The problem of scale in predicting biological responses to climate

Running Title: Scale problems in climate biology

Authors

Luca Bütikofer[a,b], lucabutikofer@gmail.com; Karen Anderson[a], Karen.Anderson@exeter.ac.uk; Daniel P. Bebber[c], D.Bebber@exeter.ac.uk; Jonathan J. Bennie[d], J.J.Bennie@exeter.ac.uk; Regan I. Early[c], regan.early@gmail.com; Ilya M. D. Maclean[a*], I.M.D.Maclean@exeter.ac.uk (+44 (0) 1326255968)

[a] Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, TR10 9FE, United Kingdom.

[b] Sustainable Agriculture Sciences, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom

[c] Department of Biosciences, University of Exeter, Exeter, EX4 4QJ, United Kingdom

[d] Centre for Geography and Environmental Science, University of Exeter, Penryn, Cornwall, TR10 9FE, United Kingdom.

[*] Corresponding author

Keywords: Climate change, resolution, model, microclimate, distribution, phenology

Abstract

Many analyses of biological responses to climate rely on gridded climate data derived from weather stations, which differ from the conditions experienced by organisms in at least two respects. First, the microclimate recorded by a weather station is often quite different to that near the ground surface, where many organisms live. Second, the temporal and spatial resolutions of gridded climate datasets derived from weather stations are often too coarse to capture the conditions experienced by organisms. Temporally and spatially coarse data have clear benefits in terms of reduced model size and complexity, but here we argue that coarse-grained data introduce errors that, in biological studies, are too often ignored. However, in contrast to common perception, these errors are not necessarily caused directly by a spatial mismatch between the size of organisms and the scale at which climate data are collected. Rather, errors and biases are primarily due to (i) systematic discrepancies between the climate used in analysis and that experienced by organisms under study and (ii) the non-linearity of most biological responses in combination with differences in climate variance between locations and time periods for which models are fitted and those for which projections are made. We discuss when exactly problems of scale can be expected to arise and highlight the potential to circumvent these by spatially and temporally down-scaling climate. We also suggest ways in which adjustments to deal with issues of scale could be made without the need to run high-resolution models over wide extents.

Introduction

Climate is among the most fundamental driving forces controlling the environment in which organisms reside (Clarke, 2017). It sets boundaries on the biological processes fundamental to their survival and reproduction, and governs the rates of processes within these boundaries. Though many ecological studies account for climate variables when explaining biological phenomena, they usually rely on data derived or modelled from weather stations, the spatial resolution of which is typically orders of magnitude larger than the organisms under study (Potter, Arthur Woods, & Pincebourde, 2013).

Conventionally, terrestrial meteorological data are collected from networks of weather stations, with variables such as temperature and humidity recorded at c.1.5-2 metres from the ground surface in locations carefully selected to be unaffected by local microclimatic influences. Observations are often subsequently interpolated to a grid at a resolution of 10^1 to 10^2 km (see for example World Meteorological Organization, 2010). Future climate predictions from regional-scale climate models and reanalyses of historical data are typically made available at a similar spatial scale. While meteorological data are frequently recorded at hourly or sub-hourly intervals, summarised data are usually in the form of daily, monthly or annual summary statistics. This standardised approach to data collection and collation is designed to capture large-scale atmospheric phenomena for the description and prediction of weather systems; the influence of very fine-scale and short-term variation is of less interest to meteorologists. Data in this form are used widely by ecologists and agronomists, not least as they offer simple and attractive means of modelling biological responses with comparative ease. For example, the WorldClim dataset (Fick & Hijmans, 2017), used very commonly in biological studies (Gardner, Maclean, & Gaston, 2019), models climate at a spatial resolution of 1 km and a temporal resolution of one month (estimated over multiple years). Nevertheless, a growing literature stresses that the microclimatic conditions that influence the growth, reproduction and survival of organisms in the environment can vary considerably from standardised meteorological data (Bramer et al., 2018; Potter et al., 2013; Suggitt et al., 2017).

Spatial and temporal variation in climate is greatest close to the ground and the surfaces of vegetation where most organisms live (Mihalakakou, Santamouris, Lewis, & Asimakopoulos, 1997). Close to the ground, or inside forests for example, most of the momentum of wind is absorbed and the air flow is thus much slower, preventing the thermal mixing that evens out temperatures at the height of weather stations. Consequently, there is much more spatial variation in ground temperature than is recorded at weather stations (Monin & Obukhov, 1954; Oke, 2002). For example, instantaneous temperatures measured a few centimetres apart just above ground (e.g. on the north and south facing sides of an anthill, or within shaded areas and underneath canopy gaps in a forest), are as variable as temperature differences over the extent of the UK measured using standard weather stations (Bramer et al., 2018).

Likewise, over just a few metres, surface water conditions can vary from permanently wet to permanently dry (Arsenault et al., 2019).

In many circumstances a biologist may seek to calculate the response of an organism to climatic variables and predict the response at times or locations with different climate. Such predictions can be made by projecting a model calibrated at a specific time and location using climate data for new times and locations. Models of this kind can be simple and correlative, for example the construction of a climate envelope encompassing the current distribution of an organism (Lembrechts et al., 2019), or the regression calculations establishing relationships between growth and accumulated temperature at different locations (McMaster & Wilhelm 1997). More complex models might include the process-based crop simulation models used in agriculture (e.g. Van Diepen et al., 1989), or mechanistic representations of plant growth in land surface models or dynamic vegetation models (e.g. Sitch et al., 2003). However, in so doing, several types of bias can arise if the resolution of climate data used is excessively coarse. Firstly, biases may result from the difference between the climate experienced by the organism and the climate data used in the model when this difference is not constant between calibration and prediction. Secondly, biases may result from the non-linearity of the biological response to climate. When climate information is spatially or temporally aggregated, a simple measure of central tendency is used to summarise the data across the aggregation (e.g. the mean temperature within a coarse-resolution grid cell). We show that, because a non-linear response to an averaged climate variable is different from the averaged response, predictions derived at one scale do not necessarily translate to those made at different scales. Moreover, even if the scale is maintained constant between calibration and prediction, when the distribution of a climate variable around its mean value varies between locations or over different time periods, biases may arise due to differences in the discrepancy between the mean response and the response to the averaged climate data.

