Examining Parameterizations of Potential Temperature Variance Across Varied Landscapes for use in Earth System Models

Tyler Waterman[1,1], Andrew D. Bragg[2,2], Gabriel G. Katul[2,2], and Nathaniel W. Chaney[2,2]

[1]Duke University Civil and Environmental Engineering
[2]Duke University

November 30, 2022

Abstract

Earth system models (ESMs) and mesoscale models have come to employ increasingly complex parameterization schemes for the atmospheric boundary layer (ABL), requiring surface boundary conditions for numerous higher order turbulence statistics. Of particular interest is the potential temperature variance (PTV), which is used not only as a boundary condition itself but also to close boundary conditions of other statistics. The existing schemes in ESMs largely rely on the assumptions of Monin-Obukhov similarity theory (MOST), and are not necessarily applicable over complex and heterogeneous surfaces where large scale circulations and roughness sub-layer effects may cause deviations from MOST. The National Ecological Network (NEON) is used here to evaluate existing parameterizations for the surface boundary of PTV, note key deficiencies, and explore possible remedies. The results indicate that existing schemes are acceptable over a variety of surface conditions provided the analysis of a priori filters out low frequency variability not associated with turbulent time scales. There was, however, significant inter-site variability in observed similarity constants and a significant bias when compared to the textbook values of these parameters. Existing models displayed the poorest performance over heterogeneous sites, and rough landscapes. Attempts to use canopy structure and surface roughness characteristics to improve the results confirmed a relation between these variables and PTV, but failed to significantly improve the predictive power of the models. The results did not find strong evidence indicating that large scale circulations caused substantial deviations from textbook models, although additional analysis is required to assess their full impacts.

Examining Parameterizations of Potential Temperature Variance Across Varied Landscapes for use in Earth System Models

T. Waterman[1], A.D. Bragg[1], G. Katul[1], N. Chaney[1]

[1]Department of Civil and Environmental Engineering, Duke University Pratt School of Engineering, Durham, NC USA

Key Points:

- Models of potential temperature variance in the surface layer based on similarity theory were evaluated using data from 39 varied sites
- Existing schemes perform well across most surfaces, although the data shows a significant bias in the values of the similarity constants
- Canopy structure and surface heterogeneity drive a large portion of inter-site variability in model performance

Corresponding author: Tyler Waterman, `tyler.waterman@duke.edu`

Abstract

Earth system models (ESMs) and mesoscale models have come to employ increasingly complex parameterization schemes for the atmospheric boundary layer (ABL), requiring surface boundary conditions for numerous higher order turbulence statistics. Of particular interest is the potential temperature variance (PTV), which is used not only as a boundary condition itself but also to close boundary conditions of other statistics. The existing schemes in ESMs largely rely on the assumptions of Monin-Obukhov similarity theory (MOST), and are not necessarily applicable over complex and heterogeneous surfaces where large scale circulations and roughness sub-layer effects may cause deviations from MOST. The National Ecological Network (NEON) is used here to evaluate existing parameterizations for the surface boundary of PTV, note key deficiencies, and explore possible remedies. The results indicate that existing schemes are acceptable over a variety of surface conditions provided the analysis of a priori filters out low frequency variability not associated with turbulent time scales. There was, however, significant inter-site variability in observed similarity constants and a significant bias when compared to the textbook values of these parameters. Existing models displayed the poorest performance over heterogeneous sites, and rough landscapes. Attempts to use canopy structure and surface roughness characteristics to improve the results confirmed a relation between these variables and PTV, but failed to significantly improve the predictive power of the models. The results did not find strong evidence indicating that large scale circulations caused substantial deviations from textbook models, although additional analysis is required to assess their full impacts.

Plain Language Summary

Modern models of the lower atmosphere, which are used to analyze climate change and weather, resolve increasingly complex characteristics of the turbulence in the atmosphere. An estimate for the value of many of these characteristics at the land surface is required to set boundary conditions for these models. An important boundary condition is the variance of very small temperature fluctuations that occur in the atmosphere due to turbulence. Currently, model estimates for these values assume the surface is flat and its characteristics do not change in space, which doesn't represent many of the conditions we wish to model over the earth. In addition, existing studies tend to only analyze data from a small number of locations. We analyzed data from a network of 39 sites and found that the current estimates work fairly well across a large variety of conditions, but that there is a bias in the constants often used and there are notable differences over forests, complex surfaces, and heterogeneous terrain. There is a clear relationship between surface characteristics such as tree canopy height and performance of the model, however it was not clear enough to improve our ability to predict the surface boundary condition.

1 Introduction

The atmospheric boundary layer (ABL) plays a fundamental role in the climate system due to its significance in bridging land surface fluxes of heat and water vapor to convection and cloud formation (Siqueira et al., 2009; Huang & Margulis, 2010; Garratt, 1992). The ABL is characterized by the coexistence of mechanically and thermally generated turbulence, which regulate mixing and transport properties and exchanges between the land surface and the lower atmosphere. The variances of turbulent quantities are of particular interest due to their emerging role in state-of-the-science Earth System Models (ESMs) and numerical weather prediction. They have accordingly received attention in the literature, although most of these studies have focused on the velocity variances. Comparatively few examine the potential temperature variance (PTV) and those that do often focus on flat homogeneous terrain (Albertson et al., 1995; Asanuma & Brutsaert, 1999; G. G. Katul & Hsieh, 1999; Mironov & Sullivan, 2016; van de Boer et al., 2014; Maronga & Reuder, 2017; Otić et al., 2005; Antonia et al., 1981; D. Li et al., 2016; Monji, 1973; Champagne et al., 1977; Kiely et al., 1996). Traditional boundary layer schemes in ESMs employed first-order or 1.5-order closure schemes (Cohen et al., 2015; Lock et al., 2000), although increasingly many higher order schemes that resolve PTV prognostically throughout the ABL are now in use, such as the Cloud Layers Unified by Binomials (CLUBB) scheme in the Community Earth System Model (CESM) and the Energy Exascale Earth System Model (E3SM) (Larson, 2017), the Mellor-Yamada-Nakanishi-Niino model (MYNN) implemented in the meso-scale Weather Research and Forecasting model and the Model for Interdisciplinary Research on Climate (MIROC) (Nakanishi & Niino, 2009), and the intermediately prognostic higher-order turbulencce closure (IPHOC) implemented in the Community Atmosphere Model, version 5 (Cheng & Xu, 2015). However, less attention has been placed on the surface boundary condition of PTV of these schemes despite their use in the aforementioned models and the fact that many higher order terms are closed based this temperature variance.

The specification of the lower boundary conditions in such schemes utilize Monin-Obukhov Similarity Theory (MOST) that rests on the assumptions of stationary and planar homogeneous, high Reynolds number flow in the absence of subsidence (Monin & Obukhov, 1954). For these idealized conditions, the turbulent fluxes are assumed to be invariant with distance from the boundary and all flow statistics can be reduced to a set of universal curves that vary with the atmospheric stability parameter (Foken, 2006). Currently, one of two parameterization schemes, both consistent with MOST (Tillman, 1972; J. Wyngaard & Coté, 1971) for unstable atmospheric conditions are used in ESMs and are given by

$$\frac{\overline{\theta'^2}}{T_*^2} = a(1 - b\zeta)^{-2/3}, \tag{1}$$

and

$$\frac{\overline{\theta'^2}}{T_*^2} = C_1(-\zeta)^{-2/3}, \tag{2}$$

where θ' is the fluctuating potential temperature, overline indicates time-averaging over a period that is sufficiently long to reliably capture the ensemble statistics of turbulence but short enough relative to variations in the mean state of the ABL, a, b, and C_1 are similarity constants, ζ is the atmospheric stability parameter defined as

$$\zeta = \frac{z - z_d}{L}, \tag{3}$$

with z_d being the zero-plane displacement height, z is the measurement height and L is the Obukhov length (Obukhov, 1946) given by

$$L = -\frac{u_*^3\overline{\theta_v}}{kg\overline{w'\theta'}}, \tag{4}$$

where $k = 0.4$ is the von Kármán constant, g is the gravitational acceleration, u_* is the friction velocity, $\overline{\theta_v}$ is the mean virtual potential temperature, $\overline{w'\theta'}$ is the kinematic turbulent sensible heat flux, and w' is the turbulent vertical velocity. Unstable atmospheric stability conditions is defined by $\zeta < 0$ whereas near-neutral atmospheric stability conditions occurs when $|\zeta| < 0.05$. The T_* is the non-dimensional temperature scale defined as

$$T_* = \frac{\overline{w'\theta'}}{u_*}. \tag{5}$$

Equations (1) and (2) converge as near-convective conditions ($-\zeta \gg 1$) are approached resulting in $ab^{-2/3} = C_1$. For these conditions, the turbulent heat flux can be linked to $\sigma_T = \sqrt{\overline{\theta'^2}}$ through the well known flux-variance expression (Tillman, 1972)

$$\overline{w'\theta'} = C_1^{-3}[kg(z-z_d)]^{1/2}\overline{\theta_v}^{-1/2}\sigma_T^{3/2}. \tag{6}$$

This expression suggests that sensible heat only depends on σ_T and $(z - z_d)$ independent of u_* as expected when convective conditions are approached. For near neutral conditions with $u_* > 0$ and $\zeta \to 0$, equation (1) ensures $\frac{\sigma_T}{T_*} \to \sqrt{a}$ whereas equation (2) suggests that σ_T is indeterminate by MOST. The two-third scaling is fixed for the purposes of this study - a reasonable assumption as it matches the logical, dimensional limits of free convection.

The 'textbook' similarity constants estimated in the literature are $a = 4, b = 8.3$, and $C_1 = 0.95$. These values were initially derived from experiments over flat, homogeneous wheat stubble in Kansas and confirmed by other studies over similarly homogeneous and largely flat terrain (Tillman, 1972; J. Wyngaard & Coté, 1971; J. C. Wyngaard & Coté, 1974; Andre et al., 1978; Albertson et al., 1995; Haugen et al., 1971; Monji, 1973). This lends some support to their supposed universal character. However, MOST is not readily generalizable for application in ESMs over more realistic landscapes, tall forests and a variety of atmospheric conditions such as those associated with significant entrainment and mesoscale phenomenon (Kroon & de Bruin, 1995; Asanuma & Brutsaert, 1999; Lloyd et al., 1991; Hang et al., 2018; Mcnaughton, 2006; van de Boer et al., 2014; Wilson, 2008; Harman, 2012; Brunet, 2020; Q. Li et al., 2018). Previous literature examining the scaling relation between ζ and non-dimensional flow statistics has focused on conditions that satisfy the assumptions of flat uniform surfaces so that the universal character suggested by MOST can be readily tested (Kader & Yaglom, 1990). However, comparatively less research has been carried out over non-idealized terrain. These few studies have found that MOST derived functions may not hold over surfaces such as sparse and open canopies and heterogenous surfaces (Lee, 2009; Kroon & de Bruin, 1995; van de Boer et al., 2014; Hang et al., 2018; Detto et al., 2008). Few studies have consistently examined PTV across a wide variety of land cover types (G. Katul et al., 1995). The latter study suggested that local similarity may still hold (i.e. a local T_* and L can explain the mathematical form of PTV) provided the similarity coefficients (e.g. C_1) are allowed to vary with land cover type. Despite these issues raised, the use of MOST scaling over various landscapes is widespread in ESMs that require it (Nakanishi & Niino, 2009; Larson, 2017; Zhao et al., 2018; Golaz et al., 2019, 2002; Cheng & Xu, 2015). To explore PTV in the atmospheric surface layer across differing landscapes and a wide range of atmospheric conditions, observations covering many ecosystems and canopy structures with appropriate parameterizations are becoming necessary and motivates the present work.

