Author's response to referee #3:

Thank you very much for the thorough, explicit, well organized, and practical comments. The review comes as an opportunity to improve the manuscript in various aspects. For convenience, the response is by order of appearance following the structure of the referee's report.

Referee's comment #1a) The motivation, strengths, and the central question of the paper could be made clearer in the introduction of the paper. As explained in Sect. 2, the region studied is interesting and different aspects are impacting the PBL height. The interesting aspects of the spatial variability of the studied region could be included in the introduction. In the light of the spatial variability, evaluating model performance on a single site would have limited value. One of the strengths of the study is the use of a network of ceilometers that can estimate the temporal development of the PBL at various locations simultaneously. This aspect deserves to be mentioned in the introduction.

Author's response: **Comment accepted.**

Author's changes in manuscript: The abstract and introduction paragraphed were changed accordingly.

Referee's comment #1b) The introduction does not provide enough information to motivate the development of a post-processing tool for the modeled PBL height. The goal to use ceilometer detected PBL height to correct for modeled PBL height could be simplified to "use A to correct B". Currently, it is demonstrated that to "use A" is possible, e.g. PBL height can, with some limitations, be retrieved from the ceilometer measurements. However, "to correct B" is neglected in the introduction. The introduction only states the need for accurate PBL estimate, but no literature on identified shortcomings, methods found for improvement or anything else that would have been done previously to evaluate or improve PBL height estimates in NWP models is presented. Do previous studies suggest that it is more feasible to correct the end product (e.g. the PBL height) than to improve model parametrizations in order to obtain a better result from the model? Do the authors envision a use for the corrected PBL height? The authors could also consider whether their main aim should be on developing a correction, or rather a rigorous evaluation of model performance in the complex region. The latter could be helpful for understanding model shortcomings and would be a more general result than a location and time specific correction.

Author's response: **Comment accepted.**

Author's changes in manuscript: The motivation of this study is to provide air pollution dispersion models with reliable input data of PBL heights. Weather models produce a high spatial and temporal resolution of PBL heights, albeit previous research has shown significant differences between the models' estimations and actual measurements. To overcome this obstacle, we established a correction tool for weather models by employing ceilometer measurements.

Referee's comment #2) One of the confusing aspects of this paper is the small number of days analyzed. The strength of the ceilometer is that data acquisition is cheap (see Sect. 1), however the small dataset is undermining this specific strength. The conclusions drawn are seriously undermined by the small sample size. For example, Sect. 6.2 seems to describe statistical results obtained from 13 data points. If possible, the authors should obtain more data. Alternatively, the study could be shifted to focus on case studies evaluating the shortcomings of the models in more detail. Although the reasons for focusing on daytime PBL only in summer are given, further selection seems to have taken place. Why are only 13 days included from August 2015, and 20 days from August 2016 (L. 292-293) in Sect. 6.1? Why does Sect. 6.2 only include 5 ceilometer sites, when Sect. 6.3 includes 8 ceilometer sites (L. 319-321 and 345-346)? Why do Sections 6.2 and 6.3 only include data from August 2015, and not from August 2016? Are the 13 days used in Sect. 6.2 a subset of the 33 days in Sect 6.1? The authors should provide an explanation for the small number of days analyzed and why certain days and sites were selected at different stages of the study.

Author's response: **Comment accepted.** The ceilometer array in Israel is a collection of ceilometers from different institutes. This study was the first attempt to gather data from all institutions. Unfortunately, some output files are missing. In other cases, the ceilometers operated for short periods. The database further narrowed down by removing days with dust storms or partial data. Eventually, we extracted the maximum days available for each ceilometer within six summer months: July-September 2015, and June-August 2016. We produced additional IFS and COSMO model runs to meet the periods available from the ceilometers. As a result, the analysis expanded from 13 specific days for 5 ceilometers to above 50 days for 6 ceilometers:

Ceilometer	# Days
Bet Dagan	91
Tel Aviv	122
Ramat David	123
Weizmann	55
Jerusalem	53
Nevatim	72

Hence, we combined sections 6.1 and 6.2 in section 5.1. Section 6.3 changed to Sect. 5.2.

Author's changes in manuscript:

The results and conclusions sections were changed considerably, as aforementioned.

Referee's comment #3) Related to the comment above about the amount of data, the authors should consider the statistical significance of the presented results. Specifically, wherever R-values are given (L. 298, Table 3, and elsewhere), the corresponding p-value should also be presented. Other techniques to analyze the statistical significance of the results are also welcomed, and the results should be discussed from the point of view of statistical significance.

Author's response: **Comment accepted.**

Author's changes in manuscript: Statistical analysis of boxplots, histograms, and tables added.

Referee's comment #4) Section 6.2 could provide possibly the most interesting results for considering model performance in terms of PBL height in complex environments. If model under- or overestimation could be connected to certain processes (e.g. the sea breeze), the results would be more generally interesting. Mountainous coastlines are not unique to Israel, and many people inhabit such areas. This section deserves a proper evaluation, and the analysis and discussion should be extended. Specifically, this section is hard to understand for someone not familiar with the geography of Israel. I would advise the authors to consider the presentation of their results. For example, the mean error at each site for each model and method could be presented with a symbol on a map having the color indicating the value. This would make any spatial structures in the mean, mean error (ME) or root mean square error (RMSE) more apparent. The authors could also plot the ME and/or RMSE as a function of the distance of the site to the shoreline and altitude above sea level (these are the two variables used for the correction in the next section). From the authors description of the situation, it seems that the sea breeze has a clear influence on the PBL height. Is it to be understood, that the model does not correctly produce the sea breeze circulation, or is the model lacking in terms of the effect of the sea breeze on PBL height? It would be interesting if the authors could evaluate the discrepancy between ceilometer and model PBL height in terms of the strength, and spatial and temporal development of the sea breeze circulation during the day. Furthermore, in Sect. 6.3 data for 9-14 UTC are used, and I suggest the authors consider including the temporal development of the PBL height in their analysis in Sect. 6.2 as well.

Author's response: **Comment accepted.**

Author's changes in manuscript: As described in response to comment #3, this section was changed dramatically following the referee's suggestions.

Referee's comment #5a) Before a correction is developed and presented, it should be made clear that a correction is needed and that there is a systematic bias that can be corrected for. Table 3 (and Section 6.1) show that the mean error of COSMOR compared to radiosondes is -3 m, which does not leave much room for improvement. Also Table 5 shows that at different sites the mean error of COSMOR is within a few tens of meters at most. (However, I would be cautious to draw conclusions from statistics comprising of 13 data points, and the authors should obtain a larger sample size if possible. See comments 2 and 3). For a 1 km deep PBL, an error of 30 m is 3%. For which application is this not good enough, and how good should the model performance be? Furthermore, considering that the definition of the planetary boundary layer is slightly ambiguous, can a perfect agreement between different methods be expected? The authors should explain why they think the model performance is not good enough and requires improvement. Furthermore, the authors could consider if the correction they presented would actually be more useful for the IFS model that shows clearly worse performance than the COSMO in terms of PBL height prediction.

Author's response: **Comment accepted.**

Author's changes in manuscript: With great effort, we obtained a larger sample size. Now, the necessity to improve the models' PBL heights is evident from the statistical analysis. The primary purpose of this study was to improve the performance of air pollution dispersion models by providing reliable PBL heights from NWP models. In some cases (Uzan et al., 2012), a height difference of 100 m between the actual PBL height and the models' assessments affect ground-level air pollution concentrations significantly. Therefore, the correction tool is useful for both regional and global models.

Uzan, L. and Alpert, P.: The coastal boundary layer and air pollution - a high temporal resolution analysis in the East Mediterranean Coast, The Open Atmospheric Science Journal, 6, 9–18, 2012.

Referee's comment #5b) Sect. 6.2 should demonstrate the basis of the correction presented in Sect. 6.3. The fact that the mean error in Tel Aviv, Beit Dagan and Weizmann are so similar suggests a spatial consistency that is more clear for COSMOR than COSMOP. (Table 5). Is this the reason COSMOR was used for the correction in Sect. 6.3 instead of COSMOP? The fact that there seems to be some spatial structure in the mean error is promising for developing a correction. The RMSE does not seem so spatially consistent.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section reveals which model and method produced the best results following the ceilometers' locations.

Referee's comment #5c) To justify the correction method presented in Sect. 6.3, it should be established that a bias exist in the models' PBL height estimation that depends on altitude and distance from shoreline, that could consequently, be corrected for. The authors should evaluate how the discrepancy between ceilometer and model PBL height depends on the topography and distance from shoreline. Furthermore, this could be done for different hours of the day, as the correction procedure is also applied for each hour separately.

Author's response: **Comment accepted.**

Author's changes in manuscript: Section 6.3 changed to 6.2, including an elaborate explanation of the correction tool performance for a single day study case between 9-14 UTC. Figures for each hour display the models' estimations, PBL heights after correction, and cross-validation examination for Bet Dagan and Jerusalem.

Referee's comment #6) Perhaps the most serious shortcoming of the manuscript is that it is not demonstrated that the model result is better after correction. The authors should include a quantitative evaluation of the improvement of the model PBL after the correction. For example, the radiosondes at Beit Dagan could be used as an independent reference for the model PBL height. Another approach would be to estimate the correction parameters using only some of the available ceilometer stations, and using the remaining stations as a references to estimate the improvement in PBL height achieved by the correction. Varying the number of stations and the locations of the stations included for fitting the correction parameters also give an indicator for how many ceilometers needs to be included, or how they need to be located, for achieving a significant improvement for the COSMOR PBL height. If the authors aim is to show that the ceilometer is a useful tool to improve the modeled PBL height, the strength of their paper relies on the extent and rigor that this kind of analysis is carried out.

Author's response: **Comment accepted.**

Author's changes in manuscript: Cross-validation analysis demonstrated the efficiency of the correction tool. The improvements were discussed in the conclusions according to the new results section (see responses to comments #2 and #5c).

Referee's comment #7a) More attention should be paid to make the reasoning understandable for readers that are not so familiar with the specific geography and climatology of the region. Firstly, the studied region and its interesting aspects could be mentioned in the introduction. The first time the location is given is the very end of the introduction, on line 97. This should be included already in the previous paragraph that outlines the purpose of the study, as well as in the abstract.

Author's response: **Comment accepted.**

Author's changes in manuscript: The spatial variability and locations were added to the abstract and introduction.

Referee's comment #7b) A topography map should be included. Global topography data is available (for example from NOAA https://doi.org/10.7289/V5C8276M) and a map can be drawn using openly available tools (such as python).

Author's response: **Comment accepted.**

Author's changes in manuscript: A topographical map was added.

Referee's comment #7c) Depending on the weight the authors want to give to the humidity (mentioned on lines 103-104) and the prevailing synoptic conditions (line 125), they could also include a map of mean precipitation and pressure in August to help the reader to follow their argumentation.

Author's response: The manuscript modifications doubled the number of figures. Therefore, we preferred to add references instead of maps.

Author's changes in manuscript: Additional references of previous research in Israel describing the dry summer season.

Referee's comment #8) L.1-2: The authors should reconsider the title of the manuscript. The current title is somewhat misleading because it implies that the correction for PBL height was considered for both models, when in the manuscript only the COSMO PBL height was corrected. Furthermore, the journal guidelines recommend avoiding the use of abbreviations in the title, so the authors might want to avoid the use of "NWP" in the title.

Author's response: **Comment accepted.** The research studies two NWP models and established a correction formula feasible for both models. Thus, we find it appropriate to mention IFS in the title as well.

Author's changes in manuscript: "NWP" was removed from the title.

Referee's comment #9) L.23-25: Here results are given for flat and elevated terrain. Consulting Tables 4 and 5 it seems that flat terrain refers to Tel Aviv, and elevated terrain to Jerusalem. The authors should consider mentioning the sites for which the numbers refer to avoid ambiguity, or at least mention that the values presented are from single stations.

Author's response: **Comment accepted.**

Author's changes in manuscript: The titles in the new results section refer to each ceilometer.

Referee's comment #10) The abstract does not mention Israel or give any other indication over the geographic locations apart from "heterogeneous area" and mention of the Beit Dagan radiosonde launch site. Location should be given.

Author's response: **Comment accepted.**

Author's changes in manuscript: Locations were added to the abstract.

Referee's comment #11) L.33-40: Considering that this paragraph states the broad motivation and importance of this study, some references would be appropriate.

Author's response: **Comment accepted.**

Author's changes in manuscript: References were added.

Referee's comment #12) L.56-57: "ceilometers obtain a wide spatial resolution per lidar" - I'm afraid I do not understand the meaning of this phrase. Perhaps the authors mean that a wider spatial resolution is achieved by ceilometers than lidars?

Author's response: **Comment accepted.**

Author's changes in manuscript: The text was changed accordingly.

Referee's comment #13) L.53-65: This paragraph seems to suggest that ceilometers are better than lidars in every aspect. It would be fair to mention a shortcoming of the ceilometer compared to a lidar.

Author's response: **Comment accepted.**

Author's changes in manuscript: The shortcomings of ceilometers were added to the introduction section.

Referee's comment #14) L.89-91: It is not obvious here why the summer season is more appropriate for an approach that is limited by precipitation. It is later explained that this season has low precipitation. This should also be mentioned here to help the readers not familiar with local climatology.

Author's response: **Comment accepted.**

Author's changes in manuscript: The meteorological conditions were added to the introduction section.

Referee's comment #15) L.92-97: It would be possible to help the reader further by outlining the structure of Sect. 6, either here or at the beginning of Sect. 6.

Author's response: **Comment accepted.**

Author's changes in manuscript: The outline was elaborated accordingly.

Referee's comment #16) L.85-86: The introduction demonstrates the strengths of ceilometers compared to other available observational techniques to estimate PBL height, but only states that ceilometers have not been used often for evaluating model performance. However, other observational techniques have, and this should be mentioned. Specifically, have other observational tools been used for evaluating PBL height in NWP models in Israel, or other mountainous coastlines?

Author's response: **Comment accepted.**

Author's changes in manuscript: Information regarding the observational tools implemented for COSMO PBL height evaluation was added in the introduction section.

Referee's comment #17) I find the extent of presenting the literature for the use of ceilometer to detect PBL height satisfactory. However, no mention of previous work using ceilometer to derive PBL height in Israel is presented. The authors should site at least Uzan et al. (2016) and any other studies employing the measurement technique in their region of study.

Author's response: **Comment accepted**. As discerned by the referee, we were first to employ ceilometers for PBL height detection in Israel (Uzan el al, 2016). Up until our research, the ceilometers' in Israel were acknowledged merely as ceiling height detectors. Thus, historical data had neither been acquired or saved. The data we received was collected following our specific request. It was the maximum amount of data available. This explains the inevitable situation of low data availability for spatial analysis limit to the summer season.

Author's changes in manuscript: Uzan et al. (2016) was cited in the introduction.

Referee's comment #18) L.106: "IMS weather reports" - The authors should provide a more specific reference, if possible.

Author's response: **Comment accepted.**

Author's changes in manuscript: **"**Israeli Meteorological Service relative humidity climate report 1995-2009, https://ims.gov.il/en/ClimateReports**".**

Referee's comment #19) L.100-103: Here could cite Fig. 1.

Author's response: **Comment accepted.**

Author's changes in manuscript: Fig 1 was cited accordingly.

Referee's comment #20) L.111: PBL height detection becomes increasingly difficult with increasing range (because of the decrease in the signal-to-noise ratio), and because of the low power of the ceilometer deep boundary layers are hard to detect. The moderate PBL height means that it is less of an issue in this study, and the authors could mention this to support their choice of instrumentation.

Author's response: **Comment accepted.**

Author's changes in manuscript: The comment was added to the text. Thank you.

Referee's comment #21) L.112-115: "Summer dust outbreaks in the eastern Mediterranean are quite rare (Alpert and Ziv 1989, Alpert et al., 2000) therefore, they were not addressed here, especially in the height levels below 1 km (Alpert et al., 2002)." - The sentence structure is unclear. Do the authors mean that especially dust outbreaks below 1 km were not addressed, or perhaps that the dust outbreaks below 1 km were especially rare and therefore not addressed? Should be clarified.

Author's response: **Comment accepted.**

Author's changes in manuscript: The sentence was clarified in the text.

Referee's comment #22) L.119: The abbreviation LST is not defined.

Author's response: **Comment accepted.**

Author's changes in manuscript: LST = UTC+2 was added to the text.

Referee's comment #23) L.116-138: This is a paragraph about PBL structure and development in the studied region based on literature. It is useful and informative, even though it is concise and provides a lot of information for someone not familiar with the region. This paragraph is crucial for understanding the results, and the authors should not be afraid to extend if necessary to better understand the results. They should also refer back to this section at later parts of the manuscript when the concepts described are discussed. Furthermore, Fig. 3b could also be referred to as an example to aid the description of the diurnal cycle.

Author's response: **Comment accepted.**

Author's changes in manuscript: Changes were made according to the figures and text of the new results section.

Referee's comment #24) L.116-138: The use of abbreviations seems excessive: SBF and RL are only used once after being introduced, and could therefore omitted. Also CBL and SBL are only used 1-2 times after this paragraph and the need for the abbreviations is questionable and does not aid readability of the manuscript.

Author's response: **Comment accepted.**

Author's changes in manuscript: The abbreviations-SBF, SBL, CBL, and RL were removed.

Referee's comment #25) L.136-138: Please provide reference(s) for nocturnal PBL in Israel, if available.

Author's response: Previous studies of the nocturnal PBL in Israel were conducted in regions not in the scope of our research, therefore, they were not cited.

Author's changes in manuscript: No change was made.

Referee's comment #26) Sect. 4.1: The placement of ceilometers in the heterogeneous research area should be described. Do the ceilometer sites adequately represent the variability of the region? Are the different regions mentioned in the text (humid, arid, coastal, complex terrain) covered by the measurements?

Author's response: **Comment accepted.**

Author's changes in manuscript: The region of the ceilometers was added in the relevant sections.

Referee's comment #27) Sect.5.3: The ceilometer backscatter profile is related to the aerosol loading, and therefore the layer that is detected is actually an aerosol layer. Implicit in the method described is the assumption that the PBL height corresponds to the height of the aerosol layer directly above ground. This assumption should be stated, and potential consequences to the results discussed. It is especially a limitation for detecting internal boundary layers which might develop due to the sea breeze circulation or katabatic winds.

Author's response: **Comment accepted.**

Author's changes in manuscript: An explanation was added to Sect. 1.

Referee's comment #28) L.143: Table 2 is mentioned before Table 1 in text, the order of the tables should be swapped.

Author's response: The explanation of the research area was moved to the introduction section therefore, it wasn't necessary to swap the table numbers.

Author's changes in manuscript: No change made.

Referee's comment #29) L.156: The authors could consider using the word "increased" rather than "improved" because it is more neutral. Although the model performance might have improved in important aspects due to increase in resolution, the computational cost likely did not.

Author's response: **Comment accepted.**

Author's changes in manuscript: The section was rephrased.

Referee's comment #30) L.163-164: "The spatial resolution of the models affects their ability to refer to the actual topography rather than a smoothed grid point." Is this the reason that the ceilometer site is used as a parameter for correction? If so, it should be clarified.

Author's response: **Comment accepted.**

Author's changes in manuscript: An explanation was aded to the new summary and conclusions section.

Referee's comment #31) L.164-165: "the models' results were corrected by the actual ground base heights for each measurement site" - Unfortunately, I cannot follow here. Presumably the correction meant here is not the correction presented in Sect. 6.3. Perhaps the authors mean that the model levels were adjusted based on the precise altitude of each ceilometer station? Clarification would be appreciated.

Author's response: **Comment accepted.**

Author's changes in manuscript: Additional text: "Therefore, the models' levels were adjusted based on the precise altitude of each ceilometer station."

