

Data Filling of Micrometeorological Variables in Complex Terrain for high-resolution nowcasting

³ Nipun Gunawardena, Eric Pardyjak*, Pierre Durand[†], Thierry Hedde[‡], and Florian Dupuy[§]

University of Utah, Salt Lake City, Utah

Should we include your code
in a supplementary material?

⁵ *Corresponding author address: Eric Pardyjak, University of Utah, 1495 E 100 S (1550 MEK),

⁶ SLC, UT 84112

⁷ E-mail: pardyjak@eng.utah.edu

⁸ [†]University of Toulouse III

⁹ [‡]Current affiliation: CEA, Cadarache, France

¹⁰ §Current affiliation: CEA, Cadarache, France

ABSTRACT

In this paper, two different methods for nowcasting/data filling spatially varying meteorological variables (wind velocity components, specific humidity, and virtual potential temperature) at the ~~microscale~~ in regions marked by complex terrain are compared. Multivariable linear regression and artificial neural networks are used to predict micrometeorological variables at ~~multiple location (up to eight)~~ ^{a small set (usually 3) of} one weather station using the measurements from ~~other~~ nearby weather stations. The models are trained using data gathered from a system of eleven automated weather stations deployed in the Cadarache Valley of southeastern France from December 2016 to June 2017. It is found that the linear regression method can be used to nowcast almost as well as artificial neural networks. It is also found that the neural networks have no sensitivity to the input stations. These methods can be used to estimate gaps in incomplete datasets or to perform in-depth field experiments using limited equipment.

Covering scales
ranging from $\sim 10^0 - 8 \text{ km}$

please include
some quantitative
results on
performance
- also any
challenges
that remain

24 **1. Introduction**

25 To study the meteorology of complex terrain such as urban or mountainous regions, researchers
26 often conduct field experiments. During these field experiments, a large amount of data is typically
27 collected with, but not limited to: tethered balloons, radiosondes, manned and unmanned aircraft,
28 remote sensing instruments (LIDAR, SODAR, RASS, etc.), meteorological towers, and small,
29 distributed weather stations (Emeis 2010). Field experiments typically last from a few weeks
30 (Miller et al. 2016), to a few months (Duine et al. 2016a), to even a few years (Philippopoulos
31 and Deligiorgi 2012). Regardless of the instrumentation used, it is often removed after the field
32 experiment is completed, ~~obviously~~ eliminating the ~~ability to make observations~~ ^{ability to make observations}
33 However, in many cases, the field experiment is conducted in an area that has permanent weather
34 stations installed. For example, the MATERHORN experiment (Fernando et al. 2015) at Dugway
35 Proving Grounds ⁱⁿ Utah, the BLLAST experiment in southern France (Lothon et al. 2014), and the
36 KASCADE experiment described herein, were all field experiments where the scientific equipment
37 used supplemented permanent ^{operational} weather stations. In such an instance, the end of the field experiment
38 doesn't necessitate a cessation of measurements, just a substantial reduction ~~of measurements~~. In
39 other words, the field experiment offers the opportunity to make dense measurements, while the
40 operational instrumentation deployment offers sparse measurements. ^{why? Inverse.}
41 In a situation such as this, it is often desirable to continue to get dense measurements after the
42 field experiment is over. While it is obviously impossible to get true meteorological measure-
43 ments from an area where a sensor is no longer deployed, the process of *nowcasting* can allow for
44 approximations of what the measurements would be. In the meteorological sense, nowcasting is
45 a process that describes the current and near-future (hourly scale) state of the atmosphere (Mass
46 2012). The process of nowcasting is also used in the field of economics, where it is defined "as the

47 prediction of the present, the very near future, and the very recent past" (Banbara et al. 2013). In
48 this paper, we present and compare two methods of nowcasting dense atmospheric measurements
49 in complex terrain using only a sparse set of deployed sensors. The two methods are multiple
50 variable linear regression and artificial neural networks, and we are specifically predicting present
51 values.

52 **2. Background**

53 There are two main reasons to study the meteorology of complex terrain. The first is to increase
54 our understanding of the flow physics associated with complex terrain that inherently violate many
55 of the meteorological assumptions usually invoked in models and analysis. However, the second

*such as
with refraction.*

1 56 main reason to study the meteorology of complex terrain is to improve ~~our~~ weather prediction
57 capabilities (Chow et al. 2013). For example, since over half of the worlds population lives in
58 cities (Fernando 2010), and air pollution is considered a serious human health risk (Kampa and
59 Castanas 2008), it is often important to predict air pollution concentrations in cities (Chow et al.
60 2013). This can involve studying phenomena such as cold air pools (Lareau et al. 2013) and valley
61 flows (Mahrt 2017), both of which affect pollution levels and fog formation (Hang et al. 2016).

62 Another reason to improve weather prediction in complex terrain is to better predict snow and
63 ice storms (Chow et al. 2013). Adams et al. ~~(Adams et al. 2004)~~ estimate that improved snow
64 prediction can potentially produce 1.3 billion US dollars of benefit annually, in addition to the
65 number of lives saved due to prevented accidents. It is important to note that the atmosphere
66 in complex terrain often exhibits high spatial and temporal variability (Acevedo and Fitzjarrald

1 67 2001; LeMone et al. 2003), which can make weather prediction difficult at ~~the~~ microscale. These

1 68 ~~microscale~~ difficulties contribute to weather prediction difficulties at larger scales, which is why
69 boundary-layer meteorology is necessary to improve large-scale weather models (Stull 2012).

70 While these are very broad reasons to study the meteorology of complex terrain, there are some
71 focused applications as well. The agricultural community is interested in airborne pollen and
72 pathogen transmission, evapotranspiration, as well as small-scale microclimate and its effects on
73 crops (Bailey et al. 2014; Mahaffee and Stoll 2016; Lin 2007; Monteith et al. 1991). The nu-
74 clear energy community is often required to consider the side-effects of a breach, and how the
75 contaminants spread (Duine et al. 2016a; Baskett et al. 1991; Stohl et al. 2012). The US military
76 is interested in dust transport and its effects on military equipment (Wilkerson 1991), while the
77 U.S. Environmental Protection Agency (EPA) is interested in dust transport because of air quality and health
78 (Pardyjak et al. 2008; Petroff et al. 2008). Finally, urban planners are interested in reducing pollu-
79 tion (Lateb et al. 2016), increasing pedestrian comfort (Shi et al. 2015), and installing clean energy
80 sources (Karthikeya et al. 2016). A good high-level overview of the benefits of complex terrain
81 meteorology studies can be found in Chow et al.'s review (Chow et al. 2013).

9 A method that is also important to meteorology (and other fields) is nowcasting. As mentioned
82 previously, nowcasting is a process that describes the current and near-future (hourly scale) state of
83 the atmosphere (Mass 2012). Nowcasting has been used in many different areas of meteorology.
84 Both Xu et al. (Xu et al. 2005) and Novak (Novak 2007) have presented systems to nowcast storms
85 using radar, and Wilson et al. (Wilson et al. 1998) published a review of storm nowcasting methods
86 prior to 1998. Rasmussen et al. (Rasmussen et al. 2001) created a nowcasting system to aid in
87 de-icing decision making, and Gultepe et al. (Gultepe et al. 2014) attempted to use nowcasting to
88 predict ice fog, though they were not successful. Yates et al. (Yates et al. 2000) used a precipitation
89 nowcasting system to improve flash flood prediction. Demetriades et al. (Demetriades and Holle
90 2006) used a lightning strike detection network to nowcast information about tropical cyclones.
91 Han et al. (Han et al. 2017) used support vector machines to nowcast radar reflectivity data,
92 showing that machine learning has a place in meteorological nowcasting.

94 The differences between the nowcasting systems presented in this paper, and those presented
95 above are significant. The nowcasting systems above are applied to large scales, sometimes over
96 hundreds of kilometers. The nowcasting systems presented here are designed to be used at small
97 scales, of the order of a few ^{*hundred meters to several*} kilometers maximum. Another large difference is the time scale.

98 All the aforementioned studies nowcast in the sense of predicting the near future, whereas our
99 systems predict present values. In addition, many of the previous studies use radar as inputs to their
100 nowcasting systems, whereas we use averaged data from distributed sensor stations. Therefore, it
101 is important to review nowcasting systems in the literature which use distributed sensor stations or
102 wireless sensor networks.

103 Distributed sensor stations ~~or~~ ^{and} wireless sensor networks (WSNs) are systems of small meteo-
104 rological stations that are easily deployed en masse, and provide excellent spatial coverage of a
105 given area. They have been used in wildlife habitat monitoring (Mainwaring et al. 2002), forest
106 fire detection (Yu et al. 2005), structural health monitoring (Kim et al. 2007), and more (Akyildiz
107 et al. 2002). Distributed sensor systems and WSNs have been used for nowcasting. Furquim et al. (2014)

108 ~~(Furquim et al. 2014)~~ used wireless sensor networks and machine learning techniques to forecast
109 flash floods. They too consider nowcasting to be predicting the near future, instead of the present

110 ~~as done here. De Haan et al. (De Haan and Stoffelen 2012)~~ use meteorological data derived from
111 deployed aircraft to improve measured wind fields, and therefore improve nowcasting techniques.

112 While aircraft are not traditional distributed sensor systems, they are deployed in great numbers
113 and offer very high spatial coverage. Furquim et al. ~~(Furquim et al. 2014)~~ used a WSN and several
114 different machine learning techniques, including artificial neural networks, to nowcast flash floods
115 in São Paulo, Brazil.

116 As outlined in Shalev-Shwartz and Ben-David's book ~~(Shalev-Shwartz and Ben-David 2014)~~,
117 artificial neural networks (ANNs) are a class of biologically-inspired algorithms that can be used

for regression or classification tasks. While there are many types of ANNs, the simplest is the multilayer perceptron, also known as a standard feedforward neural network. A feedforward neural network is a directed acyclic graph where the nodes are called *neurons*, and the edges are called *connections*. Each neuron has an associated *bias* and *activation function*. Each connection has an associated *weight*. The graph is organized into layers, where each layer contains a certain amount of neurons and is connected to the layer above and below it. The bottom most layer is called the *input layer*, and any gathered data are passed into the neurons of the input layer. The input layer has as many neurons as there are *input variables*. The topmost layer is called the output layer, and is the output of the entire feedforward network. There can be multiple outputs, and the number of neurons in the output layer is equal to the number of *target variables*. The layers in-between the input layer and the output layer are called the *hidden layers*. Therefore, data are fed into the input layer; its output is then fed into the first hidden layer, whose output goes into the second hidden layer, and so on, until the output from the last is sent to the output layer. The output of any given neuron is the weighted sum of all the neurons in the layer before it, passed through the activation function. The values of the weights and biases are found by minimizing the network's mean squared error, in a process known as *backpropagation*. The process of finding the mean squared error using known data is called *training*, while using the ANN to predict unknown values is called *testing*. More information can be found in the previously cited book by Shalev-Shwartz and Ben-David (2014).

While ANNs can be used for classification purposes (e.g., Furquim et al. (Furquim et al. 2014) used an ANN to determine whether certain conditions would lead to a flash flood or not), in this paper and in many other nowcasting contexts, ANNs are used for the purpose of regression. For example, French et al. (French et al. 1992) used an ANN to spatially nowcast rainfall up to one hour ahead. Oztopal et al. (Oztopal 2006) used an ANN to predict the wind speed at one station, given the wind

142 speeds at nine different stations. While this is similar to the work presented in herein, Oztopal et
(2006)
143 al. did it on a much larger spatial scale. Their stations were distributed on a scale of hundreds of
144 kilometers, and our work focuses on scales of a few kilometers. *✓ already mentioned several times*
n (2012)
145 Philippopoulos et al. (~~Philippopoulos and Deligiorgi 2012~~) also used ANNs to predict wind
146 speeds, though this time on the spatial scale of tens of kilometers. They compared the ANN's
147 performance to several spatial interpolation methods. In this paper, we compare our ANN perfor-
148 mance to a multiple linear regression performance, which they do not do. In addition, we show
149 that an ANN can predict variables other than wind speed, on a much shorter time scale.
(2000)

150 Benvenuto et al. (~~Benvenuto and Marani 2000~~) use an air quality monitoring network in Mestre,
151 Italy, and ANNs to nowcast pollution concentrations in the near future and to interpolate missing
152 data. While Benvenuto's work is similar to ours, there are some key differences. Benvenuto
153 nowcasts for the near future (one hour ahead and three hours ahead), while we predict the present
154 values. Benvenuto et al. also interpolates missing data using an ANN that predicts forward in
155 time. Here, we show that the ANN doesn't have to predict forward in time to interpolate missing
n
156 data. Finally, Benvenuto et al. used data gathered in an urban area, whereas we used data gathered
valley
157 in a ~~hilly~~ area. Videnova et al. (Videnova et al. 2006) also used an ANN to predict air pollution,
158 and they also predicted forward in time.

159 In this paper, we use data gathered from the Cadarache Valley, a valley with complex topogra-
160 phy located in southeast France, which has been the subject of several studies designed to better
161 forecast local-scale winds. For example, Duine et al. (~~Duine et al. 2016a~~) developed a simple
162 method, based on potential temperature differences routinely observed from a tall tower, to now-
163 cast the existence of down-valley winds. More recently, Dupuy et al. (~~Dupuy et al. 2018~~) used an
164 ANN to downscale Weather Research and Forecasting (WRF, (Skamarock et al. 2005)) forecasts.
165 Instead of using observational data as neural network inputs, Dupuy et al. used a low-resolution
site thesis, his first paper doesn't have WRF

Also mention Dupuy et al. 2019, use data from 1970 to nowcast.

on tower it nowes,
winds, I'll send you
the paper.

