



OPEN

Regression-based gap-filling methods show air temperature reductions and wind pattern changes during the 2019 total eclipse in Chile

Arno C. Hammann^{1,2}✉ & Shelley MacDonell^{2,3}✉

Singular disruptive events like solar eclipses affect the measured values of meteorological variables at the earth's surface. To quantify such an impact, it is necessary to estimate what value the parameter would have taken had the event not occurred. We design and compare several methods to perform such an estimate based on longer observational timeseries from individual meteorological surface stations. Our methods are based on regularised regressions (including a Bayesian variant) and provide both a point an associated error estimate of the disruptive event's impact. With their help, we study the effect of the total solar eclipse of July 2nd, 2019, in the Coquimbo Region of Chile, on near-surface air temperatures and winds. The observational data used have been collected by the meteorological surface station network of the Centro de Estudios Avanzados en Zonas Áridas (CEAZA). Most stations inside the eclipse's umbra registered a temperature drop of 1–2 °C, while the most extreme estimated temperature drop surpassed 6 °C. The presence of an 'eclipse cyclone' can neither be proven nor refuted. Application of the regression methods to other comparable problems like volcanic eruptions, forest fires, or simply gap filling of observational data, are conceivable.

Certain singular events, such as solar eclipses, volcanic eruptions or forest fires, cause a disruption to the solar radiation receipt at the earth's surface and throughout the atmospheric column. Other meteorological variables are affected as a result; for example, surface air temperatures may experience a temporary drop. In order to determine how much of a temperature drop (or corresponding perturbations in other variables) is caused directly by the singular event, it is necessary to estimate what the variable's value would have been in its absence. In this study, we apply and compare a number of statistical techniques to the problem of estimating the effect of a specific solar eclipse on near-surface air temperatures and winds, with a particular focus on the associated uncertainties (an aspect almost universally neglected in existing studies). The methods we explore comprise several types of regularised regressions which have not been applied to this problem before, together with some approaches which have previously been described in the literature but whose uncertainties have not been evaluated. While changes in surface air temperatures are commonly reported as a result of solar eclipses, a certain lore has also accumulated around the idea of an eclipse wind^{1,2}, and the particular theory of the eclipse-induced cyclone³ has attracted various attempts at proving or refuting it^{4,5}. Due to the fluctuating nature of wind directions, estimating eclipse-induced changes is particularly challenging. We apply the best-performing method from our tests with temperature to the problem of estimating the eclipse wind. While the methods can be applied to other variables, the quality of our records is best for temperatures and winds, which are also widely reported by previous studies (facilitating comparisons). Other applications to short-duration singular events are conceivable; we want to emphasize that longer-term changes (including climate change) cannot be addressed or would require a retooling of the methods. The specific eclipse during which we collected data occurred on July 2nd, 2019, in the Coquimbo Region of Chile, approximately 400 km north of the capital Santiago; it was an eclipse of Saros cycle 127.

Quantitative accounts of the meteorological effects of solar eclipses can be found in the literature as early as 1834¹, and include records of observations from the earth's surface layer, the stratosphere⁶ and the ionosphere⁷.

¹Asiaq - Greenland Survey, Qatserisut 8, Nuuk, Greenland. ²Centro de Estudios Avanzados en Zonas Áridas, Raúl Bitrán 1305, La Serena, Chile. ³Waterways Centre for Freshwater Management, University of Canterbury and Lincoln University, Private Bag 4800, Christchurch, New Zealand. ✉email: ach@asiaq.gl

C1	15:23
C2	16:38
Max. occultation	16:39
C3	16:40
C4	17:47
Sunset	17:56
t_{\bullet}	15:20–19:00
\bar{u}, \bar{v}	17:00–17:30

Table 1. ‘Contact’ times of the eclipse of July 2nd, 2019, calculated for La Serena ($29^{\circ}54'16.3''$ S $71^{\circ}14'56.2''$ W) on <http://xjubier.free.fr/en/> and rounded to the nearest minute, in local time (UTC-4). Also given are the time of sunset (calculated on <https://planetcalc.com/>) and time ranges referred to in the text: ‘ t_{\bullet} ’ refers to the times removed from statistical calculations to avoid training data containing effects of the eclipse, and ‘ \bar{u}, \bar{v} ’ refers to the time span over which wind data is averaged.

Many restrict themselves to documenting the evolution of measured quantities over the course of the eclipse, especially if instrumentation has been put in place only for the event itself^{6,8–14}. Attempts to diagnose what part of the system’s evolution is due to the direct impact of an eclipse have also been made, usually by comparison to its prior or posterior state^{15–23} or to a physics-based model which excludes the eclipse-specific radiative forcing. Both simpler column models of the atmospheric surface layer^{24,25} and complete numerical weather prediction models (NWP) have been used. Studies using NWPs have either compared simulations with and without eclipse-specific radiative forcing^{26,27}, or simulations without eclipse forcing and observations made during an eclipse^{2,4}.

If a sufficient number of parameters can be measured, they can be causally related through physical models and an eclipse’s impact can be directly described. For the atmospheric surface layer, important variables include radiative and turbulent fluxes of energy, and potentially also advective fluxes and some measure of the boundary layer depth. Under such conditions, the reduction in radiative heating of the earth’s surface can be connected directly to the response of the atmospheric boundary layer^{12–14,21,28–32}. Unfortunately, such a complete set of measurements is rarely available (it was not in our case), and the eclipse’s effect needs to be deduced rather than observed. Provided that the variable of interest is observed, the fundamental problem is to obtain an estimate of how this quantity would have evolved in the eclipse’s absence. We call this estimated ‘uneclipsed’ evolution the *reference*. Existing studies frequently use a quantity’s value at or before First Contact as a reference^{19,20,26,33,34}. First Contact (‘C1’) denotes the time point when the two discs of sun and moon appear to touch for the first time to an observer at a given geographic location; Fourth Contact (‘C4’) is the time point when the contact between the discs ceases. Second (‘C2’) and Third (‘C3’) Contact bracket totality and are not defined for partial eclipses. Taking into account the likely existence of a lag between radiative forcing and atmospheric response, one might estimate the eclipse’s effect as the difference between a temperature maximum reached shortly after C1 and the minimum (likely reached slightly after maximum occultation), i.e. as the temperature *range* during the eclipse³⁵. Another slight variation is to use a short-term mean of values before C1 and after C4^{8,36}. Such approaches may be appropriate outside of transition periods like sunrise¹⁶ and -set, and where the variable of interest does not typically exhibit large fluctuations on short time scales.

Complete days preceding or following the day of an eclipse are another popular choice to obtain reference values^{37–40}, and ‘similar’ days in a more general sense have also been used^{32,41}. This overcomes the issue with using persistence as estimation principle during rapid transition periods (sunrise/-set) and is useful if little information is available otherwise as to how a quantity (e.g. air pressure⁴², winds^{5,40}, ozone¹⁰, aerosol concentrations^{41,43}, radar reflectivity⁴⁴, atmospheric electric potential⁴⁵) typically evolves. An absence of large and rapid fluctuations, however, is nonetheless desirable for estimates based on single reference days to be meaningful. While solar radiation at the earth’s surface is subject to some variability due to short-term variations in aerosol optical depth⁴⁶, the impact of an eclipse is quantitatively much larger and can therefore be estimated fairly accurately (as long as clouds do not complicate the picture)^{28,40,47–49}.

If the measured quantity exhibits large intrinsic variability, some form of regularisation of the reference estimate is necessary. Averaging over several reference days can provide such regularisation^{29,43,50,51}, as can the smoothing of eclipse-day observations. Linear interpolation of the variable’s values between C1 and C4^{31,35} may be the simplest variant of such smoothing; higher-order polynomials⁵², splines⁵³, local regression (under exclusion of observations around the time of the eclipse)⁵ and fits to variable-specific parametric models⁹ have also been used.

Regularisation of the reference estimate was particularly important for the eclipse of July 2nd, 2019, which occurred so close to sunset (Table 1) that many air temperature timeseries do not show a clear rise in temperatures between C4 and sunset, as they typically do if the eclipse occurs in the middle of the day. Additionally, Chile’s Coquimbo region borders the cold coastal waters of the Humboldt current, and slight changes in wind direction can lead to sudden drops in local air temperatures due to the advection of cold air from the marine boundary layer onto land. The magnitude and rate-of-change of such advective temperature changes are indistinguishable from those registered during the eclipse.

Results

Terminology. All discussions below apply to timeseries of data from a single sensor, in particular screen level air temperature. We think of such a timeseries as a data matrix X_{dt} indexed by a calendar date d (rows; ‘date axis’) and a time t (columns; ‘time axis’). The time period affected by the eclipse will be given the symbol t_{\bullet} ; this includes the full duration of the eclipse (C1–C4), plus an additional time span beyond C4 to account for some dynamical lag in the system (see Table 1). Conversely, all times not affected by the eclipse are symbolised by t_{\circ} . Analogously, the day of the eclipse will be denoted d_{\bullet} and all other days by d_{\circ} . When we talk about an average over a number of days, we mean that the averaging is performed over the date dimension and the result is a one-dimensional array indexed by times (a daily cycle). All data used in this study comes from the meteorological surface station network operated by the Centro de Estudios Avanzados en Zonas Áridas (CEAZA) in La Serena, Chile, from here on called the CEAZAMet station network.

