A Novel Meteorological Method to Classify Wintertime Cold-Air Pool Events

Sean Colgan[1], Xia Sun[2,3] and Heather Holmes[4*]
[1] *Atmospheric Sciences Program, Department of Physics, University of Nevada, Reno, Reno, NV*
[2] *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder*
[3] *NOAA/Global Systems Laboratory, Boulder, Colorado*
[4] *Department of Chemical Engineering, University of Utah, Salt Lake City, UT*
**Corresponding Author: 50 S. Central Campus Dr. MEB 3290, Salt Lake City, UT 84112-9203, +1-801-581-3248, h.holmes@utah.edu*

Highlights:

- $PM_{2.5}$ concentration increases with CAPs but magnitude varies across locations
- Synoptic meteorology determines CAP onset and local meteorology determines duration
- New method was developed to classify CAPs relying only on local meteorology data
- Radiosondes released above the valley floor underestimate CAP strength more than 30%
- Surface station meteorological data can complement radiosonde data to quantify CAPs

Abstract:

Cold air pools (CAPs) are common in mountain valleys throughout the world (e.g., western North America, Himalayas, Alps, etc.) during winter months. Weak surface winds, cold temperatures, high humidity, and snow cover presence are common characteristics of CAPs, and in populated areas there is an increase in air pollution concentrations. Previous methods for identifying CAP events and determining their strength often rely on a combination of, radiosonde data, air pollution concentrations, and/or surface meteorological datasets. Ambient air pollution concentrations vary by location based on the local emissions sources and continually change due to regulations, therefore they are unreliable for consistent CAP quantification. Here, the bulk atmospheric stability is calculated as the valley heat deficit (VHD) using radiosonde data for 12 locations in the western U.S. over 16 winters. A new CAP classification method is developed and compared to three existing CAP classification methods. Results indicate that the new method agrees well with existing approaches but provides a more robust CAP classification because it is solely based on meteorology and not air quality. For all locations, 00Z (afternoon/early evening)

radiosondes account for roughly 20-40% of all CAP occurrences (12Z and 00Z), independent of the method used. Meaning that the stable boundary layer persists throughout the daytime in these cases often leading to persistent CAP events (PCAP). While PCAP length varies across locations, they are a similar order of magnitude because synoptic conditions that span the entire western U.S. govern CAP formation and PCAP length. Additionally, several locations (e.g., Reno, Elko, Spokane, Riverton, and Grand Junction) release radiosondes above the valley floor, often underestimating the bulk atmospheric stability by more than 30%. A method to incorporate surface station data in the bulk atmospheric stability calculation is given to reduce this underestimation.

Keywords:
Stable stratification; Mountain meteorology; Climatology; Radiosonde; Air quality; $PM_{2.5}$

1. Introduction

Cold air pools (CAPs) often occur in areas with valleys and basins in the western U.S. during winter months due to shallow, stably stratified atmospheric boundary layers (ABLs) (Whiteman et al., 2014). CAPs are defined as the topographic depression filled with cold air with warm air overlying above. CAPs can be classified as diurnal, lasting for less than one day, and persistent, lasting for days or even weeks (Whiteman et al., 2001). Diurnal CAPs are formed due to the effective surface radiative cooling at night and decay after sunrise with increased incoming solar radiation. Persistent CAPs (PCAPs) occur when the surface heating is not strong enough to mix out the stratified atmospheric layers in the afternoon. The presence of a PCAP can cause serious air pollutant accumulations in mountainous valleys. During PCAPs, $PM_{2.5}$ (particulate matter with a diameter of less than 2.5 μm) can steadily increase by 10 μg m^{-3} per day, reaching or exceeding the 24-hr $PM_{2.5}$ National Ambient Air Quality Standard (NAAQS) of 35 μg m^{-3} (Holmes et al., 2015; Silcox et al., 2012; Whiteman et al., 2014). The health impacts of air pollution exposure during CAPs have the potential to linger for a month or longer, affecting the cardiovascular and respiratory systems (Zanobetti et al., 2003).

The evolution of a CAP can be summarized as four stages: formation, disturbance, persistence, and break-up (Lareau and Horel 2015a). The formation of a CAP occurs under stagnant synoptic patterns (e.g., an incoming ridge from the west), which produce a pool of cold air at the surface with warm air aloft. The high-pressure/ridge leads to subsidence (sinking motions) that traps the cooler air near the surface and warmer air aloft. The disturbance, persistence and break-up stages are impacted by local factors, like lake breezes (e.g., in the Salt Lake Valley), shortwaves (cold air advection aloft), and turbulence in the lower/middle atmosphere (Lareau and Horel 2015b). For example, sufficiently large surface sensible heat fluxes in the afternoon can destroy the CAP by transporting heat from the ground to the ABL, leading to a breakdown of the stable stratification, namely the dry convection mechanism. In the presence of strong winds aloft, it is possible for the CAP to be removed by turbulent erosion starting from the mountain top and then penetrating downwards into the valley air over time. CAP displacement represents the re-arrangement process of CAP mass caused by dynamic and static processes, such as a mountain wave. It is sometimes the interactions of these mechanisms that lead to the final removal of a CAP. Intermittent weak perturbations between the stages occurs when these mechanisms only weaken but not fully destroy the CAP. Alternatively, snow cover strengthens a CAP with increased surface albedo and reduced surface net radiation.

Low-level clouds impact the thermodynamic structure of the ABL and the surface energy balance, and both have implications on the CAP strength and duration. Holmes et al., (2015) used the terms ‘dry CAP’ and ‘cloudy CAP’ to discuss the boundary layer cloud impacts. They found that the cloudy CAPs enhanced turbulence in the boundary layer due to the cloud top-down mixing impact. Boundary layer clouds modulate the surface radiation budget by reducing the incoming solar radiation and emitting longwave radiation. Cloud top entrainment can induce cloud internal turbulence and affect ABL structure (Deardorff 1970). Radiative cooling at the cloud top also contributes to turbulence production in the boundary layer (Shin and Ha 2009). Holmes et al., (2015) observed higher surface sensible heat fluxes under cloudy CAP with a lower valley heat deficit compared with dry CAP, suggesting a decrease in atmosphere stability. Additionally, the $PM_{2.5}$ concentrations during a cloudy CAP were slightly lower than the dry CAP due to the enhanced vertical mixing. Understanding the key factors impacting CAPs and their interactions during CAP evolution are expected to provide information for potential numerical model developments and ultimately forecasting improvements.

There have been several field campaigns focusing on CAPs. A meteorological field experiment was conducted in a small basin in the United Kingdom (UK) from October 2001 to June 2002 (Jemmet-Smith et al., 2019). Their goals were to better understand temperature and wind flow patterns during CAPs in the Clun Valley, UK (a shallow valley of 100-200-meter depth). The Persistent Cold Air Pool Study (PCAPS) was conducted in the northern Utah from December 2010 to February 2011 (Lareau et al., 2013). There were ten persistent CAPs observed during this winter field campaign with unique meteorological and surface energy balance datasets for analysis. The Yakima Air Wintertime Nitrate Study (YAWNS), focusing on the elevated air pollution levels during wintertime CAPs, was conducted Yakima, Washington from 5 to 27 January 2013 (VanReken et al., 2017). In addition, there are several individual studies focusing on CAPs in other basin areas, such as California's Central Valley in U.S. (Wilson and Fovell 2018), the Danube Valley in Germany (Zängl 2005), and the Po Valley in Italy (Hoggarth et al., 2006). Largely, these studies have been motivated by the significance of CAPs in mountainous regions during the winter months due to an abundance of unhealthy particulate matter accumulating near the surface.

Multiple CAP classification methods have been presented in literature. Each method uses different datasets and thresholds to determine the CAPs. These methods have explored using observations and reanalysis data. Yu et al., (2017) used gridded reanalysis datasets to identify valley cold pools (VCPs) using lapse rates to determine the presence of deep stable layers (Wolyn and McKee 1989) and a 10-m wind speed threshold. There have been multiple observational studies conducted in the Salt Lake Valley (SLV) over the past decade (Baasandorj et al., 2017; Lareau et al., 2013; Franchin et al., 2018; Ivey et al., 2019; Silcox et al., 2012; Sun and Holmes 2019; Whiteman et al., 2014), where industry and vehicle air pollution emissions and topography lead to pollutant accumulation and prolonged stable boundary layers under the unfavorable synoptic conditions. The elevated, wintertime air pollution concentrations in the SLV have been associated with adverse cardiorespiratory health outcomes (Pope III, 1991; Pope III et al., 2006).

In this paper, we investigate the differences between CAP classification methods from Whiteman et al., (2014), Pierce et al., (2019), and Yu et al., (2017) using data during wintertime (i.e., October-March) from 2002-2018 in 12 valley locations at western U.S. The goals of this study are to improve the vertical meteorological data record where the radiosonde is released above the valley floor, evaluate the existing methods to classify CAP events, and develop a new CAP classification method that can be applied to any location. A universal method for identifying CAPs is needed to determine the duration and strength of stagnation across multiple valleys in a region and reduce ambiguity in CAP classification methods.