Since these parameterizations were developed, there has been a significant growth in the availability of data across differing surfaces that can be used to re-examine MOST parameterizations. One example is the National Ecological Observation Network (NEON). NEON is a continent-scale network where high frequency (20 Hz) velocity and air temperature fluctuations are sampled in a consistent manner (i.e. same instrumentation, relative heights, pre- and post-processing algorithms, etc..) over 39 sites that vary in climate and land-cover across the United States. Hence, the NEON high frequency data set offers a unique opportunity to explore these similarity relations over many land cover types (ideal and non-ideal) and ζ conditions. Using this information, it is possible to explore validity and modifications to the traditional MOST PTV parameterizations. The initial focus spans near-neutral to unstable stratification ($\zeta < 0$), where the turbulence is fully developed. Stably stratified conditions are characterized by a shallow boundary layer depth and are infected with numerous non-turbulent phenomena that will require a separate investigation that is better kept for a future study.

With this large data set, the time is ripe to revisit and reevaluate traditional schemes for PTV in light of these contemporary needs of ESM. In doing so, the focus is on two deviations from the assumptions of MOST. The first is mesoscale phenomenon and outer-layer eddies that impinge onto the atmospheric surface layer, potentially introducing additional length scales not captured by ζ. The second is roughness sublayer effects, especially over forests or other forms of structured heterogeneity, which is not included as part of MOST. This study seeks to quantify the significance of the distortions from both mesoscale and roughness sublayer effects on equations (1) and (2), and examine if such distortions can be partly absorbed in the parameters a and b (or C_1). The approach that follows takes advantage of the wealth of data provided by NEON as well as remotely sensed sources, and the Random Forest (RF) method, which is a machine learning method able to classify the significance of surrogate terms such as boundary layer height, land cover type, canopy height, and other ancillary variables on $\overline{\theta'^2}/T_*^2$.

2 Data

The core turbulence data are publicly available from NEON and includes the turbulence statistics ($\overline{w'\theta'}$, θ', u_*), meteorological variables (mean temperature, humidity and wind speed), as well as site specific information (h_c, tower characteristics). Additional information from remotely sensed datasets and reanalysis data (see Table 1) colocated with the NEON site data are used in predictive models that seek to link environmental variables and land surface features to PTV.

2.1 National Ecological Observation Network

The NEON sites are located within the continental United States (CONUS), Alaska, Hawaii and Puerto Rico. Sites are centrally managed and designed, which means that sampling and post-processing high frequency data are consistent, and differences can be attributed to site characteristics rather than differences in management, methods and instrumentation as is the case for other locally managed flux tower networks such as FLUXNET or AmeriFlux (Novick et al., 2018). Moreover, the high frequency time series spanning several years are publicly available thus enabling the determination of variances and heat fluxes in a coherent manner when post-processed. Sites are also spread across different ecological domains to ensure coverage of the different landscapes and ecosystems in North America.

Each site includes a full suite of meteorological instrumentation, eddy covariance measurements from a CSAT-3 sonic anemometer recording a 20 Hz and time averaged to 30 min, and mean wind profiles throughout the canopy and above it, compiled into one dataset (National Ecological Observatory Network (NEON), 2021). Data is examined from first availability at each site, which varies by tower but is generally in mid 2017, to May 2020. Towers at sites with a canopy less than three meters are designed to be 8m tall, whereas towers at sites with a canopy greater than three meters are designed to have a height corresponding to $z_d+4(h_c-z_d)$, with canopy height h_c, to ensure that the turbulence exchange assembly samples largely above the momentum roughness layer (Metzger et al., 2019). In addition, detailed canopy structure at each site is acquired through near-annual airborne remote sensing surveys with discrete and full waveform LiDAR. Soil, vegetative and meteorological characteristics are described and continuously collected when appropriate at each site. Only the 39 CONUS sites are included in this analysis.

For illustrative purposes, eight representative sites were selected as examples of site level differences throughout the study. Wind River Experimental Forest (WREF) - a tall evergreen forest in the Pacific Northwest, Northern Great Plains Research Laboratory (NOGP) - a flat grassland site in North Dakota, Bartlett Experimental Forest (BART) - a mixed deciduous evergreen forest in New England, Soaproot Saddle (SOAP) - a conifer forest with complex terrain in the Sierra Nevada mountains, Oak Ridge National Lab (ORNL) - a deciduous forest with some pine in Appalachia, Santa Rita Experimental Range (SRER) - a semiarid scrub environment site in the Sonoran Desert, Konza Prairie

Table 1. Summary table of the remotely sensed data and reanalysis products used in this project with their native spatial and temporal resolution as well as the source of the data

Variable	Spatial Resolution	Temporal Resolution	Source
LAI	250 m	8 days	MODIS
f_{veg}	250 m	1 year	MODIS
f_{tree}	250 m	1 year	MODIS
f_{bare}	250 m	1 year	MODIS
$Land\ Cover$	30 m	N/A	NLCD (Landsat)
BLH	30 km	1 hour	ERA 5 Reanalysis
f_{cloud}	30 km	1 hour	ERA 5 Reanalysis
$CAPE$	30 km	1 hour	ERA 5 Reanalysis

Biological Station (KONZ) - a pristine prairie site in Kansas, and Disney Wilderness Preserve (DSNY) - a wetland site in the headwaters of the everglades.

2.2 Remotely Sensed Data

Three satellite remote sensing datasets are collocated with the NEON site data to complete a full coverage of vegetation and land cover at each site. This data, as well as reanalysis data discussed in section 2.3, are summarized in Table 1.

Two MODIS derived satellite products are used including MODIS Leaf Area Index (LAI) (Myneni et al., 2015), which is reported at about 250m resolution every 8 days. The site is assigned the LAI of the grid cell in which the tower is contained, and linear interpolation is used to determine LAI for days in which MODIS LAI is unavailable. MODIS Vegetative Continuous Fields (VCF)(DiMiceli et al., 2015) includes measurements of vegetation cover at about 250m resolution on a yearly basis, with linear interpolation used to fill in gaps. Similar to MODIS LAI, MODIS VCF at each site for each point is assigned based on the VCF of the grid cell in which the tower is contained. The VCF product details low lying vegetation cover (f_{veg}), tree cover (f_{tree}), and bare soil coverage (f_{bare}) percentages around each site. These products provide basic information about the vegetation structure.

The National Land Cover Dataset (NLCD), a Landsat derived product defining the land cover at 30m pixels over CONUS (Jin et al., 2019), is the third remote dataset employed. Fractional coverage of each landcover type within a 250m radius from the tower location is computed for each site, as well as the dominant NLCD land cover type.

2.3 Reanalysis Data

ERA5 (Hersbach et al., 2018) is a reanalysis dataset that combines historical observations and modelling results to generate hourly data of a variety of land surface and atmospheric characteristics. For this analysis, the boundary layer height (BLH), total cloud cover (f_{cloud}) and Convective Available Potential Energy ($CAPE$) are used to include the impacts on mesoscale phenomenon using commonly reported variables in the meteorological community. BLH is selected as the depth of the boundary layer is closely related to thermal convection strength and the size of some circulations are closely related to this value. f_{cloud} is used as it may also serve as a proxy for deep convection and identification of cloudy conditions that could impact temperature statistics. CAPE is chosen as it is related to updraft and general convection strength in the atmosphere.

3 Methods

Turbulence statistics as directly acquired from NEON includes variance information from non-turbulence sources whereas models such as CLUBB focus on variances produced by turbulent eddies. A filtering process is required to remove non-turbulent events (and lack of stationarity) before they can be used for analysis. In addition, computed z_d is required for the tower area, as values reported by NEON are suspect and represent the physical characteristics of the entire ecological site rather than the local tower footprint. These values are needed to assess the influence of surface roughness on the development of turbulence. One method of analysis to be used is the Random Forest (RF) method, which is employed to determine what physical and environmental characteristics are most significant for the development of variance without constraints imposed by similarity theory and concomitant dimensional analysis.

3.1 Filtering

One of the key assumptions for MOST parameterizations is the stationarity of the temperature time series. The data for the majority of atmospheric conditions at each site are not strictly stationary. Any computed temperature variance value captures variance associated with turbulent eddies and meso-scale disturbances as well as non-stationarity found at transitions from night to day and vice-a-versa. To fulfill the requirement of solely including PTV caused by turbulence as required by ESMs, a high pass filter with a cutoff time scale of 5 min is applied to the high frequency air temperature time series in the Fourier domain. An example application of the high pass filter is featured in figure 1. Time scales exceeding 5 minutes in the air temperature spectra are assumed to be not associated with turbulent eddies produced by mechanical or buoyant production near the surface. In fact, the choice of 5 minutes exceeds by at least one to two orders of magnitude measured peaks in the co-spectra of w' and θ' or the shear time scale $k(z-z_d)/u_*$ linked with MOST. These events do not significantly impact turbulent sensible heat flux but contribute appreciably to temperature variance. For some points the filtering resulted in reductions of variance of up to 50%, however for the majority of the points the change in variance was near negligible. Alternative filter cutoffs greater and less than 5 minutes were examined and the variances were found to have only a very small sensitivity to the exact cutoff value. The remainder of the analysis presented herein uses the filtered temperature variance.

3.2 Canopy Structure Determination

While NEON does report site level h_c and z_d, which is directly estimated from h_c and required for the analysis, these values appear to be reported as averages for the whole ecological site and not the direct tower area, windshed, or source weight function. Reported z_d in particular appear to deviate significantly from experimentally derived values at a number of sites, and since they are constant they do not reflect seasonal changes in canopy, in particular over deciduous forest. z_d also is estimated simply as $z_d = (2/3)h_c$, which while a good rule of thumb is not always accurate. As such, the z_d is estimated here from measured mean wind profile data ($u(z)$) assuming a log wind profile and that u^* is approximately constant with z (as required by MOST). We apply the following using the top three points in the mean wind profile under near-neutral conditions so that the stability correction terms can be ignored (Oke, 1987)

$$\frac{du}{dz} = \frac{u_*}{k(z - z_d)}. \tag{7}$$

The resulting heights are seasonally averaged at each site and then interpolated for use in the computation of ζ via (3). The resulting z_d, seen in figure 2, approximately follow the 2/3 relation reported in the literature (Garratt, 1992) over most sites. This relation, however, is less clear at a number of sites with short vegetation. This deviation may not