Here we describe how, why and when the use of coarse-scale climate data is problematic. First, we explain exactly what the problems are. We then discuss the extent to which biological responses would be expected to be non-linear and hence affected by issues of averaging and scale. We then present

specific examples of when such errors arise, to indicate the potential magnitude of the problem. We conclude by demonstrating how the use of high-resolution climate data can avoid these problems, and how in the absence of such data, adjustments to deal with issues of non-linearity can be made.

Climate and the problem of scale

The simplest form of error arising from coarse resolution data is the discrepancy between standard meteorological measures of climate and the climatic conditions actually experienced by an organism. Such discrepancies occur whenever the organism is poorly coupled to the surrounding atmospheric air mass. This is the case for any organism living close to the ground, where radiative heating and cooling effects affect microclimate air temperatures; for organisms in environments where latent heat exchanges buffer against temperature change, including humid environments or those near bodies of water, snow or ice (Campbell & Norman 2012); or for organisms in deep shade under a forest canopy (De Frenne et al., 2019). It is also the case where an organism itself is influenced by radiative heating and cooling. Biases in the computation of biological responses derived from these climate data can thus occur even if the response is linear.

Additional biases occur if coarse-scale climate data are used to model non-linear biological responses. The translation from fine to coarse scales is usually a form of averaging. Spatially, variables measured at precise locations are assigned a value representative of a wider area (e.g. a pixel on a raster), while temporally, data for a specific time period are assigned values representative of longer time intervals. When considering a non-linear biological response to a particular climate variable, e.g. ($f_{(x)}: y = x^2$), the mean biological response is not the same as the response to the mean of the predictor, i.e. ($\bar{x}^2 \neq \bar{y}$) (Fig. 1b). Consequently, if the average of a predictor variable is used in place of unaggregated variables, a biased prediction would be expected. Intuitively one might expect that calibrating and predicting with climate data at the same level of aggregation (e.g. monthly data at 1 km resolution) would bypass this problem. However, it is often the case that the distribution of values around the mean may differ between locations or time periods such that $\sum_{i=1}^{n} |x_i - \bar{x}|/n$ is not identical. In consequence $|\bar{x}^2 - \bar{y}|$

will also differ between locations. This error impacts predictions made to areas or periods of novel climate whenever the distribution of a climate variable represented by an average differs between calibration and prediction data (Fig. 1b). Such differences are likely to be the norm rather than the exception. In time, the amplitude of diurnal fluctuations in temperature are lower in coastal regions and reduced by cloud cover (Dai, Trenberth, & Karl, 1999), the latter influenced by elevation. In space, terrain and vegetation cover exert strong influences on heterogeneity in temperatures (Lenoir et al. 2013; Suggitt et al., 2018) implying that the models calibrated in relatively flat un-forested regions, for example, are not translatable to mountainous regions and/or forested regions and vice versa.

A more specific, but very widespread problem occurs when coarse-resolution climate data are expressed in terms of accumulated "forcing units" or "growing-degrees" per unit time interval e.g. growing-degree days. In its basic formulation (McMaster & Wilhelm, 1997), the timing of phenological events are assumed to be directly related to the accumulation of forcing units, where a forcing unit is the length of time for which the average temperature is above a specified threshold (*T*0). However, the temporal resolution of the temperature data used to compute growing-degrees plays an important role. When compared to Growing Degree-Hours (GDH), Growing Degree-Days (GDD) tend to underestimate the time at which the study organism is exposed to temperatures greater than the threshold (Gu, 2016). This is caused by the daily fluctuation of hourly temperatures around the mean. The difference between GDD and GDH is greater when the mean daily temperature is close to *T*0. When the daily mean is just below *T0* no GDDs are counted, yet the warmest hours of the day will often be above *T0* causing the accumulation of some GDH. Conversely, when the daily mean is above *T0*, GDD is assumed to accumulate over the entire day, yet for several hours in the day the temperature is below *T*0. Although this phenomenon has been noticed in the past (Baker, 1980; Merrill & Peairs, 2017; Worner, 1992), it is surprisingly commonly ignored (Chuine, Cambon, & Comtois, 2000; Chung, Mack, Yun, & Kim, 2011; Shi et al., 2017).

Another specific example relates to models that seek to determine the relationship between the occurrence of species and climate in space and time. The premise of species distribution models is that

the coarse spatial and temporal resolution climate variables used in these models are statistically meaningful predictors of probability of species occurrence (Bennie, Wilson, Maclean, & Suggitt, 2014). Thus, while the variables included in these models are not necessarily assumed to affect thermal performance directly, they are assumed to correlate with performance because the closer the mean climate is to the thermal optima of a species, the greater the prevalence of favourable climatic conditions in space and time. However, the discrepancy between the true mean thermal performance and the assumed mean estimated from aggregated temperature data will vary as a function of the distribution of temperature around the mean. Thus, while it is often assumed that projections derived from these models may be biased because of the mismatch between the size of organisms and the scale at which climate data are collected and modelled (e.g. Potter et al., 2013), this is not necessarily the case. Rather, it is the non-linear relationship between occurrence probability and climate and the likelihood that spatio-temporal variability in climate is not constant in time and space that results in the bias.

