting algorithms, only divide-and-conquer provides a reasonable approximation to the Munkres algorithm, with a median D_{R} of 1.9% across the soundings from the SGP and 3.0% across those from Nauru.

The bottom-up algorithm fails to capture the majority of the MAPE in most cases; this is expected, since the sorting procedure is designed to limit the release of CAPE from buoyant surface parcels and hence should result in a smaller vertical component of MAPE. There is still a wide range of D_{R} across the soundings, particularly in the SGP case.

Lorenz's algorithm exhibits a very similar D_{R} distribution to the top-down algorithm, with both medians around 50% for each location. The poor accuracy is due largely to the fact that both the top-down and Lorenz procedures frequently compute a negative MAPE, which is unphysical by the definition of MAPE as the difference between the enthalpies of the atmosphere and its rearranged, minimized total enthalpy state. A negative value of MAPE simply means that the "minimized" enthalpy reference state computed by the approximate algorithm in fact has a higher enthalpy than the real atmospheric state. The top-down and Lorenz algorithms compute a negative MAPE for between 30% and 40% of the soundings, for both the SGP and Nauru data. To illustrate why this occurs, Figure 3 shows the difference in enthalpy for each parcel between the original sounding and the reference state, for the Nauru sounding measured at 1200 UTC on September 24, 2001. Referring to Equation 1, the total MAPE of the sounding will be equal to the sum of this enthalpy difference $h - h_{\mathrm{ref}}$ over all parcels. The circular markers show the enthalpy difference when using the Munkres algorithm, which computes a MAPE of 6.87 J/kg. The crosses show the enthalpy difference using top-down sorting, which results in a MAPE of −19.9 J/kg. It is evident that the negative MAPE is a result of the top-down algorithm lifting parcels to low reference pressures (150–250 mb), in such a way that these particular parcels experience a large decrease in enthalpy, but the parcels at higher reference pressures exhibit a slight increase in enthalpy. When computing the overall MAPE, the smaller enthalpy decreases

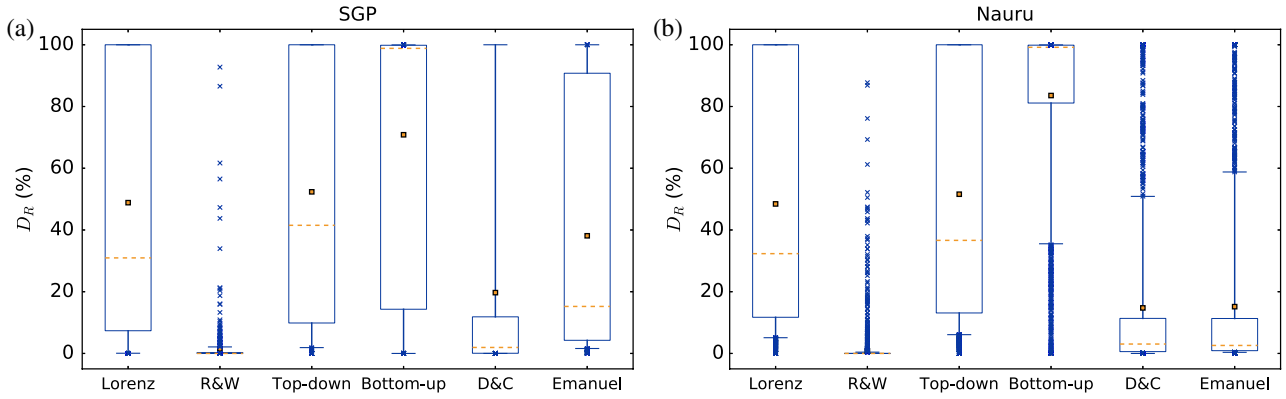

FIGURE 2 Box plots of the percentage relative difference D_R, as defined in Equation 4, between each approximate algorithm and the exact Munkres algorithm for soundings from (a) Nauru and (b) the SGP. The dashed lines denote the median D_R across the soundings, the squares the mean D_R, the boxes the 25th–75th percentiles and the whiskers the 10th–90th percentiles. Crosses represent soundings with outlying D_R values [Colour figure can be viewed at wileyonlinelibrary.com]