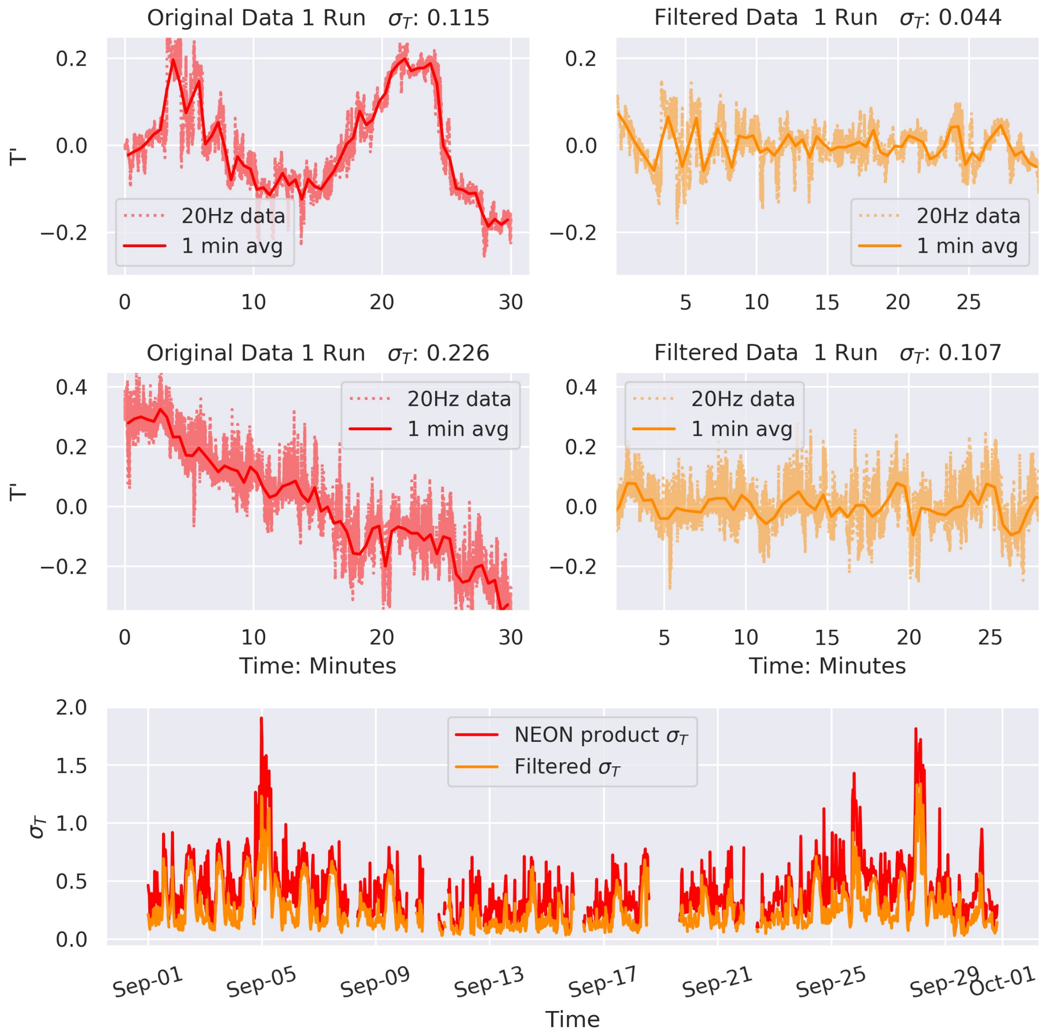

Figure 1. Illustration of filtering to reduce the effects of non-stationarity. Top row shows the raw unfiltered time series for one 30-minute run (left) and the same data after the high pass filter was employed where the nonlinear trend is removed (right). The second row shows the raw unfiltered data for another 30-minute run (left) and the same data after the high pass filter where the approximate linear trend is removed (right). Dotted lines are the actual data, and the solid line is the 1 minute average value. The bottom plot illustrates the change in over a one-month period at the ABBY site from the unfiltered NEON product and the filtered data

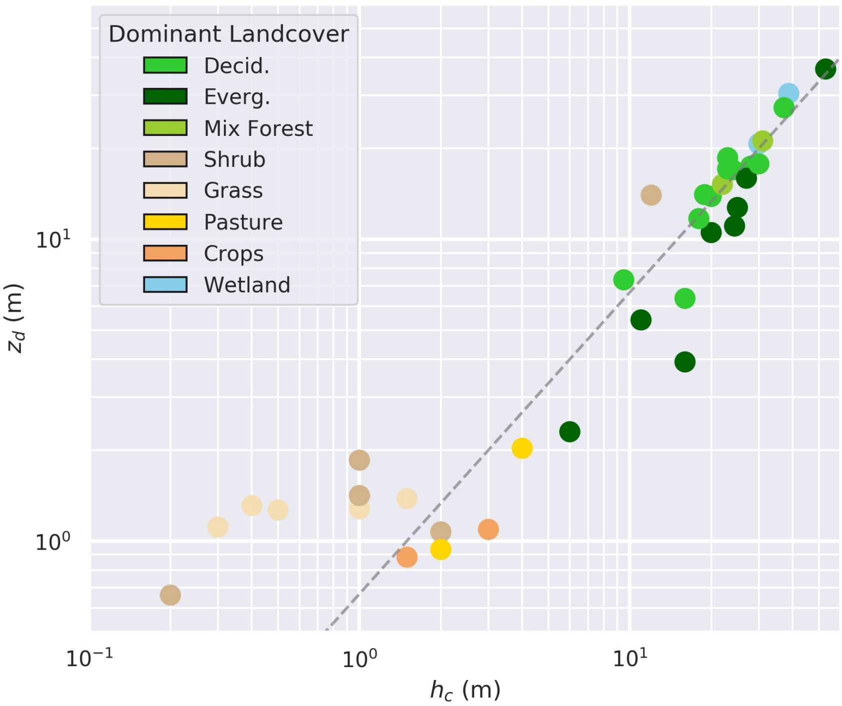

Figure 2. Comparison of NEON reported canopy height (h_c) and mean calculated zero plane displacement height from equation (7) (z_d). Colored according to the dominant land cover type at each site. Dotted line represents $z_d = (2/3)h_c$ relations.

be surprising. A basis for the $z_d = (2/3)h_c$ relation stems from an exponential mean velocity profile characterized by an extinction coefficient $a_c > 1$ inside the canopy as derived from a constant mixing length hypothesis for the turbulent eddy diffusivity (Raupach & Thom, 1981). These arguments, when combined with the drag-force centroid method to estimate z_d for (i) constant drag coefficient and leaf area density and (ii) rigid, tall and dense canopy yield

$$\frac{z_d}{h_c} = 1 - \frac{1}{2a_c} = 1 - \frac{1}{2}\frac{L_s}{h_c}, \tag{8}$$

where $L_s = u/(du/dz)$ evaluated at $z = h_c$ is known as the canopy shear length scale. For the flow near the canopy top to behave as 'mixing layers' requires an inflection point in the mean velocity profile at $z/h_c = 1$ (Raupach et al., 1996). This condition leads to a constraint on $1/2 < L_s/h_c < 1$ thereby bounding z_d/h_c to be between 1/2 and 3/4. All these assumptions (i.e. rigid, tall and dense canopy, constant mixing length within the canopy, etc..) break down for short and sparse canopies (Poggi, Porporato, et al., 2004) as evidenced by the near independence between z_d and h_c in Figure 2 for short h_c.

3.3 Quality Assurance and Quality Control

To ensure that the data are both of high quality and readily applicable, a number of quality assurance steps are applied: (1) All points that fail NEON quality assurance for air temperature are removed, (2) data where the reported energy balance has a residual greater than 20% are removed, as large residuals indicate high likelihood of significant advective fluxes and thus complicate the analysis (Mauder et al., 2020). The 20% threshold was selected to preserve as much data as possible for site-by-site analysis, and no significant difference in data quality or observed trends was noted when tightening this threshold further. (3) All data points with $\zeta > 0$ are removed as uncertainties in this range are high, data availability is relatively low, and this is not the intended focus of the study. (4) Periods with non-negligible precipitation are removed. (5) Any site which, after all previous quality control is applied, retains less than 100 half-hourly runs are removed. Quality control retained just over 32,000 half-hourly runs across 39 NEON sites, roughly equivalent to about 2 site-years at 30-min averaging.

3.4 Random Forest (RF) Method

The RF method is used to generate an initial data-driven alternative model to equations 1 and 2 with no regards to dimensional constraints as required by similarity theory. This allows us to examine empirically impacts that various environmental predictors not present in current MOST based formulations might have on the development of PTV. The RF is a machine learning method that uses ensemble decision trees for producing a regression, with each decision tree run using a random subsample of the data to generate the model (Breiman, 2001). Data are randomly split into testing (10%) and training (90%) datasets for RF as well as other fitting analysis. Accuracy is evaluated primarily using normalized root mean squared error (nRMSE). From the results of the model, we extract feature importance, a measure of which predictors play the largest role in the model fit. In this case, high feature importance indicates that the value of a given predictor is essential for describing and predicting PTV using the RF method. We have elected to use sensible heat, ζ, and u_* due to their role in MOST. Tree cover fraction, bare soil fraction, vegetative fraction, LAI, and effective drag $C_d = [u_*/u(z)]^2$ are used to potentially represent canopy structure and roughness effects. BLH, CAPE, and cloud cover fraction are selected due to their relation with mesoscale phenomenon and large scale eddies.

4 Results

The analysis begins with basic examination of the data across all sites. These are summarized as comparisons between σ_T and the environmental predictors presented in section 3.4, followed by analysis of the diurnal cycle of sensible heat and σ_T. The data are then compared to the curves of (1) and (2). Analysis continues focused on exploring site level differences, first with RF over the entire dataset as well as individually for each site. A bar plot showing the relation between predicted and observed at each site is then featured to illustrate differences between land cover types. The final section of the analysis focuses on evaluating potential model improvements leveraging the results in the previous sections. The observations are compared to Equations 1 and 2 with updated parameter values selected through curve fitting. This comparison is shown over both the overall dataset and a select few sites. Finally, select parameterizations of the b parameter in (1) based on a variety of metrics that represent canopy structure are presented, evaluated, and compared to traditional formulations.

4.1 Holistic Exploration

The data from the remotely sensed products and NEON were merged and then quality controlled as described in section 3.3. Figure 3 presents a comparison between σ_T and collocated environmental and meteorological data. The results show a clear relation between PTV and sensible heat flux $H = \rho C_p \overline{w'\theta'}$ where ρ is the mean air density and C_p is the specific heat capacity of dry air at constant pressure as well as ζ and to a lesser extent effective drag C_d. In addition, some patterns seem apparent with LAI and BLH. Other environmental variables not included in figure 3 have no significant relation with PTV. For H, there appears to be a family of curves rather than one defined shape, implying some additional parameter is influencing that relation. Effective drag, similarly,

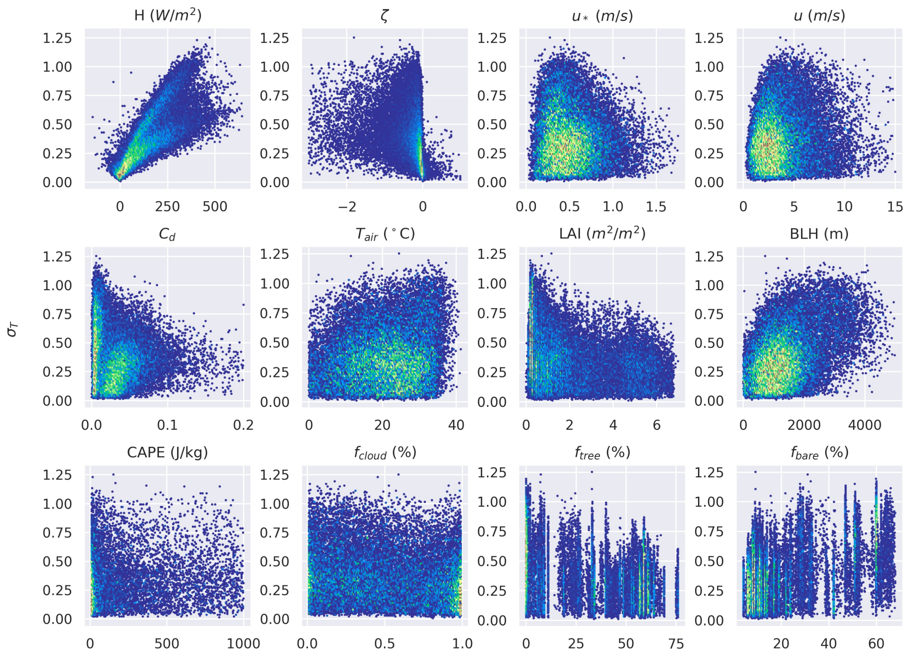

Figure 3. Relation between environmental variables and σ_T across all sites. The resulting scatterplots are binned into small hexagons; the colors illustrate the concentration of points in each hexagon where blue is low and yellow/brown is high.

has two families of curves, with the larger effective drag values arising primarily from forested sites.

The diurnal cycles of PTV and H are plotted in figure 4. In the four forested sites, the raio of H relative to σ_T is higher when compared to the four low lying sites, already suggestive of the importance of site level difference. Figure 4 is also illustrative of the differences between unfiltered and filtered PTV, with the change being most significant in the mornings when sensible heat flux is small but rapid changes in mean air temperature would artificially inflate the apparent PTV caused by turbulence only.

