



**HAL**  
open science

## Detecting Regime Transitions of the Nocturnal and Polar Near-Surface Temperature Inversion

Amandine Kaiser, Davide Faranda, Sebastian Krumscheid, Danijel Belušić,  
Nikki Vercauteren

► **To cite this version:**

Amandine Kaiser, Davide Faranda, Sebastian Krumscheid, Danijel Belušić, Nikki Vercauteren. Detecting Regime Transitions of the Nocturnal and Polar Near-Surface Temperature Inversion. *Journal of the Atmospheric Sciences*, 2020, 77 (8), pp.2921-2940. 10.1175/JAS-D-19-0287.1 . hal-02909915

**HAL Id: hal-02909915**

**<https://hal.science/hal-02909915>**

Submitted on 31 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## ABSTRACT

18 Many natural systems undergo critical transitions, i.e. sudden shifts from  
19 one dynamical regime to another. In the climate system, the atmospheric  
20 boundary layer can experience sudden transitions between fully turbulent  
21 states and quiescent, quasi-laminar states. Such rapid transitions are observed  
22 in Polar regions or at night when the atmospheric boundary layer is stably  
23 stratified, and they have important consequences in the strength of mixing  
24 with the higher levels of the atmosphere. To analyze the stable boundary layer,  
25 many approaches rely on the identification of regimes that are commonly de-  
26 noted as weakly and very stable regimes. Detecting transitions between the  
27 regimes is crucial for modeling purposes.

28 In this work a combination of methods from dynamical systems and statisti-  
29 cal modeling is applied to study these regime transitions and to develop an  
30 early-warning signal that can be applied to non-stationary field data. The pre-  
31 sented metric aims at detecting nearing transitions by statistically quantifying  
32 the deviation from the dynamics expected when the system is close to a sta-  
33 ble equilibrium. An idealized stochastic model of near-surface inversions is  
34 used to evaluate the potential of the metric as an indicator of regime transi-  
35 tions. In this stochastic system, small-scale perturbations can be amplified  
36 due to the nonlinearity, resulting in transitions between two possible equilib-  
37 ria of the temperature inversion. The simulations show such noise-induced  
38 regime transitions, successfully identified by the indicator. The indicator is  
39 further applied to time series data from nocturnal and Polar meteorological  
40 measurements.

## 41 **1. Introduction**

42 The atmospheric boundary layer (ABL) is the lowest part of the atmosphere that is directly influ-  
43 enced by the Earth's surface and across which turbulent exchanges of momentum, heat and matter  
44 between the surface and the free atmosphere occur. During daytime, surface warming leads to an  
45 unstable or convective boundary layer. During clear-sky nights, radiative cooling leads to a sur-  
46 face that is cooler than the air aloft and the ABL becomes stably stratified. The stable stratification  
47 can also arise when warm air is advected over a colder surface, which is a frequent event in Polar  
48 regions. Turbulence in the resulting stable boundary layer (SBL) is subject to buoyant damping  
49 and is only maintained through mechanical production of turbulent kinetic energy (TKE). Under-  
50 standing and modeling the SBL is essential for regional and global atmospheric models, yet there  
51 are many well-documented challenges to simulate stably stratified atmospheric flows (Sandu et al.  
52 2013; Holtslag et al. 2013; LeMone et al. 2018). One of the challenges is to develop an accu-  
53 rate understanding and representation of distinct regimes of the SBL and transitions between them  
54 (Baas et al. 2017).

55 Numerous observational and modeling studies show that the SBL can be classified, to a first  
56 approximation, in a weakly stable regime in which turbulence is continuous, and a very stable  
57 regime with patchy and intermittent turbulence, requiring a different modeling approach (Mahrt  
58 2014). The weakly stable regime typically occurs when cloud cover limits nocturnal radiative  
59 cooling at the land surface, or with strong winds associated to wind shear that produces enough  
60 TKE to sustain turbulence. The vertical mixing is therefore maintained and a well-defined bound-  
61 ary layer usually exists in which turbulent quantities decrease upwards from the surface layer  
62 following the classical model of Monin-Obukhov similarity theory and related existing concepts  
63 (Mahrt 2014). The associated temperature stratification, or temperature inversion, is weak. The

64 strongly stable regime occurs with strong stratification and weak winds and does not follow the  
65 traditional concept of a boundary layer. Transitions from weakly stable to strongly stable regimes  
66 are caused by a strong net radiative cooling at the surface which increases the inversion strength  
67 and eventually leads to suppressed vertical exchanges, unless winds are strong enough to maintain  
68 turbulence (van de Wiel et al. 2007). The reduced vertical mixing results in a decoupling from the  
69 surface, such that similarity theory breaks down (Acevedo et al. 2015). Intermittent bursts tend to  
70 be responsible for most of the turbulent transport (Acevedo et al. 2006; Vercauteren et al. 2016).  
71 Such bursts alter the temperature inversion and can sometimes drive transitions from strongly to  
72 weakly stable boundary layers.

73 Transitions between the different SBL regimes have been found by modeling studies to be dy-  
74 namically unpredictable. Based on a numerical model representing the exchanges between the  
75 surface and the SBL using a very simple two-layer scheme, McNider (1995) showed the existence  
76 of bi-stable equilibria of the system, which can thus transition between very different states under  
77 the influence of random perturbations. Interacting nonlinear processes that lead to this bi-stability  
78 partly involve thermal processes at the land surface, as was highlighted following the hypothesis  
79 that continuous turbulence requires the turbulence heat flux to balance the surface energy demand  
80 resulting from radiative cooling (van de Wiel et al. 2007, 2012b, 2017). According to this max-  
81 imum sustainable heat flux hypothesis, a radiative heat loss that is stronger than the maximum  
82 turbulent heat flux that can be supported by the flow with a given wind profile will lead to the  
83 cessation of turbulence (van de Wiel et al. 2012a) and thus to a regime transition. This concept is  
84 used by van Hooijdonk et al. (2015) to show that the shear over a layer of certain thickness can  
85 predict SBL regimes when sufficient averaging of data is considered. Based on observations, Sun  
86 et al. (2012) identify a height and site-dependent wind speed threshold that triggers a transition  
87 between a regime in which turbulence increases slowly with increasing wind speed from a regime

88 where turbulence increases rapidly with the wind speed. The change of the relationship between  
89 the turbulence and the mean wind speed occurs abruptly at the transition. Sun et al. (2016) attribute  
90 this difference to the turbulent energy partitioning between turbulent kinetic energy and turbulent  
91 potential energy (TPE): in the very stable regime, shear-induced turbulence will have to enhance  
92 the TPE in order to counter the stable stratification before enhancing the TKE.

93 The combined importance of the wind speed and of the surface thermal processes has also been  
94 evidenced by numerical studies using idealized single-column models of the atmosphere. Single-  
95 column models with a first-order turbulence closure scheme (Baas et al. 2017, 2019; Holdsworth  
96 and Monahan 2019) or a second-order closure scheme (Maroneze et al. 2019) are able to rep-  
97 resentatively simulate transitions from weakly to strongly-stable regimes. Yet, direct numerical  
98 simulations show that transitions from strongly to weakly-stable regimes can occur following a  
99 localized, random perturbation of the flow (Donda et al. 2015). Field studies have also highlighted  
100 examples of transitions induced by small-scale perturbations of the flow (Sun et al. 2012). In fact,  
101 a statistical classification scheme introduced by Vercauteren and Klein (2015) shows that the SBL  
102 flow transitions between periods of strong and weak influence of small-scale, non-turbulent flow  
103 motions on TKE production in the SBL. Such submeso-scale fluctuations of the flow (e.g. induced  
104 by various kind of surface heterogeneity) are typically not represented in models but are important  
105 in strongly stable regimes (Vercauteren et al. 2019), and may trigger regime transitions. Stochastic  
106 modeling approaches are a promising framework to analyze their impact on regime transitions.  
107 A related statistical classification of the Reynolds-averaged boundary-layer states introduced by  
108 Monahan et al. (2015) highlighted that regime transitions are a common feature of SBL dynamics  
109 around the globe (Abraham and Monahan 2019). Regime transitions typically take an abrupt char-  
110 acter (Baas et al. 2019; Acevedo et al. 2019). Predicting the transition point remains a challenge  
111 (van Hooijdonk et al. 2016).

112 Abrupt or critical transitions are ubiquitous in complex natural and social systems. The con-  
113 cept of critical transition is formally defined in dynamical systems theory and relates to the notion  
114 of bifurcation (Kuznetsov 2013). When the dynamics is controlled by a system of equations de-  
115 pending on an external parameter (often called forcing), the stability of the equilibrium solutions  
116 can change abruptly and this is also reflected on macroscopic observables of the system. Some-  
117 times, one can have early-warning signals of a transition because the systems experience some  
118 influences of the bifurcated state before actually reaching it. The motion of a particle undergo-  
119 ing random fluctuations in an asymmetric double-well energy potential  $V$  is a minimal system to  
120 detect early warnings, in which each well or local minimum of the energy potential corresponds  
121 to a stable equilibrium state of the system. For a small fixed level of noise, the control parameter  
122 is  $\Delta V$ , the depth of the well in which the particle is located, leading to an energy barrier that the  
123 particle has to overcome in order to transition to the second stable equilibrium. If  $\Delta V$  is large, the  
124 distribution of the positions of the particle will be quasi Gaussian and the autocorrelation function  
125 of the position of the particle will have an exponential decay. Conversely, if  $\Delta V$  is reduced when  
126 the particle approaches the bifurcation point, then the particle position's distribution starts to "feel"  
127 the effect of the other state and the distribution will be skewed towards the new state. Similarly, the  
128 excursions from the equilibrium position will become larger, increasing the autocorrelation time.  
129 Early-warning signals can then, e.g., be defined based on the changes of the autocorrelation time.

130 These first early-warning signs have been successfully applied to several systems with excellent  
131 results (Scheffer et al. 2009, 2012), including the present context of SBL regime transitions (van  
132 Hooijdonk et al. 2016). However, sometimes, transitions can happen without detectable early-  
133 warnings (Hastings and Wysham 2010). The main limitation of early-warning signals based on  
134 the increase of autocorrelation is that their activation does not always correspond to a bifurcation.  
135 Indeed, if a single well potential widens, as it can occur in non-stationary systems, the distribution

136 of a particle's position experiences the same increase in skewness and autocorrelation function  
137 without the need of approaching a bifurcation (Lenton et al. 2012; Faranda et al. 2014). In our  
138 context of SBL flows, non-stationarity of the energy potential governing the dynamics can be due  
139 to changes in the mean wind speed or cloud cover, for example. For these reasons, (Faranda  
140 et al. 2014) have introduced a new class of early-warning indicators based on defining a distance  
141 from the dynamics expected from a particle evolving in a single-well potential. The suggested  
142 indicator statistically quantifies the dynamical stability of the observables and was already used  
143 by Nevo et al. (2017) to show that strongly stable flow regimes are dynamically unstable and may  
144 require high-order turbulence closure schemes to represent the dynamics. Alternative new early  
145 warnings are based on the combination of statistical properties of observables when approaching  
146 the bifurcation (Chen et al. 2012).

147 In the present analysis, we investigate if the early-warning indicator introduced by Faranda  
148 et al. (2014) can be used to detect nearing transitions between SBL flow regimes, based on both  
149 simulated data and field measurements. We show that the conceptual model that was recently  
150 suggested by van de Wiel et al. (2017) to understand SBL regime transitions in terms of thermal  
151 coupling of the land surface is equivalent to a dynamical system representing the evolution of  
152 the temperature inversion evolving in a double-well energy potential. We extend this conceptual  
153 model to a stochastic model where added noise represents the effect of natural fluctuations of the  
154 temperature inversion's rate of change. The resilience of equilibria of the non-random model to  
155 perturbations as well as the bifurcation points are known analytically (as was discussed in van de  
156 Wiel et al. (2017)), and we thus use the simulated data to test our indicator. Additionally, the  
157 indicator relies on calculating statistical properties of the data with a moving window approach  
158 and is sensitive to the choice of the window length. We suggest two complementary, data-driven  
159 but physically justified approaches to define an appropriate window length for which results can

160 be trusted. Finally, the indicator is applied to nocturnal temperature inversion data from a site in  
161 Dumosa, Australia as well as from temperature inversion data from Dome C, Antarctica.

## 162 **2. Analyzing the dynamical stability of stable boundary layer regimes**

163 The goal of our study is to investigate if a statistical early-warning indicator of regime transitions  
164 can be successfully used to detect nearing regime transitions in the SBL. In section a, the concep-  
165 tual model introduced by van de Wiel et al. (2017) to study regime transitions will be introduced,  
166 along with its dynamical stability properties. In section b, the model is extended to a stochastic  
167 model in which noise represents fluctuations in the dynamics of the near-surface temperature in-  
168 version. In section c, we present a statistical indicator that was introduced in Faranda et al. (2014)  
169 and applied to SBL turbulence data in Nevo et al. (2017) to estimate the dynamical equilibrium  
170 properties of time series, based on a combination of dynamical systems concepts and stochastic  
171 processes tools. The conceptual model describes the evolution of the near-surface temperature in-  
172 version and is used to produce time series of controlled data for which the theoretical equilibrium  
173 properties are known.

### 174 *a. Model description and linear stability analysis*

175 A conceptual model was introduced by van de Wiel et al. (2017) to study regime transitions of  
176 near-surface temperature inversions in the nocturnal and polar atmospheric boundary layer. The  
177 authors were able to determine a connection between the dynamical stability of the temperature  
178 inversion and the ambient wind speed  $U$  through their model and measurements. Mathematically  
179 speaking, the model is a dynamical system represented by a first order ordinary differential  
180 equation, abbreviated ODE, which describes the time evolution of the difference between the  
181 temperature at a reference height  $T_r$  and the surface temperature  $T_s$ . Although the equilibrium

182 properties of the system and the dynamical stability properties (i.e. the resilience to perturbations)  
 183 of all equilibria states were thoroughly discussed in van de Wiel et al. (2017), for the sake of  
 184 completeness we briefly introduce the model and summarize the linear stability analysis of  
 185 equilibrium points of the resulting ODE for different values of a bifurcation parameter. The  
 186 bifurcation parameter is related to the ambient wind speed.

187  
 188 Assuming that the wind speed and temperature are constant at a given height  $z_r$ , the fol-  
 189 lowing equation describes the evolution of the near-surface inversion strength, based on a simple  
 190 energy balance at the ground surface:

$$c_v \frac{d\Delta T}{dt} = Q_n - G - H. \quad (1)$$

191 In this energy balance model,  $c_v$  is the heat capacity of the soil,  $\Delta T = T_r - T_s$  is the inversion  
 192 strength between the temperature at height  $z_r$  and at the surface  $z_s$ ,  $Q_n$  is the net long wave radiative  
 193 flux (an energy loss at the surface that will be set as a constant),  $G$  is the soil heat flux (an energy  
 194 storage term that will be parameterized as a linear term), and  $H$  is the turbulent sensible heat flux  
 195 (a non-linear energy transport term that will be parameterized in the following).