These discrepancies raise three important issues. First, models calibrated with field measurements of climate experienced by organisms cannot be applied using temperatures derived from weather stations without introducing significant biases into the model's predictions. Second, models calibrated using climatic data of one spatial or temporal resolution should not be used to derive predictions using climate data of another resolution without careful consideration (and ideally testing) of the potential to introduce bias under a given climate. Last, even if resolution is maintained constant between calibration and prediction, when applying models across regions with different climates, and possibly even between years at sites with inter-annual variation, significant biases may arise if coarse-resolution data are used and the variance around the mean is not constant.

Non-linear biological responses

Many biological processes are inherently non-linear (Archontoulis & Miguez, 2015). At the most fundamental level, the temperature dependence of the chemical reaction rates (the speed at which reactants turn into products) is described by the Arrhenius equation, which takes the form of an

exponential function. At higher levels, many biological responses are also non-linear. In plants, for example, the relationship between incident, Photosynthetically Active Radiation (PAR) flux and CO_2 intake per leaf area per time unit is a positive, linear function at low PAR values but eventually reaches an asymptote. Similarly, the internal net photosynthetic rate varies non-linearly with irradiance, showing saturation at high levels of irradiance for varying levels of the quantum efficiency of photosynthesis (Reed, Hamerly, Dinger, & Jarvis, 1976). This saturation occurs because, under moderate flux densities, the photosynthetic apparatuses are capable of processing all of the incoming radiation. Light saturation values are typically much below flux densities under clear-sky conditions, placing fully exposed plants in the non-linear portion of the curve relatively often. In cases of excessive exposure, PAR can damage the photosynthetic apparatuses, reducing CO_2 fixation. Similarly, growth rates also respond non-linearly to temperature, following a logistic function with exponential growth at the low end of the temperature range, a linear section in the middle, and a logarithmic-type gradual decrease of the growth rate at the high end of the range (Went, 1953). In consequence, models of plant growth calibrated for one location or time period do not translate to others. Even in instances where idealised linear biological responses are expected, non-linearity may result from Blackman's "law of limiting factors" (Blackman, 1905). Most biological processes are limited by more than one external factor. While relationships between growth and photosynthesis may be linear at low light levels, for example, when light becomes abundant, CO_2 becomes limiting and so the biological response becomes non-linear. This idea of multiple limiting factors is often invoked as an explanation of why idealised physical relationships are sometimes linear, but real biological relationships almost never are.

Animals too exhibit complex non-linear responses to climatic variables. At a fundamental level, the thermal energy emitted by an organism increases as a function of its temperature in Kelvin to the power of 4 and the latent heat release increases exponentially with temperature (Campbell, 1977; Kearney & Porter, 2020; Tetens, 1930). Sensible heat loss in the form of conduction and convection increases with the temperature difference between the body and the air. The body temperature of endotherms thus typically increases asymptotically with air temperature and is maintained within a narrow thermal range. For ectotherms, the metabolic rate will typically decrease with temperature until basal levels are

198 reached, but its water loss will increase exponentially (Porter & Gates, 1969). Since there are often
199 limits to the energy and water intake an organism is able to attain, the thermal performance functions
200 of organisms are usually highly non-linear, characterised by Gaussian, Beta or Wiebull functions
201 (Angilletta, 2006). Thermal performance is thus high within a definable range of ambient temperatures,
202 but declines sharply when these thresholds are exceeded. Spatially or temporally aggregated data do
203 not capture these climatic extremes and would thus be expected to over-estimate thermal performance
204 and survival (Sunday et al., 2014).

205 In addition to these passive dependencies on climate, plants and animals have also evolved more active
206 strategies to compensate for highly variable, and sometimes sub-optimal environmental conditions. In
207 plants, environmental variability impacts mainly the photosynthetic apparatus, and plants have thus
208 evolved many methods of responding to changes in their growing conditions (Walters, 2005). These
209 can manifest as long-term developmental shifts or adjustments in proteins within the photosynthetic
210 apparatus, over timescales of seconds to hours (Demmig-Adams et al., 1996). To prevent thermal
211 damage, for example, plants cool down through evapotranspiration and sensible heat loss. While
212 partially controlled passively, this occurs at the stomatal level, and through biochemical processes that
213 store heat energy into the chemical bonds of molecules (such as Isoprene) that are then released into the
214 air during hot days. Thermal acclimation is also important, and thought to affect strongly coupled
215 vegetation-atmosphere feedbacks in the global carbon cycle, especially as the climate warms (Stinziano,
216 Way, & Bauerle, 2018).

217 Animals too exhibit active strategies for maintaining body temperature. The most prevalent example of
218 this is behavioural thermoregulation. Most terrestrial ectotherms are mobile and can behaviourally
219 exploit local heterogeneity in climate to regulate their body temperatures somewhat independently of
220 local environmental temperatures — the so called “Bogert effect” (Bogert, 1949). Though the
221 physiological thermal-tolerance limits of most terrestrial ectotherms usually exceed local air
222 temperatures, their extreme operative body temperatures in exposed habitats often match or exceed
223 these thermal-tolerance limits (Sunday et al., 2014). Therefore, most ectotherms do not have a

physiological thermal-safety margin and must rely on behaviour to avoid overheating or to avoid lethal cold exposure (Sunday et al., 2014). In consequence, their biological responses are unlikely to change linearly with ambient conditions.

Thus, fundamental mechanisms driving chemical reactions, the exchange of heat between organisms and their environment, and the growth, development and survival of organisms vary non-linearly with respect to temperature and other climate variables. It is therefore better to assume non-linearity whenever there is no evidence to the contrary, and many of the issues raised in this paper are likely to be quite universal.

