Dear editor,
Dear Prof. David Topping,

Dear referees,

We would like to thank the editor for handling the manuscript and the referees for providing their review of the manuscript. We are very glad to note the appreciation of the potential of the CLASS4GL software.

Below, you can find our point-by-point response to the comments on the reviewers. We also provide the revised manuscript in two versions: one with track changes (below) and one without (see separate document).

We are hoping for your positive response.

Yours sincerely,
Dr. Hendrik Wouters
on behalf of the co-authors

Interactive comment on "Atmospheric boundary layer dynamics from balloon soundings worldwide: CLASS4GL v1.0" by Hendrik Wouters et al.
Anonymous Referee #1
Received and published: 4 April 2019

General Comments

This is a well-written manuscript which documents a powerful new software tool which the authors are making publicly available. This tool should allow researchers to perform extensive experiments related to boundary layer growth and development, including sensitivity to land surface and atmospheric inputs and parameters. The input datasets are global and extend back to 1981, allowing for easy application of experiments across climate regimes and seasons, and allowing users to test the representation of boundary layer dynamics in climate and earth system models. The authors include analysis of an initial experiment to demonstrate the first-order performance of the model. This projects looks like it could be extremely beneficial to the broad scientific community. I am fully in favor of this manuscript being published in Geoscientific Model Development. I document a few very minor suggestions below.

We would like to thank the referee for providing his/her review of the manuscript, and we are very glad to note the appreciation of the software's potential. We also acknowledge the comments, especially for improving the clarity of used methodologies, input data and results, and we provide a point-by-point answer below. Text of the revised manuscript is "quoted" in which the changes are provided in **bold** and removed text as ~~striked through~~. The location of text changes in the manuscript are provided in red. Unless specified otherwise, page and line numbers refer to the revised manuscript version with track changes.

Minor Suggestions

1. Section 2.1, line 15: "for which one also adds the entrainment flux driven by shear": I'm not quite sure what this phrase means. Do you mean that the component driven by shear is included in the 0.2(buoyancy flux) term, or that it should be added to this term? If it's the latter, you should change the word "for" to "to".*

It is the latter, namely the component driven by shear is added to the buoyancy flux afterwards. Hence, we change "for" to "to". It now reads as: "Entrainment flux is calculated as a fixed fraction (0.2) of the buoyancy flux, **to** which one also adds the entrainment flux driven by shear." P5R16

2. Line 18: Which parameters in the Penman-Monteith and other empirical equations are fixed and which are locally and/or seasonally determined from the input datasets?'

In order to answer this question, we now provide a table for the most important parameters in the Penman-Monteith and other empirical equations regarding to vegetation, soil and the air in the manuscript. The table indicates either whether it's a fixed parameter or specified from an external data source. In case of the latter, it is also indicated whether it's a static or time-dependent data. It is now included as a second table in the revised manuscript (P12 of revised manuscript without track changes), see also below:

“

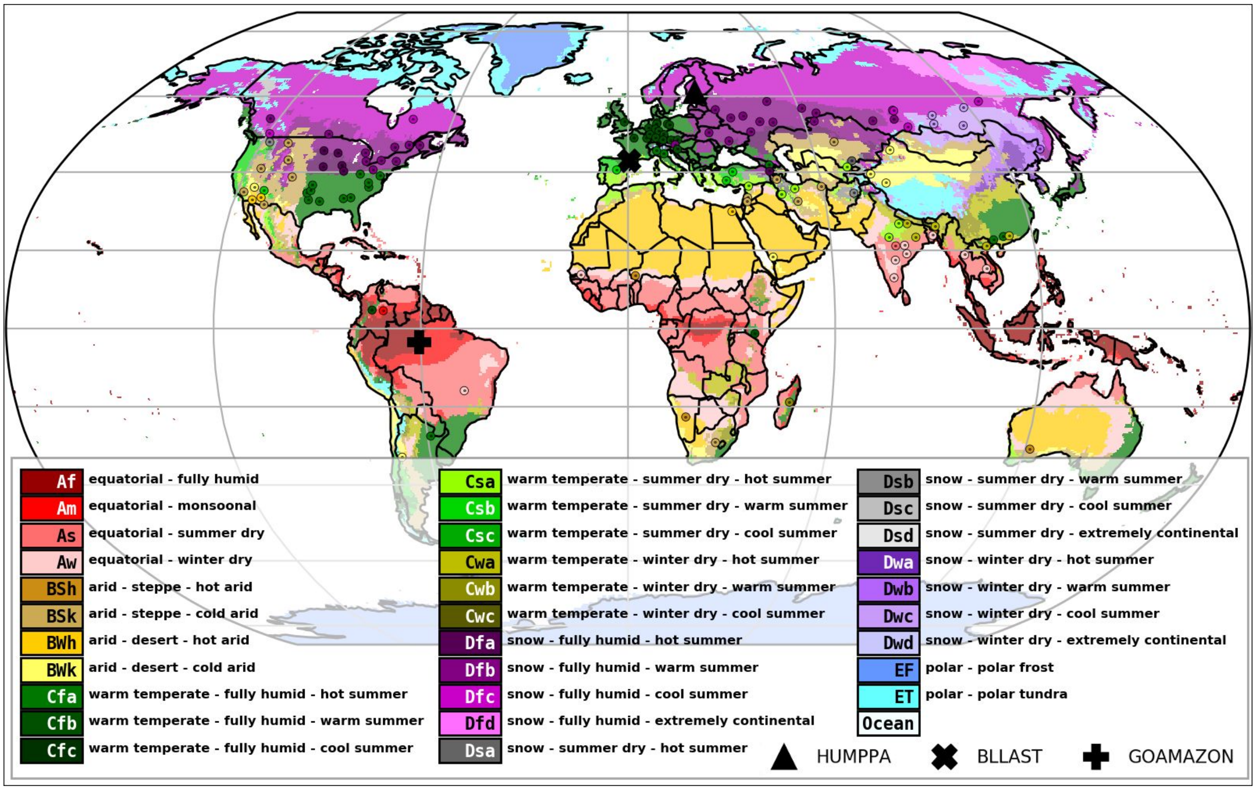
Vegetation			
Symbol	*Name*	*Unit*	*Default value or source*
LAI	Leaf area index of vegetated surface fraction	[-]	GIMSS (daily)
$r_{c,min}$	Minimum resistance transpiration	[s m^{-1}]	110
$r_{s,soil,min}$	Minimum resistance soil evaporation	[s m^{-1}]	50
g_D	Vapour pressure deficit correction factor for surface resistance	[-]	0
h_{can}	Canopy height	[m]	GLAS (static)
z_{0m}	Roughness length for momentum	[m]	0.1 x h_{can} (static)
z_{0h}	Roughness length for heat and moisture	[m]	0.1 x z_{0m} (static)
α	Surface albedo	[-]	MOD44B (static)
T_s	Initial surface temperature	[K]	ERA-Interim (3-hourly)
$T_{soil.1}$	Temperature top soil layer	[K]	ERA-Interim (3-hourly)
$T_{soil,2}$	Temperature deeper soil layer	[K]	ERA-Interim (3-hourly)
ω_{sat}	Saturated volumetric water content	[m^3 m^{-3}]	IGBP-DIS (static)
Soil			
Symbol	*Name*	*Unit*	*Default value or source*
ω_{fc}	Volumetric water content field capacity	[m^3 m^{-3}]	IGBP-DIS (static)
ω_{wilt}	Volumetric water content wilting point	[m^3 m^{-3}]	IGBP-DIS (static)
EF	Evaporative fraction	[-]	ERA5 (daily)
$\omega_{soil,1}$	Volumetric water content top soil layer	[m^3 m^{-3}]	By iterative matching of EF
$\omega_{soil,2}$	Volumetric water content deeper soil layer idem		*Idem*
c_{veg}	Vegetation fraction	[-]	MOD44B (static)
a	Clapp and Hornberger retention curve parameter	[-]	HWSD (static)
b	Clapp and Hornberger retention curve parameter	[-]	HWSD (static)
p	Clapp and Hornberger retention curve parameter	[-]	HWSD (static)
C_{Gsat}	Saturated soil conductivity for heat	[K m^{-2} J^{-1}]	HWSD (static)
$C_{2,sat}$	Coefficient force term moisture	[-]	HWSD (static)
$C_{2,ref}$	Coefficient restore term moisture	[-]	HWSD (static)
Λ	Thermal diffusivity skin layer	[-]	5.9
Air			
Symbol	*Name*	*Unit*	*Default value or source*

β	Ratio between buoyancy virtual heat and entrainment virtual heat	[-]	0.2
γ_θ	Initial lapse rate of potential temperature in the free atmosphere	[K m^{-1}]	From profile (IGRA)
γ_q	Initial lapse rate of specific humidity in the free atmosphere	[kg kg^{-1} m^{-1}]	*Idem*
γ_u	Initial lapse rate of zonal wind in the free atmosphere	[s^{-1}]	*Idem*
γ_v	Initial lapse rate of meridional wind in the free atmosphere	[s^{-1}]	*Idem*
$\Delta\theta_h$	Initial temperature jump between the mixed layer and free atmosphere	[K]	*Idem*
$\Delta\theta_q$	Initial specific humidity jump between the mixed layer and free atmosphere	[kg kg^{-1}]	*Idem*
$\Delta\theta_u$	Initial specific zonal wind jump between the mixed layer and free atmosphere	[m s^{-1}]	*Idem*
$\Delta\theta_v$	Initial specific meridional wind jump between the mixed layer and free atmosphere	[m s^{-1}]	*Idem*

Table 2. Main input parameters for CLASS4GL. The parameter specifications and source acronyms are explained in section 2.3 and table 1.
“

3. Figure 2: It took me a while to fi-nd the big X for the BLLAST experiment location, in part because much of the X is on top of country lines. Perhaps you can use a different Symbol.

The symbol for the BLLAST location is now replaced with a thick cross. It will now appear as follows, which should make it more traceable (P7):



4. Section 3, line 3: When you first mention daytime tendencies, can you clarify what time period the resultant values are averaged over? I imagine it is from sunrise through the time of the second sounding.

The tendencies are averaged from the morning sounding to the afternoon sounding. This will appear explicitly in the revised text as follows:

"The evaluation is done by comparing the modelled daytime tendencies of the mixed-layer height (*dh/dt*), potential temperature (*dθ/dt*) and specific humidity (*dq/dt*) against the corresponding tendencies observed from the balloon sounding pairs. **Observed and modelled tendencies represent the mean diurnal change from the morning sounding to the afternoon sounding.**" P14R4-7

Since the local timing of this second sounding is at different times of day in different longitudes, might this introduce a spatial bias since BL growth rates are not uniform over the course of the day?

Ideally, one would always have a sounding at the same local time in the morning and in the afternoon for every location. Since the launching times are based on UTC, this is obviously not the case. So we agree that the common local time of the sounding launch depends on longitude, and also that the ABL growth is certainly not uniform over the course of the day. However, the latter is taken into account, since the model is always initialized with the morning sounding for which the initial local model time is set equal to the sounding launch, and the same is true for the afternoon sounding. So it can be concluded that the expected tendency for each launch or site (depending on the local time window being considered in the computation of that tendency) is equivalent for observations and model simulations, hence any biases related to launching times between the two are avoided.

This will be clarified with the following additional text (it will be located just after the previous text):

"**It should be noted that the local time of the morning and afternoon soundings changes given that the launch times are often at 0 and 12 h UTC, and that the boundary-layer tendencies are not uniform over the course of the day. The resulting variety in the tendencies is taken into account in the simulations, since the model is initialized with the morning sounding while the initial solar local time in the model is set equal to the sounding launch. The same happens for the end of the simulation at the time of the afternoon sounding. Therefore, the expected tendency for each launch or site (depending on the local time window being considered in the computation of that tendency) is equivalent for observations and models, hence any biases related to launching times between the two are avoided.**" P14R7-13

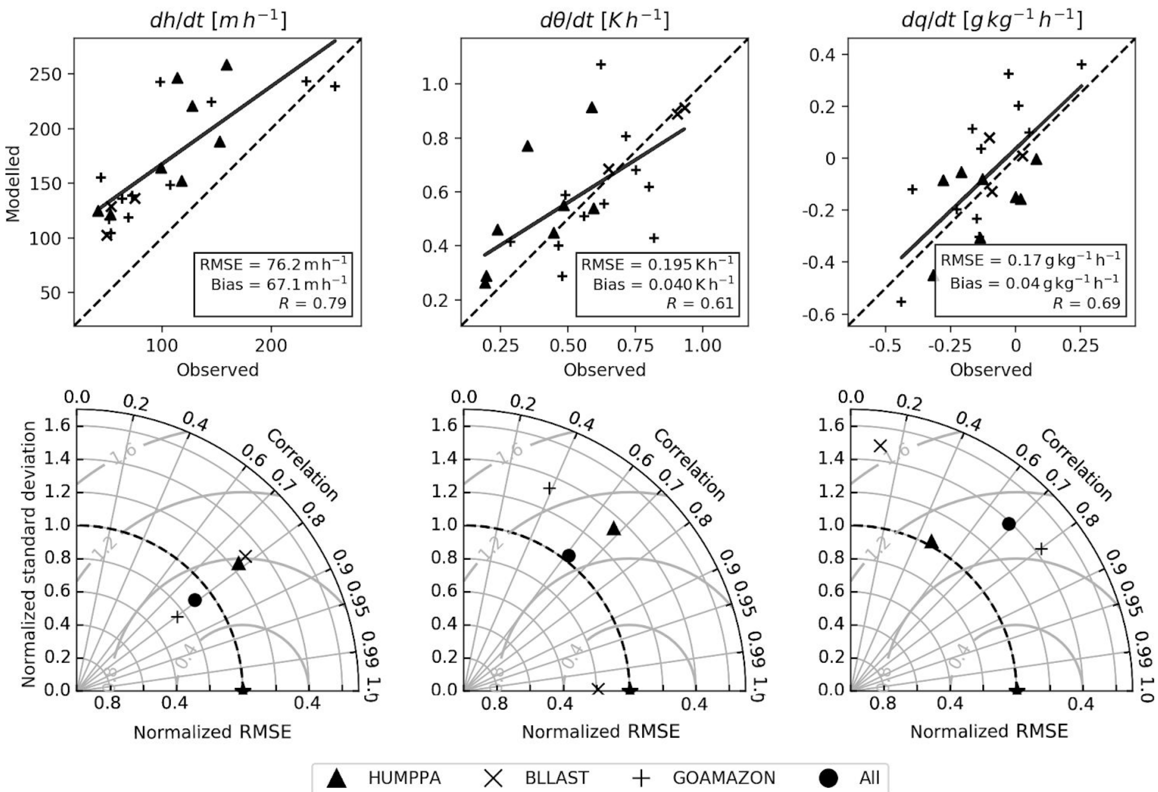
5. Page 12, line 2: Here you mention the observed daytime tendencies when you are discussing the results from the three intensive campaigns. Are the results you show actually subdaily averages since you have more than two soundings per day?

Yes, as clarified above, the results reflect the diurnal tendencies averaged over the time span between the morning and afternoon sounding, hence depend on the sounding launch times. In case there are multiple soundings retained during a particular day which is especially the case for the campaigns, the sounding closest to sunrise is taken for the initialization in the morning, and the latest sounding for the validation in the afternoon. This was not mentioned explicitly in the text. This clarification will now be added in the methodology section '2.2. Automized balloon data mining':

"... Here, the afternoon radiosonde profile on the same day needs to occur between local noon and 1 h before sunset **(defined as the time when the incoming shortwave radiation at the top of the atmosphere becomes zero)**, and at least 4 h after ~~the morning sounding~~ **the model initialization by the morning sounding so that a sufficiently large model time span is considered**. **In case there are more than two soundings retained during a particular day which especially occurs during the campaigns, the sounding closest to sunrise is taken for the initialization in the morning, and the latest sounding for the validation in the afternoon**. **so that a sufficiently large model time span is considered**. **In case there are more than two soundings retained during a particular day which especially occurs during the campaigns, the sounding closest to sunrise is taken for the initialization in the morning, and the latest sounding for the validation in the afternoon**. ..." P9R24-35

6. Figure 3: the correlation plots are quite busy, making it a touch hard to find the three symbols of interest. Maybe you can make the grey lines a little bit lighter grey so the symbols are easier to see.