196 After parameterizing the fluxes, the model has the form:

$$c_v \frac{d\Delta T}{dt} = Q_i - \lambda \Delta T - \rho c_p c_D U \Delta T f(R_b), \quad (2)$$

197 in which  $Q_i$  is the isothermal net radiation,  $\lambda$  is a lumped parameter representing all feedbacks  
 198 from soil heat conduction and radiative cooling as a net linear effect,  $\rho$  is the density of air at  
 199 constant pressure,  $c_p$  is the heat capacity of air at constant pressure,  $c_D = \left(\frac{\kappa}{\ln(z_r/z_0)}\right)^2$  is the neutral  
 200 drag coefficient with  $\kappa \approx 0.4$  the von Kármán constant,  $z_0$  the roughness length and  $z_r$  the reference  
 201 height,  $U$  is the wind speed at height  $z_r$ ,  $R_b = z_r \left(\frac{g}{T_r}\right) \frac{\Delta T}{U^2}$  is the bulk Richardson number, and  $f(R_b)$   
 202 is the stability function used in Monin-Obukhov similarity theory.

203 The lumped parameter  $\lambda$  corresponds to a linear term in the model as the soil is assumed to  
 204 respond linearly to the temperature inversion. Moreover,  $\Delta T \cdot f(R_b)$  is a non-linear term due to  
 205 the non-linear dependence of turbulent diffusion on the vertical temperature gradient.

206  
 207 Following van de Wiel et al. (2017), instead of analyzing the dynamical stability of the  
 208 energy-balance model (2) itself, we will present the linear stability analysis of a simplified system  
 209 that has a similar mathematical structure but is mathematically convenient to analyze. Using a  
 210 cutoff, linear form for the stability function, i.e.  $f(x) = 1 - x$  and  $f(x) = 0$  for  $x > 1$ , the simplified  
 211 model is

$$\frac{dx(t)}{dt} = g(x(t)), \quad \text{where} \quad g(x) = \begin{cases} Q_i - \lambda x - Cx(1-x) & \text{for } x \leq 1 \\ Q_i - \lambda x & \text{for } x > 1 \end{cases} \quad (3)$$

212 and  $x(t_0) = x_0$ . Here, up to dimensional constants,  $x$  represents  $\Delta T$ . The parameter  $C$  will be  
 213 treated as a bifurcation parameter for this simplified system. Similar types of stability functions  
 214 are typically used in numerical weather prediction tools, and the cutoff form facilitates the math-  
 215 ematical analysis of the model. Note that to be consistent with the original model, the stability  
 216 function should include a dependence on both the temperature and the wind speed via  $R_b$ . Re-  
 217 moving this dependence as it is done here changes some of the nonlinearity, however it makes  
 218 the mathematical analysis very simple and the qualitative behavior of the system is similar to the  
 219 original system - see van de Wiel et al. (2017), their Figure 8 and 10. In that sense, the model loses  
 220 some physical significance for mathematical convenience, but the qualitative nonlinear feedback  
 221 processes are maintained. This simplification also has for implication that while  $C$  is related to the  
 222 wind speed, it cannot be directly interpreted as such in the context of the energy balance model.  
 223 For a deeper discussion of the model, its simplifications and the model parameters, the reader is  
 224 referred to its thorough presentation by van de Wiel et al. (2017).

225 For fixed and physically meaningful values of  $Q_i$  and  $\lambda$ , equation (3) can have either one, two,  
226 or three possible equilibrium solutions depending on the fixed values (see illustration in van de  
227 Wiel et al. (2017), Figure 10-12, and related discussion for more details). The equilibrium solu-  
228 tions will be functions of the parameter  $C$ , which we will consider as a bifurcation parameter in  
229 the following. Physically, the case of strong thermal coupling between the surface and the atmo-  
230 sphere, corresponding to a large value of  $\lambda$ , results in one unique equilibrium solution whose value  
231 depends on  $C$ . In van de Wiel et al. (2017), it is hypothesized that such a case is representative  
232 of a grass site such as Cabauw, the Netherlands. The solution is linearly stable to perturbations,  
233 i.e. linear stability analysis shows that perturbed solutions are attracted back to the equilibrium.  
234 The case of no coupling ( $\lambda=0$ ) leads to two equilibrium solutions, one of which is linearly stable  
235 and the other unstable to perturbations (i.e. perturbed solutions are repelled by the equilibrium).  
236 A weak coupling strength, with an intermediate value of  $\lambda$  that could be representative of a snow  
237 surface or another thermally insulated ground surface, results in three possible equilibrium solu-  
238 tions. The two extreme solutions are stable to perturbations, while the middle equilibrium solution  
239 is unstable. Perturbed solutions around the middle equilibrium will thus be attracted either by  
240 the upper or the lower equilibrium. Plotting those three equilibrium solutions as a function of the  
241 bifurcation parameter  $C$  results in a back-folded curve which is qualitatively similar to observa-  
242 tions of the temperature inversion shown as a function of wind speed at Dome C, Antarctica; see  
243 Vignon et al. (2017). The bifurcation diagram is shown in Figure 1 for parameter values such that  
244  $\lambda > 0$  and  $Q_i > \lambda$ , resulting in the case with three possible equilibrium solutions. By convention,  
245 the unstable equilibrium branch is denoted by a dashed line. In the following, we will analyze  
246 transitions between the two stable equilibria. If the system undergoes random perturbations in this  
247 bi-stable context, a perturbation could drive the system sufficiently far from a stable equilibrium  
248 state so that it comes near the unstable equilibrium and finally gets attracted by the second equi-

249 librium. The three possible equilibria are denoted as  $x_{e_1}$ ,  $x_{e_2}$  and  $x_{e_3}$ . In order to study such regime  
 250 transitions induced by random perturbations, the conceptual model is extended with a noise term  
 251 in the following section.

252 *b. Extending the conceptual model by randomization*

253 The conceptual model (3) can be equivalently written in terms of a gradient system, in which the  
 254 temperature inversion represented here by  $x$  evolves according to the influence of an underlying  
 255 potential  $V(x)$ . The randomized model to be introduced will be based on this gradient structure.  
 256 Specifically, the initial value problem (3) can be written as

$$\frac{dx}{dt} = -\frac{dV}{dx}, \quad x(t_0) = x_0,$$

257 where it is easy to see that the potential is given by

$$V(x) = \begin{cases} \frac{1}{2}x^2(\lambda + C) - \frac{C}{3}x^3 - Q_i x & \text{for } x \leq 1, \\ \frac{1}{2}\lambda x^2 - Q_i x + \frac{1}{6}C & \text{for } x > 1. \end{cases} \quad (4)$$

258 The linear stability analysis discussed in the previous section can thus be understood in the sense  
 259 that the temperature inversion  $x$  equilibrates at a local minimum of a potential  $V$ . That is, an  
 260 equilibrium point  $x_e$  satisfies  $V'(x_e) = 0$ . Figure 2 sketches the form of the potential with the  
 261 exemplary parameter values  $\lambda = 2, Q_i = 2.5, C = 6.4$ . Note that  $V(x)$  is a double-well potential in  
 262 that case where each local minimum corresponds to one of the stable equilibrium points  $x_{e_1}$  and  
 263  $x_{e_3}$ , while the local maximum corresponds to the system's unstable equilibrium  $x_{e_2}$ .

264 While the conceptual model (3) has proven very insightful to explain observed sharp transitions  
 265 in temperature inversions, it only allows for regime transitions when drastic changes in the model  
 266 *parameters* (i.e., bifurcations) occur. That is, the model is overly idealized and in reality one can  
 267 expect regime transitions to also take place due to small natural fluctuations of the temperature

268 inversion itself in certain cases, e.g., when the potential barrier separating the two local minima and  
269 corresponding stable equilibria is shallow. Therefore we will consider an appropriate randomized  
270 variant of the model. Specifically, we consider the stochastic differential equation (SDE) model

$$dx = -\frac{dV(x)}{dx} dt + \sigma dB, \quad x(t_0) = x_0, \quad (5)$$

271 to account for small random perturbations to the temperature inversion's rate of change. Here,  $B$   
272 denotes a standard Brownian motion (i.e., a stochastic process) and  $\sigma > 0$  scales the intensity of the  
273 fluctuations, while the potential  $V$  is as in (4). As the randomized dynamics is characterized by the  
274 same potential, also the equilibrium points of the non-random model (3) will describe the dominant  
275 effects of the randomized model's dynamics. However, due to the presence of the noise, the stable  
276 equilibria of the non-random model (3) are not limiting points for the stochastic counter-part in (5),  
277 in the sense that the temperature inversion may still fluctuate after reaching a stable equilibrium.  
278 The reason is that in a context of two stable equilibria (i.e. for parameter values such that the  
279 model (3) exhibits two stable equilibria, denoted earlier as the thermally weakly-coupled state), the  
280 random perturbations can trigger transitions from one stable equilibrium to another one. We will  
281 therefore refer to the formerly stable states as: metastable. Note that depending on the coupling  
282 strength and noise intensity, the likelihood of regime transitions can change drastically and the  
283 system may or may not exhibit metastable states. The type of noise (additive or multiplicative for  
284 example, or noise with a Levy distribution) will also affect regime transitions. In our subsequent  
285 simulations and analyses, we will focus on the case of two metastable states with additive noise  
286 and leave other cases for future research.

287 The effect of these random perturbations to a metastable equilibrium point  $x_e$  can be understood  
288 through a localized approximation of the original dynamics. More precisely, consider a second-  
289 order Taylor approximation of the potential around an equilibrium point  $x_e$ , yielding the quadratic

290 approximate potential  $\tilde{V}$ :

$$V(x) \approx \tilde{V}(x) := V(x_e) + \frac{1}{2} \frac{d^2V}{dx^2}(x) \Big|_{x=x_e} (x - x_e)^2 .$$

291 For the same parameter values that were used to plot the original potential in Figure 2, the red line  
 292 in the same figure shows the approximate quadratic potential around the equilibrium value  $x_{e_1}$ .  
 293 Using the locally quadratic potential, we can thus define a locally approximate dynamics for the  
 294 temperature inversion by replacing  $V$  in (5) by  $\tilde{V}$ , resulting in

$$dX = -k(X - x_e) dt + \sigma dB , \quad X(t_0) = x_e ,$$

295 where  $k := \frac{d^2V}{dx^2}(x)|_{x=x_e} \in \mathbb{R}$  and  $X$  is introduced to describe the approximate dynamics of the  
 296 former  $x$ . This approximate dynamics is an example of the well-studied Ornstein–Uhlenbeck  
 297 process and it provides an accurate description of the full dynamics in the neighborhood to the  
 298 equilibrium point  $x_e$ . Discretizing the Ornstein–Uhlenbeck process  $X$  using the Euler–Maruyama  
 299 scheme with a step-size  $\Delta t := \frac{T}{L}$  for some positive integer  $L$  we furthermore find that the process  
 300 at discrete times  $t \in \{1, \dots, L\}$  approximately satisfies the difference equation

$$X_t = X_{t-1} - k(X_{t-1} - x_e)\Delta t + \sigma \left( B(t\Delta t) - B((t-1)\Delta t) \right) , \quad X_0 = x_e ,$$

301 in the sense that  $X_t \approx X(t\Delta t)$ . By defining  $\mu := kx_e \in \mathbb{R}$ ,  $\phi := (1 - k\Delta t)$  and  $w_t := \sigma [B(t\Delta t) -$   
 302  $B((t-1)\Delta t)]$  this can be written as

$$X_t = \mu + \phi X_{t-1} + w_t ,$$

303 which is a so-called autoregressive model of order one, denoted AR(1), thanks to the properties  
 304 of the (scaled) Brownian increments  $w_t$ . Consequently, we see that the discretized Ornstein–  
 305 Uhlenbeck process can be accurately approximated by an AR(1) process. This derivation can also  
 306 be found in Thomson (1987). Combining this with the observation that the Ornstein–Uhlenbeck

307 process offered an accurate approximation to the original dynamics in the vicinity of a stable  
 308 equilibrium, we can thus conclude that the local dynamics in the neighborhood of a metastable  
 309 state can be approximately described by an AR(1) process.

310 *c. Statistical indicator for the dynamical stability of time series*

311 In section a we discussed the simplified model by van de Wiel et al. (2017) which was developed  
 312 to understand regime transitions in near-surface temperature inversions. This model provides a  
 313 hypothesis that explains the existence of two possible equilibria of the temperature inversion for a  
 314 given wind speed. In agreement with the randomized conceptual model introduced in the previous  
 315 section, we say that a system exhibiting at least two metastable equilibria is called metastable.  
 316 In this section the goal is to describe a methodology for statistically detecting critical transitions  
 317 based on time series data. For the detection we apply an indicator for the dynamical stability  
 318 (i.e. the resilience to perturbations) of time series, which was defined by Faranda et al. (2014) and  
 319 applied to SBL turbulence data in Nevo et al. (2017). The indicator uses a combination of methods  
 320 from dynamical systems and from statistical modeling. In its definition, deviations from AR(1)  
 321 processes in the space of so-called autoregressive-moving-average (ARMA) models are used to  
 322 quantify the dynamical stability of a time series. A time series  $x_t, t \in \mathbb{Z}$ , is an ARMA(p,q) process  
 323 if it is stationary and can be written as

$$x_t = \nu + \sum_{i=1}^p \phi_i x_{t-i} + w_t + \sum_{j=1}^q \theta_j w_{t-j}, \quad (6)$$

324 with constant  $\nu$ , coefficients  $\phi_p, \theta_q$  and  $\{w_t\}$  being white noise with positive variance  $\sigma^2$ . The co-  
 325 efficients  $\phi_p$  and  $\theta_q$  additionally have to satisfy some constraints, see Brockwell and Davis (2016).  
 326 Notice that AR(1) = ARMA(1,0). Intuitively the parameters  $p$  and  $q$  are related to the memory lag  
 327 of the process. The longer the system takes to return to the equilibrium after a perturbation, the

328 more memory we expect to observe in the process. Examples of simple systems along with their  
329 ARMA( $p,q$ ) characteristics can be found in Faranda et al. (2014).

330 In section b, it was shown that the dynamics when the system is close to a stable equilibrium  
331 can be approximated by a AR(1) process. We will assume that far from the transition from one  
332 dynamical regime to another, the time series of a generic physical observable can be described by  
333 an ARMA( $p,q$ ) model with a reasonably low number of  $p,q$  parameters and coefficients. Indeed  
334 far from a transition, the system will tend to remain around an equilibrium despite random pertur-  
335 bations, and excursions from the equilibrium are short. The idea behind the modeling assumption  
336 is that ARMA processes are an important parametric family of stationary time series (Brockwell  
337 and Davis 2016). Their importance is due to their flexibility and their capacity to describe many  
338 features of stationary time series. Thereby, choosing ARMA( $p,q$ ) processes for modeling the dy-  
339 namics away from a stable state is a reasonable Ansatz. Close to a transition, the resilience of the  
340 system to perturbations is weak and longer excursions from the equilibrium occur. The statistical  
341 properties (such as the shape and/or the persistence of the autocorrelation function) of the system  
342 change drastically, leading to an expected increase of the value  $p + q$  (Faranda et al. 2014). Based  
343 on this, we use ARMA( $p,q$ ) models in the following to analyze the stability of a dynamical system.  
344 The dynamical stability indicator which will be defined next will be used to obtain indicators for  
345 detecting the system's proximity to a regime transition.