166 WRF output as inputs to an ANN. Their work effectively demonstrates that ANNs can be used to
167 downscale physics-based weather models.

168 One novel aspect of the work presented herein is the comparison of results from ANNs and
169 multiple linear regressions (MLRs). A multiple linear regression is a statistical tool where a set
170 of *explanatory variables* linearly models a single *target variable* (Shalev-Shwartz and Ben-David
171 2014). Linear regressions take the form of $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$, where y is the target

172 variable, x_p are the explanatory variables, and β_p are the *regression coefficients*. Since ~~multiple MCR~~
173 ~~linear regression~~ is considered a basic statistical tool, the general topic will not be explained here

174 any further. A more in-depth discussion of the method can be found in Shalev-Shwartz and Ben-
175 David's book (Shalev-Shwartz and Ben David 2014). There has been some previous work done

176 in the atmospheric science domain comparing ANN and MLR performance. For example, Sousa
177 et al. (Sousa et al. 2007) used MLRs and ANNs to predict next day hourly ozone concentrations.

178 Sousa et al. also analyzed the performance of ANNs with principal components as the inputs,
179 which is something we do not do.

180 We are predicting the measurements at any given sensor station using the measurements from
181 other stations in the area. While we (and many others) call this nowcasting, the underlying prin-
182 ciple is the same as spatial interpolation. As such, there are other methods that are capable of this

183 task that are worth mentioning. One of the most common methods, called kriging, has been used
184 by both Asa (Asa 2012) and Friedland et al. (Friedland et al. 2016) to spatially interpolate wind

185 data. The name "kriging" is in fact somewhat unique to geospatial statistics applications, and is
186 more generally known as a Gaussian process (Rasmussen and Williams 2006). Osborne et al. (Os-
187 borne et al. 2008) used Gaussian processes to interpolate sensor readings. Hart et al. (Hart et al.

188 2009) used sensor stations along with satellite imagery to spatially interpolate evapotranspiration
189 data. Finally, both Apaydin et al. (Apaydin et al. 2004) and Luo et al. (Luo et al. 2008) have writ-

190 ten spatial interpolation comparison papers, where they compare methods such as kriging, inverse
191 distance weighting, polynomial interpolation, splines, and more.

192 Ultimately, we believe this is novel work. We are nowcasting current values, while many others
193 nowcast in the near future. We also nowcast several different environmental variables, on small
194 spatial scales in complex terrain, and over short time scales. Finally, we compare ANN perfor-
195 mance to MLR performance, while many others only use ANNs.

196 **3. Methods**

197 *Experiment Overview*

198 The data used for this publication were gathered during the Katabatic winds and Stability over
199 Cadarache for Dispersion of Effluents (KASCADE) experiment of 2017. KASCADE 2017 is a
200 follow-on experiment to the KASCADE experiment conducted in 2013 (Duine et al. 2016a,b) that
201 was focused on understanding the vertical structure of the atmosphere in the Cadarache Valley at
202 night during stable atmospheric conditions.

203 KASCADE 2017 was conducted in the Cadarache Valley of the Bouches-du-Rhône department
204 in southeastern France from December 2016 through June 2017 (See Figure 2). The Cadarache
205 Valley contains the French Alternative energies and Atomic Energy Commission (CEA) research
206 center and the International Thermonuclear Experimental Reactor (ITER) is located in the adjac-
207 ent Durance Valley. The CEA performs various types of nuclear research including the study of
208 contaminant dispersion in the event of an accident. To better understand and predict dispersion
209 events, it is critical to have detailed knowledge of small-scale winds and other atmospheric vari-
210 ables. Therefore, increasing our understanding of these phenomena was the main objective of the
211 experiment.

212 As illustrated in Figure 3, the Cadarache Valley is a small ~~valley~~ about 6 km long by 1 km
9 wide. The ~~elevation~~^{valley} difference between the floor of the valley and the peaks is about 100 m. The
213 mouth of the valley is connected to the Durance valley, which runs approximately perpendicular
214 to the Cadarache Valley (Dupuy et al. 2018). The land cover and land use within the valley is
215 heterogeneous, with a combination of buildings, roads, grassy areas, and light forests. There are
216 ~~two main flow regimes present in the Cadarache Valley: thermally driven flow and synoptic flow.~~
1 ~~two main flow regimes present in the Cadarache Valley: thermally driven flow and synoptic flow.~~
217
218 During thermally driven flow events (top half of figure 3), the winds blow down the valley with
219 ~~smaller~~ downslope flow components feeding into it. During synoptic ~~flow~~^{by-forced} events (bottom half
220 of figure 3), larger-scale weather systems drive the wind ~~relatively~~^{s relatively uniformly across} ~~equally~~^{across} all the LEMS
221 locations. ~~in the Valley.~~

222 For KASCADE 2017, the Cadarache Valley and the surrounding areas were heavily instru-
223 mented. Included in the deployment were: four sonic anemometer stations, one surface flux sta-
224 tion, two SODAR stations, wind and temperature measurements from a 110 m tower, two general
225 meteorological stations, and 12 Local Energy-budget Measurement Stations (LEMS). In addition
226 to these continuous observations, radiosondes were released every three hours during fourteen
227 Intensive Observation Periods (IOPs). In this paper, we use a subset of all the data collected.
228 Namely, we use ~~the~~ data collected by the LEMS from January through March 2017.

229 LEMS are small, low-cost meteorological stations that are capable of taking surface and subsur-
230 face measurements. The LEMS used for this experiment are the second ~~version~~^{generation} of the instrument.
1 ~~version~~^{generation}
231 The first ~~version~~^{generation} was designed, built, and characterized in 2013 (Gunawardena et al. 2017). The
232 second ~~version~~^{generation} of the LEMS has a better radiation shield (the Socrima Multiplate radiation shield
~~(2008)~~
233 outlined in van der Meulen et al. ~~(van der Meulen and Brandsma 2008)~~, a better processor, and
234 updated sensors. The LEMS are open source, and information and build instructions can be found
235 at <https://github.com/madvoid/LEMSv2>.

deployed during KASCADE 2017

7 236 Each LEMS at Cadarache measures the following variables at approximately 2 m: wind speed
237 and direction, incoming shortwave solar radiation, air temperature, and air relative humidity. Baro-
238 metric pressure is measured at approximately 1 m. In addition, LEMS measure surface tempera-
239 ture, as well as soil moisture content and temperature at two different heights (≈ 5 and ≈ 25 cm)
240 below the surface. The heights of the sensors relative to the ground for each LEMS is approxi-
241 mately the same, and each LEMS has the same kind of sensor for each measurement. The wind
242 speed and direction measurement for each LEMS were made using a cup and vane anemometer.
243 Therefore, the data can be inaccurate at low wind speeds, and may also demonstrate overspeed
244 problems as observed in the literature (Kristensen 1998; Wyngaard 1981). The use of a cup and
245 vane anemometer also affects model prediction ability. Both artificial neural networks and mul-
246 tiple linear regression models will give a non-zero value as a prediction while the data is zero,
247 increasing the calculated error. The LEMS were deployed at 12 different locations around the
248 Cadarache Valley. Each location can be viewed in Figure 3, and information about each LEMS
249 location can be seen in Table 1.

250 Each LEMS station gathered data at 0.1 Hz. The data were quality controlled and averaged, with
251 an averaging period of five minutes. These 5-minute averages are used for all methods described
252 in this paper.

253 ANN Details

254 The ANNs used in this paper were implemented using MATLAB's Neural Network Toolbox
255 (Beale et al. 2017). Each ANN was a standard feedforward neural network with one hidden layer.
256 The number of hidden nodes in the hidden layer, as well as the number of inputs and outputs,
257 changed with each experiment, so those values will be presented in each of the results sections.
258 The initial values for the weights and biases of the neural networks were randomly generated, and

which should be in discussion rather

were dependent on the random seed. Since the random seed was varied across some experiments,
they will be presented alongside the results. While many ANNs use stochastic gradient descent
(SGD) as their training algorithm (Shalev-Shwartz and Ben-David 2014), we use the Levenberg-
Marquardt algorithm. The Levenberg-Marquardt algorithm is used for training since it is rec-
ommended by MATLAB as being the fastest training algorithm (MathWorks cited 2018). The
transfer function for the hidden layer is the hyperbolic tangent sigmoid function, while the transfer
function for the output layer is the linear transfer function. The ANN performance function is
the mean squared error (MSE), and normalization and regularization happens internally in MAT-
LAB. MATLAB can do preprocessing within the Neural Network Toolbox. For this application,
MATLAB's preprocessing consisted of removing constant inputs, and mapping the minimum and
maximum values of all inputs to -1 and +1 respectively. The inverse of these preprocessing steps
were taken for the output of the network. While we created our own training and testing data for
the experiments, it is important to note that MATLAB uses the dataset to create its own internal
training, testing, and validation data partitions. For all the experiments conducted, we set MAT-
LAB to split the given training data into 75% internal training data, 20% internal validation data,
and 5% internal testing data. The training and testing splits we created will be presented alongside
the results for each experiment. Since the inputs and outputs for each experiment are different,
they will also be presented alongside the results for each experiment. Finally, ensemble averaging
was frequently implemented. When ensemble averaging is utilized, multiple models with different
initial weights are trained, and their outputs averaged (Krogh et al. 1995). This ensemble model
typically has *better* performance than a single model, and is less likely to show outlier perfor-
mance. In all the tests presented here, the neural networks *within the ensemble average* varied
by their weight initialization. We didn't do ensemble averages with different inputs, outputs, or
number of hidden nodes. If an ensemble average is used, it will be specified alongside the results.

283 *MLR Details*

284 As with the ANNs, the MLRs were implemented using MATLAB. In particular, the Statistics
285 and Machine Learning Toolbox (MathWorks 2017) was used. Since the inputs, outputs, and train-
286 ing/testing splits are experiment dependent, they will be presented alongside the results. No extra
287 preprocessing was done for any of the MLRs and none of the explanatory variables are trans-
288 formed. That is, all MLR models are of the form $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$. While ANNs
289 have many hyperparameters such as batch size, learning rate, loss functions, weight initialization,
290 etc., MLRs do not. Therefore, the MLRs can be run without specifying many hyperparameters
291 beforehand.

292 **4. Results and Discussion**

1 This section is broken into four subsections. Subsections 4 and 4 include results from the ANN
293 and the MLR model tests respectively. The tests performed in both of these subsections were iden-
294 tical except for the nowcasting method used. As a result, the plots and results in these two sections
295 have identical formatting, are directly comparable, and should be viewed with the objective of
296 comparing the two different models. Subsections 4 and 4 discuss the results and conclusions from
297 both models. Specific conclusions about the hyperparameters of each models can be found there.
298

299 *ANN Results*

300 To test the effectiveness of an ANN in nowcasting microclimate parameters, a specific suite
301 of tests were conducted. In these tests, the U & V wind components (as defined in the standard
302 meteorological coordinate system), surface temperature, barometric pressure, and virtual potential
303 temperature from LEMS I, J, and K were used to predict a given parameter from each of the other
304 LEMS, excluding LEMS C. In other words, virtual potential temperature, specific humidity, or

305 individual wind speed components were predicted for LEMS A, B, D, E, F, G, H, or L using
306 the measurements from LEMS I, J, and K. These four environmental parameters were chosen as
307 outputs because they are typically available on other commercial weather stations. *For example,*
soil moisture is a measurement that is available on the LEMS but wasn't used for this analysis
308 *because most other weather stations don't have it.* LEMS I, J, and K were used as the input
309 stations because they captured the different kinds of flows present in the Cadarache Valley: slope
310 flows, valley flows, and ridge flows (this was determined from a pre-analysis of the data). Since
311 these LEMS represented the different flows present in the valley, we hypothesized that they would
312 be the strongest predictors. These LEMS were also at different elevations, so they would capture
313 any vertical stratification *of* in the atmosphere. LEMS C *wasn't* used in any of the tests because
314 a complete dataset for that station wasn't available. Since there were eight target LEMS, eight
315 different single-output ANN ensemble averages were trained. Each network within the ensemble
316 averaged model had 15 inputs (5 parameters each from 3 LEMS), and one output. Each ANN
317 had fourteen hidden nodes (see subsection 4 for more information), and each ensemble average
318 was initialized using the same random seed. Each ensemble average consisted of five randomly
319 initialized neural networks.

320 There are two test periods for each variable. These two test periods were chosen because they
321 exhibit very different flow characteristics, testing the abilities of the ANN. The test period starting
322 15 January 2017 and ending 20 January 2017 *displayed a very stable atmosphere with cold air*
323 *pooling* (thermally driven flows). The second test period, starting 27 January 2017 and ending 01
324 February 2017, displayed an atmosphere with stronger winds and broken stratification (synoptic
325 driven flows). Since synoptic-driven flows break up thermally driven flows, these two test periods
326 represent two extremes present in the Cadarache Valley.

*Cold pools
are only at night*

328 To summarize, eight different single-output ANN ensemble averages were trained for each of the
329 four output variables, for two different test periods, resulting in a total of 64 ANN tests conducted.

330 The training data is identical for all runs within a given test period: 5-minute averages of the data
331 from 16 December 2016 to 15 March 2017, excluding the test period for any given run. This
332 training period was chosen because it is the period where there was a full deployment of sensors.

333 Five-minute averages were chosen instead of raw data so noise was smoothed.

for the input variables to smooth turbulent fluctuations.

334 The results from the 64 tests are presented below. Figures 4 and 5 show time series plots com-
335 paring experimental data to the ANN output for the 15 January 2017 and 27 January 2017 test
336 periods respectively. Each figure has four sub-figures, one for each predicted variable. Out of the
337 8 LEMS that were used for prediction, only LEMS A, F, and E are shown in the plots. LEMS
338 A, F, and E were chosen to be shown for readability and because they also represent slope flow,
339 ridge flow, and valley flow respectively. Figures 4 and 5 are useful to visualize how well the ANNs
340 work.