Referee's comment #32) L.144-162: Considering that IFS provides boundary conditions for COSMO, and that the description of the COSMO model refers to IFS model parameterizations, the authors could consider switching the order of introducing the two models. e.g. move lines 156-165 before line 144.

Author's response: **Comment accepted.**

Author's changes in manuscript: The order was changed. IFS was introduced before COSMO.

Referee's comment #33) L.157: It seems that the IFS has more vertical levels, but does it have better vertical resolution in the boundary layer? Information on vertical resolution should be added in Table 2.

Author's response: **Comment accepted.**

Author's changes in manuscript: The information was added to Table 2.

Referee's comment #34) L.188-189: "In order to derive the backscatter coefficient from ceilometer measurements, signal calibrations and water vapor corrections are necessary" - It is not clear if the corrections were done (presumably not), and should be clarified.

Author's response: **Comment accepted.**

Author's changes in manuscript: The sentence was rephrased.

Referee's comment #35) L.193-194: It could be mentioned that averaging multiple profiles improves the signal-to-noise ratio and thereby is likely to also improve the detection of the PBL height.

Author's response: **Comment accepted.**

Author's changes in manuscript: The sentence was rephrased.

Referee's comment #36) L.197: The overlap effect is a well-known issue for lidar systems, however, the authors could provide a reference.

Author's response: **Comment accepted.**

Author's changes in manuscript: A reference was added: "At these heights, a constant perturbation existed due to the overlap of the emitted laser beam and the receiver's field of view (Weigner et al., 2014)".

Referee's comment #37) L.215-217: "the radiosonde's horizontal position is under 0.01° which is an order of magnitude from the models' grid resolution" - This is true for IFS but not for COSMO, which has a resolution of 0.025°. The authors should be more specific to avoid a misleading statement.

Author's response: **Comment accepted.**

Author's changes in manuscript: The text was rephrased.

Referee's comment #38) L.239-241: The method used for COSMO, why two different thresholds are needed, and how it differentiates from that used in for IFS or the radiosondes is not clear. What is the reason for applying a different criterion for COSMO than the IFS and soundings?

Author's response: **Comment accepted.** IFS adapted a single threshold of 0.25 following the conclusions of (Seidal et al.,2015). The COSMO model refers to 0.33 for stable atmospheric conditions (Wetzel, 1982), and 0.22 for unstable conditions by 0.22 (Vogelezang and Holtslag, 1996).

Author's changes in manuscript: The information was added to the text.

Referee's comment #39) L.282-283: "This height indicates the entrainment zone rather than the actual cloud top." For anything than the most optically thin clouds, the ceilometer signal attenuates before reaching the cloud top. Therefore, the ceilometer is very unlikely to be detecting cloud top.

Author's response: **Comment accepted.** We must clarify we didn't attempt to claim the ceilometer detects the cloud top. On the contrary.

Author's changes in manuscript: The sentence was rephrased to avoid the misunderstanding.

Referee's comment #40) L.292-293: Considering the change in IFS resolution between 2015 and 2016, is it appropriate to evaluate the IFS data together, or should data from 2015 be considered separately from 2016?

Author's response: In 2015 and 2016 the ceilometers were indicated by the same grid points and horizontal levels. Therefore, we did not find it necessary to separate the results. Furthermore, we ran the analysis separately for 2015 and 2016. The difference between the results was insignificant.

Author's changes in manuscript: No changes made.

Referee's comment #41) L.310-314: In the introduction it is mentioned that Ketterer et al. (2014) found poor correlation between ceilometer PBL height and the PBL height from COSMO. Why is their result so different from that found here?

Author's response: **Comment accepted.** The main difference is the research area. Ketterer et al., (2014) studied complex topography of the Swiss Alps (two sites, 3,580 m a.s.l and 2,061 m a.s.l), whist our stud region was confined between the shoreline to highest point of 830 m a.s.l.

Author's changes in manuscript: To avoid a too-long introduction section, we moved the discussion of previous research (Ketterer et al., 2014 and Collaud et al., 2014) to the results section.

Referee's comment #42) As far as I can see in Fig. 2, the gap between IFSP and RS is even larger for the data point indicated by the red rectangle in the figure below. I appreciate that the authors give an explanation to the anomalous PBL height on the 17 Aug 2016, but I'm concerned that this paragraph is slightly misleading. I'm not convinced that the difference between the IFSP and RS is the largest on 17 Aug 2016. I suggest the authors re-formulate this paragraph with the emphasis on giving an explanation for the anomalous PBL on 17 Aug 2016, rather than claiming this is the day with largest discrepancies, or alternative provide an objective measure for a "largest gap" and an explanation why the large discrepancy in IFSR is worth considering but the even larger discrepancy in IFSP on another day is omitted. Based on the next section, I could guess that these data points indicated by the red box are from 10 Aug 2015 (Fig 4b). If so, please include this information in this section of the manuscript.

Author's response: **Comment accepted.** The new results section consists of new figures according to the referee's comments 2-6.

Author's changes in manuscript: The data of Fig. 2 and new for other ceilometers was analyzed to produce new compelling figures.

Referee's comment #43) Sect 6.1. No discussion about the differences between bulk Richardson and parcel method is included. From Tables 4 and 5 it seems like IFS results are more sensitive to the choice of method. Perhaps the authors could discuss these results.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section consists of a discussion on the different methods.

Referee's comment #44) Sect 6.1: As far as I can understand, the main purpose of this chapter is to demonstrate the feasibility of ceilometer measurements to use for model evaluation. The authors could consider using this 33 point data set to compare the model results to the ceilometer to see if the results are similar than those obtained in comparison with the radiosondes to give additional confidence.

Author's response: **Comment accepted.**

Author's changes in manuscript: The results section was changed accordingly.

Referee's comment #45) L.324-330: If the 13 days evaluated in Sect 6.2. are also included in the analysis of Sect 6.1, this paragraph does not provide any new information. For the clarity of the manuscript, I would advise the authors to include all comparison of radiosonde with other data in Sect 6.1, and focus on the spatial analysis in Sect. 6.2, as indicated by the title.

Author's response: **Comment accepted.**

Author's changes in manuscript: The results section was changed accordingly.

Referee's comment #46) L.331: "By and large, COSMOR achieved the best statistical results" - This statement seems overemphasized. In terms of root mean square error, COSMOP performed better on 4 of the 5 sites presented, and the mean error was better for 2 sites.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section consists of a discussion on the results of each model by each method.

Referee's comment #47) L.336-349: "These results emphasize the advantage of high-resolution regional models such as COSMO (~2.5 km resolution) over the IFS global model (resolution of ~13 km in 2015 and ~10 km in 2016) over a diverse area." Although not necessarily surprising, this is one of the few clear results of the paper, and deserves to be discussed and possibly further analyzed. Is the poor performance of the IFS related to lacking representation of the sea breeze circulation or some local scale phenomena?

Author's response: **Comment accepted.**

Author's changes in manuscript: An explanation was added to the new summary and conclusions section.

Referee's comment #48) Sect. 6.1 and 6.2: Did the authors consider the differences between the bulk Richardson and parcel method, and whether it indicates certain shortcomings in the models description of the boundary layer structure or processes? Comparing the COSMOR and COSMOP mean errors presented in Table 5, it seems that the two methods produce more similar results more inland (Ramat David and Jerusalem) than closer to the coast (Tel Aviv, Beit Dagan, Weizmann). This seems to also hold for the IFS. Is this related to the meteorological conditions, or simply a coincidence? Again, a significantly larger data set would be desirable.

Author's response: **Comment accepted**.

Author's changes in manuscript: The new result section consists of a discussion on the differences between the models.

Referee's comment #49) Sect. 6.2: Why are only 5 sites included, if ceilometers are available at 8? No station with the description "South" is included in the analysis of spatial variability (Table 1, L. 320), do the included 5 ceilometer sites adequately represent the spatial variability of the studied region?

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section refers to these comments. See response for comment #2.

Referee's comment #50) L.342-344: "Following the conclusions of previous stages, COSMOR was chosen as the model and method that achieved the best results." In my opinion, this was not well demonstrated (see also comment 46).

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section includes a discussion of the results by models, methods, and location of the measurement sites.

Referee's comment #51) L.344: I'm guessing that the time window chosen is somehow related to the diurnal PBL height cycle that was nicely described in Sect. 2. Please provide explanation for the time chosen.

Author's response: **Comment accepted.** In the summer season, stable conditions prevail from sunset to an hour after sunrise (Stull, 1988). At this period the models' R_b profiles do not accede the relevant thresholds, and the PBL height is not detected. Subsequently, the analysis fixated on the day time hours, after sunrise and before sunset

Author's changes in manuscript: An explanation was added the results section**.**

Referee's comment #52) Fig. 4: How are daily values obtained? Is the procedure the same as in Sect. 6.1, e.g. estimating the PBL height at approximately 11 UTC? If so, it should be mentioned in the text.

Author's response: **Comment accepted.**

Author's changes in manuscript: Fig. 4 was replaced by new figures following the referee's recommendations to expand the dataset. See response for comment #2.

Referee's comment #53) L.349-357: I'm not sure I understand the correction procedure. First, the variables α, β and γ are obtained by using the mean error (ME) between model and ceilometer at each station, and the altitude and distance from shoreline as predictor variables. After α, β and γ are obtained, it is possible to estimate ME anywhere in the domain. The corrected PBL height is then the COSMOR PBL height+ the ME that is computed using altitude, distance from shoreline and α, β and γ. The same procedure is repeated for each hour, resulting in a time dependent α, β and γ. Is this a correct interpretation? The authors should clarify the description of their method.

Author's response**: Comment accepted.**

Author's changes in manuscript: The explanation of the correction tool was changed accordingly.

Referee's comment #54) L.349-357: Could the authors report the values of α, β and γ? The choice of repeating the correction for each hour of the day suggest some dependence of the correction needed on the diurnal cycle, does that exist? Do α, β and γ vary from hour to hour? What is the role of γ in the equation, and is it really needed? Presenting α and β would show whether altitude (e.g. topography) or distance from the shoreline (e.g. sea breeze circulation?) contributes more to the model discrepancy.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section provides the dependent variables α, β, and the constant γ for each hour (9-14 UTC) for three scenarios: regression by eight ceilometers, regression by seven ceilometers excluding the plain site of Bet Dagan, regression by seven ceilometers excluding the elevated site of Jerusalem.

Referee's comment #55) L.358: Is the cross-section along a fixed longitude?

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results delineate PBL heights from all ceilometers by distance from the shoreline.

Referee's comment #56) L.369-370: "The lowest value was corrected from 09 UTC (11 LST) to 14 UTC (16 LST)" - The way I understand this sentence is that the lowest value was before the correction at 9 UTC, and after the correction it was at 14 UTC. This seems to contradict Fig. 5, which shows the opposite. Comparing Figures 5 a and b, it seems that the uncorrected data had the lowest PBL height at 14 UTC (independent of longitude). After the correction, at longitudes eastward of 35.1° (where Jerusalem lies) the lowest PBL height is found at 9 UTC. It would be advisable for the authors to clarify their statement.

Author's response: **Comment accepted**. New figures in the results section clarify the results of the correction tool, hour by hour, from all ceilometer sites.

Author's changes in manuscript: New figures and explanations.

Referee's comment #57) Line.403: "which improved the description of the diurnal PBL heights" - Unfortunately, there is no evidence presented that the model performance would have improved. See comment 6.

Author's response: **Comment accepted**.

Author's changes in manuscript: The new figures and explanations provide the required evidence.

Referee's comment #58) Conclusions: The authors could discuss how the results obtained for daytime in a summer month might compare to other seasons.

Author's response: **Comment accepted.** The correction tool is relevant for all dates excluding days with precipitation or dust storms.

Author's changes in manuscript: The comment was added in the new discussion and conclusions Sect.

Referee's comment #59) Table 1: Height limit is given as 7.7 or 15.4 km, but the footnote states that the data acquisition was limited to 4.5 km. It is not clear what is the vertical extent of the measurement. Although it is not that important for the study, the presentation is confusing and could be clarified.

Author's response: **Comment accepted.** The explanation referred to the difference between the ceilometer's capabilities (hardware) to measure up to 7.7 or 15.4 km, and the actual height ranges of the database. Data acquisition is obtained by the ceilometer's software, which organizes daily profiles up to a specific height limit defined by the user. In our case, the profile height limit was 4.5 km, except for 7.7 in Bet Dagan site.

Author's changes in manuscript: Table 1 was clarified.

Referee's comment #60) Table 1: The table includes specifications for the sites such as "north", "south", "inland", "mountain", but these do not seem to be defined or used elsewhere in the manuscript. Perhaps the regions could provisionally be indicated on a map, and used in the discussion of the results.

Author's response: **Comment accepted.**

Author's changes in manuscript: A topographical map was added and reference to the regions of each site was included in the results and conclusions sections.

Referee's comment #61) Table 3: For completeness, the table could include the mean and standard deviation also from the radiosonde used as a reference.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section included the mean and standard deviation for six ceilometer sites including radiosonde Bet Dagan.

Note: the comments numbering skip from 61 to 70.

Referee's comment #70) Table 4: "The PBL heights were compared to the heights measured by the Beit Dagan ceilometer." The text states (lines 321-322) "the models' results were compared to the ceilometers' measurements in each site". These two statements seem to contradict each other, and I would ask the authors to correct one of them, or to clarify why different comparison measurements are considered in the text and in the table.

Author's response: **Comment accepted.** The clerical error was in the title of the table. Sorry about that.

Author's changes in manuscript: The tables and titles were changed.

Referee's comment #71) Tables 4 and 5: It would be interesting to also see the mean PBL height of the ceilometer (the reference) at each site.

Author's response: **Comment accepted.**

Author's changes in manuscript: The new results section included the mean and standard deviation for 6 ceilometer sites.

Referee's comment #72) Figures 1 and 6: Considering the political situation in some areas of Western Asia, the authors should carefully consult the journals guidelines regarding maps.

Author's response: **Comment accepted.**

Author's changes in manuscript: The maps were adapted accordingly.

Referee's comments:

Comment #73) Fig 3a: The figure could contain the PBL height estimated by the two methods. It would be helpful to demonstrate the performance of the two methods.

Comment #74) Fig 3b: It does not look like the data has been averaged for 30 min. Is the data presented at original 15 sec resolution? Please clarify in the caption.

Comment #75) Fig 3b: The authors should consider showing the time series of ceilometer and model based PBL height in this figure. It would be interesting to see 1) how the wavelet covariance transformation method is performing on the time series presented, 2) how the models predict the temporal development of the PBL height, and 3) whether the difference between model and ceilometer is random or the two models and two methods are consistently over or underestimating the PBL height during this one day. Although it might seem trivial to the authors, this helps the reader to gain confidence in the methods and helps with the understanding of the diurnal cycle of the PBL that is described in Sect.2.

Comment #76) Fig 3c: The results presented here are not discussed. A description of the results presented here, and the ways they help to interpret Fig. 3 a and b or other results should be added. Furthermore, the wind direction figure could be improved by shifting the x-axis so that it is centered around North (e.g. scale from 180 to 360/0 to 180 degrees).

Comment #77) Fig 4: Figure 4 is hardly mentioned in the manuscript (it is referred to in the caption of Table 4, and Fig 4b is mentioned on line 326). Consequently, it is not clear what this figure is communicating. What is the additional information provided that is not already presented in Fig. 2? The better performance of COSMO compared to IFS, and the good agreement of ceilometer and radiosonde (Fig. 4b) are already demonstrated in Sect. 6.1.

Comment #78) Fig. 5: Figure 5 could indicate the locations of the Tel Aviv and Jerusalem ceilometer stations, as well as the mean (and standard deviation) of the PBL height estimated at these sites.

Comment #79) Fig. 5 and 6: I don't think it is necessary to list the sites and number of days used for the analysis in each figure caption. In my opinion simply a reference to the text for more details would do.

Comment #80) Fig. 6: Figure 6 could include the information of the mean PBL height at the stations.

Comment #81) Fig. 6b: It is not clear what variable is presented in Fig 6b. Is it the ME estimated based on Equation 6, or one of the fitted parameters (α, β, γ)?

Author's response:

Comments accepted.

Author's changes in manuscript: Fig 1-6 were replaced.

Referee's comment #82) Citations: The authors should check their citations and list of references list. For example, Uzan et al. (2012) and Uzan et al (2018) are cited but missing from the reference list.

Author's response: **Comment accepted.**

Author's changes in manuscript: Previous citations were checked, and new citations added.

Referee's comment #83) Figures: The authors should pay attention to the quality of figures. The font size could be increased in almost all figures (especially hard to read is Fig. 3), and use of color-blind friendly colors should be considered.

Author's response: **Comment accepted.**

Author's changes in manuscript: New figures are provided.

Ceilometers as planetary boundary layer height detectors and a corrective tool for COSMO~~ECMWF~~ and ~~COSMO NWP~~IFS models

Leenes Uzan[1,2], Smadar Egert[1], Pavel Khain[2], Yoav Levi[2], Elyakom Vladislavsky[2], Pinhas Alpert[1]

[1] Porter School of the Environment and Earth Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Dept. of Geophysics, Tel-Aviv University, Tel Aviv, 6997801, Israel.
~~Tel-Aviv University, Tel Aviv, 6997801, Israel.~~

[2] The Israeli Meteorological Service, ~~Beit~~Bet Dagan, Israel.

Correspondence to: Leenes Uzan ~~(Leenesu@gmail.com)~~(Leenesu@gmail.com)

Abstract

The ~~growing importance~~significance of the planetary boundary layer (PBL) height detection is apparent in various fields, ~~from~~ especially in air pollution ~~analysis to weather prediction. Here, we demonstrate the capability of ceilometers to serve as a validation tool for the models' PBL height estimations. The study focused on the daytime summer PBL heights over a heterogeneous area. Height values from two numerical~~dispersion assessments. Numerical weather models~~, the global IFS model,~~ produce a high spatial and ~~the regional COSMO model were evaluated against actual measurements from~~ temporal resolution of PBL heights albeit, their performance requires validation. This necessity is addressed here by an array of 8 ceilometers, a radiosonde, and ~~eight~~ two models - IFS global model and COSMO regional model. The ceilometers~~. The evaluation of the PBL heights was attained by the bulk Richardson method and the parcel method. The ceilometers' backscatter profiles~~ were analyzed by the wavelet covariance transform method~~. A comparison of the PBL heights at 11 UTC on 33 summer days in Beit Dagan radiosonde launch site revealed a good agreement between the radiosonde,~~ and the ~~adjacent ceilometer (mean error = 12 m, RMSE = 97 m). Spatial analysis on 13 days~~radiosonde and models by the parcel method and the bulk Richardson method. Good agreement for PBL height was found between the ceilometer and the adjacent Bet Dagan radiosonde (33 m a.s.l) at 11 UTC launching time (N = 91 days, ME = 4 m, RMSE=143 m, R=0.83). The models' estimations were then compared to the ceilometers' results ~~from~~in an additional five ~~ceilometer sites showed COSMO evaluations by the bulk Richardson method ($COSMO_R$) produced good results for both flat (mean error = 19 m, RMSE = 203 m) and elevated terrain (mean error = -6 m, RMSE = 251 m). To correct $COSMO_R$ height estimations, a regression tool was generated based on the PBL height difference between $COSMO_R$ and~~

~~eight~~diverse regions where only ceilometers ~~from diverse sites. The independent predictor variables are the topography and the~~ operate. A correction tool was established based on the altitude (h) and distance from ~~the~~ shoreline~~. The correction factors are implemented on the COSMO~~$_{R}$ (d) of eight ceilometer sites in various climate regions, from the shoreline of Tel Aviv (h = 5 m a.s.l, d = 0.05 km), to eastern elevated Jerusalem (h = 830 m a.s.l, d = 53 km), and southern arid Hazerim (h = 200 m a.s.l, d = 44 km). The tool examined the COSMO PBL height approximations based on the parcel method. Results for August 14, 2015 case-study, between 9-14 UTC showed the tool decreased the PBL height in the shoreline and inner strip of Israel by ~ 100 m and increased the elevated sites of Jerusalem up to ~ 400 m, and Hazerim up to ~ 600 m. Cross-validation revealed good results without Bet Dagan. However, without measurements from Jerusalem, the tool underestimated Jerusalem's PBL height ~~results~~up to ~600 m difference.

