Article summary

This manuscript assesses the feasibility of planetary boundary layer height (PBLH) observations to improve the initial conditions of numerical weather prediction (NWP) models. To address this problem, off-line experiments are conducted in which Doppler lidar PBLH retrievals are combined with synthetically generated model profiles from two different simulations. An Ensemble Optimal Interpolation (EnOI) algorithm is used to assimilate the lidar-derived PBLH observations. The state vector is augmented to incorporate the prior PBLH values diagnosed by the underlying PBL scheme. Six independent radiosonde measurements are used to determine the consistency of the updated model profiles. The authors find that more significant corrections to the prior model profiles occur in the late afternoon hours when the background ensemble is (i) less accurate and (ii) contains more significant cross-variable ensemble covariances. Overall, it is also shown that the assimilation of PBLH retrievals leads to mixed results: while there are visible improvements in the analysis profiles of potential temperature and the V-component of the wind, the corrections to the water vapor mixing ratio and the U-component of the wind are not consistent with the verifying radiosonde measurements.

Decision: accept with major revisions

This proof-of-concept study provides an important benchmark for the future assimilation of ceilometer-based PBLH retrievals, so I strongly recommend its publication in AMT. I praise the authors for constructing a data assimilation system that is sophisticated enough to address the key objectives in their study, but also sufficiently simple to minimize any unnecessary computational costs. Nevertheless, there are several important aspects that the authors should address to further enhance the impact of their results. These are summarized below alongside with numerous other minor comments and stylistic/formatting suggestions.

Major comments

1. The presentation of the data assimilation results needs to be refined.

(a) Interpretation of the observation misfits

On L245-L248, the authors justify the small analysis increments in the early morning hours by noting the small deviation between the forecasted thermodynamic and kinematic profiles from the corresponding radiosonde measurements. This conclusion in not consistent with the EnOI update because the innovation $\mathrm{d} = \mathrm{y}^{o} - h(\mathrm{x}^{\mathrm{b}})$ only refers to the misfit between observed and forecasted (diagnosed) PBLH values. L248-L249 correctly state that there are relatively large misfits between the observed and prior PBLH, but the authors do not emphasize that these differences control the subsequent corrections of the prior T, MV, U and V profiles.

(b) The role of prior ensemble spread in the model space corrections

I also think that the authors need to better quantify the origin of the analysis increments in model space. From Eqs. (2) and (3), it should be clear that the magnitude of the model state corrections depends on two distinct factors: (i) the ensemble cross-covariances $\mathrm{P}^{\mathrm{f}}\mathrm{H}^{\mathrm{T}}$ and (ii) the scaled innovation $\tilde{\mathrm{d}} = \left(\mathrm{R} + \mathrm{HP}^{\mathrm{f}}\mathrm{H}^{\mathrm{T}}\right)^{-1}\mathrm{y}^{o} - h(\mathrm{x}^{\mathrm{b}})$. To justify the observed differences between the model state corrections in the early morning and late afternoon hours, the authors focus on the magnitude of the ensemble covariances (Fig. 8), but pay little attention to how the background ensemble spread in observation space ($\mathrm{HP}^{\mathrm{f}}\mathrm{H}^{\mathrm{T}}$) affects the scaled PBLH misfit $\tilde{\mathrm{d}}$. Aside from contributing to a more complete understanding of the model state corrections, examination of the prior PBLH spread will also shed some light on the underlying forecast uncertainty.

(c) Discussion of the PBLH corrections

While I agree with the authors' approach of focusing on the model state increments, a brief discussion of the PBLH corrections is also warranted. Such a discussion is important because the PBLH increments have a direct impact on the subsequent model state corrections (e.g., Anderson 2003). Intuitively, if a data assimilation system is not effective in estimating the directly observed quantities, it will also struggle to constrain the unobserved model state variables. Therefore, I encourage the authors to elaborate on the extent to which the analyzed PBLH values were able to move closer to the independent radiosonde-based PBLH retrievals.

2. I found the authors' justification on the importance of assimilating PBLH retrievals particularly appealing (L102-L104), but the literature overview on past efforts of assimilating Doppler lidar measurements (L99-L104) needs to be corrected and expanded.

