We thank the referees and Dr. Jimenez for their very thoughtful and thorough reviews. We have addressed their comments (numbered, below), with referee comments in quotes and italics, and our responses immediately after in plain text.

**Referee #1**

1. “Page 27,581, line 21. The impacts of aerosol pH on halogen-radical chemistry are not limited to coastal regions. For example, the cited investigation by Pszenny et al. (2004) evaluated influences of pH in open ocean air at Hawaii. Numerous other publications have reported results from measurement and modeling studies that characterized the pH-dependence of halogen activation over the open ocean.”

   We have made the following change to clarify this point: “…ozone formation in marine and coastal regions…”

2. “Page 27,581, line 23. Thornton et al. (2010) speculated about the influence of aerosol pH on Cl-radical chemistry in continental regions but a subsequent investigation in which Thornton and colleagues were involved confirmed the importance of solution pH and the associated thermodynamics of HCl phase partitioning in sustaining CINO2 production (Brown et al., 2013; Riedel et al., 2013; Young et al., 2013).”

   We have updated the text to read: “Recent evidence has demonstrated that aerosol pH is also a critical parameter influencing halogen chemistry in continental locations (Thornton et al., 2010; Brown et al., 2013; Young et al., 2013).”

3. “Page 27,582, lines 6-8. I know what the authors mean but the suggestion that “analysis procedures such as dilution in water prevent” characterization of aerosol solution pH is potentially misleading. As reported in several of the cited papers, direct measurements of H+ in dilute extracts of aerosol samples can be reliably extrapolated to the pH of concentrated aerosol solutions but clearly H+ concentrations do not scale proportionately with dilution. I encourage the authors to clarify this point.”

   We have made the following change to clarify this point: “Measurements are also challenged by the non-conservative nature of H+: due to buffering effects and the partial dissociation of weak acids, H+ concentrations do not scale in proportion to the level of dilution.”

4. “Page 27,582, line 26. The term “strong acidity” is not synonymous with “free (or ionized) acidity” as suggested here. Strong acidity refers to H+ contributed by acids that are essentially completely dissociated at a given pH (i.e., with pKas « solution pH). For example, in a dilute solution at pH 4, strong acidity would correspond to that contributed by sulfuric, nitric, and hydrochloric acids. Weak acidity refers to H+ contributed by acids that are partially dissociated at a given pH (i.e., with pKas in the range of solution pH). Examples for a dilute solution at pH 4 include formic and acetic acids. Free acidity refers to ionized H+ at a given pH and total acidity refers to ionized H+ plus undissociated acids. The text should be clarified and, for the benefit of readers who may not be familiar with the term, “total acidity” should be defined.”
We have changed the text to read: “This definition can take several forms, including aerosol strong acidity (H\(^+\) contributed by strong acids that dissociate completely at most pH), free acidity (dissociated H\(^+\)), or total acidity (includes free H\(^+\) and the undissociated H\(^+\) bound to weak acids), typically defined by the measurement approach.”

5. “Page 27,583, lines 2-6. The discussion here seems unnecessarily convoluted. For example, it is unclear what the authors mean by “The true quantity, however, driving much of the processes described above is pH; H\(^+\) is insufficient to truly distinguish the acidity as it scales with aerosol mass.” What is a “true quantity” and how is it used to “truly distinguish” anything in the context of scaling “with aerosol mass.” As indicated in equation 1, solution pH is simply another way of expressing the aqueous H\(^+\) activity. Here and elsewhere, the authors should clearly distinguish between solution pH or H\(^+\) per unit volume of liquid water versus H\(^+\) per unit volume of air.”

We have updated the text as follows: “The major difference between aerosol pH and the proton loading is that pH is the H\(^+\) concentration per liquid water volume (i.e., aerosol water) while the aerosol proton loading is the H\(^+\) concentration per unit volume of air. Aerosol pH is the parameter of interest for the atmospheric phenomena described above, but the proton loading is often treated as a surrogate for pH.”

6. “Page 27,583, lines 12-13 and Eq. 2. I encourage the authors to clearly state the complete underlying assumptions involved in the ion-balance approach and to evaluate results in the context of each.”

7. “Assumption 1. This approach is reliable if and only if all ionic species other than H\(^+\) are measured without bias. The approach yields biased results if (1) any ionic constituents other than H\(^+\) that contributed significantly to ion balances are not measured or (2) any of the measured ionic concentrations were inaccurate. In the context of discussing unmeasured organic anions later in the paragraph, the authors should point out that significant concentrations of unmeasured organic anions would introduce significant positive bias in H\(^+\) inferred from ion balances based on measured species.”

8. “Assumption 2. All ionic species other than H\(^+\) are measured at relatively high precision. For many types of aerosols, H\(^+\) is a minor ionic constituent of the aerosol and consequently H\(^+\) concentrations typically correspond to small differences between relatively large sums of anions and cations. In such cases, calculated H\(^+\) concentrations based on Eq. 2 will often fall within the range of accumulated analytical uncertainties for the many measured ions thereby yielding insignificant results.”

9. “Assumption 3. Reliable extrapolation of H\(^+\) concentrations inferred from ion imbalances in dilute analytical solutions (e.g., the PILS sampling medium or in water extracts of aerosols sampled on filters or impaction substrates) to the pH of concentrated aerosol solutions accounts for the large differences in liquid water volume, the buffering effects of the H\(^+\)+SO\(_4^{2-}\)$HSO\(_4^{-}\)$ equilibrium (e.g., Keene et al., 2002, 2004), and differences between concentrations and activities (Young et al., 2013).”
We have added a figure (new Figure 8) and the following clarifications and discussions: “Neglecting this effect (of dissociated organic acids) can lead to biases in the inferred H⁺.”

“All of the measured inorganic species (Section 2.5) were included in the calculation according to Equation 2 but dissociated organic acids were neglected.”