1. Introduction

In ~~this~~the era of ~~heavy industrialization~~substantial industrial development, the need to mitigate the detrimental effects of air pollution exposure is unquestionable~~.~~ (Anenberg et al., 2019, WHO, 2016, Héroux et al., 2015, Dockery et al.,1993). However,~~ in order~~ to regulate and establish environmental thresholds, a comprehensive understanding of the air pollution dispersion processes is necessary~~. One of the key~~ (Luo et al., 2014, Seidel et al., 2012, Seidel et al., 2010, Ogawa et al., 1986, Lyons, 1975). One of the critical meteorological parameters governing air pollution dispersion is the planetary boundary layer (PBL) height~~.~~ (Sharf et al., 1993, Garratt, 1992, Ludwing, 1983, Dayan et al., 1988). The PBL height is classified as the first level of the atmosphere ~~which~~that dictates the vertical dispersion extent of air pollution (Stull, 1988). ~~Consequently~~Hence, the ~~concentration level~~quality of ~~air pollution varies depending on the height~~meteorological data provided to these models is of ~~the PBL.~~great importance (Urbanski et al., 2010, Scarino et al., 2014, Su et al., 2018).

~~Applicable evaluation of PBL heights can be derived either by actual measurements or estimations based on numerical~~Numerical weather prediction (NWP) models~~. On the one hand, NWP models, such as regional models,~~ provide a high temporal and spatial ~~data~~ resolution ~~beyond the capability of actual measurements. On the other, they are~~ of PBL height based on mathematical equations with initial assumptions, and boundary conditioned set beforehand. ~~Hence~~However, the ~~models' products require a systematic~~ models display difficulty to accurately simulate the PBL creation and evolution (Luo et al., 2014, Seidal et al., 2010), and validation ~~tool based on~~against actual measurements~~.~~

~~There are two main PBL height measurement methods: in~~ is advised. In situ ~~radiosonde launches and remote sensing such as lidars and profilers. Unfortunately, radiosonde launches are~~atmospheric measurements by radiosondes are most efficient but costly as successive measurements. ~~Profilers~~Remote sensing measurements such as wind profilers and sophisticated lidars ~~produce high temporal resolution profiles but are~~ are mostly designated for specific campaigns limited in ~~space. Moreover, certain meteorological conditions may reduce their performance, such as precipitation for radio acoustic sounding system profilers (Uzan et al., 2012) and dust storms for Raman lidars (~~location and operational time (Manninen et al., 2018, Mamouri et al., 2016).

~~These limitations have led several research groups to successfully utilized ceilometers - single wavelength cloud base height detectors, as a means to recognize and determine the PBL height (Eresmaa et al., 2006, Haeffelin and Angelini, 2012, Wiegner et al., 2014). Ubiquitous~~Ceilometers, on the other hand, are ubiquitous in airports and meteorological service centers worldwide~~, ceilometers obtain a wide spatial resolution per lidar (for further information see~~ (TOPROF of COST Action ES1303 and E-PROFILE of the EUMETNET Profiling Program~~). They produce high temporal resolution profiles about every 15 s and every 10 m, up to several km, retrieved as attenuated backscatter signals. The ceilometers are low cost, easy to maintain, and operate continuously unattended~~), thus provide an advantage over the relatively scarce deployment of sophisticated lidars.

Ceilometers are single wavelength micro-lidars intended for cloud base height detection. Vaisala ceilometers produce backscatter profiles every ~15 s with a vertical resolution of 10 m and a height range up to 8 or 15 km, depending on the ceilometer type and the atmospheric conditions (Uzan et al., 2018). Unlike sophisticated lidars, ceilometers are not equipped to provide aerosol properties such as size distribution, scattering, and absorption coefficients (Ansmann et al., 2011, Papayannis et al., 2008, Ansmann et al., 2003). Nevertheless, their advantages have been recognized as low cost, easy to maintain, and continuous unattended operation under diverse meteorological conditions (Kotthaus and Grimmond, 2018). ~~These qualities reflect their advantages over high-cost, multi-wavelength sophisticated lidars, that require surveillance, calibration procedures,~~Over the years, several studies have assigned ceilometers as PBL height detectors (Eresmaa et al., 2006, Van der Kamp and ~~careful~~

maintenance. Hence, they are limited in space and operational time (Mamouri et al., 2016) and cannot achieve the spatial and temporal measurements coverage essential to validate the PBL heights generated by NWP models.

Gierens et al (2018) established a PBL height algorithm applied to the ceilometers' profiles. The PBL height was classified according to daytime convective mixing and nighttime stable surface layer accompanied by a residual layer aloft. Their research was conducted in northwestern South Africa from October 2012 August McKendry, 2010, Haeffelin and Angelini, 2012, Wiegner et al., 2014, showed good agreement with ERA-Interim reanalysis.

Another operational PBL height detection method was established by Collaud Coen et al. (2014). Their study, implemented on a two-year data set for two rural sites located on the Swiss plateau, included several remote sensing instruments (wind profiler, Raman lidar, microwave radiometer) and several algorithms (the parcel method, the bulk Richardson number method, surface-based temperature inversion, and aerosol or humidity gradient analysis). The results were validated against radio-sounding measurements and compared to the NWP model COSMO-2 (2.2 km resolution). In this research, the authors recommended using ceilometers as complementary measurements of the residual layer.

Ketterer et al. (2014) focused on the development of the PBL in the Swiss Alps by an adjacent ceilometer, wind profiler, and in-situ continuous aerosol measurements. The ceilometer's profiles were analyzed by the gradient and STRAT-2D algorithms. Good agreement was found between the PBL height derived from the ceilometer and wind profiler during the daytime and under cloud-free conditions. However, comparisons to the calculated PBL heights from the COSMO-2 model yielded low correlations.

Despite this extensive research, so far). Previous research employed ceilometers as PBL height detectors and compared them to NWP models (Collaud et al., 2014, Ketterer et al., 2014, Gierens et al., 2018). However, scarce attention has been paid to designate ceilometers asfor a correction tool for NWP PBL height assessments.heights. The main goal of this study was to evaluate the estimations of the models for the daytime summer PBL heights over complex terrain by comparing the results against remote sensing measurements from ceilometers. Ceilometers produce aerosol backscatter profiles, therefore, the evaluation of the PBL height during precipitation episodes becomes difficult (is to create this tool and improve the input data for air pollution dispersion evaluations. A description of the models and instruments applied is given in Collaud et al. 2014, Ketterer et al. 2014, Kotthaus & Grimmond 2018). Accordingly, this study focused on the summer season.

The research area and time period are explained in Sect. 2. The models and instruments applied are described in 2 and Sect. 3 and Sect. 4, respectively. Sect. The 4 presents the PBL height detection methods are presented. Spatial and temporal analysis of the PBL heights generated by the models and instruments in six sites are shown in Sect. 5. Results of NWP models compared to1. The PBL height correction tool is explained in situ radiosonde 5.2 and demonstrated by a case-study employing eight ceilometer measurements are presented in Sect. 6. Finally, summarysites. Summary and conclusions are drawn in Sect. 7 regarding the capabilities of NWP models and the evolution of the daytime summer PBL height over Israel. 6.

2. Research area

1.1 Study time and region

Located in the East Mediterranean, Israel obtains a heterogeneous research areavarious climate measurement sites in comparatively short distances, (Fig. 1). The ceilometer array (Fig. 1, Table 1) is comprised of mountains and valleys in the north and the east, a two coastline in the west and a desert in the south. This provides a range of meteorological conditions, from the humid climate on the coast to the arid south.

The Israeli summer season (June-September) is characterized by dry weather (no precipitation), high relative humidity (RH) – up to 80% in midday in the shoreline (Israeli Meteorological Service -IMS weather reports) and sporadic shallow cumulus clouds. On the synoptic scale, the summer is defined by a persistent Persian Trough (either deep, shallow or medium) followed by a Subtropical High aloft (Felix Y., sites, 40 km apart, in Hadera (10 m a.s.l), and Tel-Aviv (5 m a.s.l). Further inland, 12 km and 23 km southeast to Tel Aviv, are Bet1994, Dayan et al., 2002, Alpert et al., 2004). Combined with the sea breeze, the average PBL height is found to be quite low. For example, the average summer PBL height in Beit Dagan (33 m a.s.l and 7.5 km east from the shoreline) reaches ~900 m a.g.l after sunrise, and before the entrance of the sea breeze front (Felix) and Weizmann (60 m a.s.l), respectively. About 70 km southwest to the elevated Jerusalem site (830 m a.s.l) are Hazerim (200 m a.s.l) and Nevatim (400 m a.s.l). Ramat David (50 m a.s.l) represents the northern region 24 km inland. Y.,1994, Dayan and Rodinzki, 1999, Uzan et al., 2016, Yuval et al., 2019). Summer dust outbreaks in the eastern Mediterranean are quite rare (Alpert and Ziv 1989, Alpert et al., 2000) therefore, they were not addressed here, especially in the height levels below 1 km (Alpert et al., 2002).

Various institutions operate the ceilometers. In several cases, the ceilometers' output files were not methodically saved. In others, the ceilometers worked for limited periods. Following Kotthaus & Grimmond (2018), the analysis concentrated on the dry summer season due to the difficulty of evaluating the PBL height from backscatter signals during precipitation episodes. The database narrowed down by removing dates with partial data or during dust storm events such as the unprecedented extreme dust storm in September 2015 (Uzan et al., 2018). In general, summer dust outbreaks in the eastern Mediterranean are quite rare at the low altitudes (~ 1-2 km) of the PBL height (Alpert and Ziv 1989, Alpert et al., 2000, Alpert et al., 2002). Eventually, the analysis focused the data available from each ceilometer within six summer months: July-September 2015, and June-August 2016.

A characteristic Israeli summer has no precipitation and mainly sporadic shallow cumulus clouds (Ziv et al., 2004, Goldreich, 2003, Saaroni and Ziv, 2000). The dominant synoptic system is the persistent Persian Trough (either deep, shallow, or medium) followed by a Subtropical High aloft (Alpert et al., 1990, Feliks Y., 1994, Dayan et al., 2002, Alpert et al., 2004). Previous research describes the The average summer PBL height is under 2 km a.s.l (Dayan et al., 1988, Feliks 2004) Since backscatter signals decline with height, the conditions of low PBL heights comes as an advantage.

1.2 The summer PBL height

The formation and evolution of the Israeli summer PBL height are as a function of the synoptic and mesoscale conditions, as well as the distance from the shoreline, and the topography. Overall, the diurnal PBL height in the summer season may be portrayed in the following mannerfollows: After sunrise (~, ~ 4-5 local standard time (LST, where LST= = UTC+2)), clouds initially formed over the Mediterranean Sea are advected, advect eastward to the shoreline. As the ground warms up, the nocturnal surface boundary layer (SBL) dissipates, and buoyancy induced convective updrafts instigate the formation of the sea breeze circulation (Stull, 1988). Previous research of the PBL height in Bet Dagan (33 m a.s.l and 7.5 km east from the shoreline) revealed an average height of ~900 m a.g.l after sunrise (Koch and Dayan, 1988, Feliks Y.,1994, Dayan and Rodinzki, 1999, Uzan et al., 2016, Yuval et al., 2019). The

~~entrance of the sea breeze front (SBF) is estimated between 7-9 LST (Felix~~The sea breeze front enters between 7-9 LST (Feliks Y., 1993, Alpert and Rabinovich-Hadar, 2003, Uzan and Alpert, 2012), depending on the time of sunrise and the different synoptic modes of the prevailing system ~~–~~- the Persian Trough (Alpert et al., 2004). Cool and humid marine air hinder the convective updrafts. Clouds dissolve, and the height of the shoreline ~~convective~~ boundary layer ~~(CBL)~~ lowers by ~250 m (~~Felix~~Feliks Y., 1993, ~~Felix~~Feliks Y., 1994, Levi et al., 2011, Uzan and Alpert, 2012). Further inland, the convective thermals continue to inflate the ~~CBL~~boundary layer (Hashmonay et al., 1991, ~~Felix~~Feliks, 1993, Lieman, R. and Alpert, 1993). ~~The sea breeze circulation steers clockwise and the PBL wind speed is enhanced by the west~~West-north-west synoptic winds enhance the sea breeze wind as it steers north-west (Neumann, 1952, Neumann, 1977, Uzan and Alpert, 2012). By noontime (~11-13 LST~~)~~), maximum wind speeds further suppress the ~~CBL~~boundary layer (Uzan and Alpert, 2012). In the afternoon (~13-14 LST), the ~~SBF~~sea breeze front reaches ~30-50 km inland to the eastern elevated complex terrain (Hashmonay et al., 1991, Lieman, R. and Alpert, 1993). At sunset (~18-19 LST), as the insolation diminishes, the potential energy of the convective updrafts weakens, and the ~~CBL~~ boundary layer height drops (Dayan and Rodnizki, 1999). After sunset, as ground temperature cools down, the ~~CBL finally~~boundary layer collapses, and a residual layer ~~(RL)~~ is formed above ~~the SBL~~a surface boundary layer (Stull, 1988~~) as the ground cools down.~~). High humidity and a low ~~RL~~residual layer create low condensation levels, and shallow evening clouds are produced.

2. IFS and COSMO Models

~~The~~ IMS ~~utilizes~~capitalizes two operational models: The European Centre for Medium-range Weather Forecasts (ECMWF) Integrated Forecast System (IFS) global model~~,~~ and the ~~consortium~~COnsortium for ~~small~~Small-scale ~~modeling~~MOdeling (COSMO) regional model~~. Details of each model are given in~~ (Table 2~~.~~).

~~COSMO (~2.5km~~IFS consists of 137 vertical levels. In the years 2015 and 2016 relevant to this study, the grid resolution~~) has been running at the IMS~~ was ~13 km and ~10 km, respectively. It applies a turbulent diffusion scheme representing the vertical exchange of heat, momentum, and moisture through the sub-grid turbulence scale. A first-order K-diffusion closure based on the Monin-Obukhov (MO) similarity theory represents the surface layer turbulent fluxes. The Eddy-Diffusivity Mass-Flux (EDMF) framework (Koehler et al. 2011) describes the unstable conditions above the surface layer.

IMS runs COSMO over the Eastern Mediterranean domain (25-39 E/26-36 N) since 2013~~,~~ with boundary and initial conditions from IFS. ~~It is based on the primitive~~It consists of 60 vertical levels up to 23.5 km and a horizontal grid spacing of 2.5 km (Table 2). Primitive thermo-hydrodynamic equations ~~describing~~represent the non-hydrostatic compressible flow in a moist atmosphere (Steppeler et al., 2003, Doms et al., 2011, Baldauf et al., 2011). ~~Its vertical extension reaches 23.5 km (~30 hPa) with 60 vertical levels.~~ The model runs a two-time level integration scheme, based on a third-~~-~~order of the Runge–Kutta method, and a fifth-order of the upwind scheme for horizontal advection. Unlike ~~IFS,~~in the ~~deep convection parametrization is switched off, while~~IFS model, in the COSMO model, only ~~the~~ shallow convection is parameterized, and the deep convection is switched off (Tiedtke, 1989). The turbulence scheme~~, based on~~ of Mellor and Yamada (1982) at ~~Level~~level 2.5, uses a reduced second-order closure with a prognostic equation for the turbulent kinetic energy. ~~The transport~~Transport and local time tendency terms in ~~all~~ the ~~other~~ second-order momentum equations are neglected, and the vertical turbulent fluxes are derived diagnostically (Cerenzia I., 2017).

~~The resolution of IFS has improved from ~13 km in 2015 to ~10 km in 2016 and consists of 137 vertical levels. Its turbulent diffusion scheme represents the vertical exchange of heat, momentum, and moisture through sub-grid scale turbulence. In the surface layer, the turbulence fluxes are computed using a first-order K-diffusion closure based on the Monin-Obukhov (MO) similarity theory. Above the surface layer, a K-diffusion turbulence closure is used everywhere except for unstable boundary layers where an Eddy-Diffusivity Mass-Flux (EDMF) framework is applied to represent the non-local boundary layer eddy fluxes (Koehler et al. 2011).~~

~~The spatial resolution of the~~Both models ~~affects their ability to refer to the actual topography rather than a smoothed grid point. Therefore, the models' results were corrected~~estimate the PBL height by ~~the actual ground base heights for each measurement site (Table 1).~~

~~Concerning the time resolution,~~The bulk Richardson number method (described in Sect. 4.1). IFS produced hourly results while COSMO generated profiles every 15 min. ~~To compare PBL heights from both models, a~~A series of trials disclosed ~~that~~the COSMO profiles of the last 15 min within an hour, best represent the hourly values of the IFS~~.~~ model.

3. Instruments

43.1 Ceilometers

Vaisala ceilometers type CL31, commonly deployed worldwide, are the main is the primary research tool in this study. (Fig.1, Table 1). CL31 is a pulsed, elastic micro-lidar, employing an Indium Gallium Arsenide (InGaAs) laser diode transmitter of a 910 nm ±10 nm near-infrared wavelength of 910 nm ±10 nm at 25˚C withand a high pulse repetition rate of 10 kHz, every two seconds (Vaisala ceilometer CL31 user's guide: http://www.vaisala.com). The backscatter signals are collected by an avalanche photodiode (APD) receiver and designed as attenuated backscatter profiles at intervals of 2-120 s (determined by the user). In thisThis study, applied CL31 ceilometers were applied with the exception ofexcept for ceilometer CL51 stationed in the Weizmann Institute (Fig.1, Table 1). CL51 consists of a higher signal and signal-to-noise ratio, hence. Hence the backscatter profile measurement reaches up to ~15.4 km compared to 7~ 8 km of CL31. The ceilometers produce 10 m vertical resolution profiles every 15 or 16 sec. Half hourly backscatter profiles improved the signal to noise ratio. The second half-hour profile within each hour defined the hourly profiles.

One drawback is that calibration procedures were nonexistent in all sites, and in. In most cases, maintenance procedures (cleaning of the ceilometer window)), were not regularly carried out, with the exception ofexcept for the IMS BeitBet Dagan ceilometer. Nevertheless, the PBL height detection is based on a pronounced change of the attenuated backscatter profile. This change is attributed to variations in the aerosol content providing indications for both clouds and atmospheric layers. Therefore, the limitation of a single wavelength within the spectral range of water vapor absorption does not affect this type of detection. In order to derive the backscatter coefficient from ceilometer measurementsIn the case of the backscatter coefficients detection, signal calibrations, and water vapor corrections are necessary (Weigner et al., 2014, Wiegner and Gasteiger, 2015).

The ceilometers produce profiles every 15 or 16 sec (Table 1). In order to compare them to the models' hourly resultsHowever, the PBL height detection method employed here (Sect. 4.3), they were averaged to half-hour ones, whereas the second half-hour profile within each hour was chosen for the comparison process.

The nocturnal SBL heights in ground level ceilometer sites were detected mainly within the ceilometers' first range gates. At these heights, a constant perturbation existed due to the overlap of the emitted laser beam and the receiver's field of view. This fact limited our capability to determine the low SBL locates the height of the summer season and heightened our decision to focused on daytime CBL heights. Detailed information regarding the manufactural and technical properties of ceilometers involved in this research a pronounced change in the attenuated backscatter profile rather than a specific value. Therefore, calibration procedures are given in Uzannot mandatory (Weigner et al., 2014, Gierens et al. (., 2018).