The grid lines are now lighter which increased the visibility:

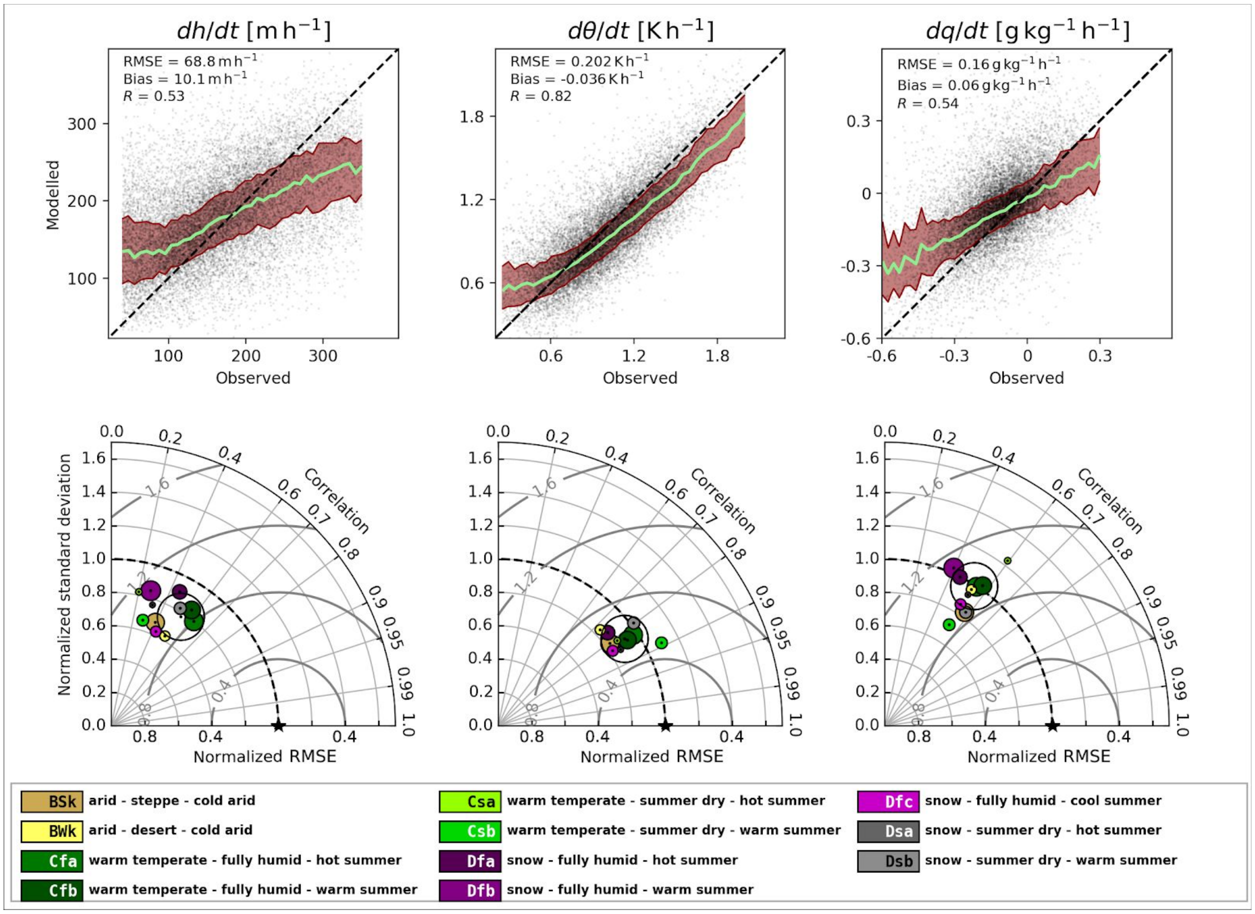


The updated figure is now included in the revised manuscript (P15).

7. Page 13, line 8: It is not clear to me where the 22% value comes from. Please clarify.

The 22% value refers to the departure of the (normalized) standard deviation of the model output from the (normalized) standard deviation of the observations. This can be seen on the Taylor plots in Fig. 4 (figure pasted below) for which the centers of the open circles are between 0.78 and 1.22 of the normalized standard deviation. This will be made clear in the text as follows:

"In addition, the overall modelled range in dh/dt, dθ/dt and dq/dt agrees well with the observed range~~, with departures from the standard deviation of the observations below 22% – see Taylor plots in Fig. 4.~~**, for which the departure of the modelled (normalized) standard deviation from the observed (normalized) standard deviation of each tendency is below 22%. This can be seen in the Taylor plots in Fig. 4 for which the centers of the open circles are between 0.78 and 1.22 of the normalized standard deviation.**" P16R12-16



Reviewer 2

Interactive comment on "Atmospheric boundary layer dynamics from balloon soundings worldwide: CLASS4GL v1.0" by Hendrik Wouters et al.
The paper presents a significant advancement in producing both useful accessible boundary layer data from radiosondes, and a nice marriage with a simple ABL model to produce continuous ABL data constrained by analyses, with open-source software.

We would like to thank the referee for providing his/her review of the manuscript, and we are very glad regarding the appreciation of the software's potential. We also appreciate the comments, especially the suggestions related to possible sources of model biases. We provide a point-by-point answer below. Text of the revised manuscript is "quoted" in which the changes are provided in **bold** and removed text as ~~striked through~~. The location of text changes in the manuscript are provided in red. Unless specified otherwise, page and line numbers refer to the revised manuscript version with track changes.

I was not able to fully run the CLASS4GL software myself. On a Mac using MacPorts, the PyYAML was not available and I downloaded directly from the website - there were issues *recognizing the CLoader option - apparently a version inconsistency. But I will follow through as I would like to use this tool.*

Thank you for testing software! The goal is to have a platform-independent software, so we strive to make it work on all platforms including Mac systems. As a solution on Mac, we would like to suggest to try either a Python environment with anaconda (as explained on https://class4gl.eu/?page_id=105) or Pycharm+homebrew. In case of pycharm+homebrew, these are the needed steps to install the CLoader module:

```
brew install libyaml-dev
pip install pyyaml
```

Please note that '`brew install libyaml`' (so without '`-dev`') will not work. A similar solution may exist in case of your current Python environment using MacPorts. The CLoader is required to read yaml files 10–100 times faster, but depends on modules written in the C language.

Regarding the manuscript, I suggest only minor changes are needed (editorial and regarding content), as outlined below:

P3 L14: Use "automates" instead of "automises". Likewise on P4 L7.
P4 L3: Change "dirunal" to "diurnal"

Thanks for identifying the typos. They are corrected in the revised manuscript at P7R1, P18R2 and P4R3.

P5, L10-11: It is a common assumption that the heat, moisture and momentum content of the ABL are perfectly mixed, but of course there will be mean vertical gradients, especially near the entrainment zone and the surface. In other words, the gradients here are a little weaker than for a well-mixed ABL, which may be compensated by other parameter choices. What would be the effect of specifying more realistic but still simple tails (e.g., exponential or even linear) of theta, q and V at the top and bottom of the ABL? This will relate to comments below regarding apparent biases.

It is true that the ABL model considers a perfectly-mixed ABL with values of potential temperature, specific humidity and wind speed that are constant throughout the mixed layer, whereas the entrainment zone is represented as a jump between the mixed-layer values of and the free atmosphere values, and the surface layer as a analytic profile between mixed-layer values and the surface values. Other gradients within of the ABL are not explicitly represented. The Monin-Obukhov similarity theory is employed for calculating analytic surface layer profiles (and gradients) and the gradient transport in surface layer in an implicit way as a replacement for a more explicit representation. For the entrainment zone, the heat entrainment ratio (β) of 0.2 (the ratio of heat entrainment to heating through the surface layer) is considered and the additional entrainment by wind shear, based on observations and large eddy simulations (Vilà-Guerau de Arellano et al., 2015). More realistic tails at the top and bottom of the ABL are not yet considered in the ABL model, hence, it is not possible to quantify their effect and the possible associated biases in the model. Therefore, one would require a dedicated study for which one needs substantial changes to the ABL model formulations. We are aware about the model limitations and associated uncertainties, and about the need for more research employing more realistic profiles. This is now mentioned more explicitly in the revised manuscript as follows:

"The use of the mixed-layer equations implies that turbulence **and vertical gradients** inside the mixed layer ~~is~~are not explicitly **re**solved, and the potential temperature (θ), specific humidity (q) and wind components are assumed to be homogeneous within the ~~ABL~~**mixed layer**. This assumption tends to be supported by the efficient turbulent mixing under convective conditions (Bauer, 1908). At the top of the ~~ABL~~**mixed layer**, the entrainment of heat and moisture is parameterized by a jump of θ, q and wind components over an infinitesimally small height, which ~~have~~ are initialized with a constant lapse rate with height in the overlying free atmosphere. Entrainment flux is calculated as a fixed fraction (0.2) of the buoyancy flux, ~~for~~ **to** which one also adds the entrainment flux driven by shear. An important feature of the model is the possibility to represent the subsidence coupled to the entrainment process at the inversion zone (Vilà-Guerau de Arellano et al., 2015). *[following text was moved up]* The surface–atmosphere turbulent exchanges for momentum, heat, and moisture **in the surface layer** are calculated considering their aerodynamic resistances. These are calculated in an iterative way assuming constant values for aerodynamic roughness lengths, while applying correction factors for non-neutral stratification of the atmospheric surface-layer (Paulson, 1970) according to the Monin–Obukhov similarity theory (Monin and Obukhov, 1954). **It should be kept in mind that more realistic profiles with explicit ABL gradients for temperature, humidity and wind speed – especially at the top (entrainment zone) and bottom (surface layer) of the mixed layer – are not yet considered by the model. In order to tackle these limitations and associated uncertainties, more research is needed employing more realistic profiles.**" P5R9-R25

P7 L4-10: Please state how many (or what percentage) of the 42,000 profiles are excluded for each reason (lacking both 00 and 12UTC soundings vs. non well-mixed profiles? The first seems a hard criterion, but exactly how well-mixed is that criterion and what if it is relaxed?

The criterion for a well-mixed profile is that the root mean square error of the profile measurements in the boundary layer is lower than 1.5°C. This information is now added to the manuscript. In addition, we provide the statistics on the reasons of profile retainment for each filtering step. Therefore, paragraph 2.2 is revised and it reads in the revised manuscript as follows (additional information is indicated in bold):

"2.2 Automized balloon data mining

Global data of weather balloon soundings are taken from the Integrated Global Radiosonde Archive (IGRA; Durre et al., 2006) which is maintained under the auspices of the National Oceanic and Atmospheric Administration (NOAA). The IGRA archive is routinely updated and currently includes more than 2700 stations covering major global climate regions. The CLASS4GL sounding database is additionally supplemented with data from intensive radiosonde campaigns from HUMPPA (Williams et al., 2011), BLLAST (Pietersen et al., 2015) and GOAMAZON (Martin et al., 2016) – see Tab. 1 and Fig. 2. Other sources of vertical profile data (from e.g., aircraft, satellites, other observation campaigns or long-term operational soundings) may be considered in future applications of the framework. As described above, CLASS requires morning sounding profiles for initialization and afternoon profiles for validation to enable a mechanistic interpretation of the diurnal ABL evolution.

[new paragraph]

All balloon sounding profiles **(~15 million profiles)** are pre-processed first **by calculating the bulk mixed-layer properties:** ~~An estimation of ABL properties is obtained for the selected profile pairs.~~ ~~First, t~~The mixed-layer height (h) is assessed as the height at which the Bulk Richardson number (RiB) exceeds a critical value (RiBc). We adopt the estimates for RiBc provided by Zhang et al. (2014): RiBc = 0.24 for strongly stable boundary layers, RiBc = 0.31 for weakly stable boundary layers, and RiBc = 0.39 for unstable boundary layers. The uncertainty range of h (used below) is determined from its interval corresponding to the RiBc range [0.24, 0.39], for which the interval is further extended to the nearest sounding records above and below. Second, the mixed-layer potential temperature (θ), specific humidity (q), zonal wind (u) and meridional wind (v) are calculated as their average values recorded within the mixed layer. The capping inversion is estimated by a linear extrapolation of the two lowest sounding measurements above h, for which its lapse rate for potential temperature ($\gamma\ \theta = d\theta/dz$), specific humidity ($\gamma\ q = dq/dz$) and wind components ($\gamma\ u = du/dz$ and $\gamma\ v = dv/dz$) are calculated. The jump values at the h for potential temperature ($\Delta\theta$), specific humidity (Δq) and wind components (Δu and Δu) are estimated from the difference between the values of the capping inversion at h and the values within the mixed-layer.

Afterwards, morning--afternoon profiles are selected that meet a series of selection criteria: the morning profiles, ie. profiles before 12 h local time, are selected first, and they amount to ~ 6 million profiles. Here, the selection of suitable morning soundings **(and the subsequent afternoon soundings after 12 h)** ~~balloon sounding (morning–afternoon) pairs~~ is ~~largely~~ based on the timing of these soundings (a): Morning **(**and afternoon**)** sounding profiles ideally should be acquired after sunrise and before sunset, respectively. However, routine sounding launches happen synchronously on a daily basis at 0 h and 12 h UTC, whereas launches at intermediate timings (3 h, 6 h, 9 h, 15 h and 18 h UTC) are rare. As a result, many launches, especially those at 0h UTC in Europe and Africa, often happen several hours before sunrise. Since the net exchanges near the surface for heat, moisture and radiation are generally low at the end of the night, the atmospheric profiles tend not to change dramatically before sunrise (unless the synoptic situation changes), being often representative for the time the convective ABL starts to emerge (van Stratum and Stevens, 2018). As such, in order to maintain a high number of soundings in our analyses, launch times within 3 h prior to sunrise are still allowed here. For these soundings, the ABL simulation starts at sunrise, assuming that the change in the atmospheric profile since the balloon launch time is negligible. Furthermore, **(b) only those soundings are retained with more than seven measurements in the vertical below 3000 m (72% of the morning soundings), (c) for which the uncertainty of the mixed-layer height is lower than 150 m (26% of the morning soundings), (d) for which the ABL is sufficiently well-mixed (ie. for which the root-mean square deviation of the temperature from the estimated mixed-layer average is lower than 1.5°C; this criterion is met by 92% of the morning soundings). We also (e) set the morning lower temperature limit to 278 K in order to minimize the chance of freezing temperatures during the course of the simulations (this criterion is met by 70% of the morning soundings). The next criterion is that (f) an afternoon sounding can be found with the same criteria as the morning sounding except regarding the uncertainty of the mixed-layer height (which is met by 24% of the filtered morning soundings). Here, the afternoon radiosonde profile** on the same day needs to occur between local noon and 1 h before sunset, and at least 4 h after the morning sounding for allowing a sufficiently large timespan of the model simulations ~~model initialization in the morning~~**. We also require that (g) all external forcing parameters are available (which is met by 8.7% of the filtered sounding pairs). The above criteria lead to 21,826 profile pairs from 134 stations**. Finally, the current version of CLASS is only capable of representing growing mixed layers. Therefore, an observed mixed-layer growth of 40 m h^{-1} is considered as a lower limit (which is met by 85% of the profile pairs), **which leads to 18,385 profile pairs from 121 stations**. For the three intensive observation campaigns, 22 out of 49 profiles are retained. An overview of the global distribution of the retained radiosondes and their corresponding climate regime is shown in Fig. 2.

The above criteria are flexible and may be reconsidered according to the intended application, since there is an obvious trade-off between sounding quality and amount of data being retained. ...” P7R1-P10R8 [*end of the section can be found in the next reply below]*

P7 L12: Change "says" to "days".

Done. P10R9

P7 L13-17: Are there clear discrepancies between the behavior and/or statistics of gap-filled (model) versus observationally driven results? I assume you have looked at this - a caveat might be warranted here.

