
Article:
Kelly, LA and Hassall, C orcid.org/0000-0002-3510-0728 (2018) The spatial ecology of phytoplankton blooms in UK canals. Inland Waters, 8 (4). pp. 422-433. ISSN 2044-2041

https://doi.org/10.1080/20442041.2018.1482152

The spatial ecology of phytoplankton blooms in UK canals

Leah A. Kelly[1]*, Christopher Hassall[1]

[1] School of Biology, Faculty of Biological Sciences, Woodhouse Lane, University of Leeds, Leeds, LS2 9JT

* Present address: Department of Animal and Plant Science, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield, S10 2TN.

Author for correspondence:

Christopher Hassall

Tel: +44 113 3435578

Email: c.hassall@leeds.ac.uk

ABSTRACT

Environmental change is expected to increase the frequency and severity of problems caused by harmful algal blooms. We investigated the ecology of phytoplankton blooms in UK canals to determine the environmental predictors and spatial structure of bloom communities. The results revealed a significant increase in bloom presence with increasing elevation. As predicted, higher temperatures were associated with a greater probability of blooms, but the relationship between temperature and bloom occurrence changed across landscapes. At the minimum level of agricultural land, the probability of bloom presence increased with increasing temperature. Conversely, at the maximum level, the probability decreased with increasing temperature. This pattern could be due to higher temperatures increasing phytoplankton growth rates despite lower nutrient concentrations at low levels of agricultural land, and nutrient depletion by rapidly growing blooms at high levels of agricultural land and temperatures. Community composition exhibited spatial autocorrelation: nearby blooms were more similar than distant blooms. Hydrological distances through the canal network showed a stronger association with community dissimilarity than Euclidean distances, suggesting a role for hydrological connectivity in driving bloom formation and composition. This new knowledge regarding canal phytoplankton bloom origin and ecology could help inform measures to inhibit bloom formation.

Keywords: algal bloom, cyanobacteria, climate, land use, health, connectivity, canal

INTRODUCTION

Algal blooms cause significant harm to humans, the economy and wildlife (Landsberg 2002). Some bloom-forming algae, particularly freshwater cyanobacteria, synthesise toxins that can cause health problems for humans and other animals (Landsberg 2002; Codd et al. 2005; Malbrouck and Kestemont 2006). Furthermore, algal blooms result in a reduction in water clarity and potential oxygen depletion, negatively impacting aquatic organisms (Paerl and Otten 2013). Consequently, removal of these toxins from water systems and prevention of future contamination are of critical importance. Although algal blooms occur naturally, human activity has significantly increased the incidence of blooms (Anderson et al. 2002). Both freshwater and marine algal blooms are exacerbated by eutrophication from N and P inputs. Previously, P has been identified as the most problematic nutrient (Smith 2003) although subsequent work has illustrated the need to control both N and P depending upon the composition of the blooms and the nature of the ecosystem (Conley et al 2009). While blooms of N_2-fixing cyanobacteria were thought to prevent N-limitation in aquatic systems, subsequent experiments have shown that this N_2-fixation is unable to compensate for N limitation (Scott et al 2010). Based on experimental studies of whole-lake ecosystems, combinations of N and P have been shown to have the greatest effect, necessitating their combined control (Paerl et al 2016). In freshwater habitats, such as lakes, rivers and streams, this N loading is often derived from fertiliser runoff and animal waste from agriculture, and in the UK P is primarily pollution from sewage (Anderson et al. 2002). The costs associated with the damage caused by freshwater eutrophication and algal blooms in England and Wales have been estimated at £75.0-114.3 million per year, with an additional £54.8 million of costs is being spent per year on policy responses in order to attend to the damage (Pretty et al. 2003). However, Pretty et al. (2003) suggest that if eutrophication was prevented before the damage occurred, the costs would be reduced.

In the absence of a predictive framework for the control of algal blooms prior to their occurrence, research has focused on how to inhibit algal blooms shortly after formation.

Some studies have found cyanophages and viruses that could be introduced into water systems to control them (Brussaard, 2004; Yoshida et al. 2006). However, as with any biocontrol method, there are highly complex potential problems that can result from the release of other species as a means of control (Simberloff and Stiling 1996). It should also be pointed out that the treatment of algal blooms with cyanophages and viruses would not fix the underlying cause of the problem. In order to do this, measures by which water pollution can be prevented need investigation. While a 37-year study investigating the efficacy of reducing N input to control algal blooms suggested that N limitation does not reduce bloom formation (Schindler et al. 2008), recent experimental work has emphasised the role of both N and P in driving eutrophication in standing waters (Scott et al 2010; Paerl et al 2016). However, efforts to reduce the concentration of P in the River Frome, UK, by inhibiting river pollution from sources such as sewage treatment works, suggested that algal blooms could be reduced using a P-focused approach (Bowes et al. 2011).

Climate change is also predicted to affect algal bloom frequency. Increases in global temperatures are expected to benefit algal development as taxa such as cyanobacteria have higher growth rates in warmer waters (Johnk et al. 2008; Paerl and Huisman 2008; O'Neil et al. 2012). In the UK, predicted reduced rainfall in summer months will result in lower concentrations of dissolved oxygen and reduced river flow, leading to an accumulation of nutrients such as P in watercourses. Furthermore, unpredictable heavy rainfall will intermittently flood watercourses with nutrients from the land (Whitehead et al. 2009; Watts et al. 2015).