43.2 Radiosonde

The IMS obtains systematic radiosonde atmospheric observations twice daily, at 23 UTC and 11 UTC, . The radiosonde launching site is adjacent to a ceilometer. Launching is performed in Beitthe Bet Dagan ceilometer (32.0 ° long, 34.8 ° lat, 33 m a.s.l), situated, 7.5 km east from the shoreline, 1112 km southeast to Tel Aviv, 45 km northwestnorth-west to Jerusalem (, see Fig.1 and Table 1). The radiosonde, type Vaisala RS41-SG, producesretrieves profiles of RHrelative humidity, temperature, pressure, wind speed, and wind direction as it ascends. Measurements are retrieved, every 10 seconds, corresponding to about(~ every 45 m, reaching), rising to ~25 km. Here, we refer to the first 2 km in about 8 minutes. The horizontal displacementfor the detection of the radiosonde depends on the intensity ofmidday summer PBL height. At this height, the ambient wind speed. The average wind speed along theat 11 UTC summer profiles is about ~5 m/s (Uzan et al., 2012). Therefore, the horizontal displacement of the radiosonde from its launch position is fairly low and is estimated at aboutis relatively low ~ 2.5 km. and neglected. Moreover, previous research showed the radiosonde position resolution is defined as 0.01°. As aforementioned, the midday PBL height in BeitBet Dagan for midday summer is estimatedis below 1 km (Dayan and Rodinzki, 1999, Uzan et al., 2016, Yuval et al., 2019). Hence, within an ascending height of 1 km, the change in the radiosonde's horizontal position is under), corresponding to horizontal displacement of ~ 0.01° which is an order of magnitude fromwell under the models' grid resolution. Thus, we assert the radiosonde profiles represent the Beit Dagan site and the displacement error of the ascending radiosonde can be neglected. of the IFS and the COSMO models.

~~5~~4. Methods

~~5~~4.1 The bulk Richardson number method

The COSMO and IFS schemes calculate the PBL height by the bulk Richardson number ~~method~~ (R_b) ~~as the most reliable technique for PBL height detection by NWP models (Zhang~~method ~~et al., 2014).~~

~~The bulk Richardson number formula~~ (Hanna R. Steven,1969, Zhang et al., 2014) ~~is~~ given in the ~~following manner~~formula below:

$$R_b = \frac{\frac{g}{\theta_v}(\theta_{vz} - \theta_{v0})(Z - Z_0)}{U^2 + V^2} \quad (1)$$

where g is the gravitational force, θ_{vz} is the virtual potential temperature at height Z, θ_{v0} is the virtual potential temperature at ground level (Z_0). U and V are the horizontal wind speed components at height Z~~.~~ (assuming U and V at surface height are insignificant, therefore negligible).

~~The IFS model defines the PBL height as the lowest height level at which the R_b (Eq. 1) reaches a critical threshold of 0.25 (ECMWF-IFS documentation – Cy43r3, Part IV: Physical Processes, July 2017). The PBL height is distinguished by scanning the bulk Richardson values from the surface level upwards. If the PBL height is found between two levels of the model, it is determined by linear interpolation.~~

~~Radiosonde's profiles were analyzed in the same manner by a R_b threshold of 0.25 to detect a specific height rather than a certain layer.~~

~~COSMO estimates the R_b based on the dynamic conditions of the first four levels (10, 34.2, 67.9, 112.3 m a.g.l.) signified by a threshold of 0.33 for stable conditions and 0.22 for unstable ones. If no level is found, then a missing value is assigned for the PBL height.~~

~~5~~The R_b threshold determines the PBL height. The IFS model has a single limit of 0.25 (Seidel et al., 2012). The COSMO model refers to 0.33 for stable atmospheric conditions (Wetzel, 1982), and 0.22 for unstable conditions by 0.22 (Vogelezang and Holtslag, 1996) in the first four levels of the model (10, 34.2, 67.9, 112.3 m a.g.l.). Linear interpolation determines the height if the detection is between two model levels. The height is assigned with a missing value if the thresholds were not reached. The models' PBL heights (given as m a.g.l.) are adjusted to the actual altitude of the ceilometer sites (Table 1). The radiosonde 11 UTC PBL heights were defined where the R_b profile values (derived every 10 sec correspond ding to ~ 45 m) altered from negative to positive. In all the dates studied, the first positive value was well above the thresholds for unstable conditions by both models (0.25 and 0.33). Therefore the PBL height was defined at the height point of the last negative value.

4.2 The parcel method

~~The PBL height is defined by the~~The parcel method ~~as~~defines the PBL height ~~aloft at which the value of~~where the virtual potential temperature aloft reaches ~~that of~~the value evaluated at the surface level (Holzworth 1964, Stull, 1988, Seidel et al., 2010). The ~~calculation~~description of the virtual potential temperature is as follows:

$$\theta_v = T_v \left(\frac{P_0}{P}\right)^{\frac{Rd}{Cp}} \qquad (2)$$

where P_0 is the ground level ~~atmospheric~~ pressure, P is the ~~at atmospheric~~ pressure at height Z, R_d is the dry air gas constant ~~of dry air~~, C_p is the heat capacity of dry air ~~in a constant pressure ($\frac{R_d}{C_p}$ = 0.286).~~. The virtual temperature (T_v) is obtained by:

$$T_v = \frac{T}{1-\frac{e}{P}(1-\varepsilon)} \qquad (3)$$

where T is the temperature at height Z, e is the actual vapor pressure, and ε is the ratio of ~~the gas constant~~molecular weight of ~~air and~~ water vapor and dry air (ε=0.622). ~~The actual vapor pressure (e) is derived by the relative humidity (RH) profile multiplied by the saturated vapor pressure (e_s). The saturated vapor pressure was derived by the temperature profile.~~

~~In this method, the value of the~~ The virtual potential temperature ~~at surface height is crucial. The first levels of IFS~~ profiles were computed based on the available meteorological parameters from the models and ~~COSMO are 10 m a.g.l~~radiosonde: mixing ratio, pressure, and ~~20 m a.g.l, respectively. Thus, evaluations of the ambient~~ temperature ~~and the~~ profiles from the IFS model. Relative humidity, pressure, and temperature profiles from the COSMO model and the radiosonde. The virtual potential temperature profiles of the models at ground level were obtained by the temperature and dew point temperature ~~(or RH) for~~at 2 m a.g.l ~~are generated by the models based on the~~. These parameters were derived from the models by the similarity theory. Finally, the PBL heights (given in m. a.s.l) were adjusted to the actual altitude of the ceilometer sites (Table 1).

54.3 The wavelet covariance transform method

The wavelet covariance transform (WCT) method is operated along the length of the backscatter profile ((Baars et al., 2008, Brooks Ian, 2003). This method) is basedimplemented on backscatter profiles by the Haar step function (Baars et al., 2012)formula given in Eq. (4) and Eq. (5) as follows4:

$$W_{f(a,b)} = \frac{1}{a}\int_{Zb}^{Zt} f(z)h(\frac{z-b}{a})dz \quad (4)$$

where $W_{f(a,b)}$ is the local maximum of the backscatter profile ($f(z)$) determined within the range of step (a).) by the Haar step function (h). The length of the step is the number of height levels (n) multiplied by the profile height resolution (Δz) from ground level (Zb) and up (Zt). In this study, Zb was defined as the height above the perturbation of the overlap function (~ 100 m), and Zt as the height with the most significant signal variance or, the first appearance of negative values. Both thresholds indicate a low signal-to-noise ratio. Zb is the lowest height among the two options. These thresholds apply under clear sky conditions. When clouds exist in the summer, they are mainly shallow cumulus clouds (Sect. 1.1). The PBL height is the height within the cloud, above the cloud base height (Wang et al. 2012, Stull 1988).

The Haar step function, given in Eq. (5),.(4) is equivalent to a derivative at height z, representing the value difference of each step (a) above and beneath a point of interest (b). HereIn this study, b is the measurement heights of the ceilometer along thebackscatter profile (every 10 m starting from 10 m a.g.l) and . The value of the step a was defined as 20 m (10 m above and beneath point b).(a) varied for each ceilometer, depending on the site location.

$$h(\frac{z-b}{a}) = \begin{Bmatrix} +1, & b-\frac{a}{2} \leq z \leq b, \\ -1, & b \leq z \leq b+\frac{a}{2} \\ 0, & elsewhere \end{Bmatrix} \quad (5)$$

To evaluate the ceilometers' PBL heights (Eq. 4), the backscatter profiles are analyzed by the WCT method between two boundaries. The lower boundary (Zb) is the height above the perturbation of the overlap function (~ 100 m, see Sect. 4.1). The upper limit (Zt) is either the height point with the largest variance within a step or the first height point with negative values indicating a low signal-to-noise ratio. The lowest height among the two aforementioned options will define the upper limit.

When clouds exist (mainly shallow cumulus clouds), the algorithm defines the PBL height as the highest measurement point of the cloud above the cloud base height. This height indicates the entrainment zone rather than the actual cloud top.

6In arid and dusty areas such as Nevatim and Hazerim, specifically on clear days, the WCT method failed to distinguish the PBL height (Van der Kamp and McKendry, 2010, Gierens et al., 2018). The analysis excluded these cases. The last stage consisted of manual inspection of the WCT results.

5. Results

In the Israeli summer season, stable PBL conditions are generated from sunset to an hour after sunrise (Stull, 1988). At this period the models' R_b profiles do not accede the relevant thresholds, and a missing value is assigned (Sect. 4.1). Additionally, the difficulty to estimate the surface boundary layer by ceilometers (Gierens et al., 2018) was associated with a constant perturbation within the first range gates due to the overlap of the emitted laser beam and the receiver's field of view (Weigner et al., 2014). **6.1 Comparison to in-situ radiosonde profiles**

In order to evaluate the daytime PBL heights produced by the models and the ceilometers, Hence, the analysis focused on the midday summer PBL heights.

5.1 Spatial and temporal analysis

The analysis was performed based on six ceilometers with available data of at least 50 days within the study period: Bet Dagan, Tel Aviv, Ramat David, Weizmann, Jerusalem, and Nevatim. In Bet Dagan, the results were compared to the radiosonde's evaluations. Consequently, the investigation was held in Beit Dagan launch site at the time of the midday launch (11 UTC). For this comparison, the ceilometer's 15 s profiles were averaged as half-hour profiles between 10:30-11:00 UTC. COSMO's results referred to the profiles of 10:45

~~UTC, and IFS estimations were given~~ radiosonde, thereupon, the analysis fixated at 11 UTC~~.~~ ~~The analysis was carried out~~ launching time. In the remaining five sites, the models compared to the ceilometers. Statistical analysis for ~~33 summer days, 13 days from August 2015, and 20 days from Aug 2016. The PBL heights were produced by the same methods: the parcel method (denoted by subscript P) and the bulk Richardson method (denoted by subscript R). These methods require meteorological parameters such as temperature and pressure profiles generated by the models and the radiosonde. Ceilometers, on the other hand, produce only backscatter signals. Therefore, they were analyzed by the WCT method. The results were statistically analyzed by~~ each site presents the mean error (ME), root mean square ~~error~~ (RMSE), ~~and~~ correlation (R~~) presented in Fig. 2 and Table 3.~~), Mean and standard deviation (STD) given in tables and plots.

Good agreement was found between the ceilometer and the radiosonde (RS) in Bet Dagan (Fig. 2 and Table 3, ME = ~~12 m~~4, RMSE = ~~97 m, and R = 0.93), although they produced~~143, R = 0.83). The IFS by the parcel method (IFS-pm) appears to overestimate the PBL ~~heights by different methods.~~height (ME = 346, RMSE = 494, R = 0.14), as well as by the Richardson method (IFS-ri, ME = 366, RMSE = 579, R = -0.13). Among the models and methods, ~~COSMO$_R$ retrieved~~the COSMO model by the parcel method derived the best results ~~of ME = -3 m, RMSE = 152 m and R = 0.83). IFS predominantly overestimated the PBL heights. The poorest results were generated by IFS$_R$ (ME = 274 m, RMSE = 432 m, R = 0.18~~(COSMO-pm, ME = -52, RMSE = 146, R = 0.84).

In the shoreline site of Tel Aviv (Fig. 3, Table 4), COSMO-pm displayed good agreement with the ceilometer measurements (ME = 17, RMSE = 183, R = 0.74), similar to COSMO-ri (ME = 18, RMSE = 187, R = 0.7). IFS-ri produced the highest overestimations (ME = 436, RMSE = 616, R = -0.03).

In Ramat David, stationed in the northern inner plain of Israel, the parcel method derived better results than the Richardson method in both models (Fig. 4, Table 5). Among the models, COSMO displayed better results (ME = 40, RMSE = 245, R = 0.55). IFS-ri generated the poorest correlation (ME = 446, RMSE = 745, R = -0.08).

In Weizmann (Fig. 5 Table 6), 11 km southeast to Bet Dagan, IFS-ri produced the poor results (ME = 430, RMSE = 604, R = -0.01), conversley to the good results by the parcel method (ME = 67, RMSE = 162, R = 0.85). The COSMO model derived similar results by both methods (COSMO-pm: ME = -106, RMSE = 207, R = 0.76, COSMO-ri: ME = 21, RMSE = 192, R = 0.72).

In the mountainous site of Jerusalem, the bulk Richardson method produced better results than the parcel method in both models (Fig. 6, Table 7). COSMO-pm derived good results (ME = -44, RMSE = 239, R = 0.70) and IFS-ri the poorest (ME = 366, RMSE = 498, R = 0.18).

In the elevated and arid site of Nevatim, overall correlations were weak (0.1-0.3) and high RMSE (369 - 488).

Main conclusions derived from Fig.2-7 are summarized below:

- Low correlation in Nevatim (0.1-0.3) demonstrates the difficulty of the models to assess the PBL height over complex terrain. Evaluation of PBL heights in complex terrain was studies by Ketterer et al. (2014) in the Swiss Alps by a ceilometer, wind profiler, and in-situ continuous aerosol measurements. The ceilometers analyzed by the gradient and STRAT-2D algorithms and the wind profiler by the range-corrected SNR method. The results compared to the COSMO-2 regional model. The results showed good agreement found between the heights derived by the ceilometer and wind profiler during the daytime cloud-free conditions (R^2=0.81). However, in most cases, the model underestimated the PBL height. The researchers presumed the grid resolution, parametrization schemes, and the surface type did not match the real topography. The comparison between a single measurement point and a grid point is not straight forward.
- The parcel method achieved better results in Ramat David, Tel Aviv, Bet Dagan, and Weizmann. In the elevated site of Jerusalem, the correlation of COSMO-ri was the highest (R=0.7).
- The COSMO model produced better results in the shoreline and plain regions (Ramat David, Tel Aviv, Bet Dagan) except for Weizmann (60 m a.s.l, 11.5 km from the coastline), where IFS-pm obtained the highest correlation (R=0.85).

IFS model based on the bulk Richardson method overestimated the PBL heights (~ 420 m) in the plain sites of Bet Dagan, Tel Aviv, Weizmann, and Ramat David. The bulk Richardson evaluation (See Sect. 4.1) includes the horizontal wind speed profiles that are less accurate and may contribute to the discrepancies. Collaud et al. An example of an analysis on a typical day is given in Fig. 3 for August 15, 2015. On this day, the PBL height at 11 UTC was estimated at 680 m a.s.l by the radiosonde. $COSMO_P$ accurately estimated the same height while $COSMO_R$ detected the height to be 100 m lower (580 m a.s.l). The ceilometer overestimated by 100 m (795 m a.s.l). IFS results were twice the value produced by the radiosonde (IFS_R=1,300 m a.s.l, IFS_P= 1,474 m a.s.l).

Among the 33 days tested, the largest gap was found between IFS$_R$ and RS$_R$ on August 17, 2016 (Fig. 2). The imprecision could be due to the fact that the Richardson method is based solely on dry thermodynamics for local turbulence (Von Engeln and Teixeira, 2013), while on August 17, 2016, the 11 UTC PBL height was determined through a multi-layer cloud (not shown).

6.2 Spatial analysis by ceilometers

After the good results generated by the WCT method imposed on the ceilometer's profiles, ceilometers were applied as PBL height detectors in sited where no other atmospheric measurements operated. The same analysis process was carried out but for five ceilometer sites (Ramat David, Tel Aviv, Beit Dagan, Weizmann, and Jerusalem), representing diverse terrain on 13 specific days available from all instruments and models. This time the models' results were compared to the ceilometers' measurements in each site. Both models defined the Tel Aviv site by a grid point mostly over the Mediterranean Sea. Therefore, we shifted the Tel Aviv coordinates to an adjacent grid point that was mostly land, representing Tel Aviv by the same height and distance from the shoreline.

A comparison to radiosonde's results was available only in Beit Dagan. Figure 4b reveals a good agreement between the radiosonde and the ceilometer's evaluations in Beit Dagan, although the different methods imposed on each instrument. A significant case on August 10, 2015, where an atmospheric layer above the PBL height denoted by the radiosonde and the ceilometer (not shown) led to the models' discrepancies.

By and large, COSMO$_R$ achieved the best statistical results (Tables 4-5) regarding flat and complex terrain, of RMSE from 175 m in Weizmann (60 m a.s.l, and 11.5 km east from the shoreline) up to 251 m in Jerusalem (830 m a.s.l and 53 km east from the shoreline), and ME between 19 m in Tel Aviv (5 m a.s.l, and 50 m from shoreline), and -26 m in Ramat David (50 m a.s.l, and 24 km east from the shoreline). IFS$_P$ produced high RMSE results starting at 180 m in Ramat David rising up to 569 m in Beit Dagan, and ME up to 497 m in Beit Dagan. These results emphasize the advantage of high-resolution regional models such as COSMO (~2.5 km resolution) over the IFS global model (resolution of ~13 km in 2015 and ~10 km in 2016) over a diverse area.

- **6.3**(2014) referred to the limitations of the bulk Richardson method of the COSMO-2 regional model (2.2 km resolution), which overestimated the convective boundary layer by 500–1000m. They explained the Richardson method is sensitive to the surface temperature, and errors and uncertainties in the model's temperature and relative humidity profiles could explain the significant bias. Also, the occurrence of clouds, which may be missing in the model, can lead to lower PBL heights.

