# Empirical Bifurcation Analysis of Atmospheric Stable Boundary Layer Regime Occupation

by

Elizabeth Ramsey
B.Sc., University of Alberta, 2015

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

© Elizabeth Ramsey, 2021
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

We acknowledge with respect the Lekwungen peoples on whose traditional territory the university stands and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Empirical Bifurcation Analysis of Atmospheric Stable Boundary Layer Regime Occupation

by

Elizabeth Ramsey
B.Sc., University of Alberta, 2015

Supervisory Committee

---

Dr. Adam Monahan, Supervisor
(School of Earth and Ocean Sciences)

---

Dr. Hansi Singh , Departmental Member
(School of Earth and Ocean Sciences)

---

Dr. Knut von Salzen, Departmental Member
(School of Earth and Ocean Sciences)

## ABSTRACT

Turbulent collapse and recovery are both observed to occur abruptly in the atmospheric stable boundary layer (SBL). The understanding and predictability of turbulent recovery remains limited, reducing numerical weather prediction accuracy. Previous studies have shown that regime occupation is the result of the net effect of highly variable processes, from turbulent to synoptic scales, making stochastic methods a compelling approach. Idealized stable boundary layer models have shown that under some circumstances, regimes can be related to the stable branches of model equilibria, and an additional unstable equilibrium is predicted. This work seeks to determine the extent to which the SBL regime occupation can be explained using a one-dimensional stochastic differential equation (SDE). The drift and diffusion coefficients of the SDE of an input time series are approximated from the statistics of its averaged time tendencies. These approximated coefficients are fit using Gaussian Process Regression. Probabilistic estimates of the system's equilibrium points are then found and used to create an empirical bifurcation diagram without making any prior assumptions on the dynamical form of the system. This data driven bifurcation diagram is compared to modelled predictions. The analysis is repeated on several meteorological towers around the world to assess the influence of local meteorological settings. This work provides empirical insights into the nature of regime dynamics and the extent to which the SBL displays hysteresis.

# Table of Contents

# List of Tables

# List of Figures

# ACKNOWLEDGEMENTS

I would like to thank:

My dog, Sir Isaac, who insisted I take regular walks.

My partner, Geordie, for constantly trying to convince me that eating and sleeping are more important than a thesis, in addition to unwavering love and support.

My parents, for their endless patience, support, encouragement and overconfidence in my abilities.

My siblings, Hart and Jordan, mostly for comedic relief.

My supervisor, Adam Monahan, for his calm, compassion, limitless knowledge and insights into every possible subject and relentless commitment to my success.

Carsten Abraham, for sharing his passion in stable boundary layers and his guidance, assistance and knowledge.

My committee members, Dr. Hansi Singh and Dr. Knut von Salzen for their guidance and insights.

The Natural Sciences and Engineering Research Council of Canada (NSERC), for providing me with the funds to conduct this research.

A number of organizations and institutions for generously sharing incredibly large volumes of well maintained data that was of utmost importance for this project: The Royal Dutch Meteorological Institute (KMNI), The French and Italian polar institutes (IPEV and PANRA), The Meteorological Institute of the University of Hamburg, The Institute for Meteorology and Climate Research of the Karlsruhe Institute of Technology (KIT), The Los Alamos National Laboratory (LANL) and The NOAA Earth System Research Laboratory's (ESRL) Physical Sciences Division (PSD).

*The woman who follows the crowd will usually go no further than the crowd. The woman who walks alone is likely to find herself in places no one has ever been before.*
Albert Einstein

# DEDICATION

George Hartford, who inspired me to follow my morals and passions above all else.

# Chapter 1

# Introduction

The atmospheric boundary layer (ABL) is the lowest layer of the atmosphere, which directly interacts with Earth's surface through the exchange of energy, mass and momentum. Properties are mixed through turbulent exchanges, with turbulence generated by buoyant instabilities and wind shear. The ABL is affected by local influences such as topography, surface roughness, surface moisture content, albedo, local winds and cloud cover as well as mesoscale processes and large-scale weather in the free atmosphere. The height of the boundary layer is typically defined by a sharp reduction of turbulence that often occurs near a capping temperature inversion, which frequently coincides with cloud base [Arya, 2001]. The diurnal cycle of radiative surface heating by the sun results in a diurnal cycle in the ABL's static stability. Under clear sky conditions, short wave radiation efficiently heats the surface during the day, creating an energetic, buoyancy-driven layer where energy and momentum are exchanged vigorously. This is known as the convective boundary layer (CBL), with the depth of this well mixed layer ranging from 500 m to 2 km. At night, the supply of short wave radiation to the surface from the sun is removed resulting in surface cooling through the emission of longwave radiation. The surface becomes cooler than air aloft creating a temperature inversion that makes the atmosphere statically stable. Vertical mixing can then only be maintained by the mechanical production of turbulence kinetic energy (TKE). The height of the SBL reduces to 10 to 500 m. The nocturnal SBL typically forms few hours before sunset. Large scale subsidence and the advection of warm air over cold surfaces also result in the formation of SBL's [Dörenkämper et al., 2015, Arya, 2001, Mahrt, 1998, Holtslag and Nieuwstadt, 1986]. The presence of low level cloud cover (LLCC), decreases the diurnal variation in wind speeds and temperature stratification [He et al., 2013].

Early predominance of daytime measurements combined with smaller scale turbulence and strong sensitivity to terrain limited early SBL observations. Consequently the understanding of the SBL has severely lagged that of the CBL [Holtslag et al., 2013, LeMone et al., 2019]. Subsequent measurements revealed heterogeneous structures and unpredictable dynamics, with seemingly spontaneous changes in turbulent structure [LeMone et al., 2019, Monahan et al., 2015, Abraham and Monahan, 2019b, van Hooijdonk et al., 2017b]. The SBL is typically classified into two regimes, the weakly stable boundary layer (wSBL) and the very stable boundary layer (vSBL) [Mahrt, 1998, 2014, Holtslag et al., 2013, Monahan et al., 2015, Abraham and Monahan, 2019a]. The weakly stable boundary layer is characterized by continuous turbulence and weak stratification. Conditions classified as wSBL have been found to be associated with cloudy skies and moderate winds [Abraham and Monahan, 2019a, Monahan et al., 2015]. The vSBL is characterized by strong stratification and collapsed turbulence with intermittent turbulent bursts. The vSBL is associated with clear skies and weak winds [Abraham and Monahan, 2019a, Mahrt, 2014].

Monin-Obukhov similarity theory (MOST) is the classical framework used to model the surface layer of the ABL [Foken, 2006, LeMone et al., 2019, Mahrt, 1998, 2014, Holdsworth et al., 2016, Grachev et al., 2013]. In MOST, it is assumed that turbulence does not vary horizontally and decreases with height, scaled by the Obukhov length scale, $L$, according to quasi-empirically derived similarity functions. In general, this forms a reasonable model for wSBL conditions, where very localized variations do not create large changes in flow dynamics. However in the vSBL, longwave radiative cooling from the surface strengthens temperature inversions, further suppressing vertical exchanges making turbulence very localized and therefore not represented in MOST frameworks. Weak winds and strong temperature inversions in the vSBL limit vertical interactions. The surface and the atmosphere can decouple creating a runaway surface cooling effect [Derbyshire, 1999, Holtslag et al., 2013]. As a result of this inhibition of vertical motion, the MOST assumption that height is a relevant scaling parameter starts to break down. Observations show intermittent turbulent bursts in the vSBL re-coupling vertical flow, preventing decoupling from occurring in the true atmosphere [Van de Wiel and Holtslag, 2003, Mahrt, 2014, Abraham and Monahan, 2019b]. These turbulent bursts are associated with processes which are often subgrid scale in operational weather and climate models such as gravity waves, low level jets,

density currents and microfronts. To prevent runaway cooling in numerical weather and climate models, boundary layer drag in the SBL is artificially increased through the use of a synthetic minimum diffusivity and long tailed similarity functions. Enhanced mixing that is not justified by observations results [Holtslag et al., 2013]. Artificially enhanced mixing results in the erosion of low level jets, overestimation of SBL depth and an underestimation of the turning height of wind in the lower atmosphere but increases forecast skill overall [Holtslag et al., 2013]. This simplified treatment of SBL dynamics in models decreases the accuracy of near surface condition forecasts such as fog, frost and air quality [Mahrt, 1998, 2014]. Transportation, agricultural and renewable energy industries increasingly rely on accurate forecasts to optimize resource utilization and improve public health and safety [Holtslag et al., 2013, Sommerfeld et al., 2019]. Therefore, improving near surface model accuracy is necessary for a safe and sustainable society.

Observations show that the wSBL regime is associated with weak temperature inversions and high wind speeds whereas the vSBL regime is associated with strong temperature inversions and low wind speeds. No single variable has been found to predict SBL regime transitions [Abraham and Monahan, 2019a, Monahan et al., 2015]. Transitions between regimes, both from wSBL to vSBL (turbulent collapse) and from vSBL to wSBL (turbulent recovery) are frequent and both types of regime transition can occur within a single night [Abraham and Monahan, 2019c]. Some nights are observed to occupy a single regime whereas other nights experience multiple regime transitions [Abraham and Monahan, 2019b]. Transitions occur abruptly and have not been reliably predicted [Abraham and Monahan, 2019c, Kaiser et al., 2020, van Hooijdonk et al., 2017b]. In a dynamical systems context, unpredictable transitions are often found to be associated with bifurcations.

While traditional formulations rely upon the Monin-Obukhov length scale, $L$, or a critical Richardson number, $R_c$ to predict regime occupation and transitions, many observational studies have shown that these do not uniquely determine regime occupation in the atmosphere [Monahan et al., 2015]. The concept of a maximum sustainable heat flux (MSHF) was introduced in van de Wiel et al. [2012a,b], offering a new conceptual framework for modeling SBL regime occupation. This framework is an outcome of a single column SBL model introduced in van de Wiel et al. [2007]. It was shown that turbulence can be maintained if turbulent heat flux, parameterized

by MOST, balances net radiative surface cooling for a given wind shear. Otherwise turbulence collapses, surface temperature decreases and stability increases, creating vSBL conditions. A critical wind speed at which turbulence collapses can then be derived. Using meteorological tower observations from Cabauw, van Hooijdonk et al. [2015] and Monahan et al. [2015] show that regimes can be reasonably separated using MSHF. However the MSHF model is highly idealized and assumes a fixed surface heat flux.

In van de Wiel et al. [2007] a conceptual single column momentum and energy budget model for the SBL was introduced and is developed further in van de Wiel et al. [2017] with the addition of an energetic surface coupling parameter, $\lambda$. This model will hereafter be referred to as the vdW17 model. The model fixes wind speed at a reference height allowing wind speed to be a constant, external parameter to the system. Flow therefore is not affected by turbulence and vertical structure is ignored. The equilibrium solutions to this conceptual model for various model parameters are considered in van de Wiel et al. [2017]. The vSBL is associated with a stable equilibrium branch (using the term "stable" in a dynamical systems context) at low wind speeds and large temperature inversions. The wSBL is associated with a stable equilibrium branch at high wind speeds and weak temperature inversions. At low values of $\lambda$, the stable equilibrium temperature inversion values differ widely and are separated by an unstable equilibrium branch (in a dynamical systems sense). For a relatively narrow range of wind speed values, both stable branches coexist with the unstable branch. For large values of $\lambda$, the stable branches are weakly separated and transitions between the two stable equilibrium branches are gradual, creating a stable transitional equilibrium branch. Similar dependencies of the equilibrium structure on model parameters are observed for changes in isothermal net radiation and surface roughness. The existence of a bifurcation point between regimes could explain the abrupt transitions observed. In this dynamical systems context, transitions are initiated by perturbations from processes not accounted for in this simplified single column model.

Kaiser et al. [2020] consider the vdW17 model for predicting regime transitions as bifurcations in the SBL using time series data. The approach outlined in Faranda et al. [2014] was considered, where early warning indicators of regime transitions are taken as changes to the parameters of an autoregressive moving average model

(ARMA) fit to time series data. This method allows for a statistical quantification of the dynamic stability of observables, where higher-order ARMA models are taken as less stable. This methodology is a continuation of the early warning system indicator method proposed in Scheffer et al. [2009], where bifurcations are detected as increases in autocorrelation times as the system approaches a critical transition. The approach was tested using a model inspired by the vdW17 model. A stochastic, additive white noise term was added to the model to allow for small fluctuations in temperature inversions due to processes not accounted for in the highly simplified model. Data from meteorological towers in Dumosa, Australia and Dome C, Antarctica were then considered. It was determined that a sampling frequency of at least 1 minute is required to reliably predict regime transitions using this methodology. The results from the Dumosa tower suggest the existence of one dynamically stable branch and one dynamically unstable branch, while the Dome C tower is found to have insufficient time resolution to detect unstable branches. The results of Kaiser et al. [2020] rely on the robustness of the association of relatively complex ARMA models with dynamically unstable equilibria.

The present work seeks to determine the extent to which an entirely data-driven approach can be used to represent observed variability of the temperature inversion as a low-dimensional stochastic dynamical system, using a methodology that is distinct from that of Kaiser et al. [2020]. Rather than fit the highly-idealized vdW17 model to observations, a generalized data-driven one dimensional stochastic differential equation (SDE) is considered instead. Assuming white noise allows the SDE to be parameterized numerically using time series data following the method proposed in Siegert et al. [1998]. It is assumed that wind speeds are quasi-stationary and therefore can be taken as external parameters on which observed inversion tendencies are conditioned. Time series data of temperature inversions at night are conditioned on wind speed bins giving drift and diffusion coefficient distributions for each binned wind speed. Drift and diffusion coefficients are then fit separately to Gaussian process regression (GPR) models. Equilibrium values at each wind speed bin are then computed from the GPR fits to drift coefficients. This analysis results in an empirical bifurcation curve of equilibrium temperature inversions as a function of binned wind speeds. The use of GPR fits allows for uncertainty quantification of the estimated equilibria. The empirically derived diffusion coefficients provide insights into the potential state-dependence of a stochastic term that would be appropriate for this

simplified one dimensional representation of the SBL. Similar to Kaiser et al. [2020], the methodology is tested by considering the vdW17 model with the addition of a stochastic term, representing neglected "subgrid-scale" processes theorized to instigate transitions between regimes.

The vdW17 model is introducted and nondimensionalized in Chapter 2, followed by a description of the estimation method proposed. In Chapter 3, the vdW17 model is used to asses this methodology and its sensitivities. Observations at six meteorological towers are then considered in Chapter 4. The meteorological tower at Cabauw, Netherlands provides a long record with a broad set of meteorological state variables. The data tower at Hamburg, Germany provides a large dataset recorded at a relatively high temporal resolutions. The Dome C meteorological tower in Antarctica provides insights into polar SBL dynamics. Meteorological towers at Los Alamos, USA, Boulder, USA and Karlsruhe, Germany are also considered. Meteorological tower derived equilibrium structures and diffusion coefficients are compared and related to methodological constraints identified in Chapter 4. Conclusions are summarized in Chapter 5.

# Chapter 2

# Methods

## 2.1 The van de Wiel (2017) Model

A key limitation of the MSHF framework introduced in van de Wiel et al. [2012a,b] is that it cannot predict SBL behaviour when wind speeds fall below a critical wind speed value. It was hypothesized that the lack of such an equilibrium state results from the fixed surface heat flux, and that negative feedback mechanisms such as soil heat transport prevents runaway longwave radiative cooling, allowing for a new thermodynamic equilibrium to be reached. This equilibrium state would correspond to the vSBL regime, whereas the equilibrium found at larger wind speeds would correspond to the wSBL regime. These feedbacks are included in the vdW17 model.

The vdW17 model considers a highly idealized one dimensional surface energy budget. It is assumed that there are no shortwave radiation contributions as this model considers only nocturnal SBL conditions. Stable boundary layers that occur during the day (due e.g. to advection) are not considered in this conceptual model. The model begins with the notion of the velocity crossing point, introduce in van de Wiel et al. [2012a] and van de Wiel et al. [2012b]. In late afternoon or at sunset, when the boundary layer is nearly neutral, the near-surface wind speed profile is approximately logarithmic. As the night progresses, turbulent momentum fluxes weaken and near surface winds slow due to reduced downward momentum transport. Aloft, winds accelerate due to flow decoupling. At some height in between, wind speeds are approximately constant at what is known as the velocity crossing point, $z_h$. The velocity crossing point can then be taken as a reference height, $z_h$ for the

model, allowing wind speed, $U$ at the velocity crossing point to be taken as an external parameter to the system. It is assumed in this model that wind speeds at the reference height are constant during the night. The temperature inversion is taken as the temperature difference between $z_h$ and the surface, $\Delta T = T_h - T_s$. The temperature, $T_h$ is assumed to be constant, so that changes in $\Delta T$ result from changes in $T_s$, as modelled by the surface energy budget. In reality, $\Delta T$ can change both due to changes in $T_s$ and $T_h$. The surface energy balance equation is expressed as,

$$C_v \frac{d\Delta T}{dt} = Q_n - G - H \tag{2.1}$$

where $C_v$ is the heat capacity of the surface per unit area (J m$^{-2}$ K$^{-1}$), $Q_n$ is net longwave radiation flux density (J m$^{-2}$ s$^{-1}$), $H$ is the surface turbulent sensible heat flux density (J m$^{-2}$ s$^{-1}$), and $G$ is the surface soil heat flux density (J m$^{-2}$ s$^{-1}$). Sign conventions are defined such that fluxes that tend to increase $T_s$ are positive. A diagram of this model is shown in Figure 2.1.

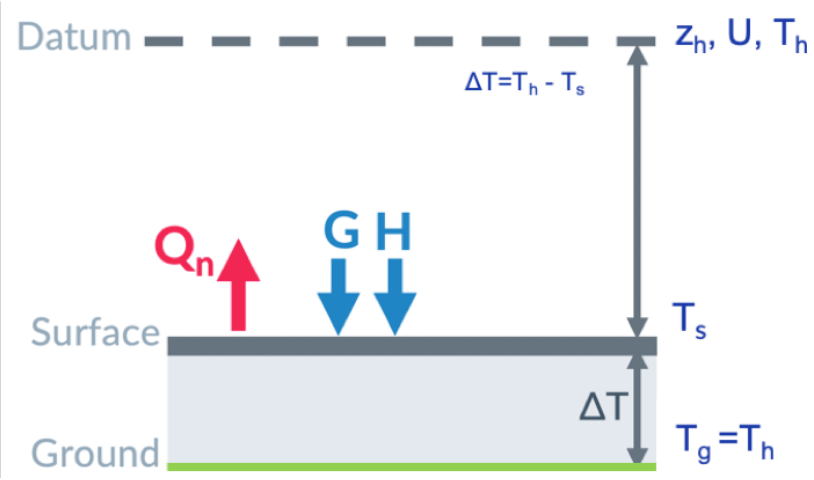



Figure 2.1: Illustration of the vdW 17 model where $Q_n$, $G$ and $H$ represent net longwave radiative flux density, surface soil heat flux density and sensible heat flux respectively. The subscript $h$ denotes the reference height, the subscript $s$ denotes the surface and $g$ denotes the subsurface (ground).

The fluxes $Q_n$ and $G$ are treated as quantifiable quantities in this model. A linear ground heat flux relationship is assumed using the surface conductance parameter, $\lambda_s$ and the temperature gradient, such that well coupled surfaces and strong temperature differences between the soil and the surface, $\Delta T_g = T_g - T_s$, result in increased

flux. The soil heat flux can then be parameterized as $G = \lambda_s \Delta T_g$ such that stronger temperature inversions are assumed to result in increased soil heat flux. It is further assumed for simplicity that the soil temperature, $T_g$ is at the same temperature as the velocity crossing point, $T_h$, and therefore $\Delta T_g = \Delta T$.

The surface and the atmosphere are modelled as gray bodies with emissivities $\varepsilon_h$ and $\varepsilon_s$ respectively. The net longwave radiation can then be derived from the Stefan-Boltzmann law,

$$Q_n = \varepsilon_s \sigma T_s^4 - \varepsilon_h \sigma T_h^4 \tag{2.2}$$

where $\sigma$ is the Stefan-Boltzmann constant. For simplicity, it is assumed that the emissivity of the surface is approximately 1, which is true of most land surfaces. The net radiative flux density, $Q_n$ is linearized with respect to $T_h$, leaving

$$Q_n \approx Q_i + \lambda_r \Delta T \tag{2.3}$$

where $Q_i$ is the so-called isothermal net radiation and $\lambda_r = 4\varepsilon_h \sigma T_h^3$ determines the rate of internal energy exchange due to the temperature inversion. The use of isothermal net radiation, $Q_i$ as opposed to the net radiative flux, $Q_n$ allows for radiative losses due to external processes to be separated from radiative exchanges due to internal processes. Surface exchange feedbacks are combined by defining a lumped surface conductance, $\lambda = \lambda_r + \lambda_s$. The parameter $\lambda$ characterizes the thermodynamic coupling between the surface and the atmosphere. Low values of $\lambda$ correspond to insulating surfaces such as ice whereas high values of $\lambda$ correspond to well coupled surfaces such as moist, vegetated surfaces.

Turbulent sensible heat flux is parameterized as

$$H = \rho c_p c_D U \Delta T f(Rb) \tag{2.4}$$

where $\rho$ (kg m$^{-3}$) and $c_p$ (J kg$^{-1}$ K$^{-1}$) are the air density and heat capacity at a constant pressure, $c_D = \left[\kappa / \ln\left(z_h / z_0\right)\right]^2$ is the neutral drag coefficient with $\kappa \approx 0.4$ and $z_0$ is the roughness length. The effect of turbulence on static stability is defined with a stability function $f(R_b)$, where $R_b$ is the bulk Richardson number, $R_b = z_h \left(g/T_h\right)\left(\Delta T / U^2\right)$. The Businger-Dyer formulation is primarily considered,

$$f(R_b) = \begin{cases} \left(1 - \left(\frac{R_b}{R_c}\right)\right)^2 & R_b \leq R_c \\ 0 & R_b > R_C. \end{cases} \tag{2.5}$$

with $R_c = 0.2$. The resulting, parameterized model is

$$C_v \frac{d\Delta T}{dt} = Q_i - \lambda \Delta T - \rho c_p c_D U \Delta T f\left(R_b\right) \tag{2.6}$$

.

Equilibrium solutions to Equation 2.6, with a Businger-Dyer stability function are shown in Figure 2.2 for a range of different $\lambda$ values.

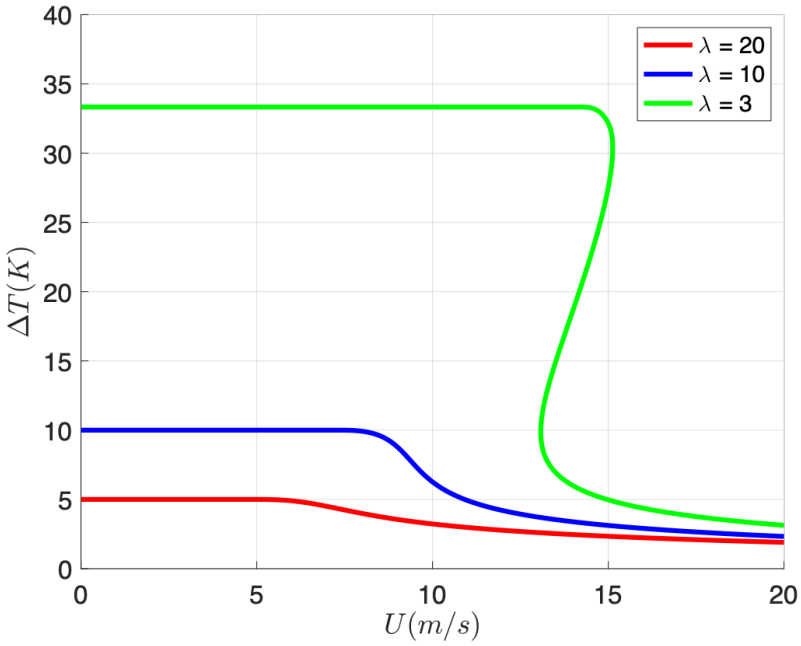



Figure 2.2: Equilibrium solutions to the vdW17 model for different values of the surface coupling parameter $\lambda$ (W m $^{-2}$ K$^{-1}$). Parameter values used to create this figure were obtained from van de Wiel et al. [2017].

A vSBL-like equilibrium branch at low wind speeds and large temperature inversions, corresponding to calm conditions, appears as well as an equilibrium branch at high wind speeds and weak inversions, corresponding to the wSBL. The vSBL equilibrium branch occurs at smaller temperature inversions when surface coupling is larger, due to the negative feedback effects of soil heat flux. As $\lambda$ decreases, the temperature inversion of the vSBL equilibrium branch increases (and is unbounded

as $\lambda \to 0$). Additionally, for sufficiently small $\lambda$ an unstable branch at intermediate wind speed values appears. From a dynamical systems context, the existence of an unstable branch in this region could explain the observed abruptness of regime transitions. Intermittent turbulent bursts have been observed in the vSBL [Abraham and Monahan, 2019a, Mahrt, 2014, Van de Wiel and Holtslag, 2003]. These bursts are hypothesized to provide the energy to excite regime transitions [van der Linden et al., 2020]. Therefore the addition of a stochastic term to the model is considered in Section 2.1.2. This work seeks to determine the extent to which this conceptual framework for the SBL holds in observations.