First, none of the cited studies (Hu et al. 2019; Coniglio et al. 2019; Degelia et al. 2019) report improvements to the PBLH; instead they examine the forecast performance with respect to convective initiation. Second, the aforementioned papers assimilate horizontal wind profiles derived from the VAD technique. This comes in contrast with other studies (e.g., Kamineni et al. 2003) which focus on the impacts from assimilating thermodynamic lidar profiles. To put this study into a broader context, I suggest that the authors place more emphasis on the aforementioned differences and further expand on their literature overview starting on L98. Apart from discussing the past efforts referenced above, the authors should also include some of the most recently published research in this area. For instance, both Degelia et al. (2020) and Chipilski et al. (2020) assimilate Doppler lidar retrievals in the context of nocturnal convective systems. Chipilski et al. (2020) additionally reveal that the assimilation of Doppler lidar wind data affects the characteristics of the stable PBL – a finding which is particularly relevant in the context of the present study.

3. A critical discussion is also needed to highlight the future challenges of assimilating PBLH retrievals.

The authors point out that their work is a "necessary first step in terms of how ensemble statistics can help to constrain profiles within the PBL", indicate the need to assimilate PBLH retrievals in real-time EnKF systems under diverse weather regimes, but provide little information on the specific problems of assimilating the PBLH variable. Given that this paper is designed to inform the future incorporation of PBLH retrievals into real-time data assimilation systems, the authors should elaborate on what the most outstanding challenges in this area are. The bullet points below provide a couple of suggestions: while I certainly do not expect the authors to follow all of them, I encourage their incorporation with some of the authors' own concerns.

- The calculation of PBLH is different in model simulations and observations: while the observed PBLH is derived from Doppler lidar measurements of turbulence intensity, horizontal wind profiles and backscatter intensity, the prior PBLH is diagnosed using the specific formulation of a particular PBL scheme. Each of these two methods has its own approximations and, perhaps even more importantly, will not yield the same PBLH value even if the simulated and observed meteorological conditions are identical. If such biases are not taken into account, the analysis estimate produced by Eq. (3) will be suboptimal. The aforementioned methodological differences also constitute a problem from the viewpoint of forecast verification.

- Treatment of observations errors. On L166, it is stated that the observation error variances are equal to the uncertainty estimates provided by the lidar retrievals. In additional to these measurement errors, operational data assimilation systems typically also consider (i) errors of representation and (ii) errors due to approximations in the observation (forward) operator. These additional error contributions might be important to consider in future efforts. For example, errors of representativity might be especially relevant within the more inhomogeneous stable PBL, whereas the methodological differences in the PBLH computation could be treated, at least to a first degree, as errors in the observation operator.
- The ability of PBLH retrievals to efficiently constrain the model state. Here I offer two different comments. The first one concerns the inability of PBLH retrievals to correct the stable PBL structure despite the large differences in the prior and observed PBLH values. The analysis offered by the authors indicates that the lack of ensemble cross-covariances is a likely explanation. This raises the question, however, whether the small corrections in stable PBLs constitutes a systematic effect induced by the formulation of current PBL schemes. Answering this question is important as one of the focal objectives of the PECAN field campaign was to understand if the information provided by ground-based PBL profilers can improve the traditionally poor forecasts of nocturnal convective systems. The second point the authors might want to consider is how effective the PBLH retrievals would be in the context of other observation networks. A past NRC report (NRC 2009) describes the potential deployment of a nation-wide network of thermodynamic and kinematic PBL profilers, quite similar to the ones employed during the PECAN field campaign. Unlike the ceilometer network that originally motivated this study, the PBL profilers will produce direct observations of model state variables that could make the ceilometer-based PBLH retrievals redundant.
- My last comment is more technical in nature and relates to the theoretical inappropriateness of current data assimilation systems to extract information from PBLH retrievals. By definition, PBLH is a non-negative quantity. As such, it faces the problems similar to those associated with the assimilation of certain moisture variables (e.g., specific humidity; see Dee and da Silva 2003). Because PBLH is a bounded quantity, its distribution will not always be Gaussian. Hence, traditional assumptions used to derive operational data assimilation schemes will be violated (e.g., Bocquet et al. 2010; Bannister et al. 2020). Such deviations from Gaussianity will be particularly visible under stable PBLs (because PBLH is closer to its lower bound of 0m), further complicating the data assimilation problems already noted in this manuscript. It might be possible to remedy the problem of boundedness by adopting certain non-Gaussian extensions of traditional data assimilation algorithms (cf. Cohn 1997; Fletcher and Zupanski 2006; Bishop 2016).