The data covers a range of stability conditions in the near neutral and unstable range, as indicated in figure 5a. The shape of the data generally follows expectations from MOST with an extensive $\zeta^{-1/3}$ scaling (Tillman, 1972) in figure 5b in the unstable range, although in the near neutral range this is less clear. Similarly, in figure 5a, there is some deviation from the established formulation in equation (1), especially as ζ increases in magnitude. Comparing equations (1) and (2) directly to the data show significant errors. Equation (1) has an nRMSE of 21.5% and a 1% bias, although the bias in equation (1) is deceptive as the model has significant negative bias at low values and a positive bias at larger values. Equation (2) performs significantly worse over that range, with an nRMSE of 27.6% and a bias of 15%.

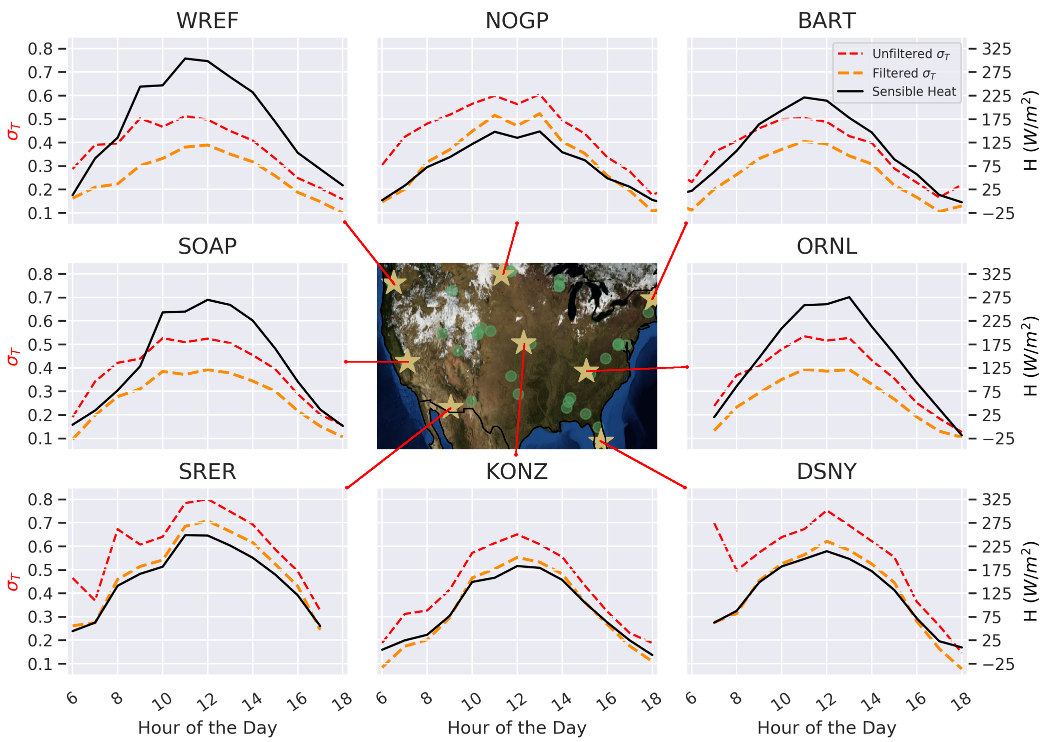

Figure 4. The diurnal cycle of sensible heat flux (H), filtered (red-dashed) and unfiltered (orange-dashed) σ_T for 8 selected sites, with their locations indicated on the central map of CONUS.

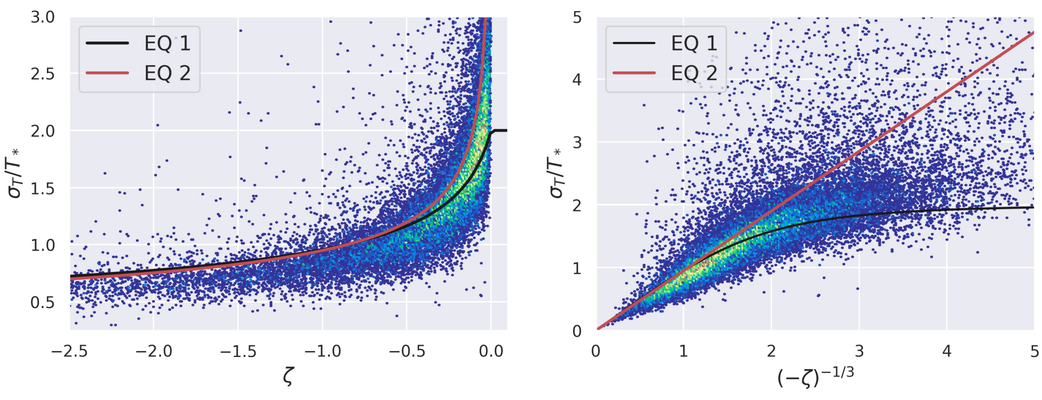

Figure 5. **a**(left): Relation between the dimensionless standard deviation of potential temperature σ_T/T_* and the stability parameter ζ for the data with the modeled values from Equation (1) in black and Equation (2) in red. **b** (right): Relation between the dimensionless standard deviation of potential temperature and $(-\zeta)^{-1/3}$. The resulting scatterplots in both panels are binned into small hexagons; the colors illustrate the number of points in each hexagon where blue is low and yellow/brown is high.

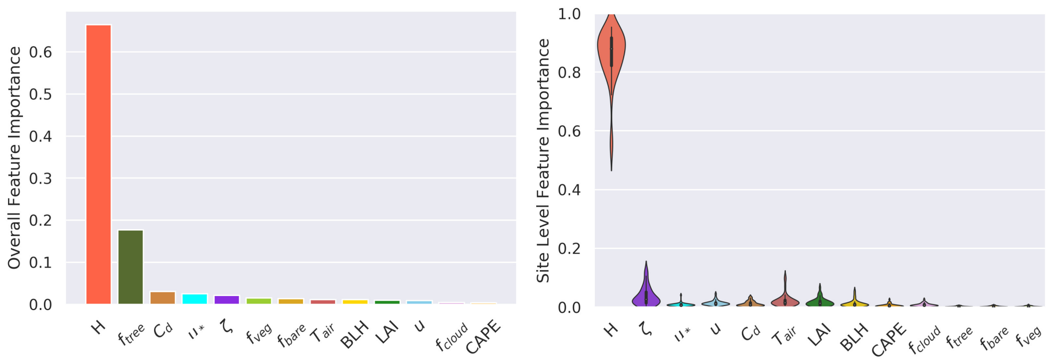

Figure 6. **a** (left): The feature importance from the random forest on the aggregate dataset sorted by overall importance. **b** (right): Results of the site level random forest feature importance. Violin plot shows distribution of site level feature importance for each predictor

4.2 Site by Site Comparison

Random Forest provides a first pass at the potential to improve upon the model when noting deviations from the data in figure 5. RF does perform significantly better than either model with nRMSE of 13.4% and a bias of less than 0.1%, although computational constraints prevent its use in ESMs. The feature importance provides dynamically interesting results as can be seen in figure 6. Sensible heat flux dominates the determination of PTV as expected from flux variance literature. The relative unimportance of friction velocity is also consistent with equation (2) and with equation (1) when the magnitude of ζ is large. The high importance of f_{tree} may be consistent with the results in figure 4 as well, further indicating that tree cover has a significant impact on the relation between sensible heat flux and PTV. Somewhat surprising is the relatively low importance for ζ. Although it is notable that since ζ is a function of H and z_d, which is related to canopy height, a significant portion of the stability effect may be captured by these two aforementioned variables.

The results shown in figures 3, 4, and 6 all indicate the possibility of variable curves for each site in the network. RF was run again, separately, for each individual site to examine these possible relations and remove any attempts by the algorithm to use a predictor as a proxy for the site. The violin plot in figure 6 shows the distribution of the feature importance of each predictor across sites. When examined site by site,H is an even more dominant predictor for PTV. The stability parameter becomes the second most important indicator, consistent with preexisting MOST formulations, although again it is small when compared to sensible heat flux.

When further exploring site level differences, key patterns begin to emerge. When comparing the PTV predicted by equation 1 and the PTV observed at each site, there is a significant variability in the slope of the best fit line of the data, which would ideally sit at 1 indicating close agreement between the observations and the model. Figure 7 illustrates how that slope changes site to site and with land cover type. Sites with slopes close to 1 are generally flat, homogeneous, and dominated by low lying vegetation, which is the ideal landscape for MOST, and matches the landscapes where the values of the parameters a, b and C_1 were originally derived. Forested sites however, especially those with significant heterogeneities, have slopes significantly higher than 1, indicating that the pre-existing model underpredicts PTV at low values and overpredicts PTV at high values. LENO, SJER and SOAP in particular are all sparsely forested sites with significant open water at LENO, an oak savannah at SJER, and sparse evergreens with vary-

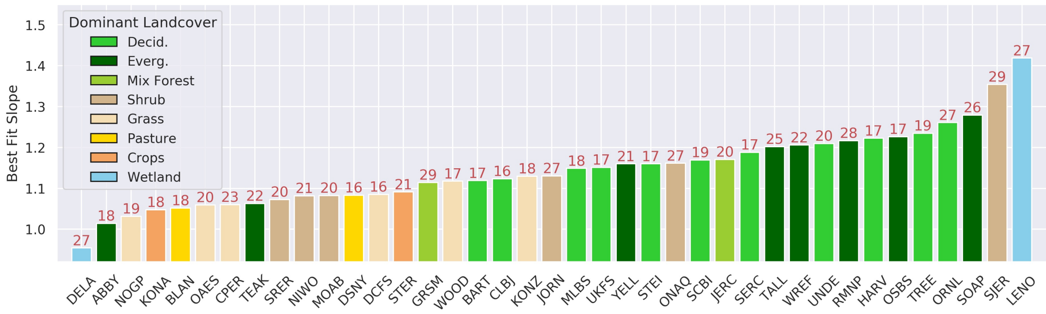

Figure 7. Bar plot showing the best fit slope between the observed and predicted temperature variance at each site using equation 1. Normalized RMSE is also listed in red above each bar. Site bars are colored by the dominant NLCD land cover

ing topography at SOAP. In addition, it is notable that ABBY is a logging site, so while it is classified as evergreen, the actual canopy is quite short, and the tower is located in a clearing. Site sensible heat and Bowen ratio were also examined, and can be seen in Figures S1 and S2 respectively in the supplementary material, however few trends were found outside of poor performance under very low Bowen ratios or sensible heat fluxes.