### 2.1.1 Nondimensionalized vdW17 Model

In this study a nondimensionalized form of the vdW17 model is considered. The dimensional scales of the model are defined as

$$\begin{aligned} T_{sc} &= T_r \text{ (temperature)} \\ L_{sc} &= \frac{c_v}{\rho c_p} \text{ (length)} \\ v_{sc} &= \left(g z_r\right)^{1/2} \text{ (speed)} \\ t_{sc} &= \frac{c_v}{\rho c_p \left(g z_r\right)^{1/2}} \text{ (time) .} \end{aligned} \tag{2.7}$$

The nondimensional form of Equation 2.6 becomes

$$\frac{dx}{ds} = \hat{Q} - \hat{\lambda} x - c_d \hat{U} f\left(\frac{x}{\hat{U}^2}\right) x \tag{2.8}$$

with nondimensional time, $s = \frac{t}{t_{sc}}$, and the nondimensional state variable, $x = \frac{\Delta T}{T_{sc}}$. Carets denote non dimensionalized parameters, summarized in Table 3.1, and $f(x/\hat{U}^2)$ represents the stability function. Note that the nondimensionalization in this study differs from that in Kaiser et al. [2020] in order to accommodate the non linear dependence of the stability function on wind speed, and to allow for $\hat{Q}$, $\hat{\lambda}$ and $\hat{U}$ to be retained as bifurcation parameters. For the latter reason none of $Q$, $\lambda$ or $U$ are used in defining the dimensional scales.

### 2.1.2 A Stochastic Generalization of the vdW17 Model

It is possible that small scale perturbations to the SBL may be related to regime transitions since no single state variable has been found to predict regime transitions [Mahrt, 2014, Monahan et al., 2015, Abraham et al., 2019, Kaiser et al., 2020]. Observations of the SBL show intermittent, stochastic turbulent bursts, particularly in the vSBL which are not explicitly accounted for in the vdW197 model [Van de Wiel and Holtslag, 2003, Mahrt, 2014]. Many other processes are not represented in the simplified form of the vdW model. This idea motivates the addition of a stochastic term to the vdW17 model to represent subgrid scale processes in the SBL. A general stochastic differential equation in Itô form is given by,

$$\frac{d}{dt}\mathbf{X}(t) = \mathbf{f}\left(\mathbf{X}(t), t\right) + \mathbf{g}\left(\mathbf{X}(t), t\right)\dot{\mathbf{w}} \tag{2.9}$$

where $\mathbf{X}(\mathbf{t})$ is a time-dependent $d$ dimensional stochastic process, $\mathbf{f}(\mathbf{X}\left(t\right), t)$ is the deterministic part of the tendency known as the drift, and $\dot{\mathbf{w}}$ represents a vector of $r$ mutually independent Gaussian white noise processes. The diffusion matrix, $\mathbf{g}\left(\mathbf{X}(t), t\right)$ is a $d \times r$ matrix defining the form of the stochastic part of the tendency. If the diffusion coefficient is a function of time only, $\mathbf{g}(t)$, then $\mathbf{X}(t)$ is said to be an additive noise process. The diffusion coefficient can also be a function of the state of the system, $\mathbf{g}(\mathbf{X}(t), t)$, defining multiplicative noise. The use of white noise in this formulation assumes each of the $r$ independent noise processes are uncorrelated between any two times. White noise is a good approximation if the timescale of the unresolved processes are short relative to the dynamical timescale. Noise processes with a non-zero autocorrelation time, $\tau_\zeta$ and variance $\sigma_r^2/2\tau_z$ can be considered by defining an Ornstein-Uhlenbeck (red noise) process, $\zeta$, as

$$\begin{aligned} \dot{\zeta} &= -\frac{1}{\tau_\zeta}\zeta + \frac{\sigma_r}{\tau_\zeta}\dot{w} \\ \frac{d}{dt}\mathbf{X}(t) &= \mathbf{f}\left(\mathbf{X}(t), t\right) + \mathbf{g}\left(\mathbf{X}(t), t\right) h(\mathbf{X}(t), \zeta(t), t) \end{aligned} \quad . \tag{2.10}$$

where $h(\mathbf{X}(t), \zeta(t), t)$ defines the red noise process. In the limit that $\tau_\zeta \to 0$, $\zeta$ becomes white noise.

The simplest extension of the vdW17 model to include stochastic processes involves addition of a diffusion term, $g(\Delta T, t)$ to Equation 2.6 as,

$$\frac{d}{dt}\Delta T = f(\Delta T, t) + g(\Delta T, t)\dot{\mathbf{w}} \tag{2.11}$$

where $f(\Delta T, t)$ is the drift term,

$$f(\Delta T, t) = \frac{1}{C_v}[Q_i - \lambda \Delta T - \rho c_p c_D U \Delta T f\left(R_b\right)] \tag{2.12}$$

Initially additive white noise is considered, similar to the model used in Kaiser et al. [2020]. The form of the additive stochastic vdW17 model in nondimensional form is

$$\frac{dx}{ds} = \hat{Q} - \hat{\lambda} x - c_d \hat{U} f\left(\frac{x}{\hat{U}^2}\right) x + \sigma \dot{w} \tag{2.13}$$

where $\sigma$ is the diffusion coefficient of the model and $\dot{w}$ is Gaussian white noise. A version of the vdW17 model including state-dependent red noise is introduced in Section 3.4.2.

## 2.2 Empirical Reconstruction Procedure

### 2.2.1 Data Driven Stochastic Parameterization

As discussed in van de Wiel et al. [2017] and Holdsworth and Monahan [2019], the vdW17 model is highly dependent on parameters that are difficult to accurately quantify, such as the heat capacity of the surface, $C_v$, the surface feedback coefficient, $\lambda$, and the roughness length. $z_0$. The model also parameterizes the effect of atmospheric stability on turbulent heat transport using the semi empirical function, $f(R_b)$, for which several forms have been proposed [Holdsworth et al., 2016, van de Wiel et al., 2017]. The functional form of this highly idealized near surface atmospheric energy model is therefore uncertain. Rather than try to directly fit this model to observations, to investigate the effective low-dimensional dynamics of the inversion strength an entirely data driven approach is taken. Estimating equilibrium curves using a data driven approach will provide empirical insights into the extent to which the SBL displays effective low dimensional dynamics similar to those of the conceptual framework of the vdW17 model. I assume only that the structure of the SBL is a function of the temperature inversion conditioned on wind speed at the velocity crossing height. The nature of noise related to the diffusion term also remains uncertain, as explored

in Abraham et al. [2019]. Deriving a diffusion term from data will help guide future conceptualizations of noise processes in this system.

A data driven method of extracting model equations for stochastic systems from time series data, without assuming functional forms for the drift and diffusion, is presented by Siegert et al. [1998] and Friedrich et al. [2000]. If the drift and diffusion terms of Equation 2.9 do not explicitly depend on time and noise is white, knowing the state of the system at one time in the past is sufficient to represent future states, defining a Markovian system. Therefore the statistics of the state variable $\mathbf{X}$ evolving from a state $\mathbf{X}(t)$ to a state $\mathbf{X}(t+\tau)$ can be used to empirically determine the drift and diffusion coefficients. It can be shown that in the limit of $\tau \to 0$ the drift coefficient is given by the conditional mean of the time tendency of the system,

$$\mathbf{f}(\mathbf{X}) = \lim_{\tau \to 0} \frac{1}{\tau} \langle \mathbf{X}(t+\tau) - \mathbf{X} \rangle \Big|_{\mathbf{X}(t)=\mathbf{X}} \tag{2.14}$$

and the coefficient of the stochastic forcing is given by the limiting conditional covariance,

$$\begin{aligned} \mathbf{g}(\mathbf{X})\mathbf{g}^{\mathrm{T}}(\mathbf{X}) = & \lim_{\tau \to 0} \frac{1}{\tau} \langle (\mathbf{X}(t+\tau) - \mathbf{X}) \\ & \times (\mathbf{X}(t+\tau) - \mathbf{X})^{\mathrm{T}} \rangle \mid_{\mathbf{X}(t)=\mathbf{X}} \end{aligned} \tag{2.15}$$

where $\langle X \rangle$ denotes the expectation operator. Rigorous proofs of equations 2.14 and 2.15 can be found in Gardiner [1990]. The reconstructed SDE is interpreted in the Itô sense, in that the system is uncorrelated with noise acting on the system at the same time. The Stratonovich form of the SDE would be more appropriate in this context, as according to the Wong-Zakai theorem real world systems approximated by white noise converge to the Stratonovich form of the SDE [Twardowska, 1996]. The SDE's are kept in Itô form as Equations 2.14 and 2.15 hold for the drift and diffusion of an Itô SDE. In principle, it is possible to transform from Itô to Stratonovich form using,

$$\begin{aligned} \tfrac{d}{dt}\mathbf{X}(t) &= \mathbf{f}(\mathbf{X}, t) - \tfrac{1}{2}\mathbf{c}(\mathbf{X}, t) + \mathbf{g}(\mathbf{X}, t)\dot{\mathbf{w}} \\ c_i(x, t) &= g_{jk}(x, t) \tfrac{\partial}{\partial x_j} g_{ik}(x, t) \end{aligned} \tag{2.16}$$

However the transformation requires calculation of the derivative of the diffusion coefficient, which may be difficult if reconstructed drifts are noisy. In the limit of additive noise, the two solutions are identical.

Given a time series of Markovian data, sampled at increments of $\Delta t$, one way of estimating the conditional averages in equations 2.14 and 2.15 is by binning $\mathbf{X}$ into $N_x$ bins of size $n_B$ using a finite difference approximation,

$$\mathbf{f}_i \simeq \frac{1}{n_B \Delta t} \sum_{b=1}^{n_B} (\mathbf{X}_{i,b+1} - \mathbf{X}_{i,b}) \tag{2.17}$$

$$\mathbf{g}_{i,j}\mathbf{g}_{j,i} \simeq \frac{1}{n_B \Delta t} \sum_{b=1}^{n_B} (\mathbf{X}_{i,b+1} - \mathbf{X}_{i,b}) \times (\mathbf{X}_{j,b+1} - \mathbf{X}_{j,b})^{\mathrm{T}} \tag{2.18}$$

.

where $\mathbf{X}_{i,b}$ represents the $\mathrm{b}^{\mathrm{th}}$ time sample of $\mathbf{X}$ in the $\mathrm{i}^{\mathrm{th}}$ bin, and $\mathbf{f}_i$ and $\mathbf{g}_i$ are the corresponding discretized drift and diffusion. The finite data sampling interval $\Delta t$, is a known limitation of this methodology [Friedrich et al., 2011]. Sura and Barsugli [2002] discuss numerical errors in drift and diffusion coefficients due to finite sampling rates by expanding the Itô Taylor series $\mathbf{X}(t + \Delta t)$ to leading order $\Delta t^2$. It is shown that finite differencing creates errors in $\mathbf{f}_i$ and $\mathbf{g}_i$ that are nonlinear and coupled with, to a leading order, a quadratic bias in $\mathbf{g}_i$. This bias will be investigated in numerical simulations in Chapter 3. Determining the appropriate number of bins to use to estimate $\mathbf{f}_i$ and $\mathbf{g}_i$ is another known limitation of this method [García et al., 2017]. Therefore, rather than bin the data over ranges of $\mathbf{X}$, conditional averages are computed through the use of Gaussian Process Regression.

### 2.2.2 Gaussian Process Regression

Gaussian Process Regression (GPR) is a probabilistic machine learning method that removes the need to bin the drift and diffusion estimates. Gaussian processes are an extension of multivariate Gaussian probability distributions to function space. From a statistical perspective, they can be thought of as defining a probability distribution over all possible functions that fit a set of points. The probabilistic nature of GPR allows for a natural uncertainty quantification of fits. Formally, a Gaussian process is an infinite collection of random variables, any finite number of which have a Gaussian distribution.

While equations 2.17 and 2.18 provide discrete estimates, $\mathbf{f_i}$ and $\mathbf{g_i}$, of the drift

and diffusion from time series data, continuous estimated functions, $\hat{f}(\mathbf{X})$ and $\hat{g}(\mathbf{X})$ are desired. GPR provides an approach with the additional advantages of producing probabilistic estimates allowing for confidence intervals to be computed, and removing the requirement to bin over values of $\mathbf{X}$. The application of GPR to stochastic parameterization is described in García et al. [2017].

Gaussian processes are completely specified by their mean $m(\mathbf{X})$ and covariance $K(\mathbf{X}, \mathbf{X}')$ functions

$$\begin{aligned} m(\mathbf{X}) &= \langle \boldsymbol{\phi}(\mathbf{X}) \rangle \\ K(\mathbf{X}, \mathbf{X}') &= \langle (\boldsymbol{\phi}(\mathbf{X}) - m(\mathbf{X}))(\boldsymbol{\phi}(\mathbf{X}') - m(\mathbf{X}')) \rangle \end{aligned} \tag{2.19}$$

giving a Gaussian process $\boldsymbol{\phi}(\mathbf{X}) \sim \mathcal{GP}(m(\mathbf{X}), K(\mathbf{X}, \mathbf{X}'))$.

The covariance function describes the dependence of predicted values, $[\phi_i, \phi_j]$ as functions of input variables, $[\mathbf{X}_i, \mathbf{X}_j]$. The covariance is defined by a kernel function, $K(\mathbf{X}_i, \mathbf{X}_j|\boldsymbol{\theta})$ which depends on hyperparameters, $\boldsymbol{\theta}$, that are estimated from the data. The choice of kernel function influences the properties of the resultant GP. The kernel function defines how smooth or rough the posterior mean will be. It is expected that the SDE drift and diffusion are smooth and functions of the state variable, however no prior knowledge of the form of the kernel function exists. It is therefore important to choose a flexible kernel that does not make strong assumptions of the form of the data. The squared exponential kernel defined as

$$K(\mathbf{X}_i, \mathbf{X}_j) = \sigma_f^2 \exp\left(-\frac{(\mathbf{X}_i - \mathbf{X}_j)^2}{2l^2}\right) + \sigma_n^2 \delta_{i,j} \tag{2.20}$$

is infinitely differentiable and therefore produces smooth realizations. For this reason it is commonly used and a suitable choice for this application [Rasmussen and Williams, 2006]. It depends on three hyperparmeters, $\boldsymbol{\theta} = [\sigma_f, l, \sigma_n]$. The quantity $\sigma_f$ defines the intrinsic standard deviation of the Gaussian Process while $\sigma_n$ represents uncorrelated measurement noise. If there is no measurement noise ($\sigma_n = 0$), then all realizations of the GP must pass through all observations. If there is uncorrelated measurement noise ($\sigma_n > 0$) then realizations of the GP should approach the observations but are not required to pass through them. The quantity $l$ is the characteristic length scale of covariance function, defining the distance over which

input variables are uncorrelated. Small values of $l$ allow predicted functions to vary substantially over small changes in input variables, creating highly variable functions, whereas large length scales create smoother, slowly varying functions. Hyperparameter values, $\boldsymbol{\theta}$ are generally estimated from data using optimization algorithms as discussed in Rasmussen and Williams [2006]. The Gaussian Process fits themselves are carried out using a Bayesian approach.

The prior mean $m(\mathbf{X})$ is often assumed to be zero. If however a prior expectation of the general structure of the data exists an explicit prior mean function $m(\mathbf{X})$ can be defined. The GP is then expressed as $\boldsymbol{\phi}(\mathbf{X}) \sim \mathcal{GP}\left(m(\mathbf{X}), K\left(\mathbf{X}, \mathbf{X}'\right)\right)$. Given a set of observations, $\{(\mathbf{x}_i, \boldsymbol{\phi}_i) \mid i = 1, \ldots, n\}$, the posterior distribution at additional prediction points, $\{\, \mathbf{x}_*, \boldsymbol{\phi}_*\}$ are desired. All variables in Gaussian processes are normally distributed, denoted $\mathcal{N}$, by definition. Therefore the joint distribution of observation and prediction points is,

$$\begin{bmatrix} \boldsymbol{\phi} \\ \boldsymbol{\phi}_* \end{bmatrix} \sim \mathcal{N}\left( \begin{matrix} m(\mathbf{X}) \\ m(\mathbf{X}_*) \end{matrix}, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K\left(X, X_*\right) \\ K\left(X_*, X\right) & K\left(X_*, X_*\right) \end{bmatrix} \right). \tag{2.21}$$

The conditional distribution of $\boldsymbol{\phi_*}$ given the distribution of $\boldsymbol{\phi_*}$ is then

$$\begin{aligned} \boldsymbol{\phi}_* \mid (\mathbf{X}_*, \mathbf{X}, \boldsymbol{\phi}) \sim \mathcal{N}(& m(\mathbf{X}) + K\left(\mathbf{X}_*, \mathbf{X}\right)\left[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I\right]^{-1}\left(\boldsymbol{\phi} - m(\mathbf{X})\right), \\ & K\left(\mathbf{X}_*, \mathbf{X}_*\right) - K\left(\mathbf{X}_*, \mathbf{X}\right)\left[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I\right]^{-1} K\left(\mathbf{X}, \mathbf{X}_*\right)\big) \end{aligned} \tag{2.22}$$

by the marginal and conditional distributions Gaussian identity.

Because the prior mean function is often unknown in advance, it can be defined using a fixed basis set, $\mathbf{b}(\mathbf{X})$ with a coefficient vector $\boldsymbol{\beta}$ that is estimated from the data. The prior on the coefficient vector, $\boldsymbol{\beta}$ is taken to be Gaussian, with mean $\bar{\boldsymbol{\beta}}$ and covariance $B$. Uncertainties in the basis function contribute to the covariance function and the corresponding prior of the GP. Before conditioning on observations, the GP then becomes $\boldsymbol{\phi}(\mathbf{x}) \sim \mathcal{GP}\left(\mathbf{b}(\mathbf{X})^\top \boldsymbol{\beta}, K\left(\mathbf{X}, \mathbf{X}_*\right) + \mathbf{b}(\mathbf{x})^\top \boldsymbol{\beta} \mathbf{b}\left(\mathbf{x}_*\right)\right)$. In the limit in which the covariance $B \to 0$, (i.e. there is no prior knowledge of $\boldsymbol{\beta}$), the mean and covariance of the GP with an explicit basis function becomes

$$
\begin{aligned}
\bar{\boldsymbol{\phi}}(\mathbf{X}_*) = \mathbf{b}(\mathbf{X}_*) &\left(B^{-1} + \mathbf{b}(\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} \mathbf{b}(\mathbf{X})^T\right)^{-1} \\
&\left(\mathbf{b}(\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} \boldsymbol{\phi} + B^{-1}\bar{\boldsymbol{\beta}}\right) \\
&+ K(\mathbf{X}_*,\mathbf{X}_*) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} \left(\boldsymbol{\phi} - \mathbf{b}(\mathbf{x})^T \bar{\boldsymbol{\beta}}\right) \\
\operatorname{cov}(\boldsymbol{\phi}_*) = K(\mathbf{X}_*,\mathbf{X}_*) &- K(\mathbf{X}_*,\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} K(\mathbf{X},\mathbf{X}_*) \\
&+ \left(\mathbf{b}(\mathbf{X}_*) - \mathbf{b}(\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} K(\mathbf{X}_*,\mathbf{X}_*\right)^T \\
&\left(B^{-1} + \mathbf{b}(\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} \mathbf{b}(\mathbf{X})^T\right) \\
&\left(\mathbf{b}(\mathbf{X}_*) - \mathbf{b}(\mathbf{X}) \left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2 I\right]^{-1} K(\mathbf{X}_*,\mathbf{X}_*)\right)
\end{aligned}
\tag{2.23}
$$

A full derivation of these equations can be found in Rasmussen and Williams [2006].

Knowing the mean and covariance, a realization of a GPR can be generated as

$$\boldsymbol{\phi}(\mathbf{X}_*) = m(\mathbf{X}_*) + \mathbf{L}\mu \tag{2.24}$$

where $\mathbf{L}$ is the Cholesky decomposition of the covariance matrix, $K(\mathbf{X},\mathbf{X}')$, and $\mu$ is a Gaussian distributed random variable of the same dimensions as $\mathbf{X}$. For this study, GPR is implemented using MATLAB's $fitgpr$ function which is discussed in Section 2.2.3.

Data sets used in this analysis have lengths of $N \approx 10^5$ to $N \approx 10^7$. GPR requires the storage and inversion of a $N \times N$ matrix, requiring $\mathcal{O}(N^3)$ operations making data sets of this size computationally prohibitive. To reduce the computational demands of this method, a subset of points $m << N$ are selected to fit the model. A variety of methods have been proposed to optimally select the $m$ points that statistically represent the $N$ data points [García et al., 2017, Rasmussen and Williams, 2006]. For this project, a subset of $m = 1000$ points that maximize the log likelihood of the prediction is used. This method showed the greatest reduction in computing time with a negligible decrease in quality of fit.

### 2.2.3 The Estimation Method

This work seeks to determine the extent to which the dynamics of the SBL temperature inversion can be empirically represented as a low-dimensional stochastic dynamical system. Having obtained a drift function, it will be possible to determine the stable and unstable equilibira of the deterministic part of the dynamics. The state-dependence of the empirically derived noise intensity can also be determined. A visualization of this approach is shown in Figure 2.2.3.

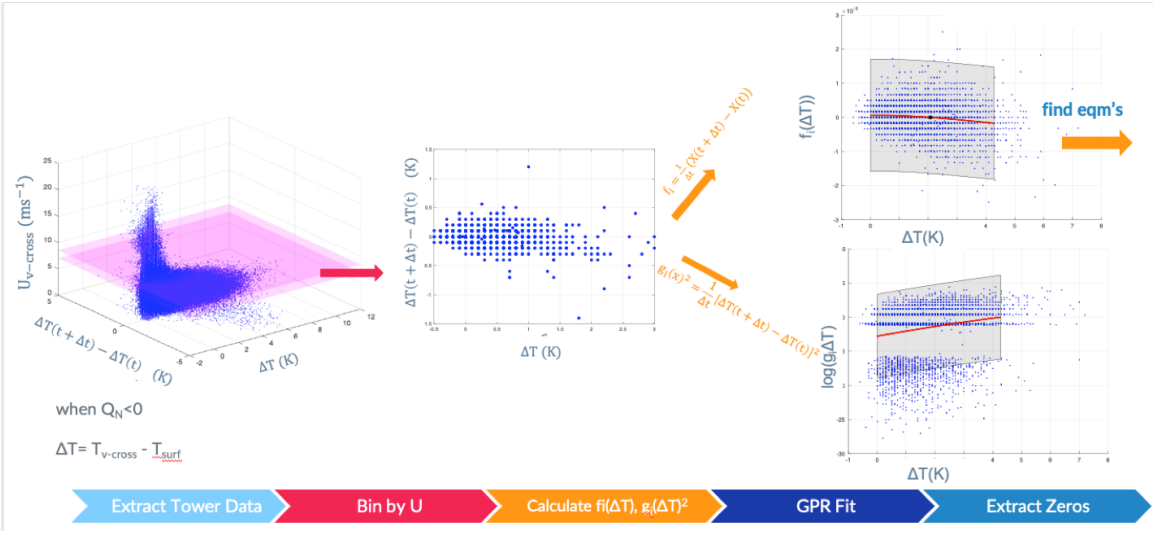



Figure 2.3: A visual schematic of the method used to estimate empirical stochastic models, discussed in Section 2.2.3.

Following the work of van de Wiel et al. [2017] the temperature inversion, $\Delta T$, is defined based on the temperature difference between the velocity crossing point, $z_h$ and the surface. Time series data of SBL $\Delta T$ and the velocity at an estimated $z_h$ are taken. Observed values of $\Delta T$ that are less than $-z_h\Gamma_d$ where $\Gamma_d$ is the dry adiabatic lapse rate are removed from the analysis as statically unstable points are not of interest. Similarly, values of $x$ modelled in the nondimensionalized stochastic vdW17 that are less than zero are removed from the analysis as the represent are statically unstable conditions. Data are conditioned by wind speed at the velocity crossing point, $U$. Values of $\Delta T$ are then binned linearly by wind speed, $U$, into $N_u$ bins across the range of observed $U$ values, using MATLAB's histogram function. The number of wind speed bins is chosen to be large enough that wind speeds do not largely vary across a bin, while still being small enough for each bin to be well

populated. The use of quantile binning was tested but resulted in very wide bins for high wind speeds, violating the assumption that wind speeds are approximately constant over a wind speed bin. Instead, a hybrid binning approach is used. Bins with greater than 1000 points are split in two to provide finer wind resolution when data is sufficient. Bin numbers of $N_u = 30$ to $N_u = 50$ were found to be most appropriate for the data set sizes considered in this study. A wide range of wind speed bin sizes were tested and results did not qualitatively differ.

Within each wind speed bin, the drift and diffusion coefficient are estimated from the time tendencies of data, where the drift is estimated as

$$f_u(t_i) = \frac{1}{\Delta t}\left[\Delta T_u(t_i + \Delta t) - \Delta T_u(t_i)\right] \tag{2.25}$$

and the diffusion is estimated as

$$g_u(t_i)^2 = \frac{1}{\Delta t^2}\left[\Delta T_u(t + \Delta t) - \Delta T_u(t)\right]^2 \tag{2.26}$$

at each measured inversion values, $\Delta T(t_i)$. Not that subscripts $u$ have been added to Equations 2.25 and 2.26 to highlight that equations are conditioned on wind speed. To obtain continuous drift and diffusion functions, two separate GP models are defined, one for the drift coefficient $\hat{f}_u \sim \mathcal{GP}\left(m_f(\Delta T), K_f(\Delta T, \Delta T')\right)$, and one for the diffusion coefficient $\hat{g}_u \sim \mathcal{GP}\left(m_g(\Delta T), K_g(\Delta T, \Delta T')\right)$. The resulting empirically derived single column SBL model can be expressed as

$$\frac{\Delta T_u}{\Delta t} = \hat{f}_u(\Delta T_u) + \hat{g}_u(\Delta T_u)\dot{w} \tag{2.27}$$

for each binned wind speed value, giving a set of $N_u$ empirical SDE's, each of which are estimated separately.