Applied examples

To illustrate the potential magnitude of errors associated with non-linear biological responses and scale, we provide two examples. In the first example GDD and GDH were calculated at multiple heights above ground, and at various spatial resolutions using the microclimate model of Maclean et al., (2019). The model was applied to derive temperatures at a grid resolution of 1 m over a 200 m by 200 m region of the Lizard Peninsula in Cornwall, UK (49.97°N, 5.22°W). To test the importance of "height above ground", GDH (base 10°C) for the period 1st Jan to 20th April 2017 were calculated from temperatures at hourly intervals at heights of 2, 5, 10, 25, 50 and 100 cm from the ground. It can be seen that the rate of increase in GDH, here for a flat surface in the centre of the study location, is much faster for temperatures near the ground (Fig. 2b). For example, temperatures at 2 cm above ground reach a GDH threshold value of 1500 on 2nd April, in contrast to temperatures 100 cm above ground, which reached the same threshold on 28th April, almost a month later. Though night-time temperatures are significantly colder near the ground (Maclean et al., 2019), this is more than compensated for by warmer daytime temperatures. Thus, the timing of phenological events for organisms living close to the ground could potentially be underestimated significantly if ambient air temperatures are used, and likewise models fitted using GDH/24 cannot be applied with daily data and vice-versa. To test the importance of the time-interval used, we compared estimates of GDH/24 with those of GDD for the same location and

period (5 cm above ground on a south-facing slope, Fig. 2c, and across the entire study region, Fig. 2i). The discrepancy was marked. The GDD estimate for the 30th of April was less than half the estimate derived by computing GDH/24. To test the effects of spatial resolution, we computed GDH/24 at grid resolutions of 5 m and 25 m using two approaches. In the first, the input climate data were coarsened, whereas in the second, we instead coarsened the cumulative degree-hour estimates (Fig. 2e-h). While at 5 m resolution only minor discrepancies were evident, at 25 m grid resolution the discrepancies were marked. When the input temperature data were averaged, spatial variation in GDH/24 was generally lower, and locations with low and high values of GDH/24 do not necessary correspond. Cleary scale is important in the estimation of GDH, and both the locations and timings of phenological events may be misrepresented when coarse spatial or temporal data are used.

In the second example, we used a slightly simplified version of the microclimate and general ectotherm models of Kearney & Porter (2017, 2020) to estimate the operative body temperature, water loss and activity budget of a the great desert skink *Liopholis kintorei* at a location in Northern Territory, Australia (23.71°S, 129.93°E) using hourly and daily climate forcing data to run the model as described in Kearney et al., (2019). The conventional model includes a suite of programs for the mechanistic modelling of heat, water, energy and mass exchange between an organism and its environment over its entire life cycle, which in turn, based on body temperature and energy and water demands, can be used to predict behaviour. In our simplified version of the model, it was assumed that the skink would bask if body temperatures do not exceed an upper thermal tolerance threshold of 45°C irrespective of energy requirements, but would retreat to burrows to seek shade if the body temperature exceeded this temperature. It was also assumed that a skink would retreat to burrows if the body temperature potentially attained in a burrow exceeded that which would be attained when basking in the open if below this upper thermal threshold, such as would be expected to occur at night (Fig. 3b). When estimated using hourly climate data, both mean daily body temperature (Fig. 3c) and water loss (Fig. 3f) were generally higher than when estimated using daily climate data. There were also marked differences is in the prediction of behaviour (Fig. 3d). Whereas the daily model predicted that skinks would spend almost all of their time basking as the upper critical threshold was not reached, and average

daily body temperature over 24 hours in open areas was higher than that which would have been attained in burrows, the hourly model predicted that skinks would spend their time basking during daylight hours only, except during the hottest periods of the day. Thus, even minor biases in the estimation of body temperatures, caused by non-linearity and temporal averaging can have a marked outcome on a predicted behavioural response.

Obtaining high-resolution climate data

Clearly, many of the issues of scale and non-linearity can be resolved through the use of higher resolution climate data, but in practical terms such data are not always readily available. Nevertheless, the issue of lack of high temporal resolution data is relatively easy to address. Sub-daily modelled estimates of historic climate have recently become available at ~30 km grid resolution through the ERA5 Atmospheric Reanalysis Project (Albergel et al., 2018). While it is inherently impossible to predict the precise climate conditions at some date and time in the distant future, reliable methods for generating synthetic time series of sub-daily or daily weather, using weather generators, are also increasingly available (e.g. Ailliot, Allard, Monbet, & Naveau, 2015). Interpolating these data to high temporal resolution is also comparatively straightforward. Simple approaches that replicate diurnal temperature cycles by fitting two terms of a Fourier series have been widely used for decades (e.g. Campbell, 1977). More complex approaches entail modelling the departure from these idealised diurnal cycles by using proxy data from alternative sources such as nearby weather stations (Luedeling 2018) or estimates of cloud cover and solar radiation (Maclean et al., 2019), but can also be applied easily. In so doing, it is also worth remembering that certain processes, such as photosynthesis, occur only during daylight hours. It is therefore important to use climate measurements that are time-restricted to the relevant periods.

The issue of spatial resolution is more problematic than temporal resolution, though a paradigm shift in the ability of the scientific community to address this issue is occurring (Lembrechts & Lenoir, 2019). Global efforts to obtain measurements of high-resolution soil temperatures are already underway

(Lembrechts et al., 2020b) and at its simplest, coarse spatial resolution data can be downscaled using spatial interpolation techniques (e.g. Wahba, 1990) or multivariate regression (e.g. Greiser, Meineri, Luoto, Ehrlén, & Hylander, 2018). Such approaches are relatively effective at capturing mesoclimatic variation, but suffer from some of the same issues associating with non-linearity, in that the environmental determinants of differences between coarse- and fine-resolution climates may not be constant in time and space. For this reason, there has been a concerted effort to develop more mechanistic approaches. These approaches, which build on the pioneering applications of physics to biology (Monin & Obukhov, 1954; Monteith, 1973; Penman, 1948), now permit both historic and future microclimate conditions to be computed anywhere on earth using freely available climate and environmental data (Kearney et al., 2019; Kearney & Porter, 2017; Maclean, 2019).