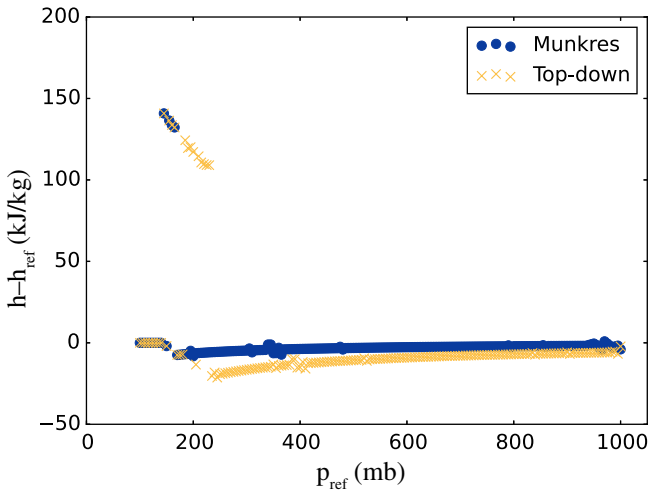

FIGURE 3 Difference between the enthalpy of each parcel from the September 24, 2001 1200 UTC Nauru sounding and its enthalpy in the reference state, plotted against the parcel's pressure in the reference state. Reference states are calculated using both the Munkres algorithm, which computes the exact MAPE, and the top-down algorithm, which computes a negative MAPE [Colour figure can be viewed at wileyonlinelibrary.com]

over many parcels at high reference pressures outweigh the large enthalpy decreases of the few parcels at low reference pressures. Parcels at high reference pressures in the Munkres reference state also show a decrease in enthalpy, but, since this decrease is smaller than in the top-down case, the net MAPE remains positive. The divide-and-conquer algorithm also computes negative MAPE for some soundings, but this does not occur as frequently as for the Lorenz and top-down algorithms (13% of SGP soundings, 6.5% of Nauru soundings).

In contrast, a non-negative value of MAPE is systematically returned by the bottom-up, Randall and Wang, and Emanuel algorithms on all soundings tested. This is expected for the latter two algorithms, since these are specifically designed to ensure that the enthalpy of the computed reference state is less than that that of the original configuration of the fluid parcels. Although no such constraint is imposed for the bottom-up algorithm, negative values are highly unlikely, because bringing fluid parcels adiabatically to higher pressures makes it nearly impossible for the fluid parcels to ever become saturated and release latent heat. It follows that, in effect, the bottom-up algorithm constructs its reference state by ordering the fluid parcels according to their virtual potential temperature θ_v, and hence any positive value returned primarily reflects the presence of an inversion in θ_v in the original sounding. For the unsaturated atmospheric soundings considered, existing inversions are in general not directly associated with the presence of CAPE, so that the positive value returned by the bottom-up algorithm in those cases is more akin to a "dry" APE than a "moist" one. In the numerical simulations studied by Wong *et al.* (2016), on the other hand, the bottom-up algorithm can occasionally pick up the APE of fluid parcels with CAPE everywhere where a local sounding has become absolutely unstable as the result of fluid parcels having overcome their CIN and reached their level of free convection. This behaviour supports the idea that, in 2D and 3D domains, the bottom-up algorithm will primarily access the horizontal component of APE and will access any "vertical" APE only when readily available to fluid parcels when the fluid has become absolutely unstable, which seems to be the simplest way to avoid including CAPE as part of MAPE when the presence of CIN or subsidence makes its release unlikely.