341 The rest of the results are quantified in tables 2 through 5. Each table shows either the root-
342 mean-square error (RMSE) or normalized root-mean-square error (NRMSE) between the model
343 and the experimental data for each run. *The* NRMSE is defined as

$$\text{NRMSE} = \frac{\sqrt{\frac{\sum_{n=1}^N (\hat{y}_n - y_n)^2}{N}}}{\max(y) - \min(y)} \quad (1)$$

344 where \hat{y} is the ANN prediction and y is the experimental data. In addition, two additional statistics
345 are displayed. The first is the RMSE or NRMSE between the experimental values range and the
346 ANN prediction range. The second is the RMSE or NRMSE between the experimental values
347 standard deviation and the ANN prediction standard deviation. Therefore “Range” refers to the
348 RMSE/NRMSE calculated between the range of all the output variable experimental values and the
349 range of all the ANN output values. Similarly, “StdDev” refers to the RMSE/NRMSE calculated

why were these stats chosen?

350 between the standard deviation of all the output variable experimental values and the standard
351 deviation of all the ANN output values. A ~~graphical~~
352 can be seen in Appendix A1.
Similar are presented
353 These tests will be shown again in section 4, except with an MLR as the prediction model instead
354 of an ANN.

b.

355 ANN Discussion

356 All neural networks have a number of hyperparameters (parameters which are not part of the
357 training process) (Claesen and Moor 2015) that can be tuned to change the model performance.
Investigate sensitivity of the model results to some
358 In this section, we present the results of tuning some, but not all, of the hyperparameters available
359 ~~in the model~~. Since the MATLAB Neural Network Toolbox was used for the implementation,
360 many of the default parameters of the toolbox are used, and not expanded upon here. Some
361 example hyperparameters that aren't discussed are the hidden node activation function, the training
362 algorithm, and the training batch size. Exploration of the variation of these hyperparameters is
363 presented in various sources in the literature (e.g., (Orr and Müller 2003)). We also do not explore
364 the number of hidden layers used in the ANNs, as there is often no practical reason to have more
365 than one hidden layer for standard feedforward neural networks (Heaton 2008).

366 One of the more important hyperparameters that can be tuned is the number of hidden nodes in
367 the hidden layer. To study this effect, we varied the number of nodes in the ANN to see how the
368 performance changed. For this test, four different environmental variables were predicted: specific
369 humidity, virtual potential temperature, the U component of the wind velocity or the V component
370 of the wind velocity. The output LEMS used was LEMS E. That specific environmental variable
371 from all the other LEMS were used as the inputs to that given ANN. For example, the specific
372 humidity from LEMS A, B, C, D, F, G, H, I, J, K, & L were used to predict the specific humidity

17 was - agreement
or change humidity to humiditirs

use specific bullet points
like we discussed
in think => use

Comments

from LEMS E. Note that the inputs for this test only consist of the wind velocity U component, whereas the inputs for the tests described in section 4a consist of the wind components, virtual potential temperature, surface temperature, and barometric pressure. This was done because we thought that using a simpler test would better illustrate the effect of changing the hidden nodes more. The testing data for each ANN were the data from 15 January 2017 to 20 January 2017. This testing period was used simply because it was studied extensively for this experiment and was used elsewhere. The training data were from 16 December 2016 to 15 March 2017, excluding the testing data period. The number of nodes were incremented from one node to 30 nodes by one. For each number of nodes, an ensemble average of five neural networks was used to calculate the performance. The output of this ensemble average was used to determine the performance of any given number of nodes. An ensemble average was used here to prevent any outlier ANNs from skewing the results.

The NRMSE was used as the error metric so the different variables can be compared, and the results can be seen in figure 1.

As seen in figure 1, the ANN performance by number of hidden nodes is relatively flat for all environmental variables predicted. Interestingly, increasing or decreasing the number of hidden nodes does not significantly change the performance of the ANN. Even more interestingly, for all variables except specific humidity, a one hidden node ANN performs as well as, or better than “more powerful” ANNs. Even for specific humidity, the one hidden node network has a low NRMSE, just not as low as the higher node networks. In fact, this result is what motivated us to try using linear regression to try and predict these environmental variables as a single node neural network is essentially the linear regression equation passed through an activation function. This also explains why we used 14 hidden nodes for the tests conducted in subsection 4: we could have used most values between two and 30 and it wouldn’t have made a very large difference in the result. The fact that a multiple linear regression can perform a task essentially as well as an

Ref.
(Dury 2018)

Section 4 *unnecessarily powerful?*
397 artificial neural network (see subsection 4) means that an ANN is too powerful of a method to use

398 for this specific task, and that hyperparameter tuning will only yield marginal gains.

399 Another hyperparameter that is relevant to *this specific usage* of ANNs is the number of output
400 nodes. ANNs are not limited to one output; they can have several outputs for any number of
401 inputs. To produce the results of section 4, eight different ANNs were trained with one output
402 each. We repeated the tests from section 4 identically, except we trained a single ANN with
403 eight outputs. We found no discernible or systematic difference between the results of the two
404 approaches. Hence, we have not presented any plots from the multiple-output test. However, we
405 did notice that the time to train one multiple-output neural network was longer than training eight
406 separate single-output networks.

407 On a more practical note, a hyperparameter that is important when using MATLAB's Neural
408 Network Toolbox is the train/test/validate data split percentage. When training the ANN, MAT-
409 LAB internally splits the data into a training subset, a testing subset, and a validation subset. By
410 default, MATLAB uses 70% of the original data as training data, 15% as validation data, and 15%
411 as test data. The test data are only there so the user can view the performance of the neural net-
412 work. The ANN performance on validation data, however, is often used by MATLAB as a training
413 stopping condition. Therefore, in this instance, having a small test split won't affect us very much,
414 as we have our own separate test data that the ANN has never "seen" before. But, having more
415 training data will increase neural network performance, and having more validation data will im-
416 prove generalization. This is why all ANNs presented in this paper have a train/validation/test split
417 of 75%/20%/15%.

an investigation of the model's sensitivity to inputs
418 The last test we performed was ~~not related to hyperparameters, but to the type of inputs~~. The
419 results of 4 always used data from the same three LEMS as inputs: LEMS I, J, and K. These LEMS
420 were chosen to be the inputs because their locations were spread across the measurement area (both

421 horizontally and vertically) and captured many phenomena associated with thermal circulation in
422 complex terrain (e.g., cold pools, slope/valley flows). A priori, one might assume that the good
423 performance exhibited by the ANNs is due to the locations of the input LEMS, and not because of
424 the inherent power of the ANN. To test this, a ~~combination~~^{combinatorial} analysis was performed, where every
425 possible combination of three LEMS was used to predict values at the other nine LEMS. The
426 inputs were identical to those from the tests from subsection 4. Specifically, the inputs were: wind
427 velocity components, surface temperature, barometric pressure, and virtual potential temperature
428 from three LEMS. The output was the virtual potential temperature of the other nine LEMS. Nine
429 single-output ANNs were trained instead of a single nine-output ANN since the training time was
430 shorter that way. While LEMS C was excluded from the tests in subsection 4, it was included here
431 to truly explore the spatial relationship between the LEMS.

432 The testing data were taken from 15 January 2017 to 20 January 2017. The training data were
433 from 12 January 2017 to 15 March 2017, excluding the testing data period (LEMS C was available
434 from 12 January on). Ensemble averages were not used and the number of hidden nodes was
435 14. There are $\frac{12!}{3!(12-3)!} = 220$ different combinations of input LEMS. Throughout this document,
436 when we refer to the “combination number”, we mean a specific combination out of the 220
437 combinations. For example, combination “1” would have LEMS A, B, and C as input LEMS, and
438 the rest as output LEMS. Since 9 different ANNs were trained for each combination, there are $9 R^2$
439 values associated with each combination; each R^2 value is computed with the difference between
440 the ANN model prediction and the experimental data. We can also calculate the same statistics
441 calculated in section 4, which are, the standard deviation of all the target variables, and the range
442 of all the target variables, for a given time step. Using these statistics, we can calculate the RMSE
443 between the statistics for the ANNs and the ~~actual~~^{experimental} data, giving us a single statistic that quantifies
444 the performance of any given combination of input LEMS.

Describe the
and metrics fully
selected.

Figure 6 shows the results of the combinatorial analysis for prediction of the virtual potential temperature. As seen in subplot (a), most combinations have very low RMSE values between both statistics, but there are some combinations of input LEMS that have very high RMSE values between both statistics. Subplot (b) shows the mean R^2 value for each combination, which correlates with the top subplot. Subplot (c) shows the median R^2 value, which shows that all combinations perform quite well, and that there are outlier R^2 values. We verified this manually. In every combination where the mean R^2 value was below zero¹, we looked at the ANN response to the input test data. For every one, there was one run (out of nine) where the ANN exhibited a highly nonlinear and non-physical response, and the prediction did not match the measured data at all. The fact that the most combinations have low RMSE values also justifies the choice of LEMS I, J, & K for the tests conducted in subsection 4. The choice of input LEMS doesn't make a very big difference to the output.

To ensure that this was not systematic, we re-ran the entire combinatorial analysis multiple times using a different random seed. Each time, the combination numbers that exhibited poor performance changed. To show this, we performed the combinatorial analysis three times with three different ANN seeds: MATLAB's default seed, the shuffled seed (where the seed is based on the current time), and a constant (the number 14). Then, the combinations from each of these tests were sorted in order of smallest to greatest range RMSE. This ranking is insightful because well-performing combinations will consistently be at the top of the rankings. For example, if combination "3" (LEMS A, B, and E as inputs) consistently performed better than other combinations, it would be near the top of the rankings for each of the three different random seeds. To note that the

¹It is commonly believed that the coefficient of determination R^2 cannot be less than 0. This is not true. The coefficient of determination is defined as $R^2 \equiv 1 - (\sum_n^N (y_n - \hat{y}_n)^2) / (\sum_n^N (y_n - \bar{y})^2)$. There is nothing that mathematically restricts this expression from being below zero. It is simply comparing the error between the model and the data to the error between the data and the mean of the data. If the model does a poorer job predicting the data than the mean of the data does, R^2 will be below zero.

466 quantify the similarity between between the three different rankings, Spearman's rank correlation
467 coefficient was calculated for the rankings. Spearman's rank correlation coefficient is the Pearson
468 correlation of two different rankings, and is an efficient way to compare rankings (Downey 2011).
469 Spearman's rank correlation coefficient will be 1 when the rankings agree completely, -1 when the
470 rankings are perfectly inverted, and 0 when there is no relationship between the rankings. Table 6
471 shows Spearman's rank correlation coefficient for the three different seeds. It is evident that there
472 is no relation between the rankings, implying that there are no combinations that consistently work
473 better than others. We believe that the bad combinations are a result of the optimization algorithm
474 behind the neural network "falling" into a local minima, greatly reducing performance. This is
475 somewhat evident from the numbers: out of $220 \cdot 9 = 1980$ ANN runs, only a handful exhibit the
476 very bad values.

were these usually non-physical values that could be rejected?

477 It is also interesting to note the difference in performance between the thermally driven flow days
478 (15 January 2017 to 20 January 2017) and the synoptic flow days (27 January 2017 to 01 February
479 2017) for the wind velocity components. On thermally driven flow days, the wind velocity between
480 stations is much less correlated with each other and behave more independently. However, on
481 synoptic flow days the wind velocity between stations is much more correlated and readings are
482 very similar. As such, the ANN predicts the wind velocity much better on synoptic flow days.

1 483 Finally, it is important to discuss the physical reasons why wind velocity is more difficult to
484 predict than virtual potential temperature and specific humidity. To begin, both virtual potential
485 temperature and specific humidity are strongly correlated with the diurnal cycle, and aren't as
486 spatially variable as wind velocity is. However, wind velocity is also difficult to predict for reasons
487 other than the underlying environmental physics. Wind velocity on the LEMS is measured by a
488 cup and vane anemometer, which doesn't measure wind speeds accurately, if at all, at wind speeds
489 less than 0.5 m/s . Therefore, many of the low magnitude readings are likely incorrect. In addition,

include example both in text.

Components are

How do you come to this conclusion, I don't think I agree.

490 the anemometer “transfer function” has a discontinuity. The wind speed measurement will be zero

491 and then it will quickly jump to $0.5 \frac{m}{s}$ or above. These discontinuities can be difficult to model

492 accurately.

To overcome these difficulties, it would be better to use
a sonic anemometer, which meas. low wind speeds more accurately

493 We conclude that for this experiment ~~and data~~, an ANN can perform nowcasting well, regard-

494 less of the input LEMS. We note that environmental variables with a strong diurnal dependence

495 are better predicted, which is why wind is not predicted as well as virtual potential temperature.

496 We also think that hyperparameter tuning is important, but not necessarily required for good per-

497 formance. We hypothesize that this is because multiple linear regressions also work for this case,

498 which is discussed in the next sections.

Let's
discuss,
I disagree

C.

MLR Results

base

500 The tests run with ANNs (section 4) were repeated identically ~~for~~ ^{to} MLRs. Therefore, 64 MLR

501 models were created: eight different MLR models for each of the four output variables for two

502 different test periods. Each MLR used the U & V wind components, surface temperature, baro-

503 metric pressure, and virtual potential temperature (θ_v) from LEMS I, J, and K (15 input variables)

504 to predict a given parameter (virtual potential temperature, specific humidity, or individual wind

505 speed components) from LEMS A, B, D, E, F, G, H, or L. The training and testing data ~~is~~ ^{are} identical

506 to that of the ANN tests. Since the ~~produced~~ MLR coefficients are identical for a given dataset,

507 ensemble averages don't increase performance and were therefore not used.