Rivers are a particular type of watercourse that pose a unique set of questions regarding algal blooms, due to the dendritic network structure of these waterways. Dendritic networks are characterised by primarily linear features separating into branches. The movement of aquatic and semiaquatic species is largely restricted within these connected channels, as they are generally unable to leave the network (Grant et al. 2007). Research involving

89 experimental microcosms found that connectivity in dendritic networks could influence the
90 transportation of species throughout these systems. In comparison to linear networks, the
91 active dispersal of six protist and one rotifer species occurred quicker in dendritic networks,
92 leading to faster colonisation of new areas (Seymour and Altermatt 2014). The flow of water
93 in natural dendritic networks, including rivers, could also potentially enable passive dispersal
94 of non-motile species of algae. Thus, the understanding of how network connectivity can
95 facilitate active or passive dispersal of species, such as algae, is important for understanding
96 the development of algal blooms in dendritic networks, such as rivers. This knowledge could
97 also be vital for invasive species research, such as in the case of the invasive freshwater
98 algae species *Didymosphenia geminata* which forms blooms in New Zealand and Canadian
99 rivers with low-nutrient conditions, as well as those with higher P and N levels (Kirkwood et
100 al. 2007; Bothwell and Kilroy 2011; Kilroy and Bothwell 2012).
101
102 Canals are another example of a dendritic network that contains algae and associated
103 blooms (Nagai et al. 2008; Zhu et al. 2015). As canals are manmade structures, they are of
104 economic importance to humans as a means of transport, for recreational activities (Willis
105 and Garrod 1991; Leuven et al. 2009), and as part of built heritage (Firth 2015). Importantly,
106 in the UK, 23 canal stretches are designated Sites of Special Scientific Interest (SSSI), some
107 of which are due to the presence of nationally rare species and habitats (Natural England
108 2016). Also, the design of canals for industrial transport means that they often flow through
109 densely populated urban areas. The study of algal blooms in canals (and, indeed, the
110 ecology of canals in general) is a neglected area, with little information known about the
111 origin and ecology of the blooms. Consequently, the conservation implications associated
112 with understanding the origin and ecology of algal blooms in canals is of some importance,
113 as such understanding has the potential to aid the protection of nationally rare species and
114 habitats. A study of the River Thames basin, UK, found that rivers that are connected to
115 canals have greater chlorophyll concentrations, indicating larger algal biomasses (Bowes et
116 al. 2012). Thus, canals may be intensifying the problem of algal blooms in rivers. Moreover,

as with rivers, canals are potentially important networks for movement of native (Kim and Mandrak 2016) and invasive species (Leuven et al. 2009; Strayer 2010; Altermatt 2013). Due to the construction of many canals occurring near urban areas and other areas of human activity, it is more likely that invasive species will be introduced into canals and subsequently disperse into rivers (Willis and Garrod 1991).

Much work has been done before on the drivers of algal and phytoplankton blooms. Instead, the main aims of this study were to investigate phytoplankton bloom ecology in canals to determine (i) the structure of the autocorrelation in the resulting residuals from models of bloom presence, and (ii) the spatial variability in the taxonomic composition of those blooms. We predict that the presence of blooms will exhibit a spatially-autocorrelated pattern, accounting for drivers of bloom formation, and that connectivity within the canal network will result in taxonomic compositions of phytoplankton blooms that are closer together geographically being more similar than those that are further apart. We test these hypotheses using a novel data source which arises from a bloom reporting system in operation in England.

METHODS

Land use data, including patterns of natural, agricultural, and urban land, were obtained from the Land Cover Map (LCM) 2007 (Centre for Ecology & Hydrology 2011) and elevation was derived from a digital elevation model (DEM) (Ordnance Survey 2016). From the UK Government's Department for Environment, Food and Rural Affairs (Defra), data concerning canals, reservoirs, locks, wharves, docks, and lakes, ponds and fisheries, were obtained. From the WorldClim dataset, two environmental BioClim variables were downloaded for the UK: Bio1 (mean annual temperature, ℃) and Bio5 (maximum temperature of the warmest month, ℃) (Hijmans et al. 2005; Haylock et al. 2008). These two variables were selected because water temperature is known to affect cyanobacterial growth, with higher temperatures causing an increase in growth rate (Johnk et al. 2008; O'Neil et al. 2012). Both

variables represent air temperature which was predicted to correlate positively with water temperature, however the uniform structure of canals (a relatively standardised depth, width, and profile) means that we might expect a spatially consistent relationship between atmospheric and water temperature. However, the maximum temperature of the warmest month may be considerably higher than all of the other months, with a cooler temperature throughout the rest of the year. Where this occurs, the mean annual temperature would be more useful as areas with a higher temperature will be warmer, on average, throughout the entire year, not just during the warmest month. Therefore, both variables were obtained as either one may influence phytoplankton bloom presence. The Canal & River Trust (CRT) and the Environment Agency (EA) provided phytoplankton bloom data for both canals and reservoirs. Both datasets originate from a bloom reporting system, and so the definition of a bloom for the purpose of this study is a visible aggregation of phytoplankton at the water surface. Since the canal network is used extensively by recreational boaters, we assume that survey effort is relatively high across the network. Water samples are collected containing phytoplankton cells and preserved in Lugol's iodine. The sample is then mixed thoroughly, and a representative subsample is transferred to a sedimentation tube. After settling, cells are identified and counted to give a density estimate for each taxon. The EA dataset includes the enumeration while the CRT dataset only contains presence/absence, and so the dataset was converted to all presence/absence to ensure that the data were comparable. While this data source does not give a standardised sample of blooms across the canal network, it provides a large number of samples from across the network that we believe represent an adequate view of where blooms occur. Details of the SSSI site canals in Great Britain were obtained from Natural England (Natural England 2016).

Initial analysis of the data was performed in ArcGIS 10.4.1 for Desktop (Esri 2016), with all layers projected in the British National Grid. In order to produce individual canal stretches in which to analyse the phytoplankton bloom data, the canal dataset was split into "pounds" (stretches of canal on the same elevation that are divided by locks) along the canals.

Subsequently, a 5 km buffer was produced around each resultant canal pound (n = 2,439). The land cover, DEM, and climate data were then clipped to these buffers and the mean, minimum, and maximum values were calculated for each buffer using R 3.3.2 (R Development Core Team 2016). The same buffers were used to extract the proportions of the areas of aggregated land cover types (woodland, grassland, agriculture, and urban). Subsequently, these proportions were arcsine square root transformed. Woodland comprised broadleaved and coniferous woodland land cover types. Grassland comprised rough grassland, neutral grassland, calcareous grassland, acid grassland, and fen, marsh and swamp land cover types. Agriculture comprised arable, horticulture, and improved grassland land cover types, and is assumed to be the main source of N entering the system. Urban comprised urban and suburban land cover types, and is assumed to be the main source of P entering the system. The locations of 279 unique sites in which phytoplankton blooms had been recorded by the EA were given in national grid references (NGRs). Northing and easting values were calculated using a converter equation in Microsoft Excel 2013 (permission granted by author, Ryan Burrell). Blooms were only included if they were identified as being within the canal network (including feeder streams and reservoirs), and any blooms located outside of the 5 km buffers were removed as they were deemed too far from the canals, leaving 93 bloom locations.