At this stage, the possibility of extracting profiles from continuous/gap-filled datasets (reanalysis, satellite-based products, and Earth system models...) is only implemented. The discrepancies between gap-filled versus observationally driven ABL model results have not been done yet. We agree that this is an important caveat, so the user should be warned that such a validation is needed as soon as a gap-filled dataset is employed. Hence, the following text will be added to the revised manuscript:

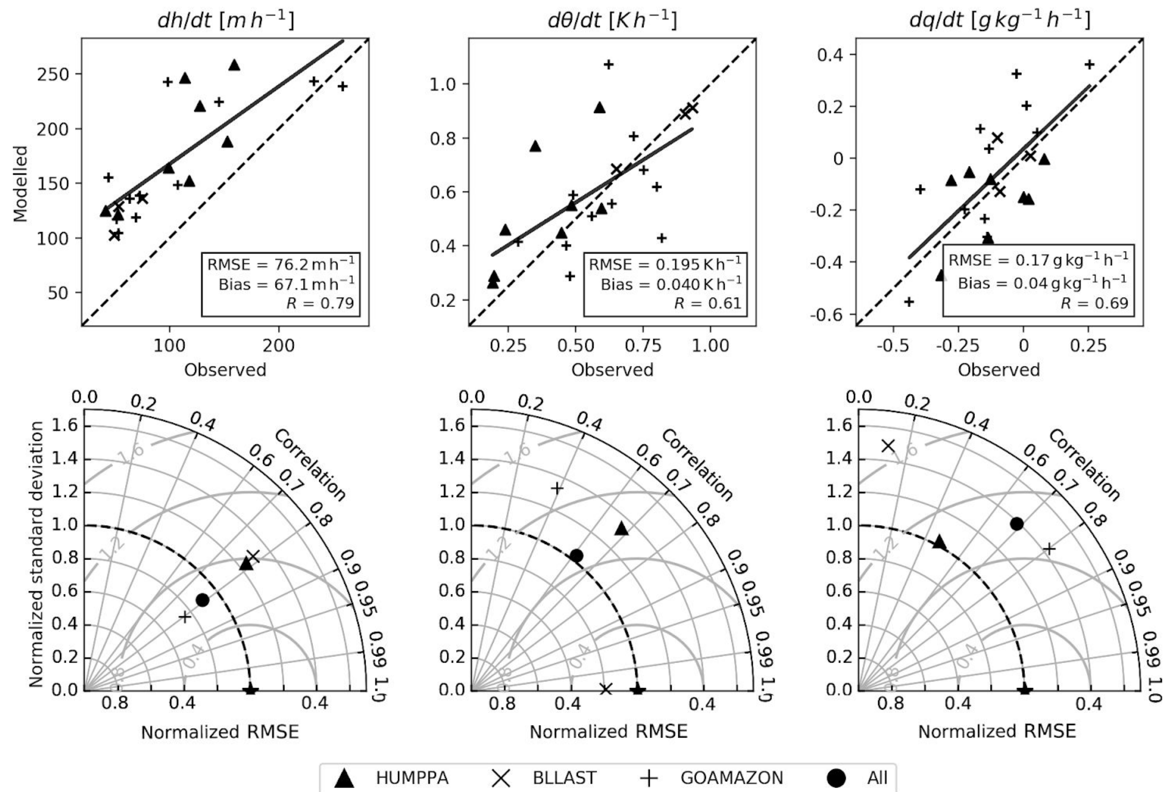
*“*It should be noted that many sites have only a few days with sounding pairs that meet the selection criteria, and most often, there are only intermittent time series available. For applications that require continuous datasets, the extraction profiles from reanalysis data, satellite-based products, or even Earth System Models, is also implemented in CLASS4GL. This alternative to the use of sounding data holds great promise for spatially-explicit climatological ABL studies and multi-annual trend assessments. **ABL model simulations using continuous/gap-filled datasets may deviate from those using the observations. Hence, an additional validation is needed as soon as such datasets are employed, in which one should compare the gap-filled datasets and the observations and the error propagation on the ABL model simulations**. Such an in-depth evaluation against the available sounding pairs can be done using the ~~default~~ present framework~~ based on the balloon sounding data~~.” P10R9-16

P10 L8: Change "reassure" to "assure".

Done. P11R7

Figure 3: There is only a circle (All) for dq/dt - not the other rates. Is something missing?

They were indeed missing. All symbols will appear in the revised manuscript, as shown in the figure below (P15):



P12 L6: Here I start wondering about the sources of biases and if you have been able to examine them. For dq/dt, a positive evaporation bias, excessive low-level moisture flux convergence (in the boundary conditions) or too little entrainment of dry air could each explain this. Has it been investigated? Is it likely a problem with the model or forcing data?

It is true that biases are found in the model, which need to be tackled during the continuous development of CLASS4GL. We also expect that the bias have multiple origins, so both the ABL model (physical concepts) and its forcing data (convergence/advection, evaporation bias, cloud cover...) but also model tuning parameters (eg., entrainment ratio) and errors in the sounding observations used to initialize and validate the model, and all of these possible errors should be investigated in the further development of CLASS4GL. We will make this clear by adding the following text in the revised manuscript as follows:

"**The bias is expected to have multiple origins, including the ABL model and its physical concepts, the forcing data (convergence/advection, evaporation bias, cloud cover…, see Table 2), model tuning parameters (such as the entrainment ratio, see Table 2) and errors in the sounding observations used to initialize and validate the model. All these possible**

error sources should be investigated in further development of CLASS4GL." P15R8-P16R1

The model performance statistics could already be when using one-hourly values (ERA5) instead of daily values (GLEAM) for evaporative fraction for the campaigns (see figure above) and the global results (see figure below, which is discussed further below). Particularly, the bias of *dq/dt* for the campaigns is now 0.04 g kg^{-1} h^{-1} (see figure above), whereas it was much higher in the previous results 0.17 g kg^{-1} h^{-1}. The clear change in performance statistics by changing the *EF* supports that the sources of biases are partly originating from errors propagating from the forcing data.

Since ERA5 provides better performance statistics, this is now used for *EF* in the results section, which is mentioned in the manuscript:

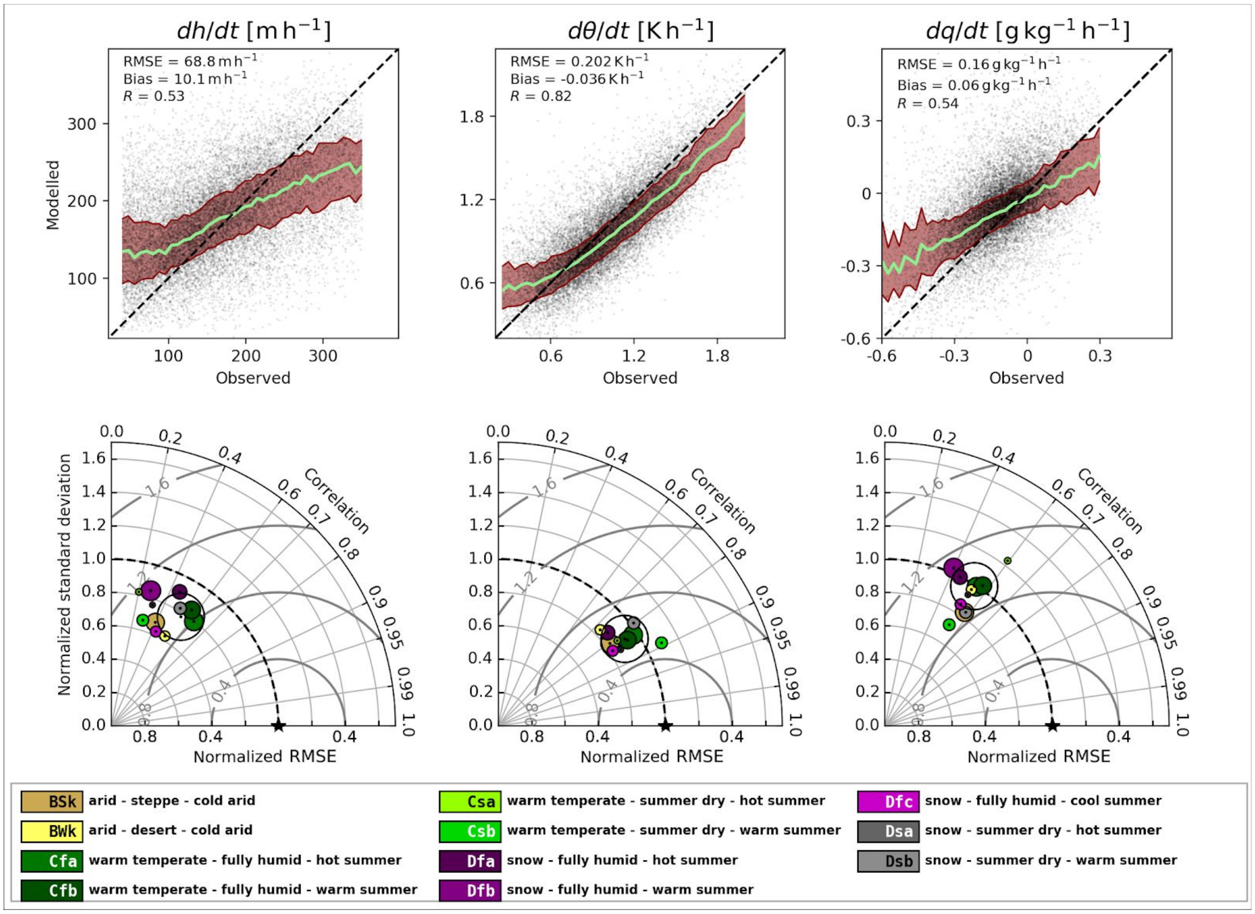
"However for the results presented in Section 3, one-hourly values of *EF* from recently released reanalysis dataset ERA5 (Copernicus Climate Change Service (C3S), 2017) are used instead of daily values from GLEAM, which provided a higher temporal resolution and better performance statistics." P11R11-13

The new results are also included in the revised manuscript. Using the updated results as shown with the figures above, we have modified the text accordingly:

"However, we could identify common model limitations over the three campaigns, see Fig. 3. ~~,~~ ~~particularly an overall positive bias in dq/dt, and a too low variability in the tendencies~~. **This includes an overall (slight) positive bias in dh/dt (dθ/dt and dq/dt), and an under(over)estimation of its (their) variability** as indicated with **a** normalized standard deviation **different from**~~ of less than~~ 1 in the Taylor plots (Fig. 3)**.**" P15R5-R8

Fig 4 and associated text: If the heating and moistening rates are converted to J/kg/h by multiplying by C_p and lambda_v respectively, we get that the heating bias is -52 J/kg/h but the positive moistening bias is 175 J/kg/h.

Using the one-hourly ERA5 1-hourly data as forcing for evaporative fraction as discussed above, we could also obtain overall better model performance statistics for the IGRA global soundings (see figure below) than with GLEAM daily input: particularly, the global negative bias in dθ/dt has been reduced from -0.052K/h (-52 J kg^{-1} h^{-1}*)* to -0.036K h^{-1} (-36 J kg^{-1} h^{-1}), and the positive bias in dq/dt has been reduced from 0.07 g kg^{-1} h^{-1} (or 175 J kg^{-1} h^{-1}) to 0.06 g kg^{-1} h^{-1} (or 150 J kg^{-1} h^{-1}). Pearson correlation coefficients have slightly improved (0.54/0.79/0.52 -> 0.53/0.82/0.54), and all RMSEs are slightly lower as well (72.3 m h^{-1} / 0.211 K h^{-1} / 0.17 g kg^{-1} h^{-1} -> 68.8m h^{-1} / 0.202 K h^{-1} / 0.16 g kg^{-1} h^{-1}).



These results are now updated in the revised manuscript, and the text and discussions will be modified accordingly:

"The simulated diurnal ABL tendencies show a similar accuracy when evaluated against the global IGRA sounding archive, despite the fact that these operational sounding have a priori lower quality standards; this applies to all three ABL tendencies shown in Fig. 4. Here, the ~~bias is larger~~ **magnitude of the bias is of the same order** for the change in θ (~~-0.052~~ **-0.036** K h^{-1}) ~~than~~ **as** for the intense observation campaigns, ~~whereas the changes~~ **but is now negative. The biases** are smaller for *h* (~~0.2~~ 10.1 m h^{-1}) compared to the campaigns and slightly larger for q (~~0.05~~**0.06** g kg^{-1}). As for the research campaigns, the model is able to reproduce the variability among the different operational sounding days, with Pearson correlation coefficients of ~~0.54, 0.79 and 0.52~~ **0.53, 0.82 and 0.54** for the diurnal tendencies of h, θ and q, respectively." P16R6-12

The discrepancy is 123 J/kg/h – ...

With the updated results mentioned above, the discrepancy has now been slightly reduced from 123 J kg^{-1} h^{-1} to 114 J kg^{-1} h^{-1}.

… again there could be multiple sources of this. First thought is net radiation, but excessive ground heat flux from the soil, advection (convergence) or entrainment could all be reasons. Any idea about the source of this net energy bias?

P13 L25-29: Related to above, a nice speculation on causes, but atmospheric models including reanalyses tend to have too much surface net radiation due to cloud errors and lack or proper representation of aerosol effects. R_Net or the input ERA-I radiation should be validated against independent data (e.g., the available CERES data) as a sanity check.

In line with the discussion above, we expect that the source of the net energy bias have multiple origins, including the ABL model and its physical concepts, the forcing data, model tuning parameters and errors in the sounding observations. We agree with the suggestions of the referee that this could be especially due to biases in the net radiation (which is calculated by the model by prescribing the cloud cover), the ground heat flux, the entrainment rates, and/or the prescribed advection. All this will be discussed in the revised manuscript as follows:

"**The negative bias in the temperature tendency and the positive bias in the humidity tendency lead to an overall net heat bias of 114 J kg^{-1} h^{-1}. Similar as for the results in the campaigns, it is expected that such global biases have multiple origins, including biases in the net radiation (which is calculated by the model by prescribing the cloud cover), underestimation of ground heat storage to the soil, the entrainment rates, and/or the prescribed advection. Further research should investigate possible errors related to input datasets and validate them against independent data (e.g., the available CERES data could be used to evaluate the net radiation).**" P16R17-22

Revised manuscript with track changes starts below.

Atmospheric boundary layer dynamics from balloon soundings worldwide: CLASS4GL v1.0

Hendrik Wouters[1], Irina Y. Petrova[1], Chiel C. van Heerwaarden[2], Jordi Vilà-Guerau de Arellano[2], Adriaan J. Teuling[2], Vicky Meulenberg[2], Joseph A. Santanello[3], and Diego G. Miralles[1]

[1]Ghent University, Laboratory of Hydrology and Water Management, Coupure Links 653, 9000 Ghent, Belgium
[2]Wageningen University and Research, Dept. of Environmental sciences, Postbox 47, 6700AA Wageningen, The Netherlands
[3]National Aeronautics and Space Administration, Goddard Space Flight Center, Hydrological Sciences Laboratory (617), Greenbelt, MD, USA

Correspondence: Hendrik Wouters (hendrik.wouters@ugent.be)

Abstract. The coupling between soil, vegetation and atmosphere is thought to be crucial in the development and intensification of weather extremes, especially meteorological droughts, heatwaves and severe storms. Therefore, understanding the evolution of the atmospheric boundary layer (ABL) and the role of land–atmosphere feedbacks is necessary for earlier warnings, better climate projection and timely societal adaptation. However, this understanding is hampered by the difficulties to attribute cause–effect relationships from complex coupled models, and the irregular space–time distribution of *in situ* observations of the land–atmosphere system. As such, there is a need for simple deterministic appraisals that systematically discriminate land–atmosphere interactions from observed weather phenomena over large domains and climatological time spans. Here, we present a new interactive data platform to study the behaviour of the ABL and land–atmosphere interactions based on worldwide weather balloon soundings and an ABL model. This software tool – referred to as CLASS4GL (http://class4gl.eu) – is developed with the objectives to (a) mine appropriate global observational data from ~~over 2~~ $\sim$ 15 million weather balloon soundings since 1981 and combine them with satellite and reanalysis data, and (b) constrain and initialize a numerical model of the daytime evolution of the ABL that serves as a tool to interpret these observations mechanistically and deterministically. As a result, it fully ~~automises~~ automizes extensive global model experiments to assess the effects of land and atmospheric conditions on the ABL evolution as observed in different climate regions around the world. The suitability of the set of observations, model formulations and global parameters employed by CLASS4GL is extensively validated. In most cases, the framework is able to realistically reproduce the observed daytime response of the ~~ABL~~ mixed-layer height, potential temperature and specific humidity from the balloon soundings. In this extensive global validation exercise, a bias of ~~0.2~~ 10.1 m h^{-1}, ~~-0.052~~ -0.036 K h^{-1} and ~~0.07~~ 0.06 g kg^{-1} h^{-1} is found for the morning-to-afternoon evolution of the ~~ABL~~ mixed-layer height, potential temperature and specific humidity. The virtual tool is in continuous development, and aims to foster a better process-understanding of the drivers of the ABL evolution and their global distribution, particularly during the onset and amplification of weather extremes. Finally, it can also be used to scrutinize the representation of land–atmosphere feedbacks and ABL dynamics in Earth system models, numerical weather prediction models, atmospheric reanalysis, and satellite retrievals, with the ultimate goal to improve local climate projections, provide earlier warning of extreme weather, and foster a more effective development of climate adaptation strategies. The tool can be easily downloaded via http://class4gl.eu and is open source.