Figures

508 As with the ANN results, plots 7 and 8 show time series plots comparing experimental data

509 to the MLR output for the 15 January 2017 and 27 January 2017 test periods respectively. Each

510 subfigure within these figures shows the prediction and the experimental data for LEMS A, F, and

511 E. The rest of the results are quantified in tables 7 through 10. A more detailed description of

4a

512 the statistics and an in-depth discussion on modeling decisions can be found in subsection 4. A
513 visualization of the data in tables 7 and 9 can be found in Appendix A1.

4.4

514 *MLR Discussion*

515 Since MLRs are much simpler than ANNs, there are no hyperparameters to tune. In addition,
516 since MLRs do not have random initialization and iterations, they produce the same result every
517 time, assuming the training data is identical. We originally didn't even consider trying an MLR,
518 because we assumed that the relationship between the measurements of the different LEMS was
519 highly non-linear. However, when conducting the ANN hidden node analysis (see section 4), we
520 noticed that a single-node ANN performed nearly as well as a multi-node ANN. Since a single-
521 node ANN is essentially a multiple linear regression whose output is passed through a sigmoidal
522 function, we decided to try an MLR as the nowcasting method. Surprisingly, as seen in 4, it worked
523 well. When comparing tables 2 and 4 to tables 7 and 9, one can see that the ANN usually performs
524 better than the MLR, but not always. In addition, the two models don't consistently perform
525 better on one set of dates compared to others. The ANN seems to have the greatest advantage
526 over the MLR when predicting wind components, which are typically more complex than the other
527 environmental variables.

528 There are two main reasons to use an MLR instead of an ANN if possible. The first is computa-
529 tional runtime. Even on powerful computers, ANNs take much longer to train when compared to
530 MLRs. The second is interpretability. ANNs tend to be black boxes that are difficult to interpret,
531 whereas MLRs are transparent and easier to interpret. However, the biggest concern about using
532 MLRs are the assumptions that need to be met. According to Poole et al. (Poole and O'Farrell
533 1971), there are six critical assumptions being made when using an MLR successfully. One of
534 these assumptions is that the independent variables are linearly independent of each other. When

they are not, collinearity exists, and the precision of the regression coefficient decreases (James et al. 2013). During our analysis, we noticed that there were high Variance Inflation Factors (VIFs) for many of the coefficients, implying multicollinearity. This makes intuitive sense. For example, when the sun sets, all stations will measure a ~~temperature drop and the radiation measurements and~~ ^{radiation drop, and the air and surface} ~~the air~~ temperature measurements will correlate. While we've shown that the MLR can nowcast successfully in this instance, the reduced accuracy of the regression coefficients due to multicollinearity means that interpreting the models must be done with care. To successfully interpret the models, one must reduce the number of explanatory variables until multicollinearity is minimized. This can also be done for the ANN, but due to the ANN's inherent "black box" nature one should be careful interpreting it anyways. Manually, multicollinearity can be minimized by calculating correlation coefficients between all explanatory variables and removing highly correlated ones. Automatically, this can be done by running algorithms such as lasso regression, which automatically remove unneeded variables. Using lasso regression will also quickly reveal which stations are most correlated with others. While some preliminary work has been done with regards

to this, it wasn't included in this paper as it felt out of the scope.

in the cases that we examined for the Cadarache Valley

In conclusion, multiple linear regression can be used to nowcast many environmental models as well as artificial neural networks ~~can~~, implying a linear spatial relationship between the environmental variables ~~measured at Cadarache Valley~~. The MLR models do not perform as well as the ANN when it comes to wind velocity nowcasting though, and this is probably due to the greater amounts nonlinearity in wind velocity. To properly interpret the MLR models, collinearity would have to be reduced either manually or with an automated method such as lasso regression.

5. Summary

To summarize, the main purpose of this paper is to demonstrate that artificial neural networks and multiple linear regression models can be used to nowcast environmental measurements in complex-terrain boundary-layer meteorology applications. Specifically, we show that this works in the Cadarache Valley located in south-east France. The valley, which is approximately 6 km long by 1 km wide was instrumented with a dozen low-cost weather stations (called LEMS) for a four month long field experiment. The weather stations measured several different variables, but the variables predicted by the models were wind velocity components, virtual potential temperature, specific humidity, and air temperature. The valley exhibited two main types of flow: thermally driven flow and synoptic flow.

The valley exhibited various types of flows including nearly pure thermally driven flows and a wide range of synoptically forced flows.

In general, both the ANN and the MLR models performed similarly well on the test data. Two test periods were used. The test period from 15 January 2017 to 20 January 2017 represents mostly slope and valley-
thermally driven flow, while the test period from 27 January 2017 to 01 February 2017 represents mostly synoptic flow. Both models performed better on synoptic-driven flow periods over thermally driven flow periods, probably because *the variability among the stations was smaller* during synoptic periods. Both models predicted specific humidity and virtual potential temperature better than the wind components, likely because the wind components are much more chaotic.

We also performed multiple tests to further understand the ANN. This involved a hidden node analysis, an output node analysis, and a combination analysis. The results indicated that the ANN performance doesn't change very much with increasing hidden nodes, and that the three LEMS chosen as input LEMS also doesn't matter very much. We also discovered that the ANN will sometimes get stuck in a local minima and produce non-physical values, and that this seemed random. We

I think
→ Light winds, which are highly variable & can be influenced by weak perturbations & measurement noise.

*for review
during this period*

Add: These promising results indicate that ANN and at times mLR can be used for dock filling after a field experiment has been completed. Proximity of complex structures, spatial information in regions of concern, and temporal variations in flow regimes are h

578 did not perform as many of these tests on the MLR model, mostly because MLRs do not have as
579 many hyperparameters, thus limiting the amount of tests that can be run.

580 Despite these tests, there are still many tests that can be run, and a lot of future work to be done.

promising

581 The purpose of this paper was to present the results from some preliminary tests of these methods,
582 not to be an exhaustive reference on the methods' capabilities. In addition, some of these tests
583 have virtually limitless configurations that can be run, meaning we had to limit the scope of the
584 tests somehow. For example, we ran the hidden node tests using a single input environmental
585 variable (for 11 LEMS). We could have ^{run} it using two environmental variables for 11 LEMS. Or
586 three environmental variables, and so on. Mainly, we chose tests that we felt exposed the behavior
587 of each model well without going down the rabbit hole of adjusting parameters. Especially for the
588 ANN, we felt that many of the possible tests we could have done were out of the scope of this
589 particular paper since we showed that a) ^{mLR} models work, and b) the number of hidden nodes
590 doesn't affect output performance very much.

Hence,

591 As such, there are many topics that we would consider as possibilities for future work. In no
592 particular order, here is a list of things we didn't try, or did try but didn't publish:

- 593 1. Add environmental variables one by one, in different orders to see how results are affected
- 594 2. Add stations one by one, in different orders to see how results are affected
- 595 3. Use different amounts of training/testing data to see how much data is needed to create a
596 successful nowcasting tool
- 597 4. Test models using data from different locations with different sensors at different spa-
598 tial/temporal scales
- 599 5. Quantify how model results change with different flow regimes

for ANN or
mLR?
MLR or ANN?
Specify
for all

600 6. Increase the interpretability of both models. For MLR, this may involve using a tool such as
601 lasso regression

602 With the completion of these and other tasks, we conclude that both artificial neural networks and
603 multiple linear regression show promise in becoming successful nowcasting tools in micrometeo-
604 rology.

605 *Acknowledgments.* We would like to thank the 4M team of the Centre National de Recherche
606 Mtorologique for kindly allowing us to use their Socrima radiation shields during KASCADE
607 2017, Eric Pique for the preparation and installation of the LEMS for KASCADE 2017, Pierre
608 Rubin for general experimental help, the United States Fulbright Program for providing funding,
609 and the University of Utah Global Change and Sustainability Center for providing funding.

610 APPENDIX

611 Supplementary Plots

612 A1. Supplementary Plots

613 The plots in this section all expand on the results presented in sections 4 and 4. Specifically, they
614 show visualizations of every result presented in tables 2, 4, 7, and 9. There are sixteen figures, each
615 with two subfigures. Eight of the figures are for the ANN results, for the four variables tested, for
616 the two different test periods, while the other eight are for the MLR results, for the same variables
617 and test periods. The first subfigure in each figure is a one-to-one plot showing the model output
618 vs. the actual data, while the second shows a time series of the two statistics calculated: range and
619 standard deviation for all the time series predicted.

620 This shows how to enter the commands for making a bibliography using BibTeX. It uses refer-
621 ences.bib and the ametsoc2014 bst file for the style.

Check all references & formats, be sure they are correct for ISMEC.
Also, you have some titles capitalized & others not. Only 1st letter of title is capitalized.

- 622 References
- 623 Acevedo, O. C., and D. R. Fitzjarrald, 2001: The Early Evening Surface-Layer Transition: Temporal and Spatial Variability. *Journal of the Atmospheric Sciences*, **58** (17), 2650–2667, doi:10.1175/1520-0469(2001)058<2650:TEESLT>2.0.CO;2, URL <http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%282001%29058%2C2650%3ATEESLT%2E2.0.CO%2B2>.
- 624
- 625
- 626
- 627
- 628 Adams, R., L. Houston, and R. Weiher, 2004: The value of snow and snow information services. *Report prepared for NOAA's National Operational Hydrological Remote Sensing Center, Chanhassen, MN, under contract DG1330-03-SE-1097*.
- 629
- 630
- 631 Akyildiz, I. F., W. Su, Y. Sankarasubramaniam, and E. Cayirci, 2002: Wireless sensor networks: a survey. *Computer networks*, **38** (4), 393–422.
- 632
- 633 Apaydin, H., F. Kemal Sonmez, and Y. E. Yildirim, 2004: Spatial interpolation techniques for climate data in the GAP region in Turkey. *Climate Research*, **28** (1), 31–40, doi:10.3354/cr028031.
- 634
- 635 Asa, E., 2012: Nonlinear Spatial Characterization and Interpolation of Wind Data. *Wind Engineering*, **36** (3), 251–272, doi:10.1260/0309-524X.36.3.251.
- 636
- 637 Bailey, B. N., R. Stoll, E. R. Pardyjak, and W. F. Mahaffee, 2014: Effect of vegetative canopy architecture on vertical transport of massless particles. *Atmospheric environment*, **95**, 480–489.
- 638
- 639 Banbara, M., D. Giannone, M. Modugno, and L. Reichlin, 2013: Now-casting and the real-time data flow. *Handbook of economic forecasting*, **2** (Part A), 195–237.
- 640
- 641 Baskett, R. L., J. S. Nasstrom, and R. Lange, 1991: Emergency Response Model Evaluation Using Diablo Canyon Nuclear Power Plant Tracer Experiments. *Air Pollution Modeling and Its Application VIII*, Springer, 603–604.
- 642
- 643

- 644 Beale, M. H., M. T. Hagan, and H. B. Demuth, 2017: MATLAB Neural Network Toolbox User's
645 Guide. Tech. Rep. R2017b, Mathworks, 512 pp.
- 646 Benvenuto, F., and A. Marani, 2000: Neural networks for environmental problems: data quality
647 control and air pollution nowcasting. *Global NEST: The International Journal*, **2** (3), 281–292.
- 648 Chow, F. K., S. F. De Wekker, and B. J. Snyder, 2013: *Mountain Weather Research and Forecast-*
649 *ing: Recent Progress and Current Challenges*. Springer Science & Business Media, 760 pp.,
650 doi:10.1007/978-94-007-4098-3.
- 651 Claesen, M., and B. D. Moor, 2015: Hyperparameter Search in Machine Learning. *CoRR*,
652 **abs/1502.0**, URL <http://arxiv.org/abs/1502.02127>, 1502.02127.
- 653 De Haan, S., and A. Stoffelen, 2012: Assimilation of high-resolution Mode-S wind and tempera-
654 ture observations in a regional NWP model for nowcasting applications. *Weather and Forecast-*
655 *ing*, **27** (4), 918–937.
- 656 Demetriades, N. W. S., and R. L. Holle, 2006: Long range lightning nowcasting applications for
657 tropical cyclones. *Preprints, Conf. Meteorology Application of Lightning Data, Atlanta*, 353–
658 365.
- 659 Downey, A. B., 2011: *Think Stats*. Green Tea Press.
- 660 Duine, G.-J., T. Hedde, P. Roubin, and P. Durand, 2016a: A Simple Method Based on Routine
661 Observations to Nowcast Down-Valley Flows in Shallow, Narrow Valleys. *Journal of Applied*
662 *Meteorology and Climatology*, **55** (7), 1497–1511, doi:10.1175/JAMC-D-15-0274.1.
- 663 Duine, G.-J., T. Hedde, P. Roubin, and P. Durand, 2016b: A Simple Method Based on Routine
664 Observations to Nowcast Down-Valley Flows in Shallow, Narrow Valleys. *Journal of Applied*
665 *Meteorology Climate*, **55** (7), 1497–1511.