4.3 Adjusting Existing Models

The results in figure (5) show that there is good predictive value in the existing schema, but also imply that operational adjustments could yield improvements. An iterative fitting process was used to determine the optimal values for the constants a, b, and C_1 over the aggregate data. This global fit resulted in small but non-trivial model improvements for equation (1) and equation (2) both in error and bias. Equation (1) after a global fit to the data nRMSE changes from 20.5% to 17% while bias remains constant and for equation (2) the error is reduced from 27.6% to 18.4% and bias shifts from 15% to -1.4%. These values are compared with additional models in Table 2. Figure 8 illustrates how the newly fitted curves describe the entire data in two ways: dimensional and dimensionless forms. The dimensional form of the comparison was selected because it does not suffer from any self-correlation. Self correlation arises here because H impacts both T_* (ordinate) and ζ (abscissa) in the stability correction function, which can lead to spurious agreement (especially in the exponent). In dimensional form (left panels 8), the σ_T is computed from measured u_*, H, z, and inferred z_d and compared to independently measured σ_T obtained after filtering the high frequency air temperature series. Fitted equation (1) performs better than fitted equation (2) though in dimensional form, this difference appears minor. This difference becomes clear when the two formulations are assessed by stability class and appear to diverge in the near neutral range. In this range, there is much greater uncertainty in the values of σ_T/T_* (though the variances themselves are small). As $T_* \to 0$ but σ_T remains finite due to entertainment of heat and due to finite signal-to-noise ratio in the measurements, σ_T/T_* becomes ill-defined or suspect in equation (1). Interestingly, equation (2) suggests that both the left-hand and right-hand side becomes unbounded as $T_* \to 0$, and thus predicts rapid increase in σ_T/T_* as $|\zeta| \to 0$. While this increase in σ_T/T_* appears to be consistent with some data sets, it is simply a statement that σ_T/T_* may be ill-defined. As such, we will be focusing our analysis on equation (1) in which σ_T/T_* is forced to approach a constant and σ_T maintains its scaling with local sensible heat flux. Last, a realizability constraint was also developed (described later) so as to illustrate a theoretical lower limit for the applicability of equations (1) and (2). Figure 8 demonstrates that the majority of the observations (in dimensionless form) as well as

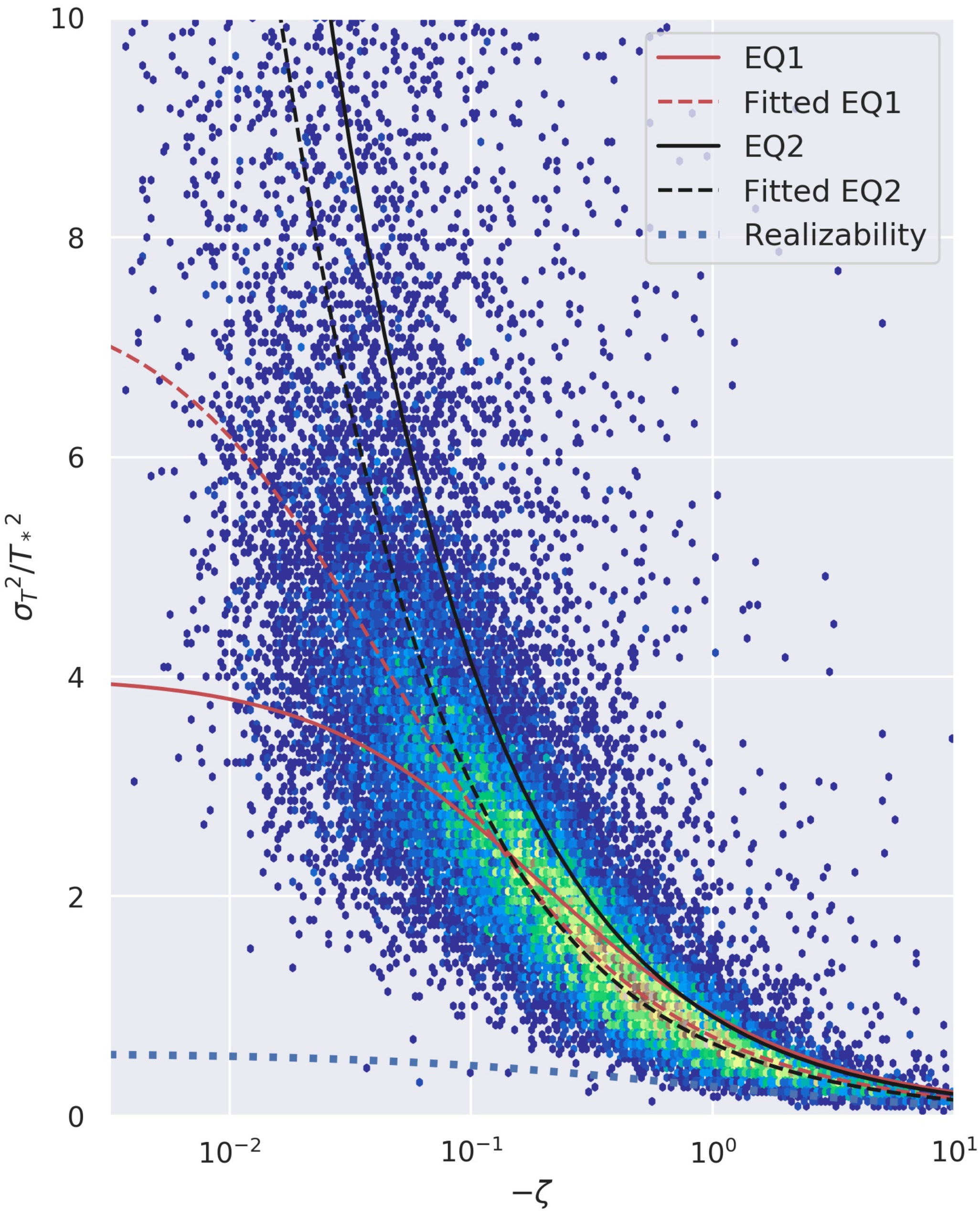

Figure 8. (upper left): Comparison between observed and predicted temperature variance by equation (1) after a global fit (a=7.5, b=34.0) over all sites. **(bottom left)**: Comparison between observed and predicted temperature variance by equation (2) after a global fit (=.812) over all sites. Note the comparisons in the left panels do not suffer from self-correlation. **(right)**: the stability correction function for the non-dimensional temperature variance σ_T^2/T_*^2. The original forms of equation (1) (a=4, b=8.3) and equation (2) (=.95) are shown as well as the fitted versions of both equations. In addition, the limit imposed by the realizability constraint is featured. The resulting scatterplots in all three are binned into small hexagons; the colors illustrate the number of points in each bin where blue is low and yellow/brown is high.

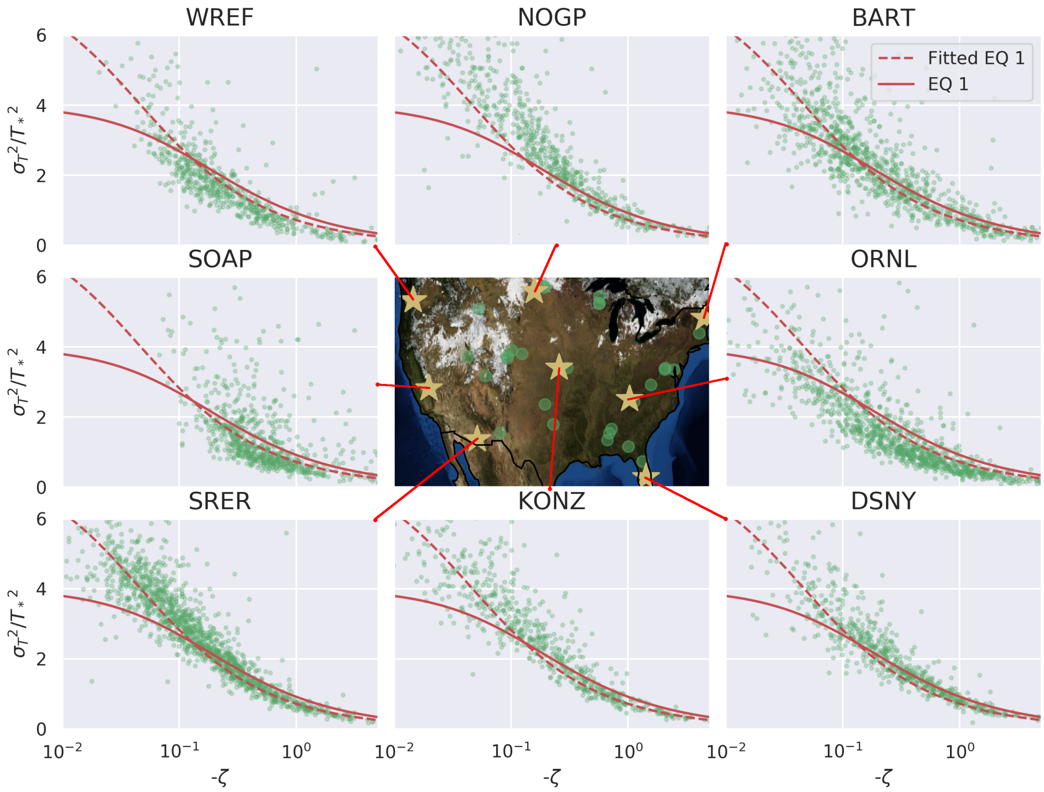

Figure 9. Stability correction function for the temperature variance at selected sites. In red, the modeled results are plotted for both the original Equation 1 (a=4, b=8.3) and the global fit Equation 1 (a=7.5, b=34) as well.

the changes in parameter values remain above the line and satisfy this realizability constraint.

The global fit does not perform universally well at each site, although most sites realize some improvements. Figure 9 shows the global fit and original equation (2) over select sites as well as scatterplots of the data, similar to the hexbin scatterplot for the aggregate data in figure 8c. At three of the sites, ORNL, WREF, and SOAP, the data lies largely below both the original equation (1) curve as well as the fitted curve. These sites all have significant forest cover, especially compared to the 4 sites where the data lies largely above the fitted curve, NOGP, SRER, KONZ and DSNY, which are all flat sites with only bare soil or low-lying vegetation. The performance of these site by site fits are presented in Table 2.

These site level differences can also be examined more quantitatively. The fitting exercise was repeated, this time doing a separate fit of the b parameter for each site while holding a constant. The a is held constant and b is adjusted because under highly convective conditions, which are the main conditions of interest, b is the dominant parameter whereas a dominates in the more uncertain near-neutral range. After comparing b and other environmental predictors, it became clear that there is a close relation between a variety of measurements of canopy structure around the tower and the best fit value of the b parameter. As shown in Figure 10, there is a clear linear relation between the cube root of b and the different measurements of vegetative structure: canopy height, leaf area density (LAI/h_c), effective drag C_d, z/z_d, LAI/z_d, and $h_c^2/[(z-z_d)^2 LAI]$.

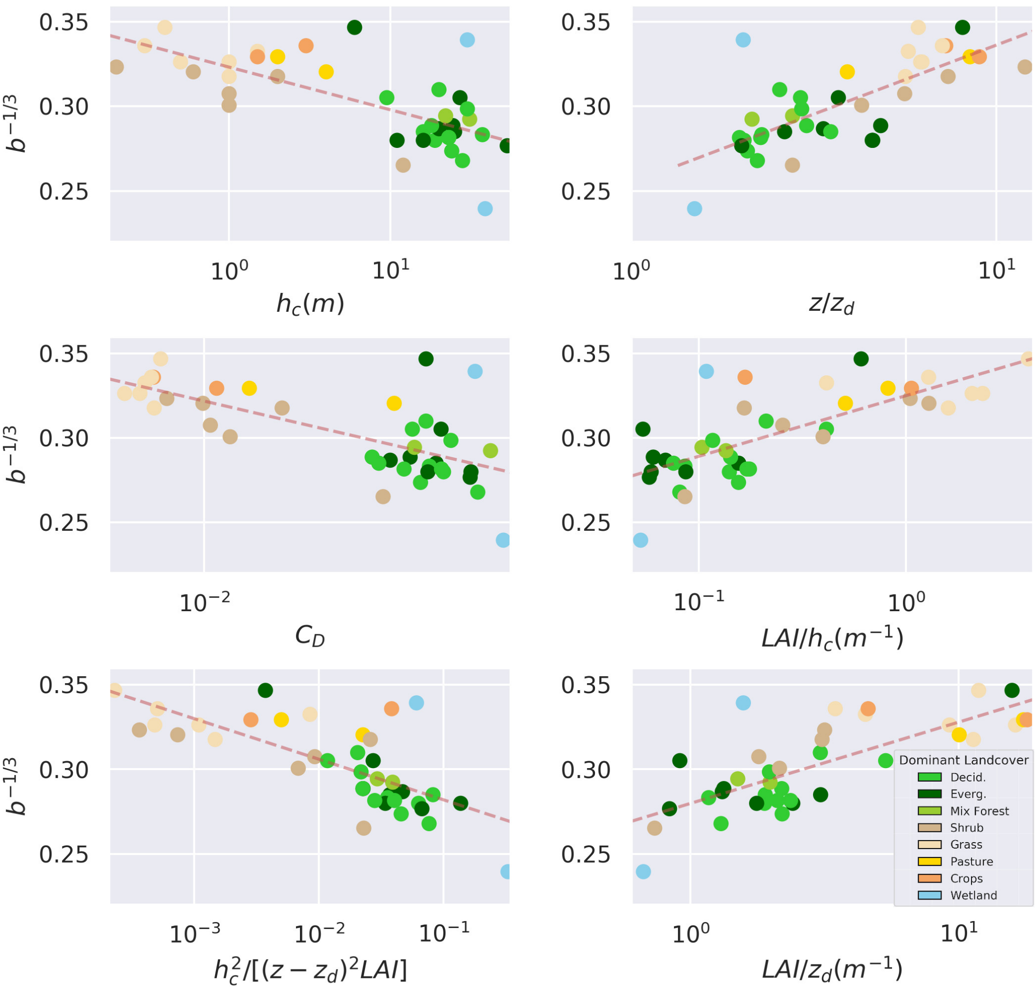

Figure 10. Scatterplot of the selected predictors at each site on a log scale compared to the best fit value for the b parameter from equation (1). The points are colored according to the dominant NLCD landcover type

Table 2. Summary table of the results of various selections for the parameters of equations 1 and 2 as well as the random forest model. The table includes the normalized RMSEs and normalized biases of the different models for PTV.