Statistical analysis

All statistical analyses were performed using the “Hmisc”, “MuMIn”, “car”, and “vegan” packages in the statistical software, R 3.3.2 (Bartoń 2015; Harrell 2016; Fox, et al. 2016; Oksanen, et al. 2017). The presence/absence of phytoplankton blooms was investigated in relation to the environmental predictor variables for each canal pound using generalised linear models (GLZs) with binomial errors. Spearman’s rank correlations performed between each of the predictor variables revealed that the mean, minimum and maximum values for the elevation, Bio1, and Bio5 variables were significantly correlated with each other ($\rho >$ 0.600, df = 2437, $P < 0.001$). Thus, only mean elevation, mean Bio1 and mean Bio5 were

retained in the models along with the proportions of the areas of the four aggregated land cover type variables. In addition, two-way interactions terms between mean Bio1 and the transformed proportion of agricultural land, and mean Bio5 and the transformed proportion of agricultural land, were included in the model. These interaction terms were included as it was predicted that a combination of the nutrient concentration derived from agricultural land and temperature would have a synergistic, as opposed to additive, effect on the presence of phytoplankton blooms.

VIF analysis of this full model and the Spearman's rank order correlations revealed multicollinearity (VIF>5) between mean Bio1 and mean Bio5, and the transformed proportions of urban and agricultural land. Consequently, mean Bio5 and the associated interaction term were removed from the model, as Bio1 is a more biologically important variable. Bio5 represents the maximum temperature of the warmest month; yet phytoplankton blooms were reported in all months, not just the summer months, likely due to peaks in chlorophyll in April-June while peak temperatures occur in August (e.g. Skidmore et al. 1998). Therefore, we argue that Bio1 is more appropriate as it represents the mean annual temperature. In addition, the transformed proportion of urban areas was removed as the elimination of agricultural areas (and the two interaction terms) from the model resulted in a higher ΔAICc value (30.9), than the elimination of urban areas (and the two interaction terms) (ΔAICc = 21.2). Hence, there is a greater decline in explanatory power when the transformed proportion of agricultural land is removed from the model. The full model included (i) mean annual temperature, (ii) elevation, and the proportions of (iii) agricultural, (iv) woodland, (v) grassland cover, and (vi) the interaction between temperature and agricultural land cover.

The dredge function ("MuMIn" package) was used on the full model to calculate the AICc values for a set of models, each containing a different possible combination of the variables. Since three models had ΔAICc < 2 compared to the top model, indicating negligible

229 difference in explanatory power, model averaging with shrinkage was performed. As the
230 odds and 95% C.I. of the resultant model could not be calculated due to model averaging,
231 the values were estimated from the top model.
232
233 To evaluate the role of distance and connectivity, we conducted three complementary spatial
234 analyses: non-spatial, pseudo-spatial, and network distance. The non-spatial model does not
235 take spatial autocorrelation into account and so represents a null model assuming all
236 locations are independent. The pseudo-spatial model used the Euclidian distance between
237 each canal pound as a measure of distance but did not take into account hydrological
238 connectivity along the network. The dist function was used on the centroid data to produce
239 pairwise geographical distances between each of the blooms. Finally, the network distance
240 used the distance along the canal network between each pair of pounds. The canal network
241 was imported into the riverdist package in R (Tyers 2017), and a hydrological distance matrix
242 was created for each pairwise distance between sites using the riverdistmat() function.
243 These three distance models were then incorporated into the analyses in order to explore
244 the spatial autocorrelation in the data. The residuals from the top GLZ model were analysed
245 for spatial autocorrelation using Moran's I based on the pseudo-spatial (Euclidian) and
246 network distance (hydrological) distance matrices. Finally, the full GLZ with binomial errors
247 was repeated, with spatial filtering performed on the model using the centroids of the canal
248 stretches for the pseudo-spatial and network distance data (Dormann et al. 2007). The
249 effectiveness of this control for autocorrelation was verified by performing Moran's I tests on
250 the model residuals with the spatial filters.
251
252 *Community composition*
253 A more conservative analysis was conducted to evaluate spatial patterns in the composition
254 of phytoplankton within each reported bloom. Bloom locations were only incorporated if they
255 were within 500m of the canal network, giving greater confidence in their location along the
256 hydrological system. Comparisons of the phytoplankton bloom community compositions in

this subset of blooms (n = 39) in connecting canal stretches were performed in relation to geographical distance. Presence-absence species-by-site matrices were transformed by Hellinger transformation using the decostand function ("vegan" package). Redundancy analysis (RDA) of the Hellinger-transformed data was computed in order to produce an ordination plot of the phytoplankton bloom sites by community compositions. The vegdist function ("vegan" package) was then used on the Hellinger transformed species data to produce pairwise Bray-Curtis dissimilarity matrices describing the ecological distance between each of the blooms. Subsequently, a Mantel test (with Spearman's rank order correlation due to non-normality of the two distance matrices (Shapiro-Wilk normality tests: $W > 0.601$, $P < 0.001$)) was performed between the community distance matrix and each of the Euclidian and hydrological distance matrices.

RESULTS

Canal phytoplankton blooms with species-level identification were reported from 1.6% (39 out of 2439) of the associated canal pounds between 1990 and 2014. The UK canal system is generally located in low-lying areas (median elevation 101.83 m; interquartile range (IQR) = 72.03 m). The temperature data revealed that there was only an approximately 3℃ difference between the sites with the highest and lowest mean annual temperatures (median = 9.33 ℃; IQR = 0.45 ℃). The landscape through which canals pass is dominated by agricultural land (median proportion cover = 0.64; IQR = 0.44), with a smaller coverage of grassland and woodland (0.04 and 0.05, respectively; IQR = 0.06 and 0.05, respectively) (for more details, see Table 1). Bloom composition varied from 1 to 127 taxa, with a mean taxonomic richness of 10.4 taxa (±1.3 SE). The most common species recorded from blooms were *Euglena* sp. (104 sites). Of particular interest are the toxic cyanobacteria *Microcystis* sp. (from 52 sites, including *M. aeruginosa* from 27 sites), *Anabaena* sp. (from 50 sites, including *A. flos-aquae* from 46 sites) and *Oscillatoria* sp. (from 67 sites, including *O. agardhii* from 28 sites). The identification of potentially toxic cyanobacteria from these samples emphasises the importance of understanding their ecology and control.