In general, the D_R distributions of the sorting-based algorithms are similar for the two locations, suggesting that we do not expect the optimum choice of sorting algorithm to change depending on the typical local atmospheric conditions. However, the accuracy of the Emanuel MAPE is very different between the locations. For the Nauru soundings, it has a median D_R of 2.6%, which is comparable to the divide-and-conquer sorting algorithm, showing that the Emanuel algorithm would be a sensible choice for estimating MAPE. In contrast, its median D_R over the SGP is 15%, which is much less accurate than either the divide-and-conquer or Randall and Wang algorithms, and so the Emanuel method would not be a good practical choice for computing MAPE

in this environment. To investigate why this difference in accuracy occurs, Figure 4 shows scatter plots of the Emanuel MAPE against the MAPE computed using the Munkres algorithm, for each location. The dashed lines display the best linear fit to the data; for the Nauru soundings we find a correlation coefficient of $r = 0.990$, while for the SGP soundings the correlation coefficient is slightly poorer, at $r = 0.977$, as expected from the higher median value of D_R. The high correlation indicates that most of the MAPE present in the ARM soundings corresponds to the CAPE of near-surface parcels, in line with the assumption of Emanuel (1994).

We can see from Figure 4 that the poorer correlation for the SGP compared with the Nauru soundings is mostly due to a number of SGP soundings that have very low Emanuel MAPE, but values of Munkres MAPE up to 60 J/kg. We find that these discrepancies arise where an unstable layer that is elevated from the surface is present in the sounding. Since the Emanuel algorithm we have used assumes that the CAPE-containing boundary layer begins at the surface, these elevated instabilities are not correctly captured. This issue could be solved by designing an algorithm that varied both Δp_b and the pressure of the boundary-layer bottom, although this would increase the computational expense. The main advantage of the Emanuel algorithm is that it provides greater physical insight into how MAPE can be converted to kinetic energy via convection, rather than relying on the physically unconstrained rearrangements of a sorting algorithm, as will be discussed further in section 5.

5 | DISCUSSION

The results presented in section 4 allow us to make a more informed assessment of which algorithms are most suitable for the computation of MAPE, based on an analysis of a wider range of soundings than in previous studies. The key challenge for MAPE algorithms stems from the fact that MAPE is ultimately a residual arising from the positive work due to the release of CAPE minus the negative work due to compensating subsidence. As a result, sorting a vertical sounding according to decreasing density, which is the approach underlying the majority of algorithms, may occasionally result in a reference state with a larger potential energy than the actual state, if the negative work exceeds the positive work. This is in contrast to the case of a dry atmosphere, for which sorting the actual state according to potential temperature always returns the state of minimum potential energy. Without an explicit procedure to forbid it, most heuristics for computing MAPE are bound to return a negative value in some cases. How frequently this might occur in practice has not been explored previously, but it is established here to be a common problem for some of the algorithms for the particular soundings analysed.

We have found that the Lorenz and top-down algorithms have nearly identical levels of accuracy, which we did not anticipate. However, our results also indicate that both algorithms are so prone to returning a negative MAPE that they are not suited to practical application. While the accuracy of such algorithms has been questioned previously by Randall and Wang (1992) and Stansifer *et al.* (2017), this is the first time that their limitation is established for a wide range of real-world data. Our results therefore confirm that these algorithms are poorly suited to the computation of MAPE in practice, and hence that their use should be avoided.

The algorithm introduced by Randall and Wang (1992) was found to be a good predictor of the exact MAPE across the soundings studied. The relative difference between the Randall and Wang and Munkres algorithms exceeded 10% for only 2% of the soundings studied. This algorithm also benefits from the fact that it is designed specifically never to return a negative MAPE. However, it is the most computationally expensive of the approximate algorithms; for a sounding with a small number of parcels, it takes even longer to run than the Munkres algorithm (see the Appendix).

As was outlined in section 4, the divide-and-conquer algorithm is the only other approximate sorting method showing reasonable accuracy over the soundings studied here. It is also the fastest of the approximate algorithms (see the Appendix). We therefore conclude that the divide-and-conquer algorithm is the best option for the approximation of MAPE, since it offers a balance between accuracy and speed, as was suggested by Stansifer *et al.* (2017). On the other hand, the relative difference between the divide-and-conquer and Munkres algorithms is greater than 50% for 11% of the total soundings studied, and divide-and-conquer sorting may result in a negative MAPE. This clearly suggests that the three test cases analysed by (Stansifer *et al.*, 2017), for which the algorithm was found to perform well, might be special cases, which are not sufficiently representative of the variety of situations that can be encountered in nature.