346 In order to quantify the local stability of a time series, we first slice the time series  $x_t$   
347 with a moving time window of fixed length  $\tau$ . In other words, we obtain subsequences  
348  $\{x_1, \dots, x_\tau\}, \{x_2, \dots, x_{\tau+1}\}, \dots, \{x_{t-\tau+1}, \dots, x_t\}$  of the original time series that overlap. By slic-  
349 ing the original time series we obtain a sequence of shorter time series for which it is reasonable  
350 to suppose that they are amenable to ARMA modeling. In detail, we assume that the subse-  
351 quences are realizations of linear processes with Gaussian white noise which then implies that the

352 process is stationary. We then fit an ARMA( $p,q$ ) model for every possible value of  $(p,q)$ , with  
 353  $p \leq p_{max}$  and  $q \leq q_{max}$ , to these subsequences, where  $p_{max}$  and  $q_{max}$  are predefined thresholds.  
 354 Afterwards we choose the best fitting ARMA( $p,q$ ) model by choosing the one with the minimal  
 355 Bayesian information criterion, BIC, (Schwarz 1978) which is a commonly used and well-studied  
 356 tool in statistical model selection. Assuming that we have the maximum likelihood estimator  
 357  $\hat{\beta} := (\hat{\nu}, \hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)$  of the fitted ARMA( $p,q$ ) model (which can be obtained using a  
 358 so-called innovation algorithm, as it is, for example, implemented in the "forecast" R package  
 359 (Hyndman et al. 2019) which is used for the analyses), the BIC is formally defined as

$$\text{BIC}(p,q) = -2 \ln L(\hat{\beta}) + \ln(\tau)(p+q+1), \quad (7)$$

360 where  $L(\hat{\beta})$  denotes the associated likelihood function evaluated at the maximum likelihood esti-  
 361 mator  $\hat{\beta}$ . The second term introduces a penalty for high-order models (i.e., those that contain more  
 362 parameters) to avoid over-fitting.

363 We reiterate that when the system is close to a metastable state, its dynamics can be well ap-  
 364 proximated by an AR(1) process. When the system is approaching an unstable point separating  
 365 multiple basins of attraction, the approximation no longer holds as the underlying potential can-  
 366 not be approximated by a quadratic potential anymore. The change in the shape of the potential  
 367 introduces new correlations in the time series, resulting in higher-order ARMA terms when fitting  
 368 such a model to data.

369 The definition of the stability indicator is based on this observation, in the sense that it assumes  
 370 that the dynamics near a metastable state can be modeled by an ARMA(1,0) or AR(1)-process.  
 371 Specifically, the stability indicator is defined as

$$\Upsilon(p,q;\tau) = 1 - \exp\left(\frac{-|\text{BIC}(p,q) - \text{BIC}(1,0)|}{\tau}\right). \quad (8)$$

372 For a stable state, the most likely statistical model is an AR(1) process and one expects that  $\Upsilon = 0$ .  
373 The indicator  $\Upsilon$  gives a normalized distance between the stable state ( $\Upsilon = 0$ ) and the state in  
374 which the system is. The limit  $\Upsilon \rightarrow 1$  corresponds to a most likely statistical model of high  
375 order and probably to a nearing transition. While a formal proof of this statement is still missing,  
376 tests performed for systems of increasing complexity in Nevo et al. (2017) showed promising  
377 results where the indicator correctly identified changes in the stability of metastable states. Note  
378 that the character of the noise present in the physical system (additive noise, multiplicative, Levy  
379 process...) will affect the ARMA model approximation and impact the values of  $\Upsilon$ . To simplify  
380 the notation we drop the dependence of  $\Upsilon$  on  $p, q$  and  $\tau$  in the following discussion.

381 The reliability of  $\Upsilon$  highly depends on the choice of  $\tau$ , the window length (which we will  
382 consider in number of discrete observations in the following), and it relates to the characteris-  
383 tic timescales of the dynamics. Intuitively, the window length, when converted to its equivalent  
384 physical duration (i.e. the number of discrete observations multiplied by the discrete sampling  
385 time), has to be shorter than the residence timescale in one basin of attraction (i.e. the time spent  
386 in the neighborhood of an equilibrium before transitioning to another one) in order to satisfy the  
387 local stationarity, but large enough so that statistical model estimation is reliable. In winter at  
388 Dome C where the Polar winter results in a near absence of daily cycle, no preferred timescale of  
389 residence around an equilibrium of the temperature inversion was observed (an equilibrium can  
390 remain for several days), however the transition between two equilibria was observed to take place  
391 over a timescale of the order of 10 hours (Baas et al. 2019). For nocturnal flows, the residence  
392 timescale is tightly connected to the daily cycle and could be of a few hours during the night,  
393 or the entire night. The transition between two equilibria typically takes place over a duration of  
394 about a half hour. For reliable statistical estimation, multiple tests showed that a minimum window

length of 20 discrete points is needed. With a sampling time of 1 minute, that means that a moving window of approximately 20 or 30 minutes may be appropriate.

In addition, the sampling frequency has to be fine enough to sample typical fluctuations of the dynamics. In the following analyses, we find a sampling frequency of one minute to be appropriate for that purpose. The characteristic timescale here is given by the timescale at which the system recovers from perturbations (which is estimated by linear stability analysis in the case where the model is known, see e.g. (van de Wiel et al. 2017)), and the time interval between two observations should be close to or smaller than this quantity so that (small-scale) local fluctuations can be resolved. Since the characteristic timescales of the system cannot be known analytically in many situations, for example when analyzing time series from atmospheric models or from field data, we suggest two data-driven approaches to select a window length:

- In the first approach, the mean residence time around each metastable state as well as the mean transition time between the two states will be estimated based on a *data clustering approach*. The observations will be clustered in the metastable regimes and an intermediate, transition regime. From the clustered data, the mean residence time in each cluster will be evaluated. This approach will provide an upper bound to select the window length.
- The second approach is motivated by the fact that the indicator  $\Upsilon$  is obtained through a statistical inference procedure through the definition of the BIC which involves fitting suitable ARMA processes to data. Specifically, a maximum likelihood approach is used, which assumes that subsequences are sampled from a normal distribution. To assess the validity of this *statistical approach*, a normality test will be implemented as a criterion to select a window length for which the normality assumption is justified and ARMA model estimation is reliable.

418 Both approaches are applicable when the data-generating model is unknown. This is important in  
419 cases where data showing signs of metastability are available, but an underlying model is unknown.  
420 A summary of the full algorithmic procedure used to calculate the statistical indicator is given in  
421 Appendix A3.

#### 422 1) CLUSTERING APPROACH: K-MEANS

423 In the first approach, we suggest to use the K-means algorithm (Hartigan and Wong (1979),  
424 see pseudo code in Appendix A1) to select a window length for the analysis. In the context of  
425 analyzing transitions in the temperature inversion, the idea is to cluster the data into three different  
426 clusters: data around each stable fixed point and data near the unstable fixed point (in other words,  
427 data covering the transition periods between two metastable states). By that, the goal is to estimate  
428 the average time needed by the system to transition between two metastable states. The mean  
429 residence time within each cluster is calculated from the time series of cluster affiliation. We  
430 choose  $\tau$  (recall that we consider it in number of discrete observations and not in physical time)  
431 such that it is smaller than the minimal mean time spent in one cluster, which should ensure that  
432 subsequences remain mostly around one equilibrium. This value is denoted by  $\tau_{Kmeans}$ . For the  
433 simulated data in the following, each simulated time series will be assigned a window length  
434  $\tau_{Kmeans}$  by this procedure. For the nocturnal dataset, we cluster the entire dataset once and obtain  
435 a length  $\tau_{Kmeans}$  of 22 points, corresponding to a duration of 22 minutes. For the Polar dataset,  
436 only one continuous time series during a Polar winter will be considered and assigned one value  
437 of  $\tau_{Kmeans}$ , namely 10 points, corresponding to a duration of 100 minutes. This window length is  
438 insufficient to obtain reliable statistical estimations.

439 Note that this clustering approach to determining a residence timescale around an equilibrium is  
440 a crude approximation and suffers from many caveats: a high density of data close to a given value

441 of the temperature inversion may not necessarily relate to the existence of metastable equilibria,  
442 but could occur due to non-stationary dynamics or complex nonlinear effects, for example. Never-  
443 theless, we use it as a first approach and future research may result in more reliable approaches. In  
444 the following analyses, the K-means procedure can be interpreted as providing an upper bound for  
445 selecting a window length for the analysis and thus, combined with the following criterion, will  
446 offer an applicability criterion for our method.

## 447 2) STATISTICAL APPROACH: ANDERSON-DARLING NORMALITY TEST

448 The K-means clustering approach described above estimates the system's physical timescales,  
449 but the statistical properties of the process should also be considered for reliable calculations. To  
450 fit ARMA models reliably and to calculate the Bayesian information criterion for ARMA model  
451 selection, we need the underlying process to follow a normal distribution. Note that the reliability  
452 of ARMA model fitting generally increases for increasing number of data points (assuming sta-  
453 tionarity remains fulfilled). In this approach, we suppose that the subsequences are sampled from  
454 a normal distribution, at least for some window length  $\tau$ . We then choose  $\tau$  as the biggest window  
455 length such that this normality assumption holds (more precisely, such that the normality hypoth-  
456 esis cannot be rejected). This value is denoted by  $\tau_{AD}$ . If we find that  $\tau_{AD}$  has to be much larger  
457 than  $\tau_{Kmeans}$  to fulfil the normality assumption, we will interpret this as a sign of undersampling  
458 in the data.

459 Specifically, the statistical test results in a p-value for each subsequence, and we choose the  
460 window length such that the median of the p-values of all subsequences is above a threshold for  
461 which the null-hypothesis cannot be rejected. The normality test applied here is the Anderson-  
462 Darling Test (Anderson and Darling 1952), abbreviated AD test, as it is, for example, more stable  
463 than the Kolmogorov–Smirnov test (Stephens 1974). Further details details of the AD test are

464 summarized in Appendix A2. Similarly to the clustering approach, the window length  $\tau_{AD}$  for  
 465 which normality cannot be rejected is selected for each analyzed time series, and this value of  $\tau_{AD}$   
 466 is then used for all subsequences of the time series. Each of the simulated dataset, the nocturnal  
 467 dataset and the Polar dataset will be assigned a single value of  $\tau_{AD}$ . The value of  $\tau_{AD}$  for the  
 468 nocturnal dataset is 19 discrete points, hence 19 minutes with a sampling frequency of one minute.  
 469 For the Polar dataset, the value of  $\tau_{AD}$  is 43 points corresponding to a duration of 430 minutes.

### 470 **3. Stability analysis of simulated and observed time series**

471 In this section we quantify the reliability of the stability indicator introduced in section 2.c.  
 472 We start by testing it on a controlled dataset generated by the simplified model for near-surface  
 473 temperature inversion (see section 2.a) and then proceed by applying  $\Upsilon$ , the stability indicator, to  
 474 observational data. In the tests we use the `auto.arima()` function from the "forecast" R package  
 475 (Hyndman et al. 2019). The `auto.arima()` function fits ARMA(p,q) models by calculating the  
 476 maximum likelihood estimators for a given model order (using the innovation algorithm mentioned  
 477 earlier). It calculates the corresponding BIC (using the definition (7)) for all ARMA(p,q) models  
 478 with  $p \leq p_{max}$  and  $q \leq q_{max}$ , where  $p_{max}$  and  $q_{max}$  are thresholds set to 10 in our application, and  
 479 then it chooses the ARMA model with the minimal BIC value. This procedure is repeated for each  
 480 subsequence of data, using the moving window approach, and the minimal BIC value leads to the  
 481 optimal ARMA(p,q) model to represent the given subsequence.

#### 482 *a. Simulated time series*

483 To generate the simulated data, we use the conceptual randomized model (5), which we recall  
 484 here for the reader's convenience:

$$dx(t) = -\frac{dV(x(t))}{dx}dt + \sigma dB(t), \quad x(t_0) = x, \quad 0 \leq t \leq T,$$

485 where  $V(x)$  is the energy potential defined in (4). That is, the data-generating model reads

$$dx = \begin{cases} (Q_i - \lambda x - Cx(1-x))dt + \sigma dB, & x \leq 1, \\ (Q_i - \lambda x)dt + \sigma dB, & x > 1. \end{cases} \quad (9)$$

486 The SDE model (9) is approximated path-wise (i.e., for each realization of the driving Brownian  
487 path) using the Euler-Maruyama scheme.

488 For the purpose of testing the accuracy of the  $\Upsilon$  indicator and its potential to detect nearing  
489 regime transitions, one realization  $\{x_t\}$  of the stochastic process is used for each fixed value of  
490 the bifurcation parameters  $C$ . Multiple fixed values of  $C$  are used, resulting in one timeseries per  
491 value of  $C$ . The initial parameters are set to  $t_0 = 0$  and  $x(t_0) = \min\{x_{e_i} | i = 1, 2, 3\}$  where  $x_{e_i}$  are the  
492 three equilibria of the system. To generate the controlled data set the model parameters are set to  
493  $\lambda = 2$ ,  $Q_i = 2.5$  and  $\sigma = 0.35$ . The value of  $C$  is varied between  $C = 5.3$  and  $C = 7.2$  with discrete  
494 increments of 0.1 and one simulation is done per value of  $C$ . The simulations are ran for  $n = 2000$   
495 time steps of size  $\Delta t = 0.01$ . The amplitude of the noise, or diffusion coefficient,  $\sigma = 0.35$  is  
496 chosen as it resulted in trajectories for which regime transition could be observed on the time  
497 interval  $[0, T = n\Delta t = 20]$ . The range for  $C$  is chosen because for these values the time series  
498 shows frequent transitions from one metastable state to another. To choose the window length  
499  $\tau$  we apply both the K-means Algorithm (section 2.c.1) and the Anderson Darling Test (section  
500 2.c.2). The length  $\tau$  is determined individually for each simulation, i.e. for each fixed value of  
501  $C$ . The K-means algorithm can be used to estimate the amount of discrete observation points  
502 covering the transition time. We set the cluster number to three as we expect three equilibria.  
503 The results of the clustering algorithm are exemplary shown in figure 3 for  $C = 6.4$ . Note that  
504  $t \in [0, n \cdot \Delta t] = [0, 20]$ , whereas we will express our window length  $\tau$  in number of discrete points  
505 in the following. In this case the equilibria are  $x_{e_1} = 0.46$  (metastable),  $x_{e_2} = 0.97$  (unstable),

506 and  $x_{e_3} = 1.25$  (metastable). The cluster centers, estimated by the K-means algorithm, are 0.46,  
 507 0.97, and 1.31 which are a close approximation of the equilibria. Therefore, we expect a good  
 508 estimation for the amount of points covering the transition. The average time spend in each cluster  
 509 are (for  $C = 6.4$ ):  $mean(T_1) = 112.2$ ,  $mean(T_2) = 94$ , and  $mean(T_3) = 286.67$ , where  $mean(T_i)$  is  
 510 the average time spent without observed transitions in cluster  $i \in \{1, 2, 3\}$  expressed in number of  
 511 discrete points. The minimal mean residence time is thus  $mean(T_2) = 94$  and provides an upper  
 512 bound to select a window length that respects the timescales of the system. The window length  $\tau$   
 513 is thus chosen such that it is smaller than the minimal average time spent in one cluster, i.e. for  
 514  $C = 6.4$  we choose  $\tau < \tau_{KMeans} := \min\{mean(T_i) | i = 1, 2, 3\} = 94$ . For all tested  $C$ , we choose  
 515  $\tau = \min\{mean(T_i) | i = 1, 2, 3\} - 5$  in order to give room for some uncertainty in the evaluation of  
 516 the time spent in each cluster, due to potential overlaps of the clusters (we recall that the minimal  
 517 mean residence time should be understood as an upper bound to select  $\tau$ ). By applying  $\Upsilon$  to the  
 518 data generated by the simplified model with  $C = 5.3$ ,  $C = 6.4$  and  $C = 7.2$  we get the results shown  
 519 in figure 4. The solid red lines correspond to the stable equilibria and the dotted red line to the  
 520 unstable one. The colors ranging from dark blue to yellow represent the stability of the points  
 521 measured by  $\Upsilon$  and we always color the last point of the subsequence. The simulation is initialized  
 522 around the stable equilibrium  $x_{e_1}$ , where short memory of the random perturbations should prevail.  
 523 As expected, the values of  $\Upsilon$  remains close to 0 (corresponding to a most likely AR(1) model) as  
 524 long as the simulation oscillates around the equilibrium. The timeseries eventually approaches  
 525 the neighborhood of the unstable equilibrium where long memory properties are to be expected  
 526 and thus higher order ARMA(p,q) models, hence larger values of  $\Upsilon$ , are more likely. This first  
 527 transition through the unstable equilibrium is well recognized with higher values of  $\Upsilon$  (green dots  
 528 after the dotted red line).