Evaluation of the regression methods with respect to 2m air temperatures. Our methods are generalisations of two types of estimates found in the literature: (1) averages over a number of ‘similar’ days, and (2) smoothed observations from the day of the eclipse with interpolated values for t_{\bullet} ^{1,2}. Approach (1) can formally be seen as a regression with selected days $d \in d_{\circ}$ serving as predictors and equal regression weights which sum to 1. Dropping the restrictions on the weights and allowing for an intercept generalises the approach, and any regression method with subset selection replaces the task of choosing ‘similar’ days with an objective procedure. Including an intercept can be interpreted as allowing for slowly-varying (synoptic) background conditions. Since we need to judge the similarity of days without considering t_{\bullet} , this regression is trained on t_{\circ} even though we are interested in t_{\bullet} , for which the model is simply evaluated. Approach (2) can equally be cast as a regression, with $t \in t_{\bullet}$ as targets and $t \in t_{\circ}$ as predictors, trained on examples $d \in d_{\circ}$. This procedure is equivalent to a smoothing along the time axis with the kernel’s weights fit by regression instead of having been determined a priori. We will call the two approaches ‘average’ and ‘smoothing’ type regressions, respectively. In both cases we can make use of the additional information we have from years of monitoring data at the stations at which we observe the eclipse; we use approximately 4 years of data preceding the eclipse, since for this time period, most CEAZAMet stations in the eclipse’s umbra have been collecting data at a 5 min sampling rate.

We use the lasso as a regression method with subset selection (for both the ‘average’ and the ‘smoothing’ type); the predictors are selected by including an ℓ^1 penalty on the vector of regression weights during the loss minimization, which effectively sets many of the weights to zero⁵⁴. For comparison, we evaluate unweighted averages over N days similar to the eclipse day (‘N-averages’ from now on), as well as standard smoothing procedures (local regression, in both linear⁵ and quadratic variants). Similarity between predictor and target days for the N-average is assessed as the pairwise Euclidean distance in the space spanned by times t_{\circ} . Uncertainties for all methods can be quantified empirically by estimating the values for all $t \in t_{\bullet}$ on all days $d \in d_{\circ}$ and subtracting the corresponding actual observations. This gives an error distribution for each time $t \in t_{\bullet}$. For the smoothing regression—the only method where the quantity of interest to us is actually the target—care needs to be taken to separate training and test data. To this end, we use four-fold cross-validation, so that approximately 1 year of data is in the test set at any time, and we calculate the error for a day only if it is in a test set.

Figures 1 and 2 show the regression-based estimates and those obtained from comparison methods, respectively, for air temperature on the day of the eclipse at a few selected stations in and around La Serena. The box plots show the temperature depression ΔT of the observation relative to the reference estimate. The plot elements representing uncertainty (shaded bands, boxes, whiskers and fliers) are constructed by adding the error distributions computed as described above to the point estimate. The regression methods’ uncertainty is much reduced relative to that of the comparison methods; some but not all estimates of ΔT fall outside of 95% of the calculated error distributions.

Besides the empirical estimates for uncertainties, we also employ a Bayesian regression technique which estimates a full posterior distribution. Certain specific regularisation penalties in classical regression correspond to particular distributions of priors over the parameters in a Bayesian setting⁵⁴, and we chose automatic relevance determination (ARD)⁵⁵ as a method that selects predictors based on the posterior distributions of their associated coefficients. The agreement between the estimates and their uncertainties obtained from the three methods (Figs. 3, 4, 5) is encouraging. Note that a Bayesian regression would be meaningless in the ‘average’ direction, since the time range of interest t_{\bullet} is not actually the target in that case.

Stronger temperature depressions due to the eclipse are generally found further inland (Figs. 3, 4), which is expected due to the thermal inertia of the sea. Only one station close to the coast shows a larger ΔT estimate (Fray Jorge Quebrada)—but it should be noted that the timing of the largest temperature drop is late there compared to most stations (Fig. 5), suggesting it may have been related to the sunset rather than the eclipse (the station is separated from the sea and setting sun by a mountain range). There is a hint of a suggestion that moving away from the center line of the eclipse reduces the experienced temperature drop, especially outside of the region of totality, but the proximity to the coast appears dominant. The largest estimated temperature drop of more than 6° C (Cachiyuyo) might raise suspicions, but it turns out that its magnitude is largely a result of its occurrence at the end of the day, when temperatures are dropping very fast even on regular days and the eclipse’s effect is more akin to an earlier onset of nightfall (cf. Fig. 7).

The eclipse wind. Encouraged by the results for the surface temperatures, we apply the smoothing type lasso also to the wind field. The average type regression gives quantitatively very similar results but is both much more costly computationally (at least if uncertainties are to be estimated) and harder to analyse since its setup differs from regular regression frameworks in that the information of interest (times t_{\bullet}) are not actually the regression’s target. The results (not shown) with respect to uncertainties are very similar to those for the tem-

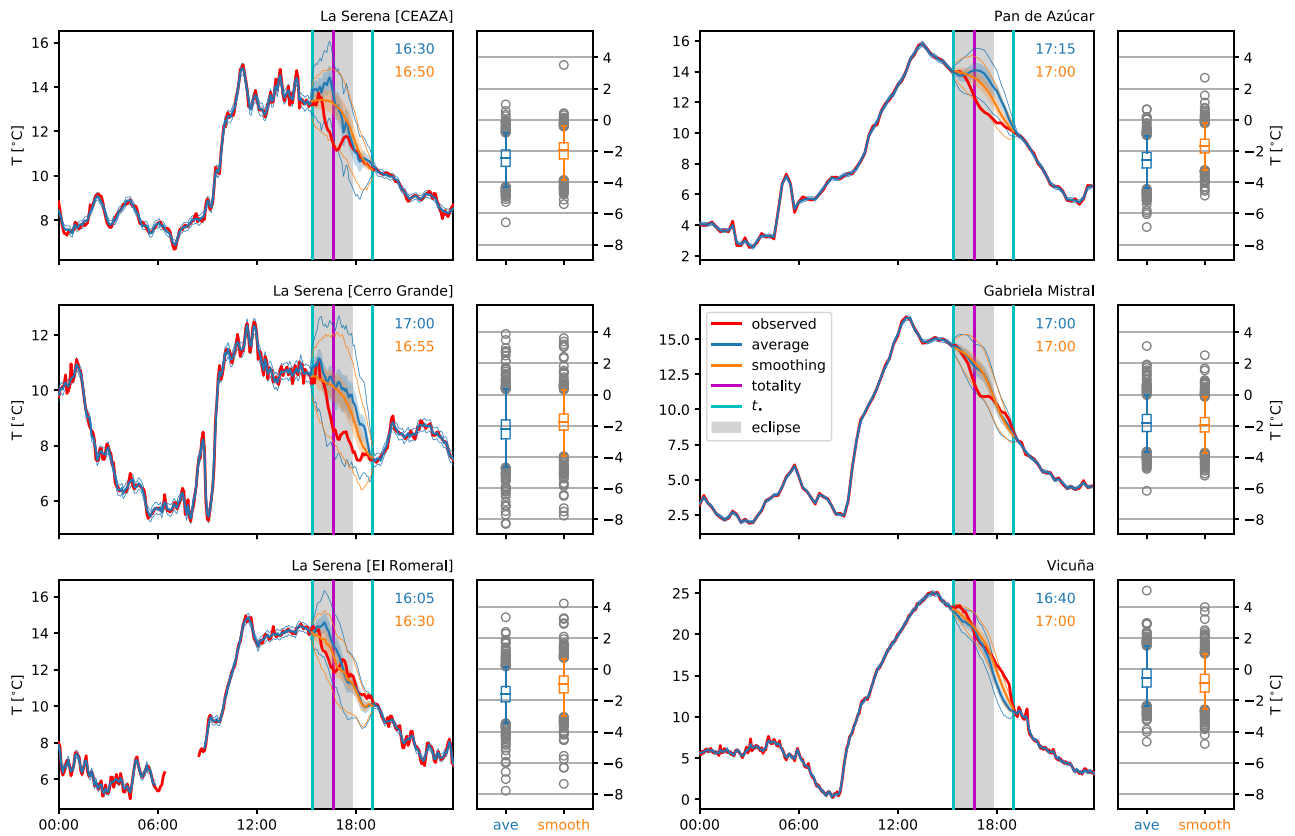


Figure 1. Observed 2m air temperatures on the day of the eclipse (July 2nd, 2019), together with estimates of the ‘unclipped’ temperature evolution, for selected stations in the vicinity of La Serena, Chile (left plot of each panel). The reference estimates are obtained from average (‘ave’) and smoothing (‘smooth’) style lasso regressions. Box plots (right) show maximum temperature depression for each estimation method, with corresponding times in the right upper corners of main plot. Boxes (right) and shaded bands (left) correspond to the IQR, whiskers (right) and thin lines (left) to 95% of the data.

perature case, with some but not all stations achieving a statistical significance of the estimated eclipse effect at 95%; we performed the ARD variant of the regression as well with comparable results.