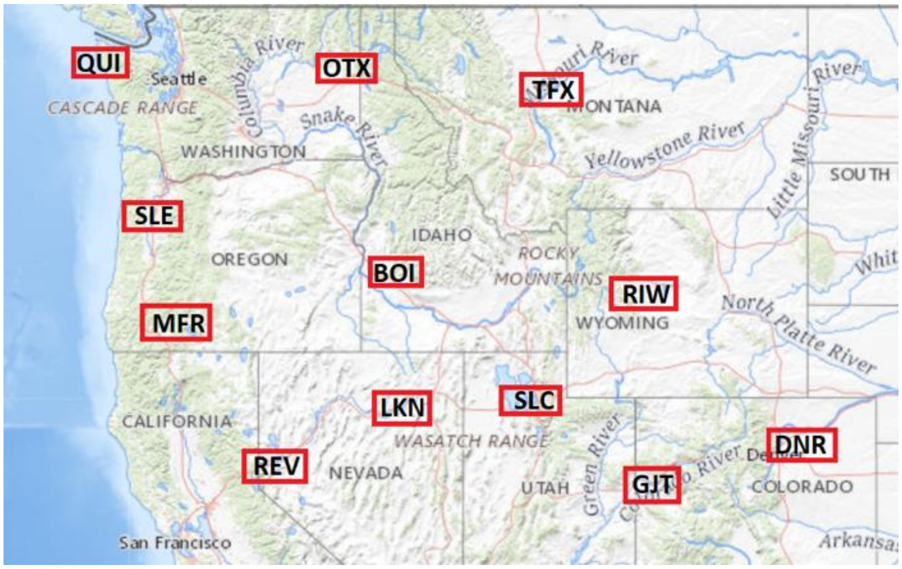

Figure 1. **Map of 12 Radiosonde Locations.** Boise, ID (BOI); Salt Lake City, UT (SLC); Reno, NV (REV); Medford, OR (MFR); Salem, OR (SLE); Denver, CO (DNR); Spokane, WA (OTX); Grand Junction (GJT); Elko, NV (LKN); Great Falls, MT (TFX); Riverton, WY (RIW); Quillayute, WA (QUI).
Map from: https://viewer.nationalmap.gov/advanced-viewer/

2. Data

Observational data from radiosondes, weather stations, and air quality monitors in 12 locations in the western U.S. (Figure 1) were used. This study extends previous CAP studies by investigating CAPs in multiple locations using observations. Using all available radiosonde data across the mountainous western U.S. provides results for locations where CAP events have not been previously studied (e.g., Elko, Salem, and Riverton). This analysis provides an observational quantification of CAP events that can be compared with CAP classifications from gridded reanalysis products in future studies.

2.1 Radiosonde

Radiosonde data from October 1, 2002 to March 31, 2018 were obtained online from the University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html). The stations used to develop and evaluate new metrics for CAP classification are shown in Figure 1 and Table 1. The elevation difference between the valley floor and radiosonde launch site (Table 1) results in missing data for the purpose of CAP identification (i.e., when the radiosonde is released above the valley floor, there is a section of the atmosphere not accounted for in the radiosonde data record). The radiosonde data, including vertical profiles of temperature, dewpoint temperature, pressure and wind, were used to classify CAP events based on the methods described below in Section 3. Mean ridge heights for each valley were determined using a topographic available online from the United States Geological Survey (https://viewer.nationalmap.gov/advanced-viewer/). Mean ridge heights are important for identifying CAPs because the depth of the stable atmospheric boundary layer is below the ridge heights due to the blocking effect caused by the upwind mountains.

Table 1. Elevations (m MSL) of valley floor, mean ridge height, and radiosonde launch site for 12 locations in the western United States and the amount of missing data in the vertical profile obtained from the radiosonde (*i.e.,* elevation difference between radiosonde site and valley floor).

Location	**Valley Floor (m MSL)**	**Mean Ridge Height (m MSL)**	**Radiosonde Site (m MSL)**	**Missing Data (m)**
BOI	830	1829	874	44
SLC	1289	2438	1289	0
REV	1342	2134	1516	174
MFR	400	1219	405	5
SLE	47	914	61	14
DNR	1611	2743	1611	0
OTX	595	1829	728	133
GJT	1397	2743	1475	78
LKN	1540	2438	1608	68
TFX	1015	2134	1134	119
RIW	1509	2743	1703	194
QUI	59	914	62	3

2.2 Surface Meteorology

Surface meteorological data were collected at the three locations where radiosondes are released above the valley floor: Reno (KRNO), Elko (KEKO), and Spokane (KSFF). These locations use Automated Surface Observing Stations (ASOS) and the ASOS data were obtained through MESOWEST (https://mesowest.utah.edu/) (Horel et al., 2002), and followed by quality control checks. For example, station pressure was recorded as mean sea level pressure prior to 2014 for all locations and the altimeter was being recorded as station pressure during that time frame. The ASOS data on the valley floor were appended to the radiosonde data as the lowest elevation at 00Z and 12Z.

2.3 Air Pollution

$PM_{2.5}$ data were used to investigate the relationship between the bulk atmospheric stability and air pollution concentrations during CAP events. $PM_{2.5}$ concentration data were collected from the U.S. Environmental Protection Agency (EPA) Air Quality System (AQS) network for cities with EPA monitors (https://www.epa.gov/outdoor-air-quality-data). Stable atmospheric boundary layer events lead to increased $PM_{2.5}$ concentrations near the surface because of the inability for air parcels to rise and the lower boundary layer heights which leads to less turbulent mixing and a decreased mixing volume, respectively (Holmes et al., 2015; Silcox et al., 2012; Whiteman et al., 2014,). Investigating the relationship between atmospheric stability and pollutant concentrations across the western U.S. is important to determine how the local topography and air population emissions influence the $PM_{2.5}$ concentrations during CAP events. Three cities (Elko, NV; Quillayute, WA; and Riverton, WY) do not have EPA monitors for the time period of this study, thus analyses of bulk atmospheric stability and $PM_{2.5}$ concentrations were not conducted for these locations.

3. Methods

Previous studies typically focused on a single valley, like the Columbia Basin, Snake River Valley, and the Salt Lake Valley (SLV), while few have focused on CAP events that occur across a large region encompassing many valleys. For example, Whiteman et al., (2014) developed a threshold using bulk stability and air pollution that is specific to the SLV. While Yu et al., (2017) used reanalysis data to determine the frequency of CAP events across the western U.S. Every valley in the western U.S. experiences similar synoptic meteorology that leads to the formation of CAPs. However, each valley is unique with respect to local topography (e.g., blocking pattern induced by topography, prevailing winds at the surface, etc.) and the amount of air pollution emissions. Here, we evaluate three CAP classification methods (Section 3.1, 3.2, and 3.3) and propose a new method (Section 3.4) that can be applicable to all locations. The new method relies on physical (i.e., meteorological) variables that are related to the CAP behavior and is independent of air pollution concentrations.

3.1 Valley Heat Deficit with $PM_{2.5}$ Threshold (Whiteman et al., 2014)

The valley heat deficit (VHD, H_h) is a measure of the bulk atmospheric stability for a valley (Whiteman et al., 1999):

$$H_h = c_p \int_{sfc}^{h} \rho(z)[\theta_h - \theta(z)]dz \qquad (1)$$

Where c_p is the specific heat of air (1005 J $kg^{-1}K^{-1}$), h is the integration height, ρ is the density of air, θ is potential temperature and z is the height above ground level. A VHD threshold for Salt Lake City was determined based on a 40-year climatology of VHD and $PM_{2.5}$ concentrations (Whiteman et al., 2014). The VHD threshold can be used to identify the onset of a CAP episode. A persistent CAP occurs when the VHD is above the threshold for at least 3 consecutive soundings (i.e., more than one day). However, the threshold in Whiteman et al., (2014) is specific to SLV. Each valley location would require its own climatological assessment of the VHD and $PM_{2.5}$ relationship to establish the VHD threshold.

Whiteman et al., (2014) used a $PM_{2.5}$ concentration of 17.5 µg m^{-3} (half of the NAAQS for $PM_{2.5}$) to determine the VHD threshold that identifies CAP events in SLV (hereafter referred to as VHD_17.5). Here, a similar approach is taken to determine if this VHD and air quality method can be applied to other valleys in the Intermountain West. The limitation of this NAAQS-based $PM_{2.5}$ threshold is that the different magnitudes of ambient air pollution concentration in each city associated with different amounts of emissions will lead to inconsistent VHD thresholds. Additionally, with the exception of 2016-2020, $PM_{2.5}$ concentrations have been declining in the U.S. due to the regulations on industrial and vehicle emissions (Clay and Muller 2019). Therefore, another VHD threshold for CAP events based on $PM_{2.5}$ statistics, rather than the absolute $PM_{2.5}$ values, is necessary.

3.2 Valley Heat Deficit with 75^{th} Percentile $PM_{2.5}$ (Pierce et al., 2019)

To establish another VHD threshold value, the 75^{th} percentile 24-hour $PM_{2.5}$ concentration is used in locations where air quality data was available. $PM_{2.5}$ above the 75^{th} percentile represent polluted values for that specific location and can attributed to CAP event during the winter months. This method was adapted from Pierce et al., (2019), to identify polluted days their study highlighted that the 75^{th} percentile $PM_{2.5}$ concentration separates typical air pollution days from days that experience elevated air pollution concentrations in Reno, NV (i.e., increased air pollution due to wildfires or CAPs). The 75th percentile $PM_{2.5}$ concentrations are used in this study as a threshold to calculate a unique VHD threshold for each location (hereafter referred to as VHD_75). The mean VHD is obtained from days when the 24-hr $PM_{2.5}$ concentration exceeds the 75th percentile $PM_{2.5}$, then this mean VHD is used as a threshold for identifying CAP events. While this approach no longer uses a static $PM_{2.5}$ concentration threshold to establish a VHD threshold for CAPs, it still relies on air pollution concentrations and not meteorological data.

3.3 Valley Cold Pools (Yu et al., 2017) based on Deep Stable Layers (Wolyn and McKee 1989)

Deep stable layers (DSLs) are abnormally deep stable nocturnal boundary layers that are of at least moderate stability with effective stagnation in the lower atmosphere (Wolyn and McKee 1989). DSLs are defined with a temperature lapse rate less than 2.5 K km^{-1} for at least 65% of the lowest 1500-m AGL layer. Requiring greater than 975 m of the lower atmosphere to reach this criterion often ignores stable layers with effective stagnation in the lower atmosphere that do not meet the definition (Wolyn and McKee 1989). Yu et al., (2017) used the North America Regional Reanalysis (NARR) dataset to identify valley cold pool (VCP) events using the criteria of DSLs in Wolyn and McKee (1989), combined with a 10-m wind speed threshold of less than 3 m s^{-1} for VCP classification.

This method is used in this study (hereafter referred to as VCP) and these criteria are applied to morning and afternoon soundings, unlike Wolyn and McKee (1989) who applied their criterion to only morning soundings. Yu et al., (2017) acknowledged that they likely underestimated the number of VCP in their study due to strict criteria applied to NARR, given the coarse resolution (32-km horizontal and 29 vertical levels) and applying this criterion to afternoon soundings. Despite that, VCP was a common weather phenomenon in the western U.S., with events longer than 7 days occurring in the Columbia Basin, Snake River Basin and Bonneville Basin (Yu et al., 2017).

3.4 Novel CAP Classification Method

Due to the limitations of prior approaches, there is a critical need to develop a universal method that relies on the physical variables of CAP meteorology and is independent of the air pollution concentrations. Ideally, the universal method can also be applied to locations without radiosonde data using girded reanalysis datasets to obtain the vertical atmospheric profiles. However, the focus of this paper is on developing and evaluating this new method using radiosonde observations, since additional assimilation uncertainties exist in reanalysis datasets.