Minor comments

1. L32: Momentum is also exchanged in land-atmosphere interactions, please add to the list.

2. L46-L47: "… since aerosols are well mixed throughout the PBL (Hicks et al., 2019)" - this statement is only valid in the context of CBL.

3. L72-L73: "… were not yet available for the campaign we are using". Before providing details on how the Doppler lidar measurements in Greensburg were obtained, the readers would benefit from a short description of the PECAN field campaign as a whole.

4. The use of "PBLH retrievals" is more appropriate than "PBLH measurements" as it emphasizes that PBLH is a derived quantity. There are a couple of instances in the manuscript where this correction needs to be applied.

5. Description of the methods to assimilate the PBLH retrievals (L107-L109). It might be better to replace the statement "either by creating an adjoint of the PBL parameterization scheme" with "either by adopting a variational data assimilation scheme" (or a semantically equivalent expression). Formulating the adjoint of a parameterisation scheme is a specific implementation aspect of variational data assimilation algorithms. If the authors desire so, they could motivate their preference for an EnKF approach in this study by pointing out that ensemble-based methods sidestep the generally difficult task of linearizing the model physics equations.

6. Please cite the original EnKF paper of Evensen (1994) and its subsequent refinement in Burgers et al. (1998) on L123.

7. L124: "… where the analysis state is the estimate with the minimum estimated errors". The original derivation of the Kalman filter minimizes the expected squared errors. Note that the latter corresponds to a minimisation of a L^2 error norm rather than the L^1 error norm implied on L124.

8. The EnOI was originally introduced by Oke et al. (2002) and also discussed by Evensen (2003), so please make sure to add these references on L133.

9. On L141, the authors mention the names of the two parameterization schemes used in the archived NU-WRF simulations. A brief justification on why these two parameterization schemes were chosen in the original PECAN runs will be helpful.

10. The abbreviation TKE on L145 is commonly used to denote "turbulent kinetic energy" in boundary layer research, so it might be best to consider rewording.

11. State vector definition (L145-L148). Please clarify whether the state variable Q refers to specific humidity, mixing ratio or another moisture variable. The authors should also stress that the state vector $\mathbf{x}$ in this study represents a vertical column, which is why P^{f} only refers to the vertical ensemble covariances.

12. Mathematical description of the EnOI algorithm (L160-L191). Here I have a couple of technical remarks aimed at refining the mathematical presentation of the EnOI algorithm.

- Because the definition of the forecasted error covariance in Eq. (1) is too general and not specific to the EnOI algorithm, it might be best to remove it from the discussion.
- Similarly, it is not necessary to write the measurement equation $\mathrm{y}^o = \mathbf{H}\mathbf{x}^{\mathrm{f}}$; instead, the authors could simply state that y^o in Eq. (2) represents the PBLH observations retrieved from the Doppler lidar. (As a side comment, the aforementioned measurement equation is only partially complete in the context of filtering theory as it should include a random noise term.)
- It will be best to avoid mixing the vectorial and scalar notations in Eqs. (2) and (3). This could be done by first writing the general form of Eqs. (2)-(3) and then describing how

these were solved by the EnOI algorithm employed in this study. Regarding the latter, the authors should highlight that y^o as well as the corresponding HP^fH^T and R matrices are scalar quantities and that both P^fH^T and $HP^fH^T$are computed from an ensemble of model profiles, as indicated by Eqs. (4) and (5).

- A brief description is needed to explain how the vertical localization mentioned on L177-L185 was implemented in this study.

13. Opening sentences in Section 3 (L193-L199). Some aspects regarding the description of the NU-WRF simulations were already discussed in the paragraph starting on L141. The missing details found on L193-L199 should be moved back to this paragraph.

14. L205: "… in the late evening to early morning (2-7 UTC)". 7 UTC does not correspond to early morning.

15. L207: "early morning and early afternoon". Please define this period in the same manner as you have done in other places within the text.

16. Instead of using the temperature as an example in Eq. (7), please use a generic variable, say X, to generalize the RMS difference formula. Moreover, instead of explaining the meaning of i=8 on L220, just mention that the index i denotes a model level.

17. Statements regarding the corrections made to the WV profile at 22 UTC (L262, L270-L271 and L302-303). The analyzed WV profile overshoots the observed one only with respect to the MYNN simulation. By contrast, the forecasted values in the MYJ WV profiles are already higher, so increasing the WV values following the DA update acts to further increase the observation misfit. Please make sure to make this distinction while describing your results.