529 The Anderson Darling Test can be used to find the biggest  $\tau$  for which we can assume that most  
530 of the subsequences are sampled from a normal distribution and hence trust the ARMA model  
531 selection and fitting. As shown in figure 5, for  $C = 6.4$  the Anderson Darling Test yields that for  
532  $\tau = 67$  the median of the p-values for all subsequences is greater than the significance level 0.05.  
533 The solid line in the gray boxes is the median of the p-values for a fixed  $\tau$  while the upper and  
534 lower border of the gray boxes refer to the upper and lower quartile of the p-values. The dotted  
535 horizontal line is the significance level. We report that the values of  $\tau_{AD}$  given by the Anderson  
536 Darling Test are ranging from 60 to 70 discrete points for all values for  $C$ .

537 Figure 6 summarizes the  $\Upsilon$  values obtained for different choices of  $\tau$  and different values of  $C$  in  
538 a bifurcation diagram. In the figure, the equilibrium solutions of the deterministic equation (3) are  
539 shown by a red line for the considered range of values of  $C$ . This is the same diagram as shown  
540 in Fig. 1, where the upper and lower branches of the equilibrium solution correspond to the two  
541 stable equilibria, while the middle one is the unstable equilibrium separating the two basins of  
542 attraction of the stable equilibria. A discontinuity in the solution is visible between the upper and  
543 middle solution branches, which is due to the discontinuity introduced by the cutoff form of the  
544 stability function. The  $\Upsilon$  values obtained for the simulations of the stochastic system (eq. (9)) for  
545 all considered values of  $C$  are then shown as a scatter plot along with the equilibrium solution, and  
546 the darker color corresponds to higher values of  $\Upsilon$ . As the initial condition for all simulations is  
547 taken at the lowest equilibrium values, the transitions are expected to occur between the lower and  
548 upper equilibrium branches when the system transitions from the basin of attraction of the lowest  
549 equilibrium value to that of the highest value. High values of  $\Upsilon$  are indeed mainly found in this  
550 region of the diagram.

551 Three methods are used to select the value of  $\tau$  and the results associated with these window  
552 sizes are shown in figure 6. In figure 6 panel a,  $\tau = \tau_{KMeans} - 5$  is used. Around the stable branches,

553 values of  $\Upsilon$  are small, denoting that stable states are detected as such. Large values of  $\Upsilon$  are found  
554 between the unstable branch and the upper stable branch of the bifurcation diagram, indicating  
555 that transitions from the lower to the upper stable branches are detected by the indicator. The  
556 fact that the high values are not exactly located around the unstable branch is due to the use of a  
557 finite window size for the calculation: in the diagram, the color is always assigned to the last point  
558 of the subsequence. For small values of  $C$ , e.g. for  $C = 5.4$ , large values of  $\Upsilon$  are occasionally  
559 inappropriately found around the upper stable branch. For small  $C$ , the potential well will be  
560 relatively steep and the system rapidly approaches the second equilibrium, so that the detection  
561 can be too slow. Figure 6 panel b shows the results for  $\Upsilon$  when choosing  $\tau$  according to the  
562 Anderson Darling test, denoted as  $\tau_{AD}$ . The figure is very similar to the one using  $\tau_{KMeans}$  except  
563 that for  $C \leq 5.6$  there are more high values of  $\Upsilon$  located around the stable branch. This is due to the  
564 fact that for these  $C$ 's the  $\tau$ 's chosen by the Anderson Darling test are larger than the ones estimated  
565 by the K-means algorithm. Consequently, the local stationarity assumption may break down. For  
566  $C \geq 5.9$  the  $\tau$ 's given by the AD Test are smaller than the ones of the K-means algorithm. In these  
567 cases  $\Upsilon$  gives a good indication for the stability. Figure 6 panel c is a bifurcation plot showing only  
568 the time series for which  $\tau$  can be chosen to satisfy both the K-means and the Anderson Darling  
569 condition, i.e.  $\tau_{AD} < \tau_{KMeans} - 5$ . Here  $\tau := \tau_{AD}$  is used for the analysis and timeseries that do  
570 not satisfy the condition are discarded from the analysis. In this case, we see that large values  
571 of  $\Upsilon$  always occur between the unstable branch and the upper stable branch, thus  $\Upsilon$  is capable of  
572 recognizing the location of unstable equilibria for all  $C$  and stable equilibria are never assigned a  
573 large value of  $\Upsilon$ . A note of caution should however be given regarding the reverse transitions from  
574 the upper stable branch to the lower stable branch in those numerical examples. Indeed, those are  
575 poorly identified. This is probably related to the asymmetry of the underlying potential, which  
576 induces different characteristic timescales in the system and hence a need to adapt the value of

577  $\tau$  locally, and not just globally for all types of transitions as it is done here. To overcome this  
578 difficulty, an adaptive tuning of  $\tau$  would be required, which will be left for future research. With  
579 these possible limitations in mind,  $\Upsilon$  will next be applied to observational data.

580 As a final remark, the values of  $\tau$  considered here can be compared to analytical results in  
581 this numerical example. Indeed, for this simple bistable example system, analytical results can  
582 provide the expected time taken by the system to transition from one of the local equilibria to  
583 the bifurcation point (Krumscheid et al. 2015) and can serve as a comparison to the statistical  
584 estimations of  $\tau$  obtained here. As a matter of fact, for  $C$  close to the bifurcation point, the results  
585 given by the Anderson Darling test are similar to those given by analytical calculations.

#### 586 *b. Analysis of regime transition in observed nocturnal and Polar temperature inversions*

587 In this section we apply the stability indicator  $\Upsilon$  on observational data obtained from one site  
588 near Dumosa, Australia for which nocturnal data are selected, and from Dome C, Antarctica for  
589 which we consider the Polar winter. When we plot  $\Delta T$  over  $U$  for both sites (see figure 7) we see  
590 a clear sign of two distinct states: one when the wind is weak and  $\Delta T$  is large and one for strong  
591 wind where  $\Delta T$  is small.

#### 592 1) DUMOSA

593 The first observational dataset consists of temperature measurements from a site near Dumosa,  
594 Victoria, Australia. The site was located in a large area with mostly homogeneous and flat terrain,  
595 covered by wheat crops, and measurements were taken during the crop season. The temperature  
596 measurements were made on the main tower at heights of 3 and 6m and the wind measurements at  
597 6m. The frequency of measurements is 1 minute. Further details about the observational site can  
598 be found in Lang et al. (2018). As we want to use data where we can expect temperature inversions

599 to take place we exclusively use evening and nighttime data from March until June 2013 (89 days).  
600 Each night of data results in a timeseries of 1020 discrete observations. Similarly to the simulated  
601 data, we use the K-means algorithm and the Anderson Darling Test to choose the window length  
602  $\tau$ . The results are shown for all nights considered together in Figure 8. According to these tests  
603 the maximal  $\tau$  for which we can assume normality ( $\tau_{AD}$ ) is 19 discrete observations (hence 19  
604 minutes with the sampling frequency of 1 minute) and the  $\tau$  which corresponds to the minimal  
605 mean residence time in one of three clusters ( $\tau_{KMeans}$ ) is 22. Hence we have  $\tau_{AD} < \tau_{KMeans}$  and  
606 choosing  $\tau_{AD}$  should be appropriate. The results for  $\Upsilon$  applied to all 89 nights with both choices of  
607 window lengths are given in figure 9. The results highlight a lower branch with low values of  $\Upsilon$ , or  
608 dynamics identified as stable, and an upper branch with high values of  $\Upsilon$ , or dynamics identified  
609 as unstable. In some cases, a proper ARMA(p,q) model cannot be fitted by the statistical methods,  
610 resulting in absence of results for some windows. Generally, a reliable ARMA(p,q) fit becomes  
611 difficult for a time series with less than 20 observations, and the estimated window lengths are  
612 on the lowest end to obtain the statistical estimations. Figure 10 shows the time evolution of  $\Delta T$   
613 when conditionally averaged for all nights with the wind speed (wsp) being in a given category.  
614 The corresponding time evolution of  $\Upsilon$  is shown for the same conditional averages. The window  
615 length here is chosen as the most restrictive criterion  $\tau = \tau_{AD} < \tau_{KMeans}$ . For low wind speeds,  
616 the  $\Upsilon$  values are high (on average), which implies that in this case we have an unstable system.  
617 Note that in this dataset, a leveling-off of the temperature inversion for low wind speeds (which  
618 could correspond to the stable equilibrium of a strong inversion according to the model of van de  
619 Wiel et al. (2017)) is not very evident from Figure 9. It could be that the temperature inversion  
620 does not have time to reach the stable equilibrium during the night, or that other instabilities which  
621 are not considered in the simplified model arise in strong stability conditions. For example, flow

622 instabilities such as submeso motions, which are favored in strongly stratified situations, could  
623 make the system dynamically unstable.

## 624 2) ANTARCTICA

625 The Dome C data was measured at the Concordia Research Station which is located on the  
626 Antarctica Plateau. It is a French-Italian research facility that was built 3233m above sea level.  
627 It is extensively described for example in Genthon et al. (2010). The Dome C dataset contains  
628 10-min averaged meteorological data from 2017. Regimes and their transitions were analyzed by  
629 Vignon et al. (2017) and Baas et al. (2019). Important for our analysis are measurements of the  
630 temperature at height 9.4 m and surface, the wind speed (m/s) at height 8 m and the radiation  
631 made in the polar night which is from March to September. We focus on the polar night during  
632 which multiple regime transitions take place. Following van de Wiel et al. (2017) the data is  
633 classified into two subcategories of radiative forcing being the sum of net shortwave and incoming  
634 longwave radiation:  $R_+ = K^\downarrow - K^\uparrow + L^\downarrow$ . Strong cooling is favored in cases of low incoming  
635 radiation and when plotting  $\Delta T = T_{9.4m} - T_s$  over the wind speed  $U_{8m}$  a back-folding of the points  
636 becomes apparent when  $R^+ < 80Wm^{-2}$  (van de Wiel et al. (2017), their Figure 6 and less clearly  
637 in our Figure 7). Therefore, we focus on the case when  $R^+ < 80Wm^{-2}$ . We apply  $\Upsilon$  to the longest  
638 consecutive time series with  $R^+ < 80Wm^{-2}$  which is from 2017-08-03 10:50 to 2017-08-24 21:50,  
639 i.e. 3091 data points. Our analyzed time series is thus shorter than the one visualized in van de  
640 Wiel et al. (2017), which explains the differences in the scatter plot. Again we choose  $\tau$  with the  
641 Anderson Darling Test and the K-Means algorithm. The value for  $\tau$  given by the Anderson Darling  
642 Test  $\tau_{AD} = 43$  observations (with a corresponding window duration of 430 minutes) is much larger  
643 than the one given by the K-means algorithm ( $\tau_{KMeans} = 10$  observations, corresponding to a  
644 window duration of 100 minutes), which is in fact too few points to expect a good fit for the

645 ARMA(p,q) model. In figure 11 we see that no transitions are recognized by  $\Upsilon$  with  $\tau = \tau_{AD}$  (left)  
646 but with  $\tau = \tau_{KMeans}$  (right) some transitions are noted. This indicates that the data frequency of  
647 this data set is not high enough to give reliable results for the stability indicator.

648 *c. Sensitivity analysis on averaged data*

649 As discussed when applying the indicator  $\Upsilon$  on the Dome C dataset, there is strong indication  
650 for the data frequency being crucial for the reliability of the results when applying the stability  
651 indicator. Observational data is often stored in block averages, e.g. measurements over 5 minute  
652 time window are averaged into 1 data point. The issue with this can be that the data frequency can  
653 be too low to sample typical fluctuations during the observed transition. In more detail, if the time  
654 taken by the system to transition from one metastable state to the next is less than approximately  
655 20 discrete measurement points (the minimum needed to have relevant statistical results according  
656 to our tests), then the approach may not be applicable. Therefore, data frequency needs to be high  
657 enough to give reliable results for  $\Upsilon$ . As a comparative study to illustrate this point, we block  
658 average the temperature measurements for the Dumosa data, such that we repeat the analysis  
659 based on 5-min averaged data instead of 1-min averages. Thereby, we reduce the length of the  
660 time series for each individual night from 1020 to 204 data points. Again we choose  $\tau$  with the  
661 Anderson Darling Test and the K-means algorithm. There is a clear distinction between the  $\tau$   
662 estimated by the Anderson Darling Test and the one given by the K-means algorithm for the 5  
663 min data, contrary to the 1 min data where both methods suggested comparable window lengths.  
664 Indeed,  $\tau_{AD} = 31$  and  $\tau_{KMeans} = 7$  for the 5 min averaged data whereas  $\tau_{AD} = 19$  and  $\tau_{KMeans} = 22$   
665 for the 1 min data. The small value for  $\tau$  given by the K-means algorithm for the 5 min averaged  
666 data suggests that there is only a small amount of points covering the transition time and we cannot  
667 fit an ARMA(p,q) model properly to subsequences this short. Moreover, as the value for  $\tau$  given

668 by the Anderson Darling Test is much bigger than the amount of points covering the transition  
669 time we do not expect reliable results for  $\Upsilon$  with this  $\tau$ . Figure 12 confirms this hypothesis. The  
670 left plot is with  $\tau = \tau_{AD}$  and the right one with  $\tau = \tau_{KMeans}$ .