In the western U.S., the onset of a CAP episode occurs when an approaching synoptic-scale ridge (i.e., surface high pressure) settles over the Intermountain West, resulting in increasing surface pressure, decreasing wind speeds near the surface, and increased VHD. These physical variables are attributed to stagnation in the lower atmosphere and can be used to identify CAP events. We use these variables to develop a new, universal CAP classification method. The new classification method still relies on the VHD from Whiteman et al., (1999), but establishes the VHD thresholds based on local meteorology, instead of air quality. We use surface pressure and wind speed to determine the VHD threshold value for CAP events. Normalized station pressure is used to determine high pressure systems by calculating the median station pressure for each location and dividing the ambient station pressure by the median station pressure. Then a median VHD threshold is obtained when the normalized pressure and 10-m wind speed are greater than 1.0 and less than 3 m s^{-1}, respectively. When the normalized pressure is greater than 1.0, a higher surface pressure than normal is observed and has potential to indicate stagnation near the surface. Since higher pressure at the surface is not always indicative of stagnation, the 10-m wind speed threshold of 3 m s^{-1} isolates the stagnation days from non-stagnation days. This new CAP threshold (hereafter referred to as VHD_met) does not rely on the air pollution concentrations and is based purely on the meteorological processes that occur in each valley.

While this method provides a new VHD threshold for CAP events the calculation still requires location specific information that impacts the VHD magnitude (e.g., integration heights in Equation 1). To extend this method and provide a VHD value that is comparable across all locations, the VHD is normalized by a standard VHD value for each sounding. To obtain a standard VHD, the moist adiabatic lapse rate (MALR) is used (e.g., Whiteman et al., 2014). The MALR VHD is calculated for each radiosonde dataset using a standard tropospheric lapse rate of 0.0065°C m^{-1} (NOAA 1976). Then the VHD is divided by the MALR VHD to provide a non-dimensional value of bulk atmospheric stability.

3.5 Station Data Appending

Calculating the bulk atmospheric stability depends on the vertical profiles of potential temperature and pressure. The bulk atmospheric stability is not representative of the valley conditions when the radiosonde is released above the valley floor. For example, in Reno the vertical extent of the missing data between the ASOS station on the valley floor and radiosonde site is 174 m (see Table 1). Station data from the valley floor are pivotal to identify CAPs, especially if the CAP is shallow (i.e., less than 200 m). Appending NWS maintained station data to radiosondes ensures the entire vertical profile of the atmosphere in the valley is well represented and is required to calculate a more accurate bulk stability to identify CAP events. In addition, 10-m wind speed on the valley floor is crucial for classifying CAP events (e.g., classification method in Yu et al., (2017)). If the 10-m wind speed is not included in the vertical profiles where the radiosonde is released above the valley floor, misclassifications of CAP events are possible due to the increasing wind speed with height.

Appending surface station data to radiosonde observations in the western U.S. yields better CAP classification results compared to methods that solely rely on radiosonde data (Green et al., 2015, Pierce et al., 2019). Spokane, Elko and Reno have NWS maintained ASOS stations on the valley floor with data spanning the entire study period; KSFF, KEKO and KRNO respectively. These data are merged with the radiosonde datasets based on time stamps and the data from surface stations are appended onto radiosonde observations, essentially adding a lower elevation point to the vertical radiosonde record. The combined, surface station and radiosonde, dataset was missing less than 5% of the data (i.e., radiosonde launch failure or maintenance on NWS stations when observations led to missing data), therefore adding the station data did not significantly reduce the data availability.

To evaluate the impact of appending the ASOS station data onto the radiosonde record, a sensitivity analysis is conducted using the Medford radiosonde data. This analysis provides an assessment for determining the effects of releasing radiosondes above the valley floor on calculating a bulk atmosphere stability. Quantifying this impact is especially important for locations where NWS maintained stations are absent (i.e., Riverton, Great Falls and Grand Junction). The sensitivity analysis includes two parts. The first, calculates the impact of missing vertical data on VHD and the second estimates the VHD errors associated with adding station data from the valley floor onto the radiosonde record.

To calculate the impact of missing data on VHD, measurements in the lowest 205 m AGL layer of the Medford radiosonde record are withheld to model a situation where the radiosonde is released 205 m above the valley floor. Using this modeled missing data radiosonde profile, a bulk atmospheric stability (VHD) is calculated for each radiosonde. When compared to the VHD calculated using the full radiosonde record, the VHD calculated using the modeled vertical profile with missing data quantifies the underestimation of bulk atmospheric stability. This provides an estimate for the underestimated VHD in Riverton, Great Falls and Grand Junction, where the station meteorology data on the valley floor are unavailable. The radiosonde in Riverton is released 194 m above the valley, less than 205 m AGL, therefore this modeled radiosonde profile from Medford provides the 'worst case scenario' estimate.

Next, to analyze the VHD errors associated with appending surface meteorology data onto to radiosonde record another sensitivity test is done using the Medford radiosonde data. In this analysis, the modeled vertical profile is similar to the case above (measurements in the lowest 205 m AGL layer of the radiosonde record are withheld) but now a surface observation is added onto the vertical record. This simulates adding the 10-m observations to the radiosonde and another modeled VHD is calculated. Then the difference between the observed VHD and the VHD calculated from appending surface station data onto the radiosonde record is calculated to estimate the VHD errors. This provides insight into the impact of the vertical integration in the VHD calculation (Equation 1) for the lowest level where the most stable layer is typically found during CAPs (i.e., does appending surface meteorology data artificially increase the VHD).

4. Results

The overall objective of this paper is to develop a new classification method for identifying CAP events in the western U.S. by analyzing previous classification methods and exploring valley meteorology for each location (i.e., Columbia Basin, Truckee Meadows, Snake River Basin, Rogue Valley, etc.). One goal of this new method is to create a CAP metric that can be used in any location that allows for comparison across locations. The first step is to highlight the importance of appending station data from the valley floor to radiosondes in locations where the radiosondes are released from elevations above the valley floor (e.g., Reno, Elko, and Spokane). The next step is to analyze the air pollution concentrations, bulk atmospheric stability (VHD), 10-m wind speeds, and lapse rates to determine a framework for identifying the meteorological processes that influence CAP events. Finally, results from the new method described above in Section 3.4 are shown, including a comparison of the CAP prevalence across 12 cities in the western U.S. for16 winters.

4.1 Evaluation of Appending ASOS to Radiosonde Data

To evaluate the impact of appending surface station data onto the radiosonde vertical profiles the Medford site is used for a data withholding comparison of the appended vertical profiles. When all soundings in Medford are considered to have a radiosonde launch elevation of 205 m AGL (withholding the data from the first 205 m above the ground), the modeled VHD is 1.47 MJ m^{-2} lower than the observed VHD over the 16 winters. Thus, missing the lowest layer of the atmosphere hinders the ability to classify a CAP. Misclassification of CAPs can hinder air quality forecasting and has implications for residents residing in the lowest part of a valley. For example, when stratifying the data withholding results by selecting only the soundings with VHD values above the new CAP VHD threshold (VHD_met results shown in Section 4.3.4), the modeled missing data VHD is 2.95 MJ m^{-2} lower than the observed VHD. This indicates that low-level temperature inversions occur near the surface and that the stable layer near the surface impacts the VHD calculation.

Estimating the errors associated with the station data appending is necessary to understand the impact on the calculated VHD. When forcing the vertical resolution of the first two data points of the radiosonde record to be 205 m (i.e., simulating the case where surface observations are appended onto the radiosonde record), the modeled VHD is 0.32 MJ m^{-2} greater than the observed VHD in Medford for all CAP soundings. The 12Z and 00Z soundings have a modeled VHD increase of 0.43 and 0.07 MJ m^{-2}, respectively. These results indicate that appending ASOS to radiosonde data is appropriate to extend the vertical profiles down to the surface.

In this study, station data are appended onto the radiosonde record in three locations. Table 2 shows the increased median and mean VHD when valley station data is appended to the radiosondes. In Reno, Spokane, and Elko the overall median VHD increased 54.8%, 25.9%, and 20.7%, respectively, throughout the study period. The Reno radiosonde represents the largest amount of missing data (174 m) and has a larger slope in the linear relationships between the two VHDs than Elko and Spokane, confirming that it has the largest increase in VHD values. The

Table 2. **Data Appending Evaluation**. Mean, median, and standard deviation VHD values (MJ m^{-2}) for three locations (REV, LKN, OTX). *Radiosonde Only* - data provided by the radiosonde. *Appended Radiosonde* - ASOS data appended to the bottom of the radiosonde data. Slope and correlation for the regression between the two VHDs.

	Mean VHD (MJ m^{-2})	**Median VHD (MJ m^{-2})**	**σ VHD (MJ m^{-2})**	**Correlation**	**Slope**
	REV				
Radiosonde Only	1.17	0.73	1.50		
Appended Radiosonde	1.98	1.13	2.55	0.968	1.691
	OTX				
Radiosonde Only	4.58	4.05	3.40		
Appended Radiosonde	5.71	5.10	4.20	0.994	1.237
	LKN				
Radiosonde Only	2.57	1.88	2.63		
Appended Radiosonde	3.10	2.27	3.15	0.997	1.199

Elko radiosonde has the least amount of missing data near the surface (68 m) and has the smallest slope when the ASOS data are included in the vertical profile. More information on the correlation of appending ASOS data to radiosonde can be found in the supplement information. For Reno, Spokane, and Elko, all of the following results in this paper are calculated using vertical profiles with station data appended to the radiosonde vertical profile.

4.2 Statistics of wintertime meteorology and air quality

Statistics of $PM_{2.5}$ concentrations, 10-m wind speed, VHD, and lapse rate percentages (i.e., percent of the lowest 1.5-km layer with a LR <=2.5°C km^{-1}) from sounding data at 00Z and 12Z for all locations over the study period are shown in Figure 2. These variables are key variables to identify CAPs in existing classification methods.

$PM_{2.5}$ concentrations are useful for identifying events (e.g., CAPs and wildfires) that contribute to poor air quality. Figure 2a shows that peak $PM_{2.5}$ concentrations are the highest in Salt Lake City, likely related to the anthropogenic air pollution emissions with a population greater than one million and industrial processes in the valley. Medford also experiences high pollution, relative to other locations in this study, despite having a lower population than Salt Lake City. The mean and median $PM_{2.5}$ concentrations are highest in Medford.