“Finally, another major limitation in estimating the aerosol H⁺ loading from the ion balance is the small H⁺ levels relative to analytical uncertainty. Figure 8a shows the relative H⁺ uncertainty (%) vs. the H⁺ level inferred from the ion balance (neq m⁻³). The H⁺ uncertainty was calculated using a standard propagation of error technique (Harris, 1999):

\[ u_{H^+} = \left( u_{Na^+}^2 + u_{NH_4^+}^2 + u_{Mg^{2+}}^2 + u_{Ca^{2+}}^2 + u_{Cl^-}^2 + u_{NO_2^-}^2 + u_{NO_3^-}^2 + u_{SO_4^{2-}}^2 \right)^{1/2} \]

Where \( u_{H^+} \) is the absolute uncertainty in H⁺ (neq m⁻³), and the terms on the right hand side represent the absolute uncertainties in each inorganic species in neq m⁻³ (Na⁺, NH₄⁺, Mg²⁺, Ca²⁺, Cl⁻, NO₂⁻, NO₃⁻, and SO₄²⁻). Figure 8a shows that the relative H⁺ uncertainty \( (u_{H^+}/H^+) \) grows with decreasing H⁺, as one would expect. A frequency distribution of the H⁺ uncertainty from the MILAGRO study shows that more than 20% of the H⁺ calculations had an associated uncertainty higher than 50%, and more than 10% of the H⁺ calculations had an associated uncertainty higher than 100%.

These results demonstrate numerous problems with the ion balance and strongly suggest that it is inadequate to represent the pH of atmospheric particles. The ion balance may provide a qualitative indication of acidic conditions when anions far exceed cations, but the variable effects of liquid water, the buffering action of HSO₄⁻/SO₄²⁻, and the effect of species activity coefficients in concentrated particles precludes its use for any quantitative means. Even when all ionic components – including the dissociated ions of organic acids – are accounted for and measured with good precision, the ion balance is unlikely to qualitatively distinguish alkaline particles from mildly acidic particles due to the propagation of uncertainties in the aerosol composition measurements.”

10. “In addition, most aerosol solution pHs that have been reliably characterized to date (including those reported by the authors in Fig. 12) yield pHs that are less than 5 in which chase OH⁻ concentrations are trivial and can be ignored. Rather than complicate the presentation with unimportant details, I encourage the authors to drop OH⁻ from the equation and to add a simple qualifier to the text noting the possibility of rare cases when explicit evaluation of OH⁻ may be necessary. If the authors wish to consider aerosol solution pHs greater than about 5, contributions from carbonic acid must also be evaluated explicitly, which should be addressed in the manuscript.”

We have dropped \([OH⁻]\) from Equation 2 as suggested and leave the conditional explanatory statement.

11. “Page 27,584, Section 2.2. The “molar ratio” approach and “ion-balance method” represent different variants for interpreting the same information to address the same issue and should be reported as such. For example, if the equivalent sum of all anions minus the sum of all cations in dilute analytical solutions equals 0.0 (implying no contributions from H⁺), then the corresponding equivalent ratio for sum all anions to sum all cations will equal 1.0 (implying “complete neutralization”). The underlying assumptions and limitations detailed above for the ion-balance approach apply equally to the molar ratio approach.”

We understand and appreciate the referee’s point here. Clearly, the data underlying the calculation of ion balance and the molar ratio are the same. However, there are some fundamental differences that are important to note. The ion balance is an extensive property that scales with sulfate and total PM2.5 mass in many environments. The molar ratio is employed as a surrogate for acidity – it is an intensive property
independent of total PM$_{2.5}$ mass. It is possible to have a very low inferred H$^+$ from the ion balance due to low PM$_{2.5}$ levels, but at the same time infer that sample to be highly acidic from the molar ratio. Figure 6 supports this point: at the vertical dashed line (where the ion balance infers neutral particles), there are points with molar ratios less than 0.7 (i.e., inferred acidic. We therefore think it is important to note the distinction and to keep the text unchanged in this case. However, we have significantly shortened the discussion of the molar ratio results (see below).

12. “While concentrations in units of micromoles can be converted to and interpreted in units of microequivalents, equations 3, 4, and 5 as written correspond to equivalent ratios not molar ratios. The molar ratio does not equal the sum of anions divided by the sum of cations in units of microequivalents as indicated in Equation 3. In addition, the units for concentrations reported in equations 4 and 5 are in micromoles not microequivalents as stated in the text. The text and equations should be corrected.”

We have made the following corrections to the text: “The molar ratio is a ratio of the total molar concentration of measured inorganic cations to the measured inorganic anions…” and deleted the phrase (pg. 27584 line 6) “…where all concentrations are in milliquivalents.”

13. “Page 27,584, lines 2-3. Both the ion balance and the molar ratio approaches can be applied to loadings in the atmosphere, loadings on filters, or concentrations in dilute solutions. As noted above, both involve the same underlying data. For example, both ion balances and ion ratios are often used in laboratories as QA diagnostics for measured aerosol loadings (i.e., in units of microequivalent per filter or impaction substrate or per liter of extract solution) before converting to concentrations per unit volume of air. Many labs also routinely employ both ion balances and ion ratios as QA diagnostics for measured ionic constituents in samples of precipitation and cloud water (in units of microequivalents per liter). There is nothing inherently different in these two approaches.”

See our response to comment #11 above.

14. “Page 27,584, equations 4 and 5. As indicated above, these equations define equivalent not molar ratios. The authors should point out here that ignoring other ionic constituents in samples introduces bias into the corresponding acidity inferred from such relationships. I encourage them use their own data to illustrate this important point explicitly.”

We have made the following corrections to the text: “The molar ratio is a ratio of the total molar concentration of measured inorganic cations to the measured inorganic anions…” and deleted the phrase (pg. 27584 line 6) “…where all concentrations are in milliquivalents.” We feel that we have addressed the referee’s other comment: see Fig. 4a and 4b, and the discussion from Pg. 27590, line 25 through Pg. 27591, line 10.

15. “Page 27,586, lines 1-10. The Henry’s Law and dissociation constants and associated temperature corrections used in these calculations should be reported. Are these values the same as those used in E-AIM and ISORROPA? If not, what are the implications for interpreting results?”
See our response to comment #18 below. We give the values and implications are discussed.

16. “Page 27,586, Equation 6. The equation as written is incorrect. It should specify activities for H+ and NO3- not concentrations. If the distinction between activities and concentrations is ignored in these calculations, the assumption that concentrations are equal activities should be stated explicitly and the implications for results assessed, particularly in light of the fact that the thermodynamic model calculations with which these results are compared are based on activities not concentrations.”