We have also demonstrated the feasibility and good performance of an algorithm exploiting Emanuel (1994)'s theoretical expression for MAPE, which does not involve any form of sorting nor require discretizing the vertical sounding into parcels of equal mass, and which by construction always returns a positive value of MAPE.

So far, the implicit assumption of the present study and others has been that it is legitimate or most useful to define the APE of a moist atmosphere in terms of the reference state that defines the absolute minimum in potential energy, but this is not necessarily the case. For a moist atmosphere, it is a priori possible to construct alternative sorted reference states that define only a local minimum in potential energy. Although such reference states would result in a lesser global value of APE, it is unclear why this would necessarily invalidate their use. In tropical cyclones, for instance, numerical simulations reveal that boundary-layer parcels away from the eyewall may have CAPE, the release of which is suppressed by subsidence in that region, as pointed out by Wong *et al.*

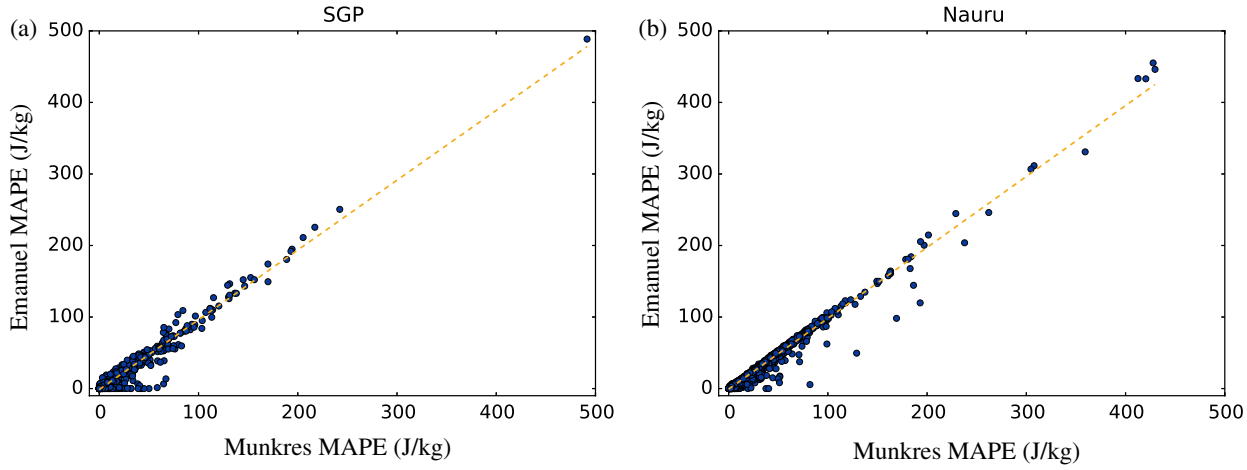

FIGURE 4 Comparison of MAPE calculated using the Emanuel algorithm with the exact MAPE calculated using the Munkres algorithm for all soundings from (a) Nauru and (b) the SGP. The dashed line in each case shows the linear best fit to the data [Colour figure can be viewed at wileyonlinelibrary.com]

(2016). Since the CAPE of such parcels can rarely, if ever, be released, it is unclear why it should be included in the definition of a tropical cyclone APE, as will normally be the case if the reference state defining a global potential energy minimum is selected. From a practical viewpoint, it is important to remark that the choice of reference state affects the overall value of APE as well as its diabatic generation rate $G(APE)$, but neither the energy conversion between APE and kinetic energy nor the general form of the APE evolution equation, given by

$$\frac{d\mathrm{APE}_i}{dt} = C(KE, APE) + G(APE)_i, \tag{5}$$

where the index i is used to indicate dependence on the reference state chosen. Equation 5 states that the conversion $C(KE, APE)$ between kinetic energy and APE always appears as a residual between the APE storage term $d\mathrm{APE}_i/dt$ and the APE generation rate $G(APE)_i$ (see Pauluis, 2007 for a discussion of how moist processes may affect the latter). From a theoretical viewpoint, Equation 5 represents a balance between three terms, of which the storage term is the least interesting or meaningful. For this reason, Wong *et al.* (2016) argued that the reference state should be chosen so as to minimize the storage term, in order potentially to make it possible to predict the APE/KE conversion from knowledge of the APE generation rate. In this regard, Wong *et al.* (2016) found the use of the bottom-up sorted reference state to yield a lower storage term that the top-down sorted reference state, but more research is required to establish whether this can be regarded as a general result.