#### 671 **4. Discussion and conclusion**

672 In this study we analyzed the potential of a statistical indicator to be used to detect the system's  
673 proximity to critical regime transitions in the near-surface temperature inversion. The statistical  
674 indicator evaluates the dynamical stability of time series resulting from a dynamical system and  
675 was initially suggested in Faranda et al. (2014). Based on idealized numerical simulations, van  
676 Hooijdonk et al. (2016) had found the presence of early-warning signs in the turbulent flow field  
677 before a transition from weakly stable to strongly stable conditions. These signs included a critical  
678 slowing down, referring to the fact that dynamical systems tend to recover slower from perturba-  
679 tions when approaching a transition point in the dynamics. This slowdown was evaluated based  
680 on fluctuations of the temperature field and the early-warning signal relied on a change in the  
681 variance. Such metrics, which are often used in studies of tipping points, can become problematic  
682 when the underlying dynamics is highly non-stationary, as an increase of variance could be due to  
683 the non-stationarity of the system without implying a transition (Lenton et al. 2012; Faranda et al.  
684 2014). The typical scatter of atmospheric field data and their inherent non-stationarity makes the  
685 application of classical critical slowdown metrics difficult. The metric presented and used here is  
686 different in that it statistically quantifies the deviation from the dynamics expected when the sys-  
687 tem is close to a stable equilibrium. Specifically, the indicator is based on ARMA modeling with a  
688 moving window for which local stationarity is assumed, and the distance from stable equilibrium  
689 dynamics is evaluated based on a Bayesian information criterion. The indicator crucially relies on  
690 an appropriate window length and we suggested two methods to select its value in a data-driven

691 manner. That is, both methods can be used when the underlying model governing the dynamics  
692 is unknown, such that those can be applied to field data with significant scatter. Our suggestion  
693 to ensure reliable results is to use a combination of both approaches. The shortest residence time  
694 around an equilibrium estimated through the K-means approach provides an upper bound to select  
695 a window length that respects the timescale of the system, i.e. a length that ensures local station-  
696 arity for ARMA model fitting. The window length should be selected as shorter or equal to this  
697 upper bound, and such that the data within individual windows mostly satisfy normality to ensure  
698 reliable Bayesian inference. The Anderson Darling normality test is appropriate, but an improve-  
699 ment of the clustering approach to estimate the residence time around an equilibrium (here done  
700 based on a simple K-means clustering approach) would be beneficial. Based on this approach,  
701 we find that a nocturnal temperature inversion dataset with a sampling frequency of 1 minute can  
702 be analyzed successfully using a window length of approximately 20 minutes. Slower sampling  
703 frequency did not lead to conclusive results.

704 The conceptual model introduced by van de Wiel et al. (2017) was developed to understand  
705 regime transitions in the near-surface temperature inversion and can support scenarios with multi-  
706 ple stable equilibria. For our purpose of identifying the system's proximity to regime transitions,  
707 it offers an ideal model for which the theoretical dynamical stability can be calculated analytically.  
708 We extended the model to include random perturbations in the dynamics and used the resulting  
709 stochastic model to provide a test dataset on which to evaluate the potential of the indicator of  
710 regime transitions. In this stochastic system, small-scale perturbations can be amplified due to  
711 the nonlinearity, resulting in transitions between the bi-stable equilibria. Our simulations show  
712 such noise-induced regime transitions, successfully identified by the indicator  $\Upsilon$ . More research  
713 would however be beneficial in order to assess the type of noise that is appropriate to represent  
714 randomized dynamics of the SBL.

715 The application to field data was done for one nocturnal dataset and one Polar dataset. In their  
716 discussion, van de Wiel et al. (2017) suggest that the strength of the thermal coupling between  
717 the soil and the atmosphere may be a key process to distinguish between cases where the temper-  
718 ature inversion has a unique stable equilibria and cases with bi-stable equilibria, separated by an  
719 unstable equilibrium. The wind-speed dependence of observational scatter is partly attributed to  
720 the existence of a dynamically unstable branch in the system in cases where the thermal coupling  
721 is weak. In both datasets considered in our analysis, a weak thermal coupling is to be expected.  
722 Clearly in the Polar dataset, the snow surface leads to a weak thermal coupling between the atmo-  
723 sphere and the soil (Vignon et al. 2017; van de Wiel et al. 2017). The nocturnal dataset originates  
724 from a wheat crop near Dumosa, Australia, probably resulting in a weak thermal coupling as well.  
725 While the Dome C data did not have the required sampling rate in order to have reliable estimates  
726 of the dynamical stability, the Dumosa data were found to have a clear signal with one dynam-  
727 ically stable branch and one dynamically unstable branch. A second dynamically stable branch  
728 corresponding to a strong inversion was not clearly observed. This data-driven result agrees with  
729 the theoretical result of van de Wiel et al. (2017), namely that a dynamically unstable branch ex-  
730 ists for a certain range of wind speeds in case of weak atmosphere-surface thermal coupling. Note  
731 that this is an idealized model and other non-represented physical processes may be at work and  
732 impact the interpretations. As an additional note of caution, if the nature of the noise was different  
733 in the two populations evident in the Dumosa data, then the value of  $\Upsilon$  could differ even if both  
734 branches were dynamically stable. Differences in the noise memory properties may also impact  
735 the results. Indeed, depending on how the noise enters the dynamics, its memory might or might  
736 not be represented well by an ARMA model. This result is nevertheless promising for the use of  
737 the indicator as an early-warning signal of regime transitions. Extending the analysis to a Polar  
738 night with an appropriate sampling frequency would be very interesting, as multiple regime tran-

739 sitions occur during the long-lived temperature inversion (Baas et al. 2019). Moreover, comparing  
740 results obtained for a site with strong atmosphere-surface thermal coupling would provide great  
741 insight to compare the dynamical stability of field data to the dynamical stability predicted by the  
742 conceptual model.

743 To be noted is the fact that the conceptual model is derived for a temperature inversion between  
744 the surface and a height at which the wind speed stays relatively constant during the night, found  
745 to be approximately 40 m at Cabauw in the Netherlands and 10 m at Dome C. The Dumosa dataset  
746 did not offer the possibility to select a height with such a constant wind speed. The measurements,  
747 taken at lower heights in this case, will be prone to submeso-scale activity, inducing perturbations  
748 of the shallow inversion which could affect the dynamical stability of the time series. In fact, the  
749 earlier application of the dynamical stability indicator to SBL data in Nevo et al. (2017) showed  
750 that higher stability corresponded to unstable dynamics of the vertical velocity fluctuations and of  
751 the wind speed. More analyses would be needed to assess the influence of the measurement height  
752 on the evaluated dynamical stability of the temperature inversion. Nevertheless, our results encour-  
753 age the use of the statistical dynamical stability as a metric to detect nearing regime transitions in  
754 the SBL. The ability to detect nearing regime transitions in atmospheric numerical weather predic-  
755 tion and climate models could offer a possibility to use a different type of SBL parameterization  
756 in those specific cases without relying on the assumption of turbulence stationarity.

757 *Acknowledgments.* The authors wish to thank Christophe Genthon and Etienne Vignon for pro-  
758 viding the Dome C data, obtained as part of the CALVA observation project supported by the  
759 French polar institute IPEV, and for lively discussions. The research was funded by the Deutsche  
760 Forschungsgemeinschaft (DFG) through grant number VE 933/2-2. A.K acknowledges funding  
761 from the DFG through the Collaborative Research Center CRC1114, "Scaling Cascades in Com-

762 plex system”. The scientific exchanges between N.V and D.F were greatly facilitated by funding  
763 through the DAAD exchange program Procope through the project ”Data-driven dynamical sta-  
764 bility of stably stratified turbulent flows”. We are grateful to anonymous reviewers for insightful  
765 comments that helped us improve the manuscript.

## 766 APPENDIX

### 767 A1. Details of the K-means clustering algorithm

768 The clustering is done using the K-means algorithm with the following steps:

- 769 • Input:  $k = \#$  number of clusters, set of points  $x_{i-\tau+1}, \dots, x_i$
- 770 • Place centroids  $c_1, \dots, c_k$  at random locations.
- 771 • Repeat until none of the cluster assignments change:
  - 772 – for each point  $x_i$  find nearest centroid  $c_j$  and assign  $x_i$  to cluster  $j$
  - 773 – for each cluster  $j = 1, \dots, k$  calculate new centroid  $c_j = \text{mean of all points } x_i \text{ assigned}$   
774 to cluster  $j$  in previous step.

### 775 A2. Details of the Anderson-Darling Normality Test

776 The the Anderson-Darling (AD) normality test statistic is based on the squared difference be-  
777 tween the empirical distribution function estimated based on the sample,  $F_n(x)$ , and the normal  
778 distribution  $F^*(x)$ . The statistic for this test is,

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 \psi(F^*(x)) dF^*(x)$$

779 where  $\psi$  is a non-negative weight function which is used to emphasize the tails of the presumed  
780 distribution. We use the modified AD statistic given by D’Agostino and Stephens (1986) which

781 takes into accounts the sample size  $n$

$$W_n^{2*} = W_n^2(1 + 0.75/n + 2.25/n^2).$$

782 The null hypothesis of this normality test is that the data are sampled from a normal distribution.

783 When the p-value is greater than the predetermined critical value ( $\alpha = 0.05$ ), the null hypothesis

784 is not rejected and thus we conclude that the data is normally distributed.

### 785 **A3. Summary of the algorithmic procedure**

786 The full procedure to apply the statistical indicator to a timeseries follows the following steps:

787 • Input: Timeseries  $x_1, \dots, x_T$

788 • Evaluate the window length  $\tau_{Kmeans}$ :

789 – Cluster the timeseries in  $k$  clusters (expected number of equilibrium states) using the

790 K-means algorithm.

791 – For each cluster, calculate the mean residence time of the timeseries.

792 – The minimal mean residence time over the  $k$  clusters provides an upper bound for

793  $\tau_{Kmeans}$ .

794 • Evaluate the window length  $\tau_{AD}$ :

795 – For a range of window lengths  $\tau_{min} < \tau < \tau_{max}$ , apply the AD test statistic to the time-

796 series in a moving window approach.

797 – For each  $\tau$ , calculate the median of the p-values obtained over all windows.

798 – Select the largest  $\tau$  for which the median p-value is greater than the predetermined

799 critical value ( $\alpha = 0.05$ ) as  $\tau_{AD}$ .

- 800 • If  $\tau_{AD} < \tau_{Kmeans}$ , select  $\tau_{AD}$  as a window length. Else, ARMA model fitting may be innap-  
801 propriate.
- 802 • Repeat for each window of length  $\tau_{AD}$ :
  - 803 – Select the best fitting ARMA(p,q) model (minimal BIC).
  - 804 – Fit an ARMA(1,0) to the window and calculate the BIC.
  - 805 – Calculate  $\Upsilon$ . High values will indicate transitions. The values themselves may depend  
806 on the dataset.

## 807 **References**

- 808 Abraham, C., and A. H. Monahan, 2019: Climatological Features of the Weakly and Very Sta-  
809 bly Stratified Nocturnal Boundary Layers. Part I: State Variables Containing Information about  
810 Regime Occupation. *Journal of Atmospheric Sciences*, **76 (11)**, 3455–3484.
- 811 Acevedo, O. C., L. Mahrt, F. S. Puhales, F. D. Costa, L. E. Medeiros, and G. A. Degrazia, 2015:  
812 Contrasting structures between the decoupled and coupled states of the stable boundary layer.  
813 *Quarterly Journal of the Royal Meteorological Society*, **142 (695)**, 693–702.
- 814 Acevedo, O. C., R. Maroneze, F. D. Costa, F. S. Puhales, G. A. Degrazia, L. G. N. Martins,  
815 P. E. S. Oliveira, and L. Mortarini, 2019: The Nocturnal Boundary Layer Transition from  
816 Weakly to Very Stable. Part 1: Observations. *Quarterly Journal of the Royal Meteorological*  
817 *Society*, **145 (725)**, 3577–3592.
- 818 Acevedo, O. C., O. L. L. Moraes, G. A. Degrazia, and L. E. Medeiros, 2006: Intermittency and  
819 the Exchange of Scalars in the Nocturnal Surface Layer. *Boundary-Layer Meteorology*, **119 (1)**,  
820 41–55.

- 821 Anderson, T. W., and D. A. Darling, 1952: Asymptotic theory of certain "goodness of fit" criteria  
822 based on stochastic processes. *The Annals of Mathematical Statistics*, **23** (2), 193–212.
- 823 Baas, P., B. J. H. van de Wiel, S. J. A. Linden, and F. C. Bosveld, 2017: From Near-Neutral to  
824 Strongly Stratified: Adequately Modelling the Clear-Sky Nocturnal Boundary Layer at Cabauw.  
825 *Boundary-Layer Meteorology*, **166** (2), 217–238.
- 826 Baas, P., B. J. H. van de Wiel, E. van Meijgaard, E. Vignon, C. Genthon, S. J. A. van der Linden,  
827 and S. R. de Roode, 2019: Transitions in the wintertime near-surface temperature inversion at  
828 Dome C, Antarctica. *Quarterly Journal of the Royal Meteorological Society*, **145** (720), 930–  
829 946.
- 830 Brockwell, P. J., and R. A. Davis, 2016: *Introduction to Time Series and Forecasting*. 3rd ed.,  
831 Springer.
- 832 Chen, L., R. Liu, Z.-P. Liu, M. Li, and K. Aihara, 2012: Detecting early-warning signals for  
833 sudden deterioration of complex diseases by dynamical network biomarkers. *Scientific reports*,  
834 **2**, 342.
- 835 Donda, J. M. M., I. G. S. van Hooijdonk, A. F. Moene, H. J. J. Jonker, G. J. F. van Heijst, H. J. H.  
836 Clercx, and B. J. H. van de Wiel, 2015: Collapse of turbulence in stably stratified channel flow:  
837 a transient phenomenon. *Quarterly Journal of the Royal Meteorological Society*, **141** (691),  
838 2137–2147.
- 839 Faranda, D., B. Dubrulle, and F. M. E. Pons, 2014: Statistical early-warning indicators based on  
840 autoregressive moving-average models. *Journal of Physics A: Mathematical and Theoretical*,  
841 **47** (25), 252001, URL [\url{https://doi.org/10.1088/1751-8113/47/25/252001}](https://doi.org/10.1088/1751-8113/47/25/252001).

842 Genthon, C., M. S. Town, D. Six, V. Favier, S. Argentini, and A. Pellegrini, 2010: Meteorological atmospheric boundary layer measurements and ecmwf analyses during summer at  
843 dome c, antarctica. *Journal of Geophysical Research: Atmospheres*, **115 (D5)**, doi:10.1029/  
844 2009JD012741.

846 Hartigan, J. A., and M. A. Wong, 1979: Algorithm as 136: A k-means clustering algorithm.  
847 *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **28 (1)**, 100–108.

848 Hastings, A., and D. B. Wysham, 2010: Regime shifts in ecological systems can occur with no  
849 warning. *Ecology letters*, **13 (4)**, 464–472.

850 Holdsworth, A. M., and A. H. Monahan, 2019: Turbulent Collapse and Recovery in the Stable  
851 Boundary Layer Using an Idealized Model of Pressure-Driven Flow with a Surface Energy  
852 Budget. *Journal of Atmospheric Sciences*, **76 (5)**, 1307–1327.