We have corrected Equation 6 to reflect activities not concentrations.

17. “Why is the partitioning equation for HNO3 used as an example when no aerosol solution pHs based on this approach are reported in the manuscript? Only pHs based on NH3 partitioning are reported so it would seem appropriate to discuss the equation for NH3 here. More generally, since the data for particulate NO3- and HNO3 are available, why are pHs not also calculated and interpreted based on the phase partitioning of HNO3? These results are directly relevant to the analysis and would provide another useful constraint on these approaches. These results may also fill in some of the large gaps in the phase partitioning of HNO3 reported in Figures 10, 11, and 12. I encourage the authors to report and interpret pHs inferred from the measured partitioning of HNO3.”

We have changed the equation so that the NH3/NH4+ equilibrium is the illustrated example, per the referee’s suggestion. We have chosen to use the NH3/NH4+ system since Fountoukis et al. (2009) have shown the best measurement-model agreement during MILAGRO for these components, compared to HNO3 and NO3-. Fountoukis et al. (2009) extensively discuss reasons for the model-measurement discrepancies in HNO3 and NO3-: we feel that using the HNO3/NO3- equilibrium would also require a similarly extensive discussion that would distract from the focus of the paper. This is also supported by Ansari and Pandis (1999), who showed excellent consistency among thermodynamic model in predicting NH3 and NH4+ under ‘ammonia rich’ conditions (similar to conditions during MILAGRO).

18. “Page 27,586, lines 15-18. The activity coefficients for the aqueous anions (NO3- and Cl-) and NH4+ employed in this approach also impact the calculated activity of H+, particularly at the relatively low RHs and high ionic strengths during this study, and should be considered explicitly for the most accurate results. Were the reported pHs inferred from the measured NH3 phase partition calculated based on activities or concentrations? If concentrations, what are the implications for results? It would be helpful to evaluate the sensitivity of results to use of activities versus concentrations. If activities were employed in the calculations, which estimates of activity coefficients were used? Two models were run in two modes each (reverse and forward) so, presumably, four discrete sets of estimates of activity coefficients were available for each sample.”

We have added the following text to clarify this point: “To predict aerosol pH from the NH3/NH4+ phase equilibrium (Eq. 6), aerosol liquid water was taken from the forward model output of ISORROPIA. The temperature-dependent $K_H$ and $K_b$ are from Chameides (1984), while the temperature-dependent $K_w$ is from Stumm and Morgan (1996). We have used liquid concentrations, not activities in the application of Equation 6 (i.e., activity coefficients for H+ and NH4+ are assumed to be unity). The implications of these assumptions are discussed below… Differences between the NH3/NH4+ phase partitioning predictions and
either thermodynamic model are possibly due to differences in equilibrium constants and/or differences in the activity coefficients, although the former is more likely. That is evident from the strong agreement seen in Fig. 4a; since E-AIM employs activity coefficient calculations for all species, the assumption of activity coefficients of unity does not appear to systematically bias the phase partitioning pH predictions. A detailed characterization of uncertainties and sensitivities of the various methods to differences in $K_H$, $K_b$, $K_w$, and activity coefficients is beyond the scope of this study, but should be explored in the future.”

19. “The authors should also specify which estimates of liquid water contents were used in calculating pHs from the measured phase partitioning. Again, two models were run in two modes each so, presumably, four discrete estimates of liquid water content were available for each sample.”

We have followed the referee’s suggestion and updated the text as follows: “To predict aerosol pH from the NH$_3$/NH$_4^+$ phase equilibrium (Eq. 6), aerosol liquid water was taken from the forward model output of ISORROPIA.”

20. “Pages 27,586-27,587, Section 2.5. The measurement frequency for each measurement technique should be specified and the approach used to create paired data for techniques that sampled at different frequencies should be described. For example, were the gas-phase data averaged over each aerosol sampling interval? In addition, the suite of ions that were characterized with the PILS and used in this analysis should be specified.”

We have included this information in Section 2.5 (Evaluation Dataset) and with the incorporation of two new tables (Table 1 and 2).

21. “Page 27,586, lines 22-23. As described in several of the cited papers, aerosol composition (including pH) typically varies significantly as a function of particle size. Consequently, the compositions of PM$_{2.5}$ sampled in bulk with the PILS (and the corresponding pH inferred from that composition) are not necessary representative of those for the size-resolved aerosol distributions with which NH$_3$, HNO$_3$, and HCl partitioned in ambient air. This is an important point that the authors should address in the context of this analysis. Their approach assumes implicitly that ambient aerosol composition and pH is invariant over particle sizes less than 2.5 micrometer diameter, which is almost certainly not the case. What are the implications for results?”

We have added the following discussion to the paper: “Given that PM$_{2.5}$ composition only was measured, no size-acidity dependence could be elucidated (Keene et al., 2002).”

22. “Page 27,587, lines 3-6. The assumption that “HCl was effectively zero” is inconsistent with the reported measurements, the associated thermodynamics of the system, and some of the cited literature. As is the case for HNO$_3$ (equation 6), the gas-aerosol phase partitioning of HCl is controlled by aerosol liquid water content and solution pH not by “high concentrations of gas phase NH$_3$. ” For example, the authors cite Young et al. (2013) who reported significant aerosol acidities and mixing ratios of gaseous HCl in the presence of high concentrations of gas phase NH$_3$. The solubility of NH$_3$ decreases with increasing pH. In addition, HNO$_3$ is infinitely soluble in alkaline solution so the presence of significant HNO$_3$ vapor clearly indicates that the aerosols
with which it equilibrated were acidic. Indeed, all reported estimates of aerosol pH that the authors consider to be reliable range from about 2 to 5 (e.g., Fig. 12). Applying the thermodynamic expression reported by the authors (equation 6) to HCl partitioning in the presence of acidic aerosol would yield significant gas phase concentrations, consistent with the cited literature. The text should be corrected.