Instead of focusing on uncertainties as in the temperature case, we simply consider the point estimates for the change in the wind field induced by the eclipse. We applied the regression to the u , v wind vectors separately and calculate an ‘eclipse wind’ following Clayton³ by subtracting the regression to the observations on eclipse day. A two-dimensional spatial picture can be constructed, again following Clayton, by plotting a station’s relative latitudinal distance to the center path of the eclipse in y , and replacing the time of an observation by a virtual distance along the eclipse’s path calculated with the assumption of a propagation speed of the eclipse of approximately 2000 km h^{-1} (Fig. 6).

Clayton hypothesised that an eclipse provided an optimal occasion for the assessment of earlier theories by Ferrel regarding the structure of cyclones with a cold core⁵⁶. The essence of the argument is that a cool surface temperature anomaly will lead to surface outflow of air and replacement by subsiding motion, with ensuing inflow aloft and possibly a closing of the vertical cell at some distance from the center. Outflow is associated with anticyclonic and inflow with cyclonic rotation, but air which has acquired cyclonic angular momentum aloft will conserve it when subsiding. The outflow therefore has to first overcome the existing cyclonic momentum before crossing a circle of zero rotation and turning anticyclonic at some distance from the center of subsidence. Clayton surmised, however, that because of friction, the cyclonic rotation of the innermost ring would be too weak to be observable, and focused on demonstrating an anticyclonic sense of rotation further away.

The wind field is very variable and a result of many interacting forces, including local to continental-scale contrasts in surface heat fluxes and the effects of synoptic weather systems. The signal-to-noise ratio of any eclipse-induced effects may be rather low, and Clayton’s results have not been reproduced convincingly^{2,5}. Instead, where changes have been demonstrated, explanations other than his proposed ‘cyclone’ have been favoured⁴. It should also be noted that Ferrel’s and Clayton’s ideas about cyclones predate the modern understanding of midlatitude and tropical cyclones as primarily the result of dynamical instabilities by half a century.

A noteworthy feature of the eclipse wind in Fig. 6 is the ‘bow shock’ on the eastern (C1) front, where air seems to move out of the way of the arriving eclipse. Since the eclipse travels with supersonic speed, it is impossible for any eclipse-caused outflow to result in air parcels moving ahead (eastward) of the eclipse. Remember, however, that while the frame of reference for the wind vector *locations* in Fig. 6 can be considered to be moving with the eclipse, the same is not true for the coordinate system of the quivers, which is both isotropic and static. The

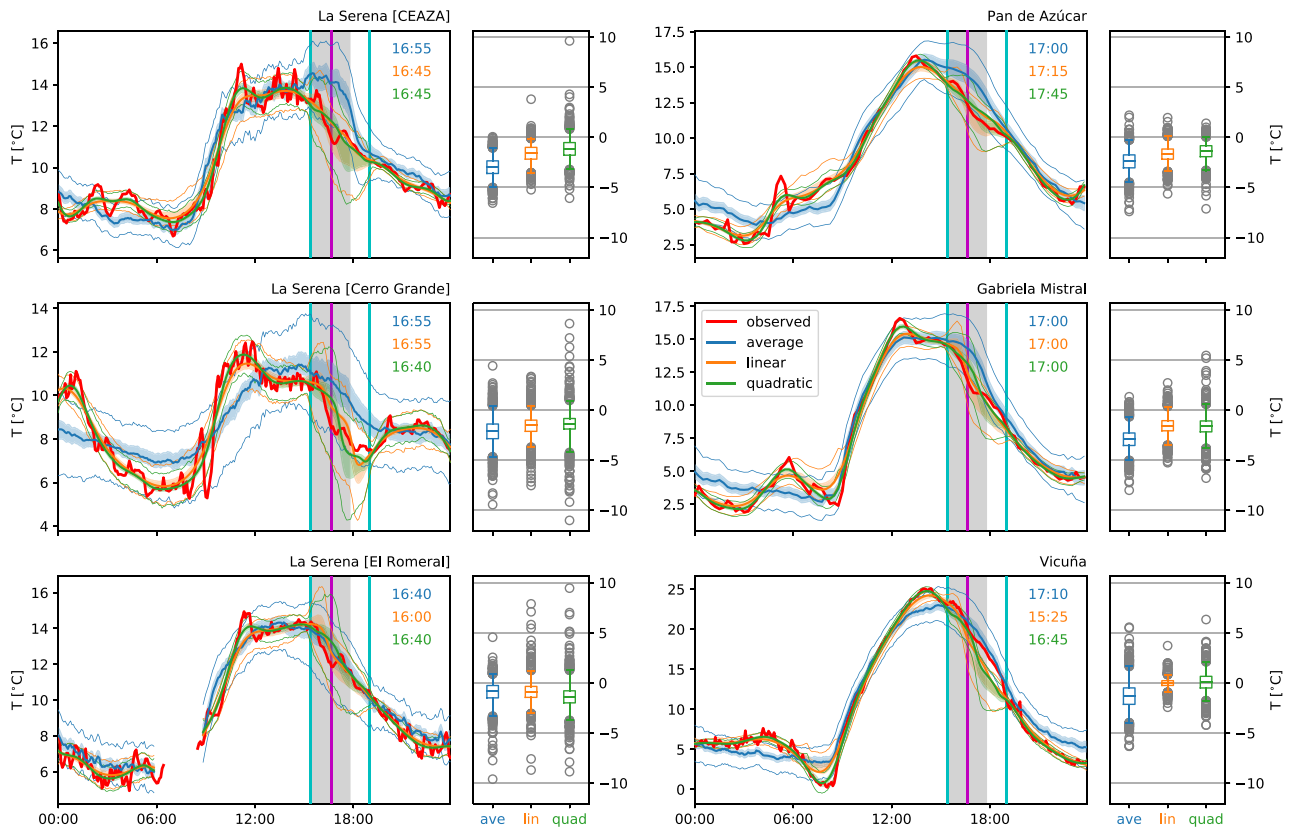


Figure 2. Observed 2m air temperatures and reference estimates as in Fig. 1, but with estimates obtained from unweighted averages over a number of ‘similar’ days and local linear and quadratic regression (a type of smoothing).

fact that eastward flow occurs at some stage during the eclipse is therefore not a contradiction. (As an aside, the strong anisotropy of the virtual along-path direction related to the eclipse’s supersonic movement casts doubt on Clayton’s original argument and his Fig. 4, which represents the eclipse cyclone as a fairly symmetric structure.) With much imagination, one might discern an anticyclonic movement of the air masses following C1, which later turns into a cyclonic movement around the time of totality. This would imply an initial anticyclonic outflow, followed by subsidence which carries with it a cyclonic vorticity stemming from the required inflow aloft. It could further be argued that the flow field appears more turbulent after the passing of totality than before, analogously to that in the wake of a fast-moving solid object.

Without more detailed analysis, which is beyond the scope of this study, we cannot confirm or refute any of the preceding interpretations; and others are equally possible: for example, it appears that zones of converging and diverging flow alternate in the penumbral zone, with convergence on the northern edge coexisting with divergence to the south and vice versa. Similarly strong gradients in surface heating as across the edge of the umbra may be found over ocean eddies⁵⁷ and fronts^{58,59}, resulting in wind field adjustments which could be compared to the eclipse case. However, the supersonic movement of the umbral shadow is likely to result in a very different dynamical adjustment of the atmosphere.

In the Coquimbo Region, the wind field is influenced by the Southeast Pacific Subtropical Anticyclone, passing frontal systems, a strong land-sea breeze and equally strong orographically driven wind systems. Around La Serena, both the land-sea and the mountain-valley systems tend to produce westerly winds during the day and easterlies at night, but differences in reversal timing may lead to complex interactions. The eclipse occurred just before sunset, and an earlier-than-usual reversal of the daily wind systems due to the associated decreased surface heating is consistent with the mostly easterly direction of the eclipse wind in Fig. 6.