Presence/absence

Model selection produced three models containing subsets of these six predictor variables that had ΔAICc < 2. Model averaging with shrinkage found that four of the predictor variables had a significant effect on the presence of phytoplankton blooms, and were found in all three models. The two other variables were only present in one model (Table 3). The results revealed a significant increase in the proportion of phytoplankton bloom presence with an increase in elevation (Table 3) (Figure 2). The estimated odds and 95% C.I. for the averaged model, revealed that the odds of phytoplankton bloom presence increased by 9% (95% C.I. 3-14%) for each 10 m increase in elevation.

As the interaction term is significant (Table 3), the effect of mean annual temperature on the presence/absence of phytoplankton blooms depends on the transformed proportion of agricultural land. As the proportion of agricultural land increases, the effect of temperature on presence/absence changes (Figure 2). At the minimum level of agricultural land, the predicted probability of phytoplankton bloom presence increases with increasing temperature. Conversely, at the maximum level of agricultural land, the predicted probability of phytoplankton bloom presence decreases with increasing temperature. At the median level of agricultural land, the predicted probability of phytoplankton bloom presence remains relatively similar with increasing temperature, with only a slight increase observed. Due to the significance of the interaction term, the single main effects cannot be interpreted in isolation. However, the transformed proportion of agricultural land and the mean annual temperature are still important.

The GLMs with spatial filtering based on Euclidean or network distances between sites showed that there were no spatial eigenvectors that explain a significant proportion of the variance in the residuals of the models. Thus, there was no spatial autocorrelation in the data. The Moran's I test confirmed that no spatial autocorrelation was present in the

residuals of the non-spatial models using either the Euclidean (Moran's I = -2.912 x 10^{-04}, s = 0.341, P = 0.367) or network distances (Moran's I = -4.287 x 10^{-04}, s = -1.963, P = 0.975). As a result, no further incorporation of spatial data into the presence/absence analysis was attempted.

Community composition

The results of the Mantel tests revealed that the compositions of phytoplankton species are more similar in blooms that are closer together than blooms that are further apart. There was a significant positive correlation between the distance between phytoplankton bloom sites and the dissimilarity of those sites when distance was measured using Euclidean distances (Mantel r statistic = 0.183, df = 38, P = 0.001), and this correlation was stronger for hydrological distance (Mantel r statistic = 0.278, df = 38, P = 0.001).

DISCUSSION

Based on the results of this study, the environmental conditions found around the canals of the UK affect the probability of phytoplankton bloom presence. Phytoplankton blooms are more likely to be present at higher elevation canals. Furthermore, it was found that as the proportion of agricultural land surrounding the canal stretches increases, the effect of temperature on the likelihood of phytoplankton bloom presence changes. These variables were found to be significant at the non-spatial level, with no spatial autocorrelation observed in the data as demonstrated by the pseudo-spatial analysis. Nevertheless, spatial analysis revealed that the community compositions of phytoplankton blooms that are closer together are more similar than those that are further apart. Hydrological connectivity seems to be more important than Euclidean distance, as would be predicted if there was a role in structuring blooms for movement of propagules through the canal network.

The reason for the increased likelihood of phytoplankton blooms at higher elevations is not obvious (Figure 1). The growth rate of phytoplankton, such as cyanobacteria, is known to be

341 greater at higher water temperatures (Johnk et al. 2008; O'Neil et al. 2012). Thus, the
342 opposite outcome would be expected as higher temperatures are generally found at lower
343 elevations (Fitter et al. 1998; Ineson et al. 1998; Tipping et al. 1999). Nevertheless, blooms
344 have been documented at high elevation sites in the past (Mwaura et al. 2004; Derlet et al.
345 2010; Anderson et al. 2014; Zhang et al. 2016). However, it should be noted that the
346 altitudinal gradient of this study area is not particularly large compared to other areas (Table
347 1), which could have affected the results. A potential reason for this unexpected result could
348 be that there is greater precipitation at high elevations (Ineson et al. 1998; Tipping et al.
349 1999); thus, larger quantities of pollutants may be washed into the canals. This effect of
350 greater run-off, combined with higher levels of N and P that have not yet been stripped from
351 the water supplies as much as downstream, could lead to higher nutrient availability for
352 blooms. Blooms are known to occur in upland reservoirs that feed into the canal network,
353 which could also result in concentrations of blooms in upland areas. However, Figure 1 also
354 shows a number of blooms that arise close to urban areas (London, West Midlands,
355 Liverpool) and which might be indicative of local P pollution via sewage entering the system.
356 A recent study found that the effect of nutrients on blooms is greater than water temperature
357 (Deng et al. 2014); hence, the potentially higher nutrient concentrations caused by greater
358 precipitation may compensate for the decrease in temperature. Furthermore, a mesocosm
359 experiment with marine phytoplankton suggested light as an important factor for bloom
360 initiation (Sommer and Lengfellner 2008). Potentially fewer or smaller trees at upland canal
361 sites may result in greater light intensity and thus, an increased likelihood of bloom presence
362 (Coomes and Allen 2007). For example, canal stretches that traverse upland moors may be
363 running through entirely deforested areas. Previous research suggests that reforestation
364 along the edges of waterways could reduce bloom growth more effectively than decreasing
365 eutrophication, by reducing light intensity (Hutchins et al. 2010). This complex spatial pattern
366 of bloom formation, combined with issues of hydrological connectivity, raises a series of
367 hypotheses that should be tested in future studies in order to inform local control measures
368 based on local problems.