Equation 1				
Model	a	b	nRMSE	nBias
Standard	4	8.3	20.5%	1.4%
Global Fit	7.5	34	17.0%	-2.7%
Site by Site Fit	7.5	20-80	15.5%	-3.0%
Equation 2				
Model	C_1		nRMSE	nBias
Standard	0.95		27.6%	15%
Global Fit	0.812		18.4%	1.4%
Random Forest				
Global	——		13.4%	$< 0.1\%$

Table 3. Summary of a variety of possible parameterizations of the b parameter in equation (1) following the form of equation (9)

χ	α	β	nRMSE	nBias	R^2
LAI/h_c	0.036	0.289	16.2%	-5.0%	0.55
LAI/z_d	0.036	0.284	16.3%	-4.8%	0.57
z/z_d	0.042	0.277	16.7%	-3.6%	0.54
C_d	-0.047	0.228	16.8%	-2.9%	0.46
$h_c^2/[(z-z_d)^2 LAI]$	-0.024	0.258	16.4%	-4.7%	0.52
h_c	-0.025	0.323	16.5%	-3.4%	0.5

Taking advantage of this relation between b and the various proxies for canopy structure around the tower, a linear model was developed for b based on a linear regression. Applying the model in Equation (6) to update the parameters in Equation (1) with the existing data yields only marginal improvements on the updated Equation (1) based on globally fit parameters, as is clear in Table 3. The relation

$$b^{-1/3} = \alpha \log(\chi) + \beta \tag{9}$$

does suggest that as canopy height increases, b also increases thereby amplifying the modulations introduced by ζ that act to reduce the dimensionless temperature variance. Hence, it appears that tall canopies make the dimensionless temperature variance more sensitive to ζ.

5 Discussion

5.1 Results Summary and Synthesis

Overall, the filtered data indicates agreement with MOST based formulations (Tillman, 1972). While this study emphasizes that there is room for improvement, it is also important to note that equation (1) holds even over non-idealized landscapes and atmospheric conditions despite the fact that the formulation and parameter values were derived over highly idealized flows. The evaluation of equation (1) over these wide-ranging landscapes offers one of the clearest pictures in the literature of its broad applicability for understanding PTV at the bottom of the surface layer (provided non-turbulent phenomenon are filtered). From the results of the random forest, and inspection of figure 2, it is evident that other local physical and meteorological characteristics that were thought to have some influence on the development of PTV in the surface boundary are largely unimportant. Heat flux and local stability continue to be the driving factors and can yield good predictions for PTV over flat landscapes using the standard parameter values. Based on model error analysis, there is additional uncertainty to be captured.

Numerous studies have shown how parameter values for various local sites can deviate from the global values described in the early literature, however few have proposed updates to models used in ESMs as these studies often include only a very small number of sites and therefore painted a limited picture of the variety that one can find in the field. Site by site fitting to the parameter values indicate that most sites have parameter values larger than those defined in the literature, and only one site was found to have parameter values smaller. Since the best fit values of the parameters across landscapes do not oscillate around these 'ideal' values, but rather are all greater than or equal to them, ESMs can benefit from alternative global parameters to cover regions with varied and heterogeneous canopies. The inter-site best fit parameter variation is quite significant, with best fit values of b ranging from 20 to 80. This suggests that, while global parameter values may be useful for broad application, localized studies will benefit most from a local, site based empirical fit, especially if they deviate from ideal (i.e. flat, homogeneous, short vegetation) surfaces. An important note with respect to these model fit values is the role of the filtering process. The filtering process described yielded closer agreement to MOST for all of these model fits; unfiltered data overall yields slightly more noise and a greater deviation from traditional MOST relations with a larger magnitude of bias, but maintains inter site trends with unfiltered data.

Attempts to use environmental predictors to capture the local variation were only marginally successful outside of a random forest model. The RF method detailed significant improvement, and was able to capture most of the inter-site variability based on the tree cover fraction. This implies that canopy structure and surface roughness characteristics are partly responsible for a significant portion of the deviations from ideal conditions. The RF method, however, is too computationally intensive for application in ESMs. As such, there was an attempt to generate a compact model for the b parameter based on environmental variables related to surface roughness and canopy. Of a long list of possible predictors to model values of b, the most successful are shown in figure 10. A plausibility argument for the inclusion of some of these variables to parameterize b may be obtained by examining qualitatively the potential temperature variance budget (PTV).

5.2 PTV Budget: A scaling analysis

Using index notation, the PTV above the canopy is given by

$$\underbrace{\frac{1}{2}\frac{\partial\overline{\theta'^2}}{\partial t}}_{\text{I}}+\underbrace{\frac{1}{2}\overline{U_j}\frac{\partial\overline{\theta'^2}}{\partial x_j}}_{\text{II}}=\underbrace{-\overline{u_j'\theta'}\frac{\partial\overline{\theta}}{\partial x_j}}_{\text{III}}-\underbrace{\frac{1}{2}\frac{\partial\overline{u_j'\theta'^2}}{\partial x_j}}_{\text{IV}}-\underbrace{\epsilon_\theta}_{\text{V}}, \tag{10}$$

where $i = 1, 2, 3$ indicates longitudinal (x_1 or x), lateral (x_2 or y), and vertical (x_3 or z) directions, respectively, repeated indicies imply summation, t is time, and u_i' are velocity fluctuations along direction x_i from the mean value $\overline{U_i}$. In this budget, term I is the local variance storage, which as an outcome of the filtering exercise, can be neglected. Term II is advection of PTV by the mean flow. In the analysis here, it is assumed that subsidence ($\overline{U_3} = 0$) is small and due to the choice of coordinate systems, the mean lateral velocity is zero ($\overline{U_2} = 0$). Term III is the production due to the turbulent heat fluxes and mean potential temperature gradients, which we asssume to be driven by the vertical direction components. Term IV is the turbulent transport term. This term is significant in some cases, but the filtering process used here will remove some portion of the non-local heat transport effects captured by this term. Hence, for simplicity, this term is momentarily ignored. Term V is the molecular dissipation of PTV, which, along with term III, tends to comprise the largest portion of the budget (Champagne et al., 1977; Monji, 1973). There are additional terms, not presented here, that represent radiative destruction, conductive diffusion and another dissipation term. These terms are generally considered negligible near the surface. With these simplifications,

$$\frac{1}{2}\overline{U}\frac{\partial\overline{\theta'^2}}{\partial x}=-\overline{w'\theta'}\frac{\partial\overline{\theta}}{\partial z}-\epsilon_\theta. \tag{11}$$

To proceed further, closure schemes and scaling analyses are needed for the mean advection and variance dissipation terms, which are given as

$$\epsilon_\theta=C_{\epsilon,\theta}\frac{\overline{\theta'^2}}{\tau};\frac{1}{2}\overline{U}\frac{\partial\overline{\theta'^2}}{\partial x}=-C_{adv,\theta}\frac{\overline{\theta'^2}}{\tau_{adv}};\tau_{adv}=\frac{L_x}{\overline{U}}, \tag{12}$$

where $C_{\epsilon,\theta}$ and $C_{adv,\theta}$ are closure constants, τ is a relaxation time scale describing how long a potential temperature excursion lasts before it gets dissipated by molecular processes, τ_{adv} is an advection time scale formed by the local mean velocity at z and a potential temperature spatial variability integral length scale L_x. This length scale reflects imprints of 'near-field' heat sources from tree crowns or vegetation upper layers not 'blended out' by turbulence within the roughness sublayer above the canopy. When the spatial imprint of these heterogeneous heat sources is entirely blended out by turbulence mixing, $L_x \to \infty$. Inserting these closure schemes into the simplified PTV budget and after some algebra results in

$$\frac{\overline{\theta'^2}}{T_*^2}=-\frac{u_*\tau}{\kappa(z-z_d)}\frac{\phi_h(\xi)}{C_{\epsilon,\theta}}\frac{1}{\left(1-\frac{C_{adv,\theta}}{C_{\epsilon,\theta}}\frac{\tau}{\tau_{adv}}\right)};\phi_h(\xi)=-\frac{d\theta}{dz}\frac{\kappa(z-z_d)}{T_*}\phi_h^*(\xi), \tag{13}$$

where $\phi_h(\xi)$ is the stability correction function for temperature defined in the roughness sublayer defined using a surface layer representation adjusted by $\phi_h^*(\xi)$, a roughness sublayer modification. The $\phi_h^*(\xi) \to 1$ as $|\xi| \to \infty$. As discussed previously, over forested terrain and those with high heterogeneity in canopy structure, equations (1) and (2) have a tendency to over predict temperature variance unless the fit is adjusted (in the case of C_1, a lower value). If the simplified version of the PTV budget is examined, then this underestimation emerges from competition between three terms: (i) the finite ratio of τ/τ_{adv} increasing the normalized PTV, (ii) the role of $\phi_h^*(\xi) \subset [0,1]$ reducing the normalized PTV, and (iii) the reduced τ relative to its surface layer value. The L_x is finite in the roughness sublayer and may be commensurate with crown size, crown-to-crown spacing, and measurement height z (i.e. more blending with increasing z), which may explain why z/z_d emerges as an explanatory variable to b. Moreover, tree spacing and canopy height are indirectly captured by canopy LAI, another variable impacting b. The finite L_x here is shown to increase normalized PTV above its surface layer similarity value, not reduce it. The two processes that act to reduce the normalized PTV are $\phi_h^*(\xi)$ and τ. The $\tau = TKE/\epsilon$ may be smaller than predicted by its surface layer value ($\propto (z - z_d)/u_*$), where TKE and ϵ are the turbulent kinetic energy and its dissipation rate. This underestimation of τ arises because of an enhancement of ϵ near the canopy top relative to predictions from extrapolations of the surface layer similarity value (Poggi, Katul, & Albertson, 2004). Likewise, roughness layer corrections to $\phi_h(\xi)$ are usually smaller than unity, meaning that $\phi_h(\xi)$ in the roughness sublayer must be smaller than its surface layer counterpart (Garratt & Segal, 1988; Harman, 2012).