Discussion

After witnessing the eclipse in La Serena on July 2nd, 2019, we were interested in finding out how much the atmosphere’s surface layer had cooled during the event. Observers reported that they felt colder during the eclipse than before, but that does not necessarily imply that the air temperature dropped considerably, since much of the sensation may be due to an experienced reduction in direct radiative heating of the body. CEAZAMet operates a monitoring network with many stations inside the belt of totality of this eclipse which provided us with near-surface air temperature data. However, no ‘obvious’ pattern of change, such as a pronounced dip in the temperature curve, is consistently discernible at all or a majority of the stations. The eclipse occurred so close to sunset that temperatures had started dropping before C1 and little radiative energy was available after C4 to raise temperatures again. Furthermore, any features of the data from the day of the eclipse which could potentially be

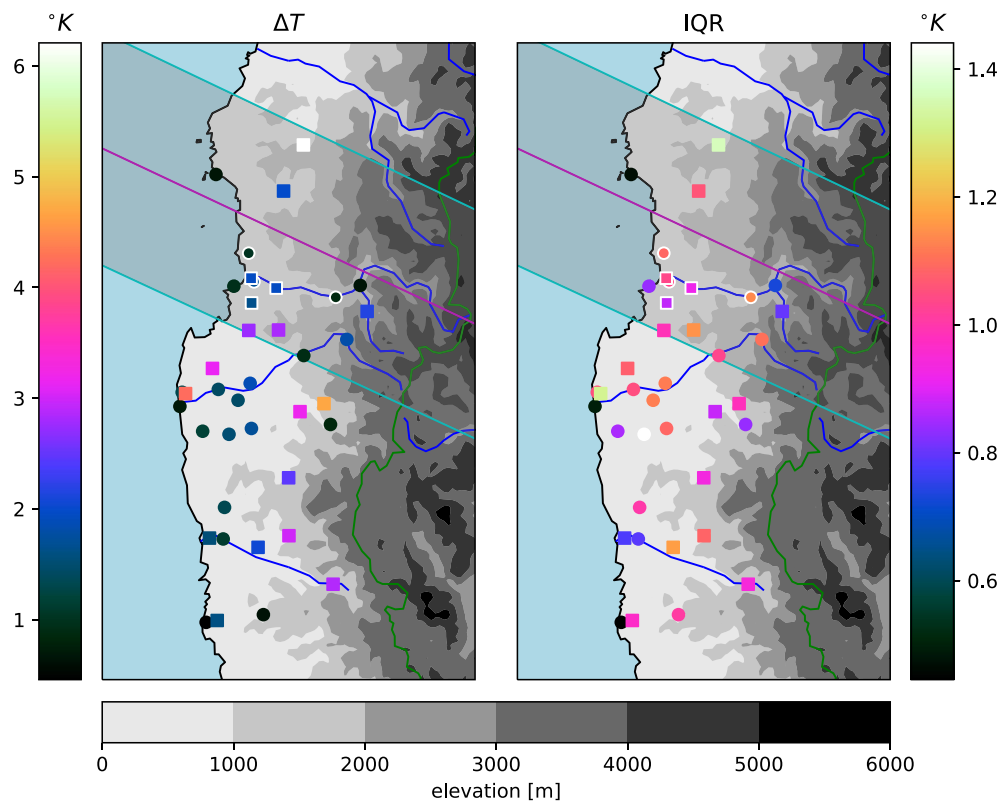


Figure 3. Estimated maximum 2m air temperature depressions (ΔT) during the eclipse (left), based on the smoothing type lasso regression, and IQR of the estimates (right). The lasso estimates for ΔT are corrected for the bias of the error distribution. Square symbols mark estimates which fall outside 95% of the test error distribution for a given station, and white marker edges correspond to the stations shown in Figs. 1 and 2. Cyan lines demarcate the edges of the band of totality and the magenta line its center⁶¹. Gray shading indicates terrain elevation⁶².

interpreted as dips in temperature do not stand out against the general background of intra-diurnal variability in the Coquimbo Region. This is because the dominantly southerly wind alternates frequently between a slight on- and offshore flow, which brings with it strong advective temperature changes since a cold ocean borders a strongly heated land surface.

We therefore chose to focus on the problem of how to separate the effect of the eclipse from other circumstantial influences on the observed meteorological variables. The reference estimate of how a meteorological variable would have behaved on the same day but in absence of the eclipse is speculative, and we cannot confidently use its point value without some idea of how uncertain it is. While some previous studies give confidence intervals on certain specific estimates^{4,5}, these are not related to the uncertainties surrounding the actual effect sizes at specific times and locations. Here, we calculate uncertainties and their bias (for the lasso estimates) empirically on the basis of longer-term records from each of the stations at which we observed the eclipse, by performing our estimation procedure for any non-eclipse day and calculating the errors with respect to the true evolution of that day. The error distribution thus pertains to any arbitrary day of the year and includes uncertainties arising from annual-scale variability; in areas with a more pronounced annual cycle than the Coquimbo region, we would expect the estimates to exhibit larger uncertainties, which in turn could be counteracted by restricting the training data set to days closer to the day to be predicted (in terms of the day of year, DOY) if longer-term records are available.

Calculating the errors associated with particular estimates allows for selecting one with a low uncertainty. The lowest test errors are achieved by the smoothing lasso and the ARD regression (Table 2), but the error for ARD is taken to be simply the posterior standard deviation and may not be entirely accurate since ARD assumes the data to be Gaussian-distributed. It is convenient that the smoothing-type estimates perform the best, since they are both more time-efficient than the average-lasso and take the form of a regular regression. Therefore different regression methods can be used interchangeably and analysed with well-established tools such as cross-validation. All the methods we have described possess an adjustable parameter which we select to approximately minimise a test error; for the lasso regressions it is the regularisation parameter α , for the unweighted average it is the number of days N , and for the local regressions it is the kernel width. The estimates are not overly sensitive to the parameter values, and a complete numerical optimization over all stations is computationally expensive for some methods. Additionally, different stations' estimates have rather different error levels and a general loss

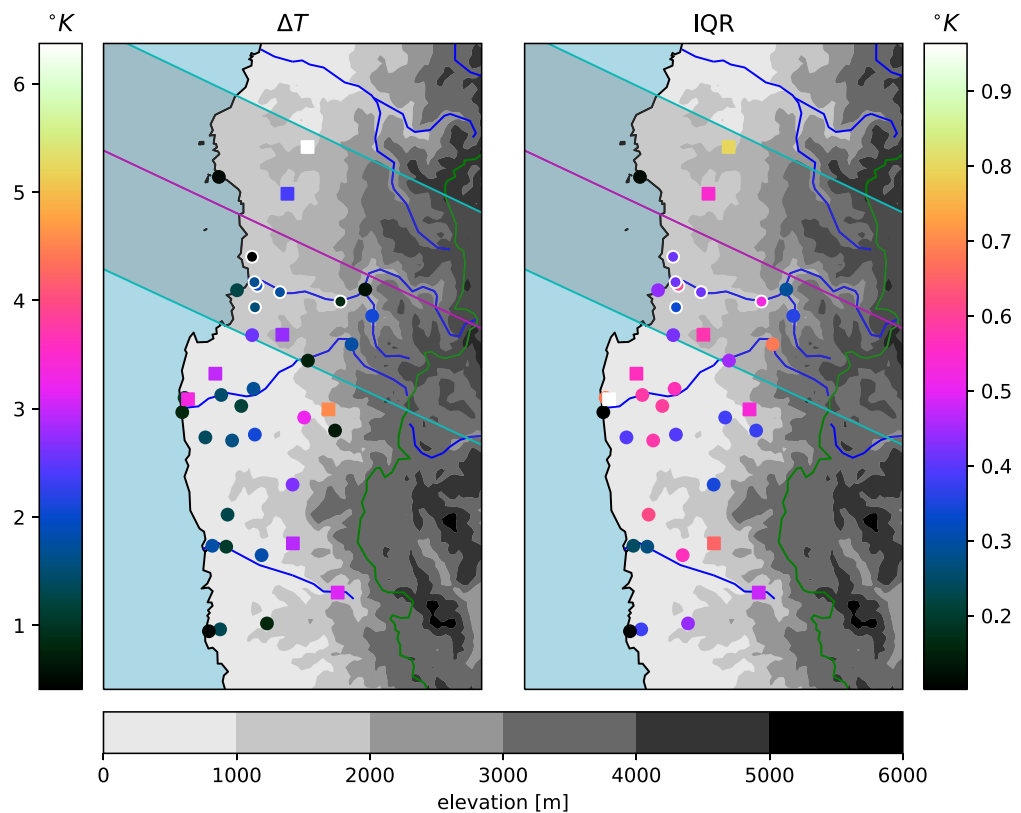


Figure 4. Estimated maximum 2m air temperature depressions as in Fig. 3 (symbols as described there), but for estimates based on the automatic relevance determination (ARD) regression. IQR here is 0.67 times the standard deviation of the predictive distribution.

function would need to be able to weight those appropriately. Therefore, we simply carry out a grid search for a selection of stations and choose roughly appropriate parameter values.