5.2 COSMO PBL height correction

~~Finally, the spatial daytime summer PBL heights were investigated. Following the conclusions of previous stages, $COSMO_R$ was chosen as the model and method that achieved the best results. Average hourly values were derived between 09-14 UTC (corresponding to 11 -16 LST) and compared to the results from eight ceilometer sites (Fig. 1, Table 1). The comparison was accomplished by all dates available for each ceilometer site on August 2015: Jerusalem - 21 days, Nevatim - 13 days, Hazerim - 20 days, Ramat David - 26 days, Weizmann - 25 days, Beit Dagan - 13 days, Hadera - 16 days, Tel Aviv - 25 days.~~

~~In order to validate and correct $COSMO_R$ results by the ceilometers' measurements, a correction tool based on a regression function was implemented for each hour (09-14 UTC), for all ceilometers' sites simultaneously by the following formula:~~

~~ME_{st}~~A correction formula for the models' PBL height employing ceilometers is given below:

$$dH_{st} = \cancel{\alpha G + \beta D +} \alpha h_{st} + \beta d_{st} + \gamma \qquad (6)$$

~~where ME_{st} is the dependent variable representing the PBL height mean error for each ceilometer station (st) compared to the results obtained by the $COSMO_R$. The independent predictor variables are the ground altitude of the ceilometer's site (G) and its distance from the shoreline (D). The correction factors α, β, and γ are implemented on the $COSMO_R$ PBL height results.~~

~~$COSMO_R$ mean PBL heights cross-section from Tel Aviv (34.8° lat) to Jerusalem (35.2° lat) is presented in Fig. 5. Before the correction (Fig. 5a), $COSMO_R$ approximated Tel Aviv PBL heights descend gradually from 750 at 09 UTC (11 LST) to 600 m a.s.l at 14 UTC (16 LST). Apparently, the correction tool reduced the height difference to ~700 m a.s.l with the exception of ~750 m a.s.l at 09 UTC (Fig. 5b). These results correspond to Uzan et al, (2012) showing~~

~~Tel Aviv site is practically on the shoreline, therefore as the sea breeze enters Tel Aviv (~ 08 UTC), it surmounts the convective thermals preventing from the mixed layer to inflate.~~

~~In Jerusalem, the summer PBL height inflates according to the insolation intensity, as the main source of the buoyancy force. Therefore, the maximum daytime PBL heights are measured at midday. In the afternoon, when the sea breeze reaches eastern Israel, the height decreases. $COSMO_R$ results before and after the correction showed the highest value at 11 UTC (13 LST), corresponding to maximum insolation at midday. The lowest value was corrected from 09 UTC (11 LST) to 14 UTC (16 LST) as insolation decreases and the cool and humid air of sea breeze front demolishes the thermals and the PBL height subsides.~~

~~Between the shoreline of Tel Aviv and the eastern mountains of Jerusalem, the overall range of PBL height values was reduced. For example, in 35° lat (between Weizmann - 60 m a.s.l and Jerusalem - 830 m a.s.l), the PBL heights of 09-14 UTC varied from 750 to 1500 m a.s.l. After the correction, the height values ranged from 1000 to 1400 m a.s.l, generating higher PBL heights for the daytime hours. Fig. 6 demonstrated the correction tool at 14 UTC disclosing a correction of ~ 300 m (Fig. 6b) over the complex terrain of Jerusalem (830 m a.s.l) and Nevatim (400 m a.s.l).~~

~~7~~Where dH_{st} is the PBL height difference between the ceilometer and the model, the altitude (h_{st}), and distance from the shoreline (d_{st}) for each measurement site (*st*). The formula runs simultaneously for all ceilometer sites to derive the dependent variables α, β, and the constant γ. The formula is suitable for both models

A case-study demonstrates the correction formula on August 14, 2015, from the COSMO model based on the parcel method (COSMO-pm). COSMO-pm is the model and method that derived good results in Sect. 5.1. The formula runs for each hour between 9-14 UTC for the daytime PBL height (See Sect. 5). Results are portrayed for each hour by a 2-D plot of the height correction within the area of ceilometers' deployment. Along with an east-west cross-section plot, corresponding to the location of the ceilometers. Cross-validation tests for Bet Dagan and Jerusalem show the effectivity of the correction formula. Main findings for each hour are as follows:

9 UTC (Fig. 8): Along the coast, the correction tool lowers the PBL height by 70 m to 670 m and increases by 90 m in the inner strip of Israel to ~ 890 m a.s.l. Cross-validation for Bet

Dagan (CV-BD) shows good results, whereas, in Jerusalem (CV-JRM), the correction tool reduced the height by 600 m.

10 UTC (Fig. 9): The correction tool distinguishes between the coastal sites of Tel Aviv and Hadera, and the inland locations of Bet Dagan and Weizmann, only ~ 10 km apart from Tel Aviv. While the correction tool increased the height of the coastal stations, a slight height decreased was performed in the inner sites. In the arid southern Hazerim, the correction tool lowered the PBL height by 400 m. In the desert south of Nevatim, the correction tool decreased the PBL height by 200 m. Cross-validation of Jerusalem (CV-JRM) underestimates the PBL height in Jerusalem by 400 m.

11 UTC (Fig. 10): A distinction between the shoreline and the inner sites is more evident, as the PBL height of Tel Aviv and Hadera is increased by ~100 m to ~700 m a.s.l, whereas, Bet Dagan and Weizmann remained ~ 800 m a.s.l. This finding corresponds to Uzan et al. (2016) analysis of the mean diurnal-cycle of the PBL height from July to August 2014, based on ceilometer measurements. A pronounced correction is visible in the elevated southern site of Hazerim by 550 m down to 1120 m a.s.l. This gap is not unexpected since NWP models have difficulty assessing the meteorological conditions over complex terrain. Here, Jerusalem cross-validation (CV-JRM) underestimates the PBL height by a comparatively lower range of 200 m.

12 UTC (Fig. 11): The correction tool increased the PBL height in the coast and inland stations, but in fact, the height is lower than an hour before. The PBL height in Hazerim is decreased by 300 m. Jerusalem cross-validation (CV-JRM) underestimates the PBL height in Jerusalem by 600 m.

13 UTC (Fig. 12): The correction tool increased the PBL heights. A substantial increase of 380 m in Jerusalem generates a height of ~1750 m a.s.l. Jerusalem cross-validation (CV-JRM) underestimates the PBL height by 550 m.

14 UTC (Fig. 13): Similar to an hour before, the correction increases the PBL height in all sites, but in fact, the PBL heights are lower than an hour earlier, except a mild increase in the coastal locations of Tel Aviv and Hadera. Jerusalem cross-validation (CV-JRM) underestimates the PBL height by ~300 m.

6. Summary and Conclusions

The primary purpose of this study was to improve the performance of air pollution dispersion models by providing applicable data of PBL heights from NWP models employing ceilometers. A correction tool using ceilometer measurements was established to validate the models' PBL height assessments. The study focused on the summer PBL heights (July-September 2015, June-August 2016) during the day hours (9-14 UTC). At this period, the highest air pollution events occur in Israel from tall stacks (Dayan et al., 1988, Uzan et al., 2012).

The study contained eight ceilometers, a radiosonde, two models - IFS and COSMO, and three PBL height analysis methods. The bulk Richardson method, the parcel method for the models and radiosonde, and the WCT method for the ceilometers. In Bet Dagan radiosonde launching site, results revealed good agreement between the ceilometer's PBL heights and the radiosonde (N = 91 days, ME = 4 m, RMSE=143 m, R=0.83). In Ramat David, Tel Aviv, Weizmann, Jerusalem, and Nevatim, the models were compared to the ceilometers. The COSMO model performed better in the plain areas of Tel Aviv (10 m a.s.l), Bet Dagan (33 m a.s.l), and Ramat David (50 m a.s.l) and the mountainous Jerusalem (830 m a.s.l). The IFS model showed good agreement with the ceilometer in Weizmann (60 m a.s.l, N=55 days, ME = 67 m, RMSE = 162 m, R=0.85). In the arid southern site of Nevatim (400 m a.s.l), overall correlations were poor. The IFS-pm produced better in Bet Dagan, Ramat David, Tel Aviv, and Weizmann (four out of five sites except for Nevatim). The COSMO-pm produced better results in Bet Dagan and Ramat David, while in Tel Aviv the results generated by both methods were similar (N = 123 days, COSMO-pm: ME = 17 m, RMSE = 183 m, R=0.74, COSMO-ri: ME = 18 m, RMSE = 180 m, R=0.80).

The PBL height correction tool for the NWP models is based on the altitude and the distance from the shoreline of the ceilometers' measurement sites. A case-study demonstrated the tool's feasibility on August 14, 2015. Moving from 9 to 14 UTC, the correction decreased the PBL height in flat terrain (Tel Aviv, Hadera, Bet Dagan, and Ramat David). This finding corresponds with Uzan et al., 2016, analyzing the diurnal PBL height of Bet Dagan and Tel Aviv in the summer of 2014. Similar results produced in Hadera describe the summer PBL height between 1997-1999 and 2002-2005 based on measurements from a wind profiler (Uzan et al., 2012). Koch and Dayan (1992) revealed air pollution episodes of sulfur dioxide increased in shallow PBL heights in the coastal plain of Israel. Uzan et al. (2012) showed an average decrease of ~ 100 m in the coastal PBL height resulted in an average increase of ~200 air pollution episodes of sulfur dioxide.

The tool increased the PBL height in the elevated site of Jerusalem (830 m a.s.l) by ~380 m. In the arid south in Hazerim (200 m a.s.l), the tool lowered the PBL height by ~ 550 m. The significant height corrections in the elevated sites are attributed to the models' difficulty to imitate local meteorological processes in complex terrain (e.g., Alpert et al., 1984). Dayan et al. (1988) presumed the diurnal cycle and the prevailing synoptic systems govern the temporal behavior of the Israeli summer PBL height. The strength of the sea breeze determines significant variations in the inner PBL heights.

Cross-validation for Bet Dagan produced excellent results. Bet Dagan is located in flat terrain 11 km north to the Weizmann site and 12 km southeast to Tel Aviv site. Without the single measurement site in Jerusalem (830 m a.s.l), the correction tool failed to generate Jerusalem's PBL height and produced lower values up to a 600 m difference. This finding shows the process of cross-validation can assist in defining the required ceilometers' deployment in the future.

In summary~~. **Summary and Conclusions**~~

~~Earlier studies have successfully employed ceilometers for PBL height detection, typically under dry conditions. However, these studies employed weather models primarily as a validation tool rather than investigating the models' predictive capabilities. Here, we tested the ability of ceilometers to serve as a correction tool for PBL height estimations derived from two operational models: the IFS global model, and the mesoscale COSMO regional model. The study focused on the daytime summer PBL heights.~~

~~Firstly, we compared the models' and the ceilometer's evaluations to actual measurements from an adjacent radiosonde in the Beit Dagan launch site. Results for 11 UTC on 33 August days revealed the promising ability of the WCT method to detect the PBL heights generated by the radiosonde by the bulk Richardson method and by the parcel method (RMSE= 97 m).~~

~~In the next stage, the investigation expanded spatially to four other diverse measuring sites, from the shoreline of Tel Aviv (5 m a.s.l) to the mountainous Jerusalem (830 m a.s.l). The same methods were applied for 13 summer days, except this time, the models' values were compared to the ceilometers' measurements in each site. The results disclosed the COSMO model based on the bulk Richardson method ($COSMO_R$) achieved the best results for both flat (Tel Aviv: RMSE=203 m, ME=19 m) and complex terrain (Jerusalem: RMSE = 251m, ME = -6 m).~~

~~Finally, the temporal and spatial evolution of the summer daytime (11-16 LST) PBL heights were examined. The heights were derived by COSMO$_R$ and compared to ceilometers measurements distributed in eight sites across Israel, providing a heterogeneous research area in comparatively short distances. A correction tool was established based on a regression function comprised of the topography of the ceilometer's site (G) and its distance from the shoreline (D) serving as the independent predictor variables. The results revealed corrections up to ~ 300 m difference which improved the description of the diurnal PBL heights.~~

~~Despite the limited database~~, our results offer a preview of the great potential of ceilometers as a ~~validation and a~~ correction tool ~~to discern~~for PBL heights derived from ~~weather~~NWP models. This tool demonstrates the benefit of deploying ceilometers, specifically in complex terrain. Future research should~~, therefore,~~ include a larger dataset to ~~evaluate whether these results are retained in the long term and to define~~create a systematic ~~validation process.~~ correction process and produce sufficient input data for mandatory air pollution dispersion assessments.

Data availability

Weather reports - Israeli Meteorological Service weather reports (in Hebrew): http://www.ims.gov.il/IMS/CLIMATE/ClimateSummary.

Radiosonde profiles ~~–~~- Israeli Meteorological Service provided by request.

Ceilometer profiles - the data is owned by ~~governmental offices~~several institutions and provided by request.

Author contribution

Leenes Uzan carried out the research and prepared the manuscript under the careful guidance of Pinhas Alpert and Smadar Egert alongside a fruitful collaboration with Yoav Levi, Pavel Khain, and Elyakom Vladislavsky. The authors declare that they have no conflict of interest.

Acknowledgments

We wish to thank the Israeli Meteorological Service, the Israeli Air Force, the Association of Towns for Environmental Protection (Sharon-Carmel), and Rafat Qubaj from the Department of Earth and Planetary Sciences at the Weizmann Institute of Science~~,~~ for their ceilometer data. We ~~are indebted to Hadas Marcus~~thank Noam Halfon from the IMS for ~~her editing assistance.~~ the topography map

References

Alpert, P. and J. Neumann, On the enhanced smoothing over topography in some mesometeorological models, Bound. Lay. Met., 30, 293-312, 1984.

Alpert P., and Ziv B.: The Sharav cyclone observations and some theoretical considerations, J. Geophys. Res., 94, 18495–18514, 1989.

Alpert, P., Abramsky, R., Neeman, B.U.: The prevailing summer synoptic system in Israel-Subtropical high, not Persian Trough, Isr, J. Earth Sci, 39, 93-102, 1990.

Alpert P., Herman J., Kaufman Y. J., and Carmona I.: Response of the climatic temperature to dust forcing, inferred from TOMS Aerosol Index and the NASA assimilation model. J. Atmos. Res., 53, 3-14, 2000.

Alpert, P., Krichak, S. O., Tsidulko, M., Shafir, H., and Joseph, J. H.: A Dust Prediction System with TOMS Initialization, Mon. Weather Rev., 130, 2335–2345, 2002.

Alpert P., and Rabinovich-Hadar M.: Pre- and post-frontal lines - A meso gamma scale analysis over south Israel, J. Atmos. Sci., 60, 2994-3008, 2003.

Alpert, P., Osetinsky, I., Ziv, B., Shafir H.: Semi-objective classification for daily synoptic systems: Application to the eastern Mediterranean climate change. Int. J. of Climatol., 24, 1001-1011, 2004.

Anenberg, S., Miller, J., Henze, D., & Minjares, R.: A global snapshot of the air pollution-related health impacts of transportation sector emissions in 2010 and 2015. ICCT report, Washington, D.C., 2019.

Ansmann et al.: Long-range transport of Saharan dust to northern Europe: The 11–16 October 2001 outbreak observed with EARLINET, J. of Geo. Phys. Research, 108, D24, 4783, doi:10.1029/2003JD003757, 2003.

Ansmann, A., Petzold, A., Kandler, K., Tegen, I.N.A., Wendisch, M., Mueller, D., Weinzierl, B., Mueller, T. and Heintzenberg, J.: Saharan Mineral Dust Experiments SAMUM–1 and SAMUM–2: what have we learned? Tellus B, 63(4), 403-429, 2011.

Baars, H., Ansmann, A., Engelmann, R., and Althausen, D.: Continuous monitoring of the boundary layer top with lidar, Atmos. Chem. Phys., 8, 7281–7296, https:// doi:10.5194/acp-8-7281-2008, 2008.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. Mon. Wea. Rev., 139, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1, ~~2001~~2011.

~~Bechtold~~Bechthold, P.: Convection parametrization. ECMWF Seminar proceedings on ~~“~~"The parametrization of subgrid physical processes~~”~~", 63-85, 2008.

Briggs, G.A.: Plume rise predictions. In: Haugen, D.A. (Ed.), Lectures on Air Pollution and Environmental Impact Analysis, American Meteorological Society, pp. 59–111, 1975.

Brooks, I.: Finding Boundary Layer Top: Application of a wavelet covariance transform to lidar backscatter profiles, J. Atmos. Ocean. Tech., 20, 1092–1105, 2003.

Cerenzia I., Challenges and critical aspects in stable boundary layer representation in numerical weather prediction modeling: diagnostic analyses and proposals for improvement, Ph.D. thesis, University of Bologna, 2017.

Collaud Coen, M., Praz, C., Haefele, A., Ruffieux, D., Kaufmann, P., and Calpini, B.: Determination and climatology of the planetary boundary layer height above the Swiss plateau by in situ and remote sensing measurements as well as by the COSMO-2 model, Atmos. Chem. Phys., 14, 13205-13221, https://doi.org/10.5194/acp-14-13205-2014, 2014.

Dayan, U., Shenhav, R., Graber, M.: The Spatial and temporal behavior of the mixed layer in Israel, J Appl Meteorol, 27, 1382- 1394, 1988.

Dayan, U., and Koch J.: A synoptic analysis of the meteorological conditions affecting dispersion of pollutants emitted from tall stacks in the coastal plain of Israel. Atmos Environ., 26A, No.14, 2537-2543,1992.

Dayan, U., Rodnizki, J.: The temporal behavior of the atmospheric boundary layer in Israel. J Appl Meteorol, 38, 830-836, 1999.

Dayan, U., Lifshitz-Goldreich B., and Pick, K.: Spatial and structural variation of the atmospheric boundary layer during summer in Israel-Profiler and rawinsonde measurements. J. Appl. Meteor., 41, 447-457,2002.

Dockery, D. W., Pope, C. A. III, Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Benjamin G. Ferris, Jr., Speizer, F. E.: An association between air pollution and mortality in six U.S. cities. New England Journal of Medicine, 329, 1753– 1759, 1993.

Doms, G., J. Förstner, E. Heise, H.-J. Herzog, D. Mironov, M. Raschendorfer, T. Reinhardt, B. Ritter, R. Schrodin, J.-P. Schulz, and G. Vogel: A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization. Deutscher Wetterdienst, Ofenbach, 154 pp, 2011.

Eresmaa, N., Karppinen, A., Joffre, S. M., Räsänen, J., and Talvitie, H.: Mixing height determination by ceilometer, Atmos. Chem. Phys., 6, 1485-1493, https://doi.org/10.5194/acp-6-1485-2006, 2006.

Feliks, Y.: A numerical model for estimation of the diurnal fluctuation of the inversion height due to a sea breeze, Bound. Layer Meteor., 62, 151-161. 1993.

Feliks, Y: An analytical model of the diurnal oscillation of the inversion base due to sea breeze, J. Atmos. Sci., 51, 991-998,1994.

Feliks, Y: Nonlinear dynamics and chaos in the sea and land breeze, J. Atmos. Sci., 61, 2169-2187, 2004.

Garratt, J. R.: The Atmospheric Boundary Layer, Cambridge Univ. Press, Cambridge, UK, 335 pp., 1992.

Gierens, R.T., Henriksson, S., Josipovic, M., Vakkari, V., Van Zyl, P.G., Beukes J.P., Wood, C.R., ~~O’Connor~~O'Connor, E.J.: Observing continental boundary layer structure and evolution over the South African savannah using a ceilometer, Theor. Appl. Climatoi., 136, 333-346, https://doi.org/10.1007/s00704-018-2484-7, 2018.

Goldreich Yair: The climate of Israel-Observations, research and applications, Springer US, Chap-5, 2003.

Haeffelin, M. and Angelini, F.: Evaluation of mixing height retrievals from automatic profiling lidars and ceilometers in view of future integrated networks in Europe, Bound. Lay. Meteorol., 143, 49–75, 2012.

Hanna, S. R.: The thickness of the planetary boundary layer, Atmos. Environ., 3, 519–536, 1969.

Hashmonay, R., Cohen, A., and Dayan, U.: Lidar observations of the atmospheric boundary layer in Jerusalem, J. Appl. Meteorol, 30, 1228-1236, 1991.

Héroux, E., Brunekreef, B., Anderson, H. R., Atkinson, R., Cohen, A., Forastiere, F., Hurley F., Katsouyanni K., Krewski D., Krzyzanowski M., Ku¨nzli N., Mills I., Querol X., Ostro B., Walton, H.: Quantifying the health impacts of ambient air pollutants: Recommendations of a WHO/Europe project. International Journal of Public Health, 60, 619–627, 2015.

Holzworth, C. G.: Estimates of mean maximum mixing depths in the contiguous United States, Mon. Weather Rev., 92, 235–242, 1964.

Ketterer, C., Zieger, P., Bukowiecki, N.: Bound. Lay. Meteo., 151, 317-334, https://doi.org/10.1007/s10546-013-9897-8, 2014.