We respectfully disagree with the referee on this comment. In ammonia-poor environments (such as the marine environment), HCl concentrations can be significant. However, in Mexico City during MILAGRO, ammonia was abundant (see discussion in the manuscript). Our assumption of low gas-phase HCl is based upon the modeling and analysis of Fountoukis et al. (2009) using the MILAGRO dataset, which we cite as our justification. Fountoukis et al. (2009) note: “Concentrations of HCl were most likely low (mean predicted value for HCl(g) = 0.03 µg m\(^{-3}\)), a consequence of having large excess NH\(_3\)(g) which tends to drive Cl\(^-\) into the aerosol.” With this assumption employed, Fountoukis et al. (2009) get excellent agreement between measured and modeled aerosol chloride (see their Fig. 2e). If their assumption was not valid, such a high level of agreement would not likely be observed.

23. “Page 27,587, lines 19-28. It would be helpful if the authors evaluated the sensitivity of results to these approaches for forcing the required ion balances. For example, the authors’ analysis suggests that pHs predicted by the thermodynamic models run in the “reverse” mode are highly sensitive to NH\(_4^+\) concentrations (e.g., Fig. 9 and associated text) and yet, for E-AIM, they add NH\(_4^+\) to balance “missing” cations that are not considered in the model. What are the implications for results if missing cations were balanced by adding H\(^+\) rather than NH\(_4^+\)?”

See also our response to referee comment #46 below. First, “For conditions of excess cations, ISORROPIA-II assumes that bicarbonate (HCO\(_3^-\)) and carbonate (CO\(_3^{2-}\)) account for the deficit and a pH limit of 7 is imposed (as in Fig. 4)… E-AIM assumes that OH\(^-\) balances any excess cations, rather than carbonate, and thus has regions of higher predicted aerosol pH (e.g., Fig. 4).” We have used this description of the models to explain why ISORROPIA and E-AIM produce different pH levels above 7 (e.g., in Fig. 4) – this does not reflect any analyses that we have undertaken or assumptions we have employed.

Second, ISORROPIA does not require ion balance among the input data in order to run (although the model uses electroneutrality in its computational procedure). Since ISORROPIA treats all of the inorganic species that were measured, there were no substitutions or procedures to balance ions prior to running the model.

Finally, E-AIM requires ion balance among the input data in order to perform a computation. E-AIM offers H\(^+\) and OH\(^-\) as species to balance any differences between cations and anions. This is what we used to achieve electroneutrality in our input data so the model could be run. The text in Pg. 27,587, lines 24-26 reflects a common procedure used for thermodynamic model applications (e.g., Jacobson 1999). E.g., how does one treat measured calcium ions in a model that does not include calcium? Clearly, using H\(^+\) would introduce bias in the acidity results, so we used equivalent sodium to account for species not treated by the model. This most definitely will influence model results (for example the hygroscopicity of calcium and sodium are different), which helps explain some of the systematic differences between E-AIM and ISORROPIA.

24. “It would also be helpful for the authors to comment on how the E-AIM models II and IV compared at RHs above 60%. Do they yield comparable results and, if not, what are the implications for results?”
The AIM-II and AIM-IV simulations above 60% RH yielded nearly identical results owing to the predominance of \( \text{NH}_4^+ \), \( \text{NO}_3^- \), and \( \text{SO}_4^{2-} \) over Na+ and Cl- (see Table 1).

25. “Note that Figure 4 is cited out of sequence.”

Pg. 27587, line 21: we have removed the reference to Figure 4 here so that the figures are introduced in the text sequentially.

26. “Page 27,588, line 6. In this regard, it would be helpful to specify the numbers or percentages of samples evaluated with E-AIM II (i.e., at RHs between 40% and 60%) versus E-AIM IV (i.e., at RHs greater than 60%). The authors should also point out that eliminating all data for periods when RH was less than 40% presumably eliminated a disproportionately large fraction of daytime samples.”

We have included this information with the incorporation of a new table (Table 2).

27. “Page 27,588, Section 3. Although the underlying measurements may have been reported elsewhere, only calculated values based on unspecified subsets of those measurements are presented in the manuscript. Without digging into the cited literature, this prevents interpretation of results in terms of either the measurements from which they were derived or ambient air quality. It would be helpful for readers if the authors included time series plots (or at minimum summary statistics) for the major chemical and physical properties of the system on which the analysis was based. In preceding sections the authors describe some data qualitatively (e.g., high concentrations of gas phase NH3) but not quantitatively. This begs the question of how high is high?”

We have included this information with the incorporation of a new table (Table 1).

28. “Page 27,588, lines 11-14 and Figure 1. The species included in the ion-balance calculations should be specified. To minimize the potential for confusion, I also suggest that the authors employ consistent terminology when describing the model modes used for the calculations. Here they refer to “: : :E-AIM corresponding to the inorganic ion balance : : :” whereas elsewhere (including the cited figure) they refer to “: : :E-AIM in the reverse mode.””

We have made the addition in Section 2.5, as suggested. We have also modified the text to read: “Figure 5b shows the aerosol pH predicted by E-AIM vs. the inorganic ion balance.”

29. “Figures 1b, 2, 3a,b, 4a,b, 5. Major conclusions of this analysis include the following: “These results suggest that the two best proxy methods for estimating aerosol pH are: (1) thermodynamic equilibrium models run using gas+aerosol inputs, and (2) the phase partitioning of ammonia.”

“The aerosol pH levels predicted by both reverse models do not agree with either the pH predictions using gas+aerosol inputs or the phase partitioning of ammonia.” Based on the above, it is unclear why the authors focus much of the analysis on the presentation (as depicted in the above mentioned figures) and interpretation of unreliable
acidities inferred from ion balances and molar ratios versus unreliable pHs predicted by the reverse models. If these three approaches are all considered unreliable, how can the primary cause(s) for variability be ascertained by comparing paired sets of corresponding results? If pHs based on the forward models and measured phase partitioning are considered the most reliable estimates, then it would seem appropriated to use one or the other or both as the primary reference(s) for these interpretations. In this regard, the authors may wish to consider the merit of reorganizing the manuscript to present the reliable approaches first and then to employ results for these approaches to illustrate the nature of problems with the other approaches.”

We appreciate the referee’s suggestion and have completely re-organized the paper. We now present the results in the following order: thermodynamic equilibrium models, phase partitioning, ion balance, and molar ratio. We have also re-ordered the figures and removed some of the early figures. We have also substantially shortened the molar ratio results section. We feel that the paper flow and focus has improved significantly.