It is conceivable that methods with better performance can be found. We limit ourselves to predicting diurnal temperature cycles by using other diurnal temperature cycles as training data. A regression framework is more typically used to relate one variable to another, and this could certainly be done here; however, contemporaneous multi-variable regressions cannot be used since additional independent variables may also be affected by the eclipse. To the extent that the values of other variables may be able to differentiate locations in state space characterised by very similar temperatures, it could be useful to include them. Similarly, it might be possible to exploit autocorrelations on an inter-diurnal scale. We experimented with non-linear regression methods (Gaussian Processes, Neural Networks) for the more appropriate smoothing setup but found no improvements, which is explained by the dominant influence of only the two data points immediately before and after the prediction period t_* . We also experimented with using the solar elevation angle β as a coordinate instead of time, since it may be expected that it normalise annual variability and guard against the impact of varying day lengths to some extent, but similarly found no improvement to the predictive accuracy of our models while significantly complicating the analysis.

A physics-based model could in principle be initialized to the state of the atmosphere at the start of the eclipse and run predictively with eclipse-unaware radiative forcing in order to provide a reference estimate. If all necessary physical parameters are measured for a specific location, this could be a relatively confined boundary layer energy balance model. However, in our case we have little knowledge of important parameters such as surface albedo, boundary layer thickness and longwave radiation. As we have previously stated, advection plays an important role in the surface energy balance, and either a crude observational estimate based on widely spaced station measurements or a full numerical weather prediction (NWP) model is needed in order to account for it. However, such a physics-based modeling exercise represents a major effort, and we devised our statistical methods primarily in order to provide a simpler alternative. It is furthermore not clear whether the results from a physical model would be less uncertain than those of a statistical one, in particular under the circumstances described for the Coquimbo Region: relatively small swings between on- and offshore winds can have major effects on local temperatures, and the horizontal resolution of NWP models is coarse compared to the complexities of the terrain. Even more importantly, adequate data for model initialization is not available.

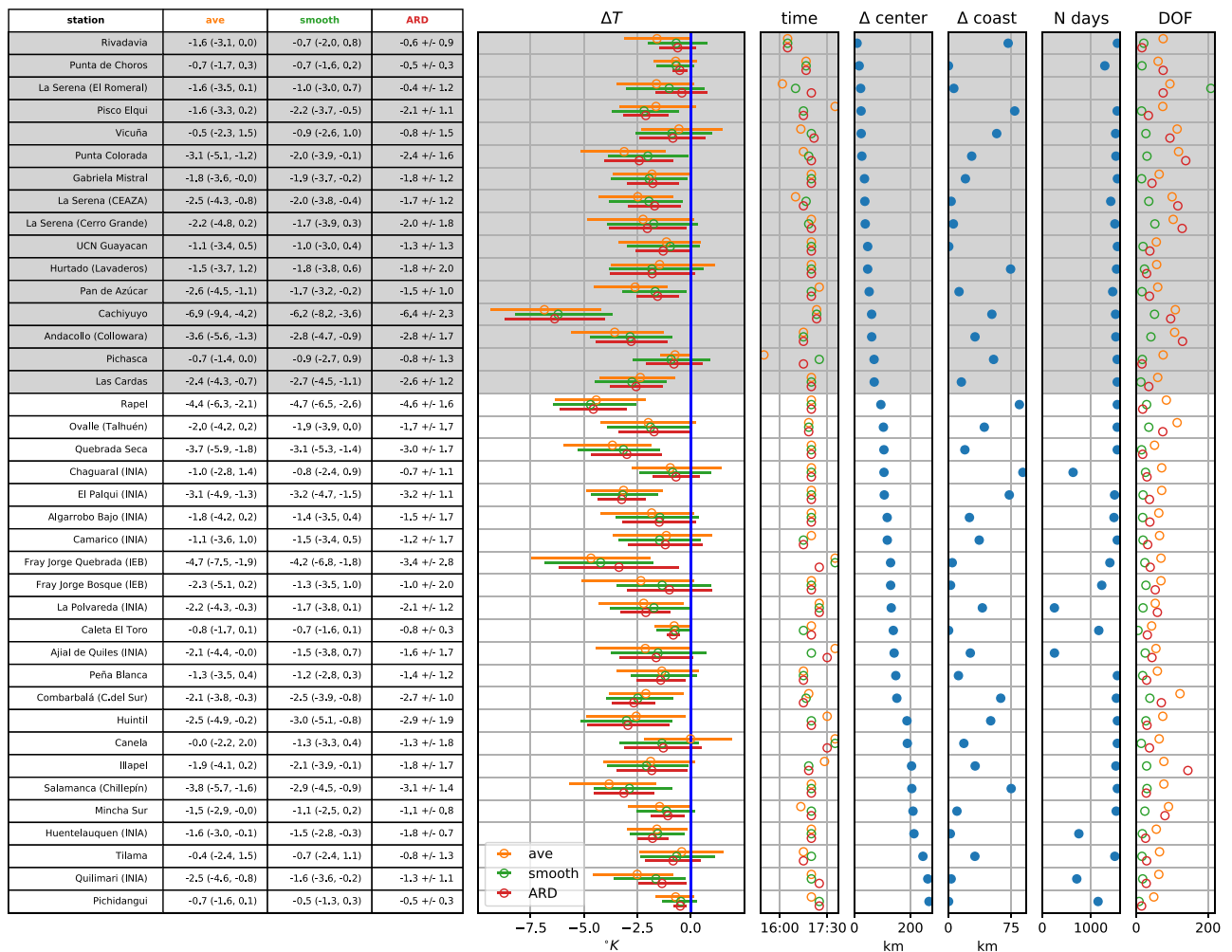


Figure 5. ΔT estimates for stations in the Coquimbo Region with at least 1 year of available data at a resolution of at least 15 min. For lasso regressions (average and smoothing), uncertainties are given as 2.5 and 97.5 percentiles of the test errors separately, while for ARD regression, one value of 1.96 times the standard deviation of the predictive distribution is given. Lasso estimates are corrected for the bias of the error distributions. Errorbars on plots correspond to these ranges. ‘Time’ is the time of maximum ΔT for each type of estimate; ‘ Δ center’ is the distance from a station to the eclipse’s center line; ‘ Δ coast’ the distance to the coast. ‘N days’ is the number of available days on record, and ‘DOF’ are the degrees of freedom (number of regression coefficients) for each estimate.

Methods

In terms of the data matrix X with rows corresponding to days and columns to times, the ‘average’ type regression estimate for d_{\bullet} is given by

$$\hat{Y}_{d_{\bullet},t} = X_{d_{\bullet},t}^T \hat{\beta}_{d_{\bullet},t}$$

where $\hat{\beta}$ is an estimate for the (column) vector of regression coefficients β and \hat{Y} is the regression estimate. This expression corresponds to an unweighted average over N (selected) days if β contains N elements with value $1/N$ and otherwise zeros. Viewed as a regression, an intercept is included by adding a column of ones to X^T , and the estimate $\hat{\beta}_{d_{\bullet},t_0}$ for β is obtained by minimising a loss function. We denote with the subscripts the fact that $\hat{\beta}$ is estimated by using the observation at d_{\bullet} but only the times t_0 (excluding the eclipse) as a target.

The ‘smoothing’ approach consists in using the columns of X corresponding to t_{\bullet} as the regression’s targets and all remaining columns X_{dt_0} as predictors:

$$\hat{Y}_{dt_{\bullet}} = X_{dt_0} \hat{\beta}_{d_{\bullet},t_{\bullet}}$$

In this case, all non-eclipse days d_0 on record serve as training samples, and estimates are obtained for any day but only for t_{\bullet} . The regression coefficients β in this case correspond to weights of a smoothing kernel applied to the day of the eclipse, where the values at t_{\bullet} have been removed from and filled in by the smoothing procedure.

