Scaling, anisotropy and complexity in near-surface atmospheric turbulence

Ivana Stiperski[1], Marc Calaf[2], Mathias W. Rotach[1]

[1]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria.
[2]Department of Mechanical Engineering, University of Utah, Salt Lake City, Utah, USA

Key Points:

- Similarity Theory
- Anisotropy Invariants
- Turbulent flows in complex terrain

Corresponding author: Ivana Stiperski, ivana.stiperski@uibk.ac.at

Abstract

The development of a unified similarity scaling has so far failed over complex surfaces, as scaling studies show large deviations from the empirical formulations developed over flat and horizontally homogeneous terrain and also large deviations between the different complex terrain datasets. However, a recent study of turbulence anisotropy for flat and horizontally homogeneous terrain has shown that separating the data according to the limiting states of anisotropy (isotropic, two-component axisymmetric and one component turbulence) improves near-surface scaling. In this paper we explore whether this finding can be extended to turbulence over inclined and horizontally heterogeneous surfaces by examining near-surface scaling for twelve different datasets obtained over terrain ranging from flat to mountainous. Although these datasets show large deviations in scaling when all anisotropy types are examined together, the separation according to the limiting states of anisotropy significantly improves the collapse of data onto common scaling relations indicating the possibility of a unified framework for turbulence scaling. The causes for the breakdown of scaling and the physical mechanisms behind the turbulence complexity encountered over complex terrain are identified and shown to be related to the distance to the isotropic state, prevalence of directional shear with height in mountainous terrain and the existence of non-isotropy in the inertial subrange. A measure of turbulence complexity is finally developed.

1 Introduction

Atmospheric surface-layer (ASL) similarity theory was developed as a unified theory of statistically stationary turbulence over horizontally homogeneous and flat terrain [HHF; e.g. *Monin and Yaglom*, 1971]. Although never meant to be employed over heterogeneous and non-flat surfaces, the lack of a better framework has thus far led to similarity theory being employed in weather prediction and climate models over all types of terrain [cf. *Rotach et al.*, 2017]. Given the prevalence of heterogeneity of the earth land surface [e.g. *Rotach et al.*, 2014], adaptations were developed by reconciling theory and application under the principle of '*local*' homogeneity in order to be able to model real flows over heterogeneous surfaces. Meaning, that over small enough regions, sampled long enough, what a-priori might resemble a heterogeneous surface, can ultimately be interpreted as homogeneous. While these practical adjustments work well for regions with weak heterogeneities [e.g. *Sfyri et al.*, 2018], similarity relationships become severely challenged in complex terrain [e.g. *Martins et al.*, 2009; *Nadeau et al.*, 2013; *Sfyri et al.*, 2018] where terrain slope, land-use characteristics and complexity of the flow itself (*e.g.* low-level jets, flow separation etc.) cause turbulence to exhibit increasing complex structure [e.g. *Nadeau et al.*, 2013; *Oldroyd et al.*, 2016; *Stiperski and Rotach*, 2016; *Grachev et al.*, 2016]). In this work, reference to complex terrain is understood as topographic perturbations that induce spatial and/or temporal perturbations to the atmospheric flow with a timescale shorter than that of the diurnal cycle or mesoscale phenomena (*e.g.,* sloped terrain, ground roughness and thermal patchiness, obstacles, etc.)

Numerous studies have illustrated the adequacy of similarity theory under idealized terrain and flow conditions [e.g. *Panofsky and Dutton*, 1984; *Wyngaard*, 2010]. Nonetheless, an important degree of scatter still exists, particularly for horizontal velocity variances that are commonly assumed not to obey surface layer scaling [e.g. *Kaimal and Finnigan*, 1994; *Wyngaard*, 2010; *Banerjee et al.*, 2015; *Chamecki et al.*, 2017]. This scatter also persists despite the advanced post-processing techniques and progressively more restrictive quality criteria imposed on the data. In an effort to overcome these challenges, *Stiperski and Calaf* [2018] employed a novel approach by examining traditional similarity scaling relations over flat and horizontally homogeneous terrain based on clustering the data according to anisotropy. Results of this work illustrated a strong dependence between the quality of the scaling fit and the characteristic topology of the turbulent flow, showing that the similarity scaling significantly improves when the turbulent flow is a-priory classified according to the anisotropy type. In essence, results illustrated that isotropic and two-component axisymmetric type turbulence scales the best (*i.e.,* show closest collapse on a scaling line), although for horizontal velocity components, the two types of anisotropy were shown to follow different scaling curves. This finding could explain the commonly encountered large scatter observed for scaled standard deviations of horizontal velocities. On the other hand, one-component turbulence strongly departs from regular scaling curves. These results together with the possibility of predicting the anisotropy type based on larger scale variables as shown in *Stiperski and Calaf* [2018], promised to be a powerful tool in improving similarity scaling relations.

In complex terrain, on the other hand, despite the progressively more severe restrictions imposed on the experimental data [cf. *Stiperski and Rotach*, 2016], significant scatter and a relevant degree of discrepancy between the experimental data and similarity relationships are more evident. Even more, all the studies examining the applicability of surface-layer scaling for data obtained over diverse complex settings [*e.g., Park and Park*, 2006; *de Franceschi et al.*, 2009; *Martins et al.*, 2009; *Nadeau et al.*, 2013; *Kral et al.*, 2014; *Babić et al.*, 2016a,b; *Grachev et al.*, 2016; *Sfyri et al.*, 2018] show that the scaling relations not only differ from the functional relations obtained over flat and horizontally homogeneous terrain, but also differ from site to site, suggesting that scaling might be inherently '*local*' (*i.e., location dependent*) and therefore no unified theory of turbulence over all types of surfaces is possible. The search for an additional scaling variable in complex terrain that could explain these discrepancies has so far been proven unsuccessful, as the only systematic study to date [*Sfyri et al.*, 2018] found no clear relationship between scaling and slope angle, at least for the scaled standard deviations of scalars. Still, the data from progressively more complex surfaces do show larger deviations from the scaling curve and generally larger scatter, even if the exact mechanism behind this finding escapes clear explanation.

In this work, and based on the earlier approach first introduced in *Stiperski and Calaf* [2018], we present a new interpretation of the a-priori mismatch of near-surface data in complex terrain and traditional scaling relations based on the anisotropy of the turbulence stress tensor. The results show that similar to flat and horizontally homogeneous terrain, separating the complex terrain data according to anisotropy significantly improves scaling, offering a pathway towards a unified theory of turbulence. In addition, we provide a novel approach that defines *complexity* as not only exclusively associated with terrain characteristics, but also to the actual resultant turbulence structure. The physical mechanisms causing this complexity are then identified. This new definition of complexity could facilitate comparison between different datasets collected in regions with different atmospheric and topographic characteristics.

The paper is organized as follows: in Section 2 the datasets and post-processing methods are presented, the anisotropy analysis is reviewed, and scaling relations introduced; Section 3 presents the relationship between similarity scaling and the anisotropy of turbulence over complex terrain; Section 4 identifies a measure of turbulence complexity and examines its relation to the physical mechanisms acting in complex terrain; an extended discussion of the results and implications for similarity theory as well as conclusions are provided in Section 5.

2 Methodology

2.1 Datasets

In this study we examine turbulence measurements from twelve flux towers located on surfaces of different complexity, ranging from flat to highly complex mountainous terrain. These are part of well known datasets and include the tower at Cabauw experimental site for atmospheric research (Cesar) of the Royal Netherlands Meteorological Institute [*e.g.*, *Beljaars and Bosveld*, 1997], the Cooperative Atmosphere – Surface Exchange Study 1999 [CASES-99; *Poulos et al.*, 2002], the Terrain-induced Rotor Experiment [T-Rex; *Grubišič et al.*, 2008], the Mountain Terrain Atmospheric Modeling and Observations [MATERHORN; *Fernando et al.*, 2015], the Second Meteor Crater Experiment [METCRAX II; *Lehner et al.*, 2016], and the Innsbruck Box [i-Box; *Rotach et al.*, 2017]. A detailed description of the datasets is given in Table 1. The dataset that conforms to the flat and horizontally homogeneous terrain the best is CASES-99. Already studied in *Stiperski and Calaf* [2018] it forms the basis of the current analysis. The data consist of a month of measurements from a 60 m tower with 7 levels of sonic anemometers. Due to issues with the anisotropy of the CSAT3 measurements during stable periods identified in *Stiperski and Calaf* [2018], here we only study the levels with ATI-K probes in stable conditions. Cabauw data can also be considered flat, however, horizontally weakly inhomogeneous [*Sfyri et al.*, 2018]. The other datasets were chosen according to their increasing terrain complexity. The Central tower from T-Rex is located at an almost flat valley floor, however, its setting within a mountain valley (*i.e.,* complex terrain) nevertheless has a profound influence on scaling [*cf. Babić et al.*, 2016b]. The i-Box0 valley floor site (see Table 1), apart from being located in a narrower valley than T-Rex, is additionally characterized by larger surface heterogeneity, given that it is surrounded by mixed agricultural land. The rest of the datasets are located on slopes of various steepness and are strongly influenced by flows associated with sloped terrain (*e.g.,* thermally-driven katabatic and anabatic flows and dynamically-driven wind systems) and/or heterogeneity. The i-Box1 station has a small slope angle, however, the influence of surface heterogeneity (corn and meadows) for this station is larger than the influence of sloping terrain because the dominant wind direction is across the slope. The same is true for T-Rex West tower, although there the wind rose also shows a large influence of katabatic flows as well as downslope windstorms [*Babić et al.*, 2016b]. On the other hand, deep katabatic flows with jet maxima between 20 and 40 m above ground level (agl) develop regularly at the METCRAX II NEAR tower [*cf. Savage et al.*, 2008; *Lehner et al.*, 2016]. Persistent shallow katabatic flows with a jet maximum at around 5 m agl are also found at MATERHORN ES4 and ES5 towers located at the top of a relatively shallow slope below a couloir [*Grachev et al.*, 2016]. Even shallower katabatic flows develop at the steeper i-Box10 and i-Box27 stations. The i-Box mountain top station (i-BoxTop) represents the most complex site due to its location on a ridge exposed to flow from all sides that, depending on wind direction, responds to very different slope angles. Operating continuous turbulence measurements at this station is challenging therefore only seven months of measurements were analyzed in this study, as opposed to the other i-Box stations where one year of data was analyzed. The datasets (in both Table 1 and future Figures) are a-priory subjectively ordered according to their slope angle, with colder colors representing gentler slopes (flat terrain being considered more ideal than the flat valley floor locations) and warmer colors progressively steeper slope angles (4 - 27°).

2.2 Data treatment and quality control

To remove the inconsistencies in data processing applied by different groups responsible for each of the datasets, we reanalyzed all the data with a processing routine described in *Stiperski and Calaf* [2018]. First, the Multi-resolution flux decomposition [MRD, *e.g., Vickers and Mahrt*, 2003] technique was used to determine the optimal averaging time for daytime and nighttime turbulence (Figure 1). As in *Stiperski and Calaf* [2018] we separate the data into strongly and weakly stable/unstable regimes to examine how the averaging time depends on the stability. The results show that for the examined datasets a 1 min averaging time for stable stratification and 30 min averaging time for unstable stratification generally capture the majority of turbulence contributions to the flux while eliminating most of the (sub)mesoscale effects. The obvious exception here is the i-Box27 steep slope station (orange line in Figure 1). At that station very stable conditions are rarely encountered, therefore 1 min average slightly underestimates the nighttime fluxes.

Table 1. Information on the datasets used in the study.