Given the computational and conceptual difficulties entailing their use, it is important to question whether sorting algorithms are really needed to study the energetics of a moist atmosphere. The idea that simpler alternatives might exist is indeed justified by the fact that some recent APE studies successfully moved away from the use of sorting algorithms by resorting to approaches using probability density functions instead, as in the case of Saenz *et al.* (2015), itself an extension of Tseng and Ferziger (2001), although it is unclear how such a method could be applied to a moist atmosphere. Also, it has long been known from the works of Andrews (1981) and Holliday and Mcintyre (1981), which were recently generalized for multicomponent compressible stratified fluids by Tailleux (2018), that it is possible to construct a local theory of APE based on an arbitrary reference state defined by a reference pressure $p_0(z,t)$ and specific volume $\alpha_0(z,t)$ in hydrostatic equilibrium. Based on Tailleux (2013) and Novak and Tailleux (2018), this would lead one to define the APE density for a moist atmosphere as the work that a fluid parcel needs to perform to move from its reference pressure p_r to its actual pressure p, namely,

$$e_\mathrm{a}(\theta_l, q_\mathrm{T}, p, t) = \int_{p_\mathrm{r}}^{p} \left[\alpha(\theta_l, q_\mathrm{T}, p') - \alpha_0(p', t)\right] dp', \tag{6}$$

where θ_l is liquid potential temperature and q_T is total water content. An alternative formulation for APE density in a compressible atmosphere, based on modified potential temperature, has been proposed by Peng *et al.* (2015). In contrast to what is often assumed, a sorting algorithm is not required to calculate the reference pressure p_r. Indeed, as shown by Tailleux (2013), if $\alpha_0(p,t)$ is known at all times as a function of pressure, p_r can be simply estimated by solving the Level of Neutral Buoyancy (LNB) equation:

$$\alpha(\theta_l, q_\mathrm{T}, p_\mathrm{r}) = \alpha_0(p_\mathrm{r}, t). \tag{7}$$

This corresponds to the use of an LNB in the Emanuel MAPE algorithm, demonstrating the link between local and global approaches to APE.

Illustrations of how to construct energy budgets in the oceans and dry atmosphere in the case where the reference density profile is defined from a horizontal or isobaric average are discussed by Tailleux (2013) and Novak and Tailleux (2018), respectively. These recent developments, combined with the physical insights brought about by Emanuel (1994)'s theoretical expression for MAPE, suggest that a satisfactory theory of available potential energy for a moist atmosphere, which has been lacking so far, might be at hand, provided that

one moves away from sorting algorithms altogether, as we hope to demonstrate in subsequent studies.

ACKNOWLEDGEMENTS

B.L.H. is funded by NERC as part of the SCENARIO Doctoral Training Partnership (NE/L002566/1). We thank Chris Holloway and Pier Luigi Vidale for interesting discussions. We thank Eric Stansifer for making the code to compute MAPE using the Munkres algorithm available and for a review that helped to clarify many points in the manuscript. All data and algorithms used to obtain the results presented in this article are available upon request (b.l.harris@pgr.reading.ac.uk).

REFERENCES

Andrews, D.G. (1981) A note on potential energy density in a stratified compressible fluid. *Journal of Fluid Mechanics*, 107, 227–236. https://doi.org/doi:10.1017/S0022112081001754.

Emanuel, K.A. (1994) *Atmospheric Convection*. Oxford: Oxford University Press.