1 Introduction

Climate and weather phenomena are largely influenced by land surface processes and the characteristics of the landscape. The interactions between soil, vegetation and atmosphere are thought to be particularly important for the evolution of extreme weather events such as droughts, heatwaves and convective thunderstorms (Santanello et al., 2018; Miralles et al., 2018; Seneviratne et al., 2010). The quantification of the drivers behind the extreme events is challenging, yet an understanding of the physical mechanisms underlying these events is highly relevant for earlier societal warning, better climate projection and timely adaptation (Sillmann et al., 2017). First efforts to quantify the relevance of land–atmosphere feedbacks date back to the late 20th century (e.g., Ek and Mahrt, 1994; Betts and Ball, 1995). However, substantial advancements have occurred in recent years after climate modelling initiatives such as the Global Land–Atmosphere Coupling Experiment (GLACE Koster et al., 2006; Guo et al., 2006; Berg et al., 2015), and observation-based studies under the umbrella of the Global Energy and Water Exchanges (GEWEX) Local Land–Atmosphere Coupling (LoCo) project (Roundy et al., 2013; Santanello et al., 2015; Tawfik et al., 2015; Santanello et al., 2018). Recent studies have highlighted the importance of soil moisture and evaporation for the occurrence of afternoon rainstorms (Findell et al., 2011; Taylor et al., 2012; Guillod et al., 2015; Petrova et al., 2018), droughts (Roundy and Santanello, 2017; Teuling et al., 2013) and extreme heat events (Fischer et al., 2007; Miralles et al., 2014). Moreover, studies on anthropogenic land-cover change and land management — such as deforestation (Akkermans et al., 2013; Lejeune et al., 2014), irrigation (Thiery et al., 2017; Lawston et al., 2015), modified croplands (Seneviratne et al., 2010) and urban expansion (Wouters et al., 2017) — have demonstrated profound influences of land conditions on local and regional climate, and specifically on the occurrence of extreme weather.

However, assessing cause–effect relationships in observational and model studies of land–atmosphere interactions remains complex, given the cross-correlation of multiple climate variables without the need of implying causation, the bidirectional interactions within the system, the various scales of variability and autocorrelation of different elements, and the unavoidable confounding effect of unobserved causal variables (Miralles et al., 2018). Likewise, the many studies of land–atmosphere interactions based on the use of global or regional climate models are model dependent and only poorly constrained by observations ~~(Orlowsky and Seneviratne, 2010)~~(Orlowsky and Seneviratne, 2010; Davin et al., 2019), and the complexity of the Earth Systems Models hampers the assessment of individual feedback processes. An intermediate compromise between statistical analysis of observations and complex climate model simulations could close this gap in process-understanding. For example, mechanistic studies based on simpler models supported by observations have yielded new insights recently (Roundy et al., 2013; Zaitchik et al., 2013; Santanello et al., 2009). Particularly, atmospheric boundary layer (ABL) models have been initialized and constrained with observations to simulate the atmospheric response to land surface conditions and the state of the free atmosphere; this way, the influence of turbulent heat fluxes, incoming radiation, subsidence, advection, or entrainment can be easily quantified within certain ranges of uncertainty (Pietersen et al., 2015; Miralles et al., 2014; Ouwersloot et al., 2012; van Heerwaarden et al., 2009). For example, ABL models have been applied to investigate soil moisture and vegetation feedbacks during combined droughts and heatwaves in Europe (Miralles et al., 2014), the different feedbacks on heatwave evolution over forests and grasslands (van Heerwaarden and Teuling, 2014), and the suppression of clouds by plants in a CO_2-rich atmosphere

(Vilà-Guerau de Arellano et al., 2012). The advantage of using these ABL bulk models is twofold: (a) unlike climate models, they can be routinely initialized and constrained by observations and easily interpretable in terms of the interaction between variables; (b) unlike merely statistical analysis of observational data, they provide unambiguous understanding on the deterministic links among the variables in the system. Yet, these mechanistic models require detailed observations describing the entire state or evolution of the soil, vegetation and atmosphere. Given this dependency, the previously-mentioned process-based studies usually focus on one particular location only, for one or a few diurnal cycles at best. A generalization of the mechanistic outcomes of these process-based case studies to other climate regions or periods of extreme events may not always be justified. By all means, the atmospheric and land surface observations are irregular in time and space, which makes it very challenging to attribute the causes of meteorological variability.

Here, an open source interactive data platform is presented based on the application of the Chemistry Land-surface Atmosphere Soil Slab model (CLASS; Vilà-Guerau de Arellano et al., 2015; van Heerwaarden et al., 2010) to balloon soundings worldwide. The platform, hereafter referred to as the CLASS4GL (CLASS model for GLobal studies), is designed to mine observations from the radio soundings, satellite remote sensing observations, reanalysis data and surface data inventories to constrain and initialize the ABL model (CLASS). It ~~automises~~ automizes mass parallel simulations of the ABL and enables global sensitivity experiments. As a result, it is designed for studies that aim to foster a better understanding of the dynamics of ABL and the development of extreme weather, and allows the attribution of changes in the state of the ABL to specific land and atmospheric conditions. A core goal of this study is to present this interactive data platform, including a summary description of the ABL model and the data mining procedure used to initialize and constrain model simulations (Sect. 2). Furthermore, the skill of the modelling framework to reproduce the daytime evolution of the ABL is evaluated against worldwide observations from specific field campaigns as well as operational balloon soundings (Sect. 3). Finally, a perspective is provided in which the potential of this framework to contribute to a better understanding of land–atmosphere interactions over different climates is discussed (Sect. 4).

2 Data and methods

The CLASS4GL platform is composed of three modules (see Fig. 1), namely an ABL model (CLASS), a data mining module, and an interface module. The ABL model is used simulate the ABL evolution and is described in detail under Sect. 2.1. It requires appropriate observations of the ABL for the initialization in the morning and for the validation in the afternoon. Meanwhile, the data mining module collects profile observations from soundings taken during research campaigns and operational activities since 1981. The intensive research campaigns offer continuous high-quality sounding profiles available for specific time periods and locations, while the operational weather balloon soundings offer more regular balloon launches at a vast amount of locations around the world, but with intermittent and varying quality. The profile database is extensive, yet spatially and temporally sparse. Quality check tools have been applied to mine the profiles that are appropriate for the ABL model in the way described in Sect. 2.2. In order to further constrain and initialize ABL model simulations with surface conditions and larger-scale atmospheric variables, the data mining module also employs ancillary data from satellite remote sensing, reanalysis and survey inventories (Sect. 2.3). File formats are NetCDF and YaML to enable the easy adoption of any (upcoming) input datasets and the exchange of profiles, parameters and model experiments among users and a central database. Finally, the interface module provides the ability to easily perform multiple simulations of the ~~dirunal~~ diurnal ABL evolution in parallel, as well as batches of sensitivity experiments. Therefore, it enables the parallelization of multiple model simulations and offers multi-processing support for both regular computer workstations as well as super-computing infrastructure. The interface module also implements a range of tools for pre-and post-processing the sparse data pool of inputs and experiments, and a data explorer with a graphical user interface. As a result, CLASS4GL ~~automises~~ automizes mechanistic assessments of the observed diurnal ABL behaviour around the globe, and allows the exploration of local land–atmosphere feedbacks and the attribution of cause–effect relationships. A detailed description of the platform is provided in the next subsections. CLASS4GL is provided as an open source Python library, it is conveyed under the GNU General Public License version 3 (GPLv3), and it can be easily downloaded via http://class4gl.eu.

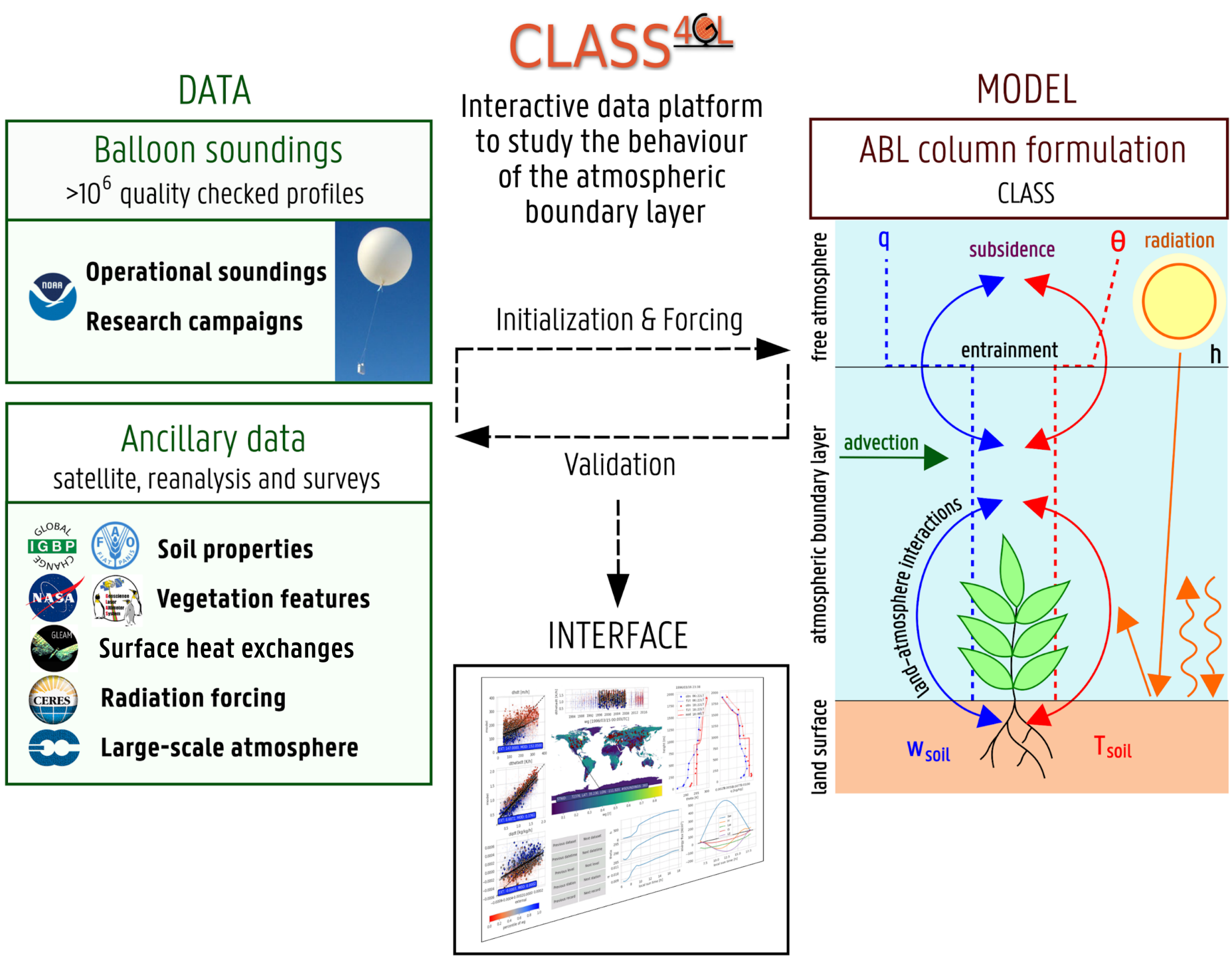


Figure 1. Schematic overview of the Chemistry Land-surface Atmosphere Soil Slab model for GLobal studies (CLASS4GL)

2.1 ABL model: CLASS

The Chemistry Land-surface Atmosphere Soil Slab model (CLASS) is a conceptual bulk model that uses a small set of differential equations to reproduce the evolution of the ABL's essential properties over diurnal time spans in response to surface and atmospheric forcings and feedbacks. The reason of choosing a bulk model is the low computational cost and the easy interpretation of the canonical ABL properties in view of extensive global experiments. However, we note that other models that allow for a more complex and explicit representation of the turbulent processes in the ABL – such as single column ABL models or large-eddy simulation models – may also be implemented.

CLASS is based on the original work by Tennekes (1973) and uses the thermodynamic equations of the ABL proposed by Tennekes and Driedonks (1981). The ABL is represented as a single model layer. The use of the mixed-layer equations implies that the turbulence ~~inside the ABL is not explicitly solved, and assumes that~~ and vertical gradients inside the mixed layer are not explicitly resolved, and the potential temperature (θ), specific humidity (q) and wind components are assumed to be homogeneous within the ~~ABL~~mixed layer. This assumption tends to be supported by the efficient turbulent mixing under convective conditions (Bauer, 1908). At the top of the ~~ABL~~mixed layer, the entrainment of heat and moisture is parameterized by a jump of θ, q and wind components over an infinitesimally small height, which ~~have~~ are initialized with a constant lapse rate with height in the overlying free atmosphere. Entrainment flux is calculated as a fixed fraction (0.2) of the buoyancy flux, ~~for~~ to which one also adds the entrainment flux driven by shear. An important feature of the model is the possibility to represent the subsidence coupled to the entrainment process at the inversion zone (Vilà-Guerau de Arellano et al., 2015). Herein, the subsidence velocity is a function of the divergence of the mean horizontal wind and the evolving ~~ABL~~ mixed-layer height. The surface–atmosphere turbulent exchanges for momentum, heat and moisture in the surface layer are calculated considering their aerodynamic resistances. These are calculated in an iterative way assuming constant values for aerodynamic roughness lengths, while applying correction factors for non-neutral stratification of the atmospheric surface-layer (Paulson, 1970) according to the Monin–Obukhov similarity theory (Monin and Obukhov, 1954). It should be kept in mind that more realistic profiles with explicit ABL gradients for temperature, humidity and wind speed – especially at the top (entrainment zone) and bottom (surface layer) of the mixed layer – are not yet considered by the model. In order to tackle these limitations and associated uncertainties, more research is needed employing more realistic profiles. The surface energy balance at the land surface is solved using the Penman–Monteith equation (Monteith, 1965), and the heat and moisture transport in the soil is described using a two-layer force–restore model (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) employing empirical relations for soil hydraulic properties of (Clapp and Hornberger, 1978). The amount of net radiation available for the sensible, latent and ground heat flux is calculated from the albedo, emissivity the incoming shortwave and long-wave radiation, the surface temperature and the cloud cover. ~~The surface-atmosphere turbulent exchanges for momentum, heat, and moisture are calculated considering their aerodynamic resistances. These are calculated in an iterative way assuming constant values for aerodynamic roughness lengths, while applying correction factors for non-neutral stratification of the atmospheric surface-layer (Paulson, 1970) according to the Monin–Obukhov similarity theory (Monin and Obukhov, 1954).~~ For the vegetation fraction, the transpiration response of vegetation to the atmospheric conditions considers the empirical stomatal resistance from Jarvis (1976) inversely proportional to the leaf area index and the plant water stress — assumed to be a linear function of the soil-moisture deficit of the deep soil layer — and also a function of the incoming shortwave radiation, the vapor pressure deficit and the air temperature. An analogous formulation is used for the resistance of the bare-soil fraction, but considering it only as an inverse linear function to the soil-moisture deficit of the shallow soil layer. More details and explicit formulations of the coupled land–atmosphere system CLASS can be found in Vilà-Guerau de Arellano et al. (2015), van Heerwaarden and Teuling (2014) and van Heerwaarden et al. (2010).

In order to exploit the full database in CLASS4GL, the ABL model is upgraded with the following features:

- A representation of advection as an additional atmospheric dynamic forcing (see below).
- A representation of the upper-air atmospheric profile that also evolves according to the external large-scale dynamic forcing of advection and subsidence, and as such accounts for varying (instead of constant) lapse rates of the capping inversion during the growth of the ~~ABL~~mixed layer.
- An implementation of alternative surface-layer transfer coefficients for momentum and heat Wouters et al. (2012), which account for additional non-neutral stability correction factors for the roughness sublayer (Ridder, 2009). The procedure uses a non-iterative approximation of the transcendental relation between the transfer coefficients and the Richardson bulk number, hence preserves numerical stability and decreases the computational cost of CLASS.
- An iterative procedure to invert soil moisture conditions based on the match of simulated evaporative fraction (*EF*) to satellite-based *EF* estimates. Details on this procedure can be found under Sect. 2.3 (eq. 3).