Koehler, M., Ahlgrimm, M. and Beljaars, A.: Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model, Q. J. R. Meteorol. Soc., 137, 43-57, 2011.

Koch J., Dayan U.: A synoptic analysis of the meteorological conditions affecting dispersion of pollutants emitted from tall stacks in the coastal plain of Israel, Atmos Environ, 26A(14), 2537-2543, 1992.

Kotthaus, S. and Grimmond, C.S.B.: Atmospheric boundary-layer characteristics from ceilometer measurements. Part 1: a new method to track mixed layer height and classify clouds, Q J R Meteorol. Soc.,144 (714), 1525–1538, https://doi.org/10.1002/qj.3299, 2018.

Levi Y., Shilo E., and Setter I.: Climatology of a summer coastal boundary layer with 1290-MHz wind profiler radar and a WRF simulation, J. Appl. Meteorol. Climatol., 50, 1815-1826, https://doi.org/10.1175/2011JAMC2598.1, 2011.

Lieman, R. and Alpert, P.: Investigation of the planetary boundary layer height variations over complex terrain, Bound. Lay. Meteorol., 62, 129-142, 1993.

Ludwing F. L.: A review of coastal zone meteorological processes important to the modeling of air pollution, Air pollution modeling and its application IV, edited by C.De Wispelaere, NATO, Challenges of modern society, VOL 7, 1983.

Luo, T., Yuan R., Wang Z.: On factors controlling marine boundary layer aerosol optical depth, J. Geophys. Res. Atmos., 119, 3321–3334, doi:10.1002/ 2013JD020936, 2014.

Lyons, W.A.: Turbulent diffusion and pollutant transport in shoreline environments. Lectures on air pollution and environmental impact analysis, D.A. Haugen (ED.), American Meteorological Society. 136-208, 1975.

Mamouri, R.E., Ansmann, A., Nisantzi, A., Solomos, S., Kallos, G., and Hadjimitsis, D.G.: Extreme dust storm over the Eastern Mediterranean in September 2015: satellite, lidar, and surface observations in the Cyprus region, Atmos. Chem. Phys., 16(21), 13711-13724, 2016.

Manninen A. J., Marke T., Tuononen M. J., O'Connor E. J.: Atmospheric boundary layer classification with Doppler lidar, Journal of Geophysical Research: Atmospheres, 123, 8172–8189, https://doi.org/10.1029/2017JD028169,2018.

Mellor M. J., and Yamada T.: Development of a turbulence closure model for geophysical fluid problems, Reviews of geophysics and space physics, 20 (4), 851-875, https://doi.org/10.1029/RG020i004p00851, 1982.

Neumann J.: Diurnal variations of the subsidence inversion and associated radio wave propagation phenomena over the coastal area of Israel. Isr. Met. Serv, 1952.

Neumann J.: On the rotation rate of the direction of sea and land breezes. J. Atmos. Sci., 34, 1913-1917, 1977.

Ogawa, Y., T. Ohara, S. Wakamatsu, P.G. Diosey, and I. Uno: Observations of lake breeze penetration and subsequent development of the thermal internal boundary layer for the Nontecore II shoreline diffusion experiment. Boundary layer meteorology, 35, 207-230,1986.

Papayannis, A., Amiridis, V., Mona, L., Tsaknakis, G., Balis, D., Bösenberg, J., Chaikovski, A., De Tomasi, F., Grigorov, I., Mattis, I. and Mitev, V.: Systematic lidar observations of Saharan dust over Europe in the frame of EARLINET (2000–2002). Journal of Geophysical Research: Atmospheres, 113(D10), 2008.

Saaroni, H., and Ziv B.: Summer Rainfall in a Mediterranean Climate – The Case of Israel: Climatological – Dynamical Analysis, Int. J. Climatol., 20, 191 – 209, 2000.

Scarino, A. J., Obland, M. D., Fast, J. D., Burton, S. P., Ferrare, R. A., Hostetler, C. A., Berg, L. K., Lefer, B., Haman, C., Hair, J. W., Rogers, R. R., Butler, C., Cook, A. L., and Harper, D. B.: Comparison of mixed layer heights from airborne high spectral resolution lidar, ground-based measurements, and the WRF-Chem model during CalNex and CARES, Atmos. Chem. Phys., 14, 5547–5560, doi:10.5194/acp-14-5547-2014, 2014.

Seidel, D. J., Ao, C. O., and Li, K.: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, J. Geophys. Res., 115, D16113, doi:10.1029/2009JD013680, 2010.

Seidel, D., Zhang, Y., Beljaars, A., Golaz, J.-C. and Medeiros, B.: Climatology of the planetary boundary layer over continental United States and Europe, J. Geophys. Res., 117, D17106, 2012.

Sharf G., Peleg M., Livnat M, and Luria M.: Plume rise measurements from large point sources in Israel. Atmos. Environ, 27A, No.11, pp 1657-1663,1993.

Steppeler J., Doms G., Schattler U., Bitzer HW, Gassmann A., Damrath U., Gregoric G.: Meso gamma scale forecasts by nonhydrostatic model LM. Meteorological Atmospheric Physics, 82, 75–96, 2003.

Stull R.B.: An introduction to boundary layer meteorology, Kluwer Academic Publishers, the Netherlands, 666 p., 1988.

Su, T., Li, Z., and Kahn, R.: Relationships between the planetary boundary layer height and 1027 surface pollutants derived from lidar observations over China, Atmos. Chem. Phys., 18, 1028 15921-15935, 2018.

Tiedtke, M., 1989: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. Mon. Wea. Rev., 117, 1779–1800, https://doi.org/10.1175/1520-0493, 1989.

Uzan, L. and Alpert, P.: The coastal boundary layer and air pollution - a high temporal resolution analysis in the East Mediterranean Coast, The Open Atmospheric Science Journal, 6, 9–18, 2012.

-Uzan, L., Egert, S., and Alpert, P.: Ceilometer evaluation of the eastern Mediterranean summer boundary layer height – first study of two Israeli sites, Atmos. Meas. Tech., 9, 4387–4398, https://doi.org/10.5194/amt-9-4387-2016, 2016.

~~Von Engeln A~~Uzan, L., Egert, S., and ~~Teixeira~~Alpert, P.: New insights into the vertical structure of the September 2015 dust storm employing eight ceilometers and auxiliary measurements over Israel, Atmos. Chem. Phys., 18, 3203-3221, https://doi.org/10.5194/acp-18-3203-2018, 2018.

Van der Kamp D., McKendry I.: Diurnal and seasonal trends in convective mixed-layer heights estimated from two years of continuous ceilometer observations in Vancouver, BC. Bound Layer Meteorol, 137(3), 459–475, 2010.

Vogelezang D.H.P., and Holtslag A.A.M.: Evaluation and model impacts of alternative boundary-layer height formulations, Boundary-Layer Meteorol. 81, 245-269,1996.

Urbanski, S., Kovalev, V.A., Hao, W.M., Wold, C., Petkov, A.: Lidar and airborne investigation of smoke plume characteristics: Kootenai Creek Fire case study, Proceedings of 25th International Laser Radar Conference, St. Petersburg, Russia, Tomsk. Publishing House of IAO SB RAS. p. 1051–1054,2010.

Wang, Z., Cao, X., Zhang, L., Notholt, J.: A., Zhou, B., Liu, R., and Zhang, B.: Lidar measurement of planetary boundary layer height climatology derived from ECMWF reanalysis data. and comparison with microwave profiling radiometer observation, Atmos. Meas. Tech., 5, 1965–1972, https://doi.org/10.5194/amt-5-1965-2012, 2012.

Wetzel P.J. Climate., 126, 6575-6590, https://doi.org/10.1175/JCLI-D-12-00385, 2013.: Toward parametrization of the stable boundary layer, J. Appl. Meteorol. 21, 7-13, 1982.

Wiegner M., Madonna F., Binietoglou I., Forkel R., Gasteiger J., Geiß A., Pappalardo G., Schäfer K., and Thomas W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, Atmos. Meas. Tech., 7, 1979–1997, https://doi.org/10.5194/amt-7-1979-2014, 2014.

Wiegner, M. and Gasteiger, J.: Correction of water vapor absorption for aerosol remote sensing with ceilometers, Atmos. Meas. Tech., 8, 3971–3984, https://doi.org/10.5194/amt-8-3971-2015, 2015.

Yuval, Dayan, U., Levy, I., & Broday, D. M: On the association between characteristics of the atmospheric boundary layer and air pollution concentrations, Atmospheric Research, doi.org/10.1016/j.atmosers.2019.104675, 2019.

Zhang, Y., Gao, Z., Li, D., Li, Y., Zhang, N., Zhao, X., and Chen, J.: On the computation of planetary boundary-layer height using the bulk Richardson number method, Geosci. Model Dev., 7, 2599-2611, https://doi.org/10.5194/gmd-7-2599-2014, 2014.

Ziv B., Saaroni H., and Alpert P.: The factors governing the summer regime of the eastern Mediterranean, Intern. J. of Climatol., 24, 1859-1871, 2004.

Table 1. Location ~~of measurements sites and ceilometer types~~and type of ceilometers

Location	~~Site~~ ~~Long/~~Terrain	Lat/Lon	Distance from ~~-~~MD[c] shoreline ~~(km)~~ (km)	Height (m a.s.l)	Ceilometer type (resolution, ~~height limit[a])~~max range[a])
Ramat David (RD)	~~North~~Plain	32.7 °/35.2 °	24	50	CL31 (10 m,16 s, up to 7.7 km)
Hadera (HD)	~~Shoreline~~Coast	32.5 °/34.9 °	3.5	10	CL31 (10 m,16 s, up to 7.7 km)
Tel Aviv (TLV)	~~Shoreline~~Coast	32.1 °/34.8 °	0.05	5	CL31 (10 m,16 s, up to 7.7 km)
~~Beit~~Bet Dagan (BD)[b]	~~Inland~~Plain	32.0 °/34.8 °	7.5	33	CL31 (10 m,15 s, up to 7.7 km)
Weizmann (WZ)	~~Inland~~Plain	31.9 °/34.8 °	11.5	60	CL51 (10 m,16 s, up to 15.4 km)
Jerusalem (~~JR~~JRM)	~~Mountain~~Mount.	31.8 °/35.2 °	53	830	CL31 (10 m,16 s, up to 7.7 km)
~~Nevatim (NV~~Hazerim (HZ)	~~South~~Arid	31.2 °/34.~~9~~6 °	44	~~400~~ 200	CL31 (10 m,16 s, up to 7.7 km)
~~Hazerim (HZ~~Nevatim (NV)	~~South~~Arid	31.2 °/~~34.7~~35.0 °	70	~~200~~ 400	CL31 (10 m,16 s, up to 7.7 km)

[a]The maximum height ~~limit depends on sky conditions and~~ decreases as the atmospheric optical density ~~(AOD)~~ increases. ~~Data acquisition was limited~~

[b]Adjacent to ~~4.5 km by the ceilometers' software (BLview), except in Beit Dagan.~~

~~[b]The location of ceilometer Beit Dagan and~~ the radiosonde launch site.

[c] Mediterranean

Table 2. ParametersDescription of the NWP models

Model	IFS	COSMO
Convection parametrization | Type | Resolution (deg) | Operation center	ECMWFModel	IMS
Mass flux Tiedtke shallow convection Global/regional	Global	Regional, boundary conditions from IFS | 0.025 | IMS | COSMO
Mass flux Tiedke-Bechtold Horizontal grid resolution | Global	0.1 0.125° in 2015 (~13km) 0.1251° in 2016 (~9 km)	ECMWF0.025° (~2.5 km) | IFS
Vertical grid resolution	137 layers up to ~79 km 23 lie within the first 3 km	60 layers up to 23.5 km 20 lie within the first 3 km
Temporal resolution of the output	Hourly profiles	15 min profiles
Convection parametrization	Mass flux Tiedke-Bechthold (Bechthold, 2008)	Deep convection resolved. Parametrization of mass flux shallow convection. (Tiedtke, 1989)

~~Table 3. Statistical analysis of the Beit Dagan PBL heights on 33 summer days (13 days on August 2015 and 20 days on August 2016) from IFS and COSMO models by the bulk Richardson method (IFS_R, $COSMO_R$), the parcel method (IFS_P, $COSMO_P$) and the WCT method for the adjacent ceilometer. The PBL heights were compared to those derived from Beit Dagan radiosonde by either the parcel or bulk Richardson methods (same results, see Fig 2).~~

Table 3. Statistical analysis of Bet Dagan PBL heights (N=91, Fig. 2a)

PBL detection	~~IFS_R~~IFS-pm	~~IFS_P~~COSMO-pm	~~$COSMO_R$~~IFS-ri	~~$COSMO_P$~~COSMO-ri	Ceilometer	RS
Mean Error (m)	~~274~~346	~~271~~-52	~~-33~~66	~~-106~~57	~~-1~~24	-
RMSE (m)	~~432~~494	~~411~~146	~~152~~579	~~-176~~193	~~-97~~143	-
R	0.~~18~~14	0.~~21~~84	-0.~~83~~13	0.~~83~~7	0.~~93~~83	-
Mean ~~PBL~~ (m a.s.l)	~~-1250~~ 1236	~~1247~~838	~~-973~~1255	~~869~~947	~~989~~894	890
~~Std PBL~~STD (m)	~~274~~290	~~-245~~237	~~-273~~346	~~-222~~232	~~259~~239	245

~~Table 4. Root mean square errors of PBL heights from five sites on 13 summer days (Fig. 4), derived by IFS and COSMO models by the bulk Richardson method (IFS_R, $COSMO_R$) and the parcel method (IFS_P, $COSMO_P$). The PBL heights were compared to the heights measured by the Beit Dagan ceilometer.~~

Table 4. Statistical analysis of Tel Aviv PBL heights (N=122, Fig. 3a)

~~Site~~PBL detection		~~IFS_R~~IFS-pm	~~IFS_P~~COSMO-pm	~~$COSMO_R$~~IFS-ri	~~$COSMO_P$~~COSMO-ri	Ceilometer
~~Ramat David~~	~~173 m~~Mean Error (m)	~~180 m~~14	~~247 m~~17	~~232 m~~436	18	-
~~Tel Aviv~~	~~276~~ RMSE (m)	~~498 m~~256	~~203 m~~183	~~182 m~~616	180	-
~~Beit Dagan~~R		~~405 m~~0.47	~~569 m~~0.74	~~235 m~~-0.03	~~171 m~~0.73	-
~~Weizmann~~	~~214 m~~Mean (m a.s.l)	~~339 m~~702	~~175 m~~706	~~209 m~~1124	707	674
~~Jerusalem~~	~~351~~ STD (m)	~~285 m~~224	~~251 m~~238	~~179 m~~337	211	258

Table 5. ~~Same as in Table 3 but for mean errors.~~Statistical analysis of Ramat David PBL heights (N=123, Fig. 4a)

~~Site~~PBL detection		~~IFS_R~~IFS-pm	~~IFS_P~~COSMO-pm	~~COSMO_R~~ IFS-ri	~~COSMO_P~~COSMO-ri	Ceilometer
~~Ramat David~~	~~-31 m~~Mean Error (m)	~~-0 m~~4	~~-26 m~~40	~~-12 m~~446	123	-
~~Tel Aviv~~	~~234~~ RMSE (m)	~~422 m~~347	~~19 m~~245	~~-35 m~~745	313	-
~~Beit Dagan~~R		~~332 m~~0.14 ~~497 m~~ ~~12 m~~	~~-~~0.55~~ m~~	-0.08	0.39	-
~~Weizmann~~	~~114 m~~Mean (m a.s.l)	~~280 m~~995	~~16 m~~1031	~~-42 m~~1437	1114	991
~~Jerusalem~~	~~298~~ STD (m)	~~243 m~~276	~~-6 m~~256	~~-1 m~~521	268	253

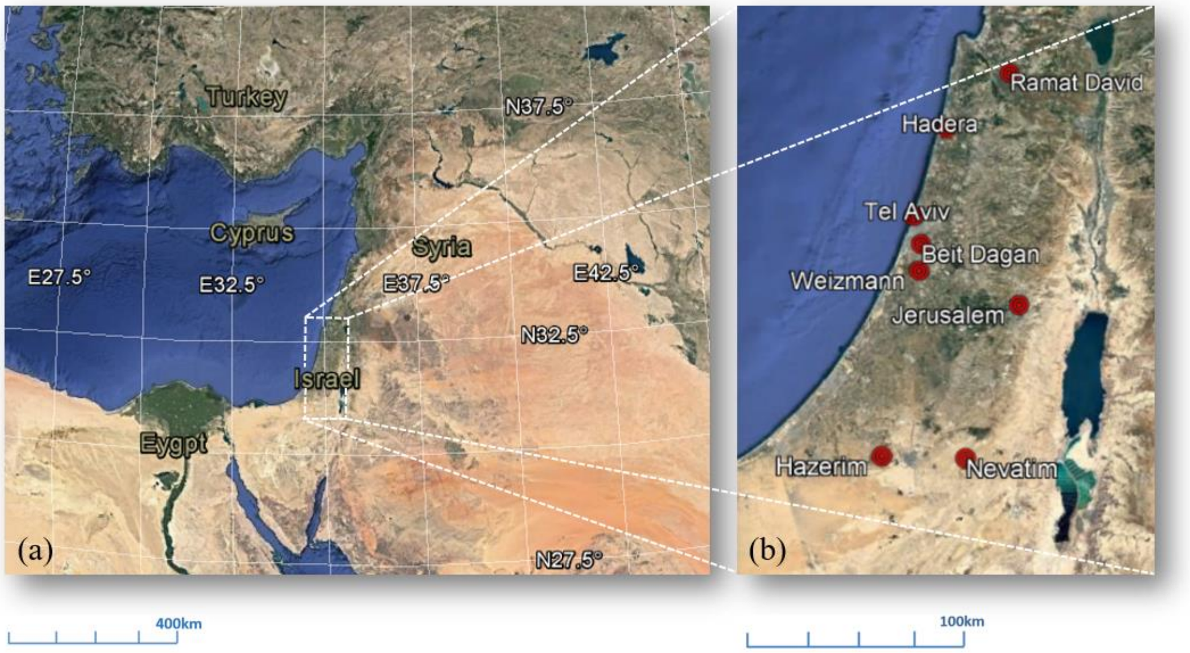


Table 6. Statistical analysis of Weizmann PBL heights (N=55, Fig. 5a)

PBL detection	IFS-pm	COSMO-pm	IFS-ri	COSMO-ri	Ceilometer
Mean Error (m)	67	-106	430	21	-
RMSE (m)	162	207	604	192	-
R	0.85	0.76	-0.01	0.72	-
Mean (m a.s.l)	892	719	1256	846	825
STD (m)	186	193	322	219	271

Table 7. Statistical analysis of Jerusalem PBL heights (N=53, Fig. 6a)

PBL detection	IFS-pm	COSMO-pm	IFS-ri	COSMO-ri	Ceilometer
Mean Error (m)	366	-129	117	-44	-
RMSE (m)	498	252	257	239	-
R	0.18	0.63	0.59	0.70	-
Mean (m a.s.l)	2239	1744	1991	1830	1874
STD (m)	276	253	258	328	250

Table 8. Statistical analysis of Nevatim PBL heights (N=72, Fig. 7a)

PBL detection	IFS-pm	COSMO-pm	IFS-ri	COSMO-ri	Ceilometer
Mean Error (m)	149	186	214	264	-
RMSE (m)	423	436	369	488	-
R	0.1	0.15	0.30	0.23	-
Mean PBL (m a.s.l)	1728	1756	1792	1843	1579
STD PBL (m)	341	352	268	394	237

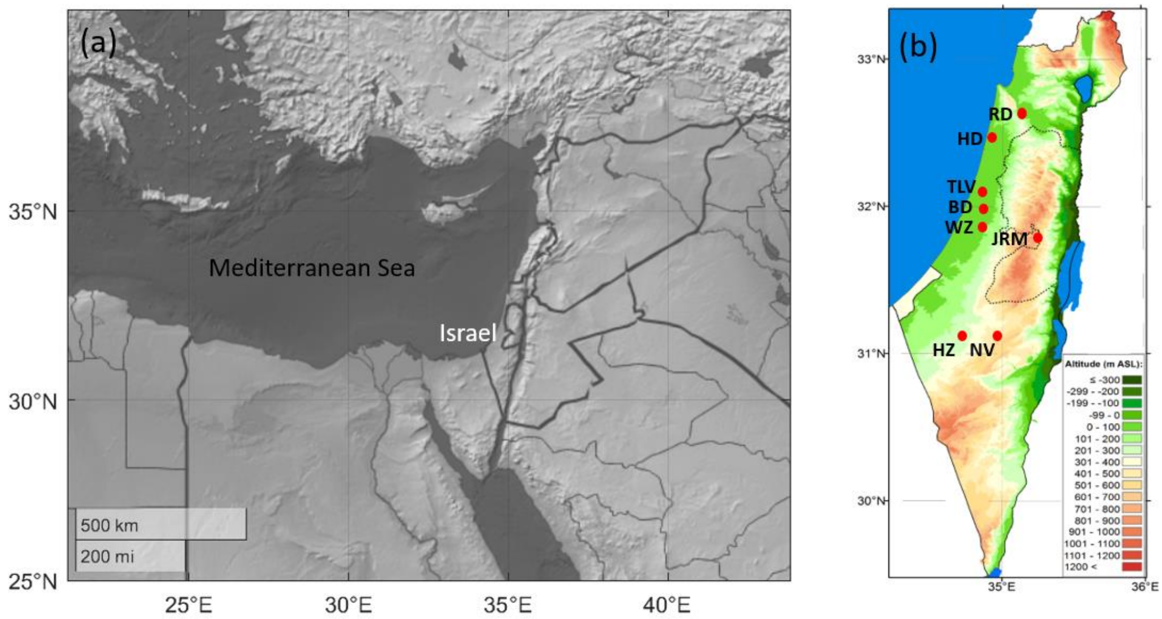


Fig. 1 Maps of ~~(a)~~ the East Mediterranean (a), and ~~(b)~~ the ~~research area including~~ study region in Israel (b), with indications of the ~~ceilometers~~ ceilometers' measurement sites (red circles~~). The Radiosonde launch site is situated in Beit~~, details given in Table 1) on a topography map adapted from © Israeli meteorological service.