MATLAB's *fitgpr* algorithm is used to fit the GPR models. Hyperparamters for the squared exponential kernel, $\boldsymbol{\theta} = [\sigma_f, l, \sigma_n]$ and basis function parameters $\boldsymbol{\beta}$ are optimized by maximizing the log likelihood, $\log P(f_i|\Delta T_i, \boldsymbol{\theta}, \boldsymbol{\beta})$ and similarly for $g_i$. For the first $U$ bin considered, the initial guesses for signal variance $\sigma_f$ and noise $\sigma_n$ are the standard deviation of drift estimates, $f_u(t)$. The length scale is first approximated as the standard deviation of input data points, $\Delta T(t)$. For subsequent $U$ bins, the optimized hyperparameters from the previous $U$ bin fit are used to initialize

the optimization. All data points are used to fit the GPR. Trained models are cross-validated to asses parameter optimization by holding out 20% of the initial data-set, selected randomly. Initial tests showed GPR fits extrapolate poorly outside of the observed values $\Delta T$. To avoid poorly-constrained parts of the GPR fits, extracted drift and diffusion functions are restricted to the range of values within the 2.5 and 97.5 percentiles of observed $\Delta T$ values in each wind speed bin.

The dynamic equilibrium temperature inversion value(s), $\Delta T_{eqm}$ at each of the $N_u$ bins are computed as the zeroes of the extracted drift coefficient $\hat{f}_u$. The dynamic stability of each equilibrium point is assessed using the sign of the slope of the drift at the intercept point. Uncertainties in the position of equilibria are assessed by calculating the range of equilibrium values obtained for 100 realizations of the GP, computed using Equation 2.24. The fraction of realizations in which a zero is obtained is calculated during the uncertainty estimation as in some cases no zero occurs within the [2.5 97.5]% range of $\Delta T$ observations obtained. In cases where multiple zeroes occur in the mean GP, the zeroes obtained from the generated realizations are associated with the closest zero of the mean GP. Repeating this analysis for each $U$ bin gives an empirical bifurcation diagram of equilibrium values of temperature inversion, $\Delta T_{eqm}$ as a function of the wind speed at the reference height $U$ analogous to the theoretical SBL single column model bifurcation diagrams for the vdW17 model.

A variety of explicit basis functions are considered for the estimated drift coefficient, $\hat{f}_u$. A first approach given the lack of certainty of the functional form of SBL dynamics is to not assume any basis functions. However, since equilibrium values for each realization drift coefficient are desired, the use of explicit basis functions may be required to regularize the data and ensure a sufficiently smooth function (without multiple spurious zero crossings) is obtained. A linear basis function is considered as the simplest regularization function, smoothing the fit while still allowing data variance to impose a more complex solution. However, this basis function is inconsistent with potential multiple equilibria. For this reason, a third-order polynomial is also considered. All three basis functions will be considered in our analysis. For the cases of linear and cubic basis functions, these bases are expressed in terms of Legendre polynomials, $[1, x]$ and $[1, x, \frac{1}{2}(3x^2-1), \frac{1}{2}(5x^3-3x)]$ respectively, to ensure normality of basis functions.

Equation 2.18 yields $\hat{g}^2$ which is a non-negative quantity. To ensure non-negativity, rather than modelling $\hat{g}_u^2$, a model for $\hat{s}_u$, where $\hat{g}_u^2 = \exp[\hat{s}_u]$ is used to fit the GPR. A constant basis function for $\hat{s}_u$ is used to allow the mean diffusion coefficient to take values other than one.

# Chapter 3

# Idealized Model Tests

## 3.1 The Idealized Model

To test the sensitivity of the proposed method for extracting effective low dimensional equilibrium structures, described in Section 2.2.3, an analysis of simulations from the nondimensionalized form of the vdW17 model (Equation 2.13) is conducted in this Chapter.

In the original formulation of the vdW17 model, wind speed is held constant. To account for the fact that wind speed is not constant in observations, variability in wind speed is added to the model. Fluctuations in along wind, $\hat{u}$ and cross wind components, $\hat{v}$ are represented as independent Ornstein–Uhlenbeck processes with a common e-folding autocorrelation timescale, $\hat{\tau_U}$ and normalized to unit variance. The autocorrelation timescale is chosen to be larger than the dynamic timescales of the vdW17 model so that winds are slowly varying and therefore equilibria are meaningful in an approximate sense. The normalized along wind component varies about a mean, $\bar{U}$ and the mean cross wind component is zero. The resulting equations are:

$$\begin{aligned} \frac{d\hat{u}}{ds} &= \frac{-\hat{u}}{\hat{\tau_U}} + \sqrt{\frac{2}{\hat{\tau_U}}}\dot{w} \\ \frac{d\hat{v}}{ds} &= \frac{-\hat{v}}{\hat{\tau_U}} + \sqrt{\frac{2}{\hat{\tau_U}}}\dot{w} \\ \hat{U} &= \sigma_U\sqrt{(\bar{U}+\hat{u})^2 + v^2} \end{aligned} \tag{3.1}$$

where $\sigma_U$ scales wind speed variance and $s$ is non-dimensionalized time. Wind speed

modelled in this way is a reasonable approximation to observations of geostrophic wind variability over land as discussed in Monahan [2014]. While true SBL winds are affected by regime transitions, such effects are ignored in this highly simplified model as wind is only considered at the velocity crossing point, where local boundary layer tendencies should be weakest relative to large scale processes. The sensitivity of results to the ratio of the timescale of wind variability, $\hat{\tau_U}$ to the dynamic timescale of the model is considered in Section 3.3.5.

Parameters used for this model are summarized in Table 3.1. Noise-free equilibrium structures for Equation 2.8 are obtained using a pseudo-arc length continuation algorithm, described in Dijkstra et al. [2014]. A kernel density estimate of the probability density function of model output for a typical set of parameter values with a Businger-Dyer stability function is shown in Figure 3.1 along with the deterministic equilibrium structure. The sigmoidal structure of the vdW17 model is reproduced, showing a gradual transition from the low $\hat{U}$, high $x$, vSBL-like branch to the high $\hat{U}$, low $x$, wSBL branch. This gradual transition region is what is expected for well-coupled surface conditions, like at Cabauw. For this additive noise model, the mean of the stationary pdf conditioned on $\hat{U}$ follows the deterministic solution.

| **Parameter** | **Description** | **Value** |
|---|---|---|
| $\hat{Q}$ | Isothermal net radiation | $3 \times 10^{-6}$ to $3 \times 10^{-5}$ |
| $\hat{\lambda}$ | Surface exchange coefficient | $8 \times 10^{-5}$ to $4 \times 10^{-4}$ |
| $c_d$ | Neutral drag coefficient | $1.3 \times 10^{-3}$ |
| $\sigma$ | Additive noise intensity | $6 \times 10^{-5}$ to $1.5 \times 10^{-3}$ |
| $\bar{U}$ | Mean wind speed | 1 |
| $\sigma_U$ | Wind speed variance | 0.7 |
| $\hat{\tau_U}$ | Wind fluctuation time scale | $1 \times 10^3$ to $3 \times 10^6$ |
| $N$ | Number of data points subsampled | $10^4$ to $10^6$ |
| $\Delta t$ | Timestep | 30 |
| $n_t$ | Sampling rate | 10 |

Table 3.1: Nondimensional Model Parameters

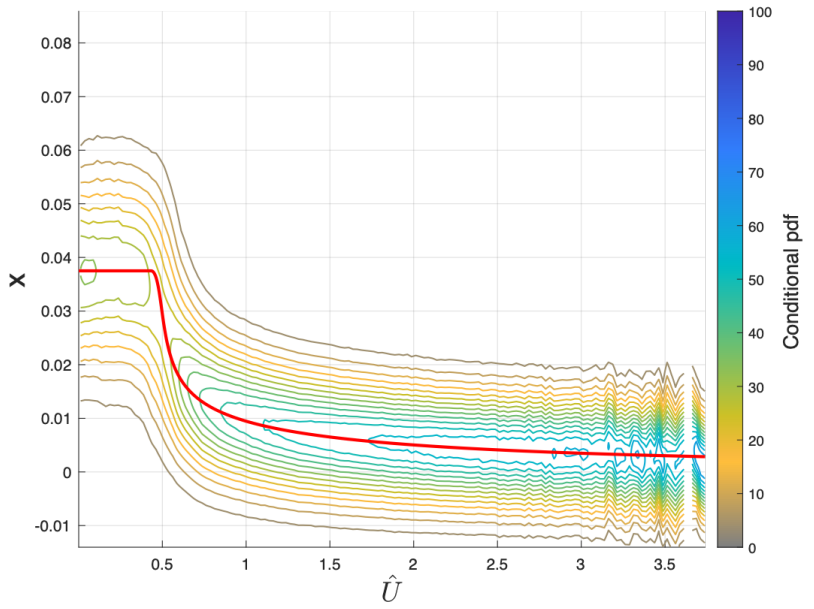



Figure 3.1: Contours show kernel density estimates of the probability density function of nondimensionalized model output with additive noise (Equation 2.13) conditioned on wind speed, $\hat{U}$, using parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$, $\hat{\lambda} = 4 \times 10^{-4}$, $\sigma = 3 \times 10^{-4}$, $\hat{\tau_U} = 3 \times 10^6$ and $N = 10^6$. The red curve shows the deterministic equilibrium structure for fixed values of $\hat{U}$.

## 3.2 Effective Dynamical Timescales

The effective dynamical timescale of the vdW17 model, $\tau_{x(\hat{U})}$, can be obtained from the coefficient of linearized dynamics about the equilibrium solution for fixed $\hat{U}$. The computed linearized timescale of the model as a function of wind speed is shown in Figure 3.2. The upper branch of the equilibrium solution, at low $\hat{U}$, high $x$ corresponds to larger $\tau_{x(\hat{U})}$ than the lower branch of the equilibrium solution, at high $\hat{U}$, low $x$. It is therefore expected that at high $\hat{U}$, the system will adjust to perturbations away from equilibrium relatively quickly. A spike in the dynamical timescale occurs where the system transitions between branches, around $\hat{U} \approx 0.5$. The long linear adjustment timescale of $x$ over a small change in $\hat{U}$ is related to a critical slowing indicating that the system is approaching a bifurcation [Scheffer et al., 2009]. When model parameters are adjusted to create an bifurcation point, the magnitude of the spike becomes infinitely large at the bifurcation point.

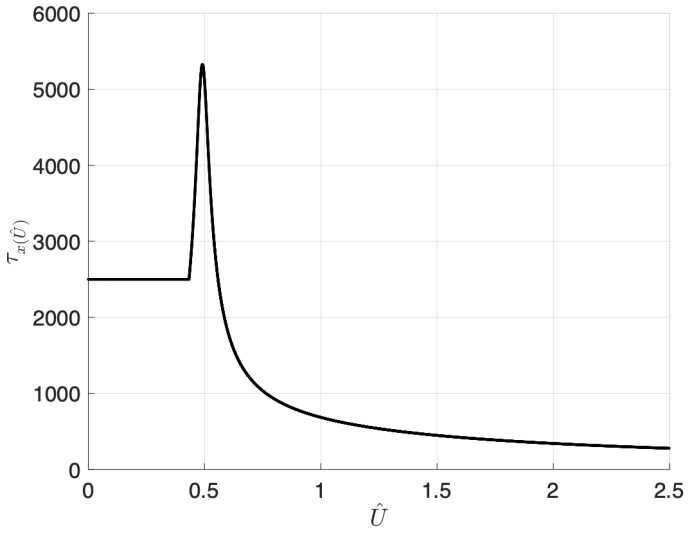



Figure 3.2: The linearized dynamic adjustment timescale, $\tau_{x(\hat{U})}$ of nondimensionalized model output with no noise (Equation 2.8), using parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$ and $\hat{\lambda} = 4 \times 10^{-4}$, corresponding to the equilibrium curve in Figure 3.1

## 3.3 Sensitivity Analysis of Reconstructed Dynamics

In this section, the proposed methodology for reconstructing dynamics from time series data is assessed using simulations of the nondimensionalized vdW17 model. Using model output with a known structure allows the performance of the method to be assessed. The sensitivity of the estimation procedure to data processing and violations of assumptions are also considered.

### 3.3.1 Cabauw-like Parameter Set

First, simulations from nondimensionalized vdW17 model with parameter values that produce a scatterplot of $x$ versus $\hat{U}$ similar to observations at Cabauw are considered. Model simulations obtained using Equations 2.13, 2.5, 3.1 and parameters in Table 3.1 with $\hat{\lambda} = 4 \times 10^{-4}$, $N = 10^5$, $\sigma = 3 \times 10^{-4}$ and $\hat{\tau_U} = 3 \times 10^6$. Model output is generated with a timestep of 30 nondimenional time units, and subsampled every 10 points giving a sample time discretization of 300. A bin size of $N_u = 50$ is used, as discussed in Section 2.2.3. Drift and diffusion coefficients are calculated using

Equations 2.17 and 2.18 respectively. At each binned $\hat{U}$ value, GPR is used to fit the drift and diffusion coefficients as functions of inversion strength. Equilibrium values for each binned $\hat{U}$ value are then computed along with their uncertainty and dynamic stability, following the methods described in Section 2.2.3. The estimated equilibrium distribution is plotted, as shown in Figure 3.3. The left, middle, and right panels show equilibria obtained using a GPR fit for the drift coefficient with no basis, a linear basis, and a cubic Legendre polynomial basis, respectively.

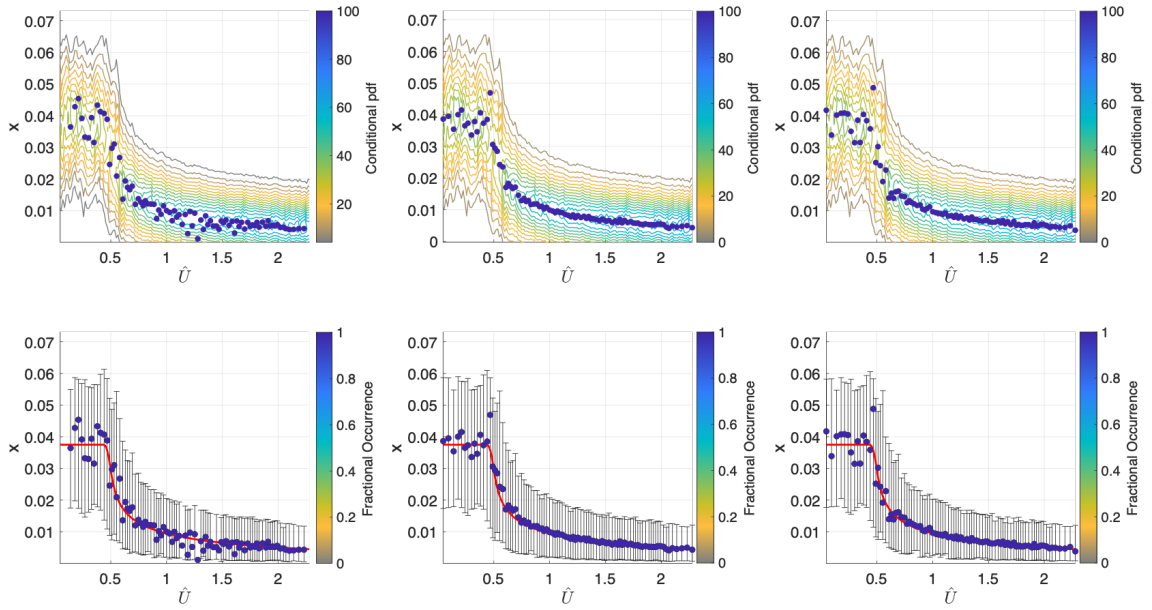



Figure 3.3: Points: empirically-estimated equilibria; and contours: kernel density estimate of a realization of the nondimensionalized model with additive noise (Equation 2.13), using parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$, $\hat{\lambda} = 4 \times 10^{-4}$, $N = 10^5$, $\sigma = 3 \times 10^{-4}$ and $\hat{\tau_U} = 3 \times 10^6$. The color of points indicates the fraction of times different realizations of the GPR find an equilibrium point over the range of data. The left, middle, and right panels respectively show equilibria obtained using a GPR fit for the drift coefficient with no basis, a linear Legendre polynomial basis, and a cubic Legendre polynomial basis. Bottom plots show estimated equilibrium values along with their 2.5 and 97.5 percentile uncertainty ranges. The red curve shows the deterministic equilibrium curve of the model.

Estimated equilibrium points from model simulations clearly follow the deterministic equilibrium structure for all three basis types. The no basis solution shows noticeably increased scatter, particularly at large $\hat{U}$ values. The linear and cubic polynomial solutions reconstruct the equilibrium curve with similar accuracy and precision. Equilibria at all values of $\hat{U}$ are correctly found to be dynamically stable. Deterministic

equilibria all lie well within the range of uncertainty computed for estimated equilibria. The range of uncertainty does not differ substantially between basis choices.

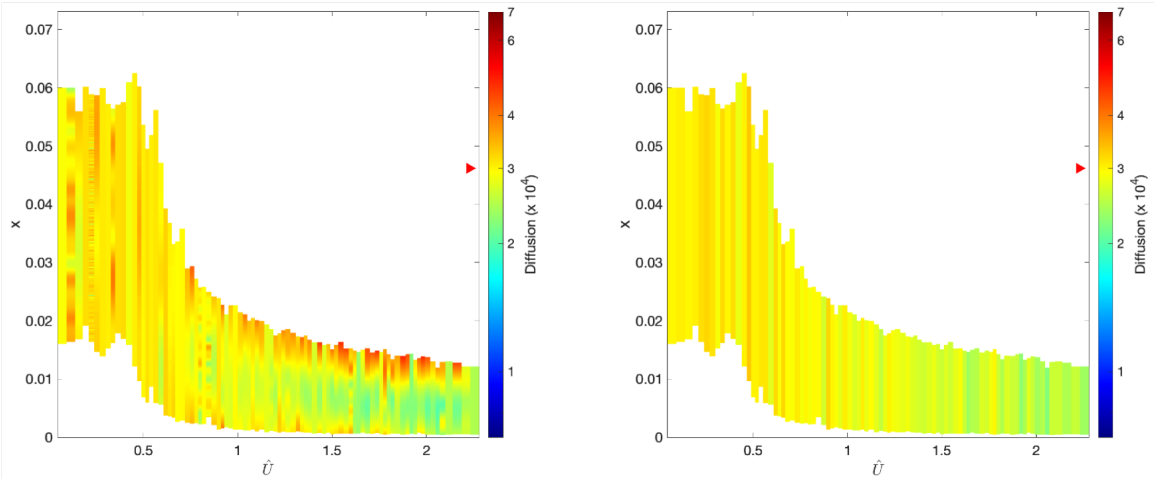



Figure 3.4: Estimated diffusion values obtained for nondimensionalized model output with additive noise (Equation 2.13), using parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$, $\hat{\lambda} = 4 \times 10^{-4}$, $N = 10^5$, $\sigma = 3 \times 10^{-4}$ and $\hat{\tau_U} = 3 \times 10^6$. The plot on the left hand size shows fits to all data within each $\hat{U}$ bin. The plot on the right shows model fits to data over the [2.5 97.5] percentile range of input data in each $\hat{U}$ bin. Note that the color scale is logarithmic. The red arrow beside the color bar indicates the true value of the diffusion.

The estimated diffusion coefficient for this model realization is shown in Figure 3.4. The plot on the right hand side shows GPR fits limited to the 2.5 to 97.5 percentile range of model data in each $\hat{U}$ bin whereas the plot on the left shows GPR fits over all model data in each bin. Note that diffusion values are fit separately over each $\hat{U}$ bin so there is some uncorrelated sampling variability between bins. Because model noise is additive in this case, the diffusion coefficient should be a constant value, corresponding to the additive noise strength, $3 \times 10^{-4}$. The left panel of Figure 3.4 shows diffusion values ranging from $2.3 \times 10^{-4}$ to $7 \times 10^{-4}$ with the largest values along the edges for large $\hat{U}$. This quadratic bias is in keeping with the findings of Sura and Barsugli [2002], who show that numerically computing diffusion coefficients through finite differencing creates errors to leading order $\Delta t^2$. The GPR fit worsens at edges, where data are sparse. Limiting the GPR fit to the 2.5 to 97.5 percentile range of data in each $\hat{U}$ bin substantially reduces this bias; estimated values are generally much closer to the true value (Figure 3.4, right panel). However the potential

influence of this bias must be considered when interpreting results.

It is clear from Figures 3.3 and 3.4 that the estimation procedure is able to reasonably reconstruct the bifurcation structure and the diffusion coefficient of Equation 2.13 with the Cabauw-like parameters from Table 3.1. The sensitivity of reconstructed dynamics to changes in parameter values, data processing and the inclusion of multiplicative red noise is explored in the subsequent sections of this chapter.

### 3.3.2 Parameter Values Resulting in Unstable Equilibria

The ability of this approach to resolve idealized model output showing multiple equilibria over a range of parameters is examined by decreasing the nondimensionalized surface exchange coefficient to $\hat{\lambda} = 8 \times 10^{-5}$. At this lower value of $\hat{\lambda}$ an unstable branch exists between approximately $\hat{U} = 0.5$ and $\hat{U} = 1$ (Figure 3.5). The resulting equilibrium structure estimated for $N = 10^5$ points for each of the different GPR basis sets considered is shown in Figure 3.5.

The no-basis reconstruction on the left hand side of Figure 3.5 resolves only one point in the transitional region, and incorrectly identifies it as stable. The linear basis locates no equilibrium points along the unstable branch, which is not surprising as the existence of multiple equilibria is inconsistent with a linear basis. The cubic polynomial correctly locates a number of unstable equilibria along or near the unstable branch. Although the resolution is poor as compared to the Cabauw-like parameter set (Figure 3.3), the overall equilibrium structure is captured.

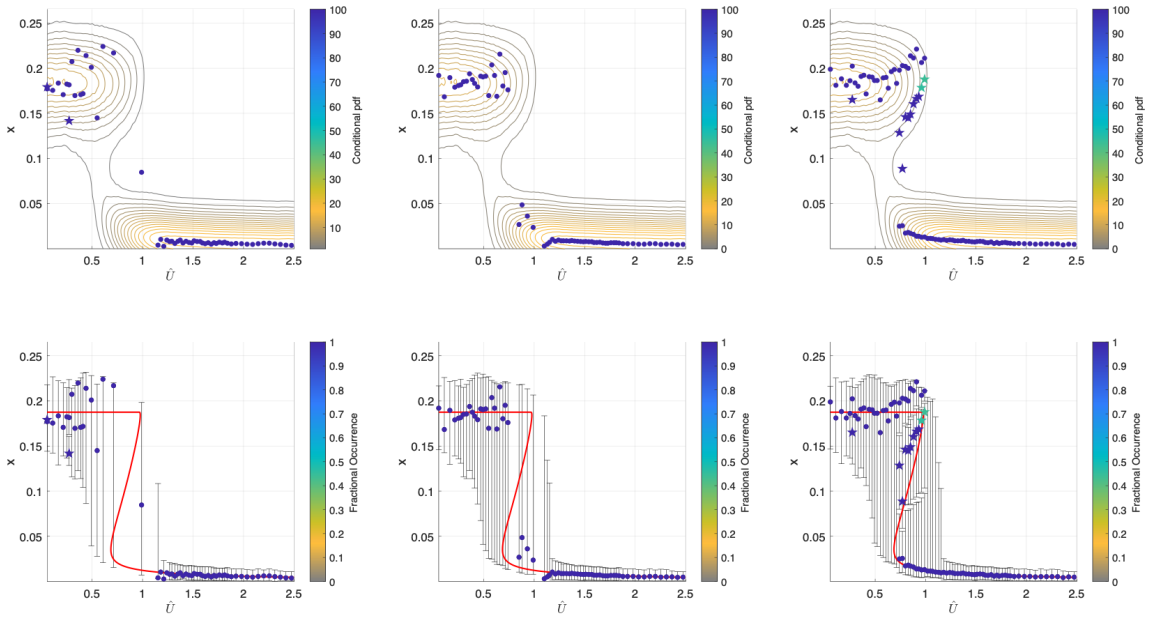



Figure 3.5: Points: empirically-estimated equilibria; and contours: kernel density estimate of a realization of the nondimensionalized model with additive noise (Equation 2.13), using parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$, $\hat{\lambda} = 8 \times 10^{-5}$, $N = 10^5$, $\sigma = 3 \times 10^{-4}$ and $\hat{\tau_U} = 3 \times 10^6$. Dynamically stable equilibria are shown as circles and dynamically unstable equilibria are shown as stars. The color of points indicates the fraction of times different realizations of the GPR find an equilibrium point over the range of data. The left, middle, and right panels respectively show equilibria obtained using a GPR fit for the drift coefficient with no basis, a linear Legendre polynomial basis, and a cubic Legendre polynomial basis. Bottom plots show estimated equilibrium values along with their 2.5 and 97.5 percentile uncertainty ranges. The red curve shows the deterministic equilibrium curve of the model.