To conclude this section, we elaborate the above-mentioned CLASS model extension to consider advection as additional ABL dynamic forcing. It is assumed that both the mean vertical wind speed and also horizontal turbulent fluxes in the ~~ABL~~ mixed layer are negligible, hence set to zero. As such, the laws of conservation of momentum, energy, and atmospheric constituents in the ~~ABL~~ mixed layer are given as follows:

$$\frac{\partial\overline{\psi}}{\partial t} + \overline{u}\frac{\partial\overline{\psi}}{\partial x} + \overline{v}\frac{\partial\overline{\psi}}{\partial y} + \frac{\partial\overline{w'\psi'}}{\partial z} = \overline{S}_\psi \qquad (1)$$

where ψ is a generic variable (θ, q, u, v), u, v, w are the wind fields, the overbars indicate the Reynolds averages, ψ' indicates the fluctuation around the averages (note the difference from eq. 2.5 of Vilà-Guerau de Arellano et al. (2015) because of the horizontal advection terms), and $\overline{S}_\psi$ represents the sum of the external forcings, sources and sinks of ψ. Integrating the last equation over the whole ~~ABL~~ mixed-layer column in both the x, y and z direction yields the tendency of the bulk mixed-layer quantity (assuming no forcings, sources or sinks in the ~~ABL~~mixed layer):

$$\frac{\partial\langle\psi\rangle}{\partial t} = \frac{1}{h}\left[\langle(w'\psi')_s\rangle - \langle(w'\psi')_e\rangle\right] - \langle\overline{\mathbf{U}}\cdot\nabla_{\mathrm{hor}}\overline{\psi}\rangle \qquad (2)$$

where $(w'\psi')_s$ is the vertical turbulent flux at the surface, and $(w'\psi')_e$ is the entrainment flux at the ~~mixed layer~~ mixed-layer height (h), l is the horizontal extent of the single column, $\langle\rangle$ indicates the bulk (slab or mixed-layer) mean value of any field over the entire ~~ABL~~ mixed-layer column, and ~~$-\overline{\mathbf{U}}\cdot\nabla_{\mathrm{hor}}\overline{\psi}$~~ $\langle-\overline{\mathbf{U}}\cdot\nabla_{\mathrm{hor}}\overline{\psi}\rangle$ is the bulk horizontal advection for $\overline{\psi}$ in the ~~ABL~~mixed layer, which is obtained from external forcing fields.

2.2 Automated Automized balloon data mining

Global data of weather balloon soundings are taken from the Integrated Global Radiosonde Archive (IGRA; Durre et al., 2006) which is maintained under the auspices of the National Oceanic and Atmospheric Administration (NOAA). The IGRA archive is routinely updated and currently includes more than 2700 stations covering major global climate regions. The CLASS4GL sounding database is additionally supplemented with data from intensive radiosonde campaigns from HUMPPA (Williams et al., 2011), BLLAST (Pietersen et al., 2015) and GOAMAZON (Martin et al., 2016) – – see Tab. 1 and Fig. 2. Other sources of vertical profile data (from e.g., aircraft, satellites, other observation campaigns or long-term operational soundings) may be considered in future applications of the framework. As described above, CLASS requires morning sounding profiles for initialization and afternoon profiles for validation to enable a mechanistic interpretation of the diurnal ABL evolution.

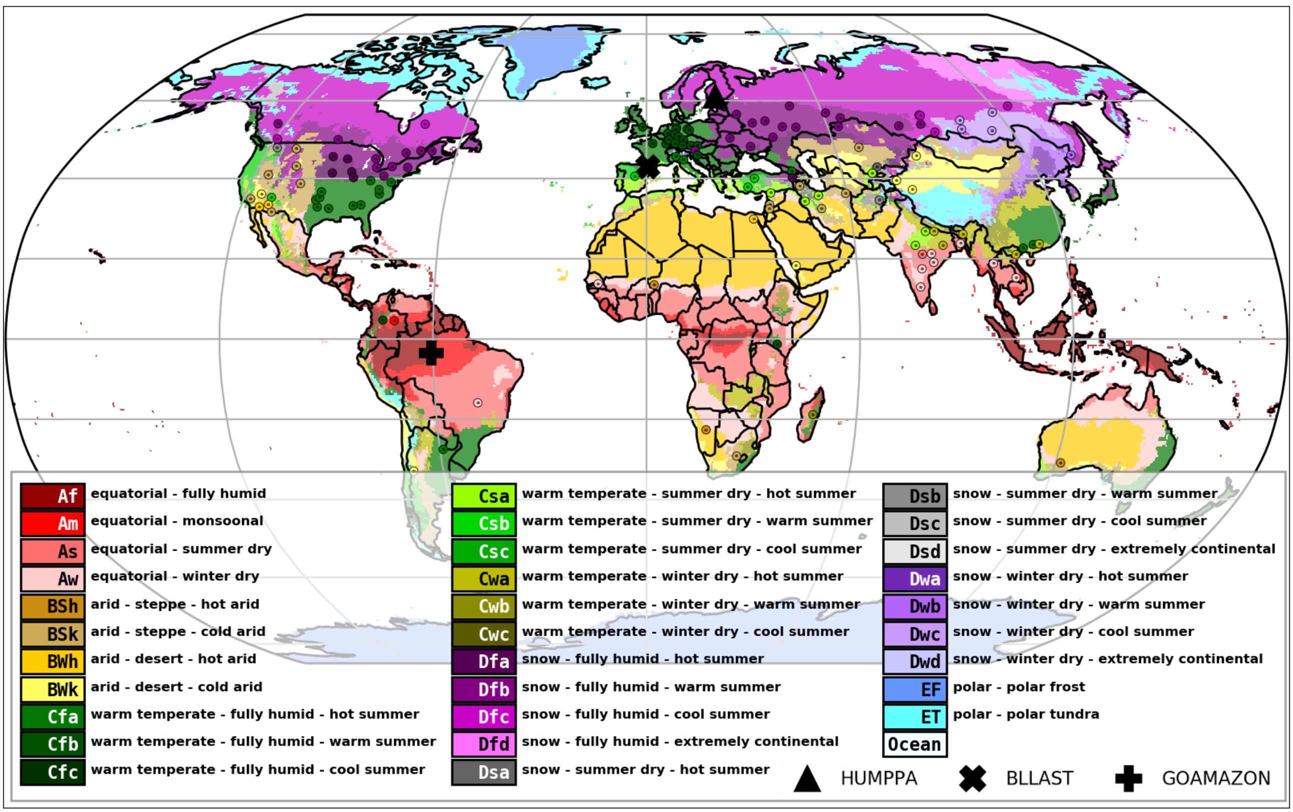


Figure 2. Distribution of the launch sites from the Integrated Global Radiosonde Archive (IGRA) retained after the profile quality selection procedure. The different climate classes are indicated with the colors according to the Köppen–Geiger climate classification. The markers indicate the locations of the three observation campaigns (i.e., HUMPPA, BLLAST and GOAMAZON). The profile quality selection is described in presented in Sect. 2.2.

All balloon sounding profiles are (∼15 million profiles) are pre-processed and only those that meet a series of selection criteria are retained. First, only days with a pair of morning and afternoon sounding profilesare considered. Second, first by calculating the bulk mixed-layer propertiesare estimated and only those soundings with a well-mixed profile are retained. These two pre-processing steps are automated in CLASS4GL, and are explained in more detail below. It should be noted that these criteria can be flexible and may be reconsidered according to the intended application, since there is an obvious trade-off between sounding quality and amount of data being retained. Overall, the selection procedure used here keeps only about 2% of the available soundings, i.e., a total of just over 42,000 profiles (from 150 sites) out of the initial 2 million (from 2,700 sites) in the global IGRA archive. For the three intensive observation campaigns, 16 out of 49 profiles are retained. An overview of the global distribution of the retained radiosondes and their corresponding climate regime is shown in Fig. 2. It should be noted that many sites have only a few says with sounding pairs that meet the selection criteria, and most often, there are only intermittent time series available. For applications that require continuous datasets, the extraction profiles from reanalysis data, satellite-based products, or even Earth System Models, is also implemented in CLASS4GL. This alternative to the use of sounding data holds great promise for spatially-explicit climatological ABL studies and multi-annual trend assessments. An in-depth evaluation against the available sounding pairs can be done using the default framework based on ballooon sounding data: The mixed-layer height (h) is assessed as the height at which the Bulk Richardson number (RiB) exceeds a critical value (RiB_c). We adopt the estimates for RiB_c provided by Zhang et al. (2014): $RiB_c = 0.24$ for strongly stable boundary layers, $RiB_c = 0.31$ for weakly stable boundary layers, and $RiB_c = 0.39$ for unstable boundary layers. The uncertainty range of h (used below) is determined from its interval corresponding to the RiB_c range $[0.24, 0.39]$, for which the interval is further extended to the nearest sounding records above and below. Second, the mixed-layer potential temperature (θ), specific humidity (q), zonal wind (u) and meridional wind (v) are calculated as their average values recorded within the mixed layer. The capping inversion is estimated by a linear extrapolation of the two lowest sounding measurements above h, for which its lapse rate for potential temperature ($\gamma_\theta = d\theta/dz$), specific humidity ($\gamma_q = dq/dz$) and wind components ($\gamma_u = du/dz$ and $\gamma_v = dv/dz$) are calculated. The jump values at the h for potential temperature ($\Delta\theta$), specific humidity (Δq) and wind components (Δu and Δv) are estimated from the difference between the values of the capping inversion at h and the values within the mixed layer.

Distribution of the IGRA sites retained after the profile quality selection procedure. The different climate classes are indicated with the colors according to the Köppen–Geiger climate classification. The markers indicate the locations of the three observation campaigns (i.e., HUMPPA, BLLAST and GOAMAZON). The profile quality selection is described in presented in Sect. 2.2.

The specific selection criteriaused here are described in more detail in the following paragraphs.

(I) As mentioned above, CLASS requires morning soundings to be initialised and afternoon profiles to be validatedAfterwards, morning–afternoon profiles are selected that meet a series of selection criteria: the morning profiles, ie. profiles before 12 h local time, are selected first, and they amount to ∼6 million profiles. Here, the selection of suitable balloon sounding (morning–afternoon) pairs is largely morning soundings (and the subsequent afternoon soundings after 12 h is based on the timing of these soundings . Morning and afternoon (a): Morning (and afternoon) sounding profiles ideally should be acquired after sunrise and before sunset, respectively. However, routine sounding launches happen synchronously on a daily basis at 0 h and 12 h UTC, whereas launches at intermediate timings (3 h, 6 h, 9 h, 15 h and 18 h UTC) are rare. As a result, many launches, especially those at 0h 0 h UTC in Europe and Africa, often happen several hours before sunrise. Since the net exchanges near the surface for heat, moisture and radiation are generally low at the end of the night, the atmospheric profiles tend not to change dramatically before sunrise (unless the synoptic situation changes), being often representative for the time the ~~convective ABL~~ mixed layer starts to emerge (van Stratum and Stevens, 2018). As such, in order to maintain a high number of soundings in our analyses, launch times within 3 h prior to sunrise are still allowed here. For these soundings, the ABL simulation starts at sunrise, assuming that the change in the atmospheric profile since the balloon launch time is negligible. ~~Finally, the afternoon radiosonde profile on the same day needs to occur between local noon and 1 h before sunset, and at least 4 h after the model initialization in the morning.~~

~~(II) An estimation of ABL properties is obtained for the selected profile pairs. First,~~ Furthermore, (b) only those soundings are retained with more than seven measurements in the vertical below 3000 m (72% of the morning soundings), (c) for which the uncertainty of the mixed-layer height ~~(h) is assessed as the height at which the Bulk Richardson number (RiB) exceeds a critical value ($RiBc$). We adopt the estimates for RiBc provided by Zhang et al. (2014) : $RiBc = 0.24$ for strongly stable boundary layers, $RiBc = 0.31$ for weakly stable boundary layers, and $RiBc = 0.39$ for unstable boundary layers. The uncertainty range of h (used below)~~ is ~~determined from its interval corresponding to the RiBc range $[0.24, 0.39]$,~~ lower than 150 m (26% of the morning soundings), (d) for which a well-mixed layer is observed (ie. for which the ~~interval is further extended to the nearest sounding records above and below. Second, the~~ root-mean square deviation of the temperature from the estimated mixed-layer ~~potential temperature (θ), specific humidity (q) , zonal wind (u)and meridional wind (v) are calculated as their average values recorded within~~ average is lower than 1.5°C; this criterion is met by 92% of the morning soundings). We also (e) set the morning lower temperature limit to 278 K in order to minimize the chance of freezing temperatures during the course of the simulations (this criterion is met by 70% of the morning soundings). The next criterion is that (f) an afternoon sounding can be found with the same criteria as the morning sounding except regarding the uncertainty of the mixed-layer ~~. The capping inversion is estimated by a linear extrapolation of the two lowest sounding measurements above h, for which its lapse rate for potential temperature ($\gamma_\theta = d\theta/dz$), specific humidity ($\gamma_q = dq/dz$) and wind components ($\gamma_u = du/dz$ and $\gamma_v = dv/dz$)are calculated. The jump values at the h for potential temperature ($\Delta\theta$), specific humidity (Δq) and wind components (Δu and Δu) are estimated from the difference between~~ height (which is met by 24% of the filtered morning soundings). Here, the ~~values of the capping inversion at h and the values within the mixed-layer. Finally, only those days in which the morning profile meets these criteria are retained: (a) the number of observations in the first 3000 m above ground level should be at least 7, and (b) the h uncertainty (specified above) needs to be below 150 m. For the afternoonsoundings we also require that the number observations below 3000 m is higher or equal than 7. Last, we also ensure that the calculated ABL properties fall within the physical range expected by CLASS. As such, since ice and snow cannot be represented by the model, we set the morning lower temperature limit to 278 K in order to minimize the chance of freezing temperatures during the course of the simulations. Furthermore~~afternoon radiosonde profile on the same day needs to occur between local noon and 1 h before sunset (defined as the time when the incoming shortwave radiation at the top of the atmosphere becomes zero), and at least 4 h after the model initialization by the morning sounding so that a sufficiently large model time span is considered. In case there are more than two soundings retained during a particular day which especially occurs during the campaigns, the sounding closest to sunrise is taken for the initialization in the morning, and the latest sounding for the validation in the afternoon. so that a sufficiently large model time span is considered. We also require that (g) all external forcing parameters described in the next Section are available for the simulation (which is met by 8.7% of the filtered sounding pairs). The above criteria lead to 21,826 profile pairs from 134 stations. Finally, the current version of CLASS is only capable of representing growing mixed layers; here. Therefore, an observed mixed-layer growth of 40 m h^{-1} is considered as a lower limit (which is met by 85% of the profile pairs), which leads to 18,385 profile pairs from 121 stations. For the three intensive observation campaigns, 22 out of 49 profiles are retained. An overview of the global distribution of the retained radiosondes and their corresponding climate regime is shown in Fig. 2. These requirements are flexible and may be reconsidered according to the intended application, since there is an obvious trade-off between sounding quality and amount of data being retained.

It should be noted that many sites have only a few days with sounding pairs that meet the selection criteria, and most often, there are only intermittent time series available. For applications that require continuous datasets, the extraction profiles from reanalysis data, satellite-based products, or even Earth System Models, is also implemented in CLASS4GL. This alternative to the use of sounding data holds great promise for spatially-explicit climatological ABL studies and multi-annual trend assessments. ABL model simulations using continuous/gap-filled datasets may deviate from those using the observations. Hence, an additional validation is needed as soon as such datasets are employed, in which should compare the gap-filled datasets with the observations and the error propagation on the ABL model simulations. Such an in-depth evaluation against the available sounding pairs can be done using the present framework.

2.3 Gridded ancillary data

In addition to balloon observations, CLASS4GL uses gridded data of the land and large-scale atmospheric state to initialize and constrain ABL model simulations. These ancillary data aim at complementing the soundings and providing context regarding the land and atmospheric conditions for which the sounding measurements take place. In total, we use four satellite-based products, two survey datasets and one reanalysis to describe soil–vegetation conditions and large-scale atmospheric forcing(. All input datasets and parameters are listed respectively in Tab. 1). and 2, and details can be found in the next paragraphs.

Static datasets are used to describe initial soil properties and land cover parameters, such as the fraction of land covered by vegetation and the surface albedo. The latter are based on the global vegetation continuous fields product from the Moderate Resolution Imaging Spectroradiometer (MODIS MOD44B; Hansen et al., 2005). Wilting point, soil porosity, field capacity and critical soil moisture are derived from the database of Global Gridded Surfaces of Selected Soil Characteristics from the International Geosphere-Biosphere Programme (IGBP-DIS, 2000). The Harmonized World Soil Database (HWSD; Nachtergaele et al., 2009) is used to provide soil classes. From this class, the Clapp and Hornberger parameters and the thermal parameters for the force–restore method are obtained via the look-up table in Noilhan and Planton (1989). The vegetation canopy height is determined from the Geoscience Laser Altimeter System (GLAS; Simard et al., 2011). Dynamic data of Leaf-Area Index (LAI) are taken from the Global Inventory Modeling and Mapping Studies (GIMMS) (Liu et al., 2012).