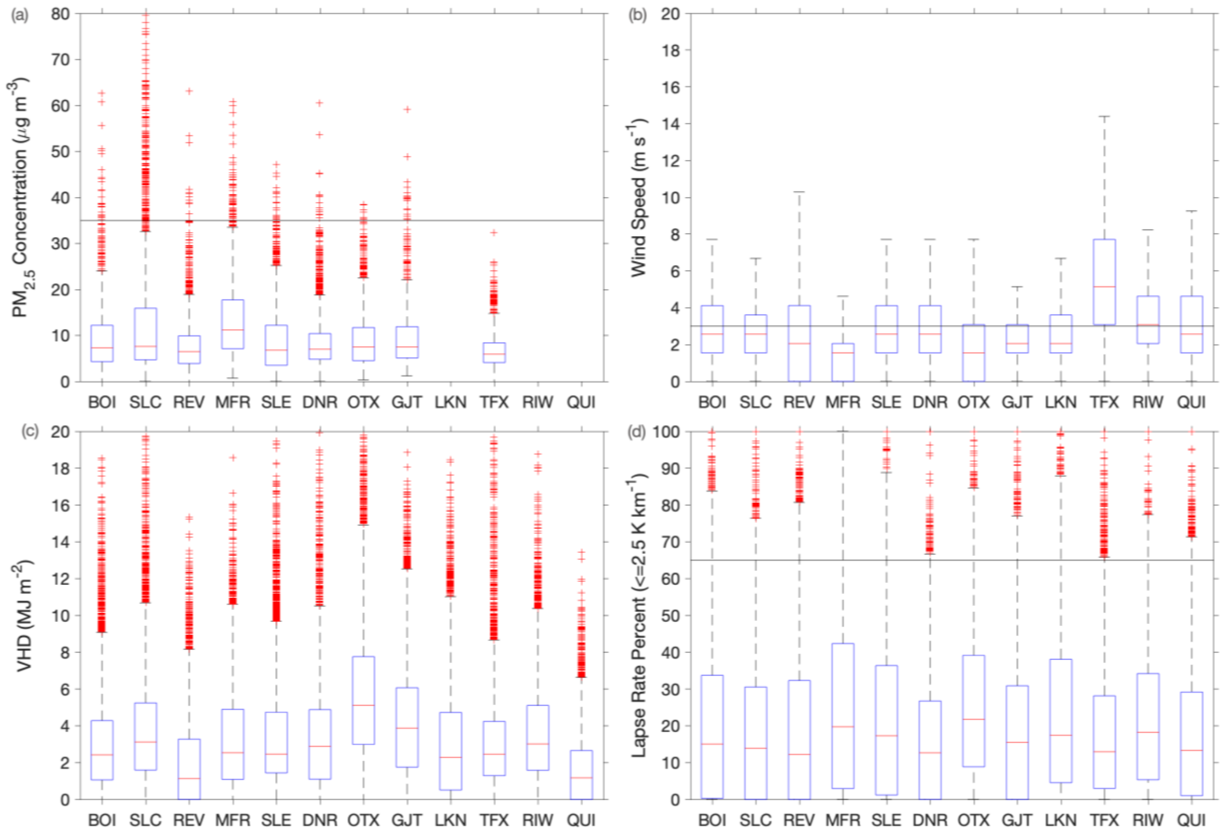

Figure 2. **Box Plots for Oct-Mar from 2002-2018.** (a) $PM_{2.5}$ concentrations line for 24-hr $PM_{2.5}$ NAAQS, (b) 10-m wind speed line for VCP threshold, (c) VHD, and (d) percentage of lapse rates less than 2.5 K km^{-1} in the lowest 1500 m MSL line for DSL threshold. Median (red line), mean (box middle), 75th and 25th percentiles (box top and bottom), 1.5 times the interquartile (dashed lines), and outliers (red plus).

The 10-m wind speed shown in Figure 2b is required in the classification criteria used in Yu et al., (2017) for VCP and the new VHD_met developed in this paper. The median wind speeds are similar at all locations ranging from 2 m s^{-1} to 3 m s^{-1}, except for Medford and Great Falls. Medford is the only location with a median 10-m wind speed less than 1m s^{-1}, which is related to its local topography that creates orographic barriers to the flow. Medford lies in the Rogue Valley with mountains ~800 m AGL in all directions. The downstream mountains have a higher elevation than the upstream mountains, which is favorable for stagnation at the surface in the Rogue Valley due to the blocking effect. CAP erosion by turbulent mixing from above plays a minor role in deep valleys, which have complex wind regimes (Zängl 2005). Great Falls and Reno have the greatest overall 10-m wind speed which can be attributed to the local meteorology influenced by terrain. The upstream mountains in both locations have a higher elevation than the downstream mountains and are oriented North-South, promoting down-sloping wind events due to the West-East oriented jet stream typically associated with an incoming trough.

The VHD for each valley in Figure 2c suggests that the bulk stability varies with each valley and is independent of $PM_{2.5}$ concentrations (only when $PM_{2.5}$ concentrations are compared across all locations). Reno, Elko, Great Falls, and Quillayute have the lowest VHD values. In arid locations, like Reno and Elko, the sensible heat flux can be an order of magnitude greater than the latent heat flux (Albertson et al., 1995). Upward surface sensible heat flux during the daytime tends to create rising parcels of air and enhance vertical mixing in the boundary layer (Holmes et al., 2015). Quillayute is on the Pacific Northwest Coast and experiences frequent synoptic systems during winter months and a lower VHD is expected, when compared to other locations surrounded by mountains that act as barriers.

Lapse rate percentages can be used to quantify the presence of a deep temperature inversion layer present near the surface. Medford and Spokane have the deepest layers of temperature inversions (Figure 2d) over the time period with the 75th percentile reaching ~40% for both locations. Denver and Quillayute have the lowest LR percentages and can be attributed to the local meteorology (lee cyclogenesis and lack of downstream mountains for Denver and active weather pattern for Quillayute). Denver lies on the front range of the Rockies, where polar continental airmasses often influence meteorology and push the cold airmass against the mountains. Quillayute lies on the coast of the Pacific Northwest, a common place for land falling troughs in the wintertime.

4.3 Cold Air Pool Classifications

Figure 3 shows number of soundings that reach the thresholds for VHD_17.5, VHD_75, VCP, and VHD_met at 12Z and 00Z. The number of days with potential persistent CAPs can be inferred from Figure 3 by comparing the number of afternoon (00Z) sounding occurrences to morning (12Z) occurrences. When the 00Z sounding threshold is met for any classification method, a stable atmosphere is present and is likely persisting from the previous morning (12Z) sounding. The 12Z soundings reached the CAP criteria more often than 00Z soundings in all four methods due to the greater atmospheric stability in the lower levels in the morning. The locations with a higher number of afternoon CAP occurrences implies that certain valleys may experience more persistent CAP events (i.e., less afternoon mixing during such events).

Table 3 shows the VHD thresholds used in Figure 3 to determine the number of soundings for CAP occurrences. Often, VHD_17.5 threshold, based on Whiteman et al., (2014)

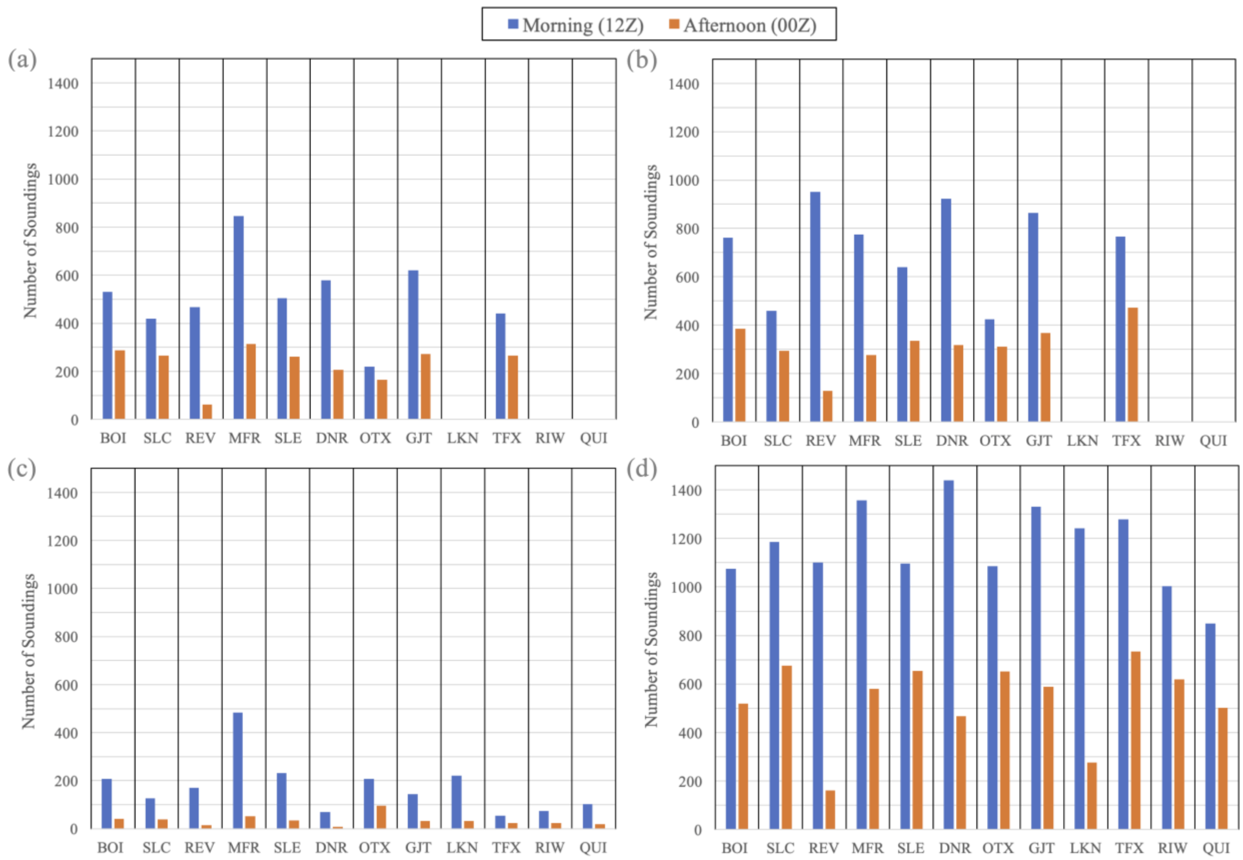

Figure 3. **CAP Occurrences for Oct-Mar from 2002-2018.** Number of CAP soundings for morning (12Z, blue) and afternoon (00Z, orange) sounding; 5,832 maximum soundings for each location. (a) VHD_17.5 based on Whiteman et al., (2014), (b) VHD_75 based on Pierce et al., (2019), (c) VCP based on Yu et al. (2017), and (d) VHD_met developed in this paper. *Note: Missing data for LKN, RIW and QUI due to lack of* $PM_{2.5}$ *data.*

is greater than the VHD_75 threshold. The exception is Medford, where the 75th percentile of $PM_{2.5}$ is greater than half of the NAAQS of 17.5 µg m^{-3}, at 17.7 µg m^{-3} so the VHD_75 threshold is larger than the VHD_17.5 threshold. Figure 3 illustrates the differences in CAP behavior across 12 locations. For example, Great Falls has the greatest number of afternoon soundings reaching the VHD_75 threshold, implying that the afternoon mixing is frequently limited. In addition, more PCAP events are possible in Great Falls because of the decreased likelihood of mixing in the afternoon compared to other locations.