We calculate a distribution over prediction errors by subtracting the observed value at t_{\bullet} for all non-eclipse days from the corresponding estimates:

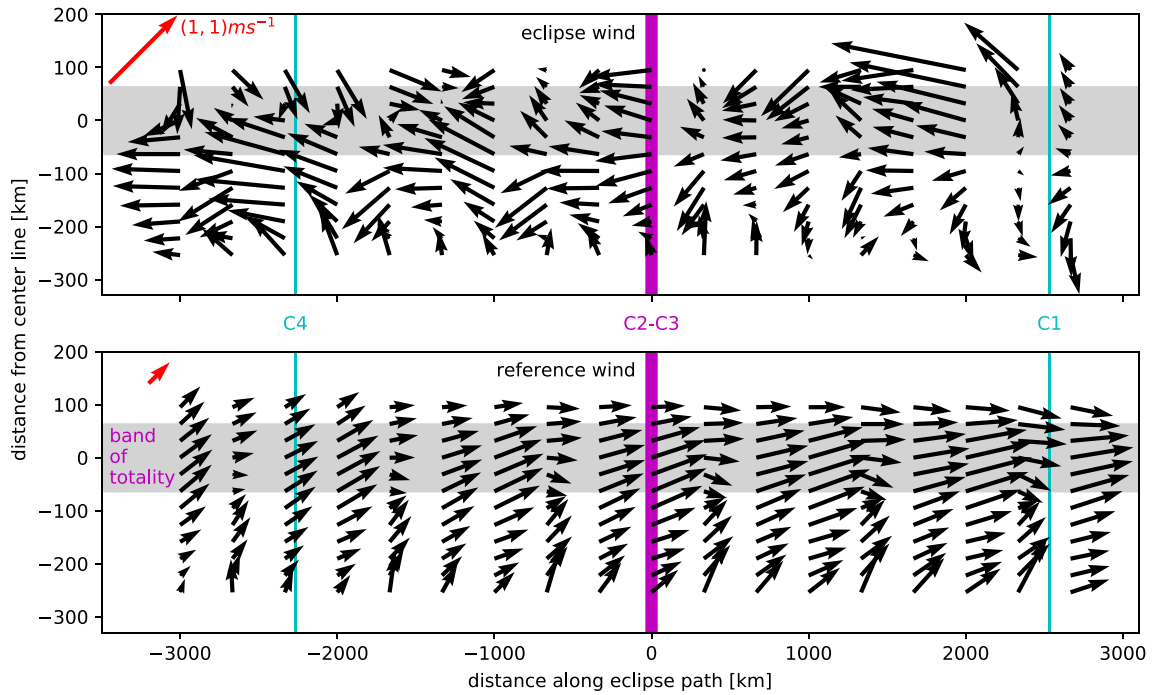


Figure 6. Vector wind as a function of distance along (x) the path of the eclipse and from the center line (y). Eclipse wind is the observed wind vector minus the reference estimate. The x distance is calculated from observation time assuming an umbral velocity of 2000 km h^{-1} ; the positive direction points west to east as usual since the eclipse’s movement is also in this sense. The field is smoothed with biquadratic radial basis functions. While the x, y locations are non-square, the arrow lengths are isotropic; red arrows show the scale of the wind with $u = v = 1 \text{ ms}^{-1}$.

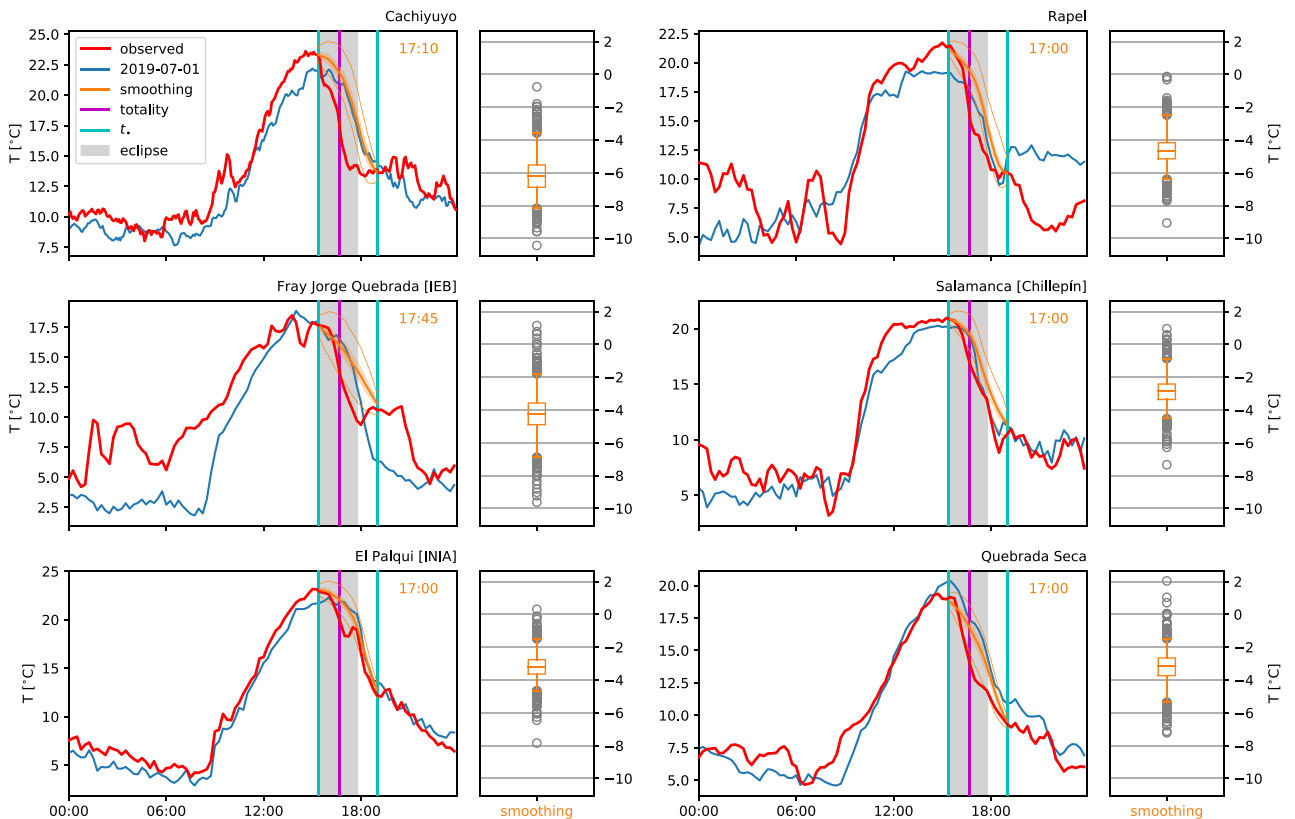


Figure 7. Observed 2 m air temperatures and smoothing lasso reference estimates for the day of the eclipse as in Fig. 1, but for the stations with the largest ΔT estimates. Additionally, the observed temperatures during July 1st, 2019 (the day preceding the eclipse) are shown, to give an idea of the realism of the estimates.

Method	Err
Unweighted average	1.15
Linear local regression	1.18
Quadratic local regression	1.0
'Average' lasso	0.9
'Smoothing' lasso	0.8
'Smoothing' ARD	0.61

Table 2. Test errors over all stations and all non-eclipse days for the methods discussed in this paper. Since we perform ARD only for a subset of times out of t_{\bullet} , we restrict all errors to this subset. The ARD error is the posterior standard deviation averaged over times and stations. The other errors are computed to be comparable with this, as the RMSE over days in d_{\circ} , averaged over times and stations.

Method	Hyperparameter	T2	u, v
Unweighted average	Number of days	10	
'Average' lasso	α	0.005	
Local linear/quadratic regression	Length scale (min)	50	
'Smoothing' lasso	α	0.1	0.05
'Smoothing' ARD	λ Threshold	10^4	10^4

Table 3. Values and types of the hyperparameters for the different estimation methods described.

$$PE = \hat{Y}_{d_{\circ}t_{\bullet}} - X_{d_{\circ}t_{\bullet}}.$$

In the 'average' type approach, the regression's target are the observations $X_{d_{\circ}t_{\bullet}}$ and information from t_{\bullet} never enters the calculations. We can hence think of all observation-prediction pairs for the eclipse as a test set and compute statistics on PE in the same way as we would on the observations themselves.

The 'smoothing' type regression is a standard setup where targets $X_{d_{\circ}t_{\bullet}}$ and estimates $\hat{Y}_{d_{\circ}t_{\bullet}}$ are paired, and we can use cross-validation to divide the date dimension d into training and test sets. Since we operate with approximately 4 years of data, we use 4-fold cross-validation, such that the test sets consist of roughly a full year of data. The distribution of PE is taken to be the set of its values for all days d_{\circ} calculated whenever a given day is in a test set. The final estimate of the variable's eclipse-unaffected value at the day of the eclipse is calculated from all days d_{\circ} .

The lasso, which we apply to both regression problems, is a regularised regression whose loss function contains a ℓ^1 penalty on the vector of regression coefficients β . It is this form of the penalty that leads to the coefficients for less influential predictors to be set to zero, and thereby to subset selection⁵⁴. The penalty is scaled by a hyperparameter, denoted α here, which controls how many predictors are culled: the larger the value of α , the more coefficients become zero, and the lower the resulting model complexity. We choose its value by approximately minimising a root mean square test error (RMSE) computed by averaging PE^2 over t_{\bullet} and d_{\circ} . The regression estimate itself is not overly sensitive to the precise value of α , and so we simply choose a single value that is approximately optimal for a number of meteorological stations (see Table 3).

We compare the lasso estimates with estimates based on simple averaging and smoothing. For the averaging, the data matrix's date dimension is first ordered according to the mean square difference over times t_{\circ} between each day and eclipse day. Then, the N closest days, in this Euclidean distance sense, are averaged. A lower value of N corresponds to a larger value of α (stronger regularisation and fewer non-zero coefficients). As with α , we select a value for N which approximately minimizes the average test error over all stations.