.

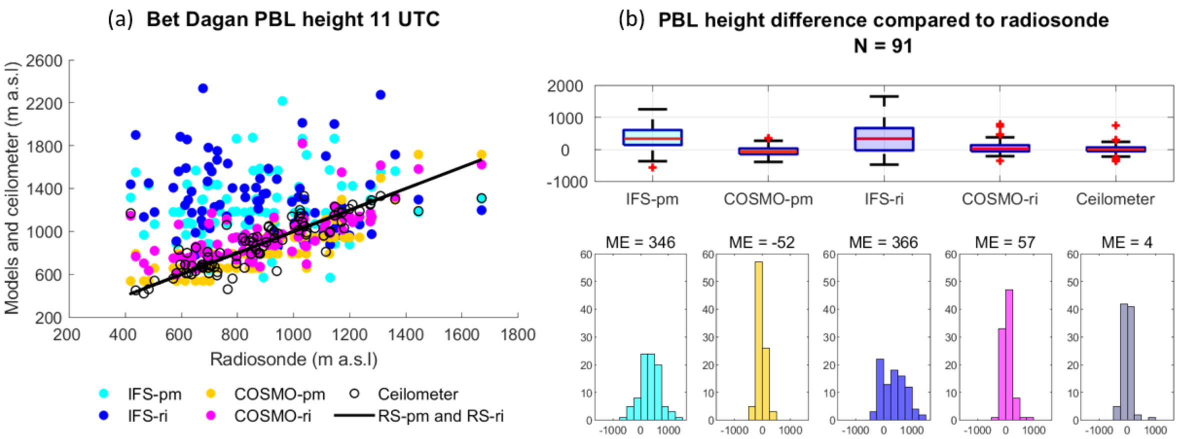


Fig. 2 PBL height from Bet Dagan~~, adjacent to the ceilometer. Adapted from © Google Maps 2019.~~ site at 11 UTC on 91 days for periods of July-September 2015 and June-September 2016. Ceilometer profiles analyzed by the WCT method. The IFS, COSMO, and radiosonde profiles analyzed by the bulk Richardson method (RS-ri, IFS-ri, COSMO-ri) and the parcel method (RS-pm, IFS-pm, COSMO-pm). The results compared to the radiosonde (RS-ri and RS-pm produced the same heights). Statistical analysis of the scatter plot (a) is given in Table 3. PBL height difference presented by boxplots and histograms (b). The edges of the boxplot are the 25th and 75th percentiles (q1 and q3), the whiskers enclose all data points not considered outliers (red crosses). A central red line indicates the median. Each boxplot is described by a histogram beneath.

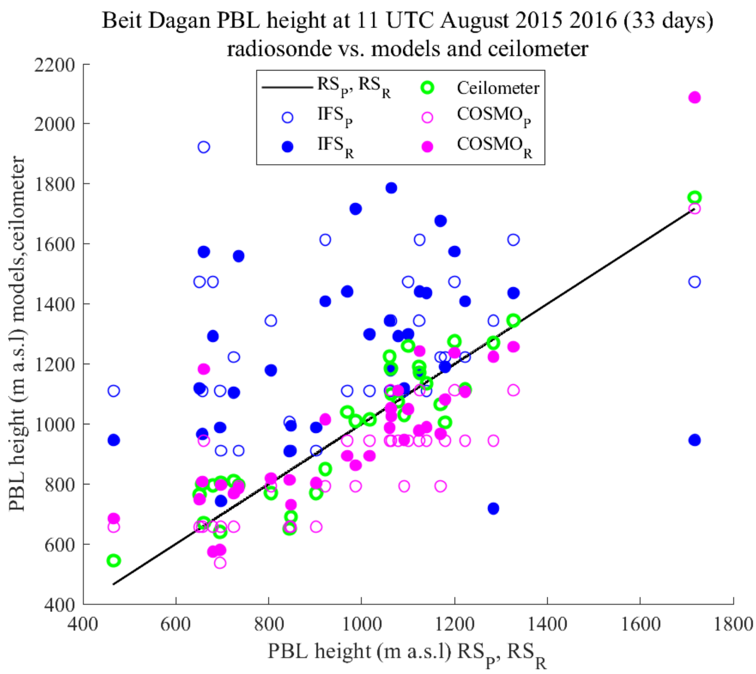
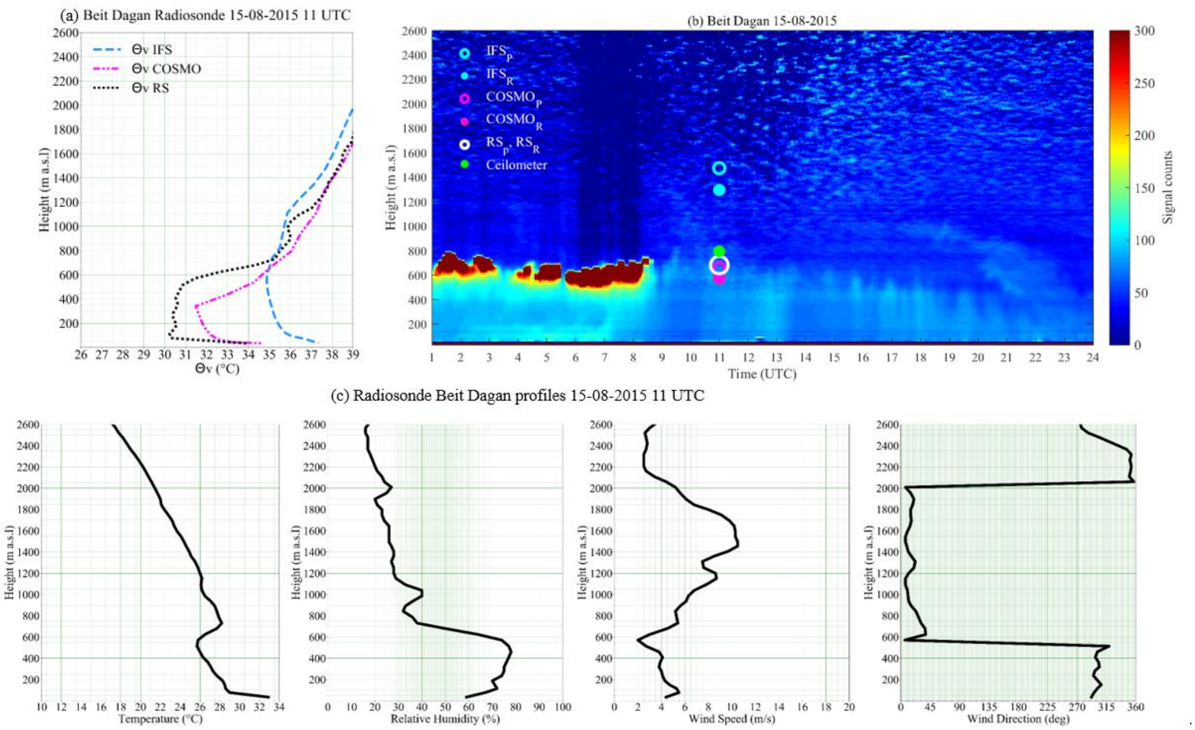
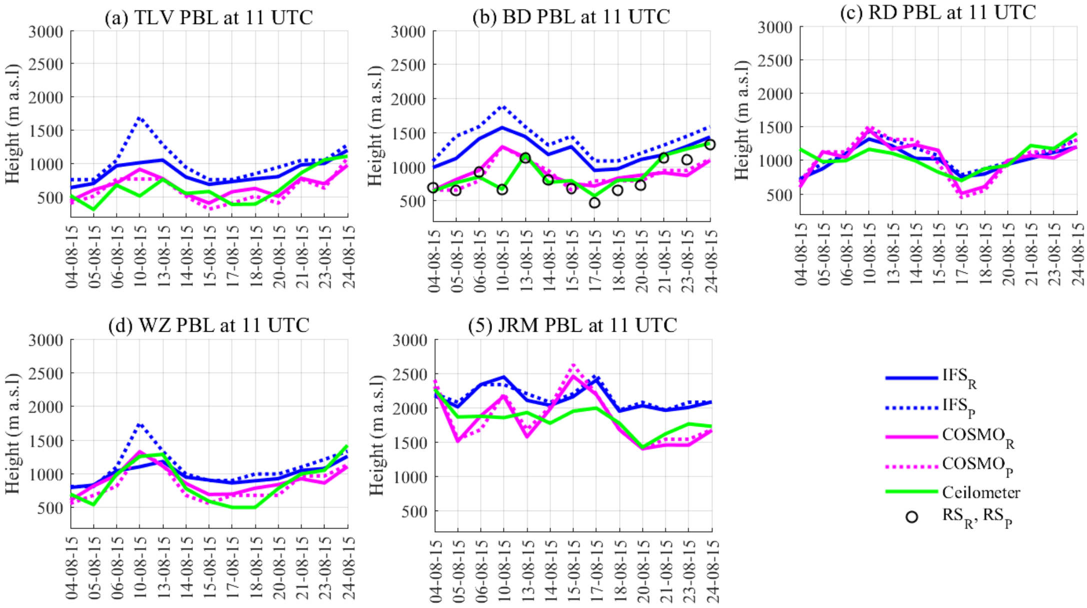


Fig 2. PBL heights over Beit Dagan site on 33 summer days (13 days on August 2015 and 20 days on August 2016), generated by the bulk Richardson method for IFS model (IFS_R, blue solid circles), COSMO model ($COSMO_R$, pink solid circles), and Beit Dagan radiosonde profiles (RS_R, black line). PBL heights generated by the parcel method for the IFS model (IFS_P, open blue circles), COSMO model ($COSMO_P$, open pink circles), and Beit Dagan radiosonde profiles (RS_P, same black line as RS_R, the results are identical). PBL heights derived from the Beit Dagan ceilometer were produced by the WCT method (green circles). Extreme Results (up to ~2, 00 m a.s.l) for August 17, 2016, are shown on the right-hand side.

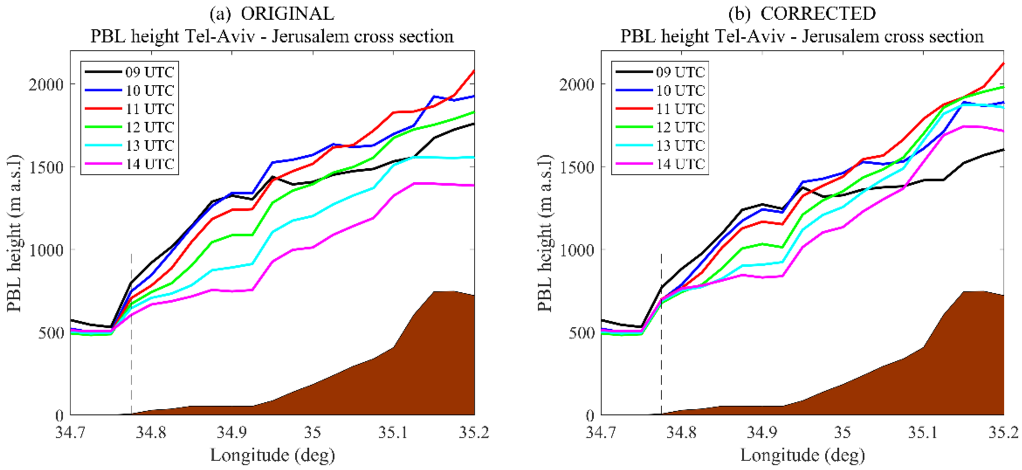


~~Fig.3 Meteorological measurements from Beit Dagan site on August 15, 2015: Virtual potential temperature profiles at 11 UTC generated from radiosonde measurements, IFS and COSMO models (a), ceilometer signal counts plot including indications of the PBL heights at 11 UTC from the models (IFS_R, IFS_P, $COSMO_R$, $COSMO_P$), radiosonde (RS_R, RS_P) and ceilometer (b). The bottom panel presents radiosonde profiles of temperature, RH, wind speed and wind direction at 11 UTC (c).~~

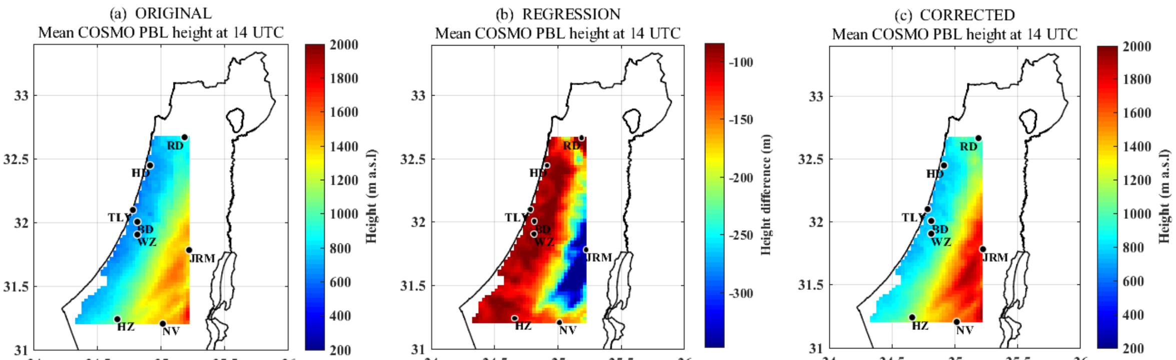
~~,~~



~~Fig. 4 PBL heights on 13 August days in 2015 from five ceilometer sites: (a) Tel Aviv (TLV), (b) Beit Dagan (BD), (c) Ramat David (RD), (d) Weizmann (WZ), and (e) Jerusalem (JRM). PBL heights were generated by the bulk Richardson method for the IFS model (IFS$_R$, blue solid line) and the COSMO model (COSMO$_R$, pink solid line). PBL heights generated by the parcel method for the IFS model (IFS$_P$, blue dashed line) and the COSMO model (COSMO$_P$, pink dashed line). Beit Dagan radiosonde profiles (RS$_R$, RS$_P$, black circles). PBL heights derived from the ceilometers (green line) were produced by the WCT method.~~



~~Fig. 5 COSMO$_R$ mean PBL height cross-section from Tel Aviv to Jerusalem before (a) and after (b) correction between 9-14 UTC. The analysis was performed on the number of available days for each site on August 2015 as follows: Jerusalem – 21 days, Nevatim – 13 days, Hazerim – 20 days, Ramat David – 26 days, Weizmann – 25 days, Beit Dagan – 13 days, Hadera – 16 days, Tel Aviv – 25 days. Indications of the seashore (dashed line) and the topography (brown area) are given.~~



~~Fig. 6 3D maps of COSMO$_R$ mean PBL heights over Israel at 14 UTC before (a), and after (c) correction. The regression (b) based on Eq. (6), depicts the height difference between the results from COSMO$_R$ and the ceilometers. The analysis was performed on the number of available days for each site on August 2015 as follows: Jerusalem – 21 days, Nevatim – 13 days, Hazerim – 20 days, Ramat David – 26 days, Weizmann – 25 days, Beit Dagan – 13 days, Hadera – 16~~

~~days, Tel Aviv - 25 days.~~

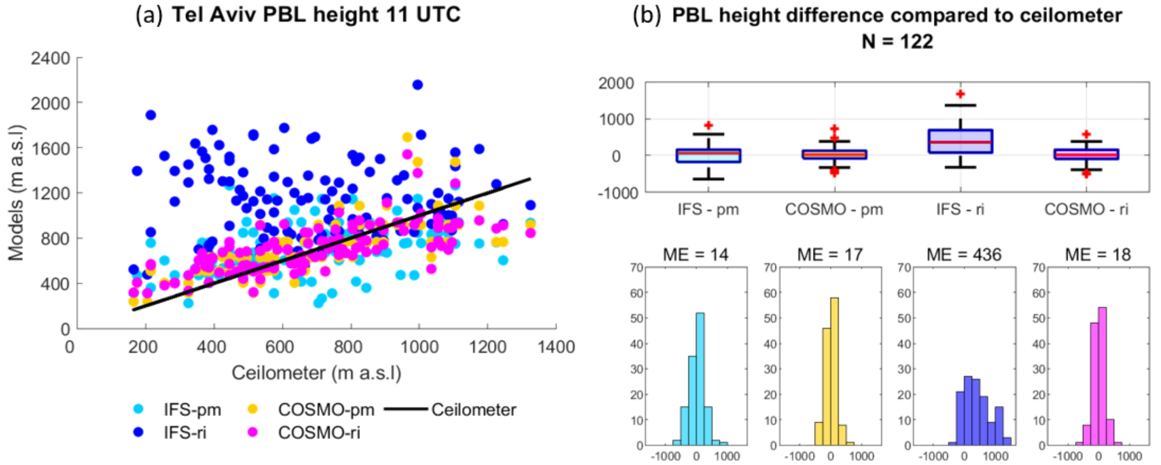


Fig. 3 Same as Fig. 2 but for Tel Aviv on 122 days. The models were compared to the ceilometer.