### 3.3.3 Sensitivity to Number of Data Points

To test the sensitivity of the reconstruction approach to the number of data points, a longer model with $N = 10^6$ points is used followed by a smaller subset of $N = 10^4$ points. This range of $N$ values is chosen as it aligns with the range data set sizes available at meteorological towers. Additionally, because GPR requires the storage and inversion of an $N \times N$ matrix, using this methodology on data sets larger than $N = 10^6$ points becomes computationally impractical. The equilibrium structure estimated for $N = 10^6$ points, for each of the different basis sets considered is shown in Figure 3.6. The no-basis equilibria obtained on the left hand side of Figure 3.6 follows the deterministic solution remarkably well, with many unstable equilibrium points identified along the unstable branch. The linear basis remains poor at locating

equilibrium points along the unstable branch. The cubic basis analysis with $N = 10^6$ improves significantly when compared to the $N = 10^5$ point case (Figure 3.5). The cubic basis resolves the overall structure albeit with fewer points on the unstable branch than the no basis case with scatter overall. The cubic basis results also show lower uncertainty than the no-basis results.

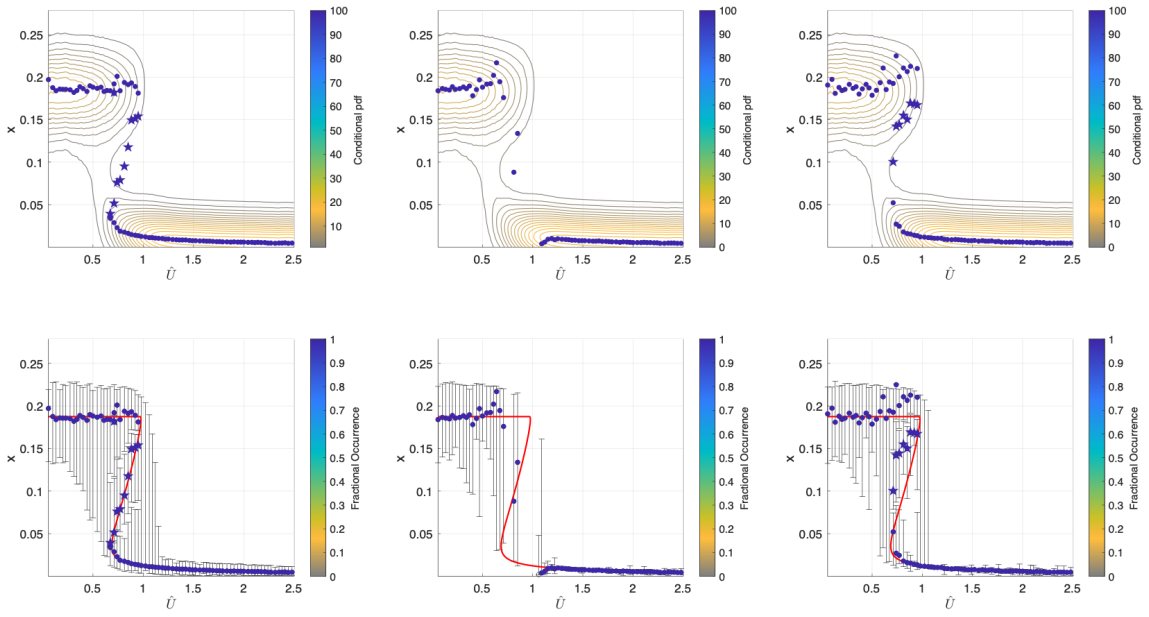



Figure 3.6: As in Figure 3.5 with $N = 10^6$.

The resulting equilibrium structure for a subset of $N = 10^4$ points is shown in Figure 3.7. The no basis solution locates few equilibrium points overall. Comparing the $N = 10^4$ solution (Figure 3.7) to the $N = 10^5$ and $N = 10^6$ solutions (Figures 3.5 and 3.6) highlights the importance of having a sufficiently large number of data points to resolve the equilibrium structure. By regularizing the drift function, imposing more structure on the GPR fit through the use of an explicit basis, the approach more effectively extracts equilibrium values, as shown by the linear and cubic basis solutions. Again the linear basis reconstruction is still unable to resolve any points along the unstable branch. The cubic polynomial solution however remains able to reconstruct the overall equilibrium structure over the unstable branch, although with increased scatter relative to the results obtained for larger values of $N$. Because it is found to produce the most robust results, the cubic basis will be primarily be shown in subsequent model tests. Smaller values of $N$ can also be accommodated using

fewer wind speed bins ($N_U$), however this results in more $\hat{U}$ variability within each wind-speed bin making the quasi-equilibrium approximation more problematic.

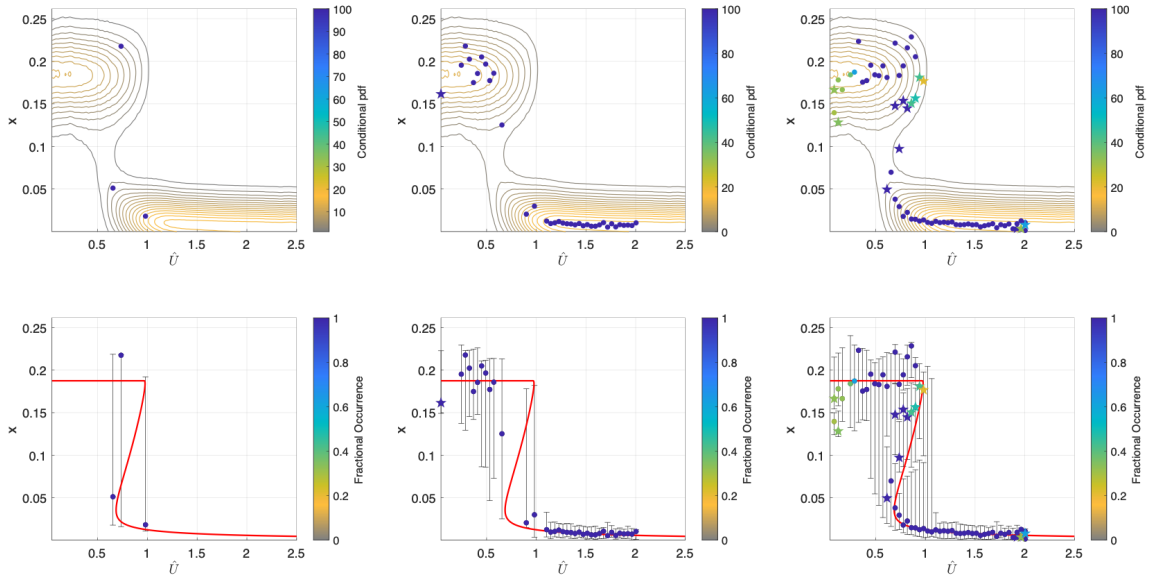



Figure 3.7: As in Figure 3.5 with $N = 10^4$.

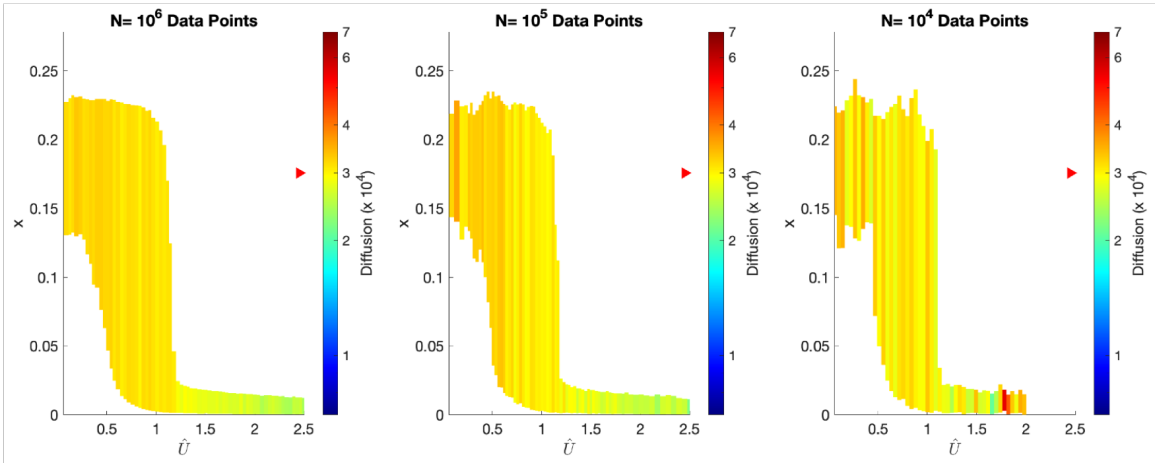



Figure 3.8: Diffusion values as in Figure 3.4 with $\hat{\lambda} = 8 \times 10^{-5}$, for different record lengths $N$. Left: $N = 10^6$, points. Center: $N = 10^5$ points. Right: $N = 10^4$ subset of the points.

The diffusion coefficients obtained for the model with $\hat{\lambda} = 8 \times 10^{-5}$ using $N = 10^6$, $N = 10^5$ and $N = 10^4$ points are shown in Figure 3.8. Diffusion values in the wSBL region are biased low for the $N = 10^6$ and $N = 10^5$ cases indicating a bias in the

estimation procedure. Diffusion estimates for the $N = 10^4$ case are cut off at lower wind speeds as high wind speeds are poorly sampled in this small data set. Diffusion values are also less consistent between wind speed bins indicating larger data sets provide more robust estimates of the diffusion.

### 3.3.4 Time Averaging

The observational data sets considered in Chapter 4 represent averages over time intervals varying from 1 minute to 30 minutes. To asses the impact of time averaging on the reconstruction of dynamics, model realizations are averaged across 10 point intervals. The same model run considered in section 3.3.2 is used with parameters from Table 3.1 with $\hat{Q} = 1.5 \times 10^{-5}$, $\hat{\lambda} = 8 \times 10^{-5}$, $N = 10^6$, $\sigma = 3 \times 10^{-4}$ and $\hat{\tau_U} = 3 \times 10^6$. After averaging, the dataset has $N = 10^5$ points. The estimated equilibrium structure obtained from the 10-point averaged model realization, assuming a cubic basis function, is shown in Figure 3.9.

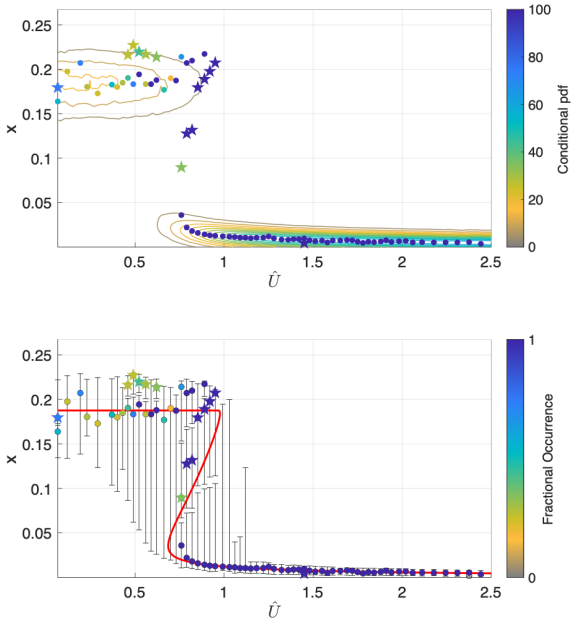



Figure 3.9: As in Figure 3.6 with points averaged over a 10 point interval using GPR fits with a cubic basis.

The cubic basis continues to find equilibrium points along the stable branches, however fewer unstable points along the unstable branch are resolved when compared to the non-averaged data in Figure 3.6. Multiple unstable equilibria on the upper, vSBL equilibrium branch are incorrectly identified. These results suggest that a consequence of time averaging can be innacurate estimates of unstable equilibria.

Because effective low-dimensional dynamics are being considered some time averaging of observations is important, as will be shown in Section 4.4. By construction, the vdW17 model output considered in this Chapter for comparison represents time averaged quantities; for example, it incorporates a (parameterized) eddy heat flux, which is a Reynolds-averaged quantity. This section addresses the effects of excessive time averaging.

The diffusion coefficient obtained from 10 point averaged data is shown in Figure 3.10. The overall magnitude of the diffusion coefficient decreases by a factor of 10, due to the time averaging acting as an effective low pass filter on high frequency noise. Reconstructed diffusion values in the low $\hat{U}$, high $x$ region, corresponding to the vSBL are higher than those in the wSBL. The dynamic restoring timescale, (Figure 3.2) is higher for the vSBL than in the wSBL. The increased diffusion values in the vSBL may therefore be a result of vSBL dynamics fluctuating at a lower frequency than the wSBL and therefore being less affected by the frequency filtering effects of time averaging. The large positive bias on the edges of the unstable branch is likely due to the sparsity of data in this region. Time averaging results in different estimates in the vSBL than in the wSBL, which could be incorrectly interpreted as evidence of multiplicative noise.

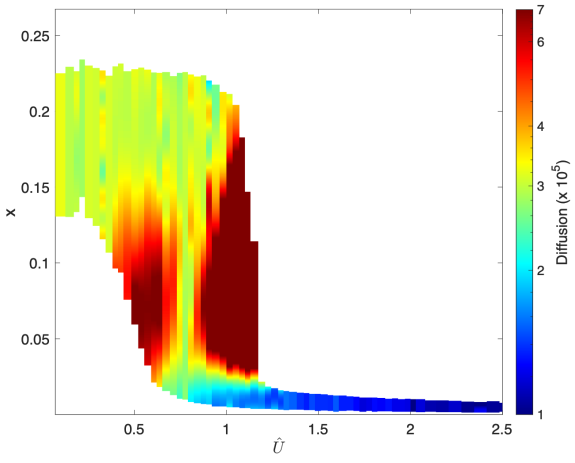



Figure 3.10: Diffusion values as in Figure 3.8 with $N = 10^6$ points time averaged over 10 point intervals.

### 3.3.5 Timescale of Wind Variability

Implicit in the equilibrium estimation procedure is the assumption that wind speeds vary over sufficiently long timescales relative to temperature inversions and that wind

speeds can be taken as quasi-stationary within each wind speed bin. To test the method's sensitivity to this simplifying assumption, calculations are repeated across simulations generated with a range of values of wind fluctuation memory, $\hat{\tau_U}$. Wind variability timescales are compared to a representative value of the system's linearized dynamic time scale, $\tau_{x(\hat{U})}$ obtained from model dynamics linearized around equilibrium points, taken to be 1000 (cf. Figure 3.2). For this analysis, the Cabauw-like set of parameter values from Section 3.3.1 are used. Plots of the cubic polynomial basis reconstructions from model realizations with $\hat{\tau_U} = 10000$, $\hat{\tau_U} = 5000$, and $\hat{\tau_U} = 1000$ are shown in Figure 3.11.

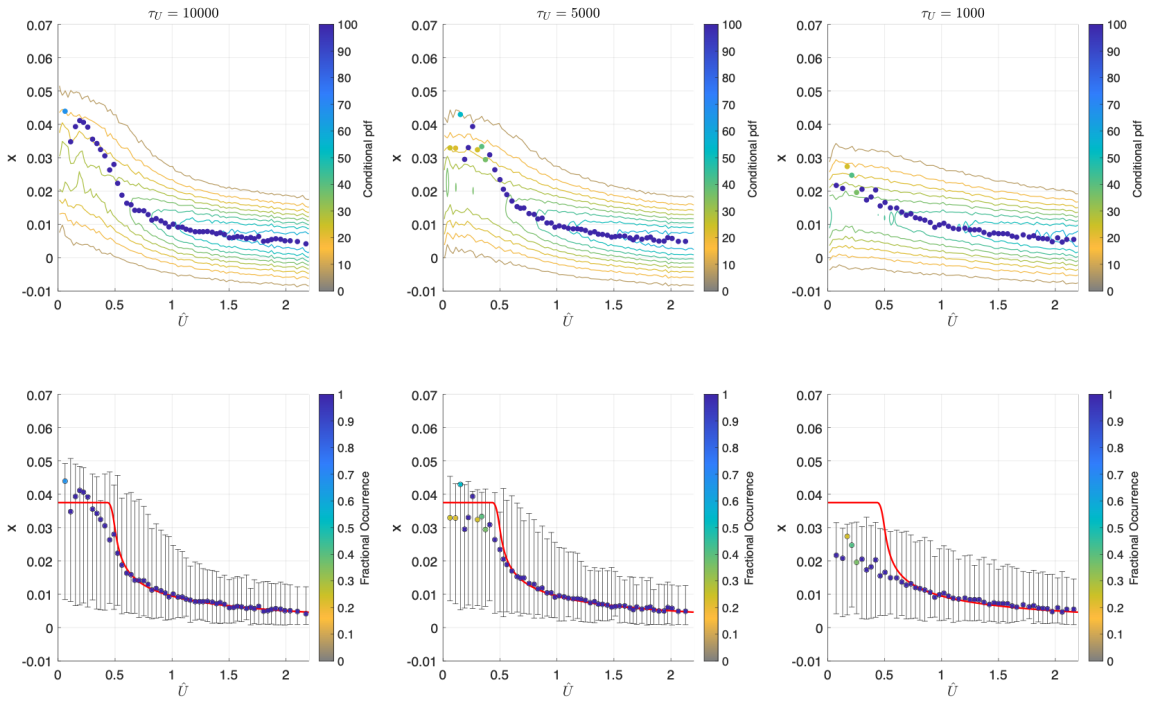



Figure 3.11: As in Figure **??** with variable $\hat{\tau_U}$ using GPR fits with a cubic basis . Left: $\hat{\tau_U} = 1 \times 10^4$. Center: $\hat{\tau_U} = 5 \times 10^4$. Right: $\hat{\tau_U} = 1 \times 10^3$.

It is clear that that as $\hat{\tau_U}$ approaches the typical scale of $\tau_{x(\hat{U})}$, the method's ability accurately reconstruct the equilibrium curve is reduced. Estimated values deviate from the deterministic solution at low $\hat{U}$, large $x$ and in the transitional region, where the linearized dynamical time scale is the largest. Approximating model wind speed as quasi stationary results in biased equilibrium reconstructions when dynamics vary at a similar timescale as wind. It is also apparent that the reconstruction is reasonably accurate even when $\hat{\tau_U}$ is not much larger than the typical value of $\tau_{x(\hat{U})}$. When

extracting equilibrium positions, assuming that the wind is approximately constant within a wind speed bin can result in weakened estimates of the strength of temperature inversions if wind speed variations are on sufficiently short timescales. This fact suggests that in real data sets, variability of the true wind speed may cause an effective dampening of the estimated equilibrium structure if the timescale of wind variability is not at least a few times longer than the dynamic timescale of temperature inversions.

Estimated diffusion coefficients for model realizations with $\hat{\tau_U} = 10000$, $\hat{\tau_U} = 5000$, and $\hat{\tau_U} = 1000$ are shown in Figure 3.12. These plots again show that the accuracy of this method is reduced when the wind varies on a timescale comparable to that of the linearized dynamics. In particular, a more pronounced quadratic bias appears as $\hat{\tau_U}$ approaches $\tau_{x(\hat{U})}$.

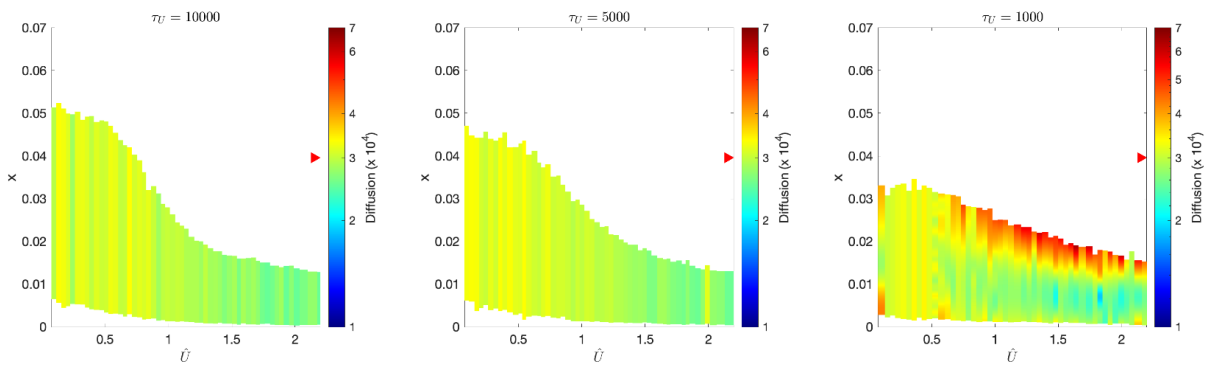



Figure 3.12: Diffusion values as in Figure 3.4 with variable $\hat{\tau_U}$ Left: $\hat{\tau_U} = 1 \times 10^4$. Center: $\hat{\tau_U} = 5 \times 10^4$. Right: $\hat{\tau_U} = 1 \times 10^3$.

### 3.3.6 Isothermal Net Radiation

Variations in cloud cover can be accounted for in the vdW17 model through changes to the isothermal net radiation, $Q_i$. Therefore to consider changes in cloud cover, the nondimensionalized model is studied with values of $\hat{Q}$ larger and smaller than those used in 3.3.1. The resulting empirical bifurcation diagrams are shown in Figure 3.13. The equilibrium estimation procedure accurately reconstructs the deterministic solution across a wide range of $\hat{Q}$ values. Low values of $\hat{Q}$, associated with cloudy conditions, result in weakened equilibrium temperature inversions, and more gradual transitions between regimes, due to enhanced downwards longwave radiation. High

values of $\hat{Q}$, associated with clear sky conditions, result in larger temperature inversions, with more abrupt transitions between regimes. The transitional wind speed is smaller under cloudy conditions than clear skies (cf. Figure 8 of Abraham and Monahan [2019b]). In broad terms, the effect on the equilibrium structure of increasing $\hat{Q}$ is similar to that of decreasing $\hat{\lambda}$.

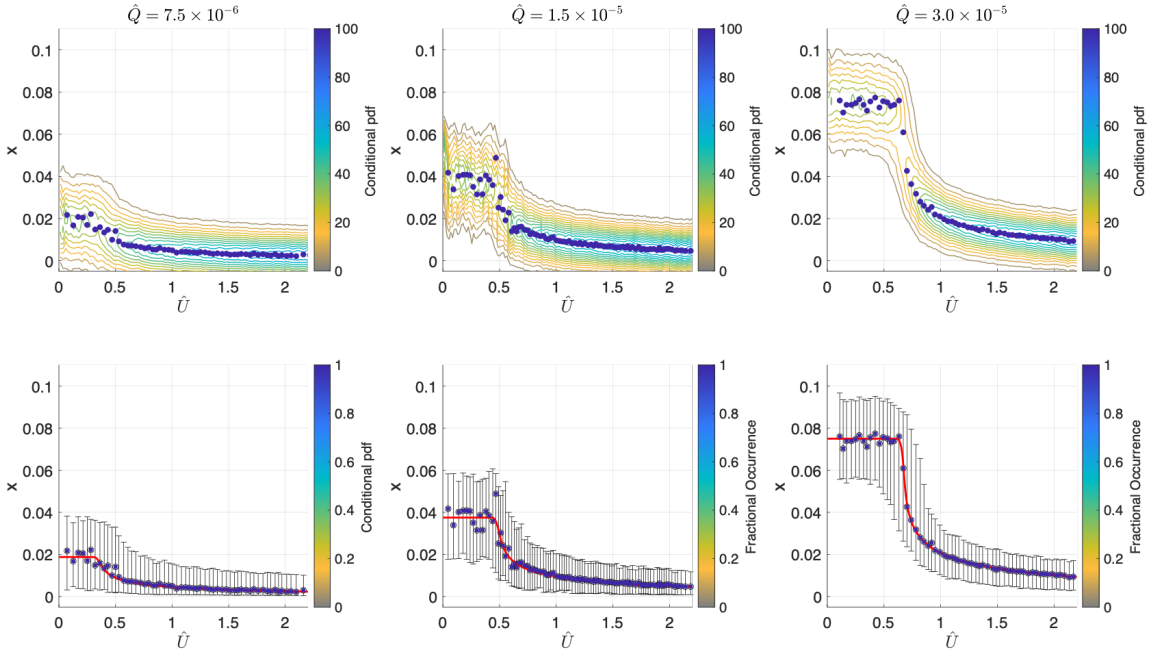



Figure 3.13: As in Figure **??** with a cubic basis and variable values of $\hat{Q}$. Left: $\hat{Q} = 7.5 \times 10^{-5}$. Center: $\hat{Q} = 1.5 \times 10^{-5}$. Right: $\hat{Q} = 3.0 \times 10^{-5}$.

## 3.4 Sensitivity to Noise Structure

### 3.4.1 Additive Noise

To examine the effects of changes in intensity of the additive noise on the system, the analysis of Section 3.3.1 is repeated with higher and lower values of $\sigma$ using the "Cabauw-like" parameter values. Figure 3.14 shows the estimated equilibrium structure for additive noise coefficients that are one fifth (left panel) and five times (right panel) the original additive noise value of $\sigma = 3 \times 10^{-4}$. The low noise case closely follows the deterministic solution with little uncertainty in the estimate of the equilibria, whereas the high noise case shows increased scatter and much larger uncertainties. It is however noteworthy that in the latter case, despite a high magnitude of additive

noise, the overall equilibrium curve is recovered to a good approximation. Equilibrium values in the high noise case also appear to be systematically over predicted, particularly in the transitional region. This could be due to a greater data density at higher $x$ values and the rejection of points where $x < 0$, as described in Section 2.2.3.