The interaction between the transformed proportion of agricultural land and the mean annual temperature was also not as predicted. Based on previous research, agricultural land is often associated with the formation of phytoplankton blooms (e.g. Bussi et al., 2016; Hamilton et al., 2016). This is due to the leaching of fertilisers and animal waste into waterways, leading to increased concentrations of N and P; two nutrients that are key drivers of phytoplankton blooms (Anderson et al. 2002; Smith 2003). Moreover, higher temperatures are known to be beneficial for phytoplankton species such as cyanobacteria due to their high thermal optima for growth rates (Johnk et al. 2008; O'Neil et al. 2012). In contrast, the interaction reveals that the effect of agricultural land on the probability of phytoplankton bloom presence differs depending on the temperature (Figure 2). At the minimum level of agricultural land, the predicted probability of bloom presence increases with increasing temperature. This can be explained by previous research regarding the effect of agricultural pollution and temperature on bloom formation (Anderson et al. 2002; Smith 2003; Johnk et al. 2008; O'Neil et al. 2012). At low levels of agricultural land, N and P may be at low concentrations, limiting the formation of phytoplankton blooms (Anderson et al. 2002). Nevertheless, as long as those low concentrations are not limiting, an increase in temperature may overcome these low concentrations by increasing the phytoplankton growth rate, leading to an increased probability of bloom formation (Johnk et al. 2008; O'Neil et al. 2012). At intermediate levels of agricultural land, nutrients may no longer be a limiting factor for phytoplankton bloom formation, as they may be present at sufficient concentrations. Hence, increasing temperature may not result in an increased probability of bloom presence, as nutrients are of greater importance than temperature and sufficient nutrients may be provided (Deng et al. 2014). However, the results suggest that at high levels of agricultural land, the predicted probability of phytoplankton bloom presence decreases with increasing temperature. The reason for this may be that when there are high concentrations of nutrients available as well as a higher temperature, the phytoplankton blooms may grow excessively leading to depletion of the nutrients available in the water

(Smayda 1998; Winder and Cloern 2010). In addition, cell sinking and consumption of algae by predators can occur as the blooms peak (Smayda 1998; Van Wichelen et al. 2010; Winder and Cloern 2010). The rate of this algae consumption is known to increase at higher temperatures (Sommer and Lengfellner 2008). Consequently, the blooms may collapse shortly after they peak (Smayda 1998; Van Wichelen et al. 2010; Winder and Cloern, 2010), resulting in fewer reported blooms at sites with both a high level of agricultural land and a higher temperature. However, other research has suggested that blooms can continue for months even when ambient concentrations of N and P are low (Paerl and Otten 2013), emphasising a role for internal nutrient cycling and regeneration.

Factors suggested as potential controls for blooms include grazing by predators, and bacterial and viral lysis. However, despite the potential controlling effect of grazing on blooms, some phytoplankton are known to survive travelling through the digestive system of grazers such as *Daphnia*, and are even capable of extracting nutrients from the gut (Porter 1976; VanDonk et al. 1997). Furthermore, the sinking of large quantities of decaying phytoplankton material can result in hypoxia, leading to death of other aquatic organisms and changes to the biogeochemical cycling of the waterway. The collapse of blooms can also release dissolved toxins into the water (Paerl and Otten 2013). Due to these problems associated with controlling blooms and bloom senescence, the prevention of algae blooming in the first place is of critical importance.

We expected that agricultural land would be related to the presence of phytoplankton blooms due to the known effect of agricultural fertilisers and animal waste on algae (Anderson et al. 2002). However, the presence of agricultural land does not necessitate the application of fertilisers. There have been efforts in recent years to try to reduce eutrophication and the associated blooms, which increased as a result of industrial and agricultural intensification (Anderson et al. 2002). For example, EU agri-environment schemes promote the termination of fertiliser application and lower livestock densities (Kleijn and Sutherland 2003). Thus, it

425 cannot be assumed that agricultural land in 21st century Great Britain leads to the
426 eutrophication of waterways. In addition, run-off of nutrients into canals may not occur in the
427 same way as natural waterways, such as rivers. The ease with which nutrients enter canals
428 could be inhibited by the material used to construct the sides of the canals, for example
429 concrete (Holland and Andrews 1998). The nutrient concentrations of the canal stretches
430 were not sampled as part of this study. Therefore, even if there is a high proportion of
431 agricultural land located around canal stretches, it does not mean that nutrients will be
432 leaching into the waterways.
433
434 The fact that phytoplankton blooms that are closer together have more similar community
435 compositions suggests that these blooms are related. It is possible that algae in upland
436 reservoirs and canal stretches are flowing down the canals and forming additional blooms in
437 other areas. This information could be useful for preventing future phytoplankton blooms by
438 identifying the origin of blooms and preventing eutrophication in these areas. Dispersal was
439 also suggested by Altermatt et al. (2013) as a reason for greater aquatic insect community
440 dissimilarity with increasing distance in dendritic river networks. Spatial along-stream
441 distances were utilised in the analysis, and those findings are corroborated by our results.
442 The research also suggested that environmental conditions could explain the community
443 similarity patterns, as elevation had a significant effect on the pattern and is a factor that
444 affects conditions such as temperature and precipitation (Altermatt et al. 2013).
445
446 Other research has found that phytoplankton bloom community compositions are dependent
447 on the environmental conditions of the waterway, such as turbidity and nutrient
448 concentrations. Different phytoplankton species have different optimal conditions and
449 therefore thrive in different environments, leading to diverse compositions of species (Smith
450 1983; Zhu et al. 2015). For example, cyanobacteria are known to take over phytoplankton
451 communities when there is a low N:P ratio. This could be due to the N_2-fixing abilities of
452 many cyanobacteria species, leading to a competitive dominance where N concentrations