- 666 Dupuy, F., G.-J. Duine, P. Durand, T. Hedde, P. Roubin, and E. R. Pardyjak, 2018: Valley-winds
667 at the local scale: Local-scale valley wind retrieval using an artificial neural network applied to
668 routine weather observations. *Journal of Applied Meteorology and Climatology*. *Submitted*
- 669 Emeis, S., 2010: *Measurement methods in atmospheric sciences: in situ and remote*. Gebr. Born-
670 traeger Science Publishers.
- 671 Fernando, H., and Coauthors, 2015: The materhorn: Unraveling the intricacies of mountain
672 weather. *Bulletin of the American Meteorological Society*, **96** (11), 1945–1967.
- 673 Fernando, H. J. S., 2010: Fluid dynamics of urban atmospheres in complex terrain. *Annual review*
674 *of fluid mechanics*, **42**, 365–389.
- 675 French, M. N., W. F. Krajewski, and R. R. Cuykendall, 1992: Rainfall forecasting in space and
676 time using a neural network. *Journal of Hydrology*, **137** (1-4), 1–31, doi:10.1016/0022-1694(92)
677 90046-X.
- 678 Friedland, C. J., T. A. Joyner, C. Massarra, R. V. Rohli, A. M. Treviño, S. Ghosh, C. Huyck, and
679 M. Weatherhead, 2016: Isotropic and anisotropic kriging approaches for interpolating surface-
680 level wind speeds across large, geographically diverse regions. *Geomatics, Natural Hazards and*
681 *Risk*, **5705 (August 2017)**, 1–18, doi:10.1080/19475705.2016.1185749, URL <http://dx.doi.org/10.1080/19475705.2016.1185749>.
- 683 Furquim, G., and Coauthors, 2014: Combining wireless sensor networks and machine learn-
684 ing for flash flood nowcasting. *Advanced Information Networking and Applications Workshops*
685 (*WAINA*), *2014 28th International Conference on*, IEEE, 67–72.
- 686 Gultepe, I., and Coauthors, 2014: Ice fog in Arctic during FRAM–Ice Fog Project: Aviation and
687 nowcasting applications. *Bulletin of the American Meteorological Society*, **95** (2), 211–226.

7
688 Gunawardena, N., E. Pardyjak, R. Stoll, and A. Khadka, 2017: Development and evaluation of
689 an open-source low-cost distributed sensor network for environmental monitoring applications.
690 *Measurement Science and Technology*. *ISSUE & PAGES?*

691 Han, L., J. Sun, W. Zhang, Y. Xiu, H. Feng, and Y. Lin, 2017: A machine learning nowcasting
692 method based on real-time reanalysis data. *Journal of Geophysical Research: Atmospheres*,
693 **122** (7), 4038–4051.

9
694 Hang, C., and Coauthors, 2016: A Case Study of the ~~the~~ Mechanisms ~~in~~ Modulating the ~~the~~ Evo-
695 lution of Valley Fog. *Pure and Applied Geophysics*, **173** (9), 3011–3030, doi:10.1007/
696 s00024-016-1370-4.

697 Hart, Q. J., M. Brugnach, B. Temesgen, C. Rueda, S. L. Ustin, and K. Frame, 2009: Daily refer-
698 ence evapotranspiration for California using satellite imagery and weather station measurement
699 interpolation. *Civil Engineering and Environmental Systems*, **26** (1), 19–33.

700 Heaton, J., 2008: *Introduction to neural networks with Java*. 2nd ed., Heaton Research, Inc.

7
701 James, G., D. Witten, T. Hastie, and R. Tibshirani, 2013: *An introduction to statistical learning*,
702 Vol. 112. Springer.

1
703 Kampa, M., and E. Castanas, 2008: Human health effects of air pollution. *Environmental pollu-*
704 *tion*, **151** (2), 362–367.

705 Karthikeya, B. R., P. S. Negi, and N. Srikanth, 2016: Wind resource assessment for urban renew-
706 able energy application in Singapore. *Renewable Energy*, **87**, 403–414.

707 Kim, S., S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, 2007: Health
708 monitoring of civil infrastructures using wireless sensor networks. *Information processing in*
709 *sensor networks, 2007. IPSN 2007. 6th international symposium on*, IEEE, 254–263.

- 710 Kristensen, L., 1998: Cup anemometer behavior in turbulent environments. *Journal of Atmospheric and Oceanic Technology*, **15** (1), 5–17.
- 712 Krogh, A., A. Krogh, J. Vedelsby, and J. Vedelsby, 1995: Neural Network Ensembles, Cross Validation, and Active Learning. *Nips*, 231–238. *Volume?*
- 714 Lareau, N. P., and Coauthors, 2013: The Persistent Cold-Air Pool Study. *Bulletin of the American Meteorological Society*, **94** (1), 51–63, doi:10.1175/BAMS-D-11-00255.1, URL <http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00255.1>.
- 717 Lateb, M., R. N. Meroney, M. Yataghene, H. Fellouah, F. Saleh, and M. C. Boufadel, 2016: On the use of numerical modelling for near-field pollutant dispersion in urban environments- A review. *Environmental Pollution*, **208**, 271–283.
- 720 LeMone, M. a., K. Ikeda, R. L. Grossman, and M. W. Rotach, 2003: Horizontal Variability of 2-m Temperature at Night during CASES-97. *Journal of the Atmospheric Sciences*, **60** (20), 2431–2449, doi:10.1175/1520-0469(2003)060<2431:HVOMTA>2.0.CO;2, URL { \% }5C{ \% }5CIBSERVER{ \% }5CID{ _ }enl{ \% }5CJournal{ _ }of{ \% }the{ \% }Atmospheric{ _ }Sciences{ \% }5C2003{ \% }5C060-20-2431.pdf.
- 725 Lin, B. B., 2007: Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture. *Agricultural and Forest Meteorology*, **144** (1-2), 85–94.
- 727 Lothon, M., and Coauthors, 2014: The BLLAST field experiment: Boundary-layer late afternoon and sunset turbulence. *Atmospheric chemistry and physics*, **14** (20), 10 931–10 960.
- 729 Luo, W., M. C. Taylor, and S. R. Parker, 2008: A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. *International journal of climatology*, **28** (7), 947–959.

732 Mahaffee, W. F., and R. Stoll, 2016: The ebb and flow of airborne pathogens: Monitoring and use
733 in disease management decisions. *Phytopathology*, **106** (5), 420–431.

734 Mahrt, L., 2017: Stably Stratified Flow in a Shallow Valley. *Boundary-Layer Meteorology*, **162** (1),
735 1–20, doi:10.1007/s10546-016-0191-4.

736 Mainwaring, A., D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, 2002: Wireless sensor net-
737 works for habitat monitoring. *Proceedings of the 1st ACM international workshop on Wireless*
738 *sensor networks and applications*, ACM, 88–97.

739 Mass, C., 2012: Nowcasting: The promise of new technologies of communication, modeling, and
740 observation. *Bulletin of the American Meteorological Society*, **93** (6), 797–809.

741 MathWorks, 2017: MATLAB Statistics and Machine Learning Toolbox User's Guide. Tech. Rep.
742 R2017b, Mathworks.

743 MathWorks, cited 2018: Levenberg-Marquardt backpropagation - MATLAB trainlm. MathWorks,
744 URL <https://www.mathworks.com/help/nnet/ref/trainlm.html>.

745 Miller, N. E., R. Stoll, W. Mahafee, T. Neill, and E. R. Pardyjak, 2016: Field-scale particle trans-
746 port in a trellised agricultural canopy during periods of row-aligned winds. *22nd Symp. Bound-*
747 *ary Layers and Turbulence*.

748 Monteith, J. L., C. K. Ong, and J. E. Corlett, 1991: Microclimatic interactions in agroforestry
749 systems. *Forest Ecology and management*, **45** (1-4), 31–44.

750 Novak, P., 2007: The Czech Hydrometeorological Institute's severe storm nowcasting system.
751 *Atmospheric research*, **83** (2), 450–457.

752 Orr, G. B., and K.-R. Müller, 2003: *Neural networks: tricks of the trade*. Springer.

- 753 Osborne, M. a., S. J. Roberts, A. Rogers, S. D. Ramchurn, and N. R. Jennings, 2008:
 754 Towards Real-Time Information Processing of Sensor Network Data Using Computation-
 755 ally Efficient Multi-output Gaussian Processes. *2008 International Conference on Infor-
 756 mation Processing in Sensor Networks (ipsn 2008)*, 109–120, doi:10.1109/IPSN.2008.25,
 757 URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4505467>{\%}5Cnhttp:
 758 //ieeexplore.ieee.org/document/4505467/.
- 759 Öztopal, A., 2006: Artificial neural network approach to spatial estimation of wind velocity data.
 760 *Energy Conversion and Management*, **47** (4), 395–406, doi:10.1016/j.enconman.2005.05.009.
- 761 Pardyjak, E. R., S. O. Speckart, F. Yin, and J. M. Veranth, 2008: Near source deposition of vehicle
 762 generated fugitive dust on vegetation and buildings: Model development and theory. *Atmo-*
 763 *spheric Environment*, **42** (26), 6442–6452, doi:10.1016/j.atmosenv.2008.04.024.
- 764 Petroff, A., A. Mailliat, M. Amielh, and F. Anselmet, 2008: Aerosol dry deposition on vegetative
 765 canopies. Part I: review of present knowledge. *Atmospheric Environment*, **42** (16), 3625–3653.
- 766 Philippopoulos, K., and D. Deligiorgi, 2012: Application of artificial neural networks for the spa-
 767 tial estimation of wind speed in a coastal region with complex topography. *Renewable Energy*,
 768 **38** (1), 75–82, doi:10.1016/j.renene.2011.07.007, URL <http://dx.doi.org/10.1016/j.renene.2011.07.007>.
- 770 Poole, M. A., and P. N. O'Farrell, 1971: The assumptions of the linear regression model. *Trans-
 771 actions of the Institute of British Geographers*, 145–158. *Volume?*
- 772 Rasmussen, C. E., and C. K. I. Williams, 2006: *Gaussian processes for machine learning*, Vol. 1.
 773 MIT press Cambridge.

- 774 Rasmussen, R., and Coauthors, 2001: Weather Support to Deicing Decision Making (WSDDM):
775 A winter weather nowcasting system. *Bulletin of the American Meteorological Society*, **82** (4),
776 579–595.
- 777 Shalev-Shwartz, S., and S. Ben-David, 2014: *Understanding machine learning: From theory to*
778 *algorithms*. Cambridge University Press.
- 779 Shi, X., Y. Zhu, J. Duan, R. Shao, and J. Wang, 2015: Assessment of pedestrian wind environment
780 in urban planning design. *Landscape and Urban Planning*, **140**, 17–28.
- 781 Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers,
782 2005: A description of the advanced research WRF version 2. Tech. rep., National Center For
783 Atmospheric Research Boulder Co Mesoscale and Microscale Meteorology Div.
- 784 Sousa, S. I. V., F. G. Martins, M. C. M. Alvim-Ferraz, and M. C. Pereira, 2007: Multiple linear
785 regression and artificial neural networks based on principal components to predict ozone con-
786 centrations. *Environmental Modelling & Software*, **22** (1), 97–103, doi:10.1016/j.envsoft.2005.
787 12.002, URL <http://linkinghub.elsevier.com/retrieve/pii/S1364815205002240>.
- 788 Stohl, A., and Coauthors, 2012: Xenon-133 and caesium-137 releases into the atmosphere from
789 the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric
790 dispersion, and deposition. *Atmospheric Chemistry and Physics*, **12** (5), 2313–2343.
- 791 Stull, R. B., 2012: *An introduction to boundary layer meteorology*, Vol. 13. Springer Science &
792 Business Media.
- 793 van der Meulen, J. P., and T. Brandsma, 2008: Thermometer screen intercomparison in De Bilt
794 (The Netherlands), Part I: Understanding the weather-dependent temperature differences. *Inter-*
795 *national Journal of Climatology*, **28** (3), 371–387.

- 796 Videnova, I., D. Nedialkov, M. Dimitrova, and S. Popova, 2006: Neural networks for air pollution
797 nowcasting. *Applied Artificial Intelligence*, **20** (6), 493–506, doi:10.1080/08839510600753741.
- 798 Wilkerson, W. D., 1991: Dust and sand forecasting in Iraq and adjoining countries. Tech. rep.,
799 AIR WEATHER SERVICE SCOTT AFB IL.
- 800 Wilson, J. W., N. A. Crook, C. K. Mueller, J. Sun, and M. Dixon, 1998: Nowcasting thunder-
801 storms: A status report. *Bulletin of the American Meteorological Society*, **79** (10), 2079–2099.
- 802 Wyngaard, J. C., 1981: Cup, propeller, vane, and sonic anemometers in turbulence research. *An-*
803 *nual Review of Fluid Mechanics*, **13** (1), 399–423.
- 804 Xu, K., C. K. Wikle, and N. I. Fox, 2005: A kernel-based spatio-temporal dynamical model
805 for nowcasting weather radar reflectivities. *Journal of the American statistical Association*,
806 **100** (472), 1133–1144.
- 807 Yates, D. N., T. T. Warner, and G. H. Leavesley, 2000: Prediction of a flash flood in complex
808 terrain. Part II: A comparison of flood discharge simulations using rainfall input from radar, a
809 dynamic model, and an automated algorithmic system. *Journal of Applied Meteorology*, **39** (6),
810 815–825.
- 811 Yu, L., N. Wang, and X. Meng, 2005: Real-time forest fire detection with wireless sensor net-
812 works. *Wireless Communications, Networking and Mobile Computing, 2005. Proceedings. 2005*
813 *International Conference on*, IEEE, Vol. 2, 1214–1217.