For smoothing, we use local linear and quadratic regression with a Gaussian kernel⁵⁴. The period t_{\bullet} is removed before the smoothing estimate is computed and the resulting gap filled in by the estimate. The governing parameter for local regressions is a length scale, which is again fit by approximately minimizing the average test error over all stations.

An additional regression type, linear Bayesian regression with automatic relevance determination (ARD), is applied to the smoothing problem. The prediction error is here estimated in the form of a full predictive distribution rather than empirically for a point estimate. Since the eclipse period t_{\bullet} is what we want to compute any errors for and the average type regression does not use it as a target, the Bayesian regression does not really provide any useful information in this case. ARD regularises complexity by pruning predictors whose distributions' precision λ (the inverse of the standard deviation) exceeds a threshold value⁵⁵. In contrast to the lasso, however, the numerical value of this threshold has almost no effect on the results and can be left at a default setting (see Table 3). Note, however, that the implementation used⁶⁰ supposes Gaussian distributions for both the data and the regression coefficients (and therefore also the predictive error distribution), an assumption that is only approximately valid in our case.

Data availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Received: 25 July 2021; Accepted: 8 February 2022

Published online: 11 May 2022

References

- Aplin, K. L., Scott, C. J. & Gray, S. L. Atmospheric changes from solar eclipses. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20150217. <https://doi.org/10.1098/rsta.2015.0217> (2016).
- Gray, S. L. & Harrison, R. G. Diagnosing eclipse-induced wind changes. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **468**, 1839–1850. <https://doi.org/10.1098/rspa.2012.0007> (2012).
- Clayton, H. H. The eclipse cyclone and the diurnal cyclones. results of meteorological observations in the solar eclipse of May 28, 1900. *Proc. Am. Acad. Arts Sci.* **36**, 307. <https://doi.org/10.2307/20021016> (1901).
- Gray, S. L. & Harrison, R. G. Eclipse-induced wind changes over the British Isles on the 20 March 2015. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20150224. <https://doi.org/10.1098/rsta.2015.0224> (2016).
- Eugster, W. *et al.* Effects of vernal equinox solar eclipse on temperature and wind direction in Switzerland. *Atmos. Chem. Phys.* **17**, 14887–14904. <https://doi.org/10.5194/acp-17-14887-2017> (2017).
- Quiroz, R. S. & Henry, R. M. Stratospheric cooling and perturbation of the meridional flow during the solar eclipse of 7 March 1970. *J. Atmos. Sci.* **30**, 480–488. [https://doi.org/10.1175/1520-0469\(1973\)0303c0480:SCAPOT3e2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)0303c0480:SCAPOT3e2.0.CO;2) (1973).
- Zhang, S.-R. *et al.* Ionospheric bow waves and perturbations induced by the 21 August 2017 solar eclipse. *Geophys. Res. Lett.* **44**, 12067–12073. <https://doi.org/10.1002/2017GL076054> (2017).
- Aplin, K. & Harrison, R. Meteorological effects of the eclipse of 11 August 1999 in cloudy and clear conditions. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **459**, 353–371. <https://doi.org/10.1098/rspa.2002.1042> (2003).
- Anderson, R., Keefer, D. & Myers, O. Atmospheric pressure and temperature changes during the 7 March 1970 solar eclipse. *J. Atmos. Sci.* **29**, 583–587 (1972).
- Kolev, N. *et al.* Aerosol Lidar and in situ ozone observations of the planetary boundary layer over Bulgaria during the solar eclipse of 11 August 1999. *Int. J. Remote Sens.* **26**, 3567–3584. <https://doi.org/10.1080/01431160500076939> (2005).
- Uddin, W., Joshi, B., Kumar, T. S., Sharma, S. & Sagar, R. Observations of total solar eclipse of 29 March 2006 and related atmospheric measurements. *Curr. Sci.* **93**, 957–959 (2007).
- Stewart, R. B. & Rouse, W. R. Radiation and energy budgets at an arctic site during the solar eclipse of July 10, 1972. *Arctic Alpine Res.* **6**, 231–236. <https://doi.org/10.2307/1550088> (1974).
- Mauder, M., Desjardins, R. L., Oncley, S. P. & MacPherson, I. Atmospheric response to a partial solar eclipse over a cotton field in central California. *J. Appl. Meteorol. Climatol.* **46**, 1792–1803. <https://doi.org/10.1175/2007JAMC1495.1> (2007).
- Foken, T. *et al.* Micrometeorological measurements during the total solar eclipse of August 11, 1999. *Meteorol. Z.* **10**, 171–178. <https://doi.org/10.1127/0941-2948/2001/0010-0171> (2001).
- Fernández, W., Castro, V. & Hidalgo, H. Air temperature and wind changes in Costa Rica during the total solar eclipse of July 11. *Earth Moon Planets* **63**(133–147), 1991. <https://doi.org/10.1007/BF00575102> (1993).
- Fernández, W., Hidalgo, H., Coronel, G. & Morales, E. Changes in meteorological variables in Coronel Oviedo, Paraguay, during the total solar eclipse of 3 November 1994. *Earth Moon Planets* **74**, 49–59. <https://doi.org/10.1007/BF00118721> (1996).
- Nymphas, E., Adeniyi, M., Ayoola, M. & Oladiran, E. Micrometeorological measurements in Nigeria during the total solar eclipse of 29 March, 2006. *J. Atmos. Solar Terr. Phys.* **71**, 1245–1253. <https://doi.org/10.1016/j.jastp.2009.04.014> (2009).
- Stoev, A. D., Stoeva, P. V., Kiskinova, N. & Stoyanov, N. Evolution of the basic micrometeorological parameters during the total solar eclipse of 29 March 2006 at Manavgat, Turkey. In *SPIE Proceedings*, Vol. 6936 (eds Matvienko, G. G. & Banakh, V. A), 69361Z. <https://doi.org/10.1117/12.783768> (2008).
- Pleijel, H. Observations of temperature and air humidity during the total solar eclipse 29 March 2006 at Side, Turkey. *Meteorol. Z.* **18**, 107–109. <https://doi.org/10.1127/0941-2948/2008/0324> (2009).
- Sjöblom, A. A solar eclipse seen from the High Arctic during the period of midnight sun: effects on the local meteorology. *Meteorol. Atmos. Phys.* **107**, 123–136. <https://doi.org/10.1007/s00703-010-0070-3> (2010).
- Venkat Ratnam, M., Shravan Kumar, M., Basha, G., Anandan, V. & Anandan, A. Effect of the annular solar eclipse of 15 on the lower atmospheric boundary layer over a tropical rural station. *J. Atmos. Solar Terr. Phys.* **72**(1393–1400), 2010. <https://doi.org/10.1016/j.jastp.2010.10.009> (2010).
- Hanna, E. *et al.* Meteorological effects of the solar eclipse of 20 March 2015: Analysis of UK Met Office automatic weather station data and comparison with automatic weather station data from the Faroes and Iceland. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20150212. <https://doi.org/10.1098/rsta.2015.0212> (2016).
- Pasachoff, J. M., Peñaloza-Murillo, M. A., Carter, A. L. & Roman, M. T. Terrestrial atmospheric responses on Svalbard to the 20, Arctic total solar eclipse under extreme conditions. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**(20160188), 2016. <https://doi.org/10.1098/rsta.2016.0188> (2015).
- Segal, M., Turner, R. W., Prusa, J., Bitzer, R. J. & Finley, S. V. Solar eclipse effect on shelter air temperature. *Bull. Am. Meteorol. Soc.* **77**, 89–99. [https://doi.org/10.1175/1520-0477\(1996\)0773c0089:SEEOSA3e2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)0773c0089:SEEOSA3e2.0.CO;2) (1996).
- Szałowski, K. The effect of the solar eclipse on the air temperature near the ground. *J. Atmos. Solar Terr. Phys.* **64**, 1589–1600. [https://doi.org/10.1016/S1364-6826\(02\)00134-7](https://doi.org/10.1016/S1364-6826(02)00134-7) (2002).
- Founda, D. *et al.* The effect of the total solar eclipse of 29 March 2006 on meteorological variables in Greece. *Atmos. Chem. Phys.* **7**, 5543–5553. <https://doi.org/10.5194/acp-7-5543-2007> (2007).
- Gross, P. & Hense, A. Effects of a total solar eclipse on the mesoscale atmospheric circulation over Europe—A model experiment. *Meteorol. Atmos. Phys.* **71**, 229–242. <https://doi.org/10.1007/s007030050057> (1999).
- Pruitt, W. O., Lourence, F. & Crawford, T. V. Radiation and energy balance changes during the eclipse of 20 July 1963. *J. Appl. Meteorol.* **4**, 272–278. [https://doi.org/10.1175/1520-0450\(1965\)0043c0272:RAEBCD3e2.0.CO;2](https://doi.org/10.1175/1520-0450(1965)0043c0272:RAEBCD3e2.0.CO;2) (1965).
- Bala Subrahmanyam, D. *et al.* over Thiruvananthapuram, India. *Bound. Layer Meteorol.* **141**, 325–332. <https://doi.org/10.1007/s10546-011-9627-z> (2011).
- Bhat, G. & Jagannathan, R. Moisture depletion in the surface layer in response to an annular solar eclipse. *J. Atmos. Solar Terr. Phys.* **80**, 60–67. <https://doi.org/10.1016/j.jastp.2012.02.025> (2012).
- Kameda, T., Fujita, K., Sugita, O., Hirasawa, N. & Takahashi, S. Total solar eclipse over Antarctica on 23 November 2003 and its effects on the atmosphere and snow near the ice sheet surface at Dome Fuji. *J. Geophys. Res.* **114**, D18115. <https://doi.org/10.1029/2009JD011886> (2009).
- Eaton, F. D. *et al.* Solar eclipse effects observed in the planetary boundary layer over a desert. *Bound. Layer Meteorol.* **83**, 331–346. <https://doi.org/10.1023/A:1000219210055> (1997).
- Hanna, E. Meteorological effects of the solar eclipse of 11 August 1999. *Weather* **55**, 430–446. <https://doi.org/10.1002/j.1477-8696.2000.tb06481.x> (2000).