18. Justification regarding the deteriorated estimates of the U profile (L266-L268). The authors correctly acknowledge that the estimation of U is a challenging task when one assimilates integrated quantities like PBLH. However, a similar inference can be made in terms of the V-component profile of the wind and, in fact, further argued that the estimation of V is more challenging due to the presence of sharp gradients in the 750-800 hPa layer (cf. lower-right panels in Figs. 6 and 7). An alternative hypothesis to explain the reported differences in the U and V estimates can be linked to the magnitude of the cross-variable ensemble covariances. The lower two panels of Fig. 8, for example, show that the PBLH-V cross-covariances are much larger than their PBLH-U counterparts. Taking into account that the performance of EnOI (and all other ensemble-based DA methods) is especially susceptible to number of ensemble members, it is quite possible that the small PBLH-U covariances are simply a manifestation of the inherent sampling noise, which would in turn act to degrade the quality of the analyzed U profiles. In theory, the above hypothesis could be tested by comparing results with different ensemble sizes (or with different number of vertical model profiles in the context of this study).

19. L272-L276: The description of how the PBLH retrievals correct the state variables should be either removed or relocated to the methodology section.

20. L284-287: Here the authors state that the "... more limited velocity corrections are largely constrained by the correlations …". This is only partially true as Fig. 8 shows that the V cross-covariances are considerably larger than their U counterparts.

21. L311: Replace “covariances” with “ensemble covariances” to emphasize the underlying computational method.

22. L311: It will be best to change “defined” to “controlled”. The ensemble covariances are defined mathematically through the sample covariance formula, but controlled by the characteristics of the T, MV, U and V profiles that enter the PBL parameterization schemes.

23. L314: Did you intend to refer to the “analysis profiles” instead of the “forecast profiles” here?

Typos and stylistic changes

1. The authors should consider segmenting their results in Section 3 with a view of enhancing the readability of their manuscript. One possibility would be to divide the results into three subsections. The first one could discuss the discrepancies between modelled and retrieved PBLH values, the second one – how the assimilated PBLH observations correct the observed and unobserved model variables, while the third might focus on interpreting the magnitude of the T, MV, U and V corrections at 04 UTC and 22 UTC.

2. The paragraph starting on L255 could be merged with the preceding one as it provides a summary of the main results.

3. Please correct the spelling of EnKF on L325 and L329.

4. There were several places where “Doppler” was not capitalized.

5. I spotted unintended word repetitions in a couple of places within the text, e.g. L252, L255 and L303.

6. L23: “leading to an increased differences” – remove “an”.

7. L36: “… rapidly transported within this layer”. It is not clear which layer is being referred to as the previous sentence mentions both the CBL and SBL.

8. L38: “The PBLH is fundamental to …”. This sentence provides a very general description on the significance of PBLH and should be mentioned earlier in the paragraph, e.g. after the sentence starting on L32.

9. L42: “penetrates the top” could be changed to “penetrates its top” to make it clear that the authors refer to the PBL top.

10. L186: “This system is solved…”. Which system? Please be more specific by listing the relevant equations.

11. L208: “… the MYJ forecasts (red triangles) both are higher than the observations”. Please confirm that “both” refers to the MYJ forecasts in the early morning and early afternoon. If this is the case, remove “both” as it is clear from the context.

12. The use of two time periods [“During the night (2-9 UTC) …”, “… in the early morning (6 and 8 UTC) …”] makes it hard to interpret the sentence starting on L225.

13. L232: “increase” should be replaced with “increases”.

14. L237: “(0.5m/2 decrease)” should be replaced with “(0.5m/s decrease)”.

15. L240: “inthe RMS differences” – please separate “in” from “the”.

16. L249-L250: “In the late afternoon (Figures 6,7) indicate …” – please remove the brackets and refine the sentence structure.

17. L260: “The WV profile is shown to be increased …”. It is the WV values that increase, not the profile itself.

18. L270: Remove “in” from “in show that”.

19. L279: “We can also analyze this …”. Not clear what “this” refers to, please clarify.

20. L321-L322: “will require the construction of an EnkF, and run over many days” – please correct the wording in this sentence.

Figures

1. It would be useful to label the figure panels with (a), (b), etc.

2. Please avoid repetitions in the title and axis labels (e.g., potential temperature in the upper-left panel of Fig. 5).