“More generally, it would also be helpful to compare results based on the same unit conventions. For example, in Figure 2, the pH values calculated with E-AIM could be converted to and plotted as H+ in units of nmol m-3 (as in Figures 3 and 10). Conversely, the H+ values inferred from the ion balances could be extrapolated to aerosol liquid water contents and expressed as pHs.”

We appreciate the referee’s point, but feel that the comparisons as presented in Fig. 2 and Fig. 5 are needed to show that the proxy methods are not at all representative of aerosol pH. There are dozens of peer-reviewed journal articles in which the ion balance and/or molar ratio are used as a proxy for acidity (pH) – without taking the steps of extrapolating the H+ using aerosol liquid water and activity coefficients to try and predict pH. Because this is so commonly employed, we feel the [H+] vs. pH comparisons are necessary and effective in showing that these methods do not approximate aerosol pH.

29. “Page 27,588, line 26 and Figure 2. Suggest adding the definition of the horizontal black lines to the figure caption.”

We have made the change as suggested (also applied to all figures).

30. “Page 27,589, line 27. I encourage the authors to elaborate on the impact of analytical uncertainties on the reliability of results based on the ion-balance approach, ideally with reference to specific quantitative examples based on their measurements. If the underlying assumptions detailed above (comment regarding page 27,583, lines 12-13 and Eq. 2) are valid, H+ concentrations in dilute analytical solutions can be reliably estimated from ion imbalances and then extrapolated to aerosol solution pHs with reasonable confidence using the same approach as used by others to extrapolate direct measurements of H+ in dilute analytical solutions to aerosol pH (e.g., Keene et al., 2002, 2004). As noted above, however, H+ concentrations inferred from ion imbalances often fall within the accumulated analytical uncertainties for the measured aerosol constituents and, thus, correspond to noise. This is the real Achilles heel of the approach when applied to many types of aerosol samples and should be emphasized in an analysis such as this. Readers would benefit if the authors employed their data to compare the relative magnitudes of H+ inferred from ion imbalances versus the corresponding estimated detection limits for H+ based on accumulated analytical
uncertainties for the many measured ions. More generally, it would be helpful is the authors clearly differentiated between the reliability of H+ concentrations inferred from this approach versus the reliability of interpreting these results in the context of aerosol acidity."

We have added a figure (New Figure 8) and the following discussion to extrapolate on this point, which was also brought up by Referee #2:

“Finally, another major limitation in estimating the aerosol H+ loading from the ion balance is the small H+ levels relative to analytical uncertainty. Figure 8a shows the relative H+ uncertainty (%) vs. the H+ level inferred from the ion balance (nmol m⁻³). The H+ uncertainty was calculated using a standard propagation of error technique (Harris, 1999):

\[ u_{H^+} = \left( u_{Na^+}^2 + u_{NH_4^+}^2 + u_{Mg^{2+}}^2 + u_{Ca^{2+}}^2 + u_{Cl^-}^2 + u_{NO_2^-}^2 + u_{NO_3^-}^2 + u_{SO_4^{2-}}^2 \right)^{1/2} \]

Where \( u_{H^+} \) is the absolute uncertainty in H+ (neq m⁻³), and the terms on the right hand side represent the absolute uncertainties in each inorganic species in neq m⁻³ (Na⁺, NH₄⁺, Mg²⁺, Ca²⁺, Cl⁻, NO₂⁻, NO₃⁻, and SO₄²⁻). Figure 8a shows that the relative H+ uncertainty (\( u_{H^+} / H^+ \)) grows with decreasing H+, as one would expect. A frequency distribution of the H+ uncertainty from the MILAGRO study shows that more than 20% of the H+ calculations had an associated uncertainty higher than 50%, and more than 10% of the H+ calculations had an associated uncertainty higher than 100%.

These results demonstrate numerous problems with the ion balance and strongly suggest that it is inadequate to represent the pH of atmospheric particles. The ion balance may provide a qualitative indication of acidic conditions when anions far exceed cations, but the variable effects of liquid water, the buffering action of HSO₄⁻/SO₄²⁻, and the effect of species activity coefficients in concentrated particles precludes its use for any quantitative means. Even when all ionic components – including the dissociated ions of organic acids – are accounted for and measured with good precision, the ion balance is unlikely to qualitatively distinguish alkaline particles from mildly acidic particles due to the propagation of uncertainties in the aerosol composition measurements.”

31. “Figure 3. Here and elsewhere, I suggest that the lines be defined on the figure captions.
In addition, the technique used to calculate linear regressions should also be specified.
For regressing paired data each of which are subject to significant uncertainty, the Reduced Major Axis procedure is the most appropriate method.”

We have made the changes (applied to all figures), as suggested.

32. “Pages 27,590-27,592, Section 3.2. As indicated above, the molar ratio approach is simply a different way of interpreting ion-balance relationships. As such, this section could be shorted considerably by referring to the same set of underlying issues raised in the preceding section. This component of the analysis may warrant a paragraph or two of text and perhaps one figure but it does not justify two pages of text along with three two-panel figures.”

Referee #2 raised a similar concern. We have removed Figures 4 and 6 from the original draft. We have also shortened Section 3.2 considerably (by more than half).

33. “Page 27,590, line 16. Here and elsewhere, the authors should specify whether they are referring to predictions based on the reverse or forward models.”
We have updated the text to reflect this change, as suggested. We also note that in all figures the mode of the thermodynamic model (forward or reverse) is explicitly stated.

34. “Page 27,591, lines 7-9. The implicit assumption that acidity can be reliably inferred from relative concentrations of NH4+ versus (Cl- + NO3- + SO42-) is not valid. The authors could use their own data to illustrate this important point quantitatively and in a more direct manner than employed for Figure 6, which depicts a plot of equivalent ratios versus equivalent imbalances. Specifically, they could directly compare the equivalent ratios of NH4+ / (Cl- + NO3- + SO42-) versus the corresponding equivalent ratios for the sum of all cations divided by the sum of all anions versus (i.e., a plot of ratios versus ratios in the same units). Alternatively, they could compare the equivalent imbalances for NH4+ - (Cl- + NO3- + SO42-) versus the equivalent imbalances for sum anions minus sum cations in units of nmol m-3.”

We have removed this from the paper.