34. Ahrens, D., Iziomon, M. G., Jaeger, L., Matzarakis, A. & Mayer, H. Impacts of the solar eclipse of 11 August 1999 on routinely recorded meteorological and air quality data in south-west Germany. *Meteorol. Z.* **10**, 215–223. <https://doi.org/10.1127/0941-2948/2001/0010-0215> (2001).
35. Clark, M. R. On the variability of near-surface screen temperature anomalies in the 20 March 2015 solar eclipse. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20150213. <https://doi.org/10.1098/rsta.2015.0213> (2016).
36. Anderson, J. Meteorological changes during a solar eclipse. *Weather* **54**, 207–215. <https://doi.org/10.1002/j.1477-8696.1999.tb06465.x> (1999).
37. Brazel, A. J., Cervený, R. S. & Trapido, B. L. Localized climatic responses during the 11 July 1991 eclipse: Phoenix AZ. *Clim. Change* **23**, 155–168. <https://doi.org/10.1007/BF01097335> (1993).
38. Chernogor, L. F. Effects of solar eclipses in the surface atmosphere. *Izvestiya Atmos. Ocean. Phys.* **44**, 432–447. <https://doi.org/10.1134/S000143380804004X> (2008).
39. Jayakrishnan, P., Babu, C. & P., S. Drastic variation in the surface boundary layer parameters over Cochin during the annular solar eclipse. Analysis using sonic anemometer data. *J. Atmos. Solar Terr. Phys.* **94**, 49–53. <https://doi.org/10.1016/j.jastp.2012.12.019> (2013).
40. Nambodiri, K. *et al.* Effects of annular solar eclipse of 15 January 2010 on meteorological parameters in the 0 to 65 km region over Thumba, India. *Meteorol. Z.* **20**, 635–647. <https://doi.org/10.1127/0941-2948/2011/0253> (2011).
41. Satyanarayana, M. *et al.* Solar eclipse induced changes in aerosol extinction profiles: A case study of 11 Aug 1999 solar eclipse using lidar at a tropical station Trivandrum, India. *Indian J. Radio Space Phys.* **31**, 82–87 (2002).
42. Anderson, R. & Keefer, D. Observation of the temperature and pressure changes during the 30 June 1973 solar eclipse. *J. Atmos. Sci.* **32**, 228–231 (1975).
43. Niranjana, K. & Thulasiraman, S. Aerosol optical depth and size distribution changes during the total solar eclipse of 24 October 1995. *Terr. Atmos. Ocean. Sci.* **9**, 255–262. [https://doi.org/10.3319/TAO.1998.9.2.255\(A\)](https://doi.org/10.3319/TAO.1998.9.2.255(A)) (1998).
44. Rabin, R. M. & Doviak, R. J. Meteorological and astronomical influences on radar reflectivity in the convective boundary layer. *J. Appl. Meteorol.* **28**, 1226–1235. [https://doi.org/10.1175/1520-0450\(1989\)0283c1226:MAAIOR3e2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)0283c1226:MAAIOR3e2.0.CO;2) (1989).
45. Manohar, G. K., Kandalgaonkar, S. S. & Kulkarni, M. K. Impact of a total solar eclipse on surface atmospheric electricity. *J. Geophys. Res.* **100**, 20805–20814. <https://doi.org/10.1029/95JD01295> (1995).
46. Peñaloza-Murillo, M. A. Optical response of the atmosphere during the Caribbean total solar eclipses of 26 February 1998 and of 3 February 1916 at Falcón State, Venezuela. *Earth Moon Planets* **91**, 125–159 (2002).
47. Kolarž, P., Šekarić, J., Marinković, B. & Filipović, D. Correlation between some of the meteorological parameters measured during the partial solar eclipse, 11 August 1999. *J. Atmos. Solar Terr. Phys.* **67**, 1357–1364. <https://doi.org/10.1016/j.jastp.2005.07.016> (2005).
48. Abbott, W. N. On certain radiometric effects during the partial solar eclipse of February 25, 1952. *Geofisica Pura e Applicata* **39**, 186–193. <https://doi.org/10.1007/BF02001144> (1958).
49. Muraleedharan, P., Nisha, P. & Mohankumar, K. Effect of January 15, 2010 annular solar eclipse on meteorological parameters over Goa, India. *J. Atmos. Solar Terr. Phys.* **73**, 1988–1998. <https://doi.org/10.1016/j.jastp.2011.06.003> (2012).
50. Bala Subrahmanyam, D. *et al.* On the atmospheric boundary layer characteristics over Thumba: A case study. *Pure Appl. Geophys.* **169**, 741–753. <https://doi.org/10.1007/s00024-011-0336-9> (2012).
51. KiranKumar, N., Purushothaman, N. & Santosh, M. Response of spectral characteristics of wind and temperature of atmospheric surface layer to the noontime annular solar eclipse. *J. Atmos. Solar Terr. Phys.* **97**, 91–98. <https://doi.org/10.1016/j.jastp.2013.02.017> (2013).
52. Tzani, C., Varotsos, C. & Viras, L. Impacts of the solar eclipse of 29 March 2006 on the surface ozone concentration, the solar ultraviolet radiation and the meteorological parameters at Athens, Greece. *Atmos. Chem. Phys.* **8**, 425–430. <https://doi.org/10.5194/acp-8-425-2008> (2008).
53. Kalafatoglu, E. C., Kaymaz, Z. & Tan, E. August 11, 1999 solar eclipse over Turkey and associated effects. In *2007 3rd International Conference on Recent Advances in Space Technologies*, 50–53. <https://doi.org/10.1109/RAST.2007.4284043> (IEEE, 2007).
54. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning*. Springer Series in Statistics (Springer, 2009).
55. Bishop, C. M. *Pattern Recognition and Machine Learning. Information Science and Statistics* (Springer, 2006).
56. Ferrel, W. *A Popular Treatise on the Winds: Comprising the General Motions of the Atmosphere, Monsoons, Cyclones, Tornadoes, Waterspouts, Hail-storms, etc.* (Wiley, 1889).
57. Chelton, D. Mesoscale eddy effects. *Nat. Geosci.* **6**, 594–595. <https://doi.org/10.1038/ngeo1906> (2013).
58. Shi, R. *et al.* Ship observations and numerical simulation of the marine atmospheric boundary layer over the spring oceanic front in the northwestern South China Sea: Observation and Simulation of MABL in SCS. *J. Geophys. Res. Atmos.* **122**, 3733–3753. <https://doi.org/10.1002/2016JD026071> (2017).
59. Shi, R., Guo, X., Wang, D., Zeng, L. & Chen, J. Seasonal variability in coastal fronts and its influence on sea surface wind in the Northern South China Sea. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **119**, 30–39. <https://doi.org/10.1016/j.dsr2.2013.12.018> (2015).
60. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
61. Espenak, F. & Meeus, J. Five Millennium Cannon of Solar Eclipses: 1999 to 3000 (2000 BCE to 3000 CE). NASA/TP-2006-214141 (2006).
62. JPL, N. NASA Shuttle Radar Topography Mission Global 1 arc second. <https://doi.org/10.5067/MEASURES/SRTM/SRTMGL1.003> (2013). Type: dataset.

Acknowledgements

SM was financed by ANID-CENTROS REGIONALES R20F0008. Fruitful discussions with fellow members of CEAZA are gratefully acknowledged during the early development of this study.

Author contributions

A.C.H. developed the scope of the paper, designed the methods, performed the analysis, and wrote the paper. S.M. contributed expertise to developing the scope and methods of the paper, and also commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.C.H. or S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022