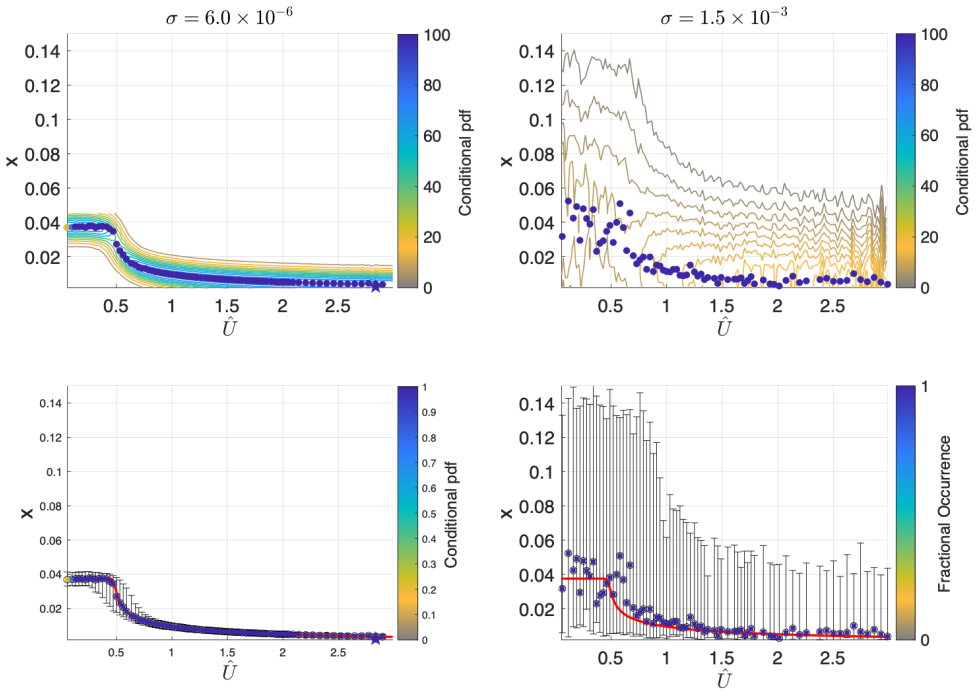



Figure 3.14: As in Figure **??** with a cubic basis and variable values of $\sigma$. Left: $\sigma = 6.0 \times 10^{-5}$. Right: $\sigma = 1.5 \times 10^{-3}$.

The estimated diffusion coefficient for the low and high additive noise model runs are shown in Figure 3.15. The input model values of $\sigma$ are both reproduced in general, in regions with high data density. Where data are sparse, a quadratic bias appears. The bias is strongest in the transitional region when there is low additive noise whereas the bias is strongest in wSBL-like region of the high noise case.

### 3.4.2 State Dependent Noise

Previous generalizations of the vdW17 models have considered only additive, white noise. The use of such noise processes assumes both that unrepresented processes represented by the noise processes are state-independent and that noise processes have timescales sufficiently small when compared to the dynamical timescales of the

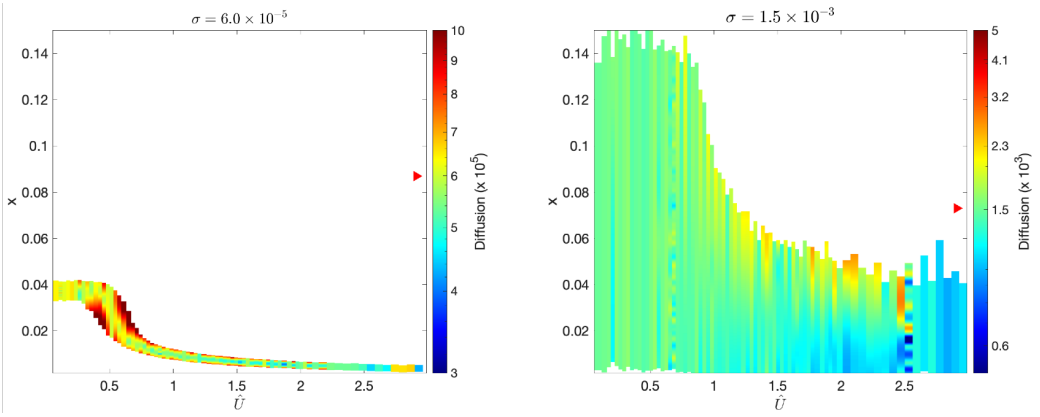



Figure 3.15: Diffusion values as in Figure 3.4 with variable $\sigma$. Left: $\sigma = 6.0 \times 10^{-5}$. Right: $\sigma = 1.5 \times 10^{-3}$.

SBL. Observations however show top-down intermittent turbulent bursts specifically in the vSBL (e.g. Abraham and Monahan [2019a], Mahrt [2014], Van de Wiel and Holtslag [2003]) , suggesting noise processes are fundamentally different in the vSBL than in the wSBL. Turbulent bursts have also been found in large-eddy simulations of the Antarctic vSBL through unstable wave growth due to shear-generated Kelvin-Helmholtz instabilities [van der Linden et al., 2020]. Therefore multiplicative noise is likely a more appropriate stochastic representation of this system. A state dependent model for the diffusion coefficient will now be considered. Parameterizing external variability as white noise is also potentially a poor approximation as intermittent turbulent bursts in the vSBL have been observed to vary at timescales of tens of minutes, which is similar to the timescale SBL regime transitions occur at [Abraham and Monahan, 2019b, Coulter and Doran, 2002, Sun et al., 2002, Deb Burman et al., 2018]. Therefore the effects of serially-dependent noise on SBL dynamics are considered, starting with a red noise process, $\zeta$, taken to be an Ornstein-Uhlenbeck process, with an autocorrelation time $\hat{\tau_z}$,

$$\frac{d\zeta}{ds} = -\frac{1}{\hat{\tau_z}}\zeta + \frac{1}{\hat{\tau_z}}\dot{w}_m. \tag{3.2}$$

where $\dot{w_m}$ is Gaussian white noise. The red noise process, $\zeta$ is Gaussian with autocovariance $\frac{1}{2\hat{\tau_z}}e^{|-s|/\hat{\tau_z}}$. The process $\zeta$ approaches white noise as $\hat{\tau_z} \to 0$. Note that $\zeta$ does not directly represent the intermittent turbulent contribution to turbulent heat fluxes. Instead a non-linear transformation of $\zeta$ is used to make the turbulent contribution "bursty."

Intermittent turbulent bursts should be small in general with occasional large values. These turbulent burst have also shown to be state-dependent, in that they are stronger in the vSBL than in the wSBL. A red noise model with such features is

$$\frac{dx}{ds} = \hat{Q} - \hat{\lambda}x - \left( c_d \hat{U} f\left(\frac{x}{\hat{U}^2}\right) + h(x)\left[(a+\zeta)^{2n} - b\right] \right) x + \sigma_w \dot{w_a} \tag{3.3}$$

where the function $h(x)$ defines the state dependence of the system. Together, these values of $a$ and $n > 1$ determine how non-Gaussian and "bursty" the noise is. For large $a$, the term $h(x)\left[(a+\zeta)^{2n} - b\right]$ becomes approximately Gaussian. For a fixed value of $a$, increasing $n$ results in a "burstier" noise process. The noise term can be taken to be mean zero with appropriate choice of $b$. Given that $\zeta$ is a zero mean Gaussian process with variance $\frac{1}{2\hat{\tau_z}}$, for $n = 1$ the noise term is mean zero for $b = a^2 + \frac{1}{2\hat{\tau_z}}$. An additional additive white noise acting on the system is scaled with intensity $\sigma_w$.

An illustrative example of the reconstructed equilibrium structure of this state dependent noise model is shown in Figure 3.16. The function defining state dependence is taken to be $h(x) = 10^{-2}x$. The value $a$ is taken to be 1.5 and $b$ is set to make the mean flux zero. The red noise process $\zeta$ has an autocorrelation time $\hat{\tau_z} = 100$ and the additive white noise intensity is lower than that used in the purely additive noise runs, $\sigma_w = 1.8 \times 10^{-4}$. These model parameters are used for illustrative purposes only and are not derived from any particular source. While the methodology used to reconstruct stochastic dynamics, described in Section 2.2.3, is only strictly valid for white noise processes, the same methods are used here for demonstrative purposes, noting that an effective state dependent drift may be diagnosed. The estimated equilibrium values are found to closely follow the deterministic solution with no noticeable bias.

The estimated diffusion coefficient is shown in Figure 3.17. Increased noise intensity is apparent in the vSBL region, as expected. The diffusion estimation methodology is able to extract an effective white-noise model with a multiplicative structure despite modelled noise being serially-dependent. A stronger quadratic bias is observed in the wSBL region.

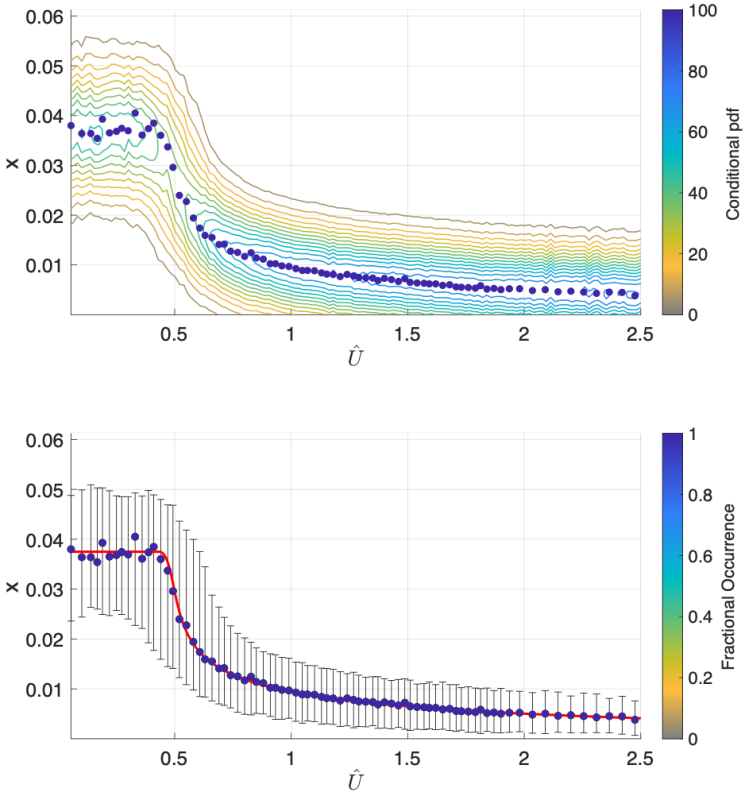



Figure 3.16: As in Figure **??** with a cubic GPR basis fit modelled with multiplicative noise using Equation 3.3 with $\sigma_w = 3 \times 10^{-4}$, $\tau_z = 100$, $h(x) = 10^{-5}\tanh(x)$, $n = 1$, $a = 1.5$ and $b = a^2 + \frac{1}{2\hat{\tau}_z}$, along with the estimated equilibria.

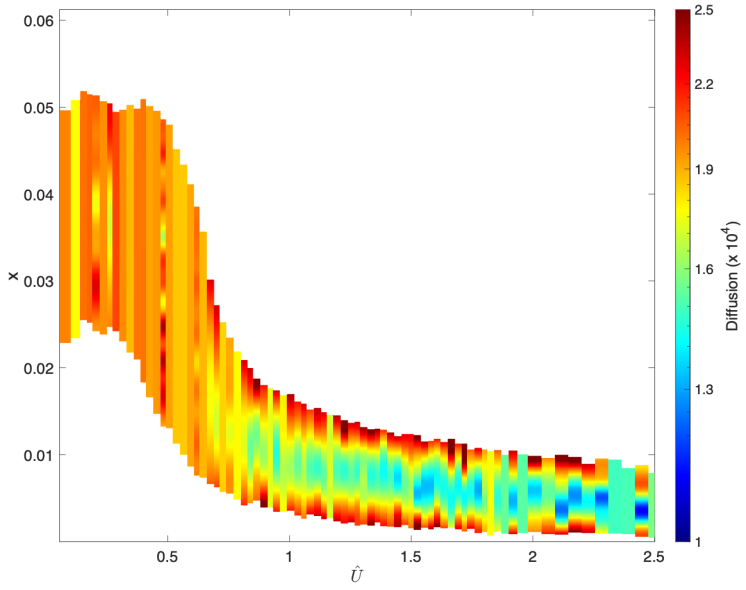



Figure 3.17: Diffusion values as in Figure 3.4 modelled with multiplicative serially dependent noise using Equation 3.3 with $\sigma_w = 3 \times 10^{-4}$, $\tau_z = 100$, $h(x) = 10^{-5}\tanh(x)$, $n = 1$, $a = 1.5$ and $b = a^2 + \frac{1}{2\hat{\tau}_z}$.

# Chapter 4

# Tower Data

## 4.1 Data

Timeseries data from six different land based meteorological towers are used for this study. Data available at each site are summarized in Table 4.1 (adapted from Abraham and Monahan [2019a]). Information regarding obtaining these data can be found in Appendix 5. The towers are located around the world in a range of locations with varying meteorological settings and surface conditions. Cabauw, Netherlands is located on flat, moist agricultural land with relatively limited surrounding topography or urban development. The Hamburg tower is located on similarly flat, moist grass land in a river valley, 8km from the urban metropolis of Hamburg, Germany. Dome C is located on flat terrain in the interior of Anatarctica, providing insight into SBL dynamics in ice-based conditions. The Karlsruhe tower is located in a small clearing of a forested valley, 10 km from the city of Kalrsruhe, Germany. Los Alamos is located on a very dry plateau downslope of the Jemez Mountains in New Mexico, U.S.A. The Boulder tower, on flat, dry farmland 25 km east of the foothills of the Rocky mountains, is also examined, although the vertical resolution of data at this tower is poor.

Most of the towers provide meteorological state variables that are Reynolds-averaged over periods of 10 to 30 minutes. Hamburg data are measured every 0.5 seconds and Reynolds-averaged over 1 minute allowing for analysis of the method's time resolution sensitivity. Hamburg and Cabauw provide the longest records of observations and therefore are used to test and calibrate the equilibrium estimation

procedure. Cabauw also provides a broad set of meteorological state variables, including radiative fluxes and low level cloud cover. Cloud cover measurements allows for analysis of the impact of cloud cover on the extracted effective low dimensional dynamics. Because this study considers SBL dynamics, only night-time data points are considered. For this study, an energetically-based definition of nocturnal data is used, defining night as those times with $Q_{net} < 0$ when radiative flux measurements are available. At stations where radiative flux measurements are not available, data one hour before sunset to one hour after sunrise are used, as in Abraham and Monahan [2019a]. van Hooijdonk et al. [2017a] analyzes the growth rate of temperature inversions around sunset at a number of land meteorological tower and concludes that development of the SBL begins two hours prior to sunset. Because this study is not interested in transitional periods a more conservative choice of one hour is used.

| Institute | References | Location | Time Period | Points | Data | avg. (min) | Measurement Heights (m) |
|---|---|---|---|---|---|---|---|
| The Boulder Atmospheric Observatory (BAO), Boulder, USA | Kaimal and Gaynor [1983], Blumen [1984] | 40.0500°N, 105.0038°W, 1584m | 2008 - 2015 | $1.7\times10^5$ | T | 10 | 10, 100, 300 |
| | | | | | U | 10 | 10, 100, 300 |
| The Royal Netherlands Meteorological Institute (KNMI), Cabauw, Netherlands | Van Ulden and Wieringa [1996], Bosveld et al. [2020] | 51.9700°N, 4.9262°E, -0.7m | 2001 - 2019 | $5.6 \times 10^5$ | T | 10 | 2, 10, 20,50, 80, 140, 200 |
| | | | | | U | 10 | 10, 20, 40, 80, 140, 200 |
| | | | | | Q | 10 | 1.5 |
| Meteorologisches Institut der Universität Hamburg (MI), | Brümmer [2012], Gryning et al. [2016] | 53.5192°N, 10.1051°E, 0.3m | 2005 - 2015 | $2.74 \times 10^6$ | T | 1 | 2, 10, 50, 110, 175, 250, 280 |
| | | | | | U | 1 | 10, 50, 110, 175, 250, 280 |
| Karlsruhe Institut fur Technologie (KIT), Karlsruhe, Germany | Barthlott et al. [2003] Kohler et al. [2018] Wenzel et al. [1997] Kalthoff and Vogel [1992] | 49.0925°N, 8.4258°E, 110.4m | 2003 - 2013 | $1.3 \times 10^5$ | T | 10 | 2, 10, 30, 60, 100, 130, 160, 200 |
| | | | | | W | 10 | 2, 20, 30, 40, 50, 60, 80, 100, 130, 160, 200 |
| Los Alamos National Laboratory (LANL), Los Alamos, USA | Bowen et al. [2000] Bruggeman [2017] | 35.8614°N, 106.3196°W, 2263m | 1995 - 2015 | $2.7 \times 10^5$ | T | 15 | 1.2, 11.5, 23, 46, 92 |
| | | | | | W | 15 | 11.5, 23, 46, 92 |
| | | | | | Q | 15 | 1.5 |
| Institut Polaire Français Paul Émile Victor (IPEV), and Programma Nazionale Ricerche in Antartide (PNRA), DomeC, Antarctica | Vignon et al. [2017b], Vignon et al. [2017a], Genthon et al. [2010], Genthon et al. [2013] | 75.1000°S, 123.3000°E, 3233m | 2011 - 2018 | $9 \times 10^4$ | T | 30 | 1.3, 2.3, 3.5, 9.0, 18.2, 25.6, 32.9, 41.3 |
| | | | | | W | 30 | 1.3, 2.3, 3.5, 9, 18.2, 25.6, 32.9, 41.3 |
| | | | | | Q | 30 | 1.5 |

Table 4.1: Summary of meteorological weather tower data used, sorted alphabetically. Detailed information about each site can be found in the cited references. This table is adapted from Abraham and Monahan [2019a]. Information on data availability can be found in Appendix 5.

## 4.2 Cabauw

### 4.2.1 Reconstruction of Stochastic Dynamics

In this section time series data from Cabauw, Netherlands are analyzed using the equilibrium estimation procedure described in Section 2.2.3. The station is located on flat, moist agricultural land, far from any large urban developments. Comprehensive reviews of this site can be found in Bosveld et al. [2020] and Van Ulden and Wieringa [1996]. Temperature and wind speed measurements from 2001 to 2019 with 0.33 Hz sampling frequency and subsequently 10-minute Reynolds-averaged are considered[Van Ulden and Wieringa, 1996]. These recorded data are subsampled to extract measurements when the surface net radiation is less than zero, resulting in a data set of $5.6 \times 10^5$ points. Temperatures at this station are measured to a precision of 0.1 K. To reduce the appearance of time invariant temperature inversion values due to precision limitations, a uniformly distributed, white in time, random noise varying from $-0.05$ K to 0.05 K is added to the data. The equilibrium estimation procedure is tested with and without this noise and results show no noticeable change. Temperature inversions are calculated between 2 m and various altitudes. Note that this definition of the inversion differs from the idealized model which uses the surface temperature. Temperature data at heights below 2 m are not available for the full 20 years of this record. Time series analysis of wind speed data at night shows that the velocity crossing point is at approximately 40 m [van de Wiel et al., 2012a]. This height is used to define the inversion in this section. Other measurement heights are considered in Section 4.2.2.

Drift and diffusion coefficients are estimated using Equations 2.17 and 2.18. Observed wind speeds are binned linearly with $N_u = 50$ bins using MATLAB's histogram function as discussed in Section 2.2.3. The drift coefficient, conditioned on wind speed, is fit to a GPR with no basis, a linear Legendre polynomial basis, and a cubic Legendre polynomial basis, as described in Section 2.2.3. Zeroes within each wind speed bin are calculated and their uncertainties are quantified. The resulting equilibrium temperature inversion values are plotted as a function of binned wind speed for assumed basis in Figure 4.1.

All GPR models tested clearly show sigmoidal equilibrium structure as in the

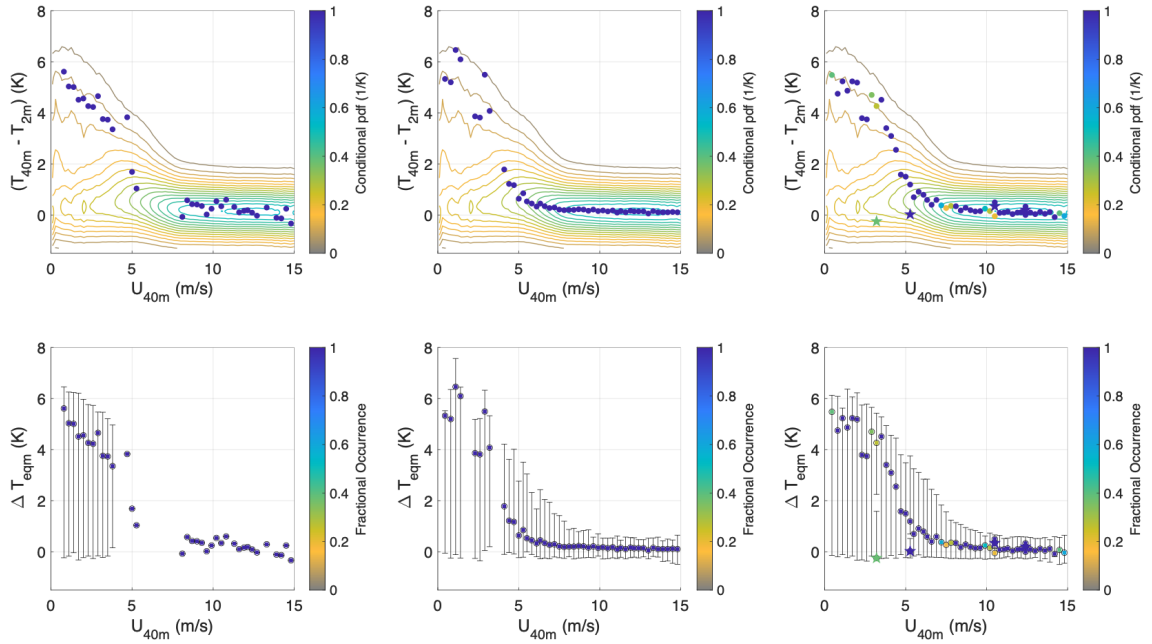



Figure 4.1: Upper row: contours of the kernel density estimate of temperature inversion as a function of wind speed from the Cabauw Meteorological data tower at a height of 40 m. Lower row: extracted equilibrium values along with uncertainties computed. Left: zeroes estimated from GPR fits to data with no basis specified. Center: GPR fits with a linear Legendre polynomial basis. Right: GPR fits with a cubic Legendre polynomial basis. Note that contours extend to statically unstable values of the stratification because of the width of kernel used in the estimation of the pdf; such points aren't present in the analysis.

vdW17 model, with large equilibrium temperature inversions at low wind speeds, and small equilibrium temperature inversions at high wind speeds (Figure 4.1), similar to results from the nondimensionlized model results in Section 3.3. The no basis model does not consistently find equilibrium values in the transitional zone, from 4 m/s to 6 m/s. Because the linear basis assumption is unable to resolve multiple equilibrium points, and there is no theoretical reason to assume linear drift distributions, the cubic polynomial will again be the focus of all subsequent analyses. In keeping with model findings in Section 3.3.3, the cubic polynomial basis produces a smoother equilibrium curve but has spurious unstable points at low $U$, high $\Delta T$ that do not appear for other basis choices. Similar results are found at other stations. These values are not physically justifiable in this framework as neutral stratification should only produce dynamical equilibiria in the presence of zero isothermal net radiation. Such points are therefore hypothesized to be an artefact of the estimation method.

The estimated diffusion coefficient from the GPR fit to the Cabauw 40 m data set is shown in Figure 4.2. A clear structure of small diffusion at small temperature inversion values and high wind speeds, and large diffusion at large temperature inversion values and low wind speeds is evident. This systematic variation in diffusion coefficient values indicates the existence of multiplicative noise in this system. Higher diffusion values in the vSBL suggests additional processes occur when the system is strongly stratified. Observations of intermittent turbulent bursts in the vSBL are consistent with this result [van der Linden et al., 2020, Van de Wiel and Holtslag, 2003].

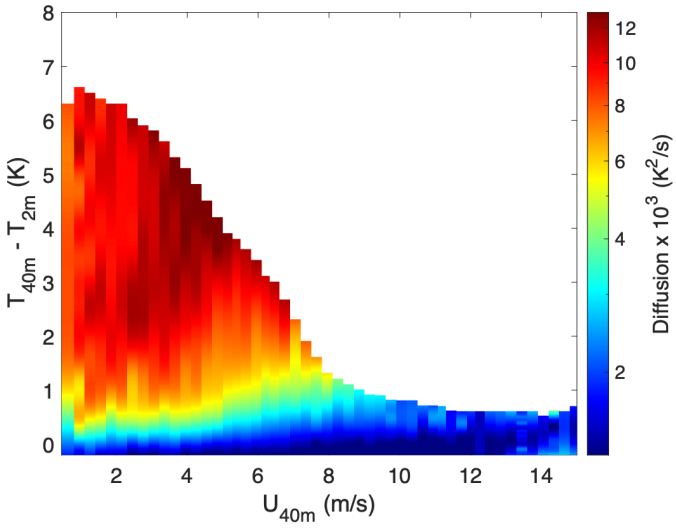



Figure 4.2: The diffusion coefficient estimated from temperature inversion values calculated between 2 and 40 m at Cabauw.