It is also worth reemphasising that a key source of the discrepancy between the conditions experienced by organisms, the temperature of the organism itself, and that of a weather station is the height above the ground at which the organism lives. Both spatial and temporal heterogeneity in temperature, and deviations from measurements made by weather stations, are most pronounced immediately above the ground. For this reason, it is important to consider the height at which temperature is measured or modelled relative to that of the organism under study. While microclimate models permit users to specify the height at which temperature is required, they are most suited to modelling conditions immediately above or below a vegetated surface, but not within a canopy itself (Bramer et al., 2018). Furthermore, in the context of within-canopy temperatures, it is worth considering the dynamic feedbacks between climate and canopy cover. Understory microclimate is influenced strongly by vegetative shading, yet the degree of shading itself varies throughout the year, partly in response to changing climatic conditions. This is of particular importance during spring and autumn, when leaf flushing, colouration and abscission change most rapidly, altering radiation transmission though the canopy and consequently understory microclimate (Villegas, Breshears, Zou, & Royer, 2010).

At what resolution are data needed?

A key question then is at what spatial resolution are climate data needed in order to avoid erroneous predictions of biological responses to climate? Potter et al. (2013) show that grid lengths in species distribution models are, on average, ~10^4-fold larger than the animals they study. Though many organisms are mobile, their temperatures are determined by heat fluxes operating on their body, averaged over time periods that scale proportionally to their thermal mass (Porter, Mitchell, Beckman, & DeWitt, 1973). However, the relationships between body temperature and air temperature are non-linear (Porter & Gates, 1969), so it is not the case that body temperatures scale simply with the average of the air temperature over the region that they roam. Moreover, many mobile organisms exhibit thermoregulatory behaviour such as basking, and therefore show strong preferences for particular microclimates within the landscape (Barton, Porter, & Kearney, 2014). At face value, the implication that there is a need to model temperatures at spatial resolutions that match the body size of organisms is worrying, as spatially explicit and accurate representation of global- or continental-extent climate at a resolution of a few centimetres to metres is impractical, even with rapid advances in computer processing power and fine spatial resolution remote sensing data, particularly if fine temporal-resolution data are also needed. However, we argue that explicit knowledge of climatic conditions at resolutions that match the body size of organisms are not necessarily needed. Instead we suggest that knowledge of the likely spatial and temporal distribution of climatic variables around the mean is more important. This in turn allows simulation of the range of conditions experienced by organisms (cf. Lembrechts et al., 2020a), which by using principles of biophysical ecology, provide direct mechanistic insight into the physiological responses and constraints and hence of thermal performance (Kearney & Porter 2009).

While it is commonly perceived that climate exerts influence on species primarily at coarser scales, and that fine-scale factors such resource availability and biotic interactions are more important (Pearson and Dawson 2003), it remains the case that organisms are most directly influenced by the climatic conditions they experience. Associations with climate at coarser scales results primarily because such data serve as proxies for the spatial and temporal variations in the microclimate that influence individual performance (Bennie et al., 2014; Gardner et al 2019). However, organisms are most directly connected to climatic conditions through exchanges of energy and mass (Porter & Gates 1969). With estimates of

the mean and range of conditions directly experienced by organisms it is possible to use principles of thermodynamics to derive mechanistic models of these processes and their physiological consequences (Kearney & Porter, 2009). Moreover, sophisticated models are now emerging to infer biotic interactions from species distribution data, but spatial scale remains one of the major challenges as biotic interactions almost invariably occur at finer spatial resolutions than those for which we have climate data (Araújo & Rozenfeld 2014). A potential solution to problems of scale is thus judicious sub-sampling. Here, instead of attempting to model climate at fine spatial and temporal resolution over wide regions, fine resolution climate data are derived at sample locations and time-periods that best represent how organisms use their environment. Such data could then either be used to simulate the direct physiological responses mechanistically, or used in place of conventional climate data when using a statistical approach by Monte Carlo simulation.

Conclusion

Many biological phenomena are studied using coarse spatial and temporal resolution climate data, but doing so introduces errors for at least two reasons. Firstly, because there may be systematic differences between the climate experienced by organisms and that measured by weather stations, and, secondly, because many responses to climate are non-linear, and the mean biological response is not the same as the response to the mean climate. Such errors are likely to be particularly pronounced when models are calibrated and projected in very different environments, such as calibrated in a lab and then applied in the field, but may manifest in any situation in which a model is projected to new time periods or locations. Most biological responses are inherently non-linear, and in the absence of evidence to the contrary it is thus safer to assume non-linearity. The problem of scale is likely to be much more ubiquitous than is commonly appreciated. We thus urge biologists to give greater consideration to this issue. Methods for downscaling climate to finer spatial and temporal resolution are now readily available and provide the tools by which to do so.

Acknowledgements

The work underpinning this paper was supported through the Agri-tech Cornwall & the Isles of Scilly Project, part-funded by the European Regional Development Fund, Cornwall Council and the Council for the Isles of Scilly (award number: 05R16P00366). We thank Jonathan Lenoir and Jonas Lembrechts for helpful comments and Katie Thurston for proof reading the manuscript.