The results shown in Figure 3 using the four CAP classification methods highlight the ambiguity of these classification methods, where each method provides a significantly different number of CAP occurrences in each location. The number of CAP occurrences using the VHD thresholds based on $PM_{2.5}$ values (VHD_17.5, VHD_75) are much higher than the number from the VCP method at each location. Another large difference comes from using air quality data

Table 3. **VHD Thresholds.** 75th percentile $PM_{2.5}$ concentrations and VHD thresholds for 2002-2018 in 12 cities. VHD_17.5: mean VHD for 24-hr $PM_{2.5}$ reaching half of the NAAQS (Whiteman et al., 2014), VHD_75: mean VHD for 24-hr $PM_{2.5}$ exceeding the 75th percentile (Pierce et al., 2019), VHD_met: median VHD based on local meteorology, MALR VHD: mean VHD using idealized lapse rate, and Normalized VHD: normalized VHD greater than 1.0 indicates that lapse rates are less than the MALR. *Missing values indicate no $PM_{2.5}$ monitors.*

Station ID	**BOI**	**SLC**	**REV**	**MFR**	**SLE**	**DNR**	**OTX**	**GJT**	**LKN**	**TFX**	**RIW**	**QUI**
$PM_{2.5}$ 75th Percentile (μg m^{-3})	12.2	15.9	9.9	17.7	12.2	10.4	11.7	11.9	--	8.4	--	--
VHD_17.5 (MJ m^{-2})	5.94	7.49	5.77	5.62	6.95	6.43	10.19	7.56	--	6.06	--	--
VHD_75 (MJ m^{-2})	4.92	7.22	4.04	5.93	6.14	5.26	8.36	6.57	--	4.65	--	--
VHD_met (MJ m^{-2})	4.03	4.40	3.49	3.92	3.96	4.12	6.88	5.23	4.41	3.97	4.71	2.72
MALR VHD (MJ m^{-2})	2.13	2.80	1.46	1.50	1.73	2.73	3.41	3.37	1.80	2.32	2.34	1.79
Normalized VHD	2.00	1.59	2.60	2.54	2.38	1.57	2.09	1.60	2.69	1.58	2.13	1.51

versus meteorology data to establish a VHD threshold. For example, Great Falls has a greater number of CAPs than Medford based on the VHD_met threshold (Great Falls n=1807 and Medford n=1144) while Medford has a greater number of CAPs based on the VCP criteria (Great Falls n=77 and Medford n=534). Therefore, the CAP classifications differ significantly when applying these criteria to different locations, even with similar synoptic meteorology. These differences will be highlighted in the following subsections.

4.3.1 VHD based on $PM_{2.5}$ NAAQS (VHD_17.5)

For all locations except Medford, VHD_17.5 identified fewer CAP soundings than VHD_75 since fewer locations have $PM_{2.5}$ concentrations greater than 17.5 µg m^{-3}. For the 12Z sounding observations, most locations reach 400-600 sounding occurrences (out of a possible 2,800) during the study period. This suggests that ~20% of the mornings in the Intermountain West experience stable conditions, while less than ~10% of the afternoons experience stable conditions. The VHD_17.5 thresholds in Table 3 are rather high, especially for Spokane, Salt Lake City and Grand Junction, which partly explains why there are fewer sounding occurrences despite previous studies that found an abundance of CAPs in these locations (Green et al., 2015; Whiteman et al., 2014; Yu et al., 2017).

To investigate this further the VHD thresholds from SLC can be used. The VHD_17.5 threshold in Table 3 for Salt Lake City, 7.49 MJ m-^{2}, is larger than the VHD threshold calculated in Whiteman et al., (2014), 4.04 MJ m^{-2}. There are two possible explanations for this. The first, is that in this study the sounding data does not have a fine scale vertical resolution (i.e., the radiosonde data were not interpolated to 10-m resolution intervals) and the integration height here was based on an average ridge height of 2439 m (instead of 2200 m). The second hypothesis is that $PM_{2.5}$ concentrations show a decreasing trend in the U.S. in recent years, thus requiring a higher VHD value for days with $PM_{2.5}$ concentrations larger than half of the NAAQS standard.

4.3.2 VHD using 75th Percentile $PM_{2.5}$ (VHD_75)

The 75th percentile $PM_{2.5}$ concentrations vary among the different valleys (Table 3). Reno, Denver and Great Falls have the lowest $PM_{2.5}$ concentrations, which is expected because the 10-m wind speed for these locations is greater (Figure 2b) than the other locations. Greater wind speed near the surface is indicative of more vertical mixing and dispersion of air pollutants near the surface, reducing the overall $PM_{2.5}$ concentrations. Medford has the highest 75th percentile $PM_{2.5}$ concentration along with the lowest 10-m wind speed. The data in Table 3 coincides with data presented in Figure 2, where there is a negative relationship between wind speed and $PM_{2.5}$ and a positive relationship between $PM_{2.5}$ and VHD.

In locations where $PM_{2.5}$ concentrations during the months from October to March are low (i.e., Reno and Great Falls), the mean VHD thresholds are low as well. This indicates that $PM_{2.5}$ concentrations that are considered to represent polluted days in Reno and Great Falls require lower bulk atmospheric stability than other locations. The opposite is true for Salt Lake City and Spokane, where a strong and/or deep stably stratified layer of atmosphere, indicated by a higher VHD, are required to have polluted days (Table 3). Implying that VHD thresholds for CAP classification based on air quality makes the comparison across locations ambiguous because local emissions can have a large impact on the results.

Figure 3d shows the number of sounding occurrences that reach the VHD_75 threshold. Reno has the least amount of afternoon soundings reaching the VHD_75 threshold, indicating that efficient afternoon mixing occurs and implies that persistent CAPs are less frequent than other locations. However, Reno has the greatest number of diurnal CAPs when compared to other locations, suggesting effective radiational warming during the early morning hours. In

Table 3, the VHD_75 for Reno is the lowest, but this value is higher than the majority of the afternoon soundings. Salt Lake City and Spokane have few morning CAP occurrences but high afternoon occurrences relative to the number of morning CAPs, likely due to a higher VHD_75 threshold value, and also indicates an increased presence of persistent CAPs in both locations.

4.3.3 Valley Cold Pool (VCP)

The CAP sounding numbers identified using the VCP method proposed by Yu et al., (2017) are shown in Figure 3c. Afternoon (00Z) soundings in Spokane reached the VCP criteria for 31.5% of the total soundings (n=302), while Denver 00Z soundings represent 11.5% (Figure 3c). This suggests that afternoon mixing is more likely in Denver while suppressed mixing occurs in Spokane, and other locations. Figure 3c shows that nocturnal inversions are common and dominant in the western U.S., which aligns with results based on the other CAP classification methods. In addition, Figure 3c shows a significant number of mornings with VCP criteria being met, while fewer afternoon soundings meeting the VCP criteria. The exceptions are Great Falls, Riverton and Quillayute which experience greater 10-m wind speed (Figure 2b), thus not meeting the VCP wind speed criteria of less than 3 m s^{-1} in the morning as well.

Medford total and morning VCP occurrences are greater than other locations by more than 200 sounding occurrences (n=534). The number of afternoon VCP occurrences in Medford is similar to other locations, possibly indicating the morning VCP occurrence could be due to valley meteorological differences when compared to other locations. 10-m wind speeds in Medford are low, with 75^{th} percentile wind speeds below the VCP threshold of 3 m s^{-1}, and deeper stable layers are observed more frequently in the morning (Figure 2b & 2d). VCP occurrences are strongly related to the deep stable criteria developed by Wolyn and Mckee (1989), apart from 10-m wind speed criteria discussed in Yu et al., (2017).

4.3.4 VHD with Meteorological Thresholds (VHD_met)

Identifying days with high surface pressure and decreased near-surface wind speed allows for identifying stagnant weather conditions associated with CAP events. Defining a VHD threshold that depends on physical variables that characterize the CAP provides a consistent method of classifying such events across valleys or locations. The median VHD_met thresholds for the 12 locations are similar, with most locations in a range of 4 ± 1 MJ m^{-2} (Table 3).

Spokane, Grand Junction, and Quillayute are outside this range with VHD thresholds of 6.88, 5.23 and 2.72 MJ m^{-2}, respectively.

This new method for classifying CAPs is independent of air pollution, thus there is less variability in the VHD_met thresholds (Table 3) and the number of CAPs (Figure 3d) across the 12 locations compared to the VHD_17.5 and VHD_75 methods. The number of soundings that meet or exceed the VHD_met threshold in each location are similar (Figure 3d) with a mean of 1706 soundings and a standard deviation of 241 soundings, when averaged across all locations. As expected, the morning (12Z) sounding frequently exceeds the VHD_met threshold due to decreased vertical mixing during the overnight and early morning hours, with fewer CAP occurrences in the afternoon (00Z) soundings. Salt Lake City, Spokane, and Great Falls have the greatest number soundings over the VHD_met thresholds in the 00Z soundings, indicating that persistent CAPs possibly persist through the diurnal cycle. Reno and Elko have the lowest number of 00Z soundings reaching their VHD_met thresholds, indicating the typical diurnal cycles likely inhibit persistent CAP formation in these locations.