While the complex atmospheric dynamics of the SBL cannot all be accounted for in this highly-simplified, empirical, one dimensional SDE, the reconstructed equilibrium structure produces effective low-dimensional dynamics consistent with the vdW17 model. It is particularly noteworthy that the estimated equilibria do not simply track the maxima of the conditional pdf of data, particularly in the vSBL, where the equilibrium temperature inversion values are far greater than the most frequently observed values. This result differs from those found in the idealized model using additive noise. While this behaviour could be a result of state dependent noise in this region, bringing true equilibria to lower temperature inversions than those predicted by the additive noise structure implicitly assumed by the reconstruction method, it is not observed for the example state dependent stochastic vdW17 model considered in Section 3.4.2. The transitional zone occurs around 5 m/s, consistent

with the minimum wind speed predicted by the MSHF for Cabauw in van Hooijdonk et al. [2015], van de Wiel et al. [2012b] and Abraham and Monahan [2019b]. These results suggest that a one dimensional SDE may capture the essential features of the complex, multidimensional SBL. The assumption of a meteorological tower having a constant, unvarying velocity crossing height for many nights is poor. Model results in Section 3.3.5 suggest that variability of wind speed may reduce the magnitude of the low wind speed, high temperature branch.

### 4.2.2 Upper Height Sensitivity Analysis

The velocity crossing height is used as the reference height in the conceptual 1D model to remove the effects of wind speed variability. However the velocity crossing point is a theoretical construct, often not well defined in reality, and varies in time [van de Wiel et al., 2012b, Abraham and Monahan, 2019c]. To test the dependence of the estimated equilibrium structure on the choice of upper height used to define the inversion strength, the analysis is repeated using $z_h$ at all other available measurement heights at Cabauw: 10 m, 20 m, 80 m, 120 m and 200 m. In each case the value of $z_h$ was also used for the conditioning wind speed. The estimated bifurcation structures at each $z_h$ are shown in Figure 4.2.2. The overall equilibrium structure is broadly reproduced at each height. Inversion strengths increase with measurement height. The transition region becomes sharper and occurs at greater wind speeds for greater measurement heights. Larger temperature inversion values are expected at higher values of $z_h$ since temperature increases with altitude. The transitional wind speed is also expected to increase with $z_h$ as wind speeds generally increase with altitude and the wSBL/vSBL distinction is characteristic of the entire boundary layer. The equilibrium curve becomes more scattered for larger $z_h$, particularly for low wind speeds, likely because these heights exceed the depth of the very shallow boundary layer under vSBL conditions. In this case, measurements are affected by free atmosphere conditions where approximating dynamics of the inversion as a one dimensional SDE may no longer be effective. Nevertheless the equilibrium estimation procedure is still able to extract a relatively clear equilibrium curve. A few unstable points with a hint of backfolding are also observed at the 140 m and 200 m levels. For $z_h = 40$ m and $z_h = 80$ m, the equilibrium structures in which range the velocity crossing point is expected, are very similar. This suggests that while a single altitude

cannot define the velocity crossing height for a tower at all nights, overall results do not appear to be sensitive to the height chosen, if $z_h$ is within the range of altitudes that the velocity crossing height is expected to fall. This height sensitivity test was repeated at all other meteorological towers and no qualitatively different results were found.

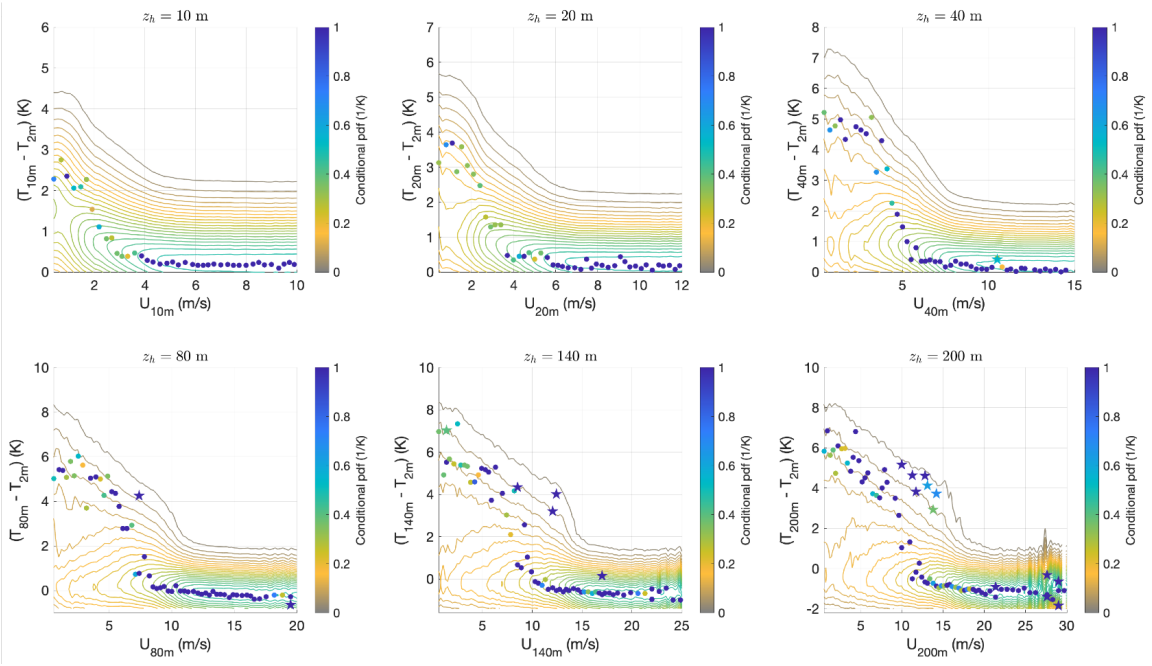



Figure 4.3: Contours show the kernel density of temperature inversion as a function of wind speed from the Cabauw meteorological data tower at various heights. Top left: estimated results at an altitude of 10 m, followed by the 20 m and 40 m heights, from left to right. Bottom left: heights of 80 m, 140 m and 200 m, from left to right. Note that the axes are different for each panel.

The estimated diffusion coefficient plots at each $z_h$ are shown in Figure 4.4. All plots show high diffusion values at low wind speeds and large temperature inversion values. The structure and magnitude of the diffusion coefficient is relatively constant for all heights except $z_h = 200$ m, which is likely outside the boundary layer. This result again suggests that diffusion values are not qualitatively very sensitive to $z_h$.

### 4.2.3 The Effect of Clouds

Ceilometer measurements are taken at Cabauw, reporting the percentage of cloud cover, from June 2008 to April 2017. While this data record is shorter than the tower

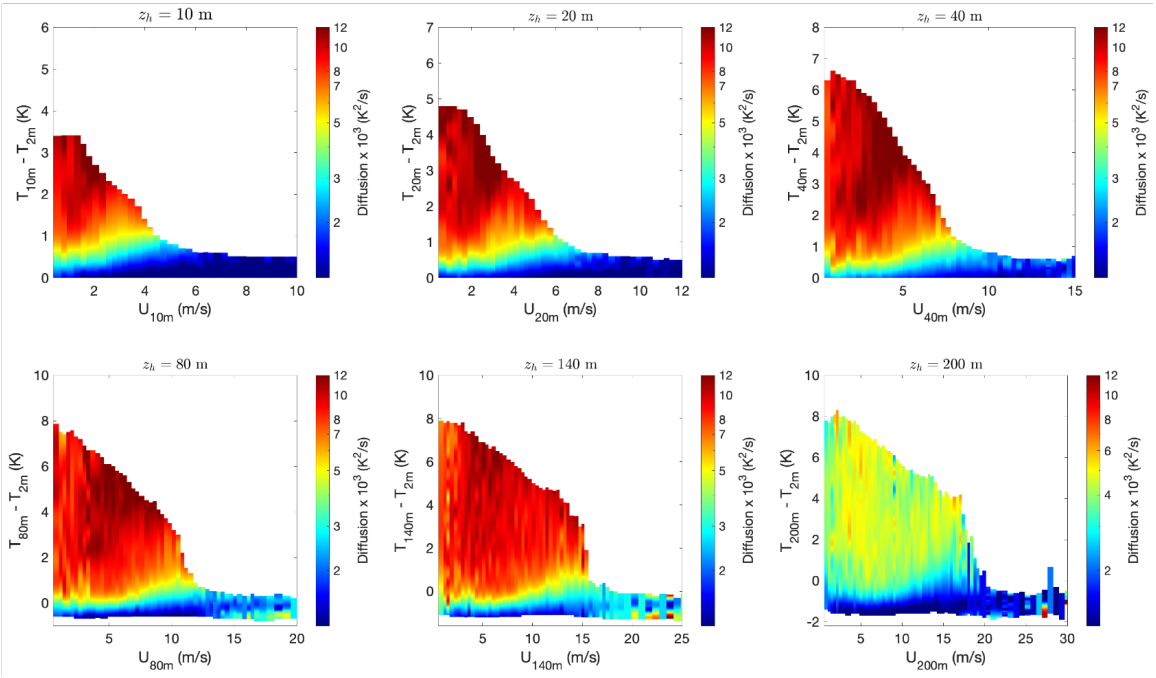



Figure 4.4: Top row: the diffusion coefficient obtained at Cabauw from an altitude of 10 m, followed by the 20 m and 40 m heights, from left to right. Bottom row: altitudes of 80 m, 140 m and 200 m, from left to right. Note that the axes are different for each panel.

temperature and wind measurements its analysis provides insights into the effects of clouds on atmospheric regime occupation. Low level cloud cover (LLCC) less than 20% is characterized as clear-sky, resulting in $9 \times 10^4$ data points. Measurements taken when LLCC is greater than 20% are clasified as cloudy, resulting in $1.5 \times 10^5$ data points. The data set is relatively insensitive to the choice of LLCC cutoff point since the cloud cover data distribution at Cabauw is highly bimodal, with the majority of measurements around 0% or 100%, skewed toward cloudy conditions. The cloudy and clear-sky data sets are analyzed separately, creating empirical bifurcation diagrams for cloudy and clear sky conditions, shown in Figure 4.5. Diffusion coefficients computed for cloudy and clear conditions are shown in Figure 4.6.

The estimated equilibrium structures under both cloudy and clear sky conditions show equilibrium structures broadly similar to those obtained at Cabauw using all the data. The structures however differ in detail. Clear sky conditions show stronger temperature inversions and a sharper transition between regimes. Cloudy conditions show weaker temperature inversions and more gradual transitions between regimes.

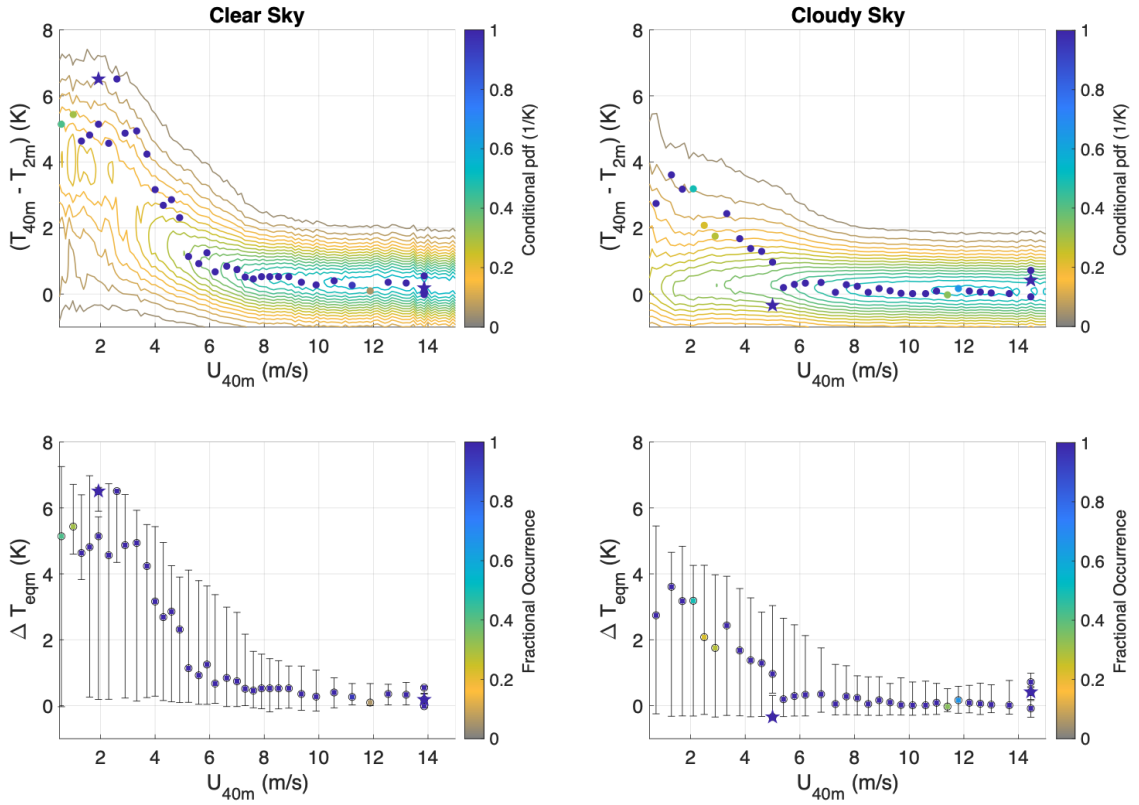



Figure 4.5: Left panels: the empirical bifurcation diagram for clear sky data from Cabauw (LLCC $< 20\%$ ). Right panels: the empirical bifurcation structure under cloudy conditions at Cabauw (LLCC $\geqslant 20\%$ ).

The transition speed under cloudy conditions ($\sim$ 5 m/s) is slightly lower than the transition speed under clear sky conditions ($\sim$ 6 m/s). This result is similar to model results with changes to the specified isothermal net radiation (Section 3.3.6), in which low $\hat{Q}$ values corresponds to weaker temperature inversions and more gradual transitions than higher $\hat{Q}$ values. There are relatively few data for strong inversions under cloudy conditions and weak inversions under clear conditions, consistent with the facts that the wSBL regime is more likely under cloudy conditions, and the vSBL regime is more likely under clear sky conditions [Abraham and Monahan, 2019b, Monahan et al., 2015]. Both cloudy and clear sky data sets admit both vSBL and wSBL states. Diffusion coefficient plots show slightly increased noise under clear skies, at low wind speeds and weak inversions (Figure 4.2), however the sparsity of data available makes it difficult to interpret these results with confidence. Other than this difference the diffusion structures are similar under clear and cloudy skies.

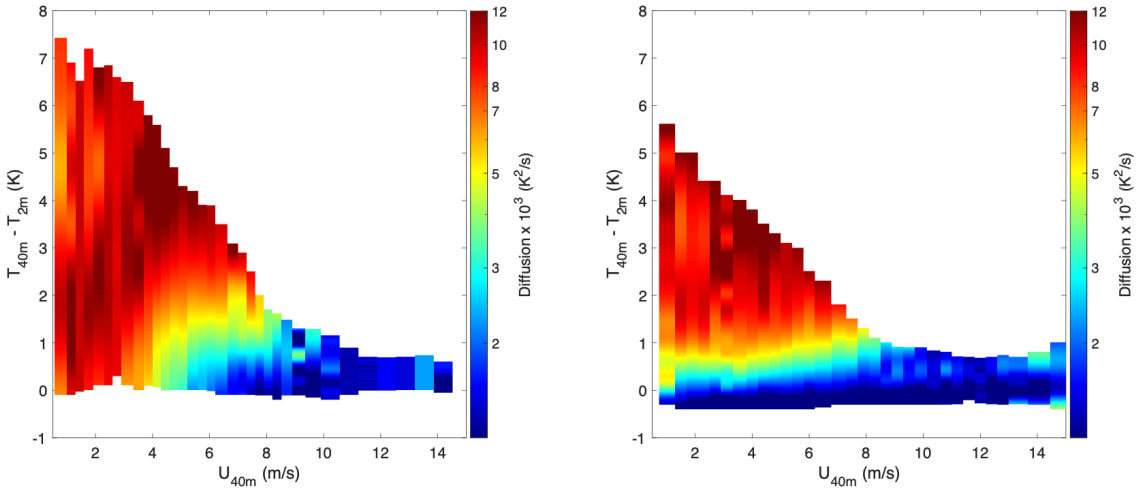



Figure 4.6: Left: diffusion coefficient estimated for the clear sky data set at Cabauw (LLCC $< 0.2$). Right: diffusion coefficients obtained for cloudy conditions (LLCC $\geqslant$ 0.2).

## 4.3 Dome C

In this section time series data from Dome C, located on the East Antarctic Plateau, are analyzed. The meteorological tower is located on a flat snow surface, 900 km from the coast. The flat terrain results in relatively weak katabatic winds. Large scale, dry southerly winds dominate resulting in predominantly clear sky conditions. Strong temperature inversions form due to strong radiative cooling during polar nights and large-scale subsidence. A detailed description of this site can be found in Vignon et al. [2017a,b]. Because snow and ice are good insulators, relatively weak surface exchanges of energy occur. The effective surface coupling parameter, $\lambda$, for this site is therefore expected to be the lowest among the sites considered. From the perspective of the vdW17 model, this location is therefore the most likely to show backfolding of the equilibrium structure. Temperature, radiation and wind speed measurements from 2011 to 2018, Reynolds-averaged over 30 minute intervals, are used in this analysis. The relatively large averaging interval and shorter observational period at this site make it the smallest of the datasets considered, with $9 \times 10^4$ data points. Fewer wind speed bins are therefore used: $N_u = 20$, 25 and 30 wind speed bins are tested. The 9 m measurement height is taken as the reference height, in keeping with the estimates of velocity crossing point in Vignon et al. [2017a,b]. Two near-surface temperature values are tested: the 1.3 m measurement and the surface temperature derived from

long-wave radiative fluxes. As in Vignon et al. [2017a], the surface temperature can be calculated using measurements of upwelling and downwelling longwave radiative flux measurements, $LW_{up}$ and $LW_{down}$, obtained from pyranometers as

$$T_{\mathrm{s}} = \left( \frac{LW_{up} + (\epsilon - 1) LW_{down}}{\epsilon \sigma} \right)^{1/4} \tag{4.1}$$

where $\epsilon = 0.99$ is the snow emissivity at Dome C from Brun et al. [2011] and $\sigma$ is the Stefan-Boltzmann constant. Equation 4.1 is derived from a simple net radiative flux equation, assuming $SW = 0$ at night. The quantity $\epsilon \sigma T_s^4$ represents radiation emitted from the surface, and $(1 - \epsilon) L_{down}$ represents reflected downwelling longwave radiation. This expression assumes surface effective absorptivity equals emissitivty [Hartmann, 2016]. Temperature inversions calculated from longwave radiative flux produce a subset of data with large wind speeds and large temperature inversions. These results are interpreted as spurious, since large temperature inversions cannot be maintained at high wind speeds, and this subset of points does not exist when 1.3 m is used as the lower temperature. Points with wind speeds greater than 10 m/s and temperature inversions greater than 7 K are therefore excluded from the analysis.

The resulting equilibrium structures estimated using $N_u = 20$, 25 and 30 wind speed bins and a cubic polynomial basis are shown for temperature inversions between 1.3 m and 9 m in in Figure 4.7, and between temperatures at the surface and 9 m in Figure 4.8. The estimated equilibrium structures at Dome C again all show the expected sigmoidal structure broadly similar to Cabauw (Figure 4.1). Similar equilibrium structures are estimated for inversions calculated using surface temperatures derived from longwave radiative flux, and using 1.3 m temperature measurements. Equilibrium temperature inversion values are higher for radiative flux derived temperatures than 1.3 m bottom temperatures as a result of the large temperature gradient near the surface. As at Cabauw, equilibria follow the maximum of the pdf only at high wind speeds. At low and intermediate wind speeds, estimated equilibria occur at temperature inversions much larger than the maximum of the pdf.

The transition between regimes occurs around 6 m/s, consistent with the critical wind speed derived from the MSHF framework in Vignon et al. [2017b]. The transitional region, between high temperature inversions at low wind speeds, and low temperature inversions at high wind speeds, occurs more abruptly than at Cabauw,

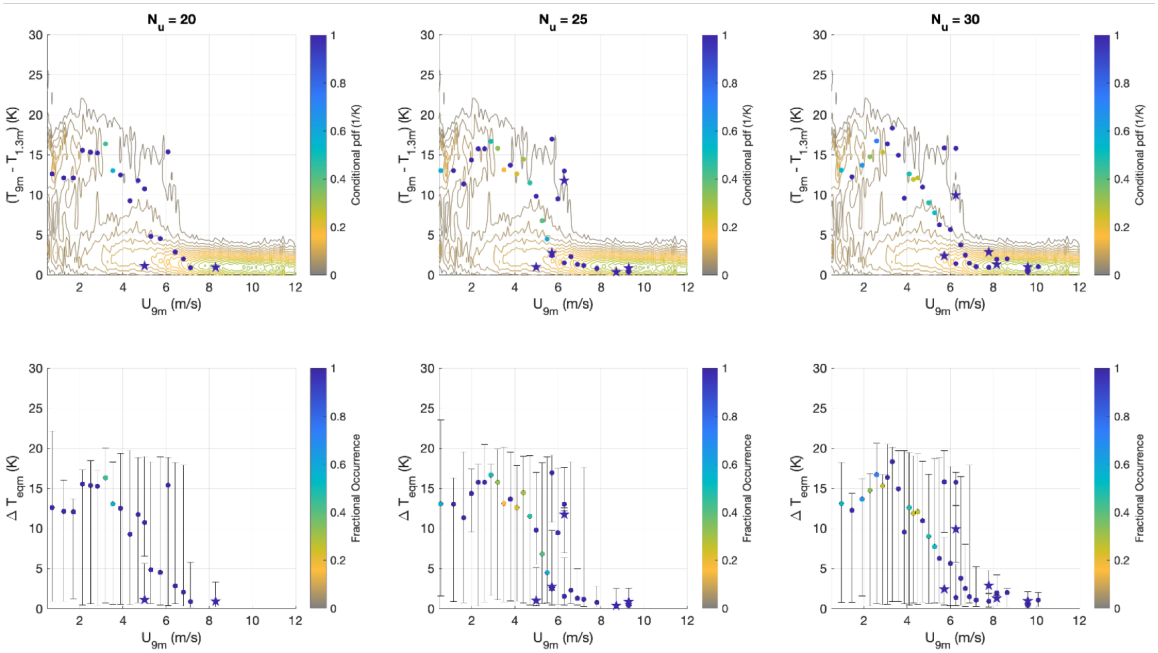



Figure 4.7: Estimated equilibrium structures for Dome C with temperature inversions calculated between 1.3 m and 9 m measurement heights. Left: 20 wind speed bins. Center: 25 wind speed bins. Right: 30 wind speed bins.

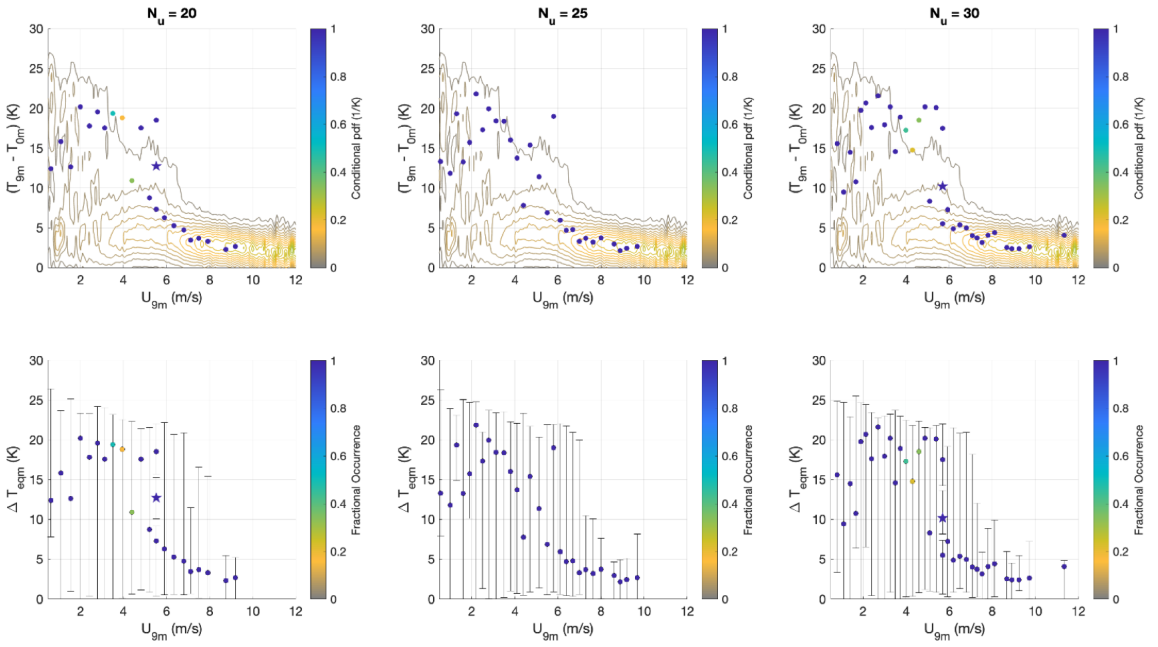
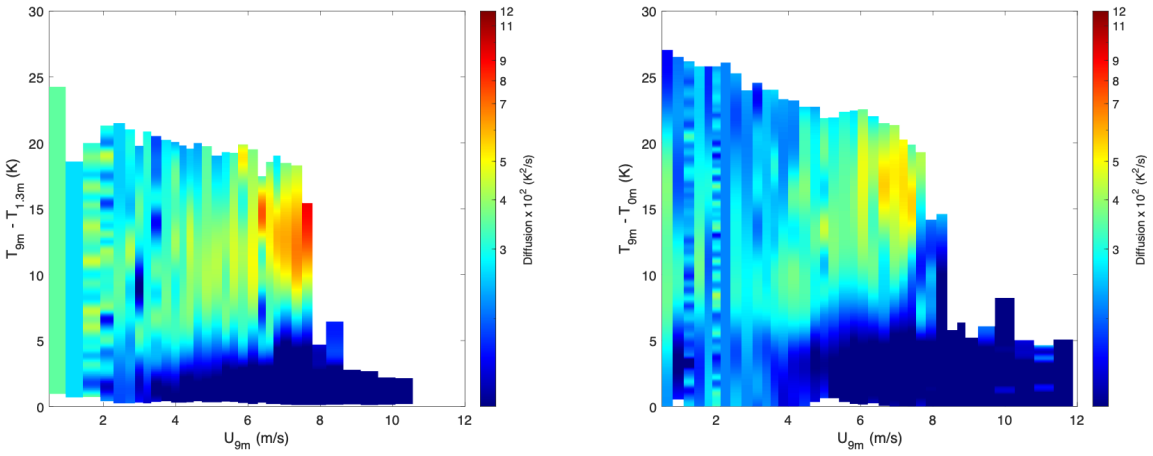



Figure 4.8: As in Figure 4.7 where temperature inversions are taken using surface temperatures calculated using Equations 4.1 and the 9 m measurement height.

and is accompanied by evidence of multiple equilibria. The distribution of equilibria in the transitional region shows evidence of backfolding, with a few unstable equilibrium values. The number, position and stability of equilibria in the transitional region varies however depending on the lower-level temperature used and number of wind speed bins. As shown in Section 3.3.2 the estimation procedure has difficulty resolving backfolding equilibrium branches, particularly in data sets with relatively few points and averaged over relatively long intervals, as is the case at Dome C. These results provide suggestive but not definitive evidence for the existence of unstable equilibria at Dome C.