814 LIST OF TABLES

| | | |
|-----|---|----|
| 815 | Table 1. Table of LEMS locations | 39 |
| 816 | Table 2. RMSE Statistics for ANN results 1-15 | 40 |
| 817 | Table 3. NRMSE Statistics for ANN results 1-15 | 41 |
| 818 | Table 4. RMSE Statistics for ANN results 1-27 | 42 |
| 819 | Table 5. NRMSE Statistics for ANN results 1-27 | 43 |
| 820 | Table 6. Table of Spearman's rank correlation coefficient for ANN combination analy- sis. | 44 |
| 822 | Table 7. RMSE Statistics for MLR results 1-15 | 45 |
| 823 | Table 8. NRMSE Statistics for MLR results 1-15 | 46 |
| 824 | Table 9. RMSE Statistics for MLR results 1-27 | 47 |
| 825 | Table 10. NRMSE Statistics for MLR results 1-27 | 48 |

TABLE 1. Table of LEMS locations

| Name | Latitude | Longitude | Elevation (m) |
|--------|----------|-----------|---------------|
| LEMS A | 43.68483 | 5.76803 | 332 |
| LEMS B | 43.68568 | 5.76885 | 347 |
| LEMS C | 43.66839 | 5.76142 | 397 |
| LEMS D | 43.67518 | 5.78671 | 328 |
| LEMS E | 43.68263 | 5.76568 | 293 |
| LEMS F | 43.66871 | 5.77791 | 383 |
| LEMS G | 43.67848 | 5.75763 | 325 |
| LEMS H | 43.69141 | 5.74918 | 276 |
| LEMS I | 43.69300 | 5.76253 | 385 |
| LEMS J | 43.69548 | 5.74323 | 262 |
| LEMS K | 43.68038 | 5.76003 | 317 |
| LEMS L | 43.68879 | 5.77071 | 368 |

Consider using tables 2 (7)
3, 8 etc
to facilitate comparison

TABLE 2. RMSE Statistics for ANN results 1-15

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|----------------------------------|---------|-------|---------|---------|---------|---------|---------|---------|---------|---------|
| $q \left(\frac{kg}{kg} \right)$ | 7.5e-05 | 7e-05 | 0.00014 | 8.7e-05 | 7.1e-05 | 8.1e-05 | 5.5e-05 | 7.3e-05 | 9.4e-05 | 3.1e-05 |
| $\theta_v(K)$ | 0.65 | 0.79 | 1.30 | 0.77 | 0.72 | 0.54 | 0.34 | 0.56 | 1.44 | 0.51 |
| $U(m s^{-1})$ | 0.24 | 0.25 | 0.93 | 1.26 | 0.52 | 0.32 | 0.41 | 0.44 | 0.68 | 0.25 |
| $V(m s^{-1})$ | 0.24 | 0.42 | 0.46 | 0.77 | 0.53 | 0.53 | 0.33 | 0.39 | 0.73 | 0.24 |

Summary

Don't use this
as stations
standard.
US Jan 2012

using LEMS 1, J, & K
as input stations.

TABLE 3. NRMSE_{~~K~~} Statistics for ANN results 1-15

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| q | 0.043 | 0.041 | 0.083 | 0.051 | 0.039 | 0.045 | 0.032 | 0.042 | 0.13 | 0.12 |
| θ_v | 0.03 | 0.04 | 0.06 | 0.04 | 0.05 | 0.03 | 0.02 | 0.03 | 0.15 | 0.14 |
| U | 0.09 | 0.09 | 0.16 | 0.23 | 0.11 | 0.06 | 0.08 | 0.11 | 0.14 | 0.16 |
| V | 0.08 | 0.08 | 0.14 | 0.15 | 0.10 | 0.08 | 0.07 | 0.15 | 0.14 | 0.15 |

summary

vs Jan 2017

using LENS IJK
as input
station

29 Jan 2017
Same case

TABLE 4. RMSE Statistics for ANN results 1-27

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|--------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| q | 9.2e-05 | 8.4e-05 | 0.00012 | 7.4e-05 | 8.3e-05 | 8.3e-05 | 4.9e-05 | 6.9e-05 | 0.00012 | 3.8e-05 |
| θ_v (K) | 0.34 | 0.31 | 0.39 | 0.22 | 0.30 | 0.22 | 0.19 | 0.27 | 0.45 | 0.14 |
| U (m s ⁻¹) | 0.23 | 0.23 | 0.44 | 0.43 | 0.36 | 0.34 | 0.36 | 0.39 | 0.50 | 0.17 |
| V (m s ⁻¹) | 0.27 | 0.36 | 0.29 | 0.35 | 0.36 | 0.44 | 0.33 | 0.33 | 0.55 | 0.17 |

TABLE 5. NRMSE Statistics for ANN results 1-27

*Done
as above*

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|------------|-------|-------|-------|-------|-------|------|-------|-------|-------|--------|
| q | 0.031 | 0.029 | 0.034 | 0.024 | 0.031 | 0.03 | 0.016 | 0.025 | 0.075 | 0.065 |
| θ_v | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | 0.06 | 0.05 |
| U | 0.05 | 0.05 | 0.06 | 0.05 | 0.05 | 0.06 | 0.05 | 0.05 | 0.10 | 0.10 |
| V | 0.06 | 0.06 | 0.08 | 0.06 | 0.08 | 0.08 | 0.06 | 0.08 | 0.09 | 0.08 |

nucleotide *mutation*

*should
be "combinatorial"*

TABLE 6. Table of Spearman's rank correlation coefficient for ANN combination analysis.

| | Default | 14 | Shuffle |
|---------|---------|-------|---------|
| Default | 1.000 | 0.026 | -0.099 |
| 14 | | 1.000 | 0.026 |
| Shuffle | | | 1.000 |

Save as above

TABLE 7. RMSE Statistics for MLR results 1-15

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|--------------------------|---------|-------|---------|---------|---------|---------|---------|---------|-------|---------|
| q | 6.4e-05 | 6e-05 | 0.00011 | 7.2e-05 | 6.8e-05 | 5.5e-05 | 5.4e-05 | 5.7e-05 | 9e-05 | 2.9e-05 |
| θ_v (K) | 0.84 | 0.85 | 1.36 | 0.79 | 0.51 | 0.43 | 0.36 | 0.76 | 1.53 | 0.56 |
| U (m s ⁻¹) | 0.28 | 0.34 | 0.81 | 0.57 | 0.49 | 0.34 | 0.49 | 0.49 | 0.62 | 0.21 |
| V (m s ⁻¹) | 0.30 | 0.72 | 0.43 | 0.50 | 0.55 | 0.60 | 0.39 | 0.39 | 0.70 | 0.23 |

Same as above

TABLE 8. NRMSE Statistics for MLR results 1-15

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| q | 0.037 | 0.035 | 0.064 | 0.042 | 0.037 | 0.031 | 0.031 | 0.033 | 0.12 | 0.11 |
| θ_v | 0.04 | 0.05 | 0.06 | 0.04 | 0.03 | 0.02 | 0.02 | 0.04 | 0.16 | 0.15 |
| U | 0.11 | 0.12 | 0.14 | 0.10 | 0.10 | 0.07 | 0.10 | 0.13 | 0.12 | 0.14 |
| V | 0.10 | 0.14 | 0.13 | 0.10 | 0.10 | 0.09 | 0.08 | 0.15 | 0.14 | 0.14 |

*Same as
above*

TABLE 9. RMSE Statistics for MLR results 1-27

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|--------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| q | 9.5e-05 | 8.6e-05 | 0.00013 | 7.6e-05 | 8.2e-05 | 8.6e-05 | 4.9e-05 | 7.1e-05 | 0.00014 | 4.2e-05 |
| θ_v (K) | 0.39 | 0.36 | 0.43 | 0.30 | 0.29 | 0.22 | 0.19 | 0.29 | 0.46 | 0.16 |
| U (m s ⁻¹) | 0.24 | 0.32 | 0.50 | 0.47 | 0.45 | 0.43 | 0.43 | 0.44 | 0.55 | 0.18 |
| V (m s ⁻¹) | 0.31 | 0.47 | 0.33 | 0.39 | 0.40 | 0.55 | 0.41 | 0.36 | 0.66 | 0.21 |

*Same
as above*

TABLE 10. NRMSE Statistics for MLR results 1-27

| Variable | A | B | D | E | F | G | H | L | Range | StdDev |
|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| q | 9.5e-05 | 8.6e-05 | 0.00013 | 7.6e-05 | 8.2e-05 | 8.6e-05 | 4.9e-05 | 7.1e-05 | 0.00014 | 4.2e-05 |
| $\theta_v(\text{K})$ | 0.39 | 0.36 | 0.43 | 0.30 | 0.29 | 0.22 | 0.19 | 0.29 | 0.46 | 0.16 |
| $U(\text{m s}^{-1})$ | 0.24 | 0.32 | 0.50 | 0.47 | 0.45 | 0.43 | 0.43 | 0.44 | 0.55 | 0.18 |
| $V(\text{m s}^{-1})$ | 0.31 | 0.47 | 0.33 | 0.39 | 0.40 | 0.55 | 0.41 | 0.36 | 0.66 | 0.21 |

826 **LIST OF FIGURES**

| | | |
|--------------------|---|----|
| 827 Fig. 1. | Hidden Node Analysis | 50 |
| 828 Fig. 2. | Map of Cadarache valley location in South-Eastern France. ©OpenStreetMap contributors. | 51 |
| 829 Fig. 3. | Map of Cadarache valley. The LEMS locations are marked by the bold letters. The top half of the figure displays an example of typical thermally driven flow in the Cadarache Valley and the bottom half of the figure displays an example of typical synoptic flow in the Cadarache Valley. The magenta arrows represent typical wind speeds and directions. Geoportail.gouv.fr [institut national de l'information géographique et forestière (IGN)] | 52 |
| 834 Fig. 4. | Time series plots for experimental data and ANN prediction for 15 January 2017 | 53 |
| 835 Fig. 5. | Time series plots for experimental data and ANN prediction for 27 January 2017 | 54 |
| 836 Fig. 6. | Example of combination analysis for the virtual potential temperature. For every combination of three input LEMS and nine output LEMS, three statistics were determined, which can be seen in the plot above. Some input combinations produce bad results, but that's only because one output (out of the nine) performed poorly. For this figure, the ANN weights were randomly initialized. | 55 |
| 841 Fig. 7. | Time series plots for experimental data and ANN prediction for 15 January 2017 | 56 |
| 842 Fig. 8. | Time series plots for experimental data and ANN prediction for 27 January 2017 | 57 |

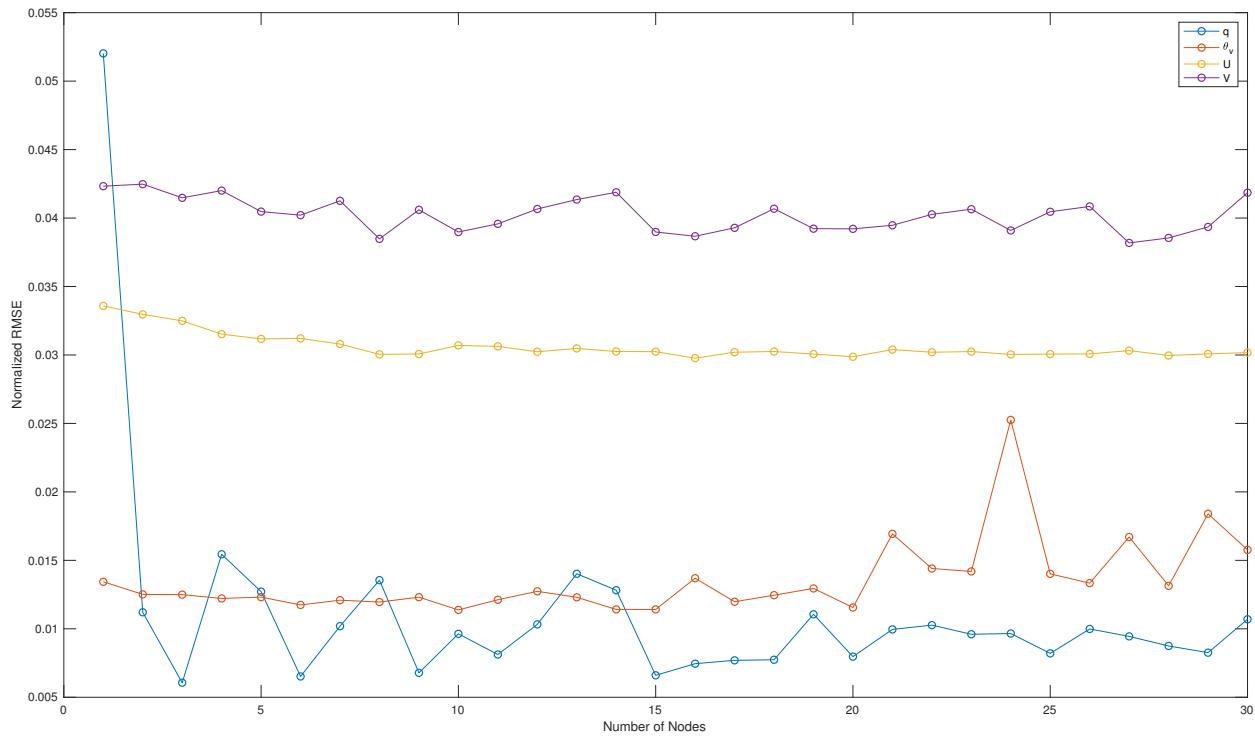


FIG. 1. Hidden Node Analysis

Illustration of the impact of adding hidden nodes to the base ANN described in Sect 4.1.



please provide
more descriptive
caption. The reader
should know what this
figure is without the
paper! for example