Station	Official Name	Short Name	Location	Terrain Complexity	Measurements heights [m]	Slope Angle [°]	Surface Type	Data length
CASES-99	CASES-99	CA	Kansas	flat	5, 10, 20, 30, 40, 50, 55	< 0.5	Grassland	October 1999
Cabauw	Cabauw	CB	Netherlands	flat	3, 60	< 0.5	Grassland	July-October 2007
T-RexC	Central Tower	TRC	California	flat valley floor	5, 10, 15, 20, 25, 30	< 0.5	Desert	March-May 2006
i-Box0	CS-VF0	iB0	Austria	flat valley floor	4, 8.7, 16.9	< 0.5	Mixed Agricultural	January-December 2015
i-Box1	CS-SF1	iB1	Austria	flat foothills	6.6	1	Alpine meadow and agricultural	January-December 2015
METCRAX II	Near Tower	MC	Arizona	Gentle Slope	5, 10, 15, 20, 25, 30, 35, 40, 45, 50	1	Desert	October 2013
T-RexW	West Tower	TRW	Califronia	Gentle Slope	5, 10, 15, 20, 25, 30	3.25	Desert	March-May 2006
MATERHORN ES4	ES4 Tower	MT4	Utah	Gentle Slope	0.47, 2.05, 5.12, 10, 20, 26.5	4	Desert	September-October 2013
MATERHORN ES5	ES5 Tower	MT5	Utah	Gentle Slope	0.55, 2.14, 5.13, 10.13, 20.08	6.4	Desert	September-October 2013
i-Box10	CS-NF10	iB10	Austria	Steep Slope	6.2	10	Alpine meadow	January-December 2015
i-Box27	CS-NF27	iB27	Austria	Steep Slope	6.8	27	Alpine meadow	January-December 2015
i-BoxTop	CS-MT21	iBTop	Austria	Mountain Top	4.67	21	High-Alpine Vegetation	April-October 2015

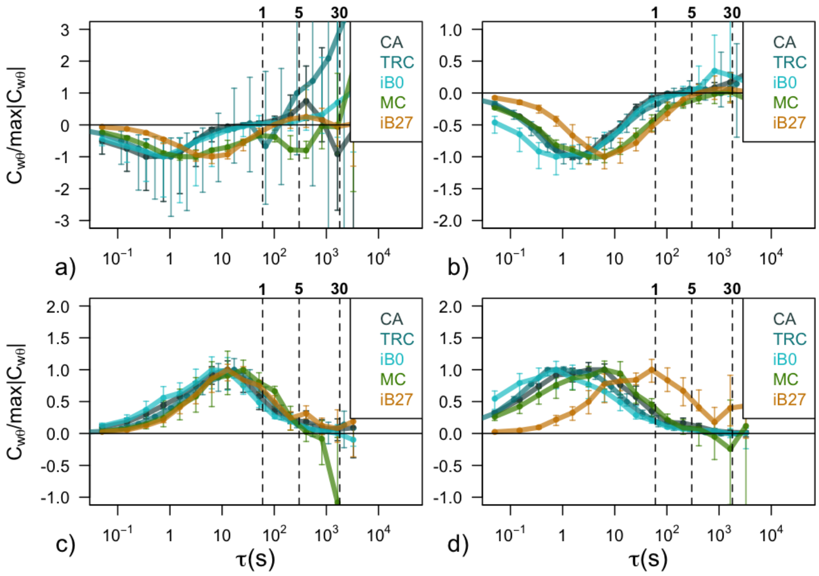


Figure 1. Multi-resolution flux decomposition of heat flux for example nights with (a) strongly stable, (b) weakly stable, (c) strongly unstable and (d) weakly unstable stratification, respectively, for the lowest measurement level of different datasets (shown in color, see Table 1 for abbreviations). The colored lines represent medians and are normalized by their maximum value within the turbulence scales for each dataset so as to eliminate differences in magnitude. Error bars represent the 25 and 75 percentile. Vertical dashed lines indicate timescales of 1 min, 5 min and 30 min, respectively.

The data were then de-trended and block averaged to the given averaging time. Double rotation was used to align the coordinates into the streamwise coordinate system. Zero-plane displacement information was applied to the stations where this information was available. For i-Box measurements, the zero-plane displacement was calculated based on the measurements of surrounding vegetation height [*Sfyri et al.*, 2018]. For the two T-Rex towers the values from the study of *Babić et al.* [2016b] were used. For the other datasets the zero-plane displacement was assumed to be zero given the generally low vegetation height.

Vertical wind shear and temperature gradients needed for calculating the gradient Richardson number R_i were determined for datasets with multiple measurement levels (*i.e.*, CASES-99, T-Rex, METCRAX II, MATERHORN and i-Box0). In order to determine the local wind speed gradient at each measurement height analytic profiles were fit through the entire tower length. Different analytic formulations were needed for each dataset due to profiles having different characteristics, particularly in case of the existence of low-level jets. The formulations used were $x = a + b\,z + c\,z^2 + d\,log(z)$ for CASES-99, $x = a + b\,z + c\,z^2 + d\,log(z) + e\,log(z)^2$ for T-Rex, $x = a + b\,z + c\,z^2 + d\,z^3 + e\,log(z) + f\,log(z)^2$ for METCRAX II, $x = a + b\,log(z) + c\,log(z)^2$ for MATERHORN, and finally $x = a + b\,log(z) + c\,log(z)^2 + d\,log(z)^3$ for i-Box0. As a quality check, only those wind speed gradients in which the root mean square error of the best fit was lower than 0.3 ms^{-1} were taken into account.

All turbulence data were required to pass the basic quality control (test of physical limits) as well as to satisfy the stationarity test given by *Foken and Wichura* [1996] at its standard 30% level. As in *Stiperski and Calaf* [2018], the stationarity criterion was dropped for small fluxes *i.e.*, for very unstable conditions stationarity of the momentum flux was not required, while for near-neutral conditions the same was true for the stationarity of the heat flux. For datasets with multiple levels, the requirement that the gradient Richardson number be smaller than 0.25 was also imposed [cf. *Grachev et al.*, 2013]. As shown in *Stiperski and Calaf* [2018], existence of unstably stratified turbulence during nighttime points to non-local sources of turbulence and cannot be expected to follow scaling. In order to filter these counter gradient fluxes, theoretical incoming shortwave radiation was used to determine sunrise and sunset times together with the conservative cross-over time of the daily cycle of sensible heat flux. This was particularly important for i-Box stations where a year of data was analyzed meaning that sunset and sunrise times varied significantly. No flux corrections were applied to the data, the same as in *Stiperski and Calaf* [2018].

2.3 Anisotropy

Traditionally, the Reynolds stresses ($\overline{u_i'u_j'}$) can be decomposed into an isotropic and anisotropic contribution, the latter being the one contributing the most to the transport of momentum [*Pope*, 2000]. The anisotropy contribution to scalar fluxes is also assumed to be important. The sum of the isotropic components of the Reynolds stress tensor is traditionally referred to as twice the turbulent kinetic energy ($2e = \overline{u_i'u_i'}$). In the above notation, a prime indicates a departure of a time-averaged quantity, and the overbar indicates the time-averaging operation. Additionally, the indices i, j vary between 1 to 3, in reference to the traditional Cartesian coordinate reference system with 1 indicating the streamwise, 2 the spanwise, and 3 the surface-normal directions, respectively.

The deviatoric anisotropy stress tensor defined as,

$$a_{ij} \equiv \overline{u_i u_j} - \frac{2}{3} e\delta_{ij}, \tag{1}$$

and in non-dimensional form (normalized by $2e$) as,

$$b_{ij} = \frac{\overline{u_i u_j}}{\overline{u_l u_l}} - \frac{1}{3}\delta_{ij}, \tag{2}$$

has long been studied in relationship to, for example, the pressure-strain correlation to develop models that capture the return-to-isotropy process once mean velocity gradients stop acting on the flow [*Rotta*, 1951; *Lumley and Newman*, 1977; *Lumley*, 1978; *Sarkar and Speziale*, 1990; *Choi and Lumley*, 2001]. Based on Lumley's work [*Lumley and Newman*, 1977; *Lumley*, 1978], it is possible to reduce the original three dimensional problem characterized by six independent terms (normalized deviatoric anisotropy stress tensor) into a simpler problem with two degrees of freedom, based on the so-called anisotropy invariants, η and ξ [*Pope*, 2000] that are also functions of the eigenvalues (λ_i, $i = 1, 2, 3$) of the anisotropy stress tensor. The first invariant (η) is positive definite and provides a measure of the degree of anisotropy in the flow field (large values indicating intense anisotropy, and small values indicating near-isotropic behavior). The second invariant (ξ) can be positive or negative, indicating that the flow is dominated by one-component turbulence when positive, and by two-component turbulence when negative. These invariants can be mathematically determined from the normalized deviatoric anisotropy stress tensor as [*Pope*, 2000]

$$6\eta^2 = b_{ij}b_{ji} \qquad \text{and} \qquad 6\xi^3 = b_{ij}b_{jk}b_{ki}. \tag{3}$$

As a result, it is possible to represent any realizable state of turbulence on a single two-dimensional non-linear map, the so-called Lumley Triangle (LT, *Lumley* [1978]; *Pope* [2000]). Here, instead, we use a modification of the original LT, the Barycentric Lumley Triangle (BLT, *Banerjee et al.* [2007], see Figure 2 and Table 2), that overcomes the complexity associated with the non-linearity of the LT by equally weighing the different limiting states of turbulence anisotropy. The corresponding coordinates (x_B, y_B)

Table 2. Summary of the special states of the Reynolds-stress tensor in terms of the invariants (η, ξ), and the eigenvalues of the anisotropy stress tensor as described by the Lumley Triangle. The fourth column introduces the corresponding ellipsoid shape described by the eigenvectors [*Pope*, 2000].

Cases	**Invariants**	**Eigenvalues**	**Shape ellipsoid**
Isotropic	$\eta = \xi = 0$	$\lambda_1 = \lambda_2 = \lambda_3 = 0$	Sphere
Two-component axisymmetric	$\eta = \frac{1}{6}, \xi = -\frac{1}{6}$	$\lambda_1 = \lambda_2 = \frac{1}{6}$	Disk
One-component	$\eta = \xi = \frac{1}{3}$	$\lambda_1 = \frac{2}{3}, \lambda_2 = \lambda_3 = -\frac{1}{3}$	Line
Axisymmetric, one large eigen-value	$\eta = \xi$	$-\frac{1}{3} \leq \lambda_1 = \lambda_2 \leq 0$	Prolate Spheroid
Axisymmetric, one small eigen-value	$\eta = -\xi$	$0 \leq \lambda_1 = \lambda_2 \leq \frac{1}{6}$	Oblate Spheroid
Two-component	$\eta = (\frac{1}{27} + 2\xi^3)^{1/2}$	$\lambda_1 + \lambda_2 = \frac{1}{3}$	Ellipse

of this linearized 2D map are related to the eigenvalues as

$$x_B \quad = C_{1c}x_{1c} + C_{2c}x_{2c} + C_{3c}x_{3c} = C_{1c} + C_{3c}\frac{1}{2}, \tag{4}$$

$$y_B \quad = C_{1c}y_{1c} + C_{2c}y_{2c} + C_{3c}y_{3c} = C_{3c}\frac{\sqrt{3}}{2}, \tag{5}$$

with the corresponding weights (C_{ic}) written as $C_{1c} = \lambda_1 - \lambda_2$, $C_{2c} = 2(\lambda_2 - \lambda_3)$, and $C_{3c} = 3\lambda_3 + 1$, with $x_{1C} = (1, 0)$, $x_{2C} = (0, 0)$, and $x_{3C} = (1/2, \sqrt{3}/2)$ indicating the limiting states of turbulence anisotropy in the BLT. Both invariant maps are equivalent given the existing relationship between the anisotropy invariants (η and ξ) and the eigenvalues of the normalized anisotropy tensor (λ_i, *Spencer* [1971]),

$$\eta^2 \quad = \tfrac{1}{3}(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) \tag{6}$$

$$\xi^3 \quad = -\tfrac{1}{2}\lambda_1\lambda_2(\lambda_1 + \lambda_2). \tag{7}$$

Finally, it is important to reiterate that the shape associated with the limiting states of anisotropy refers to the representation in eigenvalue space and not to the physical shape of turbulence itself [*Simonsen and Krogstad*, 2005]. More detail on the analysis of turbulence anisotropy can be found in [*Stiperski and Calaf*, 2018]