453 are low and P concentrations are high (Smith 1983). Furthermore, temporal changes in
454 compositions have been observed, with community succession associated with temporal
455 changes in environmental conditions, particularly nutrient concentrations (Deng et al. 2014).
456 The effect of the environment on specific phytoplankton communities may therefore allow
457 blooms to persist even when the optimal conditions for a particular composition of species
458 change, as the proportion of each species in the bloom will fluctuate (Smayda 1998). This
459 presents problems with regard to controlling phytoplankton blooms as they may be resistant
460 to environmental change. However, if the aim is to only control nuisance species such as
461 toxin-producing cyanobacteria (Landsberg 2002; Codd et al. 2005; Malbrouck and
462 Kestemont 2006), this may be possible by producing conditions that are not optimal for these
463 specific species. For example, cyanobacterial blooms could be inhibited by increasing the
464 N:P ratio (Smith 1983).
465
466 Previous research comparing the results of terrestrial, 'as the crow flies' distances (pseudo-
467 spatial analysis) with aquatic, 'as the algae flows' distances (network distance analysis)
468 found differences in the pattern of results. Network distance analysis kept more spatial
469 variables in the model compared to Euclidean distance analysis (Landeiro et al. 2011).
470 Another study also suggested that network distance analysis would account for spatial
471 autocorrelation in a way that is more appropriate for dendritic networks such as canals, than
472 pseudo-spatial analysis. Furthermore, it will prevent violation of the statistical assumption
473 that observations are independent and prevent inaccurate statistical inference, caused by
474 clustering of measurements (Isaak et al. 2014). These studies highlight the importance of
475 using network distances rather than traditional Euclidean distances for analysing species
476 data in dendritic networks such as canals, rivers and streams (Landeiro et al. 2011; Isaak et
477 al. 2014). Landeiro et al. (2011) also suggested that this method may have implications for
478 terrestrial analyses where the environment is fragmented or the dispersal of the study
479 species is limited, for example. Nevertheless, Euclidean, overland distances may still be
480 useful for studying semiaquatic or amphibiotic species (Landeiro et al. 2011).

We make use of a novel dataset derived from an algal bloom reporting system. This dataset has the advantage of broad spatial scale, detailed taxonomic information, and a growing time series of bloom locations. However, the data lack accompanying water chemistry (especially N and P) data, making certain hypotheses difficult to evaluate. However, we feel that the insights produced in the study are of value as they focus on an understudied ecosystem and present some novel findings based on the external (land use) and internal (hydrological connectivity) drivers of bloom formation and taxonomic composition that can form the basis of subsequent work. In particular, the data from the models that inform the spatial autocorrelation of bloom formation could be strengthened by the addition of other variables. First, flushing rates (or retention times) are a key predictor of bloom formation and an important method of control (Paerl et al 2011), but are complex to calculate within canal systems. In the UK, “lockage” (the frequency of opening locks) is recorded and there are some flow gauges at certain sites around the network, but it is unclear how this relates to flow in the network as a whole. Second, the retrospective nature of the study means that nutrient concentrations are not available to accompany the analysis, while previous work suggests that there are complex interactions between N and P cycling that drive cyanobacterial bloom formation and senescence (Paerl et al 2016). Finally, there may be complex interactions between land use and topography, via the impacts of slope on the rate and composition of run-off in the different canal basins (Li et al 2006). Current attempts to reforest uplands as part of natural flood management or incorporate trees into agroforestry practices may influence this relationship further (Pavlidis and Tsihrintzis 2018).

A number of canal stretches in Great Britain are designated SSSI sites, some of which are due to the presence of nationally rare species and habitats (Natural England 2016). Thus, it is of critical importance that phytoplankton blooms do not damage these sites. Bloom data analysed in this study reveal that phytoplankton blooms have occurred in at least 13 out of the 23 SSSI site canals in the past. As this study found that higher elevation is associated

509 with increased phytoplankton bloom presence, measures could be implemented to prevent
510 eutrophication in upland areas. Investigations of land use surrounding upland canal sites will
511 determine the most appropriate way to achieve this. In addition, the interaction term
512 suggests that a smaller proportion of agricultural land, and thus a lower nutrient
513 concentration in the canals, will result in a decreased probability of phytoplankton bloom
514 presence when the temperature is lower. Thus, preventing eutrophication in upland canal
515 stretches where the temperature is typically lower will hopefully inhibit the formation of
516 blooms (Smith 1983; Fitter et al. 1998; Ineson et al. 1998; Tipping et al. 1999; Schindler et
517 al. 2008; Bowes et al. 2011). This will protect downstream sites, as the community
518 composition analysis indicates that phytoplankton blooms may percolate down through the
519 network to seed further blooms at lower elevations, where conditions are appropriate.
520 Reforestation along canals could also aid with the inhibition of blooms by reducing light
521 intensity (Hutchins et al. 2010). As discussed above, it is also essential to prevent blooms
522 rather than control them once they have formed, as senescing blooms could result in the
523 release of dissolved toxins into the water and could lead to hypoxia in the canals (Paerl and
524 Otten 2013). Moreover, for invasive species such as other phytoplankton and macrophytes,
525 this information regarding the movements of cyanobacteria could prove important. This new
526 knowledge regarding the origin and ecology of canal phytoplankton blooms could therefore
527 aid with the protection of nationally rare species and habitats in SSSI site canals, as well as
528 potentially help improve other non-SSSI site canals. Furthermore, prevention of blooms in
529 canals will benefit human health through improved safety during transport and recreational
530 activities (Willis and Garrod 1991; Falconer 1999; Leuven et al. 2009).
531

532 **ACKNOWLEDGEMENTS**

533 The authors would like to thank the Environment Agency and the Canal & River Trust for
534 providing the phytoplankton bloom datasets, and two anonymous reviewers who provided
535 detailed and insightful comments that greatly improved the manuscript.
536

REFERENCES

Altermatt F. 2013. Diversity in riverine metacommunities: a network perspective. *Aquatic Ecol.* 47(3):365-377.

Altermatt F, Seymour M, Martinez N. 2013. River network properties shape alpha-diversity and community similarity patterns of aquatic insect communities across major drainage basins. *J Biogeogr.* 40(12):2249-2260.

Anderson DM, Glibert PM, Burkholder JM. 2002. Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences. *Estuaries.* 25(4B):704-726.

Anderson IJ, Saiki MK, Sellheim K, Merz JE. 2014. Differences in benthic macroinvertebrate assemblages associated with a bloom of *Didymosphenia geminata* in the lower American River, California. *Southwest Nat.* 59(3):389-395.

Bartoń K. 2015. MuMIn: Multi-model inference [R package]. Version 1.15.6. https://cran.r-project.org/web/packages/MuMIn/index.html.

Bothwell ML, Kilroy C. 2011. Phosphorus limitation of the freshwater benthic diatom *Didymosphenia geminata* determined by the frequency of dividing cells. *Freshw Biol.* 56(3):565-578.