35. “Figure 4. Suggest specifying numbers of samples depicted in panel a. What is meant by “low crustals” in panel b? Since a distinct definition for this subset is specified in the caption, I suggest that the characterization of “low crustals” be dropped from the figure.”

We have defined what is meant by ‘low crustals’ in the text (pg 27591, lines 2-4). This figure has been removed from the text.

36. “Page 27,591, line 14. Here and elsewhere, to what “crustal species” are the authors referring and what is the rationale for characterizing them as “crustal” as opposed to, for example, biomass-burning products?”

We have added the following for clarification: “(where Ca2+, Mg2+ and K+ collectively make up < 5% of inorganic aerosol mass)”.

37. “Page 27,592, line 18. It would be helpful to include a table summarizing the data subsets that were generated and used in the analysis (e.g., (1) the total number of aerosol samples characterized, (2) the subset of (1) at RH greater than 40%, (3) the subsets of (2) with simulated liquid water contents greater than 0 for each model, and (4) the subsets of (3) with paired gas-phase data). For example, Figure 9 depicts 89 results based on ISORROPIA but only 39 based on E-AIM. Why the large difference? In addition, according to section 2.5, aerosol composition was characterized with the PILS (presumably at about 5-minute intervals) over about 30 days, which would have yielded about 8640 individual measurements. The authors indicated that half of these measurements (_4320) corresponded to RHs less than 40% and were not exclude from the analysis. What are the primary causes for the large differences between the presumed numbers of aerosol observations at RHs greater than 40% and the corresponding numbers of inferred pHs?”

We have included this information with the incorporation of Tables 1 and 2.

38. “Page 27,592, line 20-23. What were the differences in estimated aerosol liquid water contents between the reverse versus forward models and which values were used to
infer pH from the phase partitioning of NH3? This issue is directly relevant to the analysis.”

Aerosol liquid water predicted by E-AIM and ISORROPIA agreed very well (slope = 1.14; R² = 0.986). This is completely consistent with past studies that show very good agreement in aerosol LWC between different thermodynamic equilibrium models (e.g., Fountoukis and Nenes (2007); Ansari and Pandis (1999)). While there are small differences in LWC (e.g., treating Ca²⁺ as equivalent Na⁺ in E-AIM changes the aerosol hygroscopicity), they are in general minor, and a deep discussion of this issue would likely distract from the main points of the paper.

39. “Figure 8, caption. Ion imbalances in range ‘a’ are less than those in range ‘b’. The end of the caption should read, “: : :demonstrate the extreme sensitivity of aerosol pH at Ib values near 0.0 (region “b”).” We have adopted the change, as suggested.

40. “Page 27,594, line 3. Suggest citing examples to support the statement that “: : :similar model sensitivities have been observed in diverse locations, as well.” We have removed this sentence from the manuscript.

41. “Page 27,594, lines 7-10 and Figure 10. I suggest that scale on the X axes in Figure 10a,b be expanded so that relative variability in H⁺ predicted by the two approaches is evident. A notation in the caption indicating that the scales on the two axes differ coupled with a 1:1 line on the figure would suffice to alert readers to these differences. Alternatively, the data could be plotted on log scales.” We have followed the referee’s suggestion and plotted the x-axes on log scales. The figure caption has also been updates as follows: “Figure 7. H⁺ levels inferred from the ion balance compared to H⁺ levels predicted by a) E-AIM (forward mode), and b) ISORROPIA (forward mode). Note the different x- and y-axis scales, and note the log scale for the x-axes.”

42. “Page 27,596, lines 9-13. The expected relationship between the pH values predicted by the reverse models and the corresponding ratio of NH₄⁺ to NH₄⁺ + NH₃ is unclear. If retained, it would be helpful to specify the range in the ratio of NH₄⁺ to NH₄⁺ + NH₃ predicted by thermodynamics at pH values in the range of 0.0. However, as questioned above, what is the rationale for interpreting paired sets of data both of which are considered unreliable?” We have added the following sentences to clarify: “Under conditions typical of the MILAGRO study (293 K, 10 µg m⁻³ aerosol liquid water, total NH₃ of 10-30 µg m⁻³), if the aerosol pH is less than 0, thermodynamics calculations predict that essentially all of the NH₃ should partition to the particle phase if equilibrium is achieved. The fact that gas-phase NH₃ was abundant suggests either: 1) the system is not at equilibrium, or 2) aerosol pH > 0.”

43. “Page 27,596, lines 16-18. While Young et al. (2013) reported good agreement between
aerosol pHs inferred from the phase partitioning of NH₃ and HNO₃, Keene et al. (2004) reported significant divergence between results based on the same approach. This inconsistency should be mentioned. Differences between results from these two studies may reflect the fact that activity coefficients were evaluated explicitly by Young et al. whereas results reported by Keene et al. were based on concentrations. As indicated above, it would be useful and relevant to evaluate the sensitivity to pHs inferred from phase partitioning based on concentrations versus activities.”

We have added the following text to clarify this point: “It should also be noted that Keene et al. (2004) observed relatively poor agreement between pH predicted by ammonia and nitric acid phase partitioning. This may be due to the explicit treatment of activity coefficients (as was done by Young et al. (2013)), or to differences in other factors such as relative ammonia levels, since this can introduce major differences in thermodynamic model predictions (Ansari and Pandis, 1999).”

44. “Page 27,597, lines 7-10. As noted above, an Achilles heel of both approaches is the generally small H⁺ signal relative to the noise associated with the overall analytical uncertainties. If signal-to-noise was high (and all ionic constituents were measured without bias), H⁺ in dilute solutions solutions can be reliably inferred from ion imbalances and extrapolated to ambient liquid water contents with reasonable confidence. This is an important distinction that should be clarified.” See our response to comment #30 above.

45. “Page 27,597, lines 18-19. What is meant by “excess ammonia”?” See response to comment #42 above. We have added in Table 1 mean, median, 5th and 95th percentiles of ammonia. We have also changed the wording of the sentence to read “…despite high concentrations of gas-phase ammonia.”

46. “Page 27,597, lines 19-21. If, as suggested by the authors, results based the reverse models are extremely sensitive to the initial inputs, then the approaches used to force ion balances for both models should be evaluated explicitly with respect to the corresponding sensitivity of results.”