853 Holtslag, A., and Coauthors, 2013: Stable Atmospheric Boundary Layers and Diurnal Cycles:  
854 Challenges for Weather and Climate Models. *Bulletin of the American Meteorological Society*,  
855 **94 (11)**, 1691–1706.

856 Hyndman, R., G. Athanasopoulos, C. Bergmeir, and et. al., 2019: Forecasting functions for time  
857 series and linear models. URL `\url{http://pkg.robjhyndman.com/forecast}`, r package version  
858 8.4.

859 Krumscheid, S., M. Pradas, G. A. Pavliotis, and S. Kalliadasis, 2015: Data-driven coarse grain-  
860 ing in action: Modeling and prediction of complex systems. *Phys. Rev. E*, **92**, 042 139, URL  
861 `\url{https://link.aps.org/doi/10.1103/PhysRevE.92.042139}`.

862 Kuznetsov, Y. A., 2013: *Elements of applied bifurcation theory*, Vol. 112. Springer Science &  
863 Business Media.

- 864 Lang, F., D. Belušić, and S. Siems, 2018: Observations of Wind-Direction Variability in the Noc-  
865 turnal Boundary Layer. *Boundary-Layer Meteorology*, **166** (1), 51–68.
- 866 LeMone, M. A., and Coauthors, 2018: 100 Years of Progress in Boundary Layer Meteorology.  
867 *Meteorological Monographs*, **59**, 9.1–9.85.
- 868 Lenton, T., V. Livina, V. Dakos, E. Van Nes, and M. Scheffer, 2012: Early warning of climate tip-  
869 ping points from critical slowing down: comparing methods to improve robustness. *Philosophical*  
870 *Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*,  
871 **370** (1962), 1185–1204.
- 872 Mahrt, L., 2014: Stably Stratified Atmospheric Boundary Layers. *Annual Review of Fluid Me-*  
873 *chanics*, **46**, 23–45.
- 874 Maroneze, R., O. C. Acevedo, F. D. Costa, F. S. Puhales, G. Demarco, and L. Mortarini, 2019:  
875 The Nocturnal Boundary Layer Transition from Weakly to Very Stable. Part 2: Numerical Sim-  
876 ulation with a Second Order Model. *Quarterly Journal of the Royal Meteorological Society*,  
877 **145** (725), 3593–3608.
- 878 McNider, R. T., 1995: Predictability of the stable atmospheric boundary layer. *Journal of the*  
879 *atmospheric sciences*, **52** (10), 1602–1614.
- 880 Monahan, A. H., T. Rees, Y. He, and N. McFarlane, 2015: Multiple Regimes of Wind, Stratifi-  
881 cation, and Turbulence in the Stable Boundary Layer. *Journal of Atmospheric Sciences*, **72** (8),  
882 3178–3198.
- 883 Nevo, G., N. Vercauteren, A. Kaiser, B. Dubrulle, and D. Faranda, 2017: Statistical-mechanical  
884 approach to study the hydrodynamic stability of the stably stratified atmospheric boundary layer.  
885 *Phys. Rev. Fluids*, **2** (8), 084 603, doi:10.1103/PhysRevFluids.2.084603.

- 886 Sandu, I., A. C. M. Beljaars, P. Bechtold, T. Mauritsen, and G. Balsamo, 2013: Why is it so  
887 difficult to represent stably stratified conditions in numerical weather prediction (NWP) models?  
888 *Journal of Advances in Modeling Earth Systems*, **5** (2), 117–133.
- 889 Scheffer, M., and Coauthors, 2009: Early-warning signals for critical transitions. *Nature*,  
890 **461** (7260), 53.
- 891 Scheffer, M., and Coauthors, 2012: Anticipating critical transitions. *science*, **338** (6105), 344–348.
- 892 Schwarz, G., 1978: Estimating the dimension of a model. *The Annals of Statistics*, **6** (2), 461–464.
- 893 Stephens, M. A., 1974: Edf statistics for goodness of fit and some comparisons. *Journal of the*  
894 *American Statistical Association*, **69** (347), 730–737, doi:10.1080/01621459.1974.10480196,  
895 URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1974.10480196>, <https://www.tandfonline.com/doi/pdf/10.1080/01621459.1974.10480196>.
- 897 Sun, J., D. H. Lenschow, M. A. LeMone, and L. Mahrt, 2016: The Role of Large-Coherent-Eddy  
898 Transport in the Atmospheric Surface Layer Based on CASES-99 Observations. *Boundary-*  
899 *Layer Meteorology*, **160** (1), 83–111.
- 900 Sun, J., L. Mahrt, R. M. Banta, and Y. L. Pichugina, 2012: Turbulence Regimes and Turbulence In-  
901 termittency in the Stable Boundary Layer during CASES-99. *Journal of Atmospheric Sciences*,  
902 **69** (1), 338–351.
- 903 Thomson, D. J., 1987: Criteria for the selection of stochastic models of particle trajectories in  
904 turbulent flows. *Journal of Fluid Mechanics*, **180** (-1), 529–556.
- 905 van de Wiel, B. J. H., A. F. Moene, and H. J. J. Jonker, 2012a: The Cessation of Continuous  
906 Turbulence as Precursor of the Very Stable Nocturnal Boundary Layer. *Journal of Atmospheric*  
907 *Sciences*, **69** (11), 3097–3115.

- 908 van de Wiel, B. J. H., A. F. Moene, H. J. J. Jonker, P. Baas, S. Basu, J. M. M. Donda, J. Sun, and  
909 A. Holtslag, 2012b: The Minimum Wind Speed for Sustainable Turbulence in the Nocturnal  
910 Boundary Layer. *Journal of Atmospheric Sciences*, **69** (11), 3116–3127.
- 911 van de Wiel, B. J. H., A. F. Moene, G. J. Steeneveld, O. K. Hartogensis, and A. Holtslag, 2007:  
912 Predicting the Collapse of Turbulence in Stably Stratified Boundary Layers. *Flow, Turbulence  
913 and Combustion*, **79** (3), 251–274.
- 914 van de Wiel, B. J. H., and Coauthors, 2017: Regime Transitions in Near-Surface Temperature  
915 Inversions: A Conceptual Model. *Journal of Atmospheric Sciences*, **74** (4), 1057–1073.
- 916 van Hooijdonk, I. G. S., J. M. M. Donda, H. J. H. Clercx, F. C. Bosveld, and B. J. H. van de  
917 Wiel, 2015: Shear Capacity as Prognostic for Nocturnal Boundary Layer Regimes. *Journal of  
918 Atmospheric Sciences*, **72** (4), 1518–1532.
- 919 van Hooijdonk, I. G. S., M. Scheffer, H. J. H. Clercx, and B. J. H. van de Wiel, 2016: Early  
920 Warning Signals for Regime Transition in the Stable Boundary Layer: A Model Study. **162** (2),  
921 1–24.
- 922 Vercauteren, N., V. Boyko, A. Kaiser, and D. Belušić, 2019: Statistical Investigation of Flow  
923 Structures in Different Regimes of the Stable Boundary Layer. *Boundary-Layer Meteorology*,  
924 **173** (2), 143–164.
- 925 Vercauteren, N., and R. Klein, 2015: A Clustering Method to Characterize Intermittent Bursts of  
926 Turbulence and Interaction with Submesoscale Motions in the Stable Boundary Layer. *Journal of the  
927 atmospheric sciences*, **72** (4), 1504–1517.

928 Vercauteren, N., L. Mahrt, and R. Klein, 2016: Investigation of interactions between scales of  
929 motion in the stable boundary layer. *Quarterly Journal of the Royal Meteorological Society*,  
930 **142 (699)**, 2424–2433.

931 Vignon, E., and Coauthors, 2017: Stable boundary-layer regimes at Dome C, Antarctica: obser-  
932 vation and analysis. *Quarterly Journal of the Royal Meteorological Society*, **143 (704)**, 1241–  
933 1253.

934 **LIST OF FIGURES**

935 **Fig. 1.** Bifurcation plot for simplified model. The lowest, middle and upper branches correspond  
936 respectively to the equilibria  $x_{e_1}$ ,  $x_{e_2}$  and  $x_{e_3}$  . . . . . 46

937 **Fig. 2.** Example of a potential  $V(x)$  (eq. 4) and its local approximation through the quadratic po-  
938 tential  $\tilde{V}(x)$  (see text for details). . . . . 47

939 **Fig. 3.** Clustered simulated time series for  $C = 6.4$  with K-means algorithm. All points with the  
940 same color correspond to the same cluster. The window length estimated by the K-means  
941 algorithm is  $\tau_{KMeans} = 94$  timesteps, with a discrete timestep  $\Delta t = 0.01$ . Hence the window  
942 duration corresponds to 0.94 time units. . . . . 48

943 **Fig. 4.**  $\Upsilon$  for simplified model with a)  $C = 5.3$ , b)  $C = 6.4$ , c)  $C = 7.2$  and added noise, estimated  
944 for a window length a)  $\tau = 54$ , b)  $\tau = 89$ , c)  $\tau = 143$  discrete timesteps. The grey dots  
945 correspond to missing values which are due to the fact that we only colour the last point of  
946 the modeled subsequence. Moreover, in rare cases the `auto.arima()` R function is not able  
947 to find an appropriate model. Red lines mark the stable equilibria, while the dotted red line  
948 marks the unstable equilibrium. . . . . 49

949 **Fig. 5.** Boxplot of the p-values from the Anderson Darling Test for simulated time series with  
950  $C = 6.4$ . The window length estimated by the Anderson Darling test is  $\tau_{AD} = 67$  discrete  
951 timesteps. . . . . 50

952 **Fig. 6.** Bifurcation diagram of the deterministic system (red) and  $\Upsilon$  calculated for simulated data  
953 (gray dots): a) for  $\tau = \min\{mean(T_i)|i = 1, 2, 3\} - 5$  (K-means), b) for  $\tau = \tau_{AD}$  (Anderson  
954 Darling Test) and c)  $\tau = \tau_{AD}$  if  $\tau_{AD} < \tau_{KMeans} - 5$ , otherwise the subseries are discarded. The  
955 red full and dotted lines show respectively the stable and unstable branches of the bifurcation  
956 diagram. . . . . 51

957 **Fig. 7.** Temperature inversion as a function of wind speed as observed at Dumosa, Australia (a) and  
958 Dome C, Antarctica (b). The color in panel b corresponds to lower and higher incoming  
959 radiation. . . . . 52

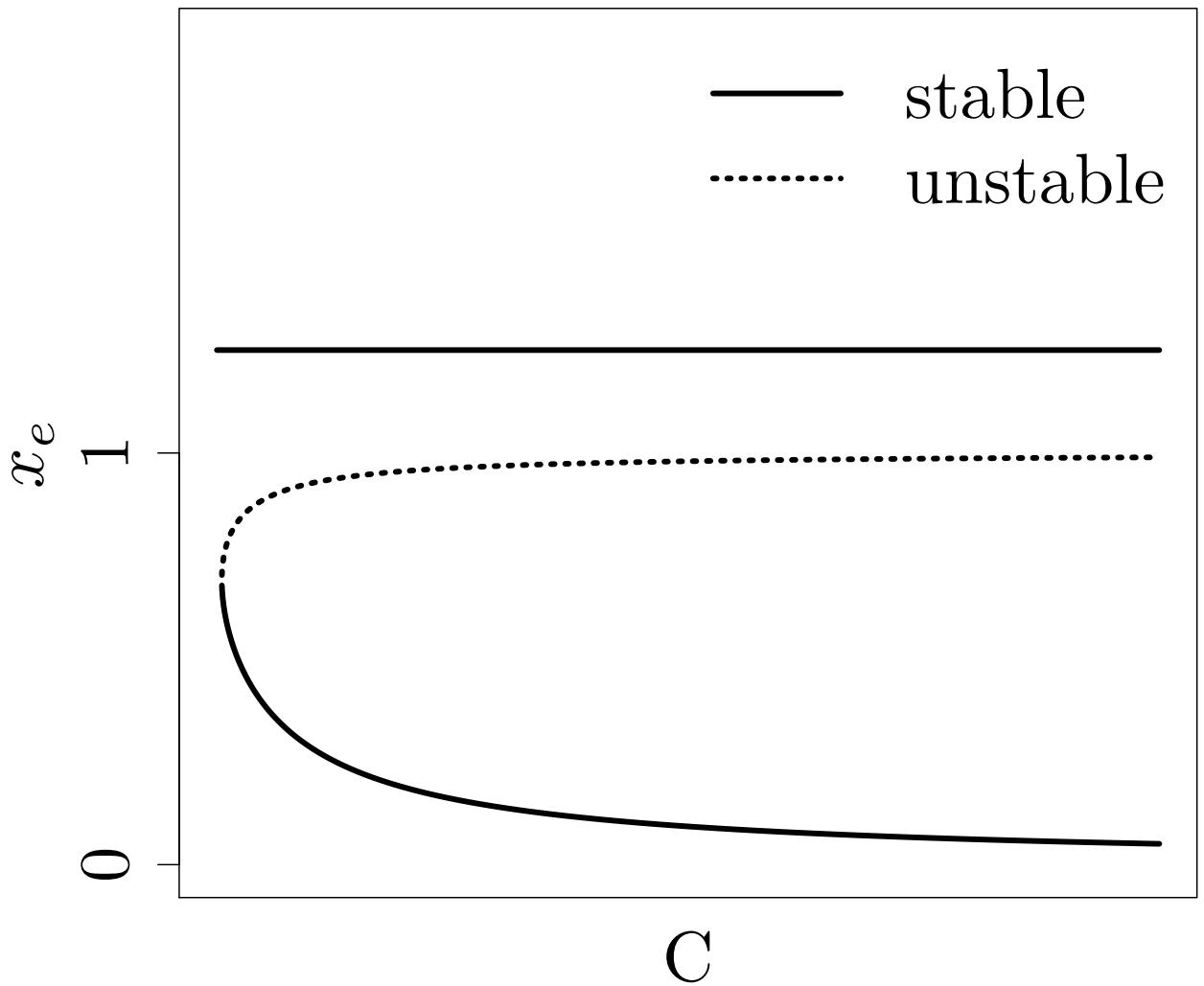
960 **Fig. 8.** a) Boxplot of the p-values from the Anderson Darling test and b) Clustered data with K-  
961 means for the Dumosa data, for all 89 nights. . . . . 53

962 **Fig. 9.** a), c): Observed temperature inversion versus wind speed relation for the Dumosa data.  
963 Colored according to  $\Upsilon$  with different window lengths  $\tau$ : a)  $\tau = \tau_{AD}$ , c)  $\tau = \tau_{KMeans}$ . The  
964 grey dots correspond to missing values which are due to the fact that we only color the last  
965 point of the modeled subsequence. Moreover, in some cases the `auto.arima()` R function is  
966 not able find an appropriate model. b), d): Histogramms of  $\Upsilon$ . The red numbers in plots a  
967 and c are the bin numbers. . . . . 54