In addition to possible reductions in C_1 due to τ, and $\phi_h(\xi)$, deviations from equation (1) due to large scale eddies was also proposed as a source of uncertainty. This effect is primarily captured by term (IV) in the PTV budget. This term was neglected in the analysis here primarily because of the filtering of temperature time series employed to minimize the effect of term I (the storage term) and likely have filtered some of the very large scale inactive eddies that can contribute to PTV but not T_*. Thus, the effect of large-eddies in the ABL is to increase the normalized PTV, not reduce it. Using the 5-min filtered series in PTV calculations, the RF algorithm here paints an unclear picture that neither confirms this hypothesis nor offers a clear refute. Figure 3 shows that there is a weak linear relation between BLH and the temperature variance, suggesting that when the ABL is large, additional external sources of variance could exist, introducing new length scales not represented in MOST for the PTV budget. The relation in Figure 3, however, is weak. Figure 6a and 6b also support this weak relation. The three predictors included in the RF to represent these effects, BLH, cloud cover fraction, and CAPE, all have a feature importance of less than 1%. These three predictors, however, do have their limitations and the weak relation shown here does not clearly refute the hypothesis. The data comes from a reanalysis product with relatively poor resolution, meaning that measurements of these values, particularly BLH are not necessarily reliable. Another important note is the specific selection of time periods where the surface energy balance is largely closed. Previous literature indicates that sub-mesoscale circulations (usually of time scales longer than several minutes) may cause the non-closure of the surface energy balance (Mauder et al., 2020), which means that by virtue of constraining the study to a closed energy balance we may be excluding the study periods where these circulations would have an impact. Likewise, the removal of time scales longer than 5 min may also ameliorate sub-mesoscale circulations. Examining the PTV budget equations, one could see how advection may increase PTV and the scatter of PTV as well if included. Upon initial examination, the data appears to support this hypothesis with greater observed scatter of PTV, although a more comprehensive analysis is outside the scope of this manuscript. Additional studies, with more reliable and locally relevant measurements of BLH such as through surface to air LIDAR as well as consideration of the surface energy balance, are required to adequately assess this hypothesis.

5.3 Realizability Constraint

Equations 1 and 2 must satisfy the realizability constraint requiring that θ' and w' must not be perfectly correlated resulting in the inequality

$$\sigma_w^2 \sigma_T^2 > (\overline{w'\theta'})^2 = (T_* u_*)^2 \tag{14}$$

and

$$\frac{\sigma_w^2}{u_*^2} \frac{\sigma_T^2}{T_*^2} > 1 \tag{15}$$

Paired with the original formulation of equation (1) are equivalent MOST consistent forms for σ_w^2/u_*^2 (Andre et al., 1978)

$$\frac{\sigma_w^2}{u_*^2} = 1.75 + 2(-\zeta)^{2/3} \tag{16}$$

Thus, combining equation (15) and equation (16) we obtain

$$\frac{\sigma_T^2}{T_*^2} > \frac{1}{\frac{\sigma_w^2}{u_*^2}} = \frac{1}{(1.75 + 2(-\zeta)^{2/3})} \tag{17}$$

Equation (17) is plotted as part of figure 8c, which clearly illustrates that the model adjustments proposed do not violate any realizability constraint.

5.4 Roughness Sub-Layer Effects

A significant source of concern when analyzing above canopy PTV and how it compares between grasslands and forested areas is the thickness of the roughness layer for heat. It is difficult to ensure that the towers are reporting flow statistics outside of the roughness layer and in the inertial layer, where MOST scaling is intended to apply. NEON towers are designed to lie above the roughness layer (i.e. in the surface layer), and mean wind profiles at the sites indicate that most sites are within the surface layer based on momentum considerations. In addition, past studies have indicated that the thickness of the momentum roughness layer and a roughness layer for scalar quantities such as potential temperature and water vapor are not necessarily the same. For scalar quantities, the roughness layer can be significantly thicker than those for momentum upon which the tower design is based (Raupach & Thom, 1981). If points are indeed interrogated inside the scalar roughness sublayer, this could yield significant changes in the values of temperature variance as the canopy and surface elements play a greater role in introducing variable heat sources and sinks. Related to this is a concern inherent in the design; towers over forested sites in NEON are designed as a factor of the canopy height, but over low-lying vegetation it is defined as simply a constant 8m. This means that the instrumentation may lie further up in a normalized profile for some of the flatter sites than the forested ones, potentially explaining some of the differences between these two categories. Initial exploration does show a poor but persistent relation between the ratio of tower height to canopy height and the model parameter values. These challenges with defining where the instrumentation lies above the canopy, however, is of lesser concern for the primary intended application of this study in earth system models, where the roughness layer is inconsistently defined across different models.

5.5 Future Work

A robust evaluation of the primary models of PTV at the surface layer was undertaken and avenues for improvements proposed. In addition to the importance of refining models over sparse canopies, as discussed in the previous section, exploration of these models and PTV more generally in stable and, to a lesser extent, near neutral atmospheric regimes is needed. The model currently assumes that the non-dimensional variance remains constant with stability, although when exploring the data this became less clear. Previous studies have suggested high errors in the near neutral range is a consequence of the non-stationarity (Kroon & de Bruin, 1995; J. Wyngaard & Coté, 1971). This effect may be partly ameliorated by spectral filtering as shown. Yet, the data scatter at the near neutral limit is undisputed. Significant scatter also exists in the mildly unstable range, implying that there are issues with the application of MOST under these conditions that require more work.

An important limitation to note is that this analysis is all based on one point within a tower reading from a flux footprint covering areas on the order of a few square kilometer. In ESMs, PTV is computed at the tile level which is intended to be representative of the landscape level but, depending upon the tower height, can be on a larger scale and is an area rather than point measurement. This scaling issue, which to the best of the authors knowledge has not been examined previously, may have an impact especially when considered in light of circulations and advection of PTV, where tiling schemes may fail to be representative of the landscape.

There are three other avenues of future research that require further exploration. First, while this study focused on the case where the energy balance is closed and there is no significant advection, unbalanced conditions where sensible heat, latent heat and ground heat flux fail to account for the energy balance constitute a large fraction of the data. Initial work shows a clear shift in the fit of the data to Equation 1 under these conditions, with lower best fit parameter values and larger scatter. Unfortunately, exploring the potential effect of significant advection require model simulations or data not available through NEON. The second avenue of future research is examining the analogous models for the other primary atmospheric scalar, water vapor. The model assumes that temperature and moisture behave similarly, with the same parameter values. Numerous studies, as well as initial examination of the NEON data, illustrate that water vapor and temperature do not behave identically (G. G. Katul & Hsieh, 1999; Asanuma & Brutsaert, 1999; De Bruin et al., 1993; Liu et al., 2021) in the surface layer as previously theorized, and as such an alternative model, or at the very least, alternative values for the a and b parameters are needed. Finally, results and previous literature have indicated that surface heterogeneity, especially in heating on scales large enough to induce circulations, can have a significant impact on MOST derived parameterizations such as the ones discussed here. A brief examination, not presented here, implies there is a complex relation between heterogeneity and temperature variance statistics, and as such additional work considering different length scales of surface heterogeneity may indicate new directions for improvement and model analysis.

6 Conclusion

High frequency time series across 39 similarly instrumented sites covering varied landscapes across CONUS were analyzed to assess the validity of existing models for temperature variance in the surface layer, note key deficiencies and recommend avenues for improvement. Results indicated that conventional flux-variance similarity formulations are largely corroborated by data in both dynamic-convective and nearly convective cases provided non-turbulent features are spectrally filtered out. This filtering reduced the temperature variance by factors of up to 2 to 3 in some cases when compared to the unfiltered runs. The most significant deviations from standard MOST formulations were observed over heterogeneous and forested sites. Site by site analysis also revealed bias towards similarity constants larger than the traditional parameter values used in the literature and ESMs. A random forest model illustrated that there is variability not captured by the traditional formulations. Results generally indicate that canopy structure, surface heterogeneity, and roughness characteristics drive a portion of the inter-site variability, although a dimensional approach was unable to illustrate superior predictive value. Future studies will expand this analysis to include situations with non-local energy balance closure as well as landscapes with sparse canopies or large surface heterogeneities. Water vapor and carbon dioxide concentration, the other primary atmospheric scalars, use the same formulation in CLUBB and other models as PTV, although the literature shows a difference in behavior. As such, any updated parameter values for temperature cannot be applied to other scalars and additional work is required to make similar improvements to their variance fluctuations.

Open Research

The ERA 5 Reanalysis data (Hersbach et al., 2018) was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. MODIS Vegetative cover data (DiMiceli et al., 2015) and Leaf Area Index (Myneni et al., 2015) is available through https://lpdaac.usgs.gov/. NEON turbulence data (National Ecological Observatory Network (NEON), 2021) is available through https://data.neonscience.org/data-products/DP4.00200.001. Finally, Land Cover types (2016 Version) from the National Land Cover Database (Dewitz, 2019) are avaialble from the Multi-Resolution Land Characteristics Consortium database https://www.mrlc.gov/data. Software used to process this data (Waterman, 2021) and generate results can be found here: https://tinyurl.com/tswneon.

Acknowledgments

This research has been supported by the National Oceanic and Atmospheric Administration, Climate Program Office grant nos. NA19OAR4310241

References

Albertson, J. D., Parlange, M. B., Katul, G. G., Chu, C., Stricker, H., & Tyler, S. (1995). Sensible Heat Flux From Arid Regions: A Simple Flux-Variance Method. *Water Resources Research*, *31*(4), 969–973. doi: 10.1029/94WR02978

Andre, J., De Moor, G., Lacarrere, P., Therry, G., & Du Vachat, R. (1978). Modeling the 24hr Evolution of the Mean and Turbulent Structures of the Planetary Boundary Layer. *Direction de la Meteorologie*, *35*, 1861–1883.

Antonia, R. A., Chambers, A. J., & Bradley, E. F. (1981). Temperature Structure in the Atmospheric Surface Layer. *Boundary-Layer Meteorology*, *20*, 293–307.

Asanuma, J., & Brutsaert, W. (1999). Turbulence variance statistics of Temperature and humidity in the unstable atmospheric surface layer above a variable pine forest. *Water Resources Research*, *35*(2), 515–521. doi: 10.1029/1998WR900051

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32. doi: doi.org/10.1023/A:1010933404324

Brunet, Y. (2020). Turbulent Flow in Plant Canopies: Historical Perspective and Overview. *Boundary-Layer Meteorology*, *177*, 315–364. doi: 10.1007/s10546-020-00560-7

Champagne, F. H., Friehe, C. A., LaRue, J. C., & Wyngaard, J. (1977). Flux Measurements, Flux Estimation Techniques, and Fine-Scale Turbulence Measurements in the Unstable Surface Layer Over Land. *Journal of Atmospheric Sciences*, *34*(3), 515-530. doi: 10.1175/1520-0469(1977)034⟨0515:FMFETA⟩2.0.CO;2

Cheng, A., & Xu, K. M. (2015). Improved low-cloud simulation from the community atmosphere model with an advanced third-order turbulence closure. *Journal of Climate*, *28*(14), 5737–5762. doi: 10.1175/JCLI-D-14-00776.1

Cohen, A. E., Cavallo, S. M., Coniglio, M. C., & Brooks, H. E. (2015). A review of planetary boundary layer parameterization schemes and their sensitivity in simulating southeastern U.S. cold season severe weather environments. *Weather and Forecasting*, *30*(3), 591–612. doi: 10.1175/WAF-D-14-00105.1

De Bruin, H. A., Kohsiek, W., & Van Den Hurk, B. J. (1993). A verification of some methods to determine the fluxes of momentum, sensible heat, and water vapour using standard deviation and structure parameter of scalar meteorological quantities. *Boundary-Layer Meteorology*, *63*(3), 231–257. doi: 10.1007/BF00710461

Detto, M., Katul, G., Mancini, M., Montaldo, N., & Albertson, J. D. (2008). Surface heterogeneity and its signature in higher-order scalar similarity relationships. *Agricultural and Forest Meteorology*, *148*(6-7), 902–916. doi: 10.1016/j.agrformet.2007.12.008

Dewitz, J. (2019). *National land cover database (nlcd) 2016 products: U.s. geological survey data release.* [Dataset] U.S. Geological Survey. doi: 10.5066/P96HHBIE

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., & Townshend, J. (2015). *Mod44b modis/terra vegetation continuous fields yearly l3 global 250m sin grid v006.* [Dataset] NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MOD44B.006

Foken, T. (2006). 50 years of the Monin-Obukhov similarity theory. *Boundary-Layer Meteorology*, *119*(3), 431–447. doi: 10.1007/s10546-006-9048-6

Garratt, J. R. (1992). *The Atmospheric Boundary Layer.* Cambridge University Press.