Hieronymus, M. and Nycander, J. (2015) Finding the minimum potential energy state by adiabatic parcel rearrangements with a nonlinear equation of state: an exact solution in polynomial time. *Journal of Physical Oceanography*, 45(7), 1843–1857. https://doi.org/doi:10.1175/JPO-D-14-0174.1.

Holliday, D. and Mcintyre, M.E. (1981) On potential energy density in an incompressible, stratified fluid. *Journal of Fluid Mechanics*, 107, 221–225. https://doi.org/doi:10.1017/S0022112081001742.

Holloway, C.E. and Neelin, J.D. (2009) Moisture vertical structure, column water vapor, and tropical deep convection. *Journal of the Atmospheric Sciences*, 66(6), 1665–1683. https://doi.org/doi:10.1175/2008JAS2806.1.

Langtangen, H.P. (2016) *A Primer on Scientific Programming with Python*, 5th edition. Berlin: Springer.

Lorenz, E.N. (1955) Available potential energy and the maintenance of the general circulation. *Tellus*, 7(2), 157–167. https://doi.org/doi:10.1111/j.2153-3490.1955.tb01148.x.

Lorenz, E.N. (1978) Available energy and the maintenance of a moist circulation. *Tellus*, 30(1), 15–31. https://doi.org/doi:10.3402/tellusa.v30i1.10308.

Lorenz, E.N. (1979) Numerical evaluation of moist available energy. *Tellus*, 31(3), 230–235. https://doi.org/doi:10.1111/j.2153-3490.1979.tb00901.x.

Munkres, J. (1957) Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1), 32–38.

Novak, L. and Tailleux, R. (2018) On the local view of atmospheric available potential energy. *Journal of the Atmospheric Sciences*, 75(6), 1891–1907. https://doi.org/doi:10.1175/JAS-D-17-0330.1.

O'Gorman, P.A. (2010) Understanding the varied response of the extratropical storm tracks to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 107(45), 19176–19180. https://doi.org/doi:10.1073/pnas.1011547107.

Pauluis, O. (2007) Sources and sinks of available potential energy in a moist atmosphere. *Journal of the Atmospheric Sciences*, 64(7), 2627–2641. https://doi.org/doi:10.1175/JAS3937.1.

Peng, J., Zhang, L. and Zhang, Y. (2015) On the local available energetics in a moist compressible atmosphere. *Journal of the Atmospheric Sciences*, 72(4), 1551–1561. https://doi.org/doi:10.1175/JAS-D-14-0181.1.

Randall, D.A. and Wang, J. (1992) The moist available energy of a conditionally unstable atmosphere. *Journal of the Atmospheric Sciences*, 49(3), 240–255. https://doi.org/doi:10.1175/1520-0469(1992)049.

Saenz, J.A., Tailleux, R., Butler, E.D., Hughes, G.O. and Oliver, K.I. (2015) Estimating Lorenz's reference state in an ocean with a nonlinear equation of state for seawater. *Journal of Physical Oceanography*, 45(5), 1242–1257. https://doi.org/doi:10.1175/JPO-D-14-0105.1.

Stansifer, E.M., O'Gorman, P.A. and Holt, J.I. (2017) Accurate computation of moist available potential energy with the Munkres algorithm. *Quarterly Journal of the Royal Meteorological Society*, 143(702), 288–292. https://doi.org/doi:10.1002/qj.2921.

Stokes, G.M. and Schwartz, S.E. (1994) The Atmospheric Radiation Measurement (ARM) Program: programmatic background and design of the cloud and radiation test bed. *Bulletin of the American Meteorological Society*, 75(7), 1201–1221. https://doi.org/doi:10.1175/1520-0477(1994)075.

Tailleux, R. (2013) Available potential energy density for a multicomponent Boussinesq fluid with arbitrary nonlinear equation of state. *Journal of Fluid Mechanics*, 735, 499–518. https://doi.org/doi:10.1017/jfm.2013.509.

Tailleux, R. (2018) Local available energetics of multicomponent compressible stratified fluids. *Journal of Fluid Mechanics*, 842, R1. https://doi.org/doi:10.1017/jfm.2018.196.