2.4 Scaling

Following *Stiperski and Calaf* [2018] we examine the influence of turbulence anisotropy on near-surface similarity within the local scaling framework [*cf. Nieuwstadt*, 1984a,b]. Turbulent quantities are therefore scaled with the fluxes obtained at the corresponding measurement height z. The local Obukhov length Λ is defined as $\Lambda = \frac{-u_*^3\theta_v}{\kappa g \overline{w'\theta'}}$, where θ_v is the mean virtual potential temperature, κ the von Karman constant, equal to 0.4, u_* the local friction velocity computed as $u_* = (\overline{u'w'}^2 + \overline{v'w'}^2)^{\frac{1}{4}}$ and $\overline{w'\theta'}$ is the local heat flux. The quantity $(z-d)/\Lambda$, where d is the displacement height, represents the local stability. We also define the local temperature scale as $\theta_* = -\frac{\overline{w'\theta'}}{u_*}$. The following functional forms of the surface-layer flux-variance similarity relationships are used for reference: for the standard deviations of velocity components (Φ_u,Φ_v,Φ_w), following *Panofsky and Dutton* [1984],

$$\Phi_w = \frac{\sigma_w}{u_*} = \begin{cases} 1.25(1 + 3\frac{z}{\Lambda})^{\frac{1}{3}} & \text{for } \frac{z}{\Lambda} > 0 \\ 1.25(1 - 3\frac{z}{\Lambda})^{\frac{1}{3}} & \text{for } \frac{z}{\Lambda} < 0 \end{cases} \tag{8}$$

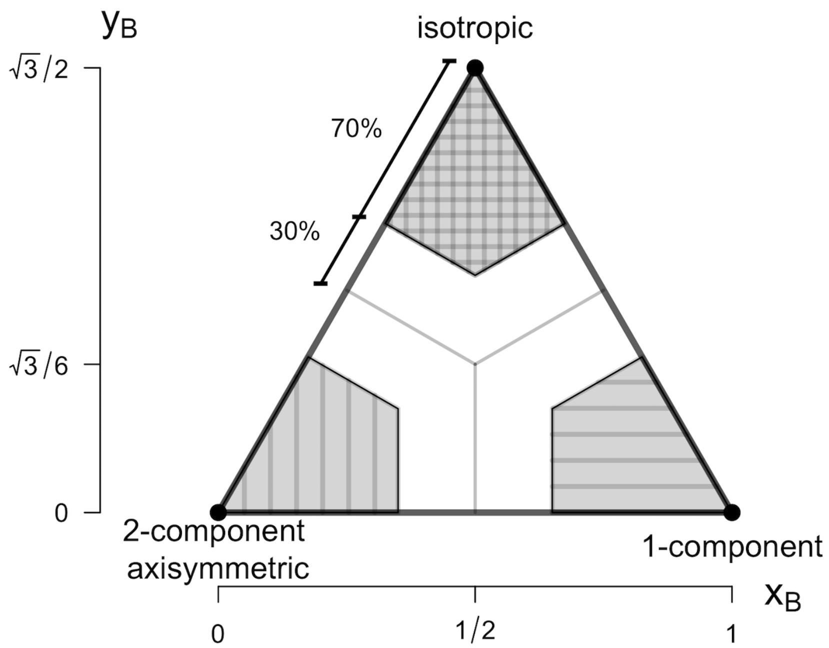


Figure 2. Barycentric Lumley Triangle as a function of the anisotropy stress tensor eigenvalues and represented through the linearized coordinates x_B and y_B. The shading indicates regions of the triangle that were selected as *pure* limiting states of anisotropy.

$$\Phi_{u,v} = \frac{\sigma_{u,v}}{u_*} = \begin{cases} 2.55(1+3\frac{z}{\Lambda})^{\frac{1}{3}} \text{ for } \frac{z}{\Lambda} > 0 \\ 2.55(1-3\frac{z}{\Lambda})^{\frac{1}{3}} \text{ for } \frac{z}{\Lambda} < 0 \end{cases} \tag{9}$$

for temperature standard deviation (Φ_θ), taking the reference curve from *Sfyri et al.* [2018],

$$\Phi_\theta = \frac{\sigma_\theta}{\theta_*} = \begin{cases} 2 + 6.7 \cdot 10^{-4} \frac{z}{\Lambda}^{-1.42} \text{ for } \frac{z}{\Lambda} > 0 \\ 1.67 - 0.016(\frac{z}{\Lambda})^{1} \text{ for } -0.05 < \frac{z}{\Lambda} < 0 \\ 1.95(0.05 - \frac{z}{\Lambda})^{\frac{1}{3}} \text{ for } \frac{z}{\Lambda} < -0.05 \end{cases} \tag{10}$$

and for turbulence dissipation rate (Φ_ε), following *Thiermann* [1990],

$$\Phi_\varepsilon = \frac{kz\varepsilon}{u_*^3} = \begin{cases} (1 + 4\frac{z}{\Lambda} + 16(\frac{z}{\Lambda})^2)^{\frac{1}{2}} \text{ for } \frac{z}{\Lambda} > 0 \\ (1-3\frac{z}{\Lambda})^{-1} - \frac{z}{\Lambda} \text{ for } \frac{z}{\Lambda} < 0 \end{cases} \tag{11}$$

In the stable z-less limit *Sorbjan* [1987] suggested the following constant values of the flux-variance relationships

$$\Phi_w = 1.6, \quad \Phi_{u,v} = 3.1. \tag{12}$$

A number of quantitative measures are used to determine the amount of scatter between the datasets and the agreement with the reference scaling curves. Absolute deviations $\Delta\Phi_y = |y - \Phi_y|$, where $y = u, v, w, \theta, \varepsilon$, are calculated as the absolute distance between the scaled data y and the corresponding reference scaling curves (Eq. 8 – 11). For stable conditions the z-less limit (Eq. 12) is used for the velocity components. For data separated according to anisotropy, the distance to the closest scaling line is employed. This implies that for standard deviations of horizontal velocity components under isotropic conditions, the reference curve used in the calculation is Φ_w, and not $\Phi_{u,v}$ [cf. *Stiperski and Calaf*, 2018]. The median of the absolute deviations (MAD) is used to quantify the mean disagreement of each dataset with the respective scaling as well as to quantify the complexity of the dataset. While the absolute value of MAD already provides an indicator of the goodness of scaling, it is sensitive to the choice of the scaling curve and therefore might lead to large values if the choice is inadequate. Therefore, as a best indicator of the improvement in scaling, *i.e.* decreased variability in MAD between the different datasets, we use the inter-quartile range (IQR). It is calculated from the twelve MADs for each scaling variable and stability as the difference between the 75 and 25 percentile. This measure provides information on the discrepancies between the different datasets and the scaling curves and in that sense quantifies how '*location dependent*' the scaling for different datasets is, and how much improvement is brought about by the new approach of *Stiperski and Calaf* [2018].

3 Results

The scaled standard deviations of high-quality stationary data for each of the twelve datasets are shown in Figure 3 in comparison with the traditional similarity relations from horizontally homogeneous and flat terrain (Eq. 8 – 12). For visualization purposes the data from each dataset are binned and the median of each bin is displayed. The spread of the data is shown as the shading and corresponds to the inter-quartile range of each bin. The absolute deviations between the data and the reference scaling curves as well as inter-dataset spread are shown in Figure 4.

The results show large scatter both within each individual dataset (large shaded area in Figure 3) as well as between the different datasets (large IQR values in Figure 4), confirming the '*location-dependent*' nature of scaling in complex terrain. The large scatter within each dataset is particularly clear for horizontal velocity components $\Phi_{u,v}$ in the very unstable region (cf. Figure 3a & b), to the point that the similarity relations can be deemed meaningless in that case. The same is true for the near-neutral regions for scaled temperature Φ_θ and less so for the scaled TKE dissipation rate Φ_ε. Notably, the vertical velocity variance exhibits good scaling behavior throughout, regardless of the dataset. This is particularly interesting given the large disparity of datasets used in this work, representative of very different terrain and flow complexities as well as vertical coordinates, which over flat terrain represent the vertical and in complex terrain the slope-normal direction. In the stable regime in general, the data scatter suggests a better collapse to a scaling curve (smaller IQR values in Figure 4 than for unstable data), particularly for weakly stable conditions of the velocity components $\Phi_{u,v,w}$. It is interesting to note, however, that the scaled standard deviation of streamwise velocity Φ_u seems to suggest an ordering according to the datasets in the near-neutral regime, with some datasets showing a higher neutral limit than others, but all exhibiting a value lower than the HHF relation. The deviations from the reference scaling curves Eq. 8 – 9 in the very stable regime are quite substantial, indicative of, though not confirming, z-less scaling. This general behavior is slightly different for Φ_θ, where large scatter, mirroring that of the very unstable region of horizontal velocities, is present under weak stability, while data seem to scale better in the strongly stable regime.

In a first approximation, one may expect the results in Figure 4 to show an increasing deviation from the traditional scaling relations with increasing 'complexity' of the underlying terrain where the data were measured (see the subjective ordering of the datasets in Table 1 and the associated colors). Therefore, it could be expected that the smallest discrepancies are found for the most ideal sites such as CASES-99 (CA), Cabauw (CB), T-RexC and i-Box0 stations, and the biggest for data measured in stations located in very inclined terrain, such is i-Box10 and i-Box27 or i-BoxTop. Figure 4, however, shows that this is not the case and there is little to no correlation between the deviation from traditional scaling and the a-priori ordering of the datasets that was based solely on the slope angle-induced complexity. In fact, it can be observed that the discrepancies in scaling are generally the largest for CASES-99 and METCRAX II (Figure 4 a–c), despite

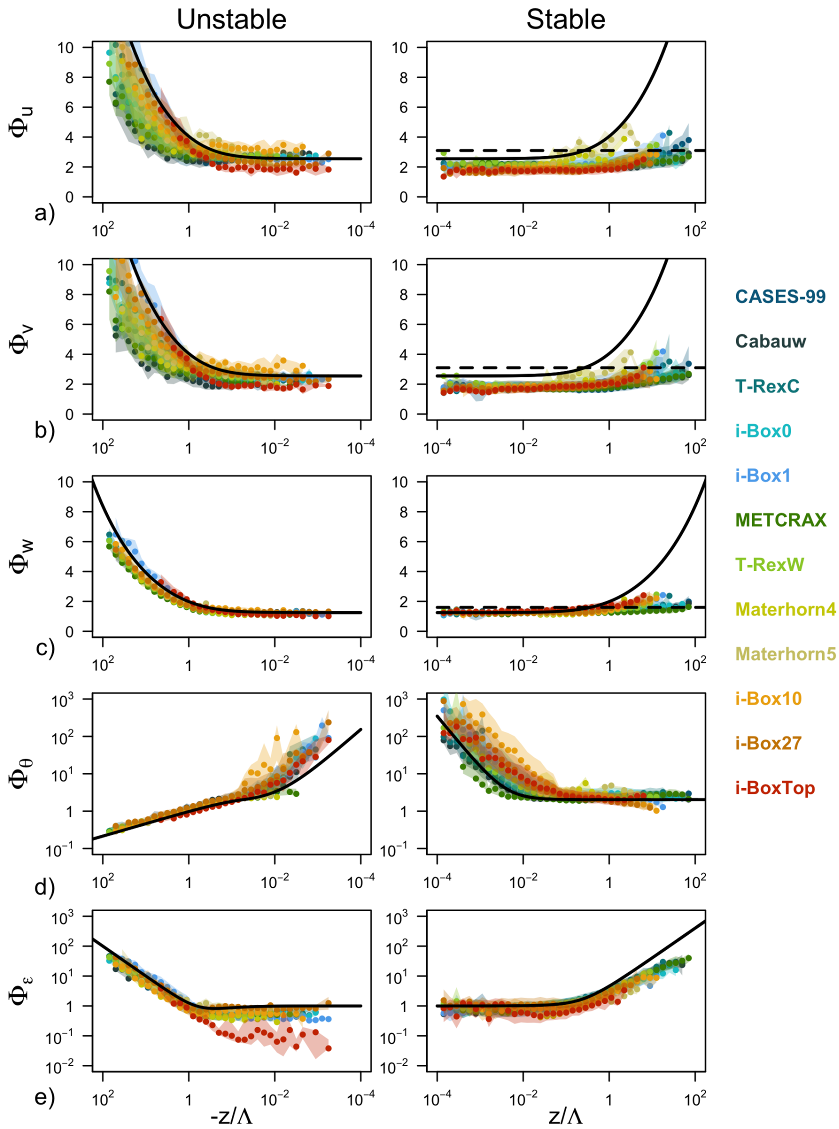


Figure 3. Scaling relations of the standard deviation of a) streamwise velocity (Φ_u), b) spanwise velocity (Φ_v), c) surface-normal velocity (Φ_w), d) temperature (Φ_θ) and d) TKE dissipation rate (Φ_ε) as a function of the local stability z/Λ for unstable (left) and stable (right) stratification. Colors represent different datasets described in Table 1. Points represent medians calculated over the bins of logarithmically-spaced z/Λ, while the shading corresponds to the inter-quartile range. The full black lines correspond to the traditional scaling relations (Eq. 8 – 11) and dashed lines to the z-less scaling for each variable (Eq. 12).