The estimated diffusion coefficient obtained for temperature inversions calculated with a base of 1.3 m and from longwave radiative flux estimates of surface temperature are shown in Figure 4.9. While results are noisy, both show increased noise at large temperature inversions, under which conditions intermittent turbulent bursts have been observed in Petenko et al. [2019] and Baas et al. [2019]. The turbulent bursts observed at Dome C have been related to Kelvin-Helmholtz instabilities resulting from accelerating winds aloft and decelerating surface friction produced in LES models that occur with a periodicity of about 10 minutes [van der Linden et al., 2020].




x

Figure 4.9: Left: diffusion coefficient estimated from temperature inversions calculated between 1.3 m and 9 m. Right: the diffusion coefficient for temperature inversions calculated using the surface temperature (calculated from Equation 4.1) and 9 m. Both diffusion distributions were calculated using 30 wind speed bins.

## 4.4 Hamburg

The Hamburg data tower measures data every 0.5 seconds and reports values Reynolds-averaged over 1 minute intervals, in contrast to the 10 minute to 30 minute averaging intervals at other stations. This data set therefore allows investigation of the effects of the recording interval on results. From examination of distributions and autocorrelation times within a night at each measured $z_h$, wind speeds are found to vary the least at 50 m (not shown). An estimated velocity crossing point of 50 m is reasonable for Hamburg given that the velocity crossing point at Cabauw, which has a similar climatology, is 40 m. The equilibrium structures obtained for Hamburg data calculated using $z_h = 50$ m and averaged over 30 minutes, 10 minute and 1 minute intervals are compared in Figure 4.10.

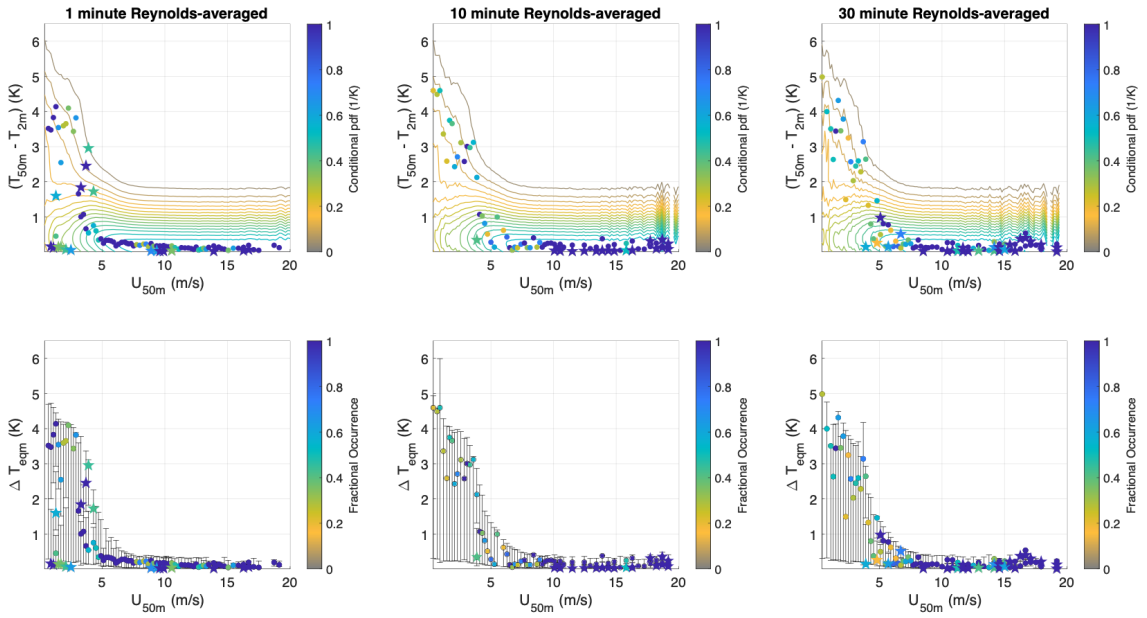



Figure 4.10: Empirical equilibrium structures for Hamburg data. Left panels: 1 minute resolution data. Center panels: data averaged over 10 minute intervals. Right panels: data averaged over 30 minute intervals.

The 10 and 30 minute Reynolds-averaged data are remarkably similar, showing a sigmoidal equilibrium curve, similar to the idealized model (Figure 3.3 ) and Cabauw (Figure 4.1) at which location 10 minute Reynolds-averaged data were used. The 1 minute Reynolds-averaged data shows a similar curve, but with substantially increased scatter and the addition of a number of unstable equilibria at low wind speeds

and low temperature inversion values. The estimation of these low wind speed, low temperature inversion equilibrium points is robust to changes of the basis function and the number of wind speed bins. These equilibrium values are not physically plausible. In particular, they are inconsistent with the surface energy budget expressed by the idealized model introduced in Section 2.1, as they would require a zero isothermal net radiation, $Q_i$. These points are not present in the reconstructions of the dynamics using 10 or 30-minute averaged data, suggesting that at the relatively fine resolution of 1-minute averages, the approximation of the inversion dynamics as an effective low-dimensional dynamical system driven by white noise is no longer reasonable. Effective low dimensional dynamics for this system appear to be more appropriate on timescales of 10 minutes and longer, than on shorter timescales. The 1 minute Reynolds-averaged data also estimates unstable equilibria in the transitional region, indicating the existence of bistability. However because so many unrealistic equilibria are also estimated, the existence of bistability cannot be concluded with any confidence. For this reason, only the equilibrium structures of 10 minute data will be considered further.

The estimated diffusion coefficients for the 1, 10 and 30 minute Reynolds-averaged data are shown in Figure 4.11. Once again, clear evidence of multiplicative noise, with enhanced noise strength in the vSBL is shown, for the 1, 10 and 30 minute averaged data. A notable difference between the diffusion plots is that the diffusion magnitude of 1 minute Reynolds-averaged data is much larger than that of the 10 and 30 minute Reynolds-averaged data, similar to time averaged model results in Section 3.3.4. This result is consistent with the interpretation that averaging reduces signal from short time scale processes not represented by the deterministic system considered and therefore appears as noise. A contrast in noise strength between the vSBL and wSBL exists for all averaged datasets, suggesting that this contrast is not an artefact of averaging like diffusion structures found for the idealized model in Section 3.3.4.

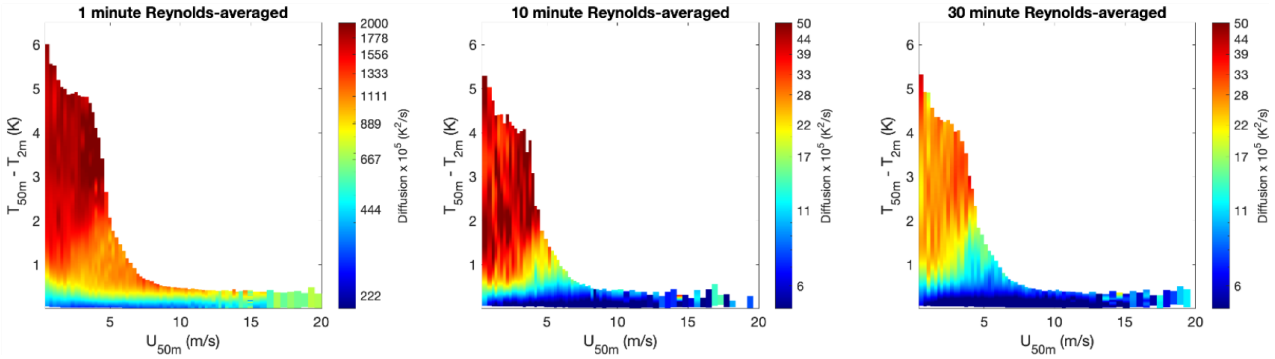



Figure 4.11: Diffusion coefficients estimated at Hamburg. Left: 1 minute Reynolds-averaged data. Center: 10 minute Reynolds-averaged data. Right: 30 minute Reynolds-averaged data.

## 4.5 Station Comparison

The estimated equilibrium structures for all meteorological towers considered (with Hamburg data averaged to 10 minutes) are shown in Figure 4.12. All towers show the sigmoidal equilibrium structure suggested by the vdW17 model, with a region of dynamically stable equilibria at low wind speeds and high temperature inversions, corresponding to the vSBL, and a region of dynamically stable equilibria at high wind speeds and low temperature inversions, corresponding to the wSBL. No evidence of backfolding is found, except at Dome C, as described in Section 4.3. The estimation procedure described in Section 2.2.3 imposes no assumption on the form of the equilibrium structure making these results non-parametric, empirical evidence that the vdW17 model captures essential features of SBL dynamics.

Note that the Boulder data tower has measurements at 10 m, 100 m and 300 m only. Temperature inversions therefore are calculated relative to the lower altitude of 10 m. Use of the 10 m height neglects the approximately logarithmic change in temperature with height at the surface, resulting in weak vSBL equilibrium temperature inversions at Boulder. Additionally, $z_h$ is taken be 100 m, which is likely higher than the ABL under vSBL conditions, potentially resulting in increased scatter, similar to that shown for the $z_h = 140$ m and $z_h = 200$ m cases in Cabauw (Section 4.2.2). A scattered set of unstable equilibrium points appears at Boulder at high wind speeds and weak temperature inversions as well, however these points are not reproduced when no basis or a linear basis is used.

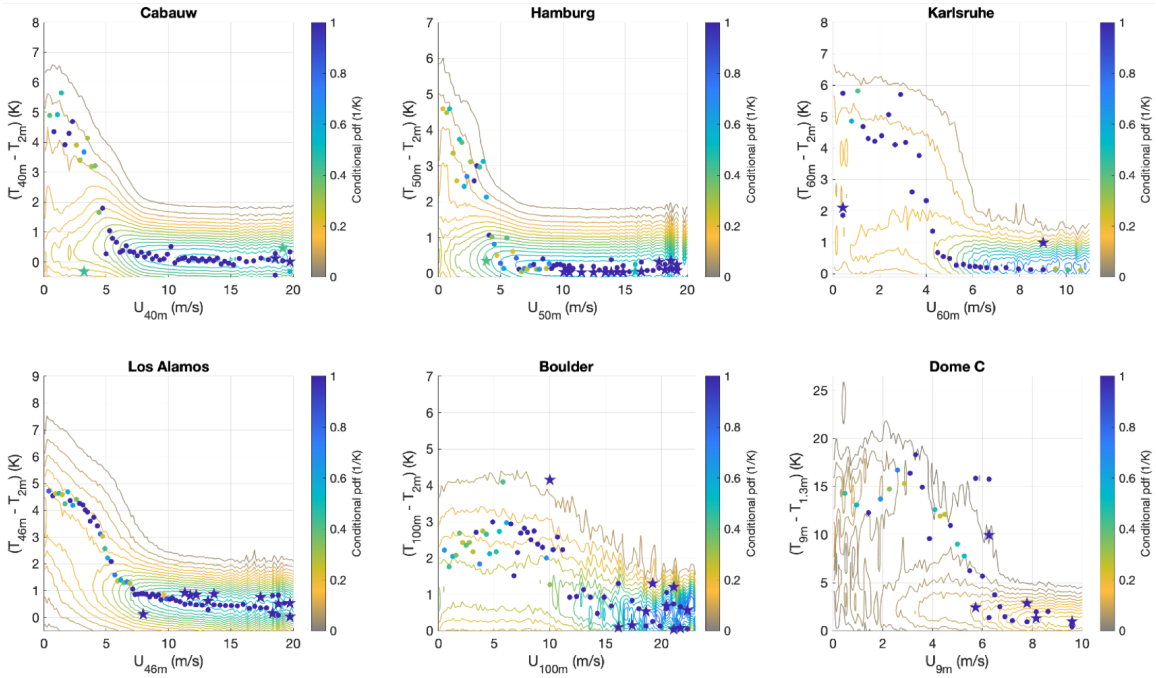



Figure 4.12: Empirical equilibrium diagrams obtained for all meteorological towers considered, along with contours of the estimated inversion pdf. Symbols are as in Figure 4.1.

The estimated diffusion coefficient plots for all towers are shown in Figure 4.13. At all sites, diffusion values are considerably lower at high wind speeds and weak temperature inversions, corresponding to the wSBL than at low wind speeds and strong temperature inversions, corresponding to the vSBL. The diffusion structures are broadly similar between sites.

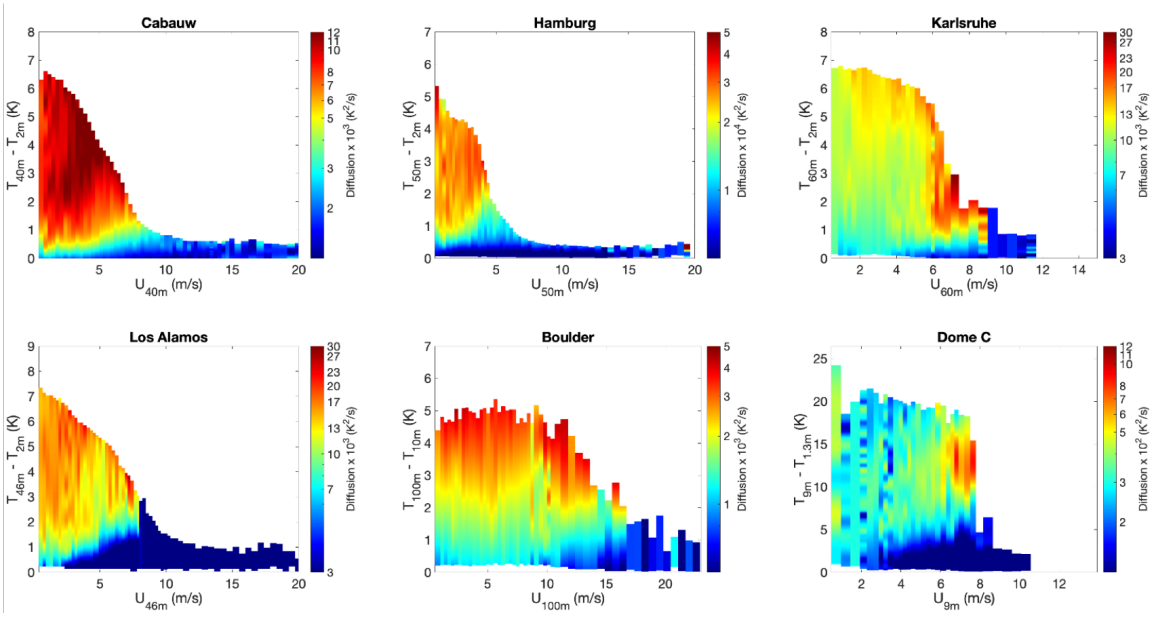



Figure 4.13: As in Figure 4.2 for diffusion coefficients estimated at all meteorological towers considered in this study.

## 4.6 Discussion

There are a number of similarities and differences among the effective stochastic dynamics of inversion strength estimated at the various sites considered. Equilibrium values do not simply follow the maximum probability distribution of observed data in the vSBL region, where noise is found to be state dependent. In the wSBL however, equilibria do follow the maximum of the probability distribution, and noise is found to be state-independent. The transitional region, marking the transition from the vSBL to the wSBL at increasing wind speeds varies by location. For Los Alamos and Cabauw, this transition occurs over a broader range of lower wind speeds than at Dome C and Boulder. The transition is hypothesized to occur at a critical wind speed, according to the vdW17 model and MSHF theory. An expression for the critical wind speed (for fixed net surface radiative heat flux and ground heat flux) is derived in van de Wiel et al. [2012b] to be

$$U_{crit} = \left\{ \frac{27}{4} \frac{g}{R_c T_0 \kappa^2} \frac{|Q_n| - |G|}{\rho c_p} \cdot z_\mathrm{h} \left[ \ln \left( z_h / z_0 \right) \right]^2 \right\}^{1/3}, \tag{4.2}$$

where $T_0$ is the reference temperature 285 K. Values of $U_{crit}$ are estimated to be 5 m/s at Cabauw in van de Wiel et al. [2012b] and 6 m/s at Dome C in Vignon et al.

[2017a], agreeing well with the observed transitional wind speeds of 5 m/s and 6 m/s, respectively.

The equilibrium structures between the six sites differ in detail due to differences in measuring altitudes, averaging time, record length, quantities measured and local conditions. Reynolds-averaging intervals vary between sites from 1 minute at Hamburg, which appears to be too fine of resolution for the effective one-dimensional stochastic dynamics approximation to be appropriate to 30 minutes at Dome C which is too course to resolve turbulent processes hypothesized to trigger regime transitions [van der Linden et al., 2020]. The sensitivity of the equilibrium estimation procedure to time averaging model output is explored in the idealized vdW17 model in Section 3.3.4 showing that fewer unstable equilibria are resolved for an equivalent record length when data are time averaged. The effect of time averaging on observational results is explored in Section 4.4, where 1-minute averaged data showed many apparently spurious unstable equilibria filtered out by averaging in the 10 and 30-minute averaged data. However, 10 and 30-minute averaged data showed fewer unstable equilibria in the transition region in which such equilibria are not physically implausible. These results indicate that the method is most effective for intermediate averaging times, that are fine enough to adequately approximate derivatives, but coarse enough that fast turbulent processes are averaged out.

Figure 4.2.2 shows that as $z_h$ increases, the transitional region becomes sharper, and the equilibrium temperature of the vSBL increases. The equilibrium estimation procedure assumes wind speeds are quasi-stationary within each wind speed bin. Model results show that this is a reasonable assumption if wind speeds vary at a timescale far greater than the systems dynamical timescale. Simulations from the idealized model indicate that the equilibrium estimation procedure is most affected by wind speed variations in the vSBL and transitional regions since the system's dynamical timescale is longest in these regions. As the wind speed's dynamical time scale approaches that of the temperature inversion, temperature inversion values decrease and transitional regions artificially broaden - potentially smearing out backfolding equilibria branches. Multiple equilibria are expected for sharper transitional regions. It is therefore possible that the weak evidence of multiple equilibria found in the tower data is a consequence of wind speed variations. However the sharpness of the transitional region is also somewhat dependent on wind-speed bin width which may in part

explain the weak evidence of unstable equilibria obtained for the towers considered

The sensitivities of estimated diffusion coefficients to the details of the estimation procedure used were considered in Chapter 3. As shown in Sura and Barsugli [2002], approximating stochastic equations by finite differenced time series data creates a quadratic bias in the estimated diffusion coefficients, proportional to the time step of data. This bias is observed in model runs and is strongest in regions where data are sparse. It is therefore is expected that such a bias exists in tower data diffusion plots, particularly where the sampling frequency and data density are low. The quadratic bias may therefore explain the anomalous high noise values in transitional regions, where there are relatively few observations. Analysis of the idealized model in Chapter 3 shows that wind speeds are modelled to vary on a similar timescale to that of the dynamical system, the quadratic bias is magnified. The quadratic bias is strongest in the wSBL when model data are used. Resolved diffusion coefficients in tower data however show higher noise in the vSBL and transitional regions, despite the variations in sampling frequency and data density. Tower data do not appear to be dominated by this quadratic bias, suggesting that the resolved state dependence of noise is real.

The SDE's reconstructed in this study are in Itô form. According to the Wong-Zakai theorem, the natural form of SDE's is Stratonovich form when representing physical systems. However converting results to Stratonovich form requires calculation of derivatives of the diffusion coefficients (Equation 2.16). The fact that derivatives must be taken of potentially noisy diffusion estimates could strongly affect estimates of zeros of the Stratonovitch drift. Therefore results were kept in Itô form, noting the potential for the presence of a noise induced drift in solutions. The reconstruction of the example state-dependant system considered in Section 3.4.2 showed no signs of such a drift. However it is difficult to generalize results from this single example. Results from tower data may be biased as a result of the use of SDE's in Itô form. Equilibrium values consistently occurring above the maximum probability distribution of observed data may be a result of a noise induced drift. Accounting for this noise-induced drift by smoothing estimated diffusion coefficients is an interesting direction of future study.

Time averaging of model output with additive noise results in lower estimated

noise in the wSBL than in the vSBL, likely as a result of the model's lower dynamical timescale in wSBL. As a result, the reconstructed noise appears state-dependent when the true noise is additive. A similar effect is observed in the comparison of 1 minute Reynolds-averaged data at Hamburg to 10 minute Reynolds-averaged data shown in Figure 4.11. This is a caveat to the interpretation of the apparent state-dependent noise intensity in the vSBL observed in the tower data. However, appearance of multiplicative noise in the 1-minute, 10-minute and 30-minute data in Hamburg provides evidence that the noise structure is not simply an artefact of time averaging.

# Chapter 5

# Conclusion

A primary goal of this study was to investigate the effectiveness of representing SBL inversion dynamics as a one dimensional dynamical system, inspired by the vdW17 model, with the addition of a stochastic term representing processes that are not modelled explicitly. Temperature inversion dynamics were reconstructed using a flexible data-driven approach, imposing relatively little parametric structure on the solution. The drift and diffusion terms of a stochastic differential equation were empirically estimated from time series observations of night-time temperature inversions, conditioned by wind speed at a reference height within the SBL. A probabilistic machine learning method, Gaussian Process Regression, was used to fit functional forms to the drift and diffusion. The equilibrium values of the estimated stochastic equation at discretized wind speed bins were then estimated as the zeroes of the estimated drift term. The estimated equilibrium structures show the general form of the equilibrium structure proposed by the vdW17 model at all six land based meteorological towers considered. Limited evidence of hysteresis between vSBL and wSBL regimes is found, however sensitivity tests to the equilibrium estimation procedure suggests that this could be a consequence of data processing and violations of empirical model assumptions. Analyses of output from a stochastic generalization of the vdW17 model show that unstable equilibria are not reliably recovered for vdW17 model runs with data records that are too short or Reynolds-averaging timescales that are too long. Furthermore, the equilibrium estimation procedure assumes wind speeds are quasistationary over binned values. It is observed in Abraham and Monahan [2019b] that rapidly changing winds occur at a range of altitudes when the SBL is nearing a transition point. In Section 3.3.5 it is shown that in model results, when wind varies on shorter timescales, estimated equilibria are lower and the transitional region is broader than

the deterministic solutions. Averaging wind speed values in bins near the transitional region therefore inhibits the resolution of unstable values. Estimated diffusion coefficients show increased noise in the vSBL at all six towers considered. Sensitivity tests in Chapter 3 show that the diffusion estimation method results in a quadratic bias, that is strongest in regions of with data density and longer Reynolds-averaging intervals. The strong contrast in estimated diffusion at all six towers considered is not in keeping with what would be expected for such a bias. This suggests that the diffusion term for an effective 1 dimensional SDE of the SBL should be state-dependent, with stronger noise in the vSBL. State-dependence can be associated with intermittent turbulent bursts observed in the vSBL only, thought to instigate regime transitions. A form for this multiplicative noise in the vdW17 model is proposed. These results indicate that the SBL can be reasonably approximated as an effective one dimensional stochastic dynamical system with state-dependent noise.

The estimation procedure is for an SDE in Itô form, implying that the system is only affected by noise from the past, and not noise processes occurring at the same instant. The Stratonovitch form of the SDE, which incorporates current and past noise processes is more appropriate for this physical system. The two forms of stochastic calculus are identical for additive white noise. However results suggest the existence of multiplicative noise, therefore a noise induced drift likely affects results. No clear evidence of a noise induced drift was found for the example vdW17 model with state-dependent noise. It is difficult to predict what affects this drift has on observations.