Bowes MJ, Gozzard E, Johnson AC, Scarlett PM, Roberts C, Read DS, Armstrong LK, Harman SA, Wickham HD. 2012. Spatial and temporal changes in chlorophyll-a concentrations in the River Thames basin, UK: are phosphorus concentrations beginning to limit phytoplankton biomass? *Sci Total Environ.* 426:45-55.

Bowes MJ, Smith JT, Neal C, Leach DV, Scarlett PM, Wickham HD, Harman SA, Armstrong LK, Davy-Bowker J, Haft M, Davies CE. 2011. Changes in water quality of the River Frome (UK) from 1965 to 2009: is phosphorus mitigation finally working? *Sci Total Environ.* 409(18):3418-3430.

Brussaard CPD. 2004. Viral control of phytoplankton populations - a review. *J Eukaryot Microbiol.* 51(2):125-138.

563 Bussi G, Whitehead PG, Bowes MJ, Read DS, Prudhomme C, Dadson SJ (2016) Impacts of
564 climate change, land-use change and phosphorus reduction on phytoplankton in the
565 River Thames (UK). *Science of the Total Environment*. 572, 1507-1519.
566 Carvalho GA, Minnett PJ, Fleming LE, Banzon VF, Baringer W. 2010. Satellite remote
567 sensing of harmful algal blooms: a new multi-algorithm method for detecting the
568 Florida Red Tide (*Karenia brevis*). *Harmful Algae*. 9(5):440-448.
569 Centre for Ecology & Hydrology. 2011. Land Cover Map 2007 Great Britain [TIFF geospatial
570 data]. Edinburgh (UK): University of Edinburgh. 1:250,000.
571 Codd GA, Morrison LF, Metcalf JS. 2005. Cyanobacterial toxins: risk management for health
572 protection. *Toxicol Appl Pharmacol.* 203(3):264-272.
573 Conley, DJ, Paerl HW, Howarth RW, Boesch DF, Seitzinger SP, Havens, KE, Lancelot C,
574 Likens GE. 2009. Controlling Eutrophication: Nitrogen and Phosphorus. *Science* 323:
575 1014-1015.
576 Coomes DA, Allen RB. 2007. Effects of size, competition and altitude on tree growth. *J Ecol.*
577 95(5):1084-1097.
578 Deng JM, Qin BQ, Paerl HW, Zhang YL, Wu P, Ma JR, Chen YW. 2014. Effects of nutrients,
579 temperature and their interactions on spring phytoplankton community succession in
580 Lake Taihu, China. *PLoS ONE.* 9(12): e113960. doi:10.1371/ journal.pone.0113960.
581 Derlet RW, Goldman CR, Connor MJ. 2010. Reducing the impact of summer cattle grazing
582 on water quality in the Sierra Nevada Mountains of California: a proposal. *J Water*
583 *Health.* 8(2):326-333.
584 Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, Davies RG, Hirzel A,
585 Jetz W, Kissling WD, Kühn I, Ohlemüller R, Peres-Neto PR, Reineking B, Schröder
586 B, Schurr FM, Wilson R. 2007. Methods to account for spatial autocorrelation in the
587 analysis of species distributional data: a review. *Ecography*. 30(5):609-628.
588 Eminniyaz A, Qiu J, Tan DY, Baskin CC, Baskin JM, Nowak RS. 2013. Dispersal
589 mechanisms of the invasive alien plant species buffalobur (*Solanum rostratum*) in
590 cold desert sites of Northwest China. *Weed Sci.* 61(4):557-563.

591 Esri. 2016. ArcGIS for Desktop [Software]. Version 10.4.1. Redlands (CA): Esri. [accessed
592 2017 Nov 7].
593 Falconer IR. 1999. An overview of problems caused by toxic blue-green algae
594 (cyanobacteria) in drinking and recreational water. *Environ Toxicol.* 14(1):5-12.
595 Firth A. 2015. Heritage Assets in Inland Waters: An Appraisal of Archaeology Underwater in
596 England's Rivers and Canals. *Hist Environ: Policy Pract.* 6(3):229-239.
597 Fitter AH, Graves JD, Self GK, Brown, TK, Bogie DS, Taylor K. 1998. Root production,
598 turnover and respiration under two grassland types along an altitudinal gradient:
599 influence of temperature and solar radiation. *Oecologia.* 114(1):20-30.
600 Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, Ellison S, Firth D, Friendly M, Gorjanc
601 G, Graves S, Heiberger R, Laboissiere R, Monette G, Murdoch D, Nilsson H, Ogle D,
602 Ripley B, Venables W, Winsemius D, Zeileis A, R-Core. 2016. car: Companion to
603 applied regression [R package]. Version 2.1-4. https://cran.r-
604 project.org/web/packages/car/index.html.
605 Grant EHC, Lowe WH, Fagan WF. 2007. Living in the branches: population dynamics and
606 ecological processes in dendritic networks. *Ecol Lett.* 10(2):165-175.
607 Hamilton DP, Salmaso N, Paerl HW (2016) Mitigating harmful cyanobacterial blooms:
608 strategies for control of nitrogen and phosphorus loads. *Aquatic Ecology.* 50, 351-
609 366.
610 Harrell FE Jr. 2016. Hmisc: Harrell miscellaneous [R package]. Version 4.0-2. https://cran.r-
611 project.org/web/packages/Hmisc/.
612 Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European
613 daily high-resolution gridded dataset of surface temperature and precipitation. *J*
614 *Geophys Res*. 113(D20119). doi:10.1029/2008JD010201.
615 Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution
616 interpolated climate surfaces for global land areas. *Int J Climatol.* 25(15):1965-1978.
617 Holland GJ, Andrews ME. 1998. Inspection and risk assessment of slopes associated with
618 the UK canal network. *Geohazards in Eng Geol.* 15:155-165.