4.4 Persistent Cold Air Pool (PCAP)

A PCAP is defined as a CAP that lasts for 3 or more consecutive twice-daily soundings (based on the definition in Whiteman et al., (2014)). Meaning a PCAP is a CAP that persists for more than one diel cycle, this same definition is used here and aligns with previous research (e.g., Yu et al., (2017) and Lareau et al., (2013)). PCAPs have significant health impacts because the CAP persists through the afternoon and leads to increasing pollutant concentrations over long periods of time. The number of PCAPs over the duration of this study varies significantly depending on the location and the CAP classification method (Figure 4). Typically, locations that are farther south and/or have an absence of downwind mountains observe fewer PCAPs, for all methods. The variability in the number of PCAPs identified by the different methods further illustrates the need for a universal classification method. The mean PCAP length determined by each method is shown in Figure 4b. The mean PCAP lengths between VHD_75, VHD_17.5, and VHD_met are similar, complementing previous research that synoptic meteorology influences PCAPs (Lareau et al., 2013; Green et al., 2015; Yu et al., 2017). However, there are subtle differences in PCAP length for each method. Where the VHD_met threshold method likely captures the PCAP length more appropriately because it is based on synoptic meteorology and does not rely on air quality

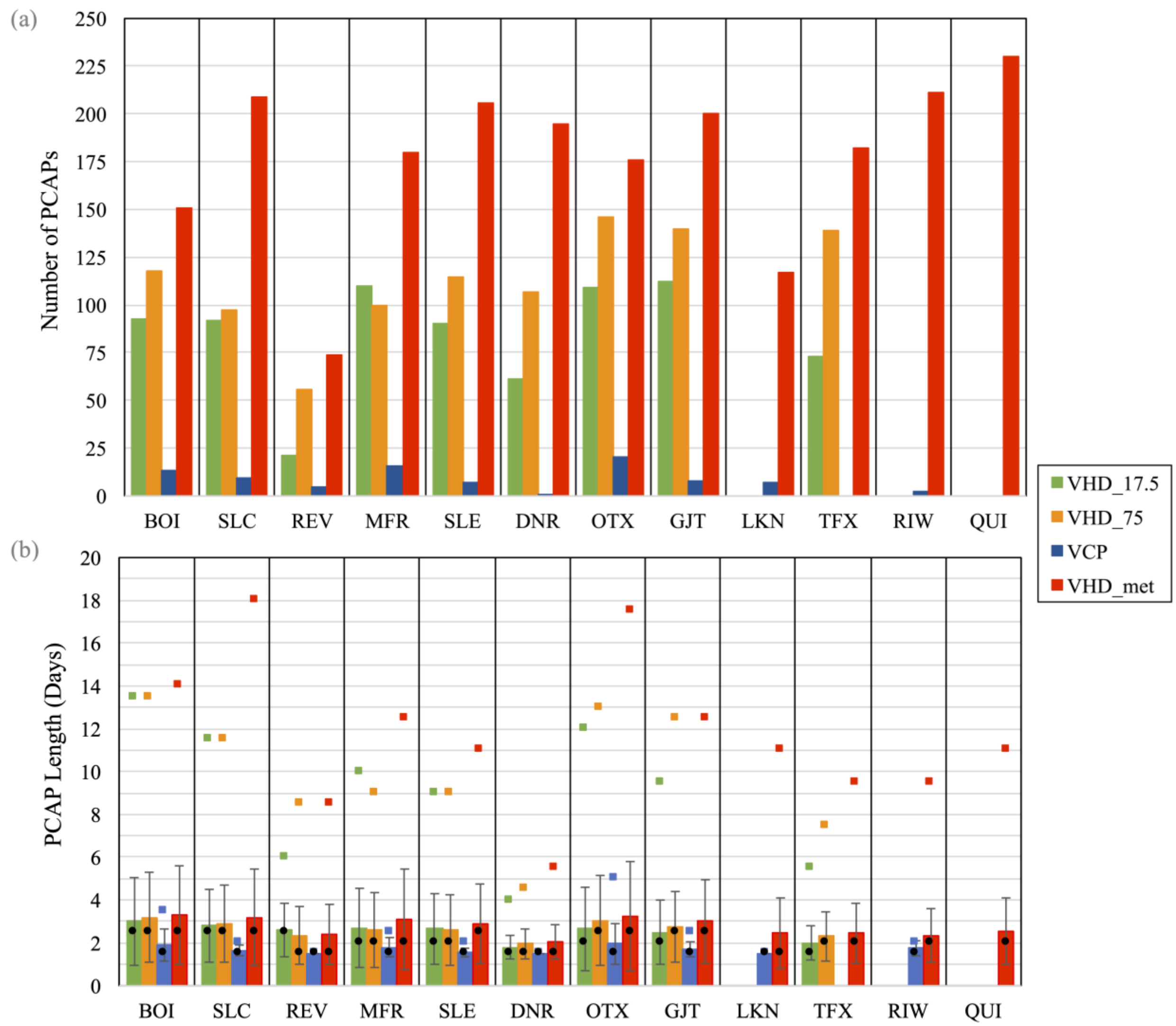

Figure 4**. PCAP Characteristics for Oct-Mar from 2002-2018**. (a) Total number of PCAPs based on VHD_75 from Whiteman et al., (2014), VHD_75 from Pierce et al., (2019), VCP from Yu et al., (2017), and the new VHD_met method. (b) Length of PCAPs in days; mean (bars), standard deviation (whiskers), median (dot), and maximum (square).

to establish a CAP threshold.

PCAP length changes with the method used, often varying by up to one day. The majority of the PCAP lengths for all locations are 4-6 consecutive soundings, or 2-3 days. Since $PM_{2.5}$ varies from year to year, VHD_17.5 may not adequately identify PCAP events over decades. The 75th percentile $PM_{2.5}$ concentration is less than half of the NAAQS at most locations, with the exception of Medford. Further, $PM_{2.5}$ data is not available in Elko, Riverton, and Quillayute, complicating the classification of PCAPs in those locations. When comparing results of the novel method (VHD_met) to the previous methods, the length of PCAPs are comparable, especially in Reno, Salt Lake City, Denver and Boise. When disregarding the method used by Yu et al., (2017), variability in the three methods is limited and indicates that the novel method can be useful for determining PCAP duration.

4.4.1 VHD_17.5 (Whiteman et al., 2014)

When using the VHD_17.5 threshold to determine the onset, duration, and breakup of PCAPs, most locations observe 80-100 PCAPs (e.g., total number of PCAPs from 2002-2018) with Reno, Denver and Great Falls having the least number of PCAPs (Figure 4a). Medford, Spokane, and Grand Junction observe up to 7 PCAPs per winter on average. Medford has the highest $PM_{2.5}$ concentrations on average while Spokane and Grand Junction have the greatest depth of integration, which may explain why these locations experience more PCAPs that other locations. Reno experiences the least number of PCAPs during the study period, this is attributed to fewer afternoon (00Z) soundings reaching the VHD_17.5 threshold (Figure 3a). Figure 4b shows that the PCAP length from VHD_17.5 varies only by a couple of days at most between all locations. Maximum PCAP lengths are not similar when comparing across all locations. This indicates that the local valley meteorology influences PCAP events, which is further discussed in the Supplemental Information.

4.4.2 VHD_75 (Pierce et al., 2019)

The total number of PCAPs increases using the VHD_75 method compared to the VHD_17.5 method for all locations, with the exception of Medford. The number of PCAPs per winter for most locations is approximately 7, with Spokane, Grand Junction and Great Falls nearing 9 PCAPs per winter. Reno still experiences the least amount of PCAPs, even with the lower VHD_75 threshold. In locations where the 75th percentile $PM_{2.5}$ concentration is close to 17.5 µg m^{-3}, the difference between the number of PCAPs using VHD_17.5 and VHD_75 is minimal (i.e., Medford and Salt Lake). Great Falls, Denver, and Reno have the lowest 75th percentile $PM_{2.5}$ concentrations and thus have significantly more PCAPs with VHD_75 than VHD_17.5 since the air pollution concentration threshold is lower using VHD_75 (i.e., requiring a lower VHD threshold on days when the 75th percentile of $PM_{2.5}$ is observed).

4.4.3 VCP (Yu et al., 2017)

Very few PCAPs are observed using the VCP method, likely due the absence of DSLs during PCAPs that maintain a depth of 975 m (65% of the lowest 1500 m AGL layer), especially during afternoon hours. Based on the VCP threshold, Boise and Medford have the greatest number of PCAPs, 13 and 16, respectively over the study period, far fewer than other methods. Other locations often have fewer than 10 PCAPs total, implying that PCAPs occur in most valleys once or twice every other year when using VCP to classify PCAPs.

When the lapse rate percentage of the lower 1500 m AGL is lowered to 30-40% (500-600 m layer depth AGL), the number of PCAPs in each location is comparable to the other methods discussed in this paper. Differences lie in the average length of the PCAPs when compared to the other methods. For example, when the lapse rate percentage is maintained at 65%, the average PCAP length is anywhere between 3-4 consecutive soundings. However, when the lapse rate percentage is decreased to ~35%, the average PCAP length increases but is still less than the average PCAP length calculated using the other methods. PCAP length using this method are expected to be less than other methods because of the DSL requirement. PCAPs typically change depth over the duration of a single event and using a criterion dependent on depth filters out shallow stable layers below the ridge height.

4.4.4 VHD_met

The total number of PCAPs is the greatest for the new meteorological classification (VHD_met) method (red bars in Figure 4). Since this method relies on synoptic meteorology, some conformity across all locations can be observed. Meaning that the results across locations are more similar using VHD_met, compared to the VHD based on air pollution, because the large-scale synoptic conditions occur over the entire western U.S. impacting multiple locations at the same time. When comparing the number of PCAPs in Elko, Riverton, and Quillayute, to the number of PCAPs in other locations with similar climates (e.g., Spokane, Reno) using this method they are now comparable. Elko and Reno have the lowest number of PCAPs, while locations farther north typically experience more PCAPs. Quillayute observed the greatest number of PCAPs, likely due to a low VHD_met threshold being easily reached during stagnant periods and/or elevated inversions, which can also be met when marine boundary layers are present. Additional investigation is required to separate the CAP events over mountains and marine ABLs using this method.