First, ISORROPIA does not require ion balance among the input data in order to run (although the model uses electroneutrality in its computational procedure). E-AIM requires ion balance among the input data in order to perform a computation. E-AIM offers H⁺ and OH⁻ as species to balance any differences between cations and anions. In this sentence, we are referring to the fact that small changes – those with a magnitude commonly associated with uncertainty in aerosol measurements – can impart major effects on predicted pH with the reverse model runs. We have changed the text to make this point more clearly: “Further, the models in reverse mode are highly sensitive to uncertainty in the measurement inputs: small deviations in major aerosol species on the order of common aerosol measurement uncertainties can induce changes in predicted pH that exceed 10 pH units.”

47. “Page 27,598, line 25. Suggest adding that the equilibration times for sub-micron-diameter particles reported by Meng and Seinfeld (1996) are on the order of seconds to minutes, which is consistent with results from other studies.”
We have updated the text to read “This sampling time should far-exceed the equilibration time for sub-micron particles, which is on the order of seconds-to-minutes (Meng and Seinfeld, 1996).”

48. “Page 27,599, lines 4-5. It is unclear what the authors mean by a “gold standard.” As indicated in the manuscript, the cited papers by Keene et al. (2002, 2004) report aerosol pHs extrapolated from a “direct measurement method.” The 2004 paper by that group indicates that pHs extrapolated from direct measurements were reasonably consistent with corresponding estimates based on the thermodynamics of HCl and HNO3 phase partitioning but not with those based on NH3 phase partitioning.”

We have changed the text to read: “These recommendations should be evaluated in other environments with different chemical characteristics and meteorology than was observed during MILAGRO. In particular, the recommendations should be evaluated in ammonia poor environments, since thermodynamic equilibrium models often diverge in their predictions under such conditions...” We also note that the results from Mexico City were consistent with those from the southeastern US (Guo et al., 2014), where the conditions were generally much more acidic owing to far lower NH3 concentrations.

Referee #2

Referee #2 emphasized many of the same comments made by Referee #1. We have addressed those comments above. In addition:

1. “Page 27591, line 22: Here and also on page 27595, line 20 and 24, correlation coefficients are given. As already pointed out by referee No.1, an explanation on which method was used to calculate these coefficients, should be given. In addition, implications and judgments should be done carefully. On page 27595, both R2 coefficients are called "high values", although a R2 value of 0.47 should not be considered as "high" especially when a value of 0.80 is also called "high".”

We have changed the text so that it now reads: “A slope of 0.98 and high R2 value (0.80) indicate excellent agreement between E-AIM and the phase partitioning approach (Fig. 4a). The median difference between these methods was only 0.4 pH units. Likewise, good agreement was observed between ISORROPIA-II in the forward mode and the phase partitioning pH (Fig. 4c, slope = 0.98 and R2 = 0.47).”

2. “Page 27588, line 23: the active term using "we" should be converted to a passive term, as was used in the whole manuscript.”

We have made the correction.

Comment by J. L. Jimenez

1. “Some conclusions may be generalized too widely or presented too strongly, and some finer nuances may be needed... Thus, unlike the abstract of this paper, the Murphy et al. work more clearly states that the molar ratio method does work for sufficiently acidic aerosols, and that it is only for neutral (or close to neutral) conditions that the large uncertainty on pH arises.”
This abstract from Murphy et al. (2011) is cited as evidence that the molar ratio method works as a proxy for aerosol pH “under high levels of strong acidity,” with 20 neq m\(^{-3}\) given as the threshold. We disagree with Dr. Jimenez’s interpretation of this abstract. We interpret the abstract to mean that thermodynamic models run in the reverse mode (using aerosol concentrations only) may provide accurate predictions of aerosol pH under sufficiently acidic conditions (“...it is possible to combine measurements of particle composition, relative humidity and temperature with calculations from aerosol thermodynamics algorithms to calculate the bulk pH. While this approach is straightforward...”). In addition, we disagree with this generalized statement, as well. For our data set, if only the subset of samples where \([H^+] > 20\) nmol m\(^{-3}\) is considered (by ion balance or reverse thermodynamic models), there is still a major difference between the reverse (mean pH = -1.50) and forward (mean pH = 2.41) thermodynamic models by approximately 3 pH units, i.e., 3 orders of magnitude in H\(^+(aq)\).

We agree with Dr. Jimenez that this issue needs to be studied in more diverse locations under widely varying conditions, but we do not feel that we have overstated our results. For one, our results are quite consistent with those from diverse locations that include the southeastern US (Guo et al., 2014), Hong Kong (Yao et al., 2006, although our interpretations differ, as we extensively discuss in the paper), and Singapore (Behera et al., 2014). We have added this last reference in the updated manuscript. In addition, the pH range in the MILAGRO data is very broad (NH\(_3\) phase partitioning predicts pH range of 0.8-5.2; ISORROPIA forward model predicts pH range of 1.7-6.8; E-AIM forward model predicts pH range of 2.1-4.8; ISORROPIA reverse model predicts pH range of -1.7-7.4; E-AIM reverse model predicts pH range of -2.2-10.1). Our recommendation against the use of the ion balance method is in line with that from a study in a marine environment off the coast of New England (Keene et al., 2004) and the southeastern U.S. (Guo et al., 2014). In the final paragraph of the updated manuscript, we have added: “These recommendations should be evaluated in other environments with different chemical characteristics and meteorology than was observed during MILAGRO. In particular, the recommendations should be evaluated in ammonia poor environments, since thermodynamic equilibrium models often diverge in their predictions under such conditions...”

“As an aside, the way Fig 8 is shown in ACPD may give the wrong impression about the uncertainty, since pHs beyond 7 or 8 are not thought to occur in the atmosphere. It seems that the inset of the figure as shown above shows the more relevant information for the atmosphere, so I suggest that the figure is updated accordingly for the ACP version.”

We understand this point; however, fresh marine particles and mineral dust particles are likely to have pH up to 9. Also, E-AIM (one of the most widely used thermodynamic models) predicts particles with pH in excess of 10 for atmospherically-relevant inputs, so we feel the range shown in the figure is appropriate.