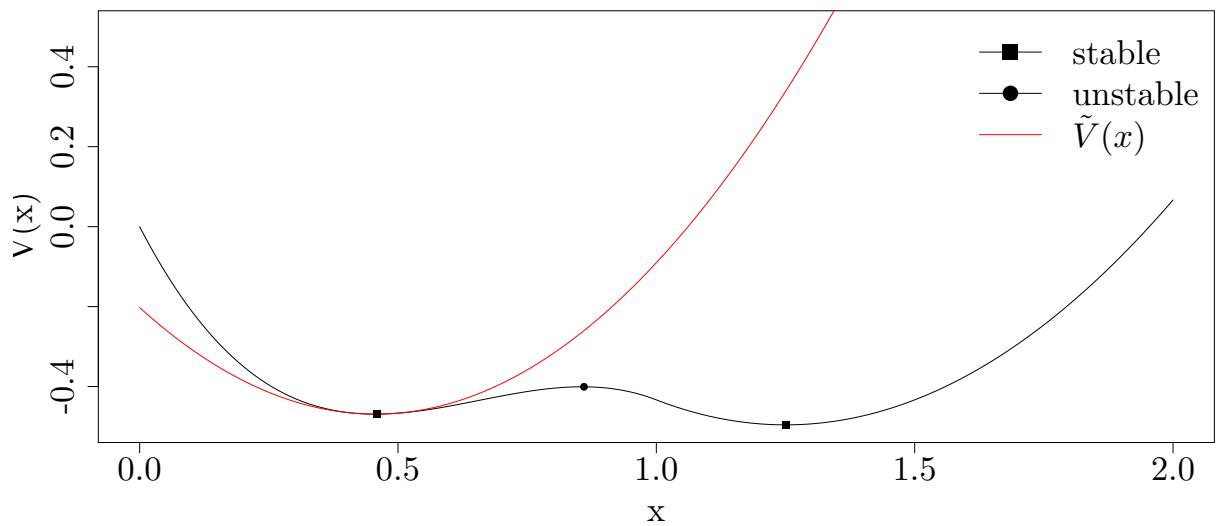
968 **Fig. 10.** Time series of  $mean(\Delta T)$  (a) and  $mean(\Upsilon)$  (b) for different wind speed (`wsp`) categories for  
969 Dumosa data, calculated with  $\tau_{KMeans}$ . The shaded area is the standard deviation. . . . . 55

970 **Fig. 11.** Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as  
971 observed at Dome C. Colored according to  $\Upsilon$  with different window lengths  $\tau$ , expressed in  
972 number of discrete observations: a)  $\tau = 10$  (K-means) and b)  $\tau = 43$  (AD Test) . . . . . 56

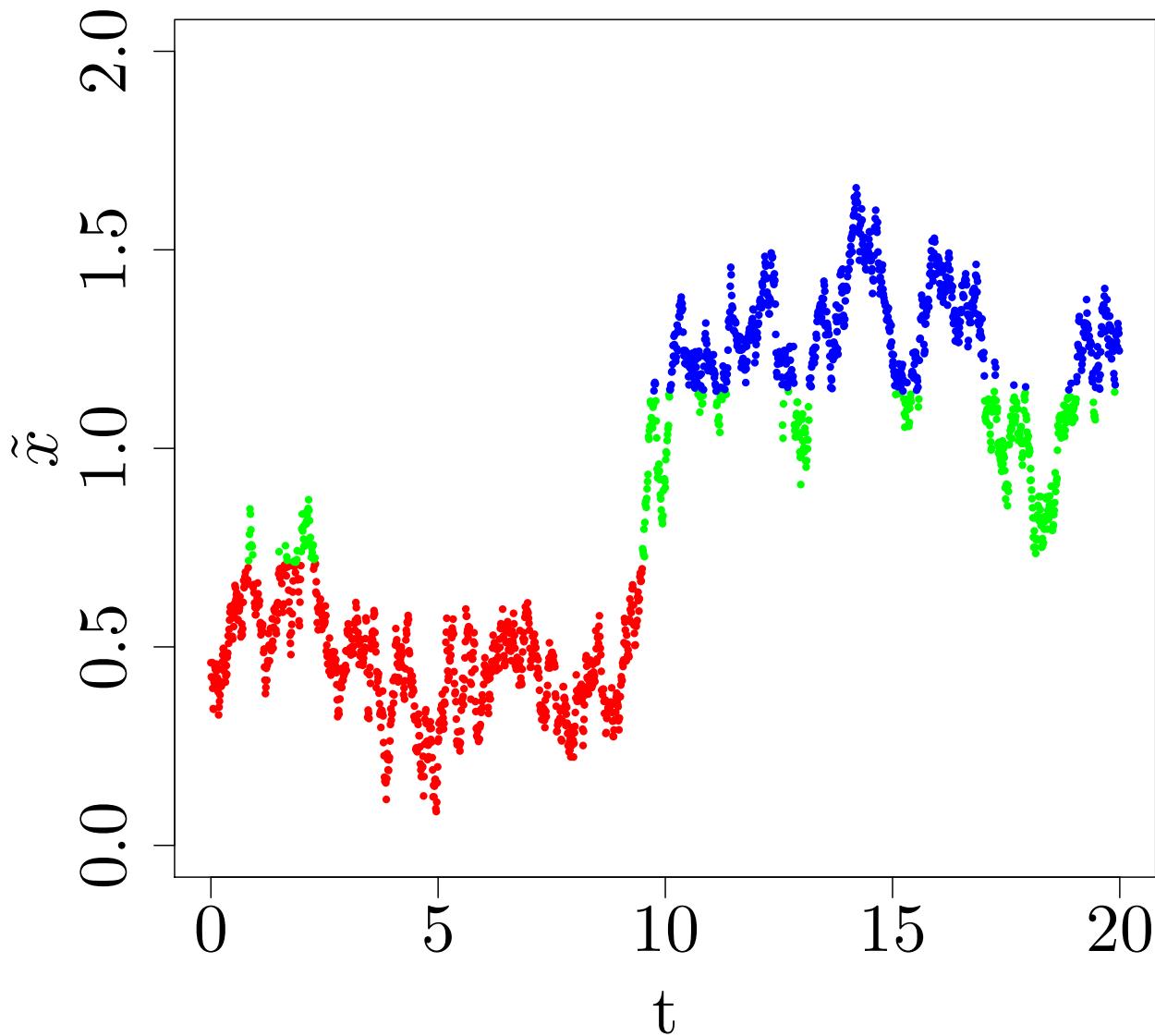
973 **Fig. 12.** Observed temperature inversion versus wind speed relation for the 5-min averaged Dumosa  
974 data. Colored according to  $\Upsilon$  with different window lengths  $\tau$ , expressed in number of  
975 discrete observations: a)  $\tau = \tau_{AD}$  and b)  $\tau = \tau_{KMeans}$ . . . . . 57



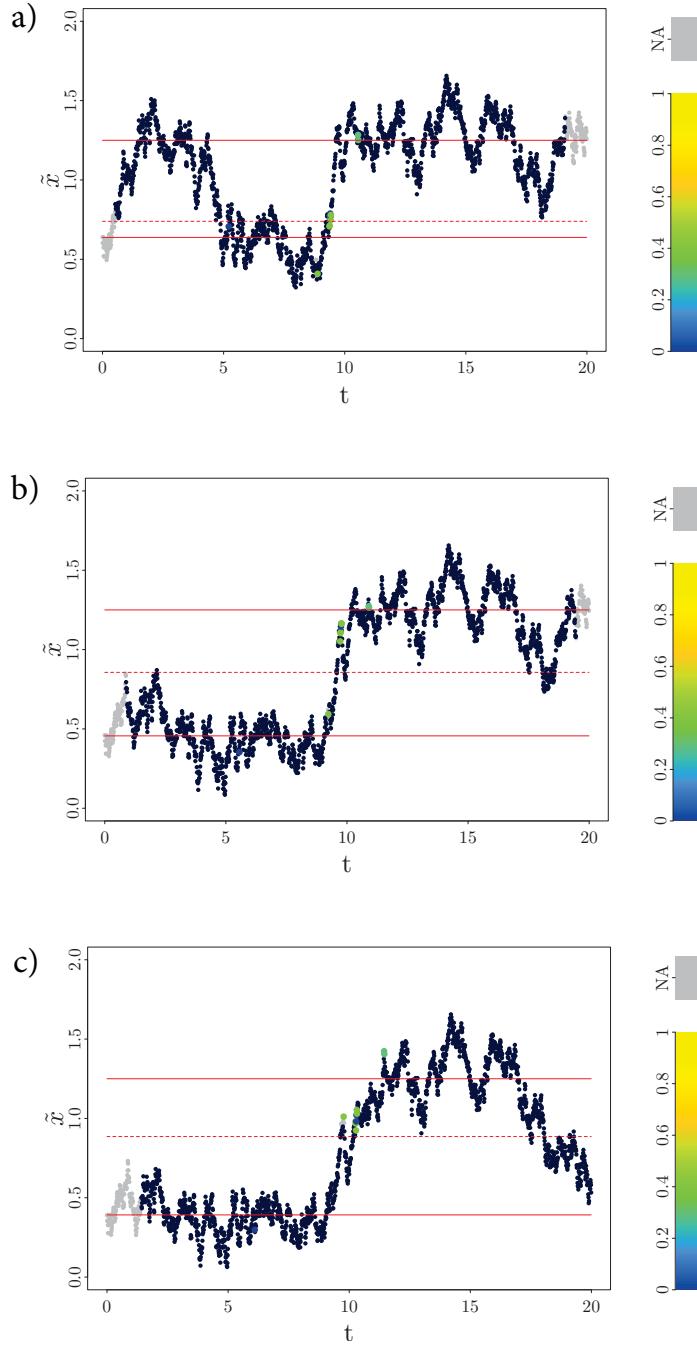
976 FIG. 1. Bifurcation plot for simplified model. The lowest, middle and upper branches correspond respectively  
 977 to the equilibria  $x_{e1}$ ,  $x_{e2}$  and  $x_{e3}$



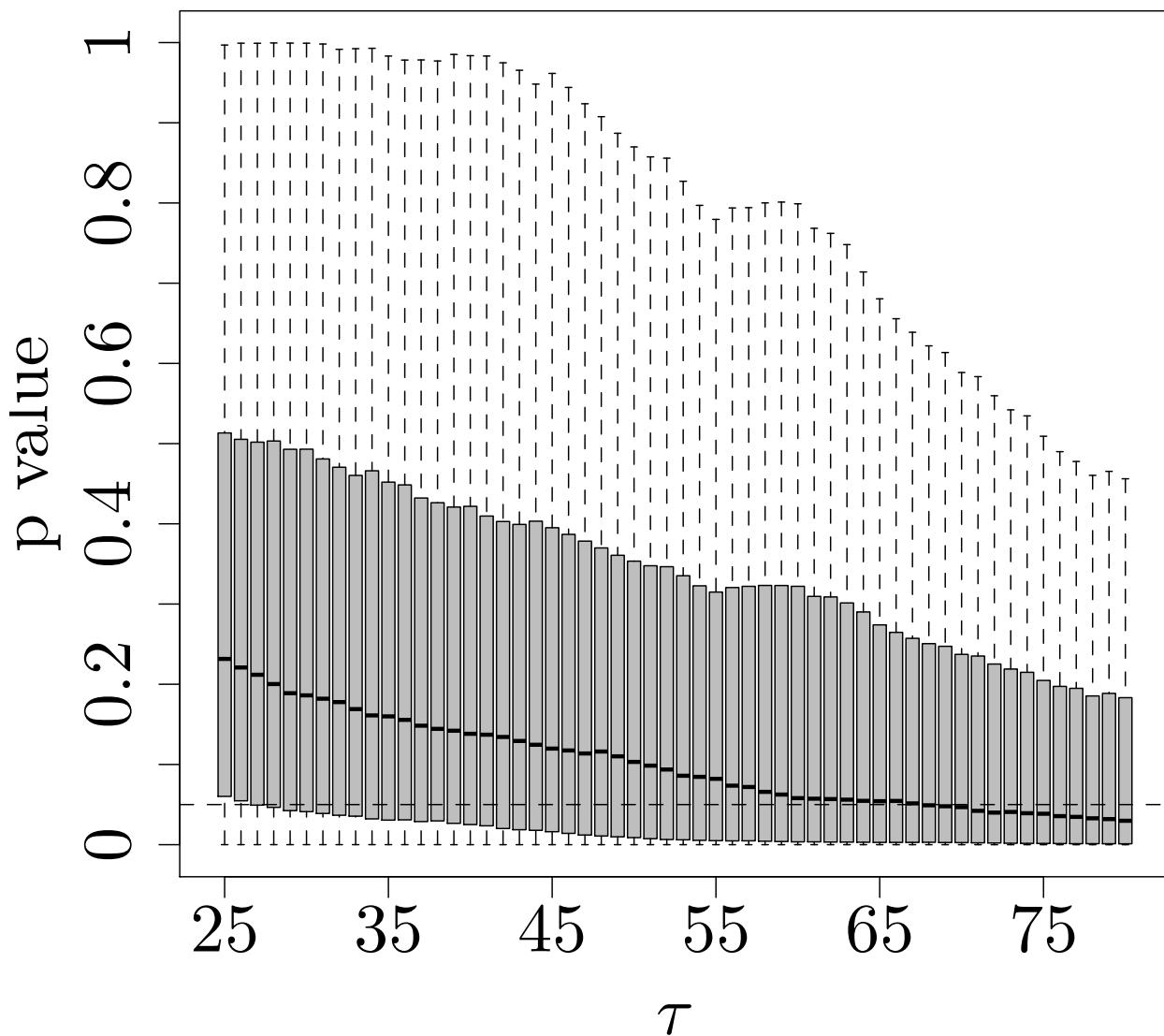
978 FIG. 2. Example of a potential  $V(x)$  (eq. 4) and its local approximation through the quadratic potential  $\tilde{V}(x)$   
 979 (see text for details).



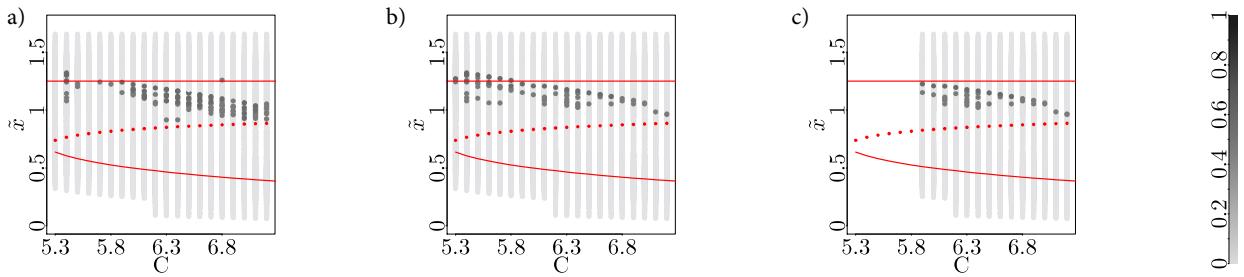
980 FIG. 3. Clustered simulated time series for  $C = 6.4$  with K-means algorithm. All points with the same  
 981 color correspond to the same cluster. The window length estimated by the K-means algorithm is  $\tau_{KMeans} = 94$   
 982 timesteps, with a discrete timestep  $\Delta t = 0.01$ . Hence the window duration corresponds to 0.94 time units.