4.5 Normalized VHD for Comparison Across Locations

To create a CAP classification metric that is comparable across multiple locations, the MALR VHD is used to normalize the VHD calculated from each radiosonde. Because the physical interpretation of the VHD magnitude varies across locations, the VHD normalized with an idealized MALR VHD provides a unitless VHD with a uniform scale for comparison. This idealized profile uses the MALR to calculate an idealized VHD. The MALR VHD is location specific and changes depending on the depth of the valley. For example, Medford has a mean MALR VHD of 1.5 MJ m^{-2} with a standard deviation of less than 0.2 MJ m^{-2}, close to values reported in Whiteman et al., (2014) for Salt Lake City. Grand Junction has a MALR VHD of 3.37 MJ m^{-2} with a standard deviation of 0.3 MJ m^{-2}. The valley depth in Medford is lower than Grand Junction (438 m difference), resulting in a larger MALR VHD (i.e., deeper integration heights increase the MALR VHD). Locations with greater integration depth (e.g., Spokane depth is ~1250 m AGL) have a greater MALR VHD because the integration includes more data points and those data points have a higher vertical resolution than at lower altitudes. When comparing valley integration depth to the MALR VHD they are correlated with an r^2 of 0.85, as integration increases, the MALR VHD also increases. Once the normalized VHD values are obtained for each radiosonde the new CAP classification method is applied at each location to calculate a new normalized VHD_met threshold (Normalized VHD in Table 3).

Using a standard MALR of 0.0065°C m^{-1}, the MALR VHD is a general VHD value that represents a normal stability for a standard atmosphere without the presence of a CAP. Therefore, normalized VHD values greater than 1.0 for any radiosonde indicate a more stable boundary layer. The normalized VHD thresholds calculated for each location range from 1.51 in Quillayute to 2.60 in Reno and are shown in Table 3. Places with climatologically higher VHD (Figure 2c), Spokane and Grand Junction, have normalized VHD values of 2.09 and 1.6, respectively, which are comparable to the normalized VHD thresholds in the other locations. When the normalized VHD is greater than the threshold values listed in Table 3, there is a CAP (i.e., diurnal or persistent) and when three consecutive soundings reach this threshold, a PCAP is identified.

The normalized VHD enables comparisons across locations. It can be used to identify CAP events over large spatial scales and to compare the duration and timing of the events across

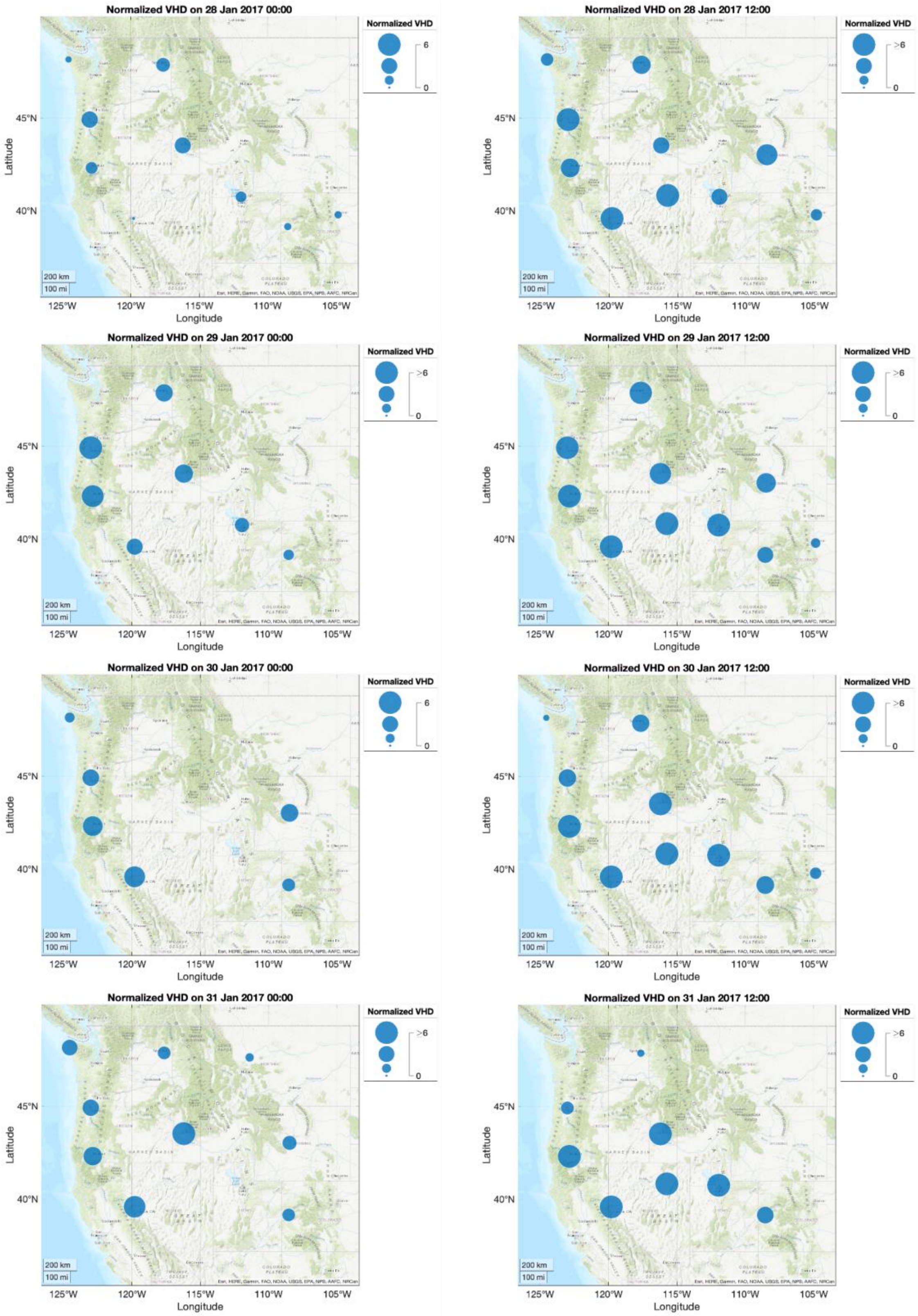

Figure 5**. Normalized VHD Values for CAPs.** Normalized VHD in each location for soundings that were above the normalized VHD threshold value, indicating a CAP for 28-31 January 2017 UTC. (*Left*) Afternoon soundings, 00Z. (*Right*) Morning soundings, 12Z. Consecutive soundings indicate a PCAPs.

locations. For example, Figure 5 shows the spatial distribution of the normalized VHD for soundings over the CAP threshold (Table 3) for four days in January 2017. It shows that large scale processes (synoptic meteorology) cause CAP formation in several valleys across the western U.S., where six to ten locations in each plot have a CAP. Additionally, the presence of PCAPs can be seen in eight locations over the four-day period (Quillayute, Medford, Salem, Spokane, Boise, Reno, Salt Lake City, and Grand Junction).

5. Discussion

CAP formation in the mountainous areas over the western U.S. occurs when there is a 500 hPa ridge system and surface cold anomaly based on climatology analysis. In previous studies, the two CAP classification criteria from literature were applied to the SLV, but only broadly to other valleys (e.g., Green et al., (2015); Yu et al., (2017)). Large scale studies have shown that CAPs occur in other valleys; however, field campaigns outside of the SLV and Columbia River Basin are rare and potentially inhibits the ability to understand the physical processes that influence CAPs in other valleys. Whiteman et al., (2001) found that PCAPs in the lower Columbia River Basin persisted for a median duration of 28.5 hours from 1989-1999 using observational station data. Local topography like upwind/downwind mountain heights, valley shape and predominant wind direction influence CAP events. Wind roses during CAP events and non-CAP events above the average ridge height provides insights to why some locations experience deeper stable layers, greater bulk stability and possibly the duration of a CAP event (see Supplemental Information for wind roses). For example, Medford typically experiences a light northerly wind at the surface and a SE wind at the average ridge height during the winter months and especially during PCAPs.

Persistent CAPs lasting longer than one diurnal cycle may be partially enhanced by radiative effects, but are predominantly forced by large-scale subsidence and/or mid-level warm air advection (Wolyn and McKee 1989; Whiteman et al., 1999, 2001; Zhong et al., 2001; Zängl 2005; Hoggarth et al., 2006; Reeves and Stensrud 2009). This is especially evident in Figure 4a with all locations experiencing a similar number of PCAPs, their proximity to large-scale subsidence can be attributed to this. The predominance of large-scale patterns that lead to these events are important to understand. The criteria discussed in this paper are mostly associated with large-scale meteorological patterns that contribute to CAPs. It should be noted that local meteorology is important for the duration and strength of the CAP, as seen in Medford, OR and Reno, NV, where deep CAPs and more shallow CAPs are dominant. Brief discussions of the local meteorological conditions at each location are given in the Supplemental Information to illustrate the influence of these factors on CAP events.

Snow impacts influence CAP behavior and can extend PCAP length. Increased surface albedo due to snow cover impacts the surface energy balance and causes more of the incoming solar radiation to be reflected from the surface. Therefore, there is less net radiation at the surface and a lower surface sensible heat flux to transfer heat from the ground to the boundary layer, which dampens the mixing and can lead to increased $PM_{2.5}$ concentrations. Green et al., (2015) found that snow-covered CAP days (22%) were nearly four times greater than no snow-covered CAP days (5.3%). Additional information about the snow impacts in specific locations are included in the local meteorological discussions in the Supplemental Information.

6. Summary

CAPs studies can be biased when using radiosonde data that comes from radiosonde launches released above the valley floor because the data does not include the lowest portion of the atmosphere where a strong stable layer can be present. Appending ASOS meteorological data on the valley floor to radiosonde data resulted in a more representative VHD to quantify the stability of the atmosphere. Stability can be underestimated by as much as 50% for locations where the radiosonde is released >150 m above the valley floor. More CAP events are identified when accounting for the vertical structure of the entire valley depth.