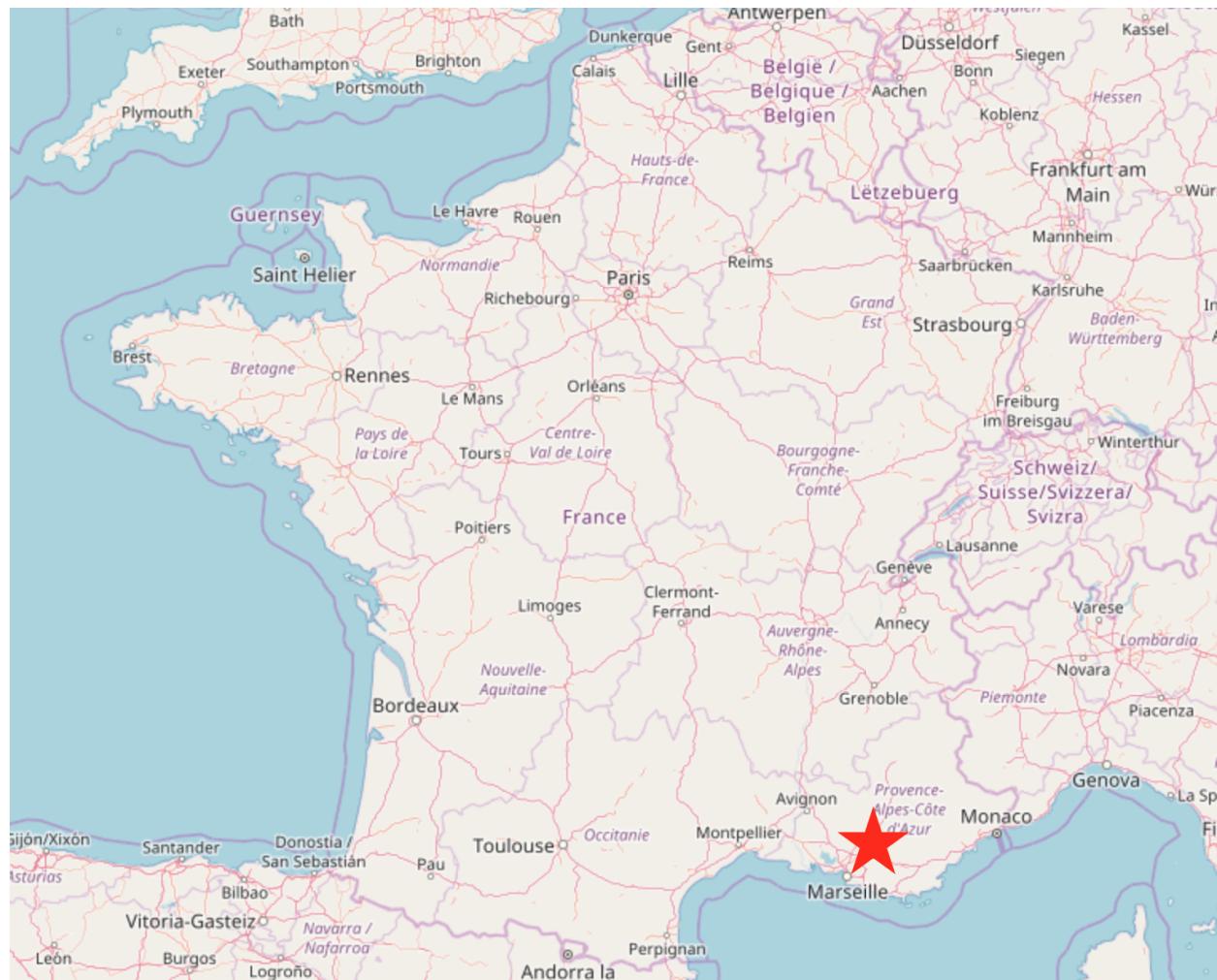


FIG. 2. Map of Cadarache valley location in South-Eastern France. ©OpenStreetMap contributors.

map of France highlighting the Cadarache
valley experimental site (red star).
in

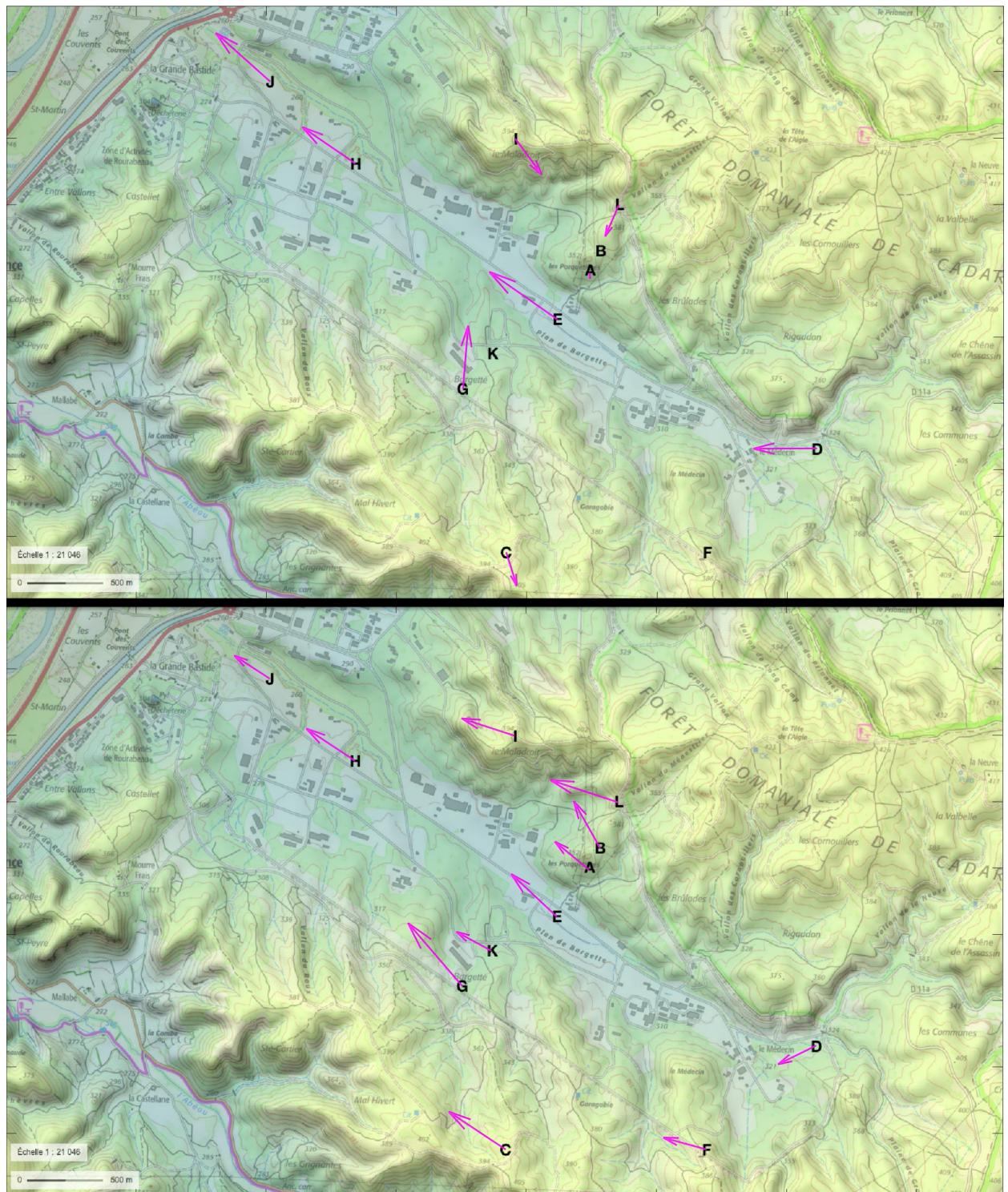


FIG. 3. Map of Cadarache valley. The LEMS locations are marked by the bold letters. The top half of the figure displays an example of typical thermally driven flow in the Cadarache Valley and the bottom half of the figure displays an example of typical synoptic flow in the Cadarache Valley. The magenta arrows represent typical wind speeds and directions. Geoportail.gouv.fr [institut national de l'information géographique et forestière (IGN)]

at what time of day?
Specify Etes
top it Jan 14th to
map from

Can 16
Can 17
Can 18

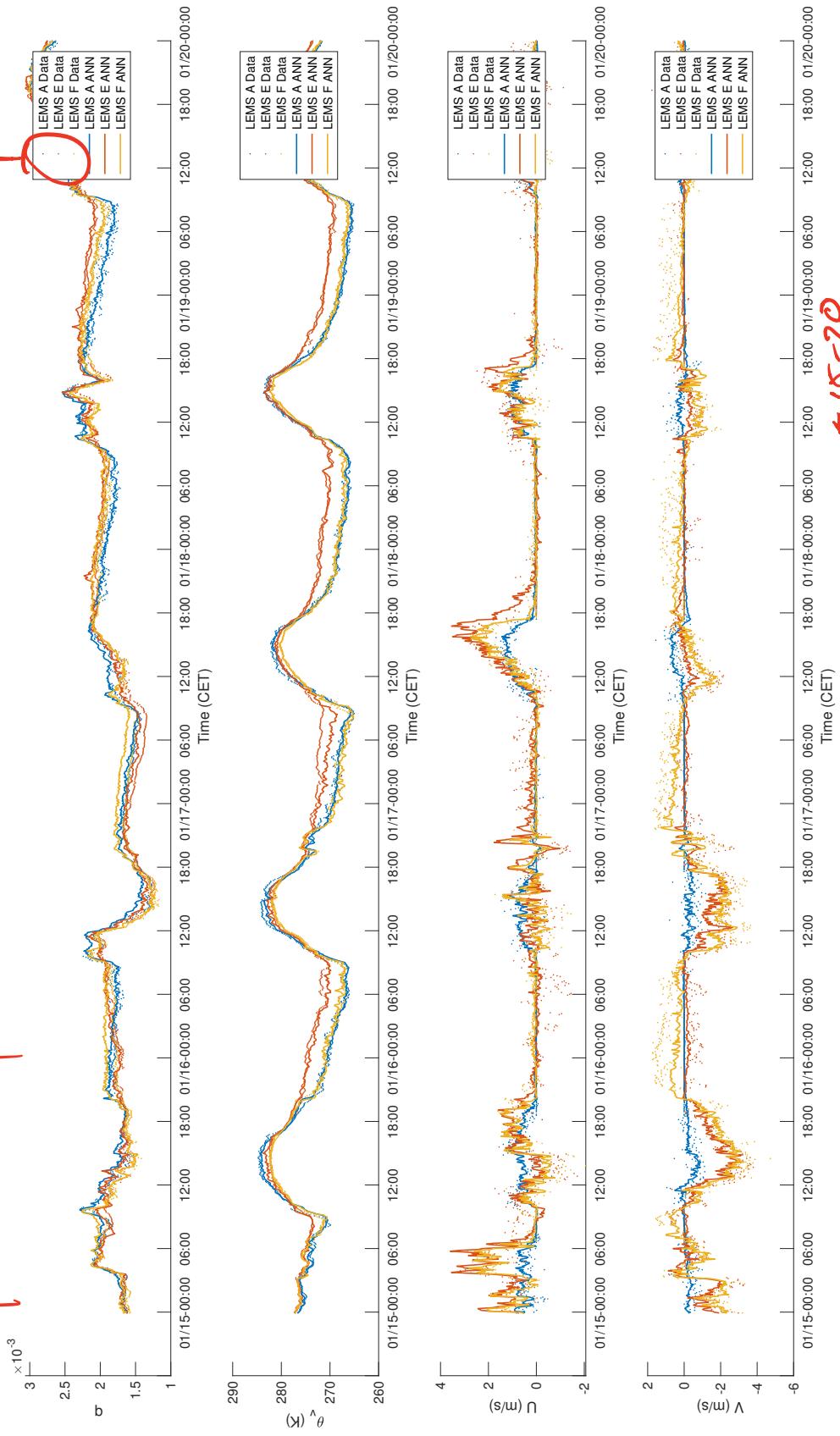


FIG. 4. Time series plots for experimental data and ANN prediction for 15-20 January 2017, our
predicted trendly forced day