Initial surface and root-zone soil moisture values are inverted using the Global Land Evaporation Amsterdam Model (GLEAM) version 3.2a (Miralles et al., 2011; Martens et al., 2017) as reference. In order to maximize the consistency between CLASS and GLEAM, the soil and vegetation ancillary data used by CLASS4GL, and described above, correspond in fact to those used by the GLEAM v3.2a. It should be noted that CLASS and GLEAM have similar but not identical surface-vegetation-atmosphere transfer schemes, hence equivalent soil moisture levels may be associated to a differing *EF*. In order to minimize potential incompatibilities, CLASS4GL inverts soil moisture values by iteratively converging to the *EF* from GLEAM, instead of using the GLEAM root-zone soil moisture directly to ~~initialise~~ initialize the ABL simulations. This iterative procedure is based on finding the zero of the following function:

$$f(w) = EF_{\mathrm{CLASS}}(w) - EF_{\mathrm{GLEAM}} \tag{3}$$

where a unique soil moisture value (w) is considered for the entire root zone in CLASS. In order to ~~reassure~~ assure convergence and reduce the number of iteration steps, two zero finding algorithms have been combined, namely the discrete midpoint method and the bisection method (Burden et al., 2016). We note that this procedure is analogous to the one used by Miralles et al. (2014) in which CLASS was steered to match the Bowen ratios (instead of the *EF*) derived from GLEAM.

However for the results presented in Section 3, one-hourly values of *EF* from recently released reanalysis dataset ERA5 (Copernicus Climate Change Service (C3S), 2017) are used instead of daily values from GLEAM, which provided a higher temporal resolution and better performance statistics. Since gridded and consistent satellite datasets of cloud properties, advection, subsidence and radiation are not available for the long time span, we also make use of reanalysis data from ERA-Interim (Dee et al., 2011). ~~The more recent reanalysis dataset ERA5 is also available as an option at .~~ Moreover, reanalysis data are also used to ~~initialise~~ initialize the soil temperature in the morning. As an alternative to the specification of cloud cover, one could also directly specify the radiation inputs based on ERA-Interim or measurements from the Clouds and Earth's Radiant Energy System (CERES) onboard Terra and Aqua (Wielicki et al., 1996), which are available globally from the year 2001 onwards on a 1° regular grid. Unless specified differently, ~~these~~ the data are used in a dynamic manner in the CLASS simulations based on the temporal resolution specified in the Table 1. Since the lateral forcing for the calculation of advection is only available at a coarse spatial (0.75°) and temporal (~~6 hourly~~6-hourly) resolution, a footprint of 1° by 1° is taken for the ABL column model — centered over the sounding location — over which average parameters of the different datasets are calculated for the model input.

Finally, the CLASS4GL framework implements a flexible interface to any NetCDF files based on the Python/Xarray software libraries that a user may be interested in adding, hence alternative datasets can easily be adopted.

Data set	Variable	type	Horizontal resolution	Temporal resolution	Period	Reference
IGRA	Vertical profiles of potential temperature, wind and specific humidity	sounding	-	daily	1900 – 2018	Durre et al. (2006)
HUMPPA	*Idem*	sounding	-	3-hourly	summer 2010	Williams et al. (2011)
BLLAST	*Idem*	sounding	-	3-hourly	summer 2011	Pietersen et al. (2015)
GOAMAZON	*Idem*	sounding	-	3-hourly	2014 – 2015	Martin et al. (2016)
GLEAM or ERA5	Initial root-zone soil moisture and evaporative fraction	satellite / reanalysis	$0.25° \times 0.25°$	daily	1980 – 2017	Martens et al. (2017); Miralles et al. (20 Martens et al. (2017); Miralles et al. (20
IGBP-DIS	Soil hydrological properties (wilting point, water field capacity, saturated water content)	survey	$0.25° \times 0.25°$	static	-	IGBP-DIS (2000)
GIMSS (AVHRR)	Leaf Area Index	satellite	$0.25° \times 0.25°$	monthly	1981 – 2015	Liu et al. (2012)
HWSD	Clapp and Hornberger parameters, force–restore thermal parameters	survey	$0.085° \times 0.085°$	static	-	Nachtergaele et al. (2009)
MOD44B (MODIS)	Fractional vegetation and albedo	satellite	250m x 250m	yearly (here considered as static)	-	Hansen et al. (2005)
GLAS	Canopy height	satellite	$0.25° \times 0.25°$	static	-	Simard et al. (2011)
ERA-Interim	Atmospheric forcing (cloud cover, advection, subsidence), initial soil temperature	reanalysis	$0.75° \times 0.75°$	6-hourly	1979 – 2018	Dee et al. (2011)
CERES L3SYN1DEGERA-Interim or CERES	radiation components	reanalysis or satellite	$1° \times 1°$	6-hourly	2001 – 2015 (CERES)	Wielicki et al. (1996)

Table 1. Ancillary data sources of CLASS4GL.

Vegetation			
Symbol	*Name*	*Unit*	*Default value or source*
LAI	Leaf area index of vegetated surface fraction	[-]	GIMSS (daily)
$r_{c,min}$	Minimum resistance transpiration	[s m^{-1}]	110
$r_{s,soil,min}$	Minimum resistance soil evaporation	[s m^{-1}]	50
g_D	Vapour pressure deficit correction factor for surface resistance	[-]	0
h_{can}	Canopy height	[m]	GLAS (static)
$z_{0\mathrm{m}}$	Roughness length for momentum	[m]	0.1 x h_{can} (static)
$z_{0\mathrm{h}}$	Roughness length for heat and moisture	[m]	0.1 x $z_{0\mathrm{m}}$ (static)
α	Surface albedo	[-]	MOD44B (static)
T_s	Initial surface temperature	[K]	ERA-Interim (three-hourly)
$T_{soil,1}$,$T_{soil,2}$	Initial temperature of the top and deep soil layer	[K]	ERA-Interim (three-hourly)
ω_{sat}	Saturated volumetric water content	[m^3 m^{-3}]	IGBP-DIS (static)

Soil			
Symbol	*Name*	*Unit*	*Default value or source*
ω_{fc}	Volumetric water content field capacity	[m^3 m^{-3}]	IGBP-DIS (static)
ω_{wilt}	Volumetric water content wilting point	[m^3 m^{-3}]	IGBP-DIS (static)
EF	Evaporative fraction	[-]	ERA5 (hourly) or GLEAM (daily)
$\omega_{\mathrm{soil,1}}$,$\omega_{\mathrm{soil,2}}$	Volumetric water content top and deep soil layer	[m^3 m^{-3}]	By iterative matching of EF
c_{veg}	Vegetation fraction	[-]	MOD44B (static)
a,b,p	Clapp and Hornberger retention curve parameters	[-]	HWSD (static)
C_{Gsat}	Saturated soil conductivity for heat	[K m^{-2} J^{-1}]	HWSD (static)
$C_{2,\mathrm{sat}}$	Coefficient force term moisture	[-]	HWSD (static)
$C_{2,ref}$	Coefficient restore term moisture	[-]	HWSD (static)
Λ	Thermal diffusivity skin layer	[-]	5.9

Air			
Symbol	*Name*	*Unit*	*Default value or source*
β	Ratio between buoyancy virtual heat and entrainment virtual heat	[-]	0.2
γ_θ	Initial lapse rate of potential temperature in the free atmosphere	[K m^{-1}]	From profile (IGRA)
γ_q	Initial lapse rates of specific humidity in the free atmosphere	[kg kg^{-1} m^{-1}]	*Idem*
	Initial lapse rates of zonal and meridional components in the free	[m^{-1}]	*Idem*

3 Results and discussion

The skill to replicate the evolution of the ABL as observed by the radiosondes is evaluated against the pre-processed and quality-controlled balloon soundings from (a) the three intensive research campaigns, and (b) the global operational IGRA dataset. The evaluation is done by comparing the modelled daytime tendencies of the mixed-layer height (dh/dt), potential temperature ($d\theta/dt$) and specific humidity (dq/dt) against the corresponding tendencies observed from the balloon sounding pairs. Observed and modelled tendencies represent the mean diurnal change from the morning sounding to the afternoon sounding. It should be noted that the local time of the morning and afternoon soundings changes given that the launch times are often at 0 and 12 h UTC, and that the boundary-layer tendencies are not uniform over the course of the day. The resulting variety in the tendencies is taken into account in the simulations, since the model is initialized with the morning sounding while the initial solar local time in the model is set equal to the sounding launch. The same happens for the end of the simulation at the time of the afternoon sounding. Therefore, the expected tendency for each launch or site (depending on the local time window being considered in the computation of that tendency) is equivalent for observations and model simulations, hence any biases related to launching times between the two are avoided. The campaign observations provide an *a priori* higher standard than the operational balloons in terms of accuracy and resolution during the balloon ascent and in terms of daytime sampling. Hence, the evaluation against the 16 22 campaign soundings in our case serves as a control experiment of the model setup, initialization and forcing employed by CLASS4GL. In turn, validation of the model results against the 42k IGRA balloon soundings 18k IGRA balloon sounding pairs serves as an overall evaluation of the suitability of CLASS4GL for the appraisal of the ABL behaviour observed and associated land–atmosphere feedbacks for different climate regimes.

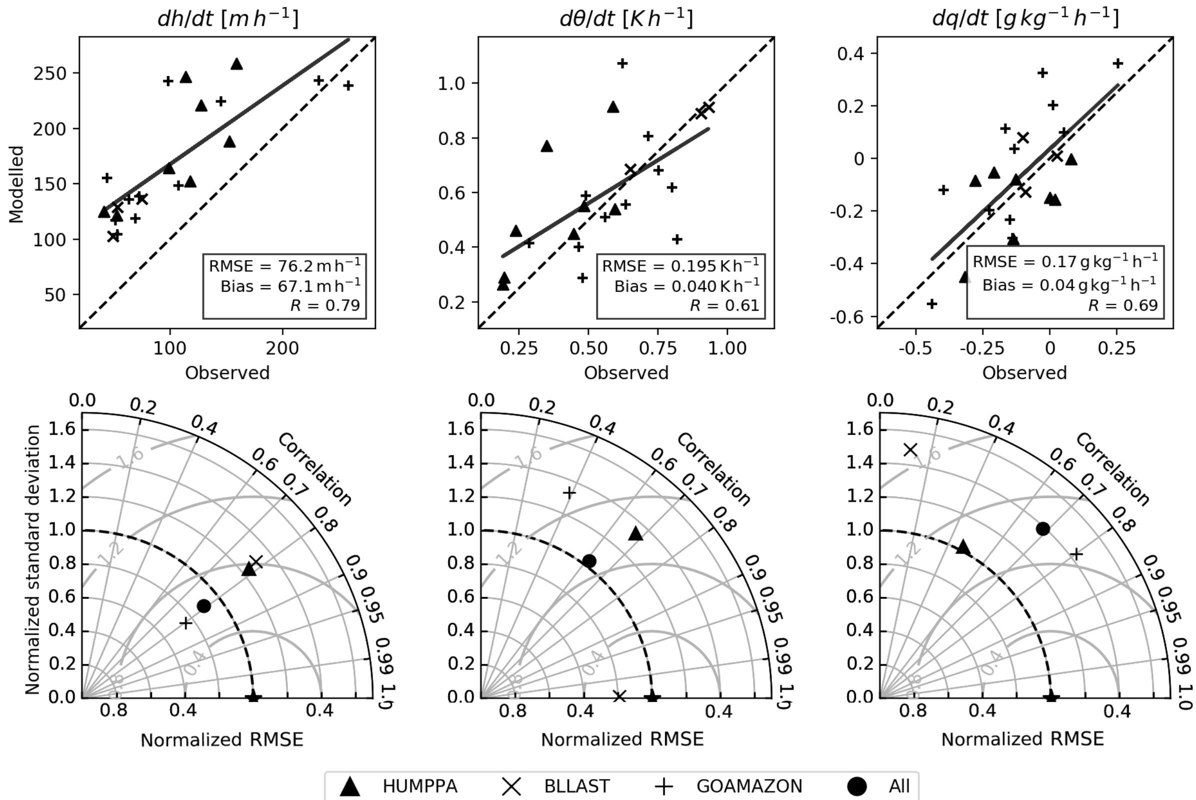


Figure 3. Performance statistics of diurnal changes in mixed-layer properties (mixed-layer height: dh/dt, potential temperature: $d\theta/dt$; specific humidity: dq/dt) during the three intensive observation campaigns: HUMPPA, BLLAST and GOAMAZON. The 1:1 line is shown as black dashed line.

The results of the model skill evaluation against the observations from the three intensive campaigns are ~~summarised~~ summarized in Fig. 3. We find that the framework can reproduce the overall magnitude of the observed daytime tendencies (dh/dt, $d\theta/dt$ and dq/dt), with a bias of ~~5.9~~ 67.1 m h^{-1}, ~~-0.007~~ 0.04 K h^{-1} and ~~0.17~~ 0.04 g kg^{-1}, respectively. The model further reproduces the overall differential response among the campaign days and sites, with a Pearson Correlation Coefficient (R) of ~~0.58, 0.81 and 0.53~~0.79, 0.61 and 0.69, for the three ABL tendencies, respectively. However, we could identify common model limitations over the three campaigns, particularly an overall (slight) positive bias in ~~dq/dt,~~ dh/dt ($d\theta/dt$ and ~~a too low variability in the tendencies,~~ dq/dt), and an under(over)estimation of its (their) variability as indicated with a normalized standard deviation ~~of less than~~ different from 1 in the Taylor plots (Fig. 3). The bias is expected to have multiple origins, including the ABL model and its physical concepts, the forcing data (convergence/advection, evaporation bias, cloud cover...), model tuning parameters (such as the entrainment ratio) and errors in the sounding observations used to initialize and validate the model. All these possible error sources should be investigated in further development of CLASS4GL. The results of the three campaigns are useful as a first check of the model performance against high-quality -observations. However, the sample size of 15 22 days over the three campaigns is too low to gain conclusive insights on whether these biases are of a systematic nature. In this respect, the validation of model performance against the 42,000 36k IGRA soundings from 150 121 different stations in different climate regions provides a more reliable and comprehensive assessment.

The simulated diurnal ABL tendencies show a similar accuracy when evaluated against the global IGRA sounding archive, despite the fact that these operational sounding have *a priori* lower quality standards than the campaign soundings; this applies to all three ABL tendencies shown in Fig. 4. Here, the bias is larger magnitude of the bias is of the same order for the change in θ (-0.052 -0.036 K h^{-1}) than as for the intense observation campaigns, whereas the changes but is now negative. The biases are smaller for h (0.2 10.1 m h^{-1}) and compared to the campaigns and slightly larger for q (0.05 0.06 g kg^{-1}). As for the research campaigns, the model is able to reproduce the variability among the different operational sounding days, with Pearson correlation coefficients of 0.54, 0.79 and 0.52 0.53, 0.82 and 0.54 for the diurnal tendencies of h, θ and q, respectively. In addition, the overall modelled range in dh/dt, $d\theta/dt$ and dq/dt -agrees well with the observed range, with departures from the for which the departure of the modelled (normalized) standard deviation from the observed (normalized) standard deviation of the observations each tendency is below 22% – see %. This can be seen in the Taylor plots in Fig. 4. 4 for which the centers of the open circles are between 0.78 and 1.22 of the normalized standard deviation. There is also a systematic underestimation of the variability for dh/dt, $d\theta/dt$, but not for dq/dt. The negative bias in the temperature tendency and the positive bias in the humidity tendency lead to an overall net heat bias of 114 J kg^{-1} h^{-1}. Similar as for the results in the campaigns, it is expected that such global biases have multiple origins, including biases in the net radiation (which is calculated by the model by prescribing the cloud cover), underestimation of ground heat storage to the soil, the entrainment rates, and/or the prescribed advection. Further research should investigate possible errors related to input datasets and validate them against independent data (e.g., the available CERES data could be used to evaluate the net radiation). Performance statistics vary slightly with the climate region under scope, with e.g., the correlations for $d\theta/dt$ ranging between 0.67 0.70 and 0.89 depending on the climate regime. The highest model skill corresponds to warm temperate climates, while the lowest skill is found for humid regions, although the variability in model skill depends on the diagnosed variable and score metric. The similar model performance for the global IGRA archive (Fig. 4) compared to the intensive field campaigns (Fig. 3) gives us the confidence that the lower quality standards of the operational measurements do not substantially hamper the framework performance, at least once the strict quality-based selection of radiosondes described above in Section 2.2 is adopted.