619 Hutchins MG, Johnson AC, Deflandre-Vlandas A, Comber S, Posen P, Boorman D. 2010.
620 Which offers more scope to suppress river phytoplankton blooms: reducing nutrient
621 pollution or riparian shading. *Sci Total Environ*. 408(21):5065-5077.
622 Ineson P, Taylor K, Harrison AF, Poskitt J, Benham DG, Tipping E, Woof C. 1998. Effects of
623 climate change on nitrogen dynamics in upland soils. 1. A transplant approach. *Glob*
624 *Chang Biol.* 4(2):143-152.
625 Isaak DJ, Peterson EE, Ver Hoef JM, Wenger SJ, Falke JA, Torgersen CE, Sowder C, Steel
626 EA, Fortin M, Jordan CE, Reusch AS, Som N, Monestiez P. 2014. Applications of
627 spatial statistical network models to stream data. *WIREs Water*. 1(3):277-294.
628 Johnk KD, Huisman J, Sharples J, Sommeijer B, Visser PM, Stroom JM. 2008. Summer
629 heatwaves promote blooms of harmful cyanobacteria. *Glob Chang Biol.* 14(3):495-
630 512.
631 Kilroy C, Bothwell ML. 2012. *Didymosphenia geminata* growth rates and bloom formation in
632 relation to ambient dissolved phosphorus concentration. *Freshw Biol.* 57(4):641-653.
633 Kim J, Mandrak NE. 2016. Assessing the potential movement of invasive fishes through the
634 Welland Canal. *J Great Lakes Res.* 42(5):1102-1108.
635 Kirkwood AE, Shea T, Jackson L, McCcauley E. 2007. *Didymosphenia geminata* in two
636 Alberta headwater rivers: an emerging invasive species that challenges conventional
637 views on algal bloom development. *Can J Fish Aquat Sci.* 64(12):1703-1709.
638 Kleijn D, Sutherland WJ. 2003. How effective are European agri-environment schemes in
639 conserving and promoting biodiversity? *J Appl Ecol.* 40(6):947-969.
640 Landeiro VL, Magnusson WE, Melo AS, Espirito-Santo HMV, Bini LM. 2011. Spatial
641 eigenfunction analyses in stream networks: do watercourse and overland distances
642 produce different results? *Freshw Biol.* 56(6):1184-1192.
643 Landsberg JH. 2002. The effects of harmful algal blooms on aquatic organisms. *Rev Fish*
644 *Sci.* 10(2):113-390.

645 Leuven R, van der Velde G, Baijens I, Snijders J, van der Zwart C, Lenders HJR, bij de
646 Vaate AB. 2009. The river Rhine: a global highway for dispersal of aquatic invasive
647 species. *Biol Invasions.* 11(9):1989-2008.
648 Li Y, Wang C, Tang H. 2006. Research advances in nutrient runoff on sloping land in
649 watersheds. *Aquatic Ecosystem Health & Management*, 9, 27-32.
650 Malbrouck C, Kestemont P. 2006. Effects of microcystins on fish. *Environ Toxicol Chem.*
651 25(1):72-86.
652 Mwaura F, Koyo AO, Zech B. 2004. Cyanobacterial blooms and the presence of cyanotoxins
653 in small high altitude tropical headwater reservoirs in Kenya. *J Water Health.* 2(1):49-
654 57.
655 Nagai T, Imai A, Matsushige K, Yokoi K, Fukushima T. 2008. Short-term temporal variations
656 in iron concentration and speciation in a canal during a summer algal bloom. *Aquatic*
657 *Sci.* 70(4):388-396.
658 Natural England. 2016. Designated Sites View. [accessed 2016 Oct 25].
659 https://designatedsites.naturalengland.org.uk/SiteList.aspx?siteName=Canal.
660 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
661 RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2017. vegan:
662 Community ecology package [R package]. Version 2.4-2. https://cran.r-
663 project.org/web/packages/vegan/index.html.
664 O'Neil JM, Davis TW, Burford MA, Gobler CJ. 2012. The rise of harmful cyanobacteria
665 blooms: the potential roles of eutrophication and climate change. *Harmful Algae.*
666 14:313-334.
667 Ordnance Survey. 2016. OS Terrain 50 DTM England and Wales [TIFF geospatial data].
668 1:50,000 [online]; [accessed 2016 Nov 7]; http://digimap.edina.ac.uk/.
669 Paerl HW, Huisman J. 2008. Blooms Like It Hot. *Science*. 320(5872):57-58.
670 Paerl HW, Otten TG. 2013. Harmful cyanobacterial blooms: causes, consequences, and
671 controls. *Microb Ecol.* 65(4):995-1010.

672 Paerl HW, Hall NS, Calandrino ES. 2011. Controlling harmful cyanobacterial blooms in a
673 world experiencing anthropogenic and climatic-induced change. *Science of the Total*
674 *Environment*, 409, 1739-1745.
675 Paerl, HW, Scott JT, McCarthy MJ, Newell SE, Gardner WS, Havens KE, Hoffman DK,
676 Wilhelm SW, Wurtsbaugh WA. 2016. It takes two to tango: When and where dual
677 nutrient (N & P) reductions are needed to protect lakes and downstream ecosystems.
678 *Environmental Science & Technology.* 50: 10805−10813.
679 Pavlidis G, Tsihrintzis VA (2018) Environmental Benefits and Control of Pollution to Surface
680 Water and Groundwater by Agroforestry Systems: a Review. *Water Resources*
681 *Management*, 32, 1-29.
682 Porter KG. 1976. Enhancement of algal growth and productivity by grazing zooplankton.
683 *Science.* 192(4246):1332-1334.
684 Pretty JN, Mason CF, Nedwell DB, Hine RE, Leaf S, Dils R. 2003. Environmental costs of
685 freshwater eutrophication in England and Wales. *Environ Sci Technol*. 37(2):201-208.
686 R Development Core Team. 2016. R: a language and environment for statistical computing
687 [Software]. Version 3.3.2. Vienna, Austria: The R Foundation for Statistical
688 Computing. [accessed 2017 Jan 12].
689 Schindler DW, Hecky RE, Findlay DL, Stainton MP, Parker BR, Paterson MJ, Beaty KG,
690 Lyng M, Kasian SEM. 2008. Eutrophication of lakes cannot be controlled by reducing
691 nitrogen input: results of a 37-year whole-ecosystem experiment. *Proc Natl Acad Sci*
692 *U S A.* 105(32):11254-11258.
693 Scott, JT; McCarthy, MJ. 2010. Nitrogen fixation may not balance the nitrogen pool in lakes
694 over timescales relevant to eutrophication management. *Limnol. Oceanogr.* 55,
695