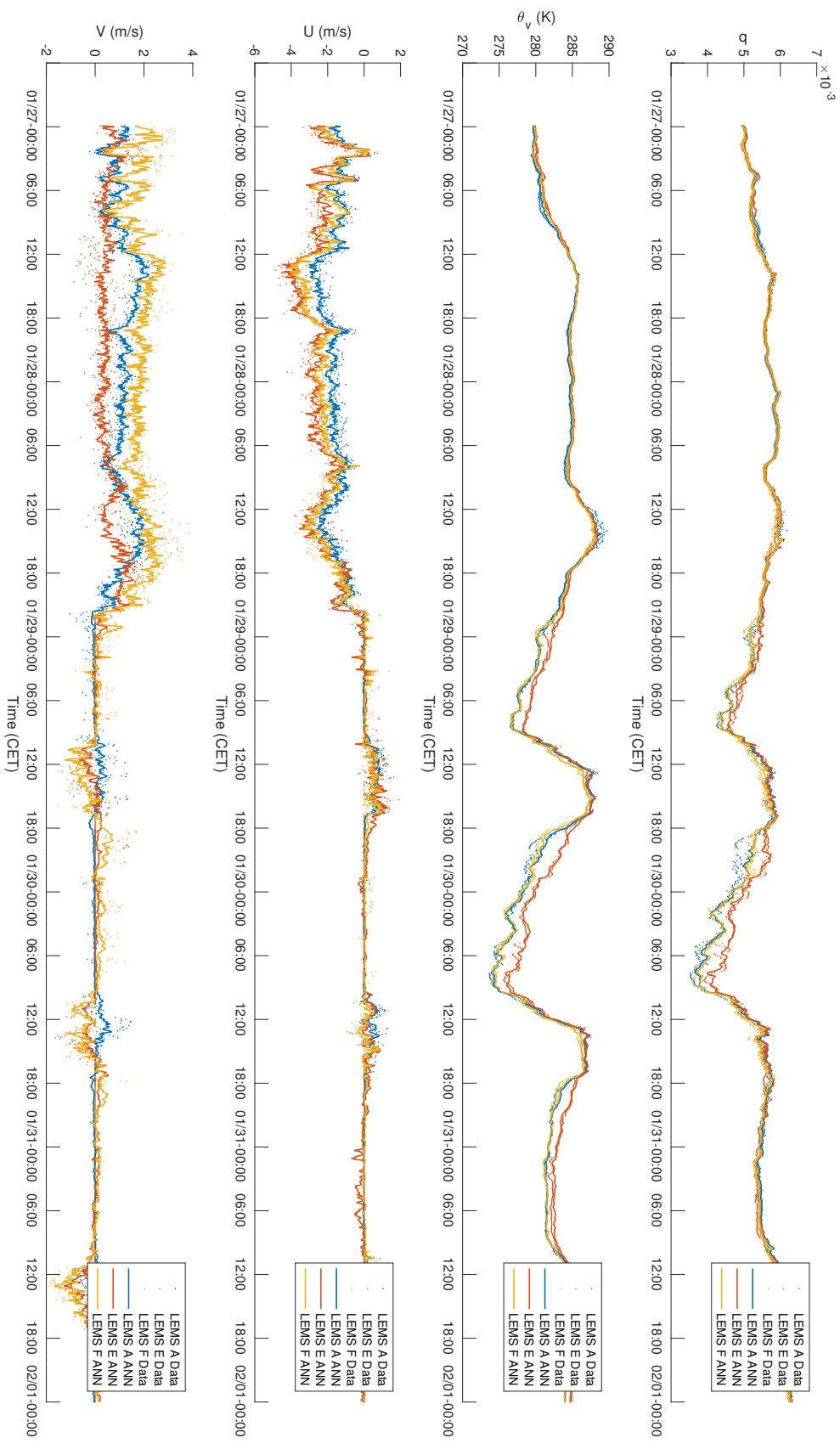
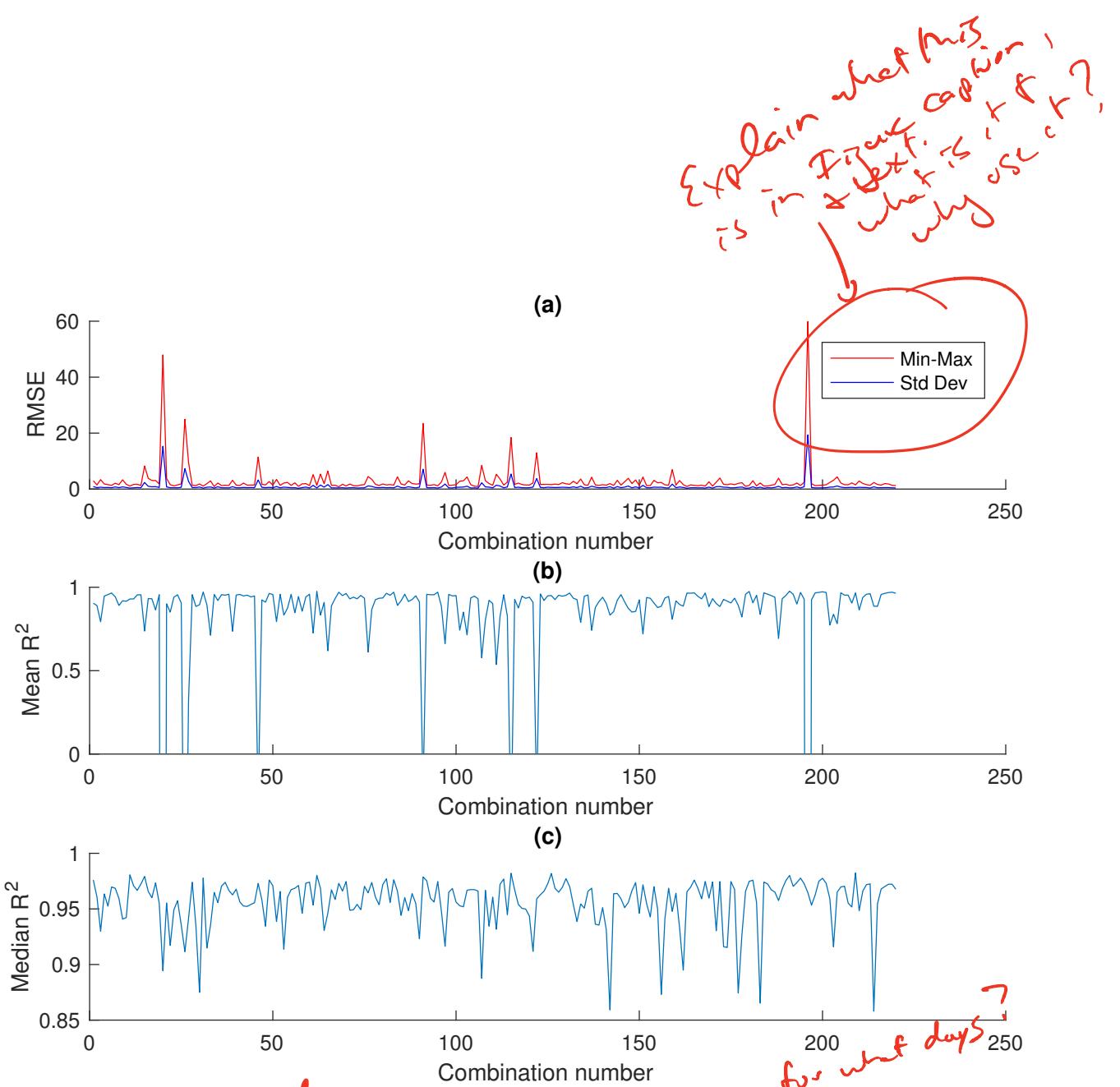


FIG. 5. Time series plots for experimental data and ANN prediction for ~~27~~January 2017

27 January - 1 February 2017



848 FIG. 6. Example of combination analysis for the virtual potential temperature. For every combination of three
 849 input LEMs and nine output LEMs, three statistics were determined, which can be seen in the plot above. Some
 850 input combinations produce bad results, but that's only because one output (out of the nine) performed poorly.
 851 For this figure, the ANN weights were randomly initialized.

I agree, but we should consider
 saying what "bad" is e.g.
 mean $R^2 < 0.75$

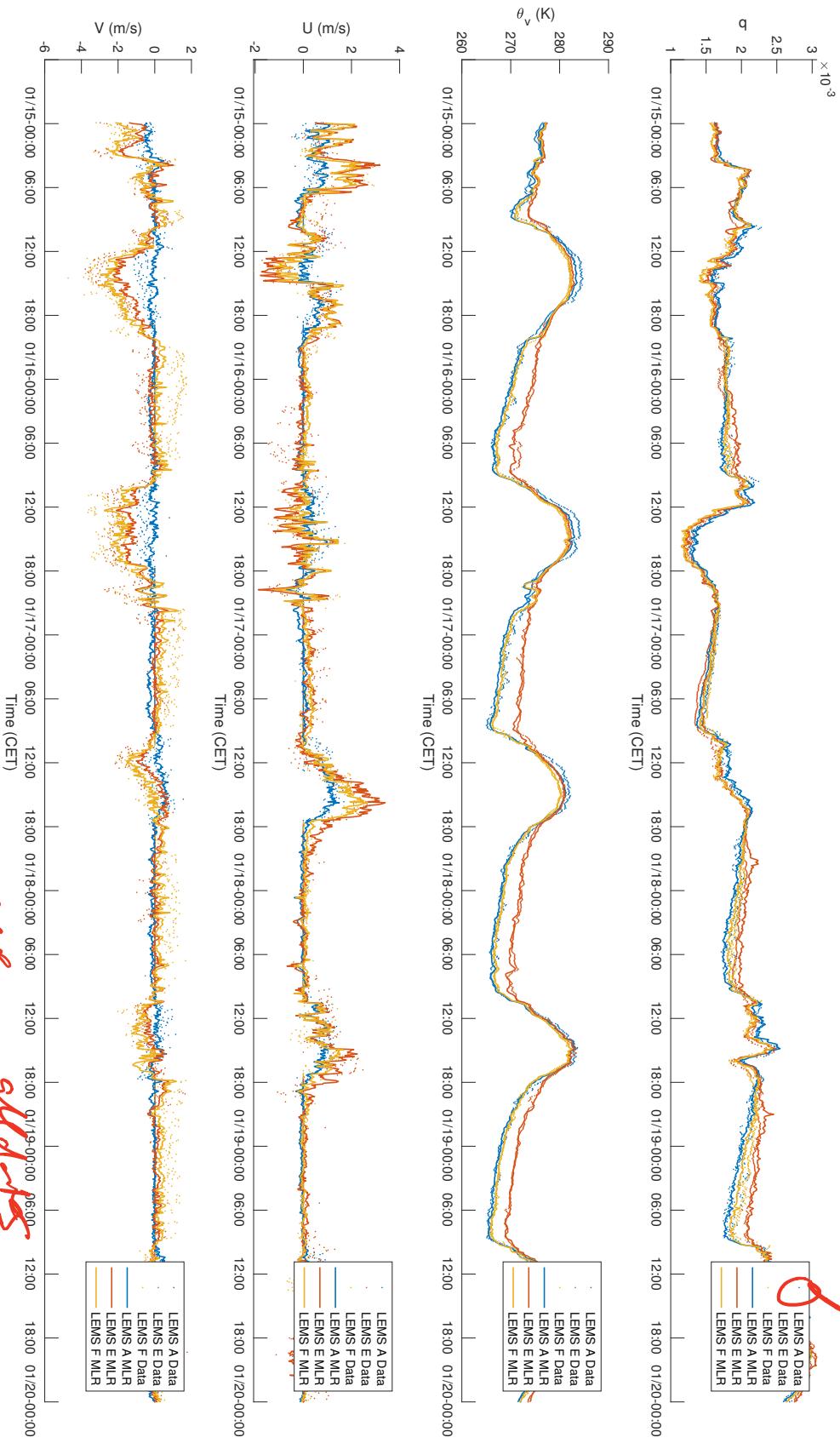


FIG. 7. Time series plots for experimental data and ANN prediction for 15 January 2017, a day with ~~weak synoptic forcing and thermally driven flow,~~ ~~meridional winds~~

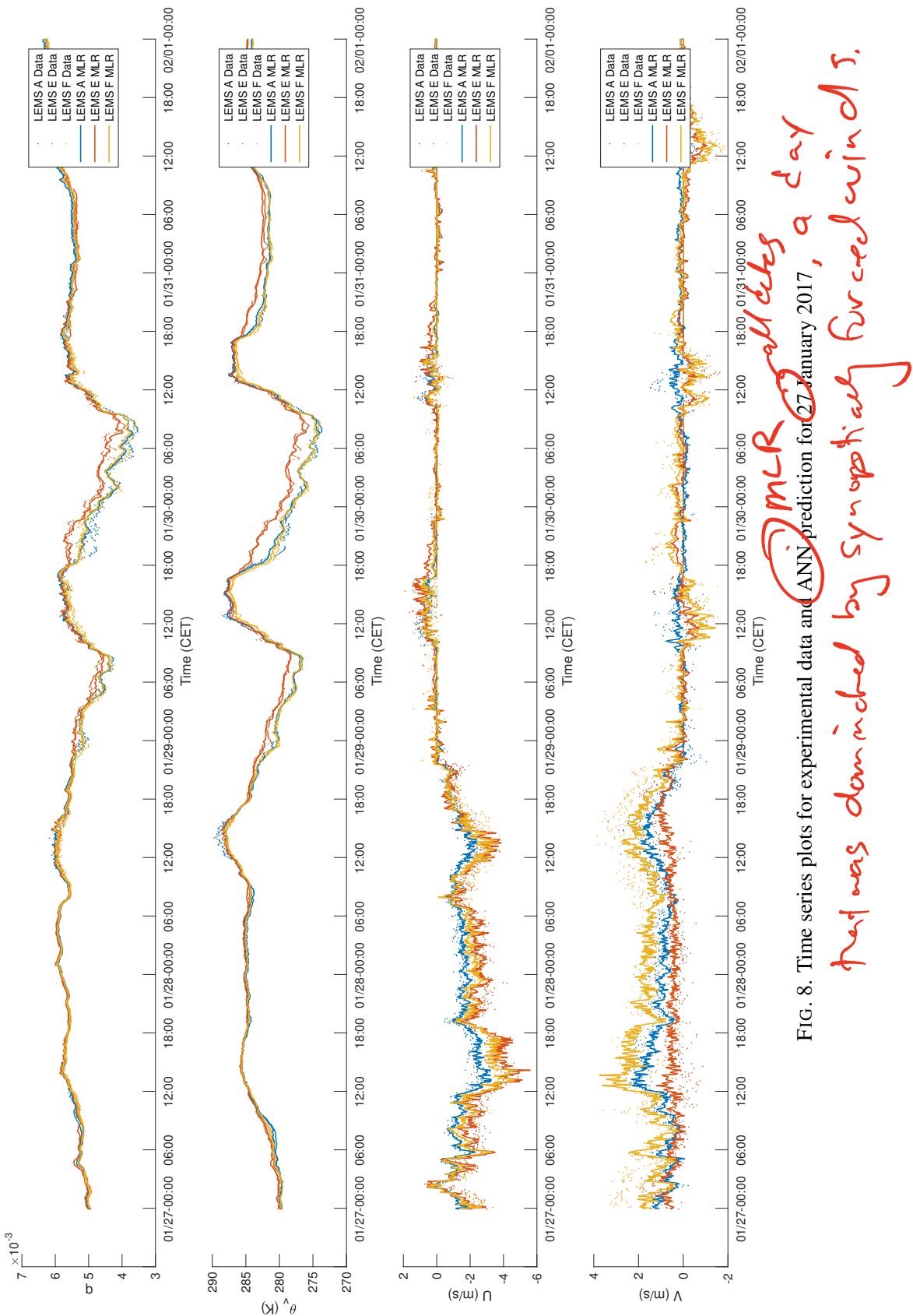


FIG. 8. Time series plots for experimental data and ANN prediction for 27 January 2017, a day
 that was dominated by synoptically forced winds.