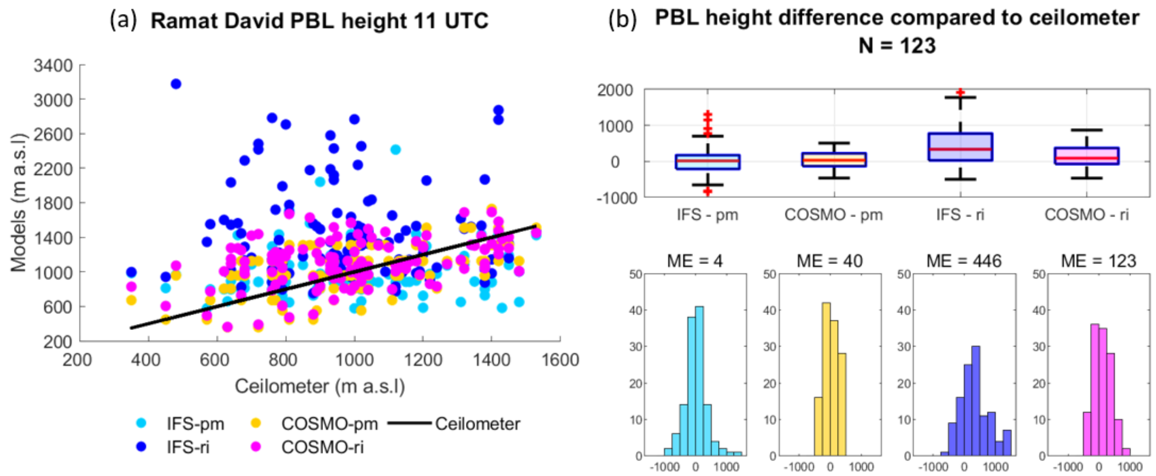


Fig. 4 Same as Fig. 2 but for Ramat David on 123 days. The models were compared to the ceilometer.

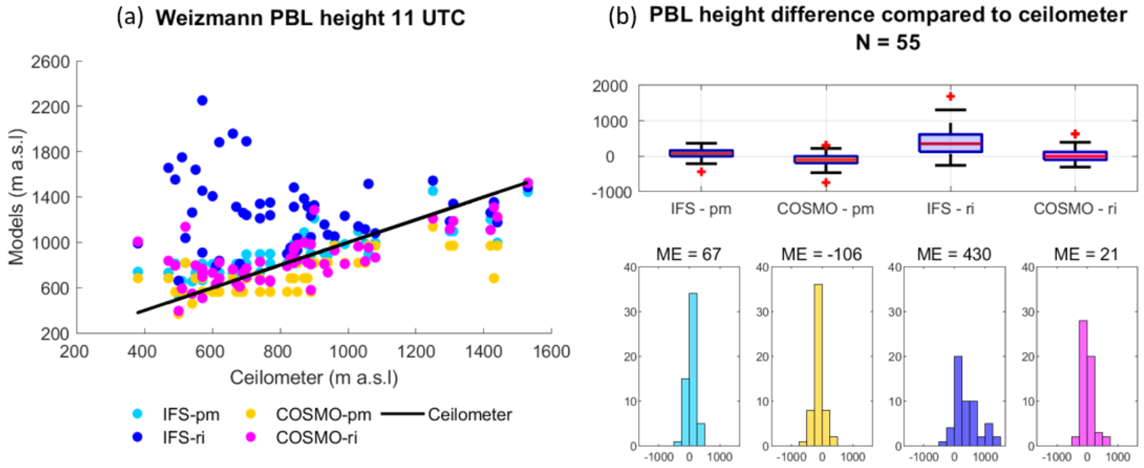


Fig. 5 Same as Fig. 2 but for Weizmann on 55 days. The models were compared to the ceilometer.

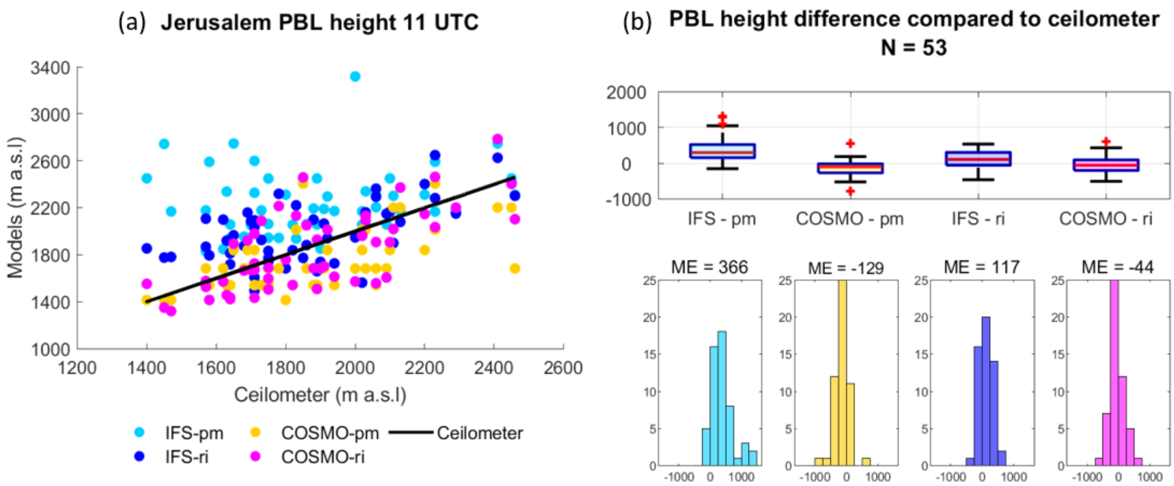


Fig. 6 Same as Fig. 2 but for Jerusalem on 53 days. The models were compared to the ceilometer.

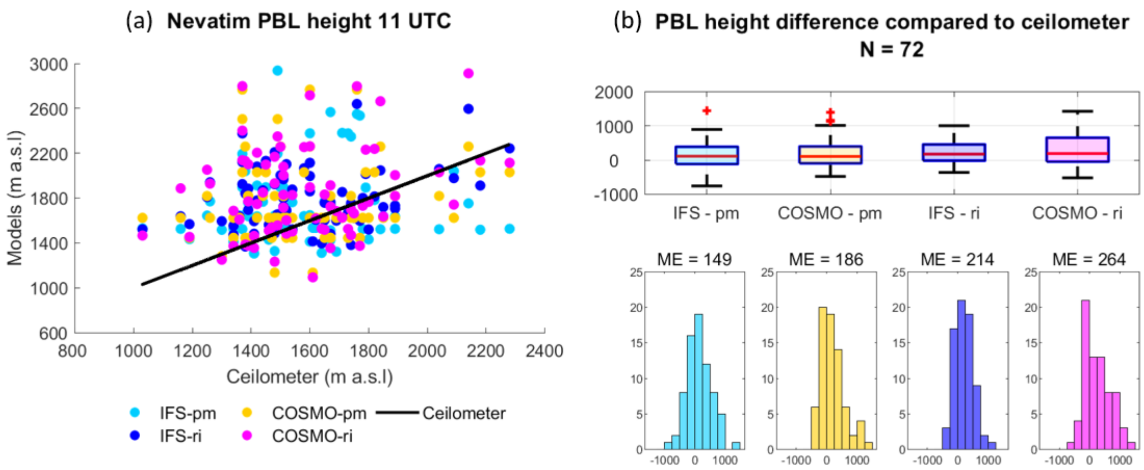


Fig. 7 Same as Fig. 2 but for Nevatim on 72 days. The models were compared to the ceilometer.

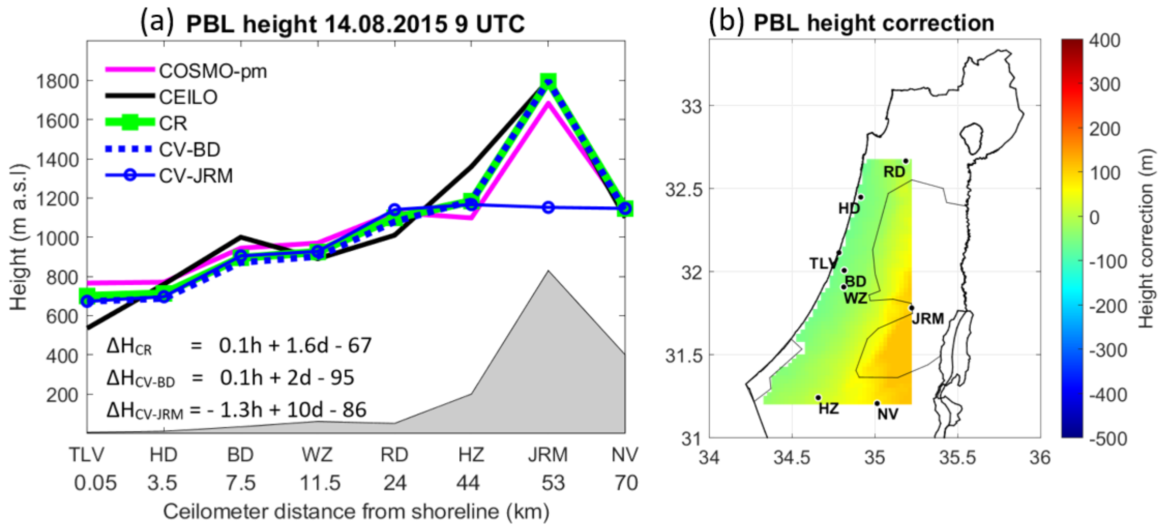


Fig. 8 PBL heights on August 14, 2015, at 9 UTC. The left panel (a) presents an east-west cross-section map, according to the ceilometers' distance from the Mediterranean shoreline. The PBL heights were derived from COSMO-pm (pink line), the ceilometers (black line), the correction tool for COSMO-pm (CR, green line), cross-validation for Bet Dagan (CV-BD, dashed blue line), and cross-validation for Jerusalem (CV-JRM, blue circles). The right panel (b) shows a 2-D map (b) of the height correction range, corresponding to figure (a).

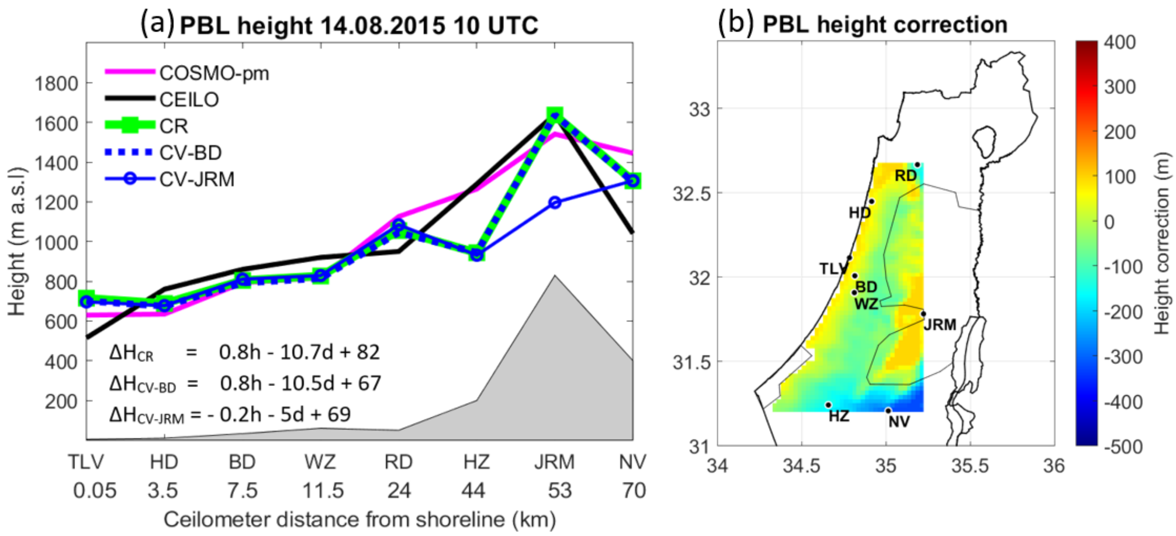


Fig. 9 Same as Fig. 8 but for 10 UTC.

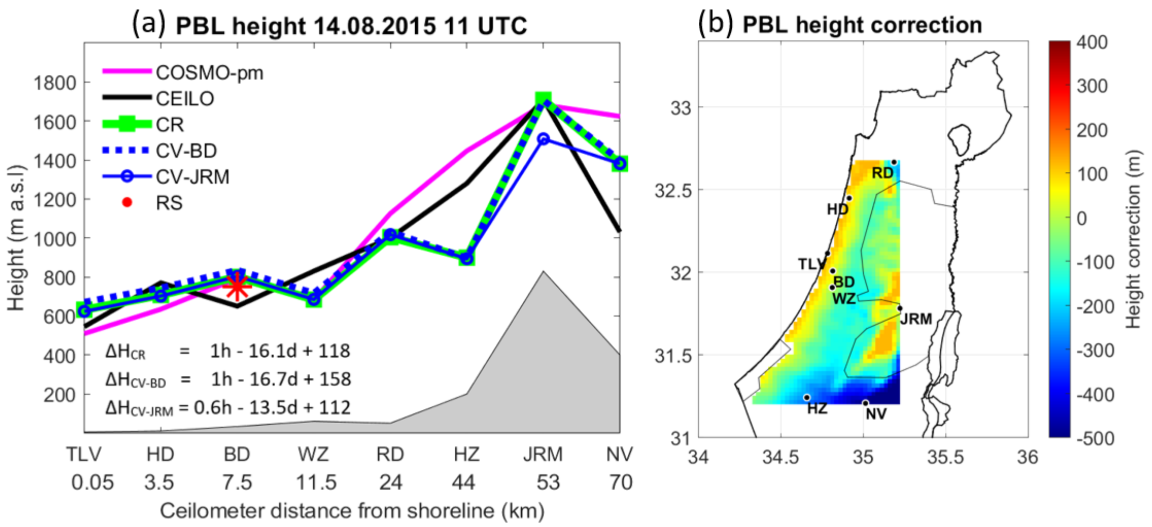
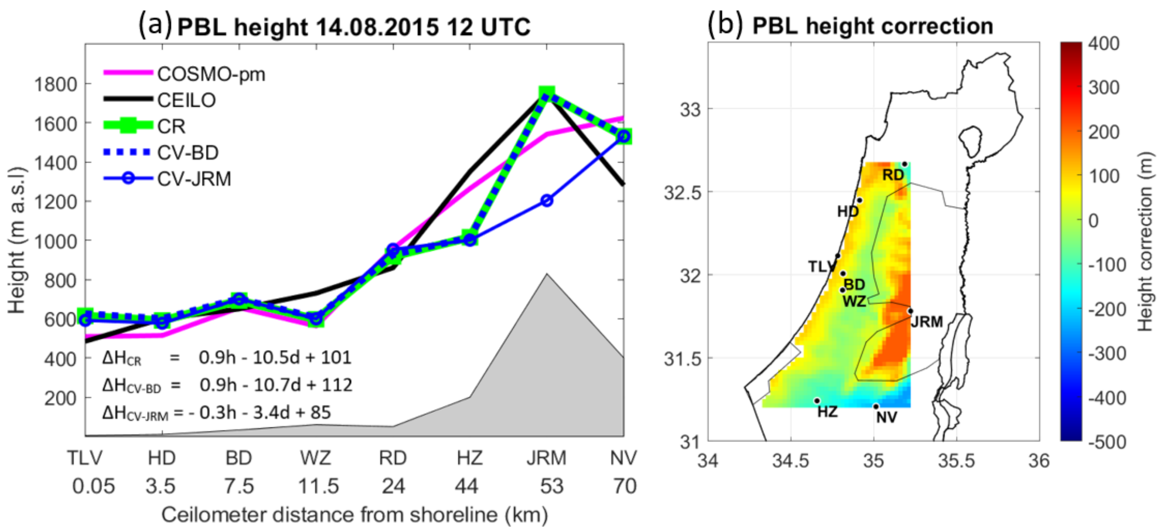
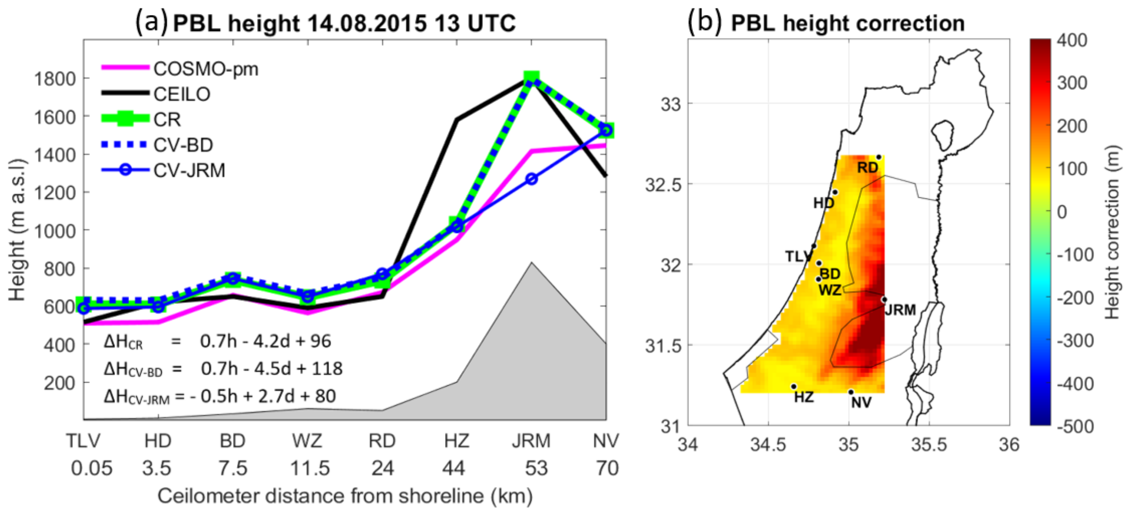


Fig. 10 Same as Fig. 8 but for 11 UTC and including the PBL height estimation from the radiosonde (red star).



1118 Fig. 11 Same as Fig. 8 but for 12 UTC.



1120 Fig. 12 Same as Fig. 8 but for 13 UTC.

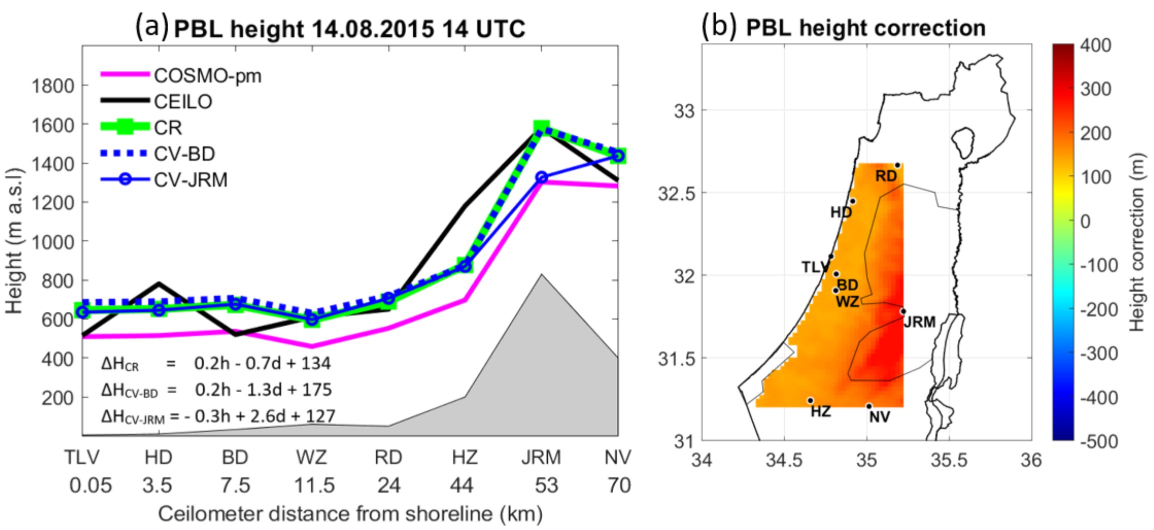


Fig. 13 Same as Fig. 8 but for 14 UTC.