The results in Figs. 3 and 4 suggest that the overall model skill in reproducing the diurnal cycle of $/theta$ is higher than the skill to reproduce the ABL mixed-layer growth and the cycle of q. Validation results also indicate an overall underestimation of the θ increase and overestimation of the change in q as the day progresses. At the same time, the high ABL mixed-layer growths are underestimated. Additional sensitivity analyses aiming at lowering the EF suggest that these deficiencies could result from an overall overestimation of surface evaporation, and an underestimation of surface sensible heat flux (not shown). This could relate in turn to errors in GLEAM the EF input dataset or in the vegetation parameters (vegetation cover and LAI) that the determine the partitioning between soil evaporation and transpiration. Moreover, the underrepresentation of extremes (e.g., very dry or very wet conditions) could relate to the coarse resolution of the surface soil parameters. Other factors, like the underestimation of the entrainment may also contribute to the cold and wet bias of the model. Particularly, the modelled ~~ABL~~ mixed-layer heat gain by entrainment is positive (0.011 K m s^{-1}; i.e., warm air entrainment) and the average ~~ABL~~ mixed-layer moisture gain by entrainment is negative (-2.24 x 10^{-5} kg kg^{-1} m s^{-1}; i.e., dry air entrainment) averaged over the whole sounding database. Intuitively, if entrainment was increased, due to (e.g.) a higher entrainment velocity, it would lead to further warming (in terms of θ) and drying (in terms of q) of the ~~ABL~~mixed layer, hence it would act to reduce the overall bias.

Overall, validation results in Fig. 3 and 4 show that CLASS4GL ~~mines data~~ provides pre-processing and modelling tools that can be used to represent the overall diurnal behaviour of the ABL. The model realistically represents the main characteristics of the global ABL diurnal evolution according to balloon soundings, including the observed variability around the globe. Several limitations are highlighted; when applying CLASS4GL, one should be aware of the fact that model parameters and assumptions, combined with input data uncertainty, can lead to the failure to simulate specific sounding profiles. The framework is in continuous development, and it is expected that results will improve with higher resolution and accuracy of the forcing datasets, and with evolving model concepts of the ABL and land–atmosphere interface. The use of local parameterisations and higher quality forcing is encouraged for applications in specific regions.

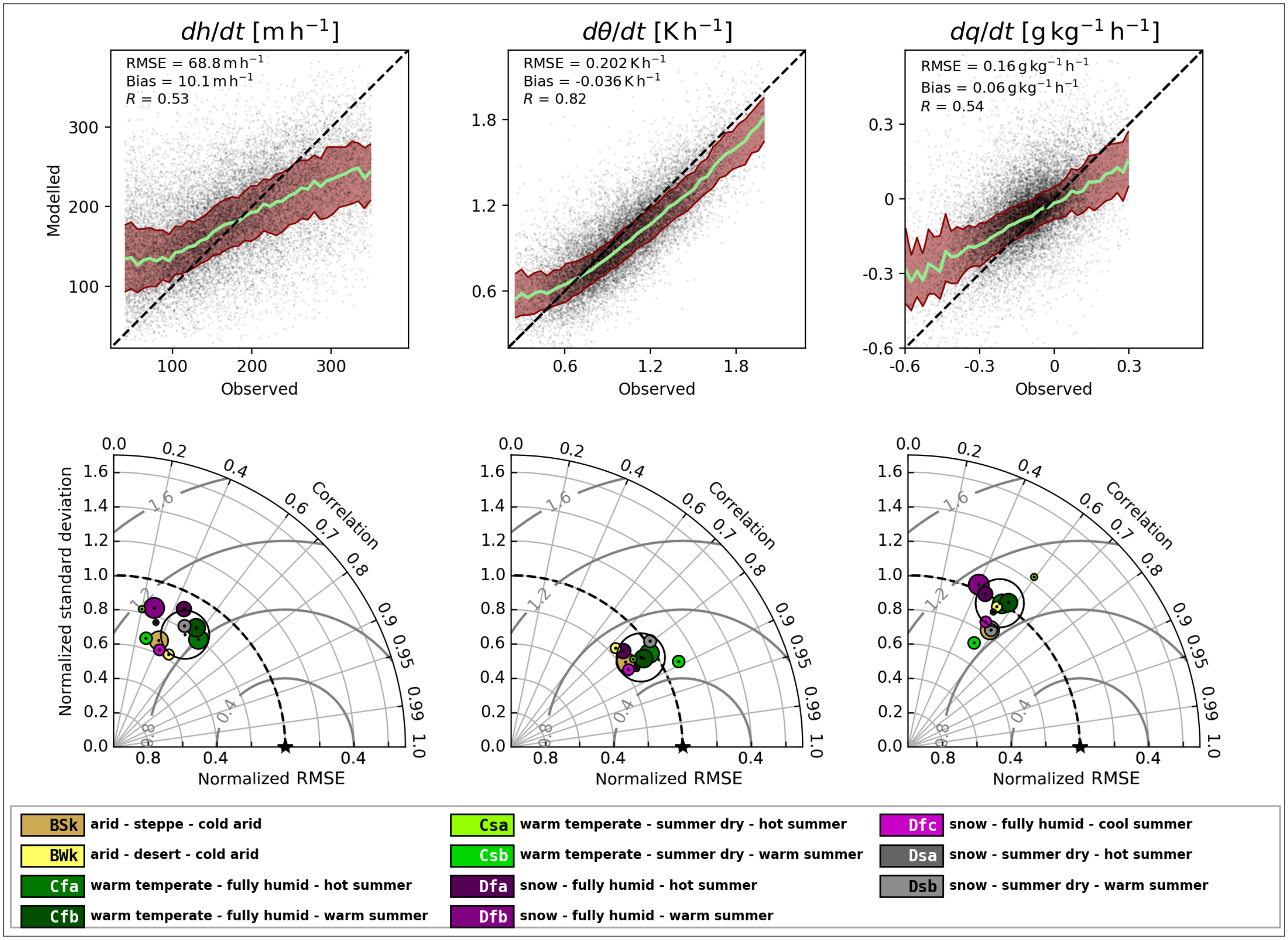


Figure 4. Model skill in reproducing diurnal changes in ABL properties. Shown are the tendencies of the mixed-layer height (dh/dt), potential temperature ($d\theta/dt$) and specific humidity (dq/dt), which are assessed by comparison of model simulations against the IGRA sounding data between 1981 and 2015. The upper plots show modelled versus observed data points (grey), and the corresponding median (green) and interquartile range (red) of the model. The 1:1 line is shown as black dashed line. The Taylor plots indicate the overall model performance as open circles, and the performance corresponding to each Köppen–Geiger climate class as colored bullets. Only climate regimes with more than > 200 soundings are illustrated. The size of the bullets is proportional to the amount of soundings for each particular climate class.

4 Conclusion and perspective

We have presented a novel interactive data platform, referred to as CLASS4GL (http://class4gl.eu), to ~~automate~~ automize the study of the diurnal ABL based on balloon radiosondes launched worldwide since the early ’80s. The framework (a) mines balloon sounding data to initialize and validate the ABL bulk model CLASS, (b) links to a predefined but expandable set of global datasets that are used to constrain the surface and large-scale atmospheric conditions in the CLASS simulations, (c) is successful in evaluating the skill to simulate the ABL globally, (d) provides a flexible and user-friendly interface for allowing extensive amount of experiments on supercomputer environments thanks to a low computational cost (e.g., a batch of 41, 000 diurnal simulations only takes ∼ 1 hour on a contemporary supercomputer infrastructure), and (e) strives for a community-driven architecture that allows to seamlessly share input datasets, experiments, analyses and model developments among the climate research community, and that also facilitates hackathons, workshops and educational activities. Validation results show an overal realistic representation of the diurnal evolution of potential temperature, specific humidityand ABL heightmixed-layer height, potential temperature and specific humidity, using data from three observation campaigns and the IGRA operational sounding dataset.

The freely-available model framework offers new perspectives to foster the study of the diurnal evolution of the ABL and the associated land—atmosphere feedbacks:

- The fast software infrastructure allows any researcher to easily employ extensive global sensitivity experiments, in which both land and atmospheric parameters can be perturbed. That can be used to investigate the effect of land conditions and large-scale atmospheric forcing on the ABL evolution worldwide. Particularly, this inititive fits well within the context of LoCo activities (http://www.gewex.org/loco/) to, e.g., construct mixing diagrams and other metrics to understand land–atmospheric feedbacks. Moreover by integrating global information of precipitation, temperature and cloud statistics, one could further investigate whether particular combinations of surface and atmospheric conditions lead to ABL properties that favor or disfavor the occurrence of convection, clouds, precipitation, or extreme temperatures.

- Using the radio sounding simulations as reference, the framework can be employed to study the climate model representation of ABL dynamics and the associated land—atmosphere feedbacks at the diurnal scales, and to evaluate satellite-based products and reanalysis data. Such a process-based validation can help improve climate models and assess the quality of satellite products intended to monitor the land–atmosphere interface.

- It could be used to challenge the added value of including novel mechanistic concepts, such as the dynamic representation of soil and vegetation interacting with the carbon cycle (CO_2 exchange, carbon stock, air CO_2 levels, etc.), atmospheric chemistry (VOCs, aerosols, ozone, etc.), vegetation dynamics and water stress (Combe et al., 2016), or urbanization (Droste et al., 2018; Wouters et al., 2016).

- Finally, the ABL evolution and the associated land–atmosphere interactions could be extracted from climate projections and land-cover climate scenarios. This way, one may determine the local drivers of shifting (extreme) weather under climate change, hence provide a better process-understanding. When integrating future land use scenarios, it may foster the development of more effective climate adaptation strategies, e.g., by quantifying the mitigatory potential of land use change that could alleviate the escalation of ABL mixed-layer temperatures during heatwaves.

Code availability

CLASS4GL freely available as a Python module and is conveyed under GNU General Public license version 3 (GPLv3). The general information, code and tutorials for running the software is maintained at http://class4gl.eu. The presented model version for CLASS4GL is 1.0. All comments, questions, suggestions and critiques regarding the functioning of the Python routine can be directed to the author of this paper.

Acknowledgements. This research was funded by the European Research Council (ERC) under Grant Agreement No. 715254 (DRY–2–DRY). BLLAST field experiment was made possible thanks to the contributions of several institutions and supports: INSU-CNRS (Institute National des Sciences de l'Univers, Centre National de la Recherche Scientifique, LEFE-IDAO program), Météo France, Observatoire Midi-Pyrénées (University of Toulouse), EUFAR (EUropean Facility for Airborne Research) and COST ES0802 (European Cooperation in the field of Science and Technology). The field experiment would not have occurred without the contribution of all participating European and American research groups, which all have contributed in a significant amount (see http://bllast.sedoo.fr/supports/). BLLAST field experiment was hosted by the instrumental site of Centre des Recherches Atmosphérique, Lannemezan, France (Observatoire Midi-Pyrénées, Laboratoire d'Aérologie). BLLAST data are managed by SEDOO, from Observatoire Midi-Pirénées. We would like to thank J. Williams et al. and S.T. Martin et al. for providing the campaign data for HUMPPA and GOAMAZON, respectively. We thank Florian Cochard and Pierre Gentine for the radiosonde and profile analysis. We also thank Diminik Schumacher for employing independent tests of the software tutorials. The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government – department EWI.

References

Akkermans, T., Rompaey, A. V., Lipzig, N. V., Moonen, P., and Verbist, B.: Quantifying successional land cover after clearing of tropical rainforest along forest frontiers in the Congo Basin, Physical Geography, 34, 417–440, https://doi.org/10.1080/02723646.2013.855698, http://dx.doi.org/10.1080/02723646.2013.855698, 2013.

Bauer, L. A.: The Relation between "Potential Temperature" and "Entropy.", Physical Review (Series I), 26, 177–183, https://doi.org/10.1103/physrevseriesi.26.177, http://dx.doi.org/10.1103/physrevseriesi.26.177, 1908.

Berg, A., Lintner, B. R., Findell, K., Seneviratne, S. I., van den Hurk, B., Ducharne, A., Chéruy, F., Hagemann, S., Lawrence, D. M., Malyshev, S., Meier, A., and Gentine, P.: Interannual Coupling between Summertime Surface Temperature and Precipitation over Land: Processes and Implications for Climate Change*, Journal of Climate, 28, 1308–1328, https://doi.org/10.1175/jcli-d-14-00324.1, http://dx.doi.org/10.1175/jcli-d-14-00324.1, 2015.

Betts, A. K. and Ball, J. H.: The FIFE surface diurnal cycle climate, Journal of Geophysical Research, 100, 25 679, https://doi.org/10.1029/94jd03121, http://dx.doi.org/10.1029/94jd03121, 1995.

Burden, R. L., Faires, D. J., and Burden, A. M.: Numerical Analysis (10th edition), 2016.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resources Research, 14, 601–604, https://doi.org/10.1029/wr014i004p00601, http://dx.doi.org/10.1029/wr014i004p00601, 1978.

Combe, M., Vilà-Guerau de Arellano, J., Ouwersloot, H. G., and Peters, W.: Plant water-stress parameterization determines the strength of land–atmosphere coupling, Agricultural and Forest Meteorology, 217, 61–73, https://doi.org/10.1016/j.agrformet.2015.11.006, http://dx.doi.org/10.1016/j.agrformet.2015.11.006, 2016.

Copernicus Climate Change Service (C3S): Fifth generation of ECMWF atmospheric reanalyses of the global climate, https://cds.climate.copernicus.eu/cdsapp{#}!/home, 2017.

Davin, E. L., Rechid, D., Breil, M., Cardoso, R. M., Coppola, E., Hoffmann, P., Jach, L. L., Katragkou, E., de Noblet-Ducoudré, N., Radtke, K., Raffa, M., Soares, P. M. M., Sofiadis, G., Strada, S., Strandberg, G., Tölle, M. H., Warrach-Sagi, K., and Wulfmeyer, V.: Biogeophysical impacts of forestation in Europe: First results from the LUCAS Regional Climate Model intercomparison, Earth System Dynamics Discussions, pp. 1–31, https://doi.org/10.5194/esd-2019-4, https://www.earth-syst-dynam-discuss.net/esd-2019-4/, 2019.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Droste, A. M., Steeneveld, G. J., and Holtslag, A. A. M.: Introducing the urban wind island effect, Environmental Research Letters, 13, 094 007, https://doi.org/10.1088/1748-9326/aad8ef, http://stacks.iop.org/1748-9326/13/i=9/a=094007?key=crossref.20cecdb1bbb8739c7811dedbf48bc43b, 2018.

Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, Journal of Climate, 19, 53–68, https://doi.org/10.1175/jcli3594.1, http://dx.doi.org/10.1175/jcli3594.1, 2006.

Ek, M. and Mahrt, L.: Daytime Evolution of Relative Humidity at the Boundary Layer Top, Monthly Weather Review, 122, 2709–2721, https://doi.org/10.1175/1520-0493(1994)122<2709:deorha>2.0.co;2, http://dx.doi.org/10.1175/1520-0493(1994)122<2709:deorha>2.0.co;2, 1994.

Findell, K. L., Gentine, P., Lintner, B. R., and Kerr, C.: Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation, Nature Geoscience, 4, 434–439, https://doi.org/10.1038/ngeo1174, http://dx.doi.org/10.1038/ngeo1174, 2011.

Fischer, E. M., Seneviratne, S. I., Lüthi, D., and Schär, C.: Contribution of land-atmosphere coupling to recent European summer heat waves, Geophysical Research Letters, 34, https://doi.org/10.1029/2006gl029068, http://dx.doi.org/10.1029/2006gl029068, 2007.

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., and Seneviratne, S. I.: Reconciling spatial and temporal soil moisture effects on afternoon rainfall, Nature Communications, 6, https://doi.org/10.1038/ncomms7443, http://dx.doi.org/10.1038/ncomms7443, 2015.

Guo, Z., Dirmeyer, P. A., Koster, R. D., Sud, Y. C., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Gordon, C. T., McGregor, J. L., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Mocko, D., Lu, C.-H., Mitchell, K., Malyshev, S., McAvaney, B., Oki, T., Yamada, T., Pitman, A., Taylor, C. M., Vasic, R., and Xue, Y.: GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis, Journal of Hydrometeorology, 7, 611–625, https://doi.org/10.1175/jhm511.1, http://dx.doi.org/10.1175/jhm511.1, 2006.

Hansen, M. C., Townshend, J. R. G., DeFries, R. S., and Carroll, M.: Estimation of tree cover using MODIS data at global, continental and regional/local scales, International Journal of Remote Sensing, 26, 4359–4380, https://doi.org/10.1080/01431160500113435, http://dx.doi.org/10.1080/01431160500113435, 2005.

IGBP-DIS: Global gridded surfaces of selection soil characteristics (IGBP-DIS), https://doi.org/10.3334/ornldaac/569, http://dx.doi.org/10.3334/ornldaac/569, 2000.