Garratt, J. R., & Segal, M. (1988). On the contribution of atmospheric moisture to dew formation. *Boundary-Layer Meteorology*, *45*(3), 209–236. doi: 10.1007/BF01066671

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., ... Zhu, Q. (2019). The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2089–2129. doi: 10.1029/2018MS001603

Golaz, J. C., Larson, V. E., & Cotton, W. R. (2002). A PDF-based model for boundary layer clouds. Part I: Method and model description. *Journal of the Atmospheric Sciences*, *59*(24), 3540–3551. doi: 10.1175/1520-0469(2002)059⟨3540:APBMFB⟩2.0.CO;2

Hang, C., Nadeau, D. F., Pardyjak, E. R., & Parlange, M. B. (2018). A comparison of near-surface potential temperature variance budgets for unstable atmospheric flows with contrasting vegetation cover flat surfaces and a gentle slope. *Environmental Fluid Mechanics*, *20*. doi: 10.1007/s10652-018-9647-z

Harman, I. N. (2012). The Role of Roughness Sublayer DynamicsWithin Surface Exchange Schemes. *Boundary-Layer Meteorology*, *142*(1), 1–20. doi: 10.1007/s10546-011-9651-z

Haugen, D. A., Kaimal, J. C., & Bradley, E. F. (1971). An experimental study of Reynolds stress and heat flux in the atmospheric surface layer. *Quarterly Journal of the Royal Meteorological Society*, *97*(412), 168–180. doi: 10.1002/qj.49709741204

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, J.-N. (2018). *Era5 hourly data on single levels from 1979 to present.* [Dataset] Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi: 10.24381/cds.adbb2d47

Huang, H. Y., & Margulis, S. A. (2010). Evaluation of a fully coupled large-eddy simulation-land surface model and its diagnosis of land-atmosphere feedbacks. *Water Resources Research*, *46*(6), 1–19. doi: 10.1029/2009WR008232

Jin, S., Homer, C., Yang, L., Danielson, P., Dewitz, J., Li, C., ... Howard, D. (2019). Overall methodology design for the united states national land cover database 2016 products. *Remote Sensing*, *11*(24). doi: 10.3390/rs11242971

Kader, B. A., & Yaglom, A. M. (1990). Mean fields and fluctuation moments in unstably stratified turbulent boundary layers. *Journal of Fluid Mechanics*, *212*(151), 637–662. doi: 10.1017/S0022112090002129

Katul, G., Goltz, S. M., Hsieh, C. I., Cheng, Y., Mowry, F., & Sigmon, J. (1995). Estimation of surface heat and momentum fluxes using the flux-variance method above uniform and non-uniform terrain. *Boundary-Layer Meteorology*, *74*(3), 237–260. doi: 10.1007/BF00712120

Katul, G. G., & Hsieh, C.-i. (1999). A note on the flux-variance similarity relationships for heat and water vapour in the unstable atmospheric surface layer. *Boundary-Layer Meteorology*, *90*(2), 327–338.

Kiely, G., Albertson, J. D., Parlange, M. B., & Eichinger, W. E. (1996). Convective scaling of the average dissipation rate of temperature variance in the atmospheric surface layer. *Boundary-Layer Meteorology*, *77*(3-4), 267–284. doi: 10.1007/BF00123528

Kroon, L. J., & de Bruin, H. A. (1995). The Crau field experiment: turbulent exchange in the surface layer under conditions of strong local advection. *Journal of Hydrology*, *166*(3-4), 327–351. doi: 10.1016/0022-1694(94)05092-C

Larson, V. E. (2017). *CLUBB-SILHS: A parameterization of subgrid variability in the atmosphere.* Retrieved from `http://arxiv.org/abs/1711.03675`

Lee, Y. H. (2009). The influence of local stability on heat and momentum transfer within open canopies. *Boundary-Layer Meteorology*, *132*(3), 383–399. doi: 10.1007/s10546-009-9405-3

Li, D., Katul, G. G., & Gentine, P. (2016). The k-1 scaling of air temperature spectra in atmospheric surface layer flows. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 496–505. doi: 10.1002/qj.2668

Li, Q., Gentine, P., Mellado, J. P., & Mccoll, K. A. (2018). Implications of nonlocal transport and conditionally averaged statistics on Monin-Obukhov similarity theory and Townsend's attached eddy hypothesis. *Journal of the Atmospheric Sciences*, *75*(10), 3403–3431. doi: 10.1175/JAS-D-17-0301.1

Liu, C., Liu, H., Huang, J., & Xiao, H. (2021). Varying Partitioning of Surface Turbulent Fluxes Regulates Temperature-Humidity Dissimilarity in the Convective Atmospheric Boundary Layer. *Geophysical Research Letters*, *48*(21), 1–10. doi: 10.1029/2021GL095836

Lloyd, C. R., Culf, A. D., Dolman, A. J., & Gash, J. H. (1991). Estimates of sensible heat flux from observations of temperature fluctuations. *Boundary-Layer Meteorology*, *57*(4), 311–322. doi: 10.1007/BF00120051

Lock, A. P., Martin, G. M., Bush, M. R., Brown, A. R., & Smith, R. N. (2000). A new boundary layer mixing scheme part II: Tests in climate and mesoscale models. *Monthly Weather Review*, *128*(9), 3200–3217. doi: 10.1175/1520-0493(2000)128⟨3200:ANBLMS⟩2.0.CO;2

Maronga, B., & Reuder, J. (2017). On the formulation and universality of Monin-Obukhov similarity functions for mean gradients and standard deviations in the unstable surface layer: Results from surface-layer-resolving large-eddy simulations. *Journal of the Atmospheric Sciences*, *74*(4), 989–1010. doi: 10.1175/JAS-D-16-0186.1

Mauder, M., Foken, T., & Cuxart, J. (2020). *Surface-Energy-Balance Closure over Land: A Review* (Vol. 177) (No. 2-3). doi: 10.1007/s10546-020-00529-6

Mcnaughton, K. G. (2006). On the kinetic energy budget of the unstable atmospheric surface layer. *Boundary-Layer Meteorology*, *118*(1), 83–107. doi: 10.1007/s10546-005-3779-7

Metzger, S., Ayres, E., Durden, D., Florian, C., Lee, R., Lunch, C., ... Zulueta, R. C. (2019). From neon field sites to data portal: A community resource for

surface-atmosphere research comes online. *Bulletin of the American Meteorological Society*, *100*(11), 2305–2325. doi: 10.1175/BAMS-D-17-0307.1
Mironov, D. V., & Sullivan, P. P. (2016). Second-moment budgets and mixing intensity in the stably stratified atmospheric boundary layer over thermally heterogeneous surfaces. *Journal of the Atmospheric Sciences*, *73*(1), 449–464. doi: 10.1175/JAS-D-15-0075.1
Monin, A. S., & Obukhov, A. M. (1954). Basic Laws of Turbulent mixing in the surface layer of the atmosphere. *Tr. Akad. Nauk SSSR Geophiz. Inst.*, *24*(151), 163–187.
Monji, N. (1973). Budgets of Turbulent Energy and Temperature Variance in the Transition Zone from Forced to Free Convection. *J.Meteor.Soc.Japan*, *51*(April), 133–145.
Myneni, R., Park, T., & Knyazikhin, Y. (2015). *Mcd15a2h modis/terra+aqua leaf area index/fpar 8-day l4 global 500m sin grid v006.* [Dataset] NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD15A2H.006
Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan*, *87*(5), 895–912. doi: 10.2151/jmsj.87.895
National Ecological Observatory Network (NEON). (2021). *Bundled data products - eddy covariance (dp4.00200.001).* [Dataset] National Ecological Observatory Network (NEON). Retrieved from `https://data.neonscience.org/data-products/DP4.00200.001/RELEASE-2021` doi: 10.48443/BWAY-HC74
Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J., Scott, R. L., & Torn, M. S. (2018). The AmeriFlux network: A coalition of the willing. *Agricultural and Forest Meteorology*, *249*(October 2017), 444–456. doi: 10.1016/j.agrformet.2017.10.009
Obukhov, A. M. (1946). Turbulentnost' v temperaturnoj – neodnorodnoj atmosfere (turbu- lence in an atmosphere with a non-uniform temperature). *Trudy Inst. Theor. Geofiz. AN SSSR*, *1*, 95–115.
Oke, T. R. (1987). *Boundary layer climates.* Psychology Press.
Otić, I., Grötzbach, G., & Wörner, M. (2005). Analysis and modelling of the temperature variance equation in turbulent natural convection for low-Prandtl-number fluids. *Journal of Fluid Mechanics*, *525*, 237–261. doi: 10.1017/S0022112004002733
Poggi, D., Katul, G., & Albertson, J. (2004). Momentum transfer and turbulent kinetic energy budgets within a dense model canopy. *Boundary-Layer Meteorology*, *111*(3), 589–614.
Poggi, D., Porporato, A., Ridolfi, L., Albertson, J., & Katul, G. (2004). The effect of vegetation density on canopy sub-layer turbulence. *Boundary-Layer Meteorology*, *111*(3), 565–587.
Raupach, M. R., Finnigan, J. J., & Brunet, Y. (1996). Coherent eddies and turbulence in vegetation canopies: the mixing-layer analogy. *Boundary-Layer Meteorology*, *78*(3-4), 351–382. doi: 10.1007/BF00120941
Raupach, M. R., & Thom, A. S. (1981). Turbulence in and above plant canopies. *Annual review of fluid mechanics*, *3*, 97–129.
Siqueira, M., Katul, G., & Porporato, A. (2009). Soil moisture feedbacks on convection triggers: The role of soil-plant hydrodynamics. *Journal of Hydrometeorology*, *10*(1), 96–112. doi: 10.1175/2008JHM1027.1
Tillman, J. (1972). The Indirect Determination of Stability, Heat and Momentum Fluxes in the Atmospheric Boundary Layer from Simple Scalar Variables During Dry Unstable Conditions. *Journal of Applied Meteorology and Climatology*, *11*(5), 1–27. doi: 10.1175/1520-0450(1972)011⟨0783:TIDOSH⟩2.0.CO;2
van de Boer, A., Moene, A. F., Graf, A., Schüttemeyer, D., & Simmer, C. (2014). Detection of Entrainment Influences on Surface-Layer Measurements and Extension of Monin-Obukhov Similarity Theory. *Boundary-Layer Meteorology*,

152(1), 19–44. doi: 10.1007/s10546-014-9920-8

Waterman, T. (2021). *Neon collocation and processing code: Version 1.0.* [Software] Github. Retrieved from `https://github.com/tswater/neon_code`

Wilson, J. D. (2008). Monin-Obukhov functions for standard deviations of velocity. *Boundary-Layer Meteorology*, *129*(3), 353–369. doi: 10.1007/s10546-008-9319-5

Wyngaard, J., & Coté, O. (1971). Local Free Convection, Stability, and Budget of Shear Stress and Heat Flux. *Journal of Atmospheric Sciences*, *28*(7), 1–27. doi: 10.1175/1520-0469(1971)028⟨1171:LFCSAT⟩2.0.CO;2

Wyngaard, J. C., & Coté, O. R. (1974). The evolution of a convective planetary boundary layer - A higher-order-closure model study. *Boundary-Layer Meteorology*, *7*(3), 289–308. doi: 10.1007/BF00240833

Zhao, M., Golaz, J. C., Held, I. M., Guo, H., Balaji, V., Benson, R., … Xiang, B. (2018). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 2. Model Description, Sensitivity Studies, and Tuning Strategies. *Journal of Advances in Modeling Earth Systems*, *10*(3), 735–769. doi: 10.1002/2017MS001209