3. Legends are sitting atop some of the curves in Figs. 4-7. Please make sure that all data is displayed in the revisited figures.

4. Fig. 8: Please include the ensemble covariance units on the x-axis. Please also replace “for PBL physics model MYHH” to “for the MYNN PBL scheme”.

References

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

Bannister, R. N., H. G. Chipilski, and O. Martinez-Alvarado, 2020: Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Q. J. R. Meteorol. Soc.*, 1–48, https://doi.org/10.1002/qj.3652.

Bishop, C. H., 2016: The GIGG-EnKF: Ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. *Q. J. R. Meteorol. Soc.*, **142**, 1395–1412, https://doi.org/10.1002/qj.2742.

Bocquet, M., C. A. Pires, and L. Wu, 2010: Beyond gaussian statistical modeling in geophysical data assimilation. *Mon. Wea. Rev.*, **138**, 2997–3023, https://doi.org/10.1175/2010MWR3164.1.

Burgers, G., P. J. Van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Chipilski, H. G., X. Wang, and D. B. Parsons, 2020: Impact of assimilating PECAN profilers on the prediction of bore-driven nocturnal convection: A multiscale forecast evaluation for the 6 July 2015 case study. *Mon. Wea. Rev.*, **148**, 1147–1175, https://doi.org/10.1175/MWR-D-19-0171.1.

Cohn, S., 1997: An Introduction to Estimation Theory. *J. Meteorol. Soc. Japan*, **75**, 257–288, https://doi.org/10.1248/cpb.37.3229.

Coniglio, M. C., G. S. Romine, D. D. Turner, and R. D. Torn, 2019: Impacts of Targeted AERI and Doppler Lidar Wind Retrievals on Short-Term Forecasts of the Initiation and Early Evolution of Thunderstorms. *Mon. Wea. Rev.*, **147**, 1149–1170, https://doi.org/10.1175/MWR-D-18-0351.1.

Dee, D. P., and A. M. da Silva, 2003: The Choice of Variable for Atmospheric Moisture Analysis. *Mon. Wea. Rev.*, **131**, 155–171, https://doi.org/10.1175/1520-0493(2003)131<0155:TCOVFA>2.0.CO;2.

Degelia, S. K., X. Wang, and D. J. Stensrud, 2019: An Evaluation of the Impact of Assimilating AERI Retrievals, Kinematic Profilers, Rawinsondes, and Surface Observations on a Forecast of a Nocturnal Convection Initiation Event during the PECAN Field Campaign. *Mon. Wea. Rev.*, **147**, 2739–2764, https://doi.org/10.1175/mwr-d-18-0423.1.

——, ——, ——, and D. D. Turner, 2020: Systematic evaluation of the impacts of assimilating a network of ground-based remote sensing profilers for forecasts of nocturnal convection initiation during PECAN. *Mon. Wea. Rev.*, **in press**, https://doi.org/https://doi.org/10.1175/MWR-D-20-0118.1.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, https://doi.org/10.1029/94jc00572.

Evensen, G., 2003: The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, https://doi.org/10.1007/s10236-003-0036-9.

Fletcher, S. J., and M. Zupanski, 2006: A data assimilation method for log-normally distributed observational errors. *Q. J. R. Meteorol. Soc.*, **132**, 2505–2519.

Hu, J. U. N. J. U. N., N. Yussouf, D. D. Turner, T. A. Jones, and X. U. G. U. A. N. G. Wang, 2019: Impact of ground-based remote sensing boundary layer observations on short-term probabilistic forecasts of a tornadic supercell event. *Wea. Forecasting*, **34**, 1453–1476, https://doi.org/10.1175/WAF-D-18-0200.1.

Kamineni, R., T. N. Krishnamurti, R. A. Ferrare, S. Ismail, and E. V. Browell, 2003: Impact of High Resolution Water Vapor Cross-Sectional Data on Hurricane Forecasting. *Geophys. Res. Lett.*, **30**, 1234, https://doi.org/10.1029/2002GL016741.

NRC, 2009: Observing Weather and Climate from the Ground Up. National Academies Press, 250 pp.

Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, and P. M. Kosro, 2002: Assimilation of surface velocity data into a primitive equation coastal ocean model. *J. Geophys. Res. Ocean.*, **107**, 1–25, https://doi.org/10.1029/2000jc000511.