Author contributions

LB and IMDM wrote the manuscript, with contributions from KA, JJB, RIE and DPB. IMDM performed the analyses underpinning figures. All authors jointly conceived the ideas presented.

Data availability

The digital elevation and climate forcing data used to generate Figure 2 are including with the R package microclima (Maclean et al., 2019). The climate forcing data used for generating Figure 3 were sourced using the climate data download tools added to version 1.1.2 of R package microclima (Keanrey et al., 2019).

References

Ailliot, P., Allard, D., Monbet, V., & Naveau, P. (2015). Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistique, 156*(1), 101-113.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., Rosnay, P. d., & Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better? *Hydrology and Earth System Sciences, 22*(6), 3515-3532. doi:10.5194/hess-22-3515-2018.

Angilletta Jr, M. J. (2006). Estimating and comparing thermal performance curves. *Journal of Thermal Biology, 31*(7), 541-545. doi:10.1016/j.jtherbio.2006.06.002.

Araújo, M. B. and A. Rozenfeld (2014). The geographic scaling of biotic interactions. *Ecography, 37*(5), 406-415. doi:10.1111/j.1600-0587.2013.00643.x.

Archontoulis, S. V., & Miguez, F. E. (2015). Nonlinear regression models and applications in agricultural research. *Agronomy Journal, 107*(2), 786-798. doi:10.2134/agronj2012.0506.

Arsenault, J., Talbot, J., Moore, T. R., Beauvais, M.-P., Franssen, J., & Roulet, N. T. (2019). The spatial heterogeneity of vegetation, hydrology and water chemistry in a peatland with open-water pools. *Ecosystems, 22*(6), 1352-1367. doi:10.1007/s10021-019-00342-4.

Baker, C. (1980). Some Problems in Using Meteorological Data to Forecast the Timing of Insect Life Cycles. *EPPO Bulletin, 10*(2), 83-91. doi:10.1111/j.1365-2338.1980.tb02628.x.

Barton, M., Porter, W., & Kearney, M. (2014). Behavioural thermoregulation and the relative roles of convection and radiation in a basking butterfly. *Journal of Thermal Biology, 41*, 65-71. doi:10.1016/j.jtherbio.2014.02.004.

Bennie, J., Wilson, R. J., Maclean, I., & Suggitt, A. (2014). Seeing the woods for the trees-when is microclimate important in species distribution models? *Global Change Biology, 20*(9), 2699-2700. doi:10.1111/gcb.12525.

Blackman, F. F. (1905). Optima and Limiting Factors. *Annals of Botany, 19*(74), 281–295.

Bogert, C. M. (1949). Thermoregulation in reptiles, a factor in evolution. *Evolution, 3*(3), 195-211. doi:10.2307/2405558.

Bramer, I., Anderson, B. J., Bennie, J., Bladon, A. J., De Frenne, P., Hemming, D., . . . Korstjens, A. H. (2018). Advances in monitoring and modelling climate at ecologically relevant scales. In D.A. Bohan, A. J. Dumbrell, G. Woodward, M. Jackson (Eds.), *Advances in ecological research* (Vol. 58, pp. 101-161): Elsevier. doi:10.1016/bs.aecr.2017.12.005.

Campbell, G. S., & Norman, J. M. (2012). *An introduction to environmental biophysics (2nd ed).* New York: Springer.

Chuine, I., Cambon, G., & Comtois, P. (2000). Scaling phenology from the local to the regional level: advances from species-specific phenological models. *Global Change Biology, 6*(8), 943-952. doi: 10.1046/j.1365-2486.2000.00368.x.

Chung, U., Mack, L., Yun, J. I., & Kim, S.-H. (2011). Predicting the timing of cherry blossoms in Washington, DC and mid-Atlantic states in response to climate change. *PLoS One, 6*(11), e27439. doi: 10.1371/journal.pone.0027439.

Clarke, A. (2017). Principles of Thermal Ecology: Temperature, Energy, and Life. : Oxford University Press.

Dai, A., Trenberth, K. E., & Karl, T. R. (1999). Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range. *Journal of Climate, 12*(8), 2451-2473. doi: 10.1175/1520-0442.

De Frenne, P., Zellweger, F., Rodriguez-Sanchez, F., Scheffers, B. R., Hylander, K., Luoto, M., Velend, M., Verheyen K. & Lenoir, J. (2019). Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution, 3*(5), 744-749. doi: 10.1038/s41559-019-0842-1.

Demmig-Adams, B., Adams III, W. W., Barker, D. H., Logan, B. A., Bowling, D. R., & Verhoeven, A. S. (1996). Using chlorophyll fluorescence to assess the fraction of absorbed light allocated to

thermal dissipation of excess excitation. *Physiologia Plantarum, 98*(2), 253-264. doi: 10.1034/j.1399-3054.1996.980206.x.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology, 37*(12), 4302-4315. doi: 10.1002/joc.5086.

Gardner, A. S., Maclean, I. M., & Gaston, K. J. (2019). Climatic predictors of species distributions neglect biophysiologically meaningful variables. *Diversity and Distributions, 25*(8), 1318-1333. doi: 10.1111/ddi.12939.

Greiser, C., Meineri, E., Luoto, M., Ehrlén, J., & Hylander, K. (2018). Monthly microclimate models in a managed boreal forest landscape. *Agricultural and Forest Meteorology, 250,* 147-158. doi: 10.1016/j.agrformet.2017.12.252.

Gu, S. (2016) Growing degree hours - a simple, accurate, and precise protocol to approximate growing heat summation for grapevines. *International Journal of Biometeorology*, *60,* 1123-1134. doi: 10.1007/s00484-015-1105-8.