2. “An important detail is that although the pH value cannot be determined accurately with the molar ratio method for neutralized aerosols, a lower limit of the pH can nevertheless be derived, e.g. pH > 1 or so for the figure above. This is very useful information, as some chemical processes may need lower pH to proceed quickly, and thus the information from the molar ratio method may still be able to establish whether some chemical processes can / cannot be active.”

We mostly disagree, as we have shown that the molar ratio method is able to provide some qualitative information about the aerosol regime (acidic or alkaline) but cannot be used for a pH surrogate, as it does not consider liquid water content variations at all. Hence, under ‘Conclusions’ we have stated: “When species accounting for greater than 95% of inorganic aerosol mass were included in the analysis, the molar ratio appears to reliably distinguish acidic from alkaline particles; however, the molar ratio should not be treated as a surrogate for aerosol pH.”
3. “On a narrower topic, I was surprised to see the paper conclude that “Similarly, no relationship is observed between the cation / anion molar ratio and predicted aerosol pH.” We had done similar work before based on the Pittsburgh 2002 AMS measurements (where aerosols were often very acidic) [Zhang et al. 2007] and we had shown that as long as T & RH are kept constant, the molar ratio and AIM-estimated pH are tightly correlated, see panel (c) below from Figure S1 in that paper: As panel (d) shows, a large amount of scatter appears only when the instantaneous T & RH are used in the calculation. A question then arises on what is more relevant for a given application of estimated pH, the estimates under constant T & RH or the ones under the instantaneous T & RH. This will depend on the timescale of the process of interest. For processes with a timescale of a day or longer one can argue that the pH estimated under constant T & RH may be a better surrogate of the pH that those aerosols have experienced over the last few days. This is relevant to e.g. IEPOX uptake to make IEPOX-SOA over the SE US, as the concentration of IEPOX-SOA observed at a given site likely has been formed over a period of several days.”

First, Figure S1-d from Zhang et al. (2007) quite strongly supports our findings and recommendation against the use of the molar ratio as a surrogate for pH. As Dr. Jimenez notes, for a given molar ratio, the predicted aerosol pH spans 2 pH units (or more) when the ambient T and RH are used. pH expresses the thermodynamic state of an aqueous phase, with a typical readjustment timescale of less than a second (Seinfeld and Pandis, 2006); the instantaneous values of RH and T are therefore the relevant quantities here. Although beyond the scope of our discussion, the link of the “instantaneous pH” to a specific process depends on the context. In the example brought forth, instantaneous aerosol pH (e.g., determined by Guo et al., 2014) in the SE US is relevant for IEPOX-SOA, because i) the aerosol throughout the SE US is consistently acidic over the region, ii) emissions of isoprene are embathed in this aerosol and processed rapidly in the BL (Xu et al., 2015), and, iii) the diurnal profiles of acidity, isoprene emissions, organic aerosol, etc. are strikingly reproducible, regardless of origin of airmass (Xu et al., 2015).

4. “Partially some of the confusion may arise because only the Mexico City MILAGRO dataset from T1 has been used here. There are two features of that dataset that may make the analysis more difficult than elsewhere: (a) the aerosols during MILAGRO were almost always neutralized within experimental error due to very high emissions of NH3. At least this was the case at the nearby T0 site (see Fig S6 of Aiken et al. [2009]). Thus some of the nominally acidic conditions in this dataset may have a substantial influence from the measurement uncertainties (also see (b) below). Thus much of the data used in this paper is in the regime where the molar ratio method does not work, and this may be part of the explanation for the low correlations observed. (b) There were substantial dust concentrations during MILAGRO, and especially the T1 site was described by the researchers there as very dusty. This may lead to complexities in the measurements, their interpretation, and the modeling. E.g. as was concluded by a previous MILAGRO paper “The bulk equilibrium approach fails to reproduce the observed coarse nitrate and overpredicts the fine nitrate” [Karydis et al., 2011]. Thus even if the mineral ions are included in the equilibrium model, they likely have not had time to reach equilibrium in the real atmosphere, potentially leading to model-measurement deviations. Given (a) and (b), it does not seem possible to draw near universal conclusions (as certainly the end of the abstract sounds like) from a limited dataset that may not span the dynamic range observed over wider regions of the atmosphere.”

Although we strongly agree that an analysis of other locations is warranted (see our response to comment #1 above), the conclusions derived from Mexico City aerosol are widely applicable because the pH range (forward model predicted range from 1.7 to 6.8), encompasses most of the dynamic range of the atmosphere. We thank Prof. Jimenez for this comment, because it brought forth another realization: even though ammonia concentrations were consistently high, all three methods (ion balance, molar ratio, reverse thermodynamic equilibrium models) routinely indicated highly acidic particles. As a consequence, an important point emerges: these three methods are overwhelmingly employed when accompanying gas-phase measurements are not made. So, if one needs to know the atmospheric
ammonia levels in order qualify the use of one of these methods (which is far from certain), then the more accurate phase partitioning approach or forward modeling can be conducted.

Regarding the high dust concentrations, Figure 4b in the discussion paper shows that even under low crustal influence (where $\text{NH}_4^+$, Cl$^-$, $\text{NO}_3^-$, and $\text{SO}_4^{2-}$ together accounted for greater than 95% of inorganic aerosol ion mass), the molar ratio was not a good surrogate for aerosol pH. Further constraining the analysis to inferred highly acidic conditions (those with $\text{NH}_4^+/(\text{Cl}^- + \text{NO}_3^- + 2\times \text{SO}_4^{2-})$ molar ratios less than 0.75) and with low crustals did not improve the correlation between the molar ratio and predicted pH ($R^2 = 0.01$, Fig. 5).

5. “It would have been desirable that the authors also include another dataset in which acidic conditions were prevalent in this paper, which would allow a fairer and more complete evaluation of the methods. I understand that is too much to ask for the revision of this paper, but I do hope that either the authors or others in the community perform and present such work in the future. We note that the relevant datasets including NH3, HNO3, and aerosol composition are likely available e.g. for some campaigns of the NOAA P3 (see http://www.esrl.noaa.gov/csd/field.html), and likely for other campaigns as well.”

The evaluations carried out in this paper are sound and fair, and the results speak for themselves. More analysis should be done, but we do not anticipate the main conclusions (i.e., that ion balance and molar ratio to correlate with acidity) to change. See our response to comments #1 and #4 above.

References