Kaiser et al. [2020] similarly used a stochastic extension of the vdW17 model to consider SBL regime occupation and transitions. Only additive white noise is considered in their study. Changes to ARMA fits to time series SBL data were used to predict regime transitions and determine dynamical stability. Quasi-stationarity of wind speed within a bin is not required in their analysis, however a window-length over which the statistical dynamical system is in equilibrium must be specified. Estimation of such a time-window added uncertainty to their results and required finely time sampled data for accuracy, since it is difficult to asses the stationary of dynamics in data averaged over longer time intervals. These results are in contrast to the methods considered in this study, where it is concluded that an intermediate time resolution of approximately 10 minutes is required. Kaiser et al. [2020] were able to confidently resolve an unstable equilibrium branch using 1-minute resolution data

using a data tower in Dumosa, Australia however the data at Dome C has insufficient time resolution for the methodology. It would be interesting to consider the Dumosa data set using this methodology, and to consider the Hamburg data set using the Kaiser et al. [2020] method.

This work demonstrated that the equilibrium structures of vdW17 model are reproduced in observationally derived equilibrium structures. To determine the extent to which the actual SBL displays hysteresis a more thorough approach to wind speed variations, data averaging and multiplicative noise modelling is required. Incorporating the early warning system metric proposed in Kaiser et al. [2020] to filter wind speeds for times at which the system appears to be approaching an unstable state may help facilitate this. A formulation for the noise term in the vdW17 model has been presented. More testing of model parameters is needed to determine if such a formulation can accurately reproduce the observed diffusion values. Detailed analysis of the meteorological processes that affect diffusion values would help with this. The effect of horizontal variations on effective equilibrium structures would also be interesting to consider. Repeating a similar analysis to the network of meteorological towers at Los Alamos, following the work of Abraham and Monahan [2020] may allow for such considerations.

# Appendix

## Data Acknowledgments

The Royal Dutch Meteorological Institute (KMNI) provided tower data from the Cabauw Experimental Site for Atmospheric Research (CSESAR) which can be downloaded here `https://ruisdael-observatory.nl/cesar/`. Data from the Dome C observatory at the Concordia station in Antarctica are provided through a collaboration with the French Polar Institute Paul-Émile Victor (IPEV) and the Italian National Research Program for Antarctica (PNRA). Data are available through the Concordia Station website `http://www.concordiastation.aq/home-1/`. In particular I would like to thank Etienne Vignon for providing additional data and lending his insights into this topic. Hamburg data are provided from the Wettermast, Hamburg of the Meteorological Institute of the University of Hamburg by Felix Ament and Ingo Lange through Carsten Abraham. Martin Kohler and the Institute for Meteorology and Climate Research of the Karlsruhe Institute of Technology (KIT) generously provided Karlsruhe data. The Los Alamos National Laboratory (LANL) provided data from the Environmental Monitoring Plan (EMP) which can be downloaded here `https://weathermachine.lanl.gov/TA6.asp`. The NOAA Earth System Research Laboratory's (ESRL) Physical Sciences Division (PSD) provided data for the Boulder Atmospheric Observatory (BAO) which can be accessed here:`https://psl.noaa.gov/technology/bao/site/`. This work would not be possible without the careful collection, maintenance and accessibility of these datasets generously provided at no cost.

# Bibliography

C. Abraham and A. H. Monahan. Climatological features of the weakly and very stably stratified nocturnal boundary layers. part i: State variables containing information about regime occupation. *Journal of the Atmospheric Sciences*, 76(11): 3455 – 3484, 2019a. doi: 10.1175/JAS-D-18-0261.1. URL https://journals.ametsoc.org/view/journals/atsc/76/11/jas-d-18-0261.1.xml.

C. Abraham and A. H. Monahan. Climatological features of the weakly and very stably stratified nocturnal boundary layers. part ii: Regime occupation and transition statistics and the influence of external drivers. *Journal of the Atmospheric Sciences*, 76(11):3485 – 3504, 2019b. doi: 10.1175/JAS-D-19-0078.1. URL https://journals.ametsoc.org/view/journals/atsc/76/11/jas-d-19-0078.1.xml.

C. Abraham and A. H. Monahan. Climatological features of the weakly and very stably stratified nocturnal boundary layers. part iii: The structure of meteorological state variables in persistent regime nights and across regime transitions. *Journal of the Atmospheric Sciences*, 76(11):3505 – 3527, 2019c. doi: 10.1175/JAS-D-18-0274.1. URL https://journals.ametsoc.org/view/journals/atsc/76/11/jas-d-18-0274.1.xml.

C. Abraham and A. H. Monahan. Spatial dependence of stably stratified nocturnal boundary-layer regimes in complex terrain. *Boundary-Layer Meteorology*, 177(1): 19–47, 2020. doi: 10.1007/s10546-020-00532-x. URL https://doi.org/10.1007/s10546-020-00532-x.

C. Abraham, A. M. Holdsworth, and A. H. Monahan. A prototype stochastic parameterization of regime behaviour in the stably stratified atmospheric boundary layer. *Nonlinear Processes in Geophysics*, 26(4):401–427, 2019. doi: 10.5194/npg-26-401-2019. URL https://npg.copernicus.org/articles/26/401/2019/.

S. P. Arya. *Introduction to micrometeorology*, volume 79. Academic Press, San Diego, 2nd edition, 2001. ISBN 2001271043008;0120593548;9780120593545;.

P. Baas, B. J. H. van de Wiel, E. van Meijgaard, E. Vignon, C. Genthon, S. J. A. van der Linden, and S. R. de Roode. Transitions in the wintertime near-surface temperature inversion at dome c, antarctica. *Quarterly Journal of the Royal Meteorological Society*, 145(720):930–946, 2019. doi: https://doi.org/10.1002/qj.3450. URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3450`.

C. Barthlott, N. Kalthoff, and F. Fiedler. Influence of high-frequency radiation on turbulence measurements on a 200 m tower. *Meteorologische Zeitschrift*, 12(2): 67–71, 04 2003. doi: 10.1127/0941-2948/2003/0012-0067. URL `http://dx.doi.org/10.1127/0941-2948/2003/0012-0067`.

W. Blumen. An observational study of instability and turbulence in nighttime drainage winds. *Boundary-Layer Meteorology*, 28(3):245–269, 1984. doi: 10.1007/BF00121307. URL `https://doi.org/10.1007/BF00121307`.

F. C. Bosveld, P. Baas, A. C. M. Beljaars, A. A. M. Holtslag, J. V.-G. de Arellano, and B. J. H. van de Wiel. Fifty years of atmospheric boundary-layer research at cabauw serving weather, air quality and climate. *Boundary-Layer Meteorology*, 2020. doi: 10.1007/s10546-020-00541-w. URL `https://doi.org/10.1007/s10546-020-00541-w`.

B. M. Bowen, J. A. Baars, and G. L. Stone. Nocturnal Wind Direction Shear and Its Potential Impact on Pollutant Transport. *Journal of Applied Meteorology*, 39(3): 437–445, Mar. 2000. doi: 10.1175/1520-0450(2000)039⟨0437:NWDSAI⟩2.0.CO;2.

D. A. Bruggeman. Los alamos climatology 2016 update. Technical Report LA-UR-17-21060, Los Alamos National Lab. (LANL), USDOE, 02 2017.

I. K. H. Brümmer, Burghard; Lange. Atmospheric boundary layer measurements at the 280 m high hamburg weather mast 1995-2011: mean annual and diurnal cycles. *Meteorologische Zeitschrift*, 21(4):319–335, 2012. doi: 10.1127/0941-2948/2012/0338.

E. Brun, D. Six, G. Picard, V. Vionnet, L. Arnaud, E. Bazile, A. Boone, A. Bouchard, C. Genthon, V. Guidard, P. L. Moigne, F. Rabier, and Y. Seity. Snow/atmosphere coupled simulation at dome c, antarctica. 57(204):721–736, 2011. doi:

DOI:10.3189/002214311797409794. URL https://www.cambridge.org/core/article/snowatmosphere-coupled-simulation-at-dome-c-antarctica/C8479CDFA2E818F4F49F4F275FE256F5.

R. L. Coulter and J. C. Doran. Spatial and temporal occurrences of intermittent turbulence during cases-99. *Boundary-Layer Meteorology*, 105(2):329–349, 2002. doi: 10.1023/A:1019993703820. URL https://doi.org/10.1023/A:1019993703820.

P. K. Deb Burman, T. V. Prabha, R. Morrison, and A. Karipot. A case study of turbulence in the nocturnal boundary layer during the indian summer monsoon. *Boundary-Layer Meteorology*, 169(1):115–138, 2018. doi: 10.1007/s10546-018-0364-4. URL https://doi.org/10.1007/s10546-018-0364-4.

S. H. Derbyshire. Boundary-layer decoupling over cold surfaces as a physical boundary-instability. *Boundary-Layer Meteorology*, 90(2):297–325, 1999. doi: 10.1023/A:1001710014316. URL https://doi.org/10.1023/A:1001710014316.

H. A. Dijkstra, F. W. Wubs, A. K. Cliffe, E. Doedel, I. F. Dragomirescu, B. Eckhardt, A. Y. Gelfgat, A. L. Hazel, V. Lucarini, A. G. Salinger, E. T. Phipps, J. Sanchez-Umbria, H. Schuttelaars, L. S. Tuckerman, and U. Thiele. Numerical Bifurcation Methods and their Application to Fluid Dynamics: Analysis beyond Simulation. *Communications in Computational Physics*, 15(1):1–45, Jan. 2014. doi: 10.4208/cicp.240912.180613a.

M. Dörenkämper, M. Optis, A. Monahan, and G. Steinfeld. On the offshore advection of boundary-layer structures and the influence on offshore wind conditions. *Boundary-Layer Meteorology*, 155(3):459–482, 2015. doi: 10.1007/s10546-015-0008-x. URL https://doi.org/10.1007/s10546-015-0008-x.

D. Faranda, B. Dubrulle, and F. M. E. Pons. Statistical early-warning indicators based on autoregressive moving-average models. *Journal of Physics A: Mathematical and Theoretical*, 47(25):252001, jun 2014. doi: 10.1088/1751-8113/47/25/252001. URL https://doi.org/10.1088/1751-8113/47/25/252001.

T. Foken. 50 years of the monin–obukhov similarity theory. *Boundary-Layer Meteorology*, 119(3):431–447, Jun 2006. ISSN 1573-1472. doi: 10.1007/s10546-006-9048-6. URL https://doi.org/10.1007/s10546-006-9048-6.

R. Friedrich, S. Siegert, J. Peinke, S. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, and G. Pfister. Extracting model equations from experimental data. *Physics Letters A*, 271(3):217 – 222, 2000. ISSN 0375-9601. doi: https://doi.org/10.1016/S0375-9601(00)00334-0. URL http://www.sciencedirect.com/science/article/pii/S0375960100003340.

R. Friedrich, J. Peinke, M. Sahimi, and M. Reza Rahimi Tabar. Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87 – 162, 2011. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2011.05.003. URL http://www.sciencedirect.com/science/article/pii/S0370157311001530.

C. A. García, A. Otero, P. Félix, J. Presedo, and D. G. Márquez. Nonparametric estimation of stochastic differential equations with sparse gaussian processes. *Physical Review E*, 96(2), Aug 2017. doi: 10.1103/physreve.96.022104.

C. Gardiner. *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 1990. ISBN 9780387113579. URL https://books.google.ca/books?id=ZfwaAQAAIAAJ.

C. Genthon, M. S. Town, D. Six, V. Favier, S. Argentini, and A. Pellegrini. Meteorological atmospheric boundary layer measurements and ecmwf analyses during summer at dome c, antarctica. *Journal of Geophysical Research: Atmospheres*, 115(D5), 2020/09/24 2010. doi: 10.1029/2009JD012741. URL https://doi.org/10.1029/2009JD012741.

C. Genthon, D. Six, H. Gallée, P. Grigioni, and A. Pellegrini. Two years of atmospheric boundary layer observations on a 45-m tower at dome c on the antarctic plateau. *Journal of Geophysical Research: Atmospheres*, 118(8):3218–3232, 2020/09/24 2013. doi: 10.1002/jgrd.50128. URL https://doi.org/10.1002/jgrd.50128.

A. A. Grachev, E. L. Andreas, C. W. Fairall, P. S. Guest, and P. O. G. Persson. The critical richardson number and limits of applicability of local similarity theory in the stable boundary layer. *Boundary-Layer Meteorology*, 147(1):51–82, 2013. doi: 10.1007/s10546-012-9771-0. URL https://doi.org/10.1007/s10546-012-9771-0.

S.-E. Gryning, R. Floors, A. Peña, E. Batchvarova, and B. Brümmer. Weibull wind-speed distribution parameters derived from a combination of wind-lidar and

tall-mast measurements over land, coastal and marine sites. *Boundary-Layer Meteorology*, 159(2):329–348, 2016. doi: 10.1007/s10546-015-0113-x. URL https://doi.org/10.1007/s10546-015-0113-x.

D. L. Hartmann. Chapter 4 - the energy balance of the surface. In D. L. Hartmann, editor, *Global Physical Climatology (Second Edition)*, pages 95 – 130. Elsevier, Boston, second edition edition, 2016. ISBN 978-0-12-328531-7. doi: https://doi.org/10.1016/B978-0-12-328531-7.00004-9. URL http://www.sciencedirect.com/science/article/pii/B9780123285317000049.

Y. He, A. H. Monahan, and N. A. McFarlane. Diurnal variations of land surface wind speed probability distributions under clear-sky and low-cloud conditions. *Geophysical Research Letters*, 40(12):3308–3314, 2013. doi: https://doi.org/10.1002/grl.50575. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50575.

A. M. Holdsworth and A. H. Monahan. Turbulent collapse and recovery in the stable boundary layer using an idealized model of pressure-driven flow with a surface energy budget. *Journal of the Atmospheric Sciences*, 76(5):1307–1327, 2019. doi: 10.1175/JAS-D-18-0312.1. URL https://doi.org/10.1175/JAS-D-18-0312.1.

A. M. Holdsworth, T. Rees, and A. H. Monahan. Parameterization sensitivity and instability characteristics of the maximum sustainable heat flux framework for predicting turbulent collapse. *Journal of the Atmospheric Sciences*, 73(9):3527–3540, 2016. doi: 10.1175/JAS-D-16-0057.1. URL https://doi.org/10.1175/JAS-D-16-0057.1.

A. A. M. Holtslag and F. T. M. Nieuwstadt. Scaling the atmospheric boundary layer. *Boundary-Layer Meteorology*, 36(1):201–209, 1986. doi: 10.1007/BF00117468. URL https://doi.org/10.1007/BF00117468.

A. A. M. Holtslag, G. Svensson, P. Baas, S. Basu, B. Beare, A. C. M. Beljaars, F. C. Bosveld, J. Cuxart, J. Lindvall, G. J. Steeneveld, M. Tjernström, and B. J. H. Van De Wiel. Stable atmospheric boundary layers and diurnal cycles: Challenges for weather and climate models. *Bulletin of the American Meteorological Society*, 94(11):1691–1706, 2013. doi: 10.1175/BAMS-D-11-00187.1. URL https://doi.org/10.1175/BAMS-D-11-00187.1.

J. C. Kaimal and J. E. Gaynor. The boulder atmospheric observatory. *J. Climate Appl. Meteor.*, 22(5):863–880, 1983.

A. Kaiser, D. Faranda, S. Krumscheid, D. Belušić, and N. Vercauteren. Detecting Regime Transitions of the Nocturnal and Polar Near-Surface Temperature Inversion. *Journal of the Atmospheric Sciences*, 77(8):2921–2940, 07 2020. ISSN 0022-4928. doi: 10.1175/JAS-D-19-0287.1. URL https://doi.org/10.1175/JAS-D-19-0287.1.

N. Kalthoff and B. Vogel. Counter-current and channelling effect under stable stratification in the area of karlsruhe. *Theoretical and Applied Climatology*, 45(2):113–126, 1992. doi: 10.1007/BF00866400. URL https://doi.org/10.1007/BF00866400.

M. Kohler, J. Metzger, and N. Kalthoff. Trends in temperature and wind speed from 40years of observations at a 200-m high meteorological tower in southwest germany. *International Journal of Climatology*, 38(1):23–34, 2018. doi: 10.1002/joc.5157.

M. A. LeMone, W. M. Angevine, C. S. Bretherton, F. Chen, J. Dudhia, E. Fedorovich, K. B. Katsaros, D. H. Lenschow, L. Mahrt, E. G. Patton, J. Sun, M. Tjernström, and J. Weil. 100 Years of Progress in Boundary Layer Meteorology. *Meteorological Monographs*, 59:9.1–9.85, 07 2019. ISSN 0065-9401. doi: 10.1175/AMSMONOGRAPHS-D-18-0013.1. URL https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0013.1.

L. Mahrt. Stratified atmospheric boundary layers and breakdown of models. *Theoretical and Computational Fluid Dynamics*, 11(3):263–279, Jun 1998. ISSN 1432-2250. doi: 10.1007/s001620050093. URL https://doi.org/10.1007/s001620050093.

L. Mahrt. Stably stratified atmospheric boundary layers. *Annual Review of Fluid Mechanics*, 46(1):23–45, 2014. doi: 10.1146/annurev-fluid-010313-141354. URL https://doi.org/10.1146/annurev-fluid-010313-141354.

A. H. Monahan. *Encyclopedia of Natural Resources: Air*, volume 2, chapter Wind Speed Probability Distribution, pages 1084–1088. Taylor Francis, 1 edition, Oct 2014. doi: 10.1201/9780203757611.

A. H. Monahan, T. Rees, Y. He, and N. McFarlane. Multiple regimes of wind, stratification, and turbulence in the stable boundary layer. *Journal of the At-*

*mospheric Sciences*, 72(8):3178–3198, 2015. doi: 10.1175/JAS-D-14-0311.1. URL https://doi.org/10.1175/JAS-D-14-0311.1.

I. Petenko, S. Argentini, G. Casasanta, C. Genthon, and M. Kallistratova. Stable surface-based turbulent layer during the polar winter at dome c, antarctica: Sodar and in situ observations. *Boundary-Layer Meteorology*, 171(1):101–128, 2019. doi: 10.1007/s10546-018-0419-6. URL https://doi.org/10.1007/s10546-018-0419-6.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006. doi: ISBN0-262-18253-X.].

M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009. doi: 10.1038/nature08227. URL https://doi.org/10.1038/nature08227.

S. Siegert, R. Friedrich, and J. Peinke. Analysis of data sets of stochastic systems. *Physics Letters A*, 243(5):275 – 280, 1998. ISSN 0375-9601. doi: https://doi.org/10.1016/S0375-9601(98)00283-7. URL http://www.sciencedirect.com/science/article/pii/S0375960198002837.

M. Sommerfeld, C. Crawford, A. Monahan, and I. Bastigkeit. Lidar-based characterization of mid-altitude wind conditions for airborne wind energy systems. *Wind Energy*, 22(8):1101–1120, 2019. doi: 10.1002/we.2343. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/we.2343.

J. Sun, S. P. Burns, D. H. Lenschow, R. Banta, R. Newsom, R. Coulter, S. Frasier, T. Ince, C. Nappo, J. Cuxart, W. Blumen, X. Lee, and X.-Z. Hu. Intermittent turbulence associated with a density current passage in the stable boundary layer. *Boundary-Layer Meteorology*, 105(2):199–219, 2002. doi: 10.1023/A:1019969131774. URL https://doi.org/10.1023/A:1019969131774.

P. Sura and J. Barsugli. A note on estimating drift and diffusion parameters from timeseries. *Physics Letters A*, 305(5):304 – 311, 2002. ISSN 0375-9601. doi: https://doi.org/10.1016/S0375-9601(02)01474-3. URL http://www.sciencedirect.com/science/article/pii/S0375960102014743.

K. Twardowska. Wong-zakai approximations for stochastic differential equations. *Acta Applicandae Mathematica*, 43(3):317–359, 1996. doi: 10.1007/BF00047670. URL https://doi.org/10.1007/BF00047670.

A. F. M. O. K. H. H. A. R. D. B. Van de Wiel, B. J. H. and A. A. M. Holtslag. Intermittent turbulence in the stable boundary layer over land. part iii: A classification for observations during cases-99. *J. Atmos. Sci.*, 60(20):2509–2522, 2003. doi: doi.org/10.1175/1520-0469(2003)060⟨2509:ITITSB⟩2.0.CO;2.

B. J. H. van de Wiel, A. F. Moene, G. J. Steeneveld, O. K. Hartogensis, and A. A. M. Holtslag. Predicting the collapse of turbulence in stably stratified boundary layers. *Flow, Turbulence and Combustion*, 79(3):251–274, 2007. doi: 10.1007/s10494-007-9094-2. URL https://doi.org/10.1007/s10494-007-9094-2.

B. J. H. van de Wiel, A. F. Moene, and H. J. J. Jonker. The cessation of continuous turbulence as precursor of the very stable nocturnal boundary layer. *Journal of the Atmospheric Sciences*, 69(1):3097–3115, 2012a. doi: https://doi.org/10.1175/JAS-D-12-064.1.

B. J. H. van de Wiel, A. F. Moene, H. J. J. Jonker, P. Baas, S. Basu, J. M. M. Donda, J. Sun, and A. A. M. Holtslag. The minimum wind speed for sustainable turbulence in the nocturnal boundary layer. *Journal of the Atmospheric Sciences*, 69(11):3116 – 3127, 2012b. doi: 10.1175/JAS-D-12-0107.1. URL https://journals.ametsoc.org/view/journals/atsc/69/11/jas-d-12-0107.1.xml.

B. J. H. van de Wiel, E. Vignon, P. Baas, I. G. S. van Hooijdonk, S. J. A. van der Linden, J. Antoon van Hooft, F. C. Bosveld, S. R. de Roode, A. F. Moene, and C. Genthon. Regime transitions in near-surface temperature inversions: A conceptual model. *Journal of the Atmospheric Sciences*, 74(4):1057–1073, 2017. doi: 10.1175/JAS-D-16-0180.1. URL https://doi.org/10.1175/JAS-D-16-0180.1.

S. J. A. van der Linden, B. J. H. van de Wiel, I. Petenko, C. C. van Heerwaarden, P. Baas, and H. J. J. Jonker. A businger mechanism for intermittent bursting in the stable boundary layer. *Journal of the Atmospheric Sciences*, 77(10):3343 – 3360, 2020. doi: 10.1175/JAS-D-19-0309.1. URL https://journals.ametsoc.org/view/journals/atsc/77/10/jasD190309.xml.

I. G. S. van Hooijdonk, J. M. M. Donda, H. J. H. Clercx, F. C. Bosveld, and B. J. H. van de Wiel. Shear capacity as prognostic for nocturnal boundary

layer regimes. *Journal of the Atmospheric Sciences*, 72(4):1518 – 1532, 2015. doi: 10.1175/JAS-D-14-0140.1. URL https://journals.ametsoc.org/view/journals/atsc/72/4/jas-d-14-0140.1.xml.

I. G. S. van Hooijdonk, H. J. H. Clercx, C. Abraham, A. M. Holdsworth, A. H. Monahan, E. Vignon, A. F. Moene, P. Baas, and B. J. H. van de Wiel. Near-Surface Temperature Inversion Growth Rate during the Onset of the Stable Boundary Layer. *Journal of the Atmospheric Sciences*, 74(10):3433–3449, 10 2017a. ISSN 0022-4928. doi: 10.1175/JAS-D-17-0084.1. URL https://doi.org/10.1175/JAS-D-17-0084.1.

I. G. S. van Hooijdonk, A. F. Moene, M. Scheffer, H. J. H. Clercx, and B. J. H. van de Wiel. Early warning signals for regime transition in the stable boundary layer: A model study. *Boundary-Layer Meteorology*, 162(2):283–306, 2017b. doi: 10.1007/s10546-016-0199-9. URL https://doi.org/10.1007/s10546-016-0199-9.

A. P. Van Ulden and J. Wieringa. Atmospheric boundary layer research at cabauw. *Boundary-Layer Meteorology*, 78(1):39–69, 1996. doi: 10.1007/BF00122486. URL https://doi.org/10.1007/BF00122486.

E. Vignon, C. Genthon, H. Barral, C. Amory, G. Picard, H. Gallée, G. Casasanta, and S. Argentini. Momentum- and heat-flux parametrization at dome c, antarctica: A sensitivity study. *Boundary-Layer Meteorology*, 162(2):341–367, 2017a. doi: 10.1007/s10546-016-0192-3. URL https://doi.org/10.1007/s10546-016-0192-3.

E. Vignon, B. J. H. van de Wiel, I. G. S. van Hooijdonk, C. Genthon, S. J. A. van der Linden, J. A. van Hooft, P. Baas, W. Maurel, O. Traullé, and G. Casasanta. Stable boundary-layer regimes at dome c, antarctica: observation and analysis. *Quarterly Journal of the Royal Meteorological Society*, 143(704):1241–1253, 2017b. doi: 10.1002/qj.2998. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2998.

A. Wenzel, N. Kalthoff, and V. Horlacher. On the profiles of wind velocity in the roughness sublayer above a coniferous forest. *Boundary-Layer Meteorology*, 84(2):219–230, 1997. doi: 10.1023/A:1000444911103. URL https://doi.org/10.1023/A:1000444911103.