Jarvis, P. G.: The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field, Philosophical Transactions of the Royal Society B: Biological Sciences, 273, 593–610, https://doi.org/10.1098/rstb.1976.0035, http://dx.doi.org/10.1098/rstb.1976.0035, 1976.

Koster, R. D., Sud, Y. C., Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., Chan, E., Verseghy, D., Cox, P., Davies, H., Kowalczyk, E., Gordon, C. T., Kanae, S., Lawrence, D., Liu, P., Mocko, D., Lu, C.-H., Mitchell, K., Malyshev, S., McAvaney, B., Oki, T., Yamada, T., Pitman, A., Taylor, C. M., Vasic, R., and Xue, Y.: GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview, Journal of Hydrometeorology, 7, 590–610, https://doi.org/10.1175/jhm510.1, http://dx.doi.org/10.1175/jhm510.1, 2006.

Lawston, P. M., Santanello, J. A., Zaitchik, B. F., and Rodell, M.: Impact of Irrigation Methods on Land Surface Model Spinup and Initialization of WRF Forecasts, Journal of Hydrometeorology, 16, 1135–1154, https://doi.org/10.1175/jhm-d-14-0203.1, http://dx.doi.org/10.1175/jhm-d-14-0203.1, 2015.

Lejeune, Q., Davin, E. L., Guillod, B. P., and Seneviratne, S. I.: Influence of Amazonian deforestation on the future evolution of regional surface fluxes, circulation, surface temperature and precipitation, Climate Dynamics, 44, 2769–2786, https://doi.org/10.1007/s00382-014-2203-8, http://dx.doi.org/10.1007/s00382-014-2203-8, 2014.

Liu, Y., Liu, R., and Chen, J. M.: Retrospective retrieval of long-term consistent global leaf area index (1981-2011) from combined AVHRR and MODIS data, Journal of Geophysical Research: Biogeosciences, 117, n/a–n/a, https://doi.org/10.1029/2012jg002084, http://dx.doi.org/10.1029/2012jg002084, 2012.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, http://dx.doi.org/10.5194/gmd-10-1903-2017, 2017.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmospheric Chemistry and Physics, 16, 4785–4797, https://doi.org/10.5194/acp-16-4785-2016, http://dx.doi.org/10.5194/acp-16-4785-2016, 2016.

Miralles, D. G., Holmes, T. R. H., Jeu, R. A. M. D., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453–469, https://doi.org/10.5194/hess-15-453-2011, http://dx.doi.org/10.5194/hess-15-453-2011, 2011.

Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C., and Vilà-Guerau de Arellano, J.: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature Geoscience, 7, 345–349, https://doi.org/10.1038/ngeo2141, http://dx.doi.org/10.1038/ngeo2141, 2014.

Miralles, D. G., Gentine, P., Seneviratne, S. I., and Teuling, A. J.: Land-atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges, Annals of the New York Academy of Sciences, https://doi.org/10.1111/nyas.13912, http://dx.doi.org/10.1111/nyas.13912, 2018.

Monin, A. S. and Obukhov, A. M.: Basic laws of turbulent mixing in the surface layer of the atmosphere, Contributions of the Geophysical Institute of the Slovak Academy of Sciences, 24, 1954.

Monteith, J. L.: Radiation and Crops, Experimental Agriculture, 1, 241, https://doi.org/10.1017/s0014479700021529, http://dx.doi.org/10.1017/s0014479700021529, 1965.

Nachtergaele, F., van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, K., van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., Wiberg, D., and Shi, X.: Harmonized World Soil Database, Tech. rep., Food and Agriculture Organization of the United Nations, 2009.

Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterisation scheme, Global and Planetary Change, 13, 145–159, https://doi.org/10.1016/0921-8181(95)00043-7, http://dx.doi.org/10.1016/0921-8181(95)00043-7, 1996.

Noilhan, J. and Planton, S.: A Simple Parameterization of Land Surface Processes for Meteorological Models, Monthly Weather Review, 117, 536–549, https://doi.org/10.1175/1520-0493(1989)117<0536:aspols>2.0.co;2, http://dx.doi.org/10.1175/1520-0493(1989)117<0536:aspols>2.0.co;2, 1989.

Orlowsky, B. and Seneviratne, S. I.: Statistical Analyses of Land–Atmosphere Feedbacks and Their Possible Pitfalls, Journal of Climate, 23, 3918–3932, https://doi.org/10.1175/2010jcli3366.1, http://dx.doi.org/10.1175/2010jcli3366.1, 2010.

Ouwersloot, H. G., Vilà-Guerau de Arellano, J., Nölscher, A. C., Krol, M. C., Ganzeveld, L. N., Breitenberger, C., Mammarella, I., Williams, J., and Lelieveld, J.: Characterization of a boreal convective boundary layer and its impact on atmospheric chemistry during HUMPPA-COPEC-2010, Atmospheric Chemistry and Physics, 12, 9335–9353, https://doi.org/10.5194/acp-12-9335-2012, http://dx.doi.org/10.5194/acp-12-9335-2012, 2012.

Paulson, C. A.: The Mathematical Representation of Wind Speed and Temperature Profiles in the Unstable Atmospheric Surface Layer, Journal of Applied Meteorology, 9, 857–861, https://doi.org/10.1175/1520-0450(1970)009<0857:tmrows>2.0.co;2, http://dx.doi.org/10.1175/1520-0450(1970)009<0857:tmrows>2.0.co;2, 1970.

Petrova, I., Miralles, D., van Heerwaarden, C., and Wouters, H.: Relation between Convective Rainfall Properties and Antecedent Soil Moisture Heterogeneity Conditions in North Africa, Remote Sensing, 10, 969, https://doi.org/10.3390/rs10060969, http://dx.doi.org/10.3390/rs10060969, 2018.

Pietersen, H. P., Vilà-Guerau de Arellano, J., Augustin, P., van de Boer, A., de Coster, O., Delbarre, H., Durand, P., Fourmentin, M., Gioli, B., Hartogensis, O., Lohou, F., Lothon, M., Ouwersloot, H. G., Pino, D., and Reuder, J.: Study of a prototypical convective boundary layer observed during BLLAST: contributions by large-scale forcings, Atmospheric Chemistry and Physics, 15, 4241–4257, https://doi.org/10.5194/acp-15-4241-2015, http://dx.doi.org/10.5194/acp-15-4241-2015, 2015.

Ridder, K. D.: Bulk Transfer Relations for the Roughness Sublayer, Boundary-Layer Meteorology, 134, 257–267, https://doi.org/10.1007/s10546-009-9450-y, http://dx.doi.org/10.1007/s10546-009-9450-y, 2009.

Roundy, J. K. and Santanello, J. A.: Utility of Satellite Remote Sensing for Land–Atmosphere Coupling and Drought Metrics, Journal of Hydrometeorology, 18, 863–877, https://doi.org/10.1175/jhm-d-16-0171.1, http://dx.doi.org/10.1175/jhm-d-16-0171.1, 2017.

Roundy, J. K., Ferguson, C. R., and Wood, E. F.: Temporal Variability of Land–Atmosphere Coupling and Its Implications for Drought over the Southeast United States, Journal of Hydrometeorology, 14, 622–635, https://doi.org/10.1175/jhm-d-12-090.1, http://dx.doi.org/10.1175/jhm-d-12-090.1, 2013.

Santanello, J. A., Peters-Lidard, C. D., Kumar, S. V., Alonge, C., and Tao, W.-K.: A Modeling and Observational Framework for Diagnosing Local Land–Atmosphere Coupling on Diurnal Time Scales, Journal of Hydrometeorology, 10, 577–599, https://doi.org/10.1175/2009jhm1066.1, http://dx.doi.org/10.1175/2009jhm1066.1, 2009.

Santanello, J. A., Roundy, J., and Dirmeyer, P. A.: Quantifying the Land–Atmosphere Coupling Behavior in Modern Reanalysis Products over the U.S. Southern Great Plains, Journal of Climate, 28, 5813–5829, https://doi.org/10.1175/jcli-d-14-00680.1, http://dx.doi.org/10.1175/jcli-d-14-00680.1, 2015.

Santanello, J. A., Dirmeyer, P. A., Ferguson, C. R., Findell, K. L., Tawfik, A. B., Berg, A., Ek, M., Gentine, P., Guillod, B. P., van Heerwaarden, C., Roundy, J., and Wulfmeyer, V.: Land–Atmosphere Interactions: The LoCo Perspective, Bulletin of the American Meteorological Society, 99, 1253–1272, https://doi.org/10.1175/bams-d-17-0001.1, http://dx.doi.org/10.1175/bams-d-17-0001.1, 2018.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004, http://dx.doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L. V., Hegerl, G., Seneviratne, S. I., Vautard, R., Zhang, X., and Zwiers, F. W.: Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities, Weather and Climate Extremes, 18, 65–74, https://doi.org/10.1016/j.wace.2017.10.003, http://dx.doi.org/10.1016/j.wace.2017.10.003, 2017.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, Journal of Geophysical Research, 116, https://doi.org/10.1029/2011jg001708, http://dx.doi.org/10.1029/2011jg001708, 2011.

Tawfik, A. B., Dirmeyer, P. A., and Santanello, J. A.: The Heated Condensation Framework. Part I: Description and Southern Great Plains Case Study, Journal of Hydrometeorology, 16, 1929–1945, https://doi.org/10.1175/jhm-d-14-0117.1, http://dx.doi.org/10.1175/jhm-d-14-0117.1, 2015.

Taylor, C. M., de Jeu, R. A. M., Guichard, F., Harris, P. P., and Dorigo, W. A.: Afternoon rain more likely over drier soils, Nature, 489, 423–426, https://doi.org/10.1038/nature11377, http://dx.doi.org/10.1038/nature11377, 2012.

Tennekes, H.: A Model for the Dynamics of the Inversion Above a Convective Boundary Layer, Journal of the Atmospheric Sciences, 30, 558–567, https://doi.org/10.1175/1520-0469(1973)030<0558:amftdo>2.0.co;2, http://dx.doi.org/10.1175/1520-0469(1973)030<0558:amftdo>2.0.co;2, 1973.

Tennekes, H. and Driedonks, A. G. M.: Basic entrainment equations for the atmospheric boundary layer, Boundary-Layer Meteorology, 20, 515–531, https://doi.org/10.1007/bf00122299, http://dx.doi.org/10.1007/bf00122299, 1981.

Teuling, A. J., Van Loon, A. F., Seneviratne, S. I., Lehner, I., Aubinet, M., Heinesch, B., Bernhofer, C., Grünwald, T., Prasse, H., and Spank, U.: Evapotranspiration amplifies European summer drought, Geophysical Research Letters, 40, 2071–2075, https://doi.org/10.1002/grl.50495, http://doi.wiley.com/10.1002/grl.50495, 2013.

Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., and Seneviratne, S. I.: Present-day irrigation mitigates heat extremes, Journal of Geophysical Research: Atmospheres, 122, 1403–1422, https://doi.org/10.1002/2016jd025740, http://dx.doi.org/10.1002/2016jd025740, 2017.

van Heerwaarden, C. C. and Teuling, A. J.: Disentangling the response of forest and grassland energy exchange to heatwaves under idealized land–atmosphere coupling, Biogeosciences, 11, 6159–6171, https://doi.org/10.5194/bg-11-6159-2014, http://dx.doi.org/10.5194/bg-11-6159-2014, 2014.

van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Moene, A. F., and Holtslag, A. A. M.: Interactions between dry-air entrainment, surface evaporation and convective boundary-layer development, Quarterly Journal of the Royal Meteorological Society, 135, 1277–1291, https://doi.org/10.1002/qj.431, http://dx.doi.org/10.1002/qj.431, 2009.

van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Gounou, A., Guichard, F., and Couvreux, F.: Understanding the Daily Cycle of Evapotranspiration: A Method to Quantify the Influence of Forcings and Feedbacks, Journal of Hydrometeorology, 11, 1405–1422, https://doi.org/10.1175/2010jhm1272.1, http://dx.doi.org/10.1175/2010jhm1272.1, 2010.

van Stratum, B. J. H. and Stevens, B.: The Impact of Vertical Mixing Biases in Large-Eddy Simulation on Nocturnal Low Clouds, Journal of Advances in Modeling Earth Systems, 10, 1290–1303, https://doi.org/10.1029/2017ms001239, http://dx.doi.org/10.1029/2017ms001239, 2018.

Vilà-Guerau de Arellano, J., van Heerwaarden, C. C., and Lelieveld, J.: Modelled suppression of boundary-layer clouds by plants in a CO2-rich atmosphere, Nature Geoscience, 5, 701–704, https://doi.org/10.1038/ngeo1554, http://dx.doi.org/10.1038/ngeo1554, 2012.

Vilà-Guerau de Arellano, J., van Heerwaarden, C. C., van Stratum, B. J. H., and van den Dries, K.: The Atmospheric Boundary Layer, vol. Cambridge University Press, 2015.

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Louis Smith, G., Cooper, J. E., Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., III, R. B. L., Smith, G. L., and Cooper, J. E.: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment, Bulletin of the American Meteorological Society, 77, 853–868, https://doi.org/10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0477{%}281996{%}29077{%}3C0853{%}3ACATERE{%}3E2.0.CO{%}3B2, 1996.

Williams, J., Crowley, J., Fischer, H., Harder, H., Martinez, M., Petäjä, T., Rinne, J., Bäck, J., Boy, M., Maso, M. D., Hakala, J., Kajos, M., Keronen, P., Rantala, P., Aalto, J., Aaltonen, H., Paatero, J., Vesala, T., Hakola, H., Levula, J., Pohja, T., Herrmann, F., Auld, J., Mesarchaki, E., Song, W., Yassaa, N., Nölscher, A., Johnson, A. M., Custer, T., Sinha, V., Thieser, J., Pouvesle, N., Taraborrelli, D., Tang, M. J., Bozem, H., Hosaynali-Beygi, Z., Axinte, R., Oswald, R., Novelli, A., Kubistin, D., Hens, K., Javed, U., Trawny, K., Breitenberger, C., Hidalgo, P. J., Ebben, C. J., Geiger, F. M., Corrigan, A. L., Russell, L. M., Ouwersloot, H. G., Vilà-Guerau de Arellano, J., Ganzeveld, L., Vogel, A., Beck, M., Bayerle, A., Kampf, C. J., Bertelmann, M., Köllner, F., Hoffmann, T., Valverde, J., González, D., Riekkola, M.-L., Kulmala, M., and Lelieveld, J.: The summertime Boreal forest field measurement intensive (HUMPPA-COPEC-2010): an overview of meteorological and chemical influences, Atmospheric Chemistry and Physics, 11, 10 599–10 618, https://doi.org/10.5194/acp-11-10599-2011, http://dx.doi.org/10.5194/acp-11-10599-2011, 2011.

Wouters, H., Ridder, K. D., and van Lipzig, N. P. M.: Comprehensive Parametrization of Surface-Layer Transfer Coefficients for Use in Atmospheric Numerical Models, Boundary-Layer Meteorology, 145, 539–550, https://doi.org/10.1007/s10546-012-9744-3, http://dx.doi.org/10.1007/s10546-012-9744-3, 2012.

Wouters, H., Demuzere, M., Blahak, U., Fortuniak, K., Maiheu, B., Camps, J., Tielemans, D., and van Lipzig, N. P. M.: The efficient urban canopy dependency parametrization (SURY) v1.0 for atmospheric modelling: description and application with the COSMO-CLM model for a Belgian summer, Geoscientific Model Development, 9, 3027–3054, https://doi.org/10.5194/gmd-9-3027-2016, http://dx.doi.org/10.5194/gmd-9-3027-2016, 2016.

Wouters, H., De Ridder, K., Poelmans, L., Willems, P., Brouwers, J., Hosseinzadehtalaei, P., Tabari, H., Vanden Broucke, S., van Lipzig, N. P. M., and Demuzere, M.: Heat stress increase under climate change twice as large in cities as in rural areas: A study for a densely populated midlatitude maritime region, Geophysical Research Letters, 44, 8997–9007, https://doi.org/10.1002/2017gl074889, http://dx.doi.org/10.1002/2017gl074889, 2017.

Zaitchik, B. F., Santanello, J. A., Kumar, S. V., and Peters-Lidard, C. D.: Representation of Soil Moisture Feedbacks during Drought in NASA Unified WRF (NU-WRF), Journal of Hydrometeorology, 14, 360–367, https://doi.org/10.1175/jhm-d-12-069.1, http://dx.doi.org/10.1175/jhm-d-12-069.1, 2013.

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, Geoscientific Model Development Discussions, 7, 4045–4079, https://doi.org/10.5194/gmdd-7-4045-2014, http://dx.doi.org/10.5194/gmdd-7-4045-2014, 2014.