Kearney, M. R., Gillingham, P. K., Bramer, I., Duffy, J. P., & Maclean, I. M. (2019). A method for computing hourly, historical, terrain-corrected microclimate anywhere on Earth. *Methods in Ecology and Evolution*. *11*(1), 38-43. doi: 10.1111/2041-210X.13330.

Kearney, M. & Porter, W. (2009), Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters, 12*(4), 334-350. doi:10.1111/j.1461-0248.2008.01277.x

Kearney, M.R. & Porter, W.P. (2017). NicheMapR – an R package for biophysical modelling: the microclimate model. *Ecography*, *40*(5), 664-674. doi:10.1111/ecog.02360.

Kearney, M. R., & Porter, W. P. (2020). NicheMapR–an R package for biophysical modelling: the ectotherm and Dynamic Energy Budget models. *Ecography, 43*(1), 85-96. doi: doi.org/10.1111/ecog.04680.

Lembrechts, J. J., Broeders, L., De Gruyter, J., Radujković, D., Ramirez-Rojas, I., Lenoir, J., & Verbruggen, E. (2020a). A framework to bridge scales in distribution modeling of soil microbiota. *FEMS Microbiology Ecology, 96*(5), fiaa051. doi: 10.1093/femsec/fiaa051.

Lembrechts, J. J., & Lenoir, J. (2019). Microclimatic conditions anywhere at any time! *Global Change Biology*, *26*(2), 337-339. doi: 10.1111/gcb.14942.

Lembrechts, J. J., Nijs, I., & Lenoir, J. (2019). Incorporating microclimate into species distribution models. *Ecography*, *42*(7), 1267-1279. doi: 10.1111/ecog.03947.

Lembrechts, J.J., Aalto, J., Ashcroft, M.B., De Frenne, P., Kopecký, M., Lenoir, J., …, Nijs, I. (2020b). SoilTemp: A global database of near-surface temperature. *Global Change Biology*, in press, doi: 10.1111/gcb.15123.

Lenoir, J., Graae, B.J., Aarrestad, P.A., Alsos, I.G., Armbruster, W.S., Austrheim, G., …, Svenning, J.-C. (2013) Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across Northern Europe. *Global Change Biology, 19*(5): 1470-1481. doi:10.1111/gcb.12129

Luedeling, E. (2018). Interpolating hourly temperatures for computing agroclimatic metrics. *International Journal of Biometeorology, 62*(10), 1799–1807. doi:10.1007/s00484-018-1582-7.

Maclean, I. M. D. (2019). Predicting future climate at high spatial and temporal resolution. *Global Change Biology, 26*(2), 1003-1011. doi: 10.1111/gcb.14876.

Maclean, I. M.D., Mosedale, J. R., & Bennie, J. J. (2019). Microclima: An r package for modelling meso-and microclimate. *Methods in Ecology and Evolution, 10*(2), 280-290. doi: 10.1111/2041-210X.13093.

Mcmaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology, 87*(4), 291–300. doi:10.1016/S0168-1923.

Merrill, S. C., & Peairs, F. B. (2017). Temperature variability is a key component in accurately forecasting the effects of climate change on pest phenology. *Pest Management Science, 73*(2), 380-388. doi: 10.1002/ps.4320.

Mihalakakou, G., Santamouris, M., Lewis, J., & Asimakopoulos, D. (1997). On the application of the energy balance equation to predict ground temperature profiles. *Solar Energy, 60*(3-4), 181-190. doi: 10.1016/S0038-092X(97)00012-1.

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR, 24*(151), 163-187.

Monteith, J. (1973). *Principles of environmental physics.* London : Edward Arnold.

Oke, T. (2002). Boundary Layer Climates. London: Routledge, doi:10.4324/9780203407219.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?. *Global Ecology and Biogeography*, *12*(5), 361-371. doi:10.1046/j.1466-822X.2003.00042.x

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 193*(1032), 120-145. doi: doi.org/10.1098/rspa.1948.0037.

Porter, W., Mitchell, J., Beckman, W., & DeWitt, C. (1973). Behavioral implications of mechanistic ecology. *Oecologia, 13*(1), 1-54. doi: 10.1007/BF00379617.

Porter, W. P., & Gates, D. M. (1969). Thermodynamic equilibria of animals with environment. *Ecological Monographs, 39*(3), 227-244. doi: 10.2307/1948545.

Potter, K. A., Arthur Woods, H., & Pincebourde, S. (2013). Microclimatic challenges in global change biology. *Global Change Biology, 19*(10), 2932-2939. doi: 10.1111/gcb.12257.

Reed, K., Hamerly, E., Dinger, B., & Jarvis, P. (1976). An analytical model for field measurement of photosynthesis. *Journal of Applied Ecology*, 13(3), 925-942. doi: 10.2307/2402267.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T. & Thonicke, K. (2003) Evaluation of ecosystem dynamics, plant

515 geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global*
516 *Change Biology*, *9*(2), 161-185. doi: 10.1046/j.1365-2486.2003.00569.x.

517 Stinziano, J. R., Way, D. A., & Bauerle, W. L. (2018). Improving models of photosynthetic thermal
518 acclimation: Which parameters are most important and how many should be modified? *Global*
519 *Change Biology*, *24*(4), 1580-1598. doi: 10.1111/gcb.13924.

520 Suggitt, A. J., Platts, P. J., Barata, I. M., Bennie, J. J., Burgess, M. D., Bystriakova, N., . . . Harper, A.
521 B. (2017). Conducting robust ecological analyses with climate data. *Oikos*, *126*(11), 1533-
522 1541. doi: 10.1111/oik.04203.

523 Suggitt, A. J., Wilson, R. J., Isaac, N. J., Beale, C. M., Auffret, A. G., August, T., . . . Fox, R. (2018).
524 Extinction risk from climate change is reduced by microclimatic buffering. *Nature Climate*
525 *Change*, *8*(8), 713-717. doi: 10.1038/s41558-018-0231-9.