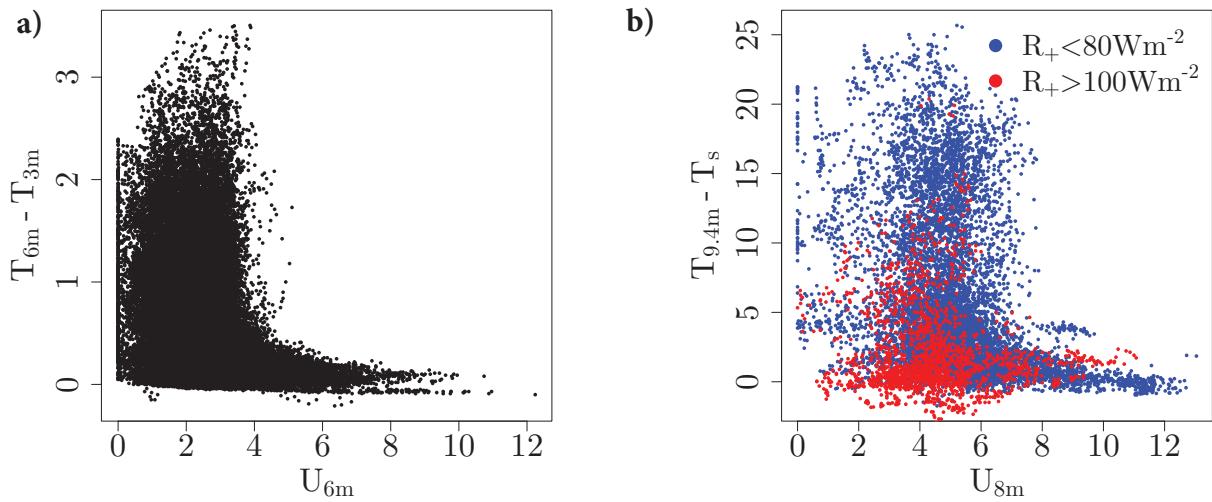
983 FIG. 4.  $\Upsilon$  for simplified model with a)  $C = 5.3$ , b)  $C = 6.4$ , c)  $C = 7.2$  and added noise, estimated for a  
 984 window length a)  $\tau = 54$ , b)  $\tau = 89$ , c)  $\tau = 143$  discrete timesteps. The grey dots correspond to missing values  
 985 which are due to the fact that we only colour the last point of the modeled subsequence. Moreover, in rare cases  
 986 the `auto.arima()` R function is not able to find an appropriate model. Red lines mark the stable equilibria, while  
 987 the dotted red line marks the unstable equilibrium.



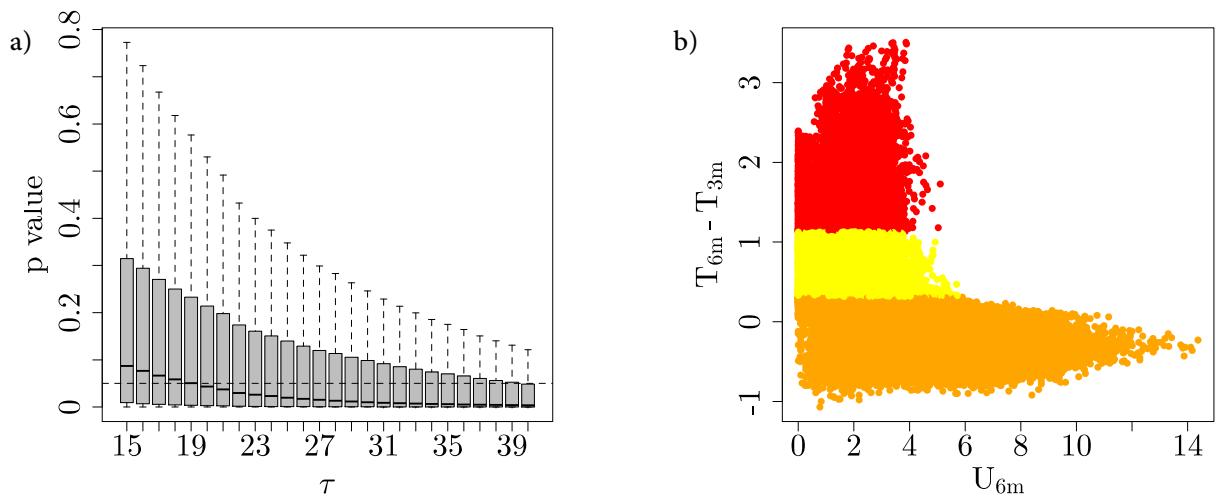
988 FIG. 5. Boxplot of the p-values from the Anderson Darling Test for simulated time series with  $C = 6.4$ . The  
 989 window length estimated by the Anderson Darling test is  $\tau_{AD} = 67$  discrete timesteps.



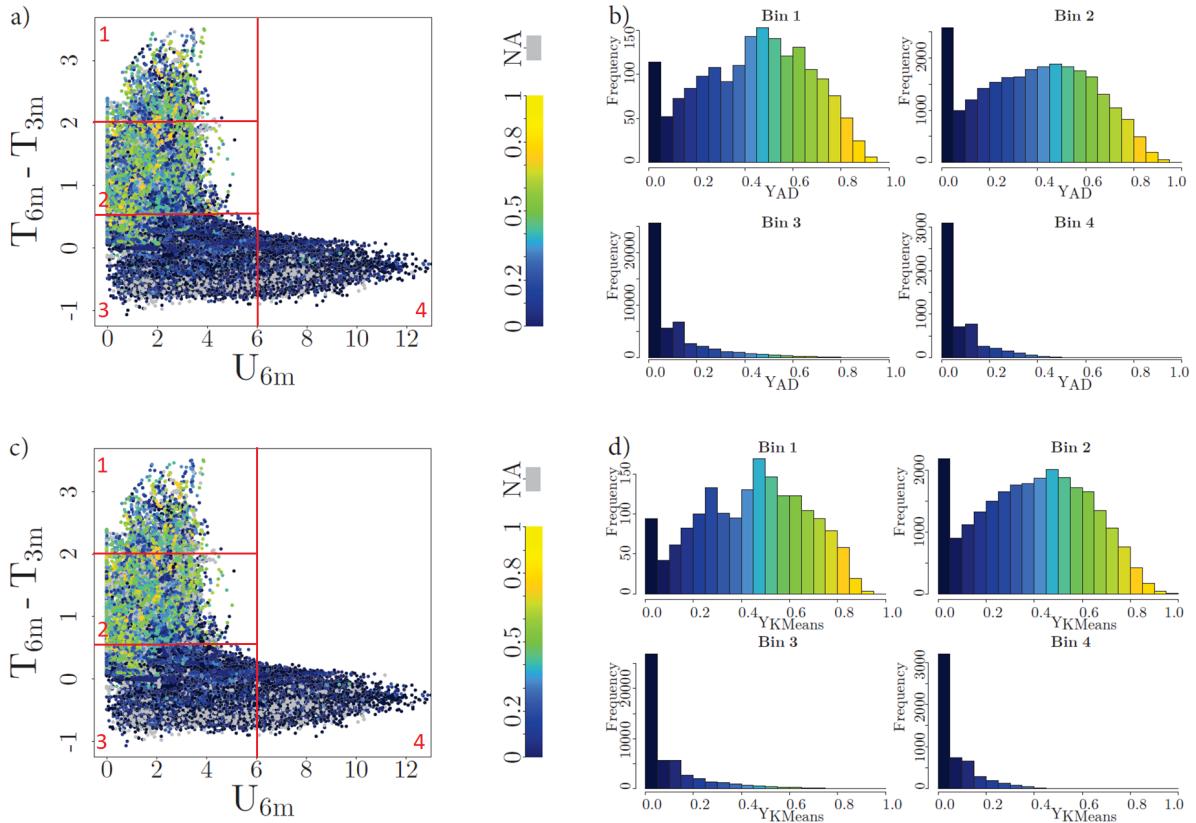
990 FIG. 6. Bifurcation diagram of the deterministic system (red) and  $\Upsilon$  calculated for simulated data (gray dots):  
 991 a) for  $\tau = \min\{mean(T_i)|i = 1, 2, 3\} - 5$  (K-means), b) for  $\tau = \tau_{AD}$  (Anderson Darling Test) and c)  $\tau = \tau_{AD}$  if  
 992  $\tau_{AD} < \tau_{KMeans} - 5$ , otherwise the subseries are discarded. The red full and dotted lines show respectively the  
 993 stable and unstable branches of the bifurcation diagram.



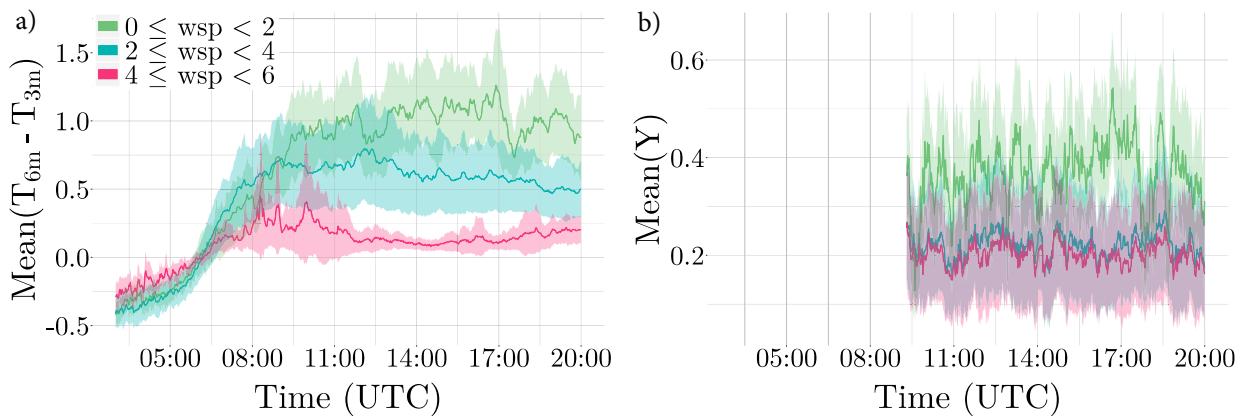
994 FIG. 7. Temperature inversion as a function of wind speed as observed at Dumosa, Australia (a) and Dome C,  
 995 Antarctica (b). The color in panel b corresponds to lower and higher incoming radiation.



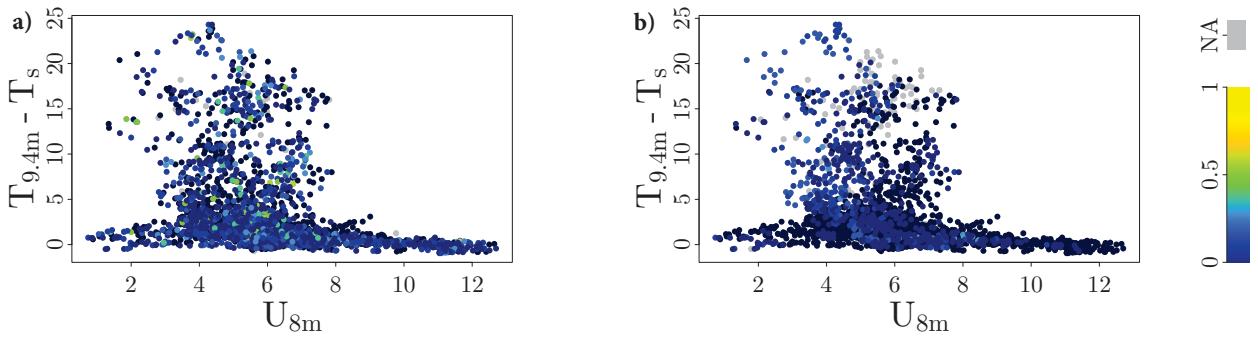
996 FIG. 8. a) Boxplot of the p-values from the Anderson Darling test and b) Clustered data with K-means for the  
 997 Dumosa data, for all 89 nights.



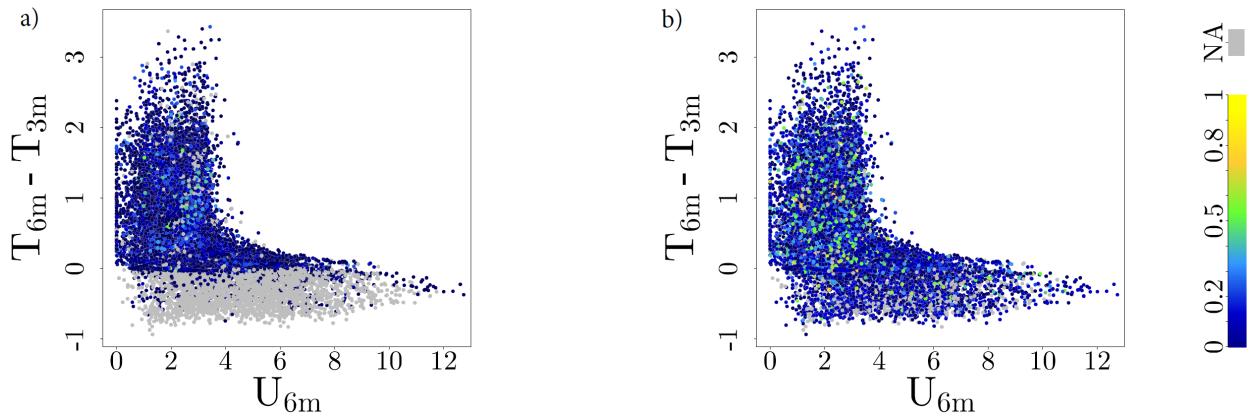
998 FIG. 9. a), c): Observed temperature inversion versus wind speed relation for the Dumosa data. Colored  
 999 according to  $Y$  with different window lengths  $\tau$ : a)  $\tau = \tau_{AD}$ , c)  $\tau = \tau_{KMeans}$ . The grey dots correspond to missing  
 1000 values which are due to the fact that we only color the last point of the modeled subsequence. Moreover, in  
 1001 some cases the `auto.arima()` R function is not able find an appropriate model. b), d): Histogramms of  $Y$ . The red  
 1002 numbers in plots a and c are the bin numbers.



1003 FIG. 10. Time series of  $\text{mean}(\Delta T)$  (a) and  $\text{mean}(Y)$  (b) for different wind speed (wsp) categories for Dumosa  
 1004 data, calculated with  $\tau_{KMeans}$ . The shaded area is the standard deviation.



1005 FIG. 11. Temperature inversion between 9.4 m and the surface, as a function of wind speed at 8m as ob-  
 1006 served at Dome C. Colored according to  $\Upsilon$  with different window lengths  $\tau$ , expressed in number of discrete  
 1007 observations: a)  $\tau = 10$  (K-means) and b)  $\tau = 43$  (AD Test)



1008 FIG. 12. Observed temperature inversion versus wind speed relation for the 5-min averaged Dumosa data.  
 1009 Colored according to  $Y$  with different window lengths  $\tau$ , expressed in number of discrete observations: a)  
 1010  $\tau = \tau_{AD}$  and b)  $\tau = \tau_{KMeans}$ .