Tailleux, R. and Grandpeix, J.Y. (2004) On the seemingly incompatible parcel and globally integrated views of the energetics of triggered atmospheric deep convection over land. *Quarterly Journal of the Royal Meteorological Society*, 130(604), 3223–3243. https://doi.org/doi:10.1256/qj.03.140.

Tseng, Y.H. and Ferziger, J.H. (2001) Mixing and available potential energy in stratified flows. *Physics of Fluids*, 13(5), 1281–1293. https://doi.org/doi:10.1063/1.1358307.

Wong, K., Tailleux, R. and Gray, S.L. (2016) The computation of reference state and APE production by diabatic processes in an idealized tropical cyclone. *Quarterly Journal of the Royal Meteorological Society*, 142(700), 2646–2657. https://doi.org/doi:10.1002/qj.2854.

How to cite this article: Harris BL, Tailleux R. Assessment of algorithms for computing moist available potential energy. *Q J R Meteorol Soc.* 2018;144:1501–1510. https://doi.org/10.1002/qj.3297

APPENDIX: ALGORITHM RUNTIMES

The most time-consuming stage of each of the algorithms used to compute MAPE in this work is the calculation of the temperatures of air parcels when they are moved reversibly adiabatically to a given pressure level. For the approximate sorting algorithms, this calculation is required in order to sort the parcels by density (or equivalently specific volume) at each pressure level of the sorting process. The Emanuel algorithm requires the calculation in order to find the specific

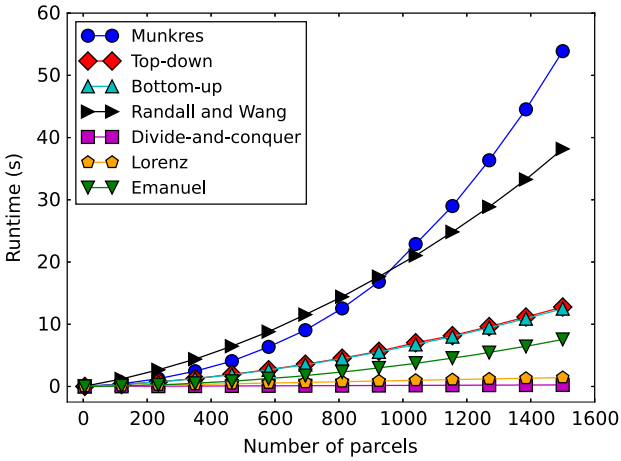

FIGURE A1 Time taken for each MAPE algorithm to compute the MAPE of an n-parcel sounding linearly interpolated from the data at Nauru on April 1, 2001, 1200 UTC [Colour figure can be viewed at wileyonlinelibrary.com]

volume of lifted parcels for the calculation of CAPE (see Equation 3). For the Munkres algorithm, it is required in order to compute the *cost matrix* **C**, where c_{ij} is the enthalpy of the ith parcel if it were moved to the jth pressure level.

Our algorithms are implemented using a vectorized bisection method to calculate the temperatures of many parcels at once (see Langtangen, 2016 for details on both vectorization with NumPy and an example of the bisection method). To illustrate the speed of the algorithms, Figure A1 shows the time taken for each algorithm to compute MAPE for a sounding with a varying number of parcels. For each number of parcels n we create the test sounding by taking the ARM Nauru data from 1200 UTC on April 1, 2001 and linearly interpolating it to n pressure levels.

As expected, the exact Munkres procedure is slowest for large numbers of parcels. We note that the bisection method is poorly suited for application to the Lorenz algorithm, since the specific volume must be repeatedly calculated for just two parcels at the midpoint between pressure levels. This results in our Lorenz algorithm being slower than previous implementations: it is slower by almost a factor of 10 than the implementation of Stansifer *et al.* (2017), whereas our other algorithms are similar in speed to their previous implementations. The speed of the divide-and-conquer algorithm shown here, combined with the accuracy illustrated in section 4, makes it a good choice amongst sorting algorithms for computing MAPE.