This paper uses four methods (VHD_17.5, VHD_75, VCP, and VHD_met) to classify CAP events in the western U.S. using observational data from radiosondes and ASOS data where needed. The new metric introduced in this paper (VHD_met) suggests that criteria related to synoptic meteorology results in similar CAP VHD threshold values across the western U.S., especially where valley depths are similar. Previous classification methods (VHD_17.5 and VHD_75) based on air quality have CAP VHD threshold values that vary significantly by location and are influenced by air pollution emissions. Calm winds near the surface, high pressure, mid-level warming, higher RH, and snow cover are characteristics of a PCAP, regardless of population. While air pollution increases during a PCAP, the increased $PM_{2.5}$ concentrations are valley specific and have the potential to change significantly across decades.

The new CAP classification metric provides a method for classifying PCAPs in any location that has vertical profiles of meteorology data. In addition, this new method provides a way for health officials, emergency managers, and government officials to determine when CAP events are present to further understand atmospheric chemistry, adverse health impacts, and more during these events. Additional valley specific meteorology needs further investigation including drainage flows, differences in why diurnal CAPs do not evolve into persistent CAPs, PCAPs that are resilient, and vertical profiles of turbulent fluxes to understand boundary layer mixing.

Acknowledgements

The authors would like to thank former University of Nevada, Reno graduate students Dr. S. Marcela Loría-Salazar and KC King for their assistance with troubleshooting and developing python codes to process the large amounts of radiosonde data used in this study.

References

Albertson, J.D., Parlange, M.B., Katul, G.G., Chu, C.-R., Stricker, H., & Tyler, S., 1995. Sensible heat flux from arid regions: A simple flux-variance method. *Water Resources Research* **31**, 969–973. https://doi.org/10.1029/94WR02978

Baasandorj, M., Hoch, S. W., Bares, R., Lin, J. C., Brown, S. S., Millet, D. B., Martin, R., Kelly, K., Zarzana, K. J., Whiteman, C. D., Dube, W. P., Tonnesen, G., Jaramillo, I. C., & Sohl, J. 2017. Coupling between chemical and meteorological processes under persistent cold-air pool conditions: Evolution of wintertime $PM_{2.5}$ pollution events and N_2O_5 observations in Utah's Salt Lake Valley. *Environmental Science and Technology* **51,** 5941–5950.

Clay, K. & Muller, N., 2019. Recent Increases in Air Pollution: Evidence and Implications for Mortality (No. w26381). National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w26381

Deardorff, J.W., 1970. Convective velocity and temperature scales for the unstable planetary boundary layer and for Rayleigh convection. *Journal of Atmospheric Science*. **27**, 1211–1213. https://doi.org/10.1175/1520-0469(1970)027<1211:CVATSF>2.0.CO;2

Franchin, A., Fibiger, D. L., Goldberger, L., McDuffie, E. E., Moravek, A., Womack, C. C., Crosman, E. T., Docherty, K. S., Dube, W. P., Hoch, S. W., Lee, B. H., Long, R., Murphy, J. G., Thornton, J. A., Brown, S. S., Baasandorj, M. & Middlebrook, A. M. 2018. Airborne and ground-based observations of ammonium-nitrate-dominated aerosols in a shallow boundary layer during intense winter pollution episodes in northern Utah. *Atmospheric Chemistry and Physics* **18,** 17259–17276.

Green, M.C., Chow, J.C., Watson, J.G., Dick, K., & Inouye, D., 2015. Effects of snow cover and atmospheric stability on winter PM2.5 concentrations in western U.S. valleys. *Journal of Applied Meteorology and Climatology*. **54**, 1191–1201. https://doi.org/10.1175/JAMC-D-14-0191.1

Hoggarth, A.M., Reeves, H.D., & Lin, Y.-L., 2006. Formation and Maintenance Mechanisms of the Stable Layer over the Po Valley during MAP IOP-8. *Monthly Weather Review* **134**, 3336–3354. https://doi.org/10.1175/MWR3251.1

Holmes, H.A., Sriramasamudram, J.K., Pardyjak, E.R., & Whiteman, C.D., 2015. Turbulent fluxes and pollutant mixing during wintertime air pollution episodes in complex terrain. *Environmental Science and Technology*. **49**, 13206–13214. https://doi.org/10.1021/acs.est.5b02616

Horel, J., Splitt, M., Dunn, L., Pechmann, J., White, B., Ciliberti, C., Lazarus, S., Slemmer, J., Zaff, D., & Burks, J. 2002 Mesowest: Cooperative mesonets in the western United States. *Bulletin of the American Meteorological Society* **83**, 211-226.

Ivey, C. E., Balachandran, S., Colgan, S., Hu, Y. & Holmes, H. A. 2019. Investigating fine particulate matter sources in Salt Lake City during persistent cold air pool events. *Atmospheric Environment* **213,** 568–578 (2019).

Jemmett-Smith, B.C., Ross, A.N., Sheridan, P.F., Hughes, J.K., & Vosper, S.B., 2019. A case-study of cold-air pool evolution in hilly terrain using field measurements from COLPEX. *Quarterly Journal of the Royal Meteorological Society* **145**, 1290–1306. https://doi.org/10.1002/qj.3499

Lareau, N.P., Crosman, E., Whiteman, C.D., Horel, J.D., Hoch, S.W., Brown, W.O.J., & Horst, T.W., 2013. The Persistent Cold-Air Pool Study. *Bulletin of the American Meteorological Society* **94**, 51–63.

Lareau, N.P. & Horel, J.D., 2015a. Dynamically Induced Displacements of a Persistent Cold-Air Pool. *Boundary-Layer Meteorology* **154**, 291–316.

Lareau, N.P. & Horel, J.D., 2015b. Turbulent Erosion of Persistent Cold-Air Pools: Numerical Simulations. *Journal of the Atmospheric Sciences* **72**, 19.

NOAA, 1976: U.S. Standard Atmosphere, 1976. NOAA Tech. Rep. S/T 76-1562, National Oceanic and Atmospheric Administration, Washington, D.C. [Available online at https://ntrs. nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19770009539.pdf].

Pierce, A.M., Loría-Salazar, S.M., Holmes, H.A., & Gustin, M.S., 2019. Investigating horizontal and vertical pollution gradients in the atmosphere associated with an urban location in complex terrain, Reno, Nevada, USA. *Atmospheric Environment* **196**, 103–117. https://doi.org/10.1016/j.atmosenv.2018.09.063

Pope III, C.A., 1991. Respiratory Hospital Admissions Associated with PM10 Pollution in Utah, Salt Lake, and Cache Valleys. *Archives of Environmental Health: An International Journal* **46**, 90–97. https://doi.org/10.1080/00039896.1991.9937434

Pope III, C.A., Muhlestein, J.B., May, H.T., Renlund, D.G., Anderson, J.L., & Horne, B.D.,

2006. Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution. *Circulation* **114**, 2443–2448.

Reeves, H.D. & Stensrud, D.J., 2009. Synoptic-Scale Flow and Valley Cold Pool Evolution in the Western United States. *Weather and Forecasting* **24**, 1625–1643.

Shin, S.-H. & Ha, K.-J., 2009. Implementation of turbulent mixing over a stratocumulus-topped boundary layer and its impact in a GCM. *Advances in Atmospheric Sciences* **26**, 995–1004. https://doi.org/10.1007/s00376-009-8145-0

Silcox, G.D., Kelly, K.E., Crosman, E.T., Whiteman, C.D., & Allen, B.L., 2012. Wintertime $PM_{2.5}$ concentrations during persistent, multi-day cold-air pools in a mountain valley. *Atmospheric Environment* **46**, 17–24.

Sun, X. & Holmes, H. Surface turbulent fluxes during persistent cold-air pool events in the Salt Lake Valley, Utah. Part I: Observations. *Journal of Applied Meteorology and Climatology* **58,** 2553–2568 (2019).

VanReken, T. M., Dhammapala, R. S., Jobson, B. T., Bottenus, C. L., VanderSchelden, G. S., Kaspari, S. D., Gao, Z., Zhu, Q., Lamb, B. K., Liu, H., & Johnston, J. 2017. Role of persistent low-level clouds in mitigating air quality impacts of wintertime cold pool conditions. *Atmospheric Environment* **154**, 236–246.

Whiteman, C.D., Bian, X., & Zhong, S., 1999. Wintertime evolution of the temperature inversion in the Colorado Plateau Basin. *Journal of Applied Meteorology* **38**, 1103–1117. https://doi.org/10.1175/1520-0450(1999)038<1103:WEOTTI>2.0.CO;2

Whiteman, C.D., Hoch, S.W., Horel, J.D., & Charland, A., 2014. Relationship between particulate air pollution and meteorological variables in Utah's Salt Lake Valley. *Atmospheric Environment* **94**, 742–753. https://doi.org/10.1016/j.atmosenv.2014.06.012

Whiteman, C.D., Zhong, S., Shaw, W.J., Hubbe, J.M., Bian, X., & Mittelstadt, J., 2001. Cold pools in the Columbia Basin. *Weather and Forecasting* **16**, 432–447.

Wilson, T.H. & Fovell, R.G., 2018. Modeling the Evolution and Life Cycle of Radiative Cold Pools and Fog. *Weather and Forecasting* **33**, 203–220. https://doi.org/10.1175/WAF-D-17-0109.1

Wolyn, P.G. & McKee, T.B., 1989. Deep Stable Layers in the Intermountain Western United States. *Monthly Weather Review* **117**, 461–472.

Yu, L., Zhong, S., & Bian, X., 2017. Multi-day valley cold-air pools in the western United States as derived from NARR. *International Journal of Climatology* **37**, 2466–2476. https://doi.org/10.1002/joc.4858

Zängl, G., 2005. Wintertime Cold-Air Pools in the Bavarian Danube Valley Basin: Data Analysis and Idealized Numerical Simulations. *Journal of Applied Meteorology and Climatology* **44**, 1950–1971.

Zanobetti, A., Schwartz, J., Samoli, E., Gryparis, A., Touloumi, G., Peacock, J., Anderson, R.H., Le Tertre, A., Bobros, J., Celko, M., Goren, A., Forsberg, B., Michelozzi, P., Rabczenko, D., Hoyos, S.P., Wichmann, H.E., & Katsouyanni, K., 2003. The temporal pattern of respiratory and heart disease mortality in response to air pollution. *Environmental Health Perspectives* **111**, 1188–1193.

Zhong, S., Whiteman, C.D., Bian, X., Shaw, W.J., & Hubbe, J.M., 2001. Meteorological processes affecting the evolution of a wintertime cold air pool in the Columbia Basin. *Monthly Weather Review* **129**, 2600–2613.