526 Sunday, J. M., Bates, A. E., Kearney, M. R., Colwell, R. K., Dulvy, N. K., Longino, J. T., & Huey, R.
527 B. (2014). Thermal-safety margins and the necessity of thermoregulatory behavior across
528 latitude and elevation. *Proceedings of the National Academy of Sciences*, *111*(15), 5610-5615.
529 doi: 10.1073/pnas.1316145111.

530 Tetens, O. (1930). Uber einige meteorologische Begriffe. *Zeitschrift fur Geophysik, 6*, 297-309.

531 Van Diepen, C.A, Wolf, J., Van Keulen, H., & Rappoldt, C. (1989) WOFOST: a simulation model of
532 crop production. *Soil Use and Management* *5*(1), 16-24. doi: 10.1111/j.1475-
533 2743.1989.tb00755.x.

534 Villegas, J. C., Breshears, D. D., Zou, C. B., & Royer, P. D. (2010). Seasonally pulsed heterogeneity in
535 microclimate: phenology and cover effects along deciduous grassland–forest continuum.
536 *Vadose Zone Journal*, *9*(3), 537-547. doi: 10.1111/j.1475-2743.1989.tb00755.x

537 Wahba, G. (1990). *Spline models for observational data*. Philadelphia: Society for Industrial and
538 Appied Mathematics. doi:pdf/10.1137/1.9781611970128.bm.

539 Walters, R. G. (2005). Towards an understanding of photosynthetic acclimation. *Journal of*
540 *Experimental Botany*, *56*(411), 435-447. doi: 0.1093/jxb/eri060.

541 Went, F. (1953). The effect of temperature on plant growth. *Annual Review of Plant Physiology*, *4*(1),
542 347-362. doi: 10.1146/annurev.pp.04.060153.002023.

543 World Meteorological Organization (2010). *Manual on the Global Observing System*. Retrieved from
544 https://www.wmo.int/pages/prog/www/OSY/Manuals_GOS.html.

545 Worner, S. P. (1992). Performance of phenological models under variable temperature regimes:
546 consequences of the Kaufmann or rate summation effect. *Environmental Entomology*, *21*(4),
547 689-699. doi: 10.1093/ee/21.4.689.

548

549

Figure legends

Fig.1. Effects of averaging on non-linear data. In (a) a hypothetical linear biological response given by 0.5·Temperature + 5 is shown. Here the mean response and response to mean temperature are identical. In (b) a hypothetical non-linear biological response to temperatures in the range 0-10, given by response = temperature^2 is shown. Here, the mean response (solid horizontal line) is not the same as mean of temperature^2 (dashed horizontal line). In (c) a hypothetical biological response given by 0.1 x temperature^2 is shown for two temperature datasets with different means and distributions, but identical sample sizes. The difference between the mean response (solid horizontal lines) and 0.1 x the mean of temperature^2 (dashed horizontal lines) differs between the two datasets, demonstrating that when models are fitted using aggregated data and then projected to new locations or different time periods, errors will result unless the distribution of data around the mean remains identical.

Fig. 2. Effects of height above ground and resolution on the derivation of growing-degree days (GDD) and hours (GDH) on the Lizard Peninsula in the south-west of the United Kingdom (a). The microclimate model of Maclean et al. (2019) was used to derive temperatures at multiple heights in April 2017 for a 200 m by 200 m location in Cornwall, UK (49.97°N, 5.22°W). In (b) GDH/24 (base 10°C) was calculated for temperatures at various heights above ground. In (c) comparisons between GDH/24 and GDD are shown as a function of time (south-facing slope, 5 cm above ground). In (d) spatial variability in GDH/24, modelled at 1 m grid resolution (5 cm above ground) is shown. In (e-h) the effects of spatial coarsening are shown. GDH/24 at 5 cm above ground was derived at 5 m (e, f) and 25 m (g, h) resolution using two methods: first by coarsening the input temperature data (e, g) and second by coarsening the output growing-degree estimates (f, h). In (i) spatial differences in GDH/24 and GDD on 30th April (5 cm above ground) are shown. The colour scale is the same for figures d-h, as depicted by the colour bar to the right of (f).

Fig. 3. Body temperature, water loss and activity budget of the great desert skink *Liopholis kintorei* in Nov-Dec 2019 calculated using a simplified version of the general ectotherm model of Kearney et al. (2020) for a location in central Australia (a, 23.71°S, 129.93°E). In (b) temperatures were modelled at hourly intervals and it was assumed the skink will bask if body temperatures (green) did not exceed an upper thermal tolerance threshold of 45°C, and body temperature was calculated as the operative body temperature (grey). If temperatures exceed this threshold, or the body temperature that would be attained in burrows was higher than would be attained by basking, it was assumed that the skink underwent thermoregulatory behaviour and sought refuge in burrows and the body temperature equilibrates with the temperature of the burrow (purple). In (c) mean daily body temperature is shown derived by averaging the inputs (purple) and outputs (green) demonstrating that body temperatures were typically estimated to be warmer when outputs were averaged. In (d) the cumulative basking time is shown, indicating that when inputs were averaged, the predicted humid operative temperature was usually warmer than burrow temperatures, but colder than the upper critical threshold of 45°C and was therefore predicted to bask over the entire 24 period, whereas hourly data predicted basking behaviour only for part of the day. In (e) hourly water loss with (blue) and without (grey) thermoregulatory behaviour are shown, and in (f) daily water loss calculated by averaging the inputs (purple) and outputs (green) are shown. In (c) and (f) thermoregulatory behaviour was assumed.