Characteristics of the Springtime Alpine Valley Atmospheric Boundary Layer Using Self-Organizing Maps

MARWAN KATURJI, BOB NOONAN, PEYMAN ZAWAR-REZA, TOBIAS SCHULMANN, AND ANDREW STURMAN
Centre for Atmospheric Research, University of Canterbury, Christchurch, New Zealand

(Manuscript received 5 December 2014, in final form 27 July 2015)

ABSTRACT
Vertical profiles of wind velocity and air temperature from a sound detection and ranging (sodar) radio acoustic sounding system (RASS)-derived dataset within an alpine valley of the New Zealand Southern Alps were analyzed. The data covered the month of September 2013, and self-organizing maps (SOM; a data-clustering approach that is based on an unsupervised machine-learning algorithm) are used to detect topological relationships between profiles. The results of the SOM were shown to reflect the physical processes within the valley boundary layer by preserving valley boundary layer dynamics and its response to wind shear. By examining the temporal evolution of ridgetop wind speed and direction and SOM node transitions, the sensitivity of the valley boundary layer to ridgetop weather conditions was highlighted. The approach of using a composite variable (wind speed and potential temperature) with SOM was successful in revealing the coupling of dynamics and atmospheric stability. The results reveal the capabilities of SOM in analyzing large datasets of atmospheric boundary layer measurements and elucidating the connectivity of ridgetop wind speeds and valley boundary layers.

1. Introduction
The atmospheric boundary layer (ABL; the lowest 1–2 km of the troposphere) responds rapidly to radiative, heat, and momentum flux exchanges with the surface throughout its depth (Stull 1988). The structure of the ABL over complex mountainous terrain can be significantly different than over flat surfaces and has been the focus of major research programs (Rotach and Zardi 2007). While our understanding of the ABL over flat areas is relatively advanced (Garratt 1994; Weigel et al. 2007), knowledge of its behavior over mountainous areas needs further research. We believe that ABL research in mountainous areas could benefit from applying newer analytical tools.

The sloped terrain under the mountain boundary layer modifies it dynamically and thermally. For example, mountain boundary layers facilitate exchange between air masses with the free troposphere (Kossmann et al. 1998) by mountain venting. Venting proceeds as upslope winds thicken the boundary layer height locally and trigger the vertical exchange of boundary layer air into the free troposphere. The thermal modification of the boundary layer by terrain not only induces slope flows (Whiteman 2000) but is also shown to introduce large biases in local temperatures from regional projections under climate change, particularly in areas conducive to the formation of stable boundary layers (SBL) (Daly et al. 2010).

The diurnal cycle of the boundary layer is characterized by variations between daytime and nighttime vertical profiles of potential temperature and wind velocity. Near-target remote sensing measurement techniques such as integrated sound detection and ranging (sodar) and radio acoustic sounding system (RASS) systems are able to measure vertical profiles of temperature and wind velocity, which can be used to determine the boundary layer height and its stability at high temporal resolution. However, it is challenging to reduce the wealth of data to characterize the long-term behavior of the atmospheric boundary layer.

Self-organizing maps (SOM) are constructed using a neural network technique based on an unsupervised machine-learning algorithm to detect topological relationships between input variables with any dimension.
and group the organized output into bins or clusters. This reduces the dimensionality of the data and enhances the relationships between variables (Kohonen 1988, 2001). In meteorology, SOM is used to summarize large spatial patterns of temperature, precipitation, or cluster atmospheric circulation patterns from model simulations (Cassano et al. 2006, 2007; Cassano and Cassano 2010; Alexander et al. 2010; Nigro et al. 2011; Richardson et al. 2003). In this paper, we apply the technique to time-varying one-dimensional data from a sodar–RASS measurement system.

The paper is structured as follows: a description is provided of the method, including the experimental setup, instrumentation, and data preparation for the analysis using SOM. The results are divided into an overview of the meteorology during the sampling period followed by an investigation of the intermontane boundary layer (the first 500 m AGL) using SOM. We then apply SOM to investigate the response of nighttime SBL profiles to wind speed profile changes. A discussion and conclusions section is presented that includes the implications of this research.

2. Method

a. Experimental setup

The dataset used in this research was collected using a Scintec sodar–RASS and an automatic weather station. The sodar–RASS was situated on the floor of the Cass River basin at 600 m above mean sea level (MSL), located in the middle of the South Island of New Zealand (43.046°S, 171.768°E) on the eastern side of the Southern Alps. The basin is approximately 7 km wide and aligned in a northwest–southeast direction (Fig. 1). The sodar–RASS system records 10-min averages of the 3D wind velocity and air temperature at a vertical resolution of 5 m—from 10 m AGL for wind velocity and 50 m AGL for air temperature to approximately 500 m AGL. The sodar–RASS system measures virtual potential temperature and also converts to standard air temperature given a user-defined relative humidity and station pressure input. The standard air temperature was then converted to potential temperature accounting for the dry-adiabatic lapse rate. An automatic weather station was positioned at 1400 m MSL on a nearby ridgeline (point B in Fig. 1). The data used for this analysis were obtained during a period of continuous sampling in September 2013 (early spring). Sampling specifications are provided in Table 1.

b. Self-organizing maps

The SOM algorithm used in this research is the SOM Program Package (SOM_PAK; Sheridan and Lee 2011). Initially, SOM is applied to the lapse rate of the vertical potential temperature profiles, which are determined by subtracting the mean temperature of each profile from
each level in the profile. The nodes of the SOM were initialized using a random selection of the profiles and the algorithm was trained twice, reducing the initial learning rate parameter and increasing the number of iterations for the second training but retaining the neighborhood radius as the smallest of the matrix dimensions. SOM iterates through the input dataset, matching each input to the SOM node that is closest in terms of its Euclidean distance, and then adjusts the node and its neighbors to incorporate the input data. The learning rate parameter controls the rate at which the SOM absorbs the information from the input data while the neighborhood radius determines which other nodes, within a topological distance, are affected by the input data. The learning rate decreases to zero and the neighborhood radius decreases to one as the algorithm iterates through the dataset. Several matrix sizes were tested, and for each matrix size the number of iterations and learning rate function types were adjusted, with the aim of reducing the quantization error or the mean Euclidean distance between the input data and the SOM; bigger matrices generally exhibit lower errors. When the change in the error is minimal (known as the elbow criterion), the process can be considered complete. A criterion referred to as the “Sammon Map” is used to see that each node of the SOM has more in common with its neighboring nodes than with nonneighboring nodes. If not, the SOM algorithm was rerun with adjusted parameters. A $3 \times 4$ matrix was selected for the SOM size, as it captured unique profiles in the sodar–RASS dataset without the profiles being too general as with a smaller matrix or containing many similar profiles as with a larger matrix. The SOM algorithm was also used to construct an SOM from a composite wind speed and potential temperature profile anomalies with the aim of investigating the relationships between them.

The sodar–RASS excludes data that do not satisfy the instrument’s signal-to-noise ratio thresholds. We investigated SOM sensitivity to these missing data by constructing an SOM from complete profiles and then gradually removing data level by level from top of the profile to the bottom. The reason we have chosen to remove data in this fashion is that sodar–RASS datasets usually have the highest percentage of missing data in the upper range gates because of sound reflection attenuation. Each time data from a level were removed, we compared the new dataset with the original SOM to see if and how the match of the input data to the SOM nodes had changed (i.e., investigating change in the best matching units). We conclude that the SOM is sensitive to missing data, but for the period and levels (below 350 m) in which we are interested the sensitivity is not significant. The results of the test show that there will be at least a 70% match in the SOM nodes despite the missing data.

### 3. Results and discussion

#### a. Meteorological setting

The midlatitude location of the South Island exposes it to eastward-propagating cyclones and anticyclones, which influence the ABL structure in the mountains. Cold fronts are preceded by intense northwesterly winds, and the passage of the fronts manifest as southerly shifts in wind direction followed by ridging that generally leads to stagnant wind conditions in the mountains. Local thermal circulations can develop when pressure gradients are weak, and with radiative heating and cooling of the valley surface, up- and downvalley winds are typically the observed flows (Sturman 1983, 1987).

Ridgetop weather station data (Fig. 2) indicate that during September the dominant wind direction was

### Table 1. Relevant technical details of the instrumentation used.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Description</th>
<th>Measured variable</th>
<th>Range (accuracy)</th>
<th>Sampling rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaisala, Inc., weather transmitter WXT510</td>
<td>Ridgetop weather station</td>
<td>Air temperature, relative humidity, and barometric pressure</td>
<td>From $-52^\circ$ to $+60^\circ$ C ($\pm 0.3^\circ$ C), 0%–100% ($\pm 3%$), and 600–1100 hPa ($\pm 0.5$ hPa)</td>
<td>15 s, then averaged over 10 min</td>
</tr>
<tr>
<td>R. M. Young Co. 05103 alpine wind sensor</td>
<td>Ridgetop weather station</td>
<td>Wind speed and wind direction</td>
<td>0–100 m s$^{-1}$ (±0.3 m s$^{-1}$) and 0°–360° ($\pm 3^\circ$)</td>
<td>3 s, then averaged over 10 min</td>
</tr>
<tr>
<td>Scintec AG RASS Basin floor sound system (radio frequency 1290 MHz)</td>
<td>Virtual direction</td>
<td>From $-35^\circ$ to $+50^\circ$ C ($\pm 0.2$ C)</td>
<td>10 s for complete nine-beam cycle, then averaged over 10 min</td>
<td></td>
</tr>
<tr>
<td>Scintec Flat Array Sodar (SFAS)</td>
<td>Basin floor sound detection and ranging velocity profiler (audible frequency range 2525–4850 Hz)</td>
<td>$u$, $v$, and $w$ from 10 to 250 m AGL at 5-m vertical resolution</td>
<td>From $-65$ to $+65$ m s$^{-1}$ (±0.1–0.3 m s$^{-1}$ for $u$, $v$) and ±10 m s$^{-1}$ (±0.03–0.1 m s$^{-1}$ for $w$)</td>
<td>~10 s for complete nine-beam cycle, then averaged over 10 min</td>
</tr>
</tbody>
</table>
from the northwest (Fig. 2a) with wind speeds often above 10 m s$^{-1}$. The shift of wind direction from the dominant northwest to the intermittent southeast is clearly depicted in Fig. 2a; it is also associated with generally weaker winds than from the northwest, and in the case of high pressure ridging large oscillations in the diurnal air temperature (Fig. 2b around 15 September) are indicative of cloudless skies. The sampling period also had a cold outbreak associated with a southerly flow that brought air temperatures down to below freezing levels ($-6^\circ$C) on 5 September.

b. Vertical temperature structure

We constructed a $3 \times 4$ SOM matrix for potential temperature (PT) anomalies alongside the hourly frequency distribution and the number of times each 10-min profile transitioned within or between nodes (Fig. 3). Each PT profile (blue lines in Fig. 3) is clustered into a representative SOM node (black lines).

The SOM profiles revealed that for the whole period the boundary layer is dominated by stable to neutral atmospheric conditions, as none of the PT nodes fall within the unstable category. The unclustered data (blue lines in Fig. 3a) show periods of superadiabatic conditions close to the surface; however, as these were uncommon they were not recognized by the SOM as a prominent feature. The stability of the lower boundary layer as revealed by the SOM is questionable, as it is known that the RASS tends to overestimate positive lapse rates in the first 100 m AGL because of cross-talk interference between the transmitter and the receiver (Petenko 1999; Argentini et al. 2008). Bias in the near-surface temperature profile could result from measurement errors in vertical velocity by the sodar as well, as this influences the RASS-derived temperature. To our knowledge, a suitable correction algorithm is not yet available. The shallow stable layer (50–135 m) could also include instances that reflect the true atmospheric stability and result from the anomalously wet and cold conditions for this month. A 13-yr climatological analysis (not shown here) from a long-term weather station on the basin floor suggests that soil moisture for September 2013 was anomalously high and mean monthly air temperature (10 m AGL) was the third coldest out of the 13-yr period. An independent assessment of the near-surface lapse rate approximated by the difference between the lowest usable range gate of the RASS (135 m AGL) and a nearby surface weather station (5 m AGL and 50 m away from the sodar–RASS) indicated that with this bulk approximation stable stratification was achieved for most cases (85% of the entire sampling period), while 15% (all fell within daytime periods) revealed unstable cases. Since the results from 50 to 135 m AGL do not alter the conclusions of this research, the data were excluded from the analysis. Hereinafter, all mention of the boundary layer stability is for the atmospheric thickness from 135 to 500 m AGL.

There is no physical significance associated with the order of the SOM nodes, although they might appear to have been sorted in order of decreasing atmospheric stability from nodes 1 to 12. This ordering is a feature of
the SOM algorithm. Aligned with each PT node is a frequency of occurrence plot (Fig. 3b) that indicates the time of the day at which each of the PT nodes occurs. For example, node 1, which has a relatively strong stable profile, occurs mostly during the evening and morning periods. Node 12 is an example of a well-mixed atmospheric layer. The frequency of occurrence of node 12 shows that this particular profile could occur at any time of the day and suggests that the dominant mechanism causing the well-mixed layer is either one or both
turbulent kinetic energy generation processes (shear or buoyancy).

The internodal transition matrix (Fig. 3c) illustrates the transition frequencies and persistence of each stability type. To aid interpretation, the intensity of the background node color (gray hue) is directly proportional to the frequency. The diagonal of this matrix represents the number of transitions that occurred from each node to itself. Most of the transitions occur in the less stable regimes (nodes 8–12). This finding aligns with the meteorological conditions throughout the sampling period, which were relatively high wind speeds with brief stagnant periods (see discussion in section 3a). Node 1 represents a relatively strong stable profile. As it is an uncommon feature in this dataset there were few transitions from or to it (excluding its self-transitions). Most of the transitions occur between a weakly stable boundary layer and a well-mixed boundary layer.

c. Nighttime SBL dynamics

1) Single-variable approach to SOM

In this section, the impact of the wind speed profile and ridgetop wind speed on mountain ABL stability is the prime focus. Figure 4 shows an SOM of nighttime PT profiles and the corresponding mean wind speed profiles derived from the unclustered data. The PT profile nodes in Fig. 4a all vary from a very stable atmosphere (node 1) to a neutral atmosphere (node 12), and this node intervariability is closely related to the variance in wind speed profiles. The two extreme nodes (nodes 1 and 12) provide a useful example. Node 1 is a relatively strong stability profile, and the wind speeds from the surface up to 200 m AGL are minimal with a linearly increasing wind speed profile above 200 m AGL. Node 12 is a well-mixed profile and it corresponds to strong mean wind speeds (~10 m s\(^{-1}\)). Another example is the progression from nodes 1 to 2 to 3 representing a decrease in stability. This destabilizing of the lower section of the ABL results from the downward transfer of momentum because of an increase in wind speed aloft.

The sensitivity of the mountain boundary layer to ridgetop weather conditions can be highlighted by examining the time series evolution of ridgetop wind speed and direction and node transitions. Figure 4c shows the response of mountain boundary layer stability during the transition period between the two weather systems on 26–27 September 2013 and is an example of an SBL being established. During this period wind speeds decelerated rapidly after the end of a cold southerly outbreak, which brought snow and freezing temperatures to the basin floor. A rapid advance of a high pressure ridge followed resulting in settled weather. As ridgetop wind speed decelerated from around 0230 to 0400 local time (LT) (decelerating at a rate of 2.6 m s\(^{-1}\) h\(^{-1}\)) the PT profile within the basin responded by cooling and stabilizing (moving between nodes 4 and 1). During this period, near-surface wind speeds measured by the sodar at 40 m AGL (not shown here) were less than 2 m s\(^{-1}\), which suggests that the response of the PT profile to ridgetop wind speed deceleration is a decoupling of the near-surface layer from disturbances aloft. The second example (Fig. 4d) did not involve a transition to an SBL, but rather the movement of a cold air mass or front (slowly increasing in depth to envelope the ridgetop station) from the east-southeast direction into the valley-basin system. This transition occurred between 0200 and 0300 LT (see shift in wind direction in Fig. 4d) with an increase in wind speed from 1 to 4 m s\(^{-1}\), followed by relatively constant wind speed throughout the rest of the night. The SOM nodes, in response to this airmass incursion into the basin, oscillated within weakly stable nodes (8–12). This night highlights the fact that although ridgetop wind speeds are important to atmospheric stability in mountain areas, shallow airmass exchanges (below ridgeline heights) occurring on the valley depth scale could play a role in modifying these relationships.

2) Composite-variable approach to SOM

To further explore the potential of the SOM technique, we have reorganized both the wind speed and PT variables for the nighttime period into one composite variable. The purpose here is to allow for the reconstruction of the SOM using two variables that are dynamically interlinked through the effect of wind shear on atmospheric stability. The fundamental difference between this composite SOM approach and the single-variable approach used in Fig. 4 is that SOM will treat both the wind speed and PT parameter as one variable array for every time step of sampling. Previously, only one variable was used for the SOM and the other was a derived quantity from the statistics of the unclustered dataset (an example is Figs. 4a,b). The main rationale behind this approach is to support the conclusions made in the previous section using the one variable SOM approach and demonstrate the success of SOM in performing a composite analysis. This approach could lead to other composite parameter SOM analyses, which is beneficial as the study of atmospheric thermodynamics is often about the correlation of two or more different variables. Figure 5 shows the SOM results of a 3 × 4 matrix for the composite nighttime variable (PT and wind speed anomalies). Note that the node numbers...
(red numerals in Fig. 5) do not refer to the SOMs in the single-variable SOM analysis (Figs. 3a and 4a). The anomalies represent the deviation from the mean of the entire profile. For example, occurrence of a positive surface anomaly of PT and a negative upper-level anomaly indicates a destabilizing atmospheric layer, as the vertical profile of PT is tending toward a neutral profile.

One good example from Fig. 5 is the progressive destabilization that occurs between nodes 1, 2, and 3 where the wind shear is increasing from nodes 1 to 3 (upper-level winds have positive anomalies, while lower-level winds have negative anomalies) and the PT profile is responding to this change by producing a tendency toward a neutral profile (node 3) from a relatively stable profile (node 1) because of the increase of mechanical mixing. The increasing wind shear, moving from left to right for all the nodes, seems to be a persistent dynamic feature of the atmospheric boundary layer of this location at this time of the year. Figure 5 reiterates and strengthens the findings of Figs. 4a and 4b. The SOM technique is successful as an objective tool to study atmospheric stability regimes as a function of wind shear.

4. Conclusions

Wind velocity and air temperature profiles were measured with a sodar–RASS within an alpine valley of the New Zealand Southern Alps. SOM was used to summarize relationships in the raw time series that would be challenging to elicit otherwise. The results of SOM are
Fig. 5. SOM nodes, $3 \times 4$, for a composite potential temperature–wind speed anomaly variable. Potential temperature is below the zero $y$-axis line, and wind speed is above the zero $y$-axis line. The zero $y$-axis line represent 500 m AGL for the PT profile (bottom half of each node), and the zero $y$-axis line also represents the first gate level above the ground for the wind speed profile (top half of each node). Blue is the SOM-derived profile, thick red lines are the mean profiles of the nonclustered data, and thin red line is the mean $\pm 1$ std dev. Note that node numbers have no connection to earlier SOMs in Figs. 3a and 4a.
realistic because they reflect the physical processes within the valley boundary layer and its response to wind shear. The results reveal the capabilities of the SOM method in analyzing large datasets of atmospheric boundary layer measurements and elucidating the connectivity of ridgetop and valley boundary layers. The results from two cases investigating the relationship between ridgetop weather conditions and the response of the valley boundary layer reveal that care needs to be exercised in this complex environment, as the connectivity between ridgetop weather disturbances and valley atmospheric stability could be masked by shallow airmass exchanges occurring on the valley depth scale. Also, the approach of using a composite variable (like in wind speed and potential temperature) for SOM may provide greater detail about the coupling of dynamic and atmospheric stability. SOM could be especially useful for analyzing longer sampling periods that capture extended mountainous boundary layer climates.

Acknowledgments. We thank the technical staff (Justin Harrison and Nicholas Key) at the University of Canterbury for their assistance with the fieldwork. We also acknowledge Dr. John Cassano and Dr. Melissa Nigro for the valuable and inspirational discussions that benefited this research.

REFERENCES