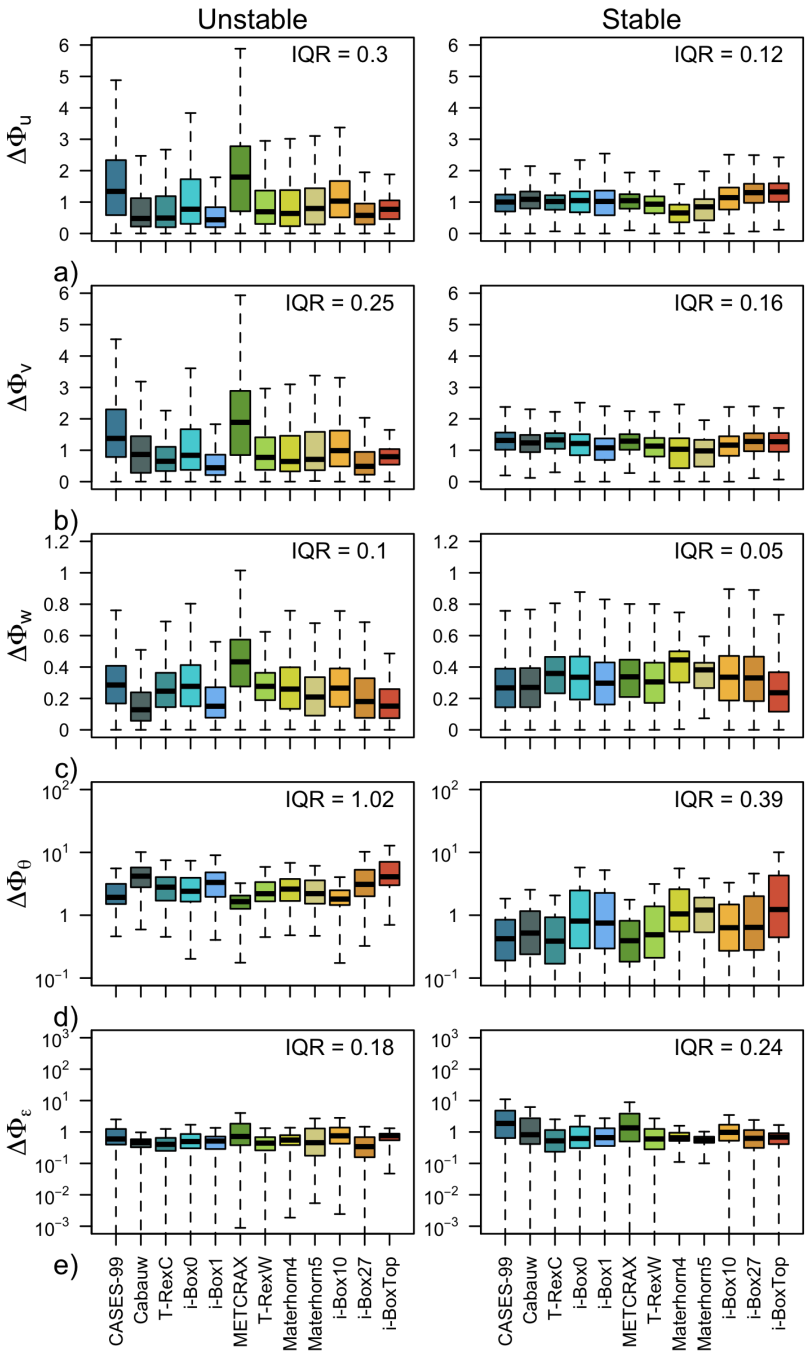


Figure 4. Box plots of absolute deviations a) – e) $\Delta\Phi_y$ where $y = u, v, w, \theta$ and ε, respectively, of the scaled data from the corresponding scaling relations (Eq. 8 – 12): as a function of the dataset (color), for unstable (left column) and stable (right column) stratification. Note the different vertical axes for each variable. The value of the inter-quartile range (IQR) is listed in each panel.

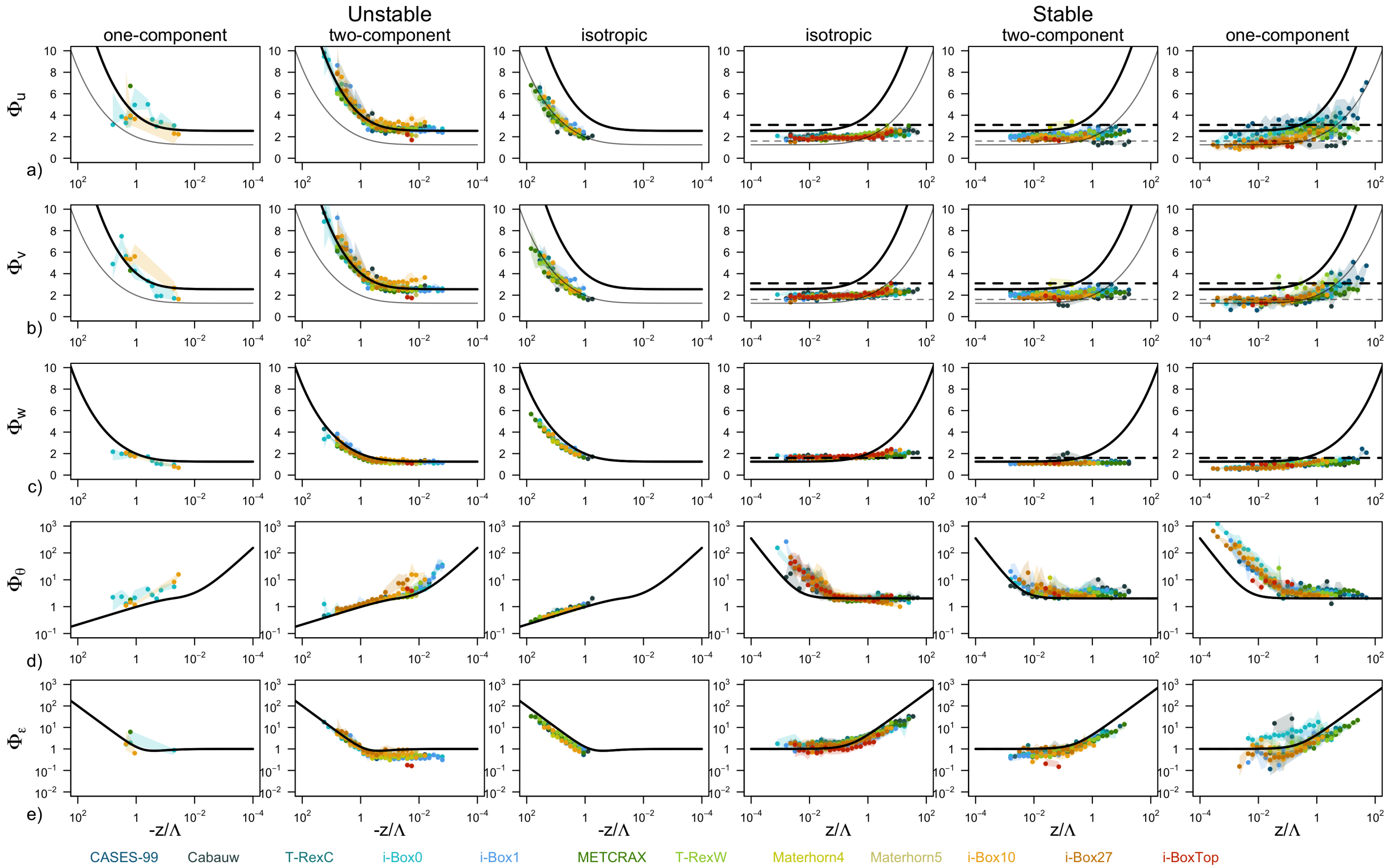


Figure 5. The same as Figure 3 but for data separated according to the pure states of anisotropy (isotropic, two-component axisymmetric and one component). The additional thin lines for the standard deviations of horizontal velocity components in a) and b) correspond to the scaling curve for Φ_w.

these being some of the most a-priori ideal locations (see Table 1). This mismatch with ordering of the datasets is present in all scaling variables, and particularly in unstable but also in stable stratification, suggesting that our simple classification of datasets based on slope angle is not corresponding to the true nature of their complexity.

Following the approach developed in *Stiperski and Calaf* [2018], next we separate the data according to the three limiting states of anisotropy: isotropic, two-component axisymmetric and one-component turbulence, before revisiting the scaling relations (Figure 5). Figure 6 shows the MAD and IQR as objective measures of the improvement of the scaling. Mirroring the results for flat terrain (CASES-99 in *Stiperski and Calaf* [2018]), accounting for turbulence topology drastically improves scaling (*i.e.,* decreases MAD and IQR) for all datasets regardless of the complexity induced by terrain and local weather conditions. For all variables, the most consistent reduction in IQR and overall the best scaling behavior is apparent for isotropic turbulence both under unstable and stable stratification, and for unstable two-component axisymmetric turbulence (Figure 6). Compared to Figure 4, the largest improvement is obtained for the horizontal velocity components under unstable stratification with MAD and IQR reduced by up to 60%. As already found for CASES-99 by *Stiperski and Calaf* [2018], isotropic and two-component axisymmetric turbulence for horizontal velocity variances follow two distinctly different scaling lines. We come to the same conclusion as *Stiperski and Calaf* [2018] that this clustering of the data to two different scaling curves is the leading cause of the large scatter illustrated in Figure 3 a&b and Figure 4 a&b in the very unstable region. It is interesting to note that for all datasets, irrespective of complexity, isotropic turbulence occupies the same region of z/Λ as for flat terrain (*i.e.* $z/\Lambda < -1$). This is a clear sign of TKE production in isotropic turbulence being thermally dominated [*Stiperski and Calaf*, 2018] irrespective of terrain complexity. On the other hand, in complex terrain, two-component axisymmetric turbulence is found also in very unstable conditions, contrary to the results over flat terrain (CASES-99 in *Stiperski and Calaf* [2018]). The small amount of data points corresponding to unstable one-component turbulence does not allow us to reach definite conclusions about similarity of this type of turbulence. The results do seem to suggest a larger degree of scatter and therefore a lack of proper scaling.

In the stable regime, the scaled standard deviations of all velocity components follow the z-less scaling up to $z/\Lambda \approx 1$ whereupon the data start to deviate from their constant value. This is opposite to the general expectations for z-less scaling and might be an indication of self-correlation [*Klipp and Mahrt*, 2004]. The least scatter between the datasets (smallest IQR) is observed for isotropic turbulence. This is true even for the standard deviation of temperature, but only over a very limited range of stability. The near-neutrally stable region of Φ_θ shows large scatter, which is likely another indication of self-correlation as found by *Sfyri et al.* [2018]. In addition, the negative slope of the Φ_θ scaling only appears at larger stability for both isotropic and one-component turbulence than suggested by the HHF curve. This *shifted* linear decrease in the relation between temperature variability and heat flux explains the large $\Delta\Phi_\theta$ (Figure 4d), i.e., the deviation of the datasets from the scaling relations. The two-component axisymmtric and one component turbulence show progressively more scatter (larger IQR) compared to isotropic turbulence but also a reduction of the value of the scaled standard deviations of velocity components ($\Phi_{u,v,w}$) and TKE dissipation rate (Φ_ε), and an increase in the scaled standard deviation of temperature (Φ_θ). This suggests that, at least for stable stratification, i.e. when mechanically produced turbulence is being damped by negative buoyancy, anisotropic turbulence tends to have a larger temperature variance but smaller velocity variances in relation to the respective fluxes. One component turbulence is again the turbulence topology that exhibits most scatter, particularly for Φ_u and Φ_ε. It is particularly interesting to note that the station with the most complex conditions (iBoxTop) shows largest deviations from scaling for unstable stratification, but on the other hand does not exhibit systematic deviations for stable stratification but rather falls within those of the other datasets.

These results confirm our initial hypothesis that anisotropy is the key variable missing from scaling relations. Indeed, anisotropy seems to provide a direction towards a unifying framework for turbulence in conditions where the assumptions of Monin-Obukhov Similarity Theory are generally violated, such as complex terrain or other sources of complex weather patterns that might affect the local flow.

Because separating the data according to anisotropy does significantly improve scaling, we first attempt to explain the large scatter between and within the original datasets observed in Figures 3 and 4 by focusing on the frequency of occurrence of a given pure state of anisotropy. This is done to examine whether local topographic dissimilarities between the locations where the datasets were taken, cause different types of turbulence topologies to occur more or less frequently and thus converge towards different scaling curves leading to large scatter if examined together. Figure 7 shows the number of averaging periods for each pure turbulence state n (separated according to anisotropy and stratification) divided by the total number of averaging periods that are unstable or stable n_{tot}. One can first note that turbulence states classified within the '*purely*' isotropic, two- and one-component regimes (cf. Figure 2) only represent a small fraction of the overall turbulence states. On average, the pure states of anisotropy jointly occur less than 40% of the time for unstable stratification and less than 10% for stable stratification. This means that only a smaller part of the data originally shown in Figure 3 fulfills the more restrictive criterion for the pure states of anisotropy (cf. Figure 5). However, it also illustrates that the more pure states of anisotropy are those that have a stronger impact on similarity scaling, attracting the data towards different scaling curves as seen in Figure 5. For example, both CASES-99 and METCRAX II have the largest proportion of isotropic turbulence, which accounts for the largest scatter in scaled horizontal velocities in Figures 3 and 4 as mentioned above. *Stiperski and Calaf* [2018] already showed that unstable isotropic turbulence occurs mostly under conditions of free convection away from the surface, which coincides with the fact that all stations with prevailing isotropic turbulence are indeed located in areas that can be expected to frequently experience conditions supportive of free convection and generally have taller towers (flat terrain and more desert-like location, *e.g.* CASES-99, METCRAX II, MATERHORN, i-Box0). On the contrary, in complex terrain and close to the surface, isotropic turbulence hardly ever occurs (*e.g.*, i-Box27). The reason for this is the fact that in complex terrain thermally driven flows, characterized by strong horizontal and vertical wind shear [cf. *Goger et al.*, 2018], develop in conditions that in flat terrain would support free convection (*i.e.,* weak shear). Given that different datasets were not only measured over different surfaces but also in different weather conditions we cannot isolate the influence of terrain on the frequency of pure anisotropic states by examining all the datasets together. Therefore, we focus next only on data from datasets obtained from multiple towers in close proximity to each other and therefore experiencing similar weather conditions (*e.g.* T-Rex, MATERHORN, i-Box). For example, T-Rex Central and West tower appear to have almost identical percentages of pure states, thus suggesting that the slope angle does not play a major role on the anisotropy type there, at least not in unstable conditions. In stable conditions, the West tower on the slope has a marginally higher prevalence of isotropic data than the Central tower on the valley floor, suggestive of more developed turbulence there. Interesting are also the two MATERHORN towers both experiencing katabatic winds during nighttime [*cf. Grachev et al.*, 2016], however MT4 has a larger incidence of pure states than the MT5 tower, possibly due to its location in a less constrained topographic surrounding (open slope). For i-Box sites, the frequency of unstable two-component axisymmetric turbulence (stable isotropic turbulence) appears to decrease (increase) with increasing terrain complexity.

Still, even this classification approach fails to identify patterns that connect anisotropy and turbulence complexity. The available methodology consequently appears to be inadequate to correctly describe the complexity of turbulence caused by both the terrain complexity (slope angle, heterogeneity, land use) and complexity of the flow conditions.

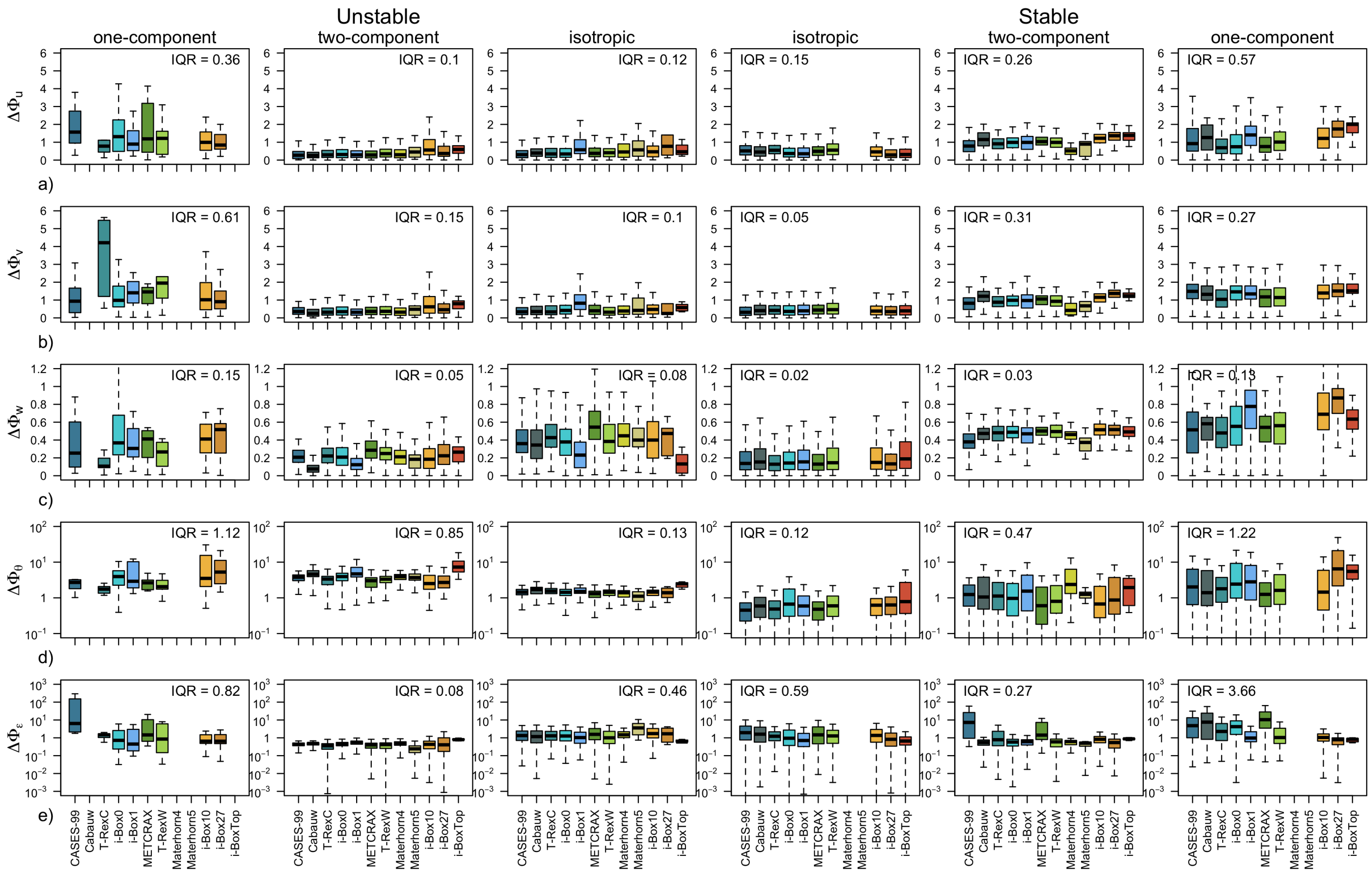


Figure 6. The same as Figure 4 but for turbulence separated according to the pure states of anisotropy. Note that here for isotropic turbulence Φ_w scaling relation is used as the reference.

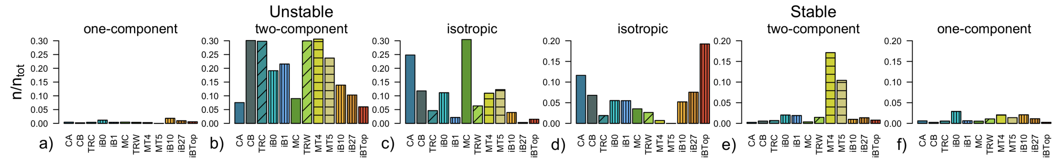


Figure 7. Frequency of occurrence (n/n_{tot}) of pure states of anisotropy (one-component, two-component axisymmetric and isotropic) for a)–c) unstable and d) – f) stable stratification. Here n is the number of averaging periods that are unstable/stable and at the same time belong to one of the pure anisotropy states, while n_{tot} is the total number of periods that are unstable/stable.

Therefore we propose instead to use the deviation from traditional similarity scaling relations (MAD) as a measure of complexity of a given dataset. The encouraging results presented above make us confident that anisotropy is the dominant process causing the departure from scaling for unstable stratification and that including information on it would improve scaling relations.We therefore hypothesize that these deviations from scaling exist due to the fact that scaling relations are estimated in the physical, streamwise Cartesian coordinate system, whereas anisotropy is defined in the eigenvector reference frame. Similarly, *Klipp* [2018] suggested calculating the friction velocity in the eigenvector space as a means of improving scaling relations. Here, we apply a different approach and instead use anisotropy as one of the explanatory variables that causes deviations from scaling curves. For stable stratification where anisotropy fails to improve scaling for two- and one-component turbulence, we identify other physical mechanisms that could be responsible for the existence of complexity. We believe, that such an objective measure of complexity will not only allow better comparison between datasets but it can provide a pathway for developing new universal scaling relations.

4 Quantifying Complexity

We now take a step back and instead of focusing only on the pure states of anisotropy, we look at all the turbulence states (including mixed states) of anisotropy together. Figure 8 shows where the centre of mass in the BLT resides for each dataset. The size of the colored triangle represents the spread of the data and is calculated as the 75th percentile in x and y direction. The centres of mass show that, for unstable stratification, data are mostly centered between the isotropic and two-component axisymmetric states, whereas for stable conditions they are more evenly spread between the two- and one-component states but generally closer to the isotropic limit. The same as with frequency of occurrence of pure states, the information on the centre of mass does not provide a conclusive information on the causes of turbulence complexity.

Complexity in the atmospheric boundary layer can be caused by a number of processes acting on a range of scales. While we use the departure from the scaling curve as a measure of complexity, the causes of this departure have to be identified manually from a number of possible processes known to be relevant in complex terrain [cf. *Serafin et al.*, 2018]. These include (but are not limited to) the influence of terrain, where the easiest measure of terrain influence is the slope angle α. Although in the analysis so far, slope angle did not show a systematic influence on scaling, the inclination of terrain can still act indirectly and this influence is therefore examined. Heterogeneity, although a significant source of complexity due to the formation of internal boundary layers as well as secondary circulations, is hard to quantify from experimental data and is therefore not examined here. Secondly, given the success of anisotropy in improving scaling in the results of the previous section, we examine anisotropy itself as a dominant variable influencing complexity. We use the coordinates of the BLT as scalar measures of anisotropy that encompases all types of anisotropy (cf. Figure 8). Here y_B represents the shortest

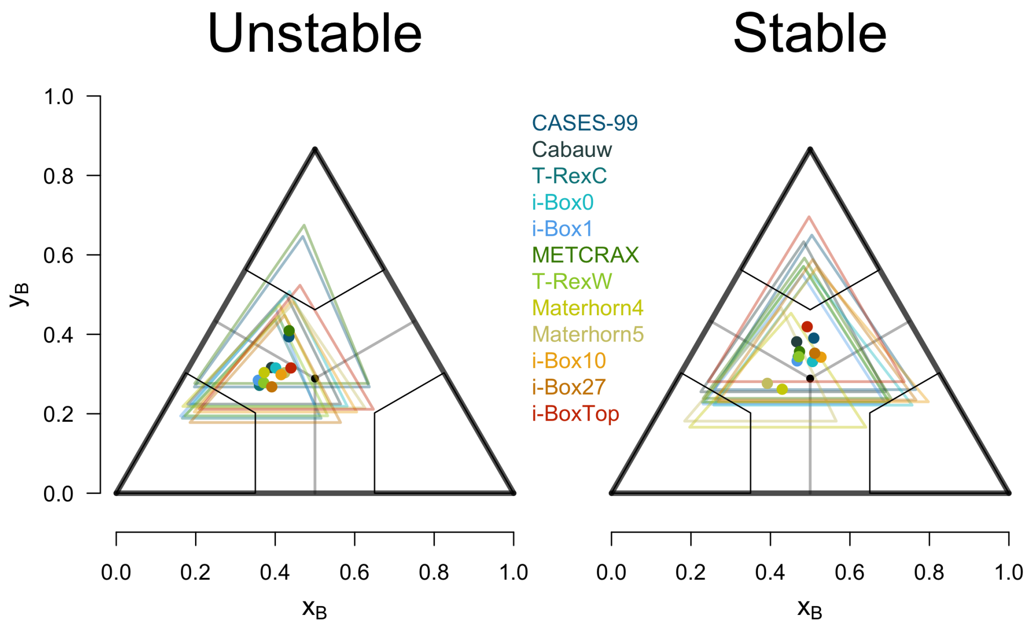


Figure 8. Barycentric Lumley Triangle showing the centre of mass for all the data points within the triangle for each dataset and stability. Different datasets are shown in color. Colored triangles represent the amount of spread of the data (calculated from the 75th percentile, and are calculated around the centre of mass.

distance to pure isotropy similarly to what was used in [*Brugger et al.*, 2018], while x_B shows where in between two-component axisymmetric and one-component state the turbulence is situated. Two mesoscale processes are ubiquitous in complex terrain: thermally-driven flows (up/down-valley, up/down-slope) and shallow water effects such as gravity waves. Thermally-driven flows are characterized by significant wind turning with height [*Rotach et al.*, 2008]. The impact of this directional shear can be measured through the angle between the streamwise $\overline{u'w'}$ and spanwise $\overline{v'w'}$ momentum flux components, defined as

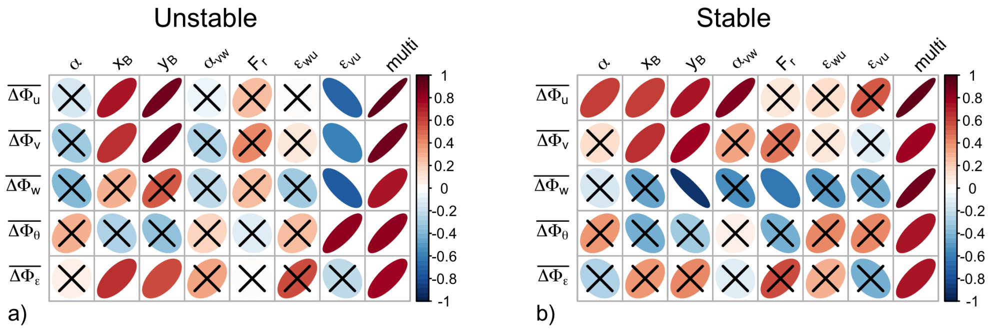
$$\alpha_{vw} = tan^{-1}(\frac{\overline{v'w'}}{\overline{u'w'}}). \tag{13}$$

If there is no directional shear, $\overline{v'w'} = 0$, then α_{vw} will also be zero as all the turbulence exchange of momentum will occur along the streamwise direction (recall that the double rotation orientates the coordinate system into the direction of the mean wind speed). The effect of wind turning on turbulence is therefore indirect, since it does not depend on driving parameters at the level where the momentum flux is measured but at heights below and above. This measure is convenient since it provides information on wind turning even if measurements are available only at one measurement level. Shallow water modes, such as gravity waves, may also affect turbulence [*Sun et al.*, 2015] and can be quantified through the Froude number

$$F_r = \frac{U}{\sqrt{gH}}, \tag{14}$$

where H is the layer depth. Given that we have no way of determining the depth scale H from the available measurements, it has to be parameterized. For this purpose we use a modified boundary layer height following *Zilitinkevich et al.* [2012],

$$H = \frac{\overline{w'w'}}{\sqrt{|f\overline{w\theta}g/\theta|}}, \tag{15}$$



Vars	Linear Regression	Multilinear Regression
$\overline{\Delta\Phi_u}$	$\mathbf{y_B}(0.79)$, $\mathbf{x_B}(0.54)$, $\varepsilon_{\mathbf{vu}}(0.5)$	$\mathbf{y_B} + \varepsilon_{\mathbf{vu}}(0.91)$
$\overline{\Delta\Phi_v}$	$\mathbf{y_B}(0.8)$, $\mathbf{x_B}(0.45)$, $\varepsilon_{vu}(0.45)$	$\mathbf{y_B}(0.8)$
$\overline{\Delta\Phi_w}$	$\varepsilon_{\mathbf{vu}}(0.53)$, $y_B(0.31)$	$\varepsilon_{\mathbf{vu}}(0.53)$
$\overline{\Delta\Phi_\theta}$	$\varepsilon_{\mathbf{vu}}(0.66)$, $y_B(0.1)$	$\varepsilon_{\mathbf{vu}}(0.66)$
$\overline{\Delta\Phi_\varepsilon}$	$\mathbf{x_B}(0.43)$, $\mathbf{y_B}(0.35)$, $\varepsilon_{wu}(0.32)$	$\mathbf{y_B} + \varepsilon_{\mathbf{wu}}(0.56)$

c)

Vars	Linear Regression	Multilinear Regression
$\overline{\Delta\Phi_u}$	$\alpha_{\mathbf{vw}}(0.74)$, $\mathbf{y_B}(0.57)$, $\mathbf{x_B}(0.405)$, $\alpha(0.39)$	$\alpha_{\mathbf{vw}} + \mathbf{y_B}(0.87)$
$\overline{\Delta\Phi_v}$	$\mathbf{y_B}(0.62)$, $\mathbf{x_B}(0.45)$, $F_r(0.2)$	$\mathbf{y_B}(0.62)$
$\overline{\Delta\Phi_w}$	$\mathbf{y_B}(0.81)$, $\mathbf{F_r}(0.41)$, $\alpha_{vw}(0.3)$	$\mathbf{y_B}(0.81)$
$\overline{\Delta\Phi_\theta}$	$F_r(0.19)$, $x_B(0.18)$, $\varepsilon_{wu}(0.18)$	$\mathbf{F_r} + \varepsilon_{\mathbf{wu}}(0.53)$
$\overline{\Delta\Phi_\varepsilon}$	$F_r(0.31)$, $x_B(0.18)$, $\varepsilon_{vu}(0.17)$	$\mathbf{y_B} + \varepsilon_{\mathbf{vu}}(0.52)$

d)

Figure 9. a) - b) Correlation matrix between the observed deviations from scaling ($\overline{\Delta\Phi_y}, y = u, v, w, \theta, \varepsilon$) and the relevant variables representing physical processes for a) unstable and b) stable stratification. The correlation coefficients are shown in color and shape so that perfectly circular shape means zero correlation and perfect line has correlation coefficient equal to 1. Crosses signify correlations that are not statistically significant at a 5 % level. c) - d) List of statistically significant ($p < 0.05$) variables and their respective R^2 values from the linear and multilinear regression with $\overline{\Delta\Phi_y}, y = u, v, w, \theta, \varepsilon$ for c) unstable and d) stable stratification. Variables that are statistically significant are shown in bold.

where f is the Coriolis parameter, and we use the vertical velocity variance instead of friction velocity, following *Monti et al.* [2002].

Finally, the influence of the smaller-scale anisotropy *i.e.* anisotropy in the inertial subrange [cf. *Katul et al.*, 1995; *Toschi et al.*, 2000; *Poggi et al.*, 2003], can be estimated through the ratio of turbulence dissipation rates

$$\varepsilon_{vu} = \varepsilon_v/\varepsilon_u, \varepsilon_{wu} = \varepsilon_w/\varepsilon_u. \tag{16}$$

Here ε_i is the dissipation rate as determined from the spectral density in the inertial subrange of the velocity component i. *Babić and Rotach* [2018] have shown that over heterogeneous surfaces these ratios, particularly ε_{wu}, can deviate strongly from one.

In order to identify which of these processes are relevant in complex terrain and therefore causing largest departures from scaling we first individually employ the linear regression approach to determine the correlation between the departure from scaling $\overline{\Delta\Phi_y}, y = u, v, w, \theta, \varepsilon$ and the corresponding predictor variables ($x_B, y_B, \alpha, \alpha_{vw}, F_r, \varepsilon_{vu}, \varepsilon_{wu}$). The correlation coefficients are shown in Figure 9 for unstable and stable stratification. Secondly, a multlinear regression with the relevant variables was performed to estimate the joint influence of these processes on the departure from scaling. How many and which variables were chosen for the multilinear regression was determined in a step-wise manner. From the step-wise procedure we choose as final the combination of statistically significant variables ($p < 0.05$) with the largest R^2 and smallest Bayesian Information Criterion (*BIC*; *Wilks* [2011]). The variables in linear and multilinear regression are shown

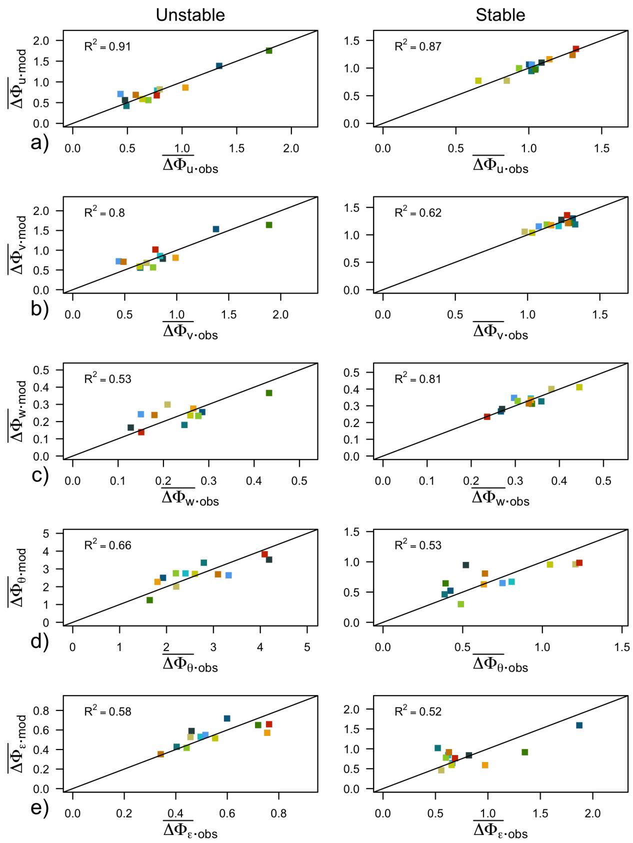


Figure 10. Comparison of observed deviations from scaling ($\overline{\Delta\Phi_{y_{obs}}}$) and those modelled by multilinear regression ($\overline{\Delta\Phi_{y_{mod}}}$) using variables from Figure 9c & d (Multilinear regression). Here $y = u, v, w, \theta, \varepsilon$.

in Figure 9 c&d, and the observed departures from scaling and the ones predicted by multilinear regression are shown in Figure 10. The largest limitation of this approach is of course the assumption of a linear relationship between predictors and the explanatory variable.

The results of both linear and multilinear regression (Figure 9) show that y_B and therefore the shortest distance to isotropic state is truly one of the most important explanatory variables both in unstable and stable stratification, confirming the results of the previous section. Indeed, y_B can explain up to 80% of variability between the complex terrain scaling relations for certain variables and stability ranges. The correlation with y_B is positive indicating that the increase of complexity coincides with the departure from isotropic conditions. In the unstable regime the other important process influencing complexity is the small-scale turbulence anisotropy shown through ε_{vu}. Although the value of this ratio is within the 20% margin of one and therefore could be considered almost constant (not shown), the individual values show a negative correlation with increasing complexity (apart from scaled temperature), indicating interestingly that the turbulence with a higher degree of horizontal small-scale isotropy is found in more complex conditions. These results are also contrary to *Babić and Rotach* [2018] where ε_{wu} diverges more from one. In the stable regime where anisotropy was successful in improving scaling only in isotropic conditions, the mesoscale processes appear to be more important than in the unstable regime. Therefore, wind turning with height (α_{vw}) and to a lesser degree the Froude number F_r appear to additionally explain an important part of the observed complexity.

The clear connection between $\Delta\Phi_{u,v}$ and y_B comes as no surprise given previous evidence that isotropic and two-component turbulence occupy different scaling curves, so that the distance to isotropy clearly delineates stations that have different percentages of these pure states and therefore cluster around them. On the other hand, in stable stratification, it is the deviation of vertical velocity variance $\Delta\Phi_w$ that (anti)correlates best with y_B suggesting that unlike in unstable stratification, it is in the vertical velocity component that anisotropy shows largest differences. This is intuitive given that as stratification increases, the vertical velocity variance decreases from isotropic towards one component with a progressively lower neutral limit, as observed in Figure 5. The importance of distinguishing between the very anisotropic states (two and one-component, *i.e.*, x_B) is also clearly identified as important for stable stratification, where one-component turbulence occurs more frequently. The existence of wind turning with height appears to have the largest influence on $\Delta\Phi_u$ in stable conditions, whereas scalar variances appear to be most affected by gravity waves. The fact that the results are dependent on the ratio of the dissipation rates points to the scale-dependence of anisotropy and the persistence of anisotropy to very small scales. *Toschi et al.* [2000] have shown that this might be due to the effect of wind shear persisting across all scales.

Figure 9 shows that the multlinear combination of the above identified processes explains the majority of the variance for the standard deviation of velocity components, and to a lesser degree of the scalar variances (temperature and TKE dissipation rate) indicating the dissimilarity between the momentum and scalars *Brutsaert* [1982], but also pointing towards missing processes that have either not been identified by our limited list or are non-linear and therefore are not detected by the linear method.

5 Discussion and conclusions

The results of the previous sections have highlighted the importance of anisotropy in shaping the scaling relations, and have therefore shown that the approach of *Stiperski and Calaf* [2018] accounting for anisotropy, significantly improves scaling even over highly complex terrain. The large site-to-site variability in turbulence structure commonly found over complex terrain was then shown to be due to the differences in the frequency of occurrence of each anisotropy type, causing large scatter in the data, as different anisotropy states follow different scaling curves. For unstable stratification, anisotropy itself was found to be the dominant processes causing the failure of scaling among the various datasets. Whether the isotropic state or the two-component turbulence should actually be taken as the reference state in unstable stratification remains a question - especially since the atmospheric boundary layer turbulence is determined by the interplay between shear- and buoyancy-dominated turbulence (and the former is by definition anisotropic). Given the fact that two-component turbulence is more prevalent in unstable stratification and the data fit classic scaling relations better, suggests that the two-component limit is the reference state for unstable stratification. For stable stratification, isotropic turbulence is clearly the reference turbulence state, corresponding to weakly stable boundary layers with well-developed turbulence [cf. *Stiperski and Calaf*, 2018]. In stable stratification, however, we see that anisotropy itself cannot explain all the variability observed. It was therefore shown that physical mechanisms, such as wind directional shear, as well as effects of mean wind speed gradients persisting to the smallest scales and affecting turbulence in the inertial subrange, cause the complexity of turbulence. While these processes obviously occur also over flat and homogeneous terrain, they are more frequent and their effects more pronounced in complex terrain.

Another relevant issue associated with turbulence in complex terrain is the depth of the boundary layer [*Rotach and Zardi*, 2007; *De Wekker and Kossmann*, 2015; *Lehner and Rotach*, 2018], and consequently of the surface layer. Whereas we are employing local scaling (and not the Monin-Obukhov similarity scaling) and therefore do not require that the measurement levels are strictly within the surface layer, it is important to investigate the validity of this hypothesis. We therefore examine the median absolute deviations from scaling ($\overline{\Delta\Phi_y}$) as a function of measurement height for data separated according to anistropy (Figure 11). Compering the influence of height (stability) on scaling, shows that the height dependence within a given tower is on the same order as the site to site variability. Indeed, the deviations from scaling show almost no height dependence for isotropic turbulence, particularly for $\Phi_w, \Phi_\theta, \Phi_\varepsilon$. There is a larger but non-systematic variability for the horizontal velocity components under conditions of two-component axisymmetric turbulence, particularly for stable stratification, suggestive of intermittent conditions and layering associated with this type of very stable turbulence.

These results suggest that the physics represented by the newly introduced variables $y_B, \varepsilon_{vu}, \alpha_{vw}, F_r$ should be considered when working on the development of new scaling relations. Additionally, the present results suggest that the multilinear regression expressions such as found in this work could be used more effectively when determining complexity of a given dataset than traditional measures, such as slope angle. It is quite likely, however, that at least part of the remaining scatter (*cf.*, Figure 9 will be associated with spatial (horizontal) heterogeneity, which is inherent in complex terrain but difficult to assess from single-tower observations. Before expressions for accounting for complexity in scaling relations themselves are used, this study would have to be extended to more than twelve datasets to improve its statistical significance. Of particular need in this respect would be horizontally distributed and long-term turbulence measurements from tall towers over very complex mountainous terrain. It is, however, already clear that this methodology provides a direction towards a unified theory of near-surface turbulence in terrains of all kinds of complexity. This will have particularly large implications for numerical modelling of weather and climate, where turbulence is parametrized by using scaling relations developed and hence only valid over flat and horizontally homogeneous terrain.

References

Babić, K., and M. W. Rotach (2018), Turbulence kinetic energy budget in the stable boundary layer over a heterogeneous surface, *Quarterly Journal of the Royal*

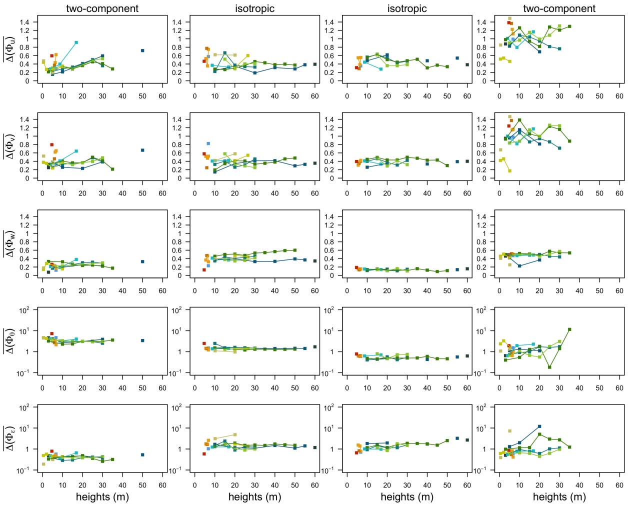


Figure 11. Median absolute deviations from scaling $\overline{\Delta\Phi_y}, y = u, v, w, \theta, \varepsilon$ as function of measurement height for unstable two-component axisymmetric and isotropic turbulence and stable isotropic and two-component axisymmetric turbulence.

Meteorological Society, doi:10.1002/qj.3274.

Babić, K., M. W. Rotach, and Z. B. Klaić (2016a), Evaluation of local similarity theory in the wintertime nocturnal boundary layer over heterogeneous surface, *Agricultural and Forest Meteorology*, *228 - 229*, 164 – 179.

Babić, N., Ž. Večenaj, and S. F. J. De Wekker (2016b), Flux–variance similarity in complex terrain and its sensitivity to different methods of treating non-stationarity, *Boundary-Layer Meteorology*, *159*(1), 123–145.

Banerjee, S., R. Krahl, F. Durst, and C. Zenger (2007), Presentation of anisotropy properties of turbulence, invariants versus eigenvalue approaches, *Journal of Turbulence*, *8*, N32.

Banerjee, T., G. G. Katul, S. T. Salesky, and M. Chamecki (2015), Revisiting the formulations for the longitudinal velocity variance in the unstable atmospheric surface layer, *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1699–1711.

Beljaars, A. C. M., and F. C. Bosveld (1997), Cabauw data for the validation of land surface parameterization schemes, *Journal of Climate*, *10*(6), 1172–1193, doi:10.1175/1520-0442(1997)010¡1172:CDFTVO¿2.0.CO;2.

Brugger, P., G. G. Katul, F. De Roo, K. Kröniger, E. Rotenberg, S. Rohatyn, and M. Mauder (2018), Scalewise invariant analysis of the anisotropic reynolds stress tensor for atmospheric surface layer and canopy sublayer turbulent flows, *Phys. Rev. Fluids*, *3*, 054,608, doi:10.1103/PhysRevFluids.3.054608.

Brutsaert, W. (1982), *Evaporation into the Atmosphere. Theory, History, and Applications*, first ed., Kluwer Academic Publishers, The Netherlands.

Chamecki, M., N. L. Dias, S. T. Salesky, and Y. Pan (2017), Scaling laws for the longitudinal structure function in the atmospheric surface layer, *Journal of the Atmospheric Sciences*, *74*(4), 1127–1147, doi:10.1175/JAS-D-16-0228.1.

Choi, K. S., and J. L. Lumley (2001), The return to isotropy of homogeneous turbulence, *Journal of Fluid Mechanics*, *436*, 59–84, doi:10.1017/S002211200100386X.

de Franceschi, M., D. Zardi, M. Tagliazucca, and F. Tampieri (2009), Analysis of second – order moments in surface layer turbulence in an alpine valley, *Quarterly Journal of the Royal Meteorological Society*, *135*(644), 1750–1765, doi: 10.1002/qj.506.

De Wekker, S. F. J., and M. Kossmann (2015), Convective boundary layer heights over mountainous terraina review of concepts, *Frontiers in Earth Science*, *3*, 77, doi:10.3389/feart.2015.00077.

Fernando, H. J. S., E. R. Pardyjak, S. D. Sabatino, F. K. Chow, S. F. J. D. Wekker, S. W. Hoch, J. Hacker, J. C. Pace, T. Pratt, Z. Pu, W. J. Steenburgh, C. D. Whiteman, Y. Wang, D. Zajic, B. Balsley, R. Dimitrova, G. D. Emmitt, C. W. Higgins, J. C. R. Hunt, J. C. Knievel, D. Lawrence, Y. Liu, D. F. Nadeau, E. Kit, B. W. Blomquist, P. Conry, R. S. Coppersmith, E. Creegan, M. Felton, A. Grachev, N. Gunawardena, C. Hang, C. M. Hocut, G. Huynh, M. E. Jeglum, D. Jensen, V. Kulandaivelu, M. Lehner, L. S. Leo, D. Liberzon, J. D. Massey, K. McEnerney, S. Pal, T. Price, M. Sghiatti, Z. Silver, M. Thompson, H. Zhang, and T. Zsedrovits (2015), The materhorn: Unraveling the intricacies of mountain weather, *Bulletin of the American Meteorological Society*, *96*(11), 1945–1967, doi:10.1175/BAMS-D-13-00131.1.

Foken, T., and B. Wichura (1996), Tools for quality assessment of surface-based flux measurements, *Agricultural and Forest Meteorology*, *78*(1), 83 – 105.

Goger, B., M. W. Rotach, A. Gohm, O. Fuhrer, I. Stiperski, and A. A. M. Holtslag (2018), The impact of three-dimensional effects on the simulation of turbulence kinetic energy in a major alpine valley, *Boundary-Layer Meteorology*, *168*(1), 1–27, doi:10.1007/s10546-018-0341-y.

Grachev, A. A., E. L. Andreas, C. W. Fairall, P. S. Guest, and P. O. G. Persson (2013), The critical richardson number and limits of applicability of local similarity theory in the stable boundary layer, *Boundary-Layer Meteorology*, *147*(1), 51–82.

Grachev, A. A., L. S. Leo, S. D. Sabatino, H. J. S. Fernando, E. R. Pardyjak, and C. W. Fairall (2016), Structure of turbulence in katabatic flows below and above the wind-speed maximum, *Boundary-Layer Meteorology*, *159*(3), 469–494.

Grubišič, V., J. D. Doyle, J. Kuettner, R. Dirks, S. A. Cohn, L. L. Pan, S. Mobbs, R. B. Smith, C. D. Whiteman, S. Czyzyk, S. Vosper, M. Weissmann, S. Haimov, S. F. J. D. Wekker, and F. K. Chow (2008), The terrain-induced rotor experiment, *Bulletin of the American Meteorological Society*, *89*(10), 1513–1533, doi:10.1175/2008BAMS2487.1.

Kaimal, J., and J. Finnigan (1994), *Atmospheric boundary layer flows: their structure and measurement*, Oxford University Press.

Katul, G. G., M. B. Parlange, J. D. Albertson, and C. R. Chu (1995), Local isotropy and anisotropy in the sheared and heated atmospheric surface layer, *Boundary-Layer Meteorology*, *72*(1), 123–148, doi:10.1007/BF00712392.

Klipp, C. (2018), Turbulent friction velocity calculated from the reynolds stress tensor, *Journal of the Atmospheric Sciences*, *75*(4), 1029–1043, doi:10.1175/JAS-D-16-0282.1.

Klipp, C. L., and L. Mahrt (2004), Fluxgradient relationship, self-correlation and intermittency in the stable boundary layer, *Quarterly Journal of the Royal Meteorological Society*, *130*(601), 2087–2103.

Kral, S. T., A. Sjöblom, and T. Nygård (2014), Observations of summer turbulent surface fluxes in a high arctic fjord, *Quarterly Journal of the Royal Meteorological Society*, *140*(679), 666–675, doi:10.1002/qj.2167.

Lehner, M., and M. W. Rotach (2018), Current challenges in understanding and predicting transport and exchange in the atmosphere over mountainous terrain, *Atmosphere*, *9*, 276, doi:doi.org/10.3390/atmos9070276.

Lehner, M., C. D. Whiteman, S. W. Hoch, E. T. Crosman, M. E. Jeglum, N. W. Cherukuru, R. Calhoun, B. Adler, N. Kalthoff, R. Rotunno, T. W. Horst, S. Semmer, W. O. J. Brown, S. P. Oncley, R. Vogt, A. M. Grudzielanek, J. Cermak, N. J. Fonteyne, C. Bernhofer, A. Pitacco, and P. Klein (2016), The metcrax ii field experiment: A study of downslope windstorm-type flows in arizonas meteor crater, *Bulletin of the American Meteorological Society*, *97*(2), 217–235, doi:10.1175/BAMS-D-14-00238.1.

Lumley, J., and G. Newman (1977), The return to isotropy of homogeneous turbulence., *J. Fluid Mech.*, *82*(1), 161–178.

Lumley, J. L. (1978), Computational modeling of turbulent flows., *Advances in Applied Mechanics*, *18*(213).

Martins, C. A., O. L. L. Moraes, O. C. Acevedo, and G. A. Degrazia (2009), Turbulence intensity parameters over a very complex terrain, *Boundary-Layer Meteorology*, *133*(1), 35–45.

Monin, A., and A. Yaglom (1971), *Statistical fluid mechanics*, vol. Volume 1, The MIT Press, Cambridge, Massachusetts.

Monti, P., H. J. S. Fernando, M. Princevac, W. C. Chan, T. A. Kowalewski, and E. R. Pardyjak (2002), Observations of flow and turbulence in the nocturnal boundary layer over a slope, *Journal of the Atmospheric Sciences*, *59*(17), 2513–2534, doi:10.1175/1520-0469(2002)059¡2513:OOFATI¿2.0.CO;2.

Nadeau, D. F., E. R. Pardyjak, C. W. Higgins, and M. B. Parlange (2013), Similarity scaling over a steep alpine slope, *Boundary-Layer Meteorology*, *147*(3), 401–419, doi:10.1007/s10546-012-9787-5.

Nieuwstadt, F. T. M. (1984a), Some aspects of the turbulent stable boundary layer, *Boundary-Layer Meteorology*, *30*(1), 31–55.

Nieuwstadt, F. T. M. (1984b), The turbulent structure of the stable, nocturnal boundary layer, *Journal of the Atmospheric Sciences*, *41*(14), 2202–2216.

Oldroyd, H. J., E. R. Pardyjak, C. W. Higgins, and M. B. Parlange (2016), Buoyant turbulent kinetic energy production in steep-slope katabatic flow, *Boundary-Layer Meteorology*, *161*(3), 405–416, doi:10.1007/s10546-016-0184-3.

Panofsky, H. A., and J. A. Dutton (1984), *Atmospheric turbulence: models and methods for engineering applications*, 397 pp., Wiley, New York.

Park, M.-S., and S.-U. Park (2006), Effects of topographical slope angle and atmospheric stratification on surface-layer turbulence, *Boundary-Layer Meteorology*, *118*(3), 613–633, doi:10.1007/s10546-005-7206-x.

Poggi, D., A. Porporato, and L. Ridolfi (2003), Analysis of the small-scale structure of turbulence on smooth and rough walls, *Physics of Fluids*, *15*(1), 35–46, doi:10.1063/1.1521728.

Pope, S. B. (2000), *Turbulent flows*, Cambridge University Press.

Poulos, G. S., W. Blumen, D. C. Fritts, J. K. Lundquist, J. Sun, S. P. Burns, C. Nappo, R. Banta, R. Newsom, J. Cuxart, E. Terradellas, B. Balsley, and M. Jensen (2002), Cases-99: A comprehensive investigation of the stable nocturnal boundary layer, *Bulletin of the American Meteorological Society*, *83*(4), 555–581.

Rotach, M. W., and D. Zardi (2007), On the boundarylayer structure over highly complex terrain: Key findings from map, *Quarterly Journal of the Royal Meteorological Society*, *133*(625), 937–948, doi:10.1002/qj.71.

Rotach, M. W., M. Andretta, P. Calanca, A. P. Weigel, and A. Weiss (2008), Boundary layer characteristics and turbulent exchange mechanisms in highly complex terrain, *Acta Geophysica*, *56*(1), 194–219, doi:10.2478/s11600-007-0043-1.

Rotach, M. W., G. Wohlfahrt, A. Hansel, M. Reif, J. Wagner, and A. Gohm (2014), The world is not flat: Implications for the global carbon balance, *Bulletin of the American Meteorological Society*, *95*(7), 1021–1028, doi:10.1175/BAMS-D-13-00109.1.

Rotach, M. W., I. Stiperski, O. Fuhrer, B. Goger, A. Gohm, F. Obleitner, G. Rau, E. Sfyri, and J. Vergeiner (2017), Investigating exchange processes over complex topography: The innsbruck box (i-box), *Bulletin of the American Meteorological Society*, *98*(4), 787–805, doi:10.1175/BAMS-D-15-00246.1.

Rotta, J. (1951), Statistische theorie nichthomogener turbulenz, *Zeitschrift für Physik*, *129*(6), 547–572, doi:10.1007/BF01330059.

Sarkar, S., and C. G. Speziale (1990), A simple nonlinear model for the return to isotropy in turbulence, *Physics of Fluids A*, *2*(1), 84–93, doi:10.1063/1.857694.

Savage, L. C., S. Zhong, W. Yao, W. J. O. Brown, T. W. Horst, and C. D. Whiteman (2008), An observational and numerical study of a regional-scale downslope flow in northern arizona, *Journal of Geophysical Research: Atmospheres*, *113*(D14), n/a–n/a, doi:10.1029/2007JD009623, d14114.

Serafin, S., B. Adler, J. Cuxart, S. F. J. De Wekker, A. Gohm, B. Grisogono, N. Kalthoff, D. J. Kirshbaum, M. W. Rotach, J. Schmidli, I. Stiperski, . Veenaj, and D. Zardi (2018), Exchange processes in the atmospheric boundary layer over mountainous terrain, *Atmosphere*, *9*(3).

Sfyri, E., M. W. Rotach, I. Stiperski, F. C. Bosveld, L. Manuela, and F. Obleitner (2018), Scalar-flux similarity in the layer near the surface layer over mountainous terrain, *Boundary-Layer Meteorology*.

Simonsen, A. J., and P. Krogstad (2005), Turbulent stress invariant analysis: clarification of existing terminology, *Physics of Fluids.*, *17*, 088,103.

Sorbjan, Z. (1987), An examination of local similarity theory in the stably stratified boundary layer, *Boundary-Layer Meteorology*, *38*(1), 63–71.

Spencer, A. (1971), Part iii. theory of invariants, *Continuum physics.*, *1*, 239–353.

Stiperski, I., and M. Calaf (2018), Dependence of near-surface similarity scaling on the anisotropy of atmospheric turbulence, *Quarterly Journal of the Royal Meteorological Society*, doi:10.1007/BF00712392.

Stiperski, I., and M. W. Rotach (2016), On the Measurement of Turbulence Over Complex Mountainous Terrain, *Boundary-Layer Meteorology*, *159*(1), 97–121.

Sun, J., C. J. Nappo, L. Mahrt, D. Belui, B. Grisogono, D. R. Stauffer, M. Pulido, C. Staquet, Q. Jiang, A. Pouquet, C. Yage, B. Galperin, R. B. Smith, J. J. Finnigan, S. D. Mayor, G. Svensson, A. A. Grachev, and W. D. Neff (2015), Review of wave-turbulence interactions in the stable atmospheric boundary layer, *Reviews of Geophysics*, *53*(3), 956–993, doi:10.1002/2015RG000487.

Thiermann, V. (1990), Optische messungen turbulenter flüsse und vorhersage der optischen turbulenz aus einfachen grenzschichtparametern, Ph.D. thesis, University of Hamburg.

Toschi, F., E. Lévêque, and G. Ruiz-Chavarria (2000), Shear effects in nonhomogeneous turbulence, *Phys. Rev. Lett.*, *85*, 1436–1439, doi:10.1103/PhysRevLett.85.1436.

Vickers, D., and L. Mahrt (2003), The cospectral gap and turbulent flux calculations, *Journal of Atmospheric and Oceanic Technology*, *20*(5), 660–672.

Wilks, D. S. (2011), *Statistical Methods in Atmospheric Sciences*, 704 pp., Academic Press.

Wyngaard, J. C. (2010), *Turbulence in the atmosphere*, 406 pp., Cambridge University Press.

Zilitinkevich, S. S., S. A. Tyuryakov, Y. I. Troitskaya, and E. A. Mareev (2012), Theoretical models of the height of the atmospheric boundary layer and turbulent entrainment at its upper boundary, *Izvestiya, Atmospheric and Oceanic Physics*, *48*(1), 133–142, doi:10.1134/S0001433812010148.

Acknowledgments

We would like to thank the providers of various datasets: the managers of the database of the Cabauw Experimental Site for Atmospheric Research (CESAR), NCARs Earth Observing Laboratory for the access to the CASES-99, METCRAX II and T-Rex datasets, Andrey Grachev and Eric Pardyjak for the access to the MATERHORN tower datasets and all the collaborators of Department of Atmospheric and Cryospheric Sciences (ACINN), who worked for many years to make the reliable and continuous data base of the i-Box study area available. The authors would also like to thank Gabriel G. Katul at the Department of Civil and Environmental Engineering, Duke University for his support and Manuela Lehner for fruitful discussions. This research was funded by Austrian Science Fund (FWF) grant T781-N32 awarded to Ivana Stiperski. Marc Calaf also acknowledges the Mechanical Engineering Department at University of Utah for start-up funds. The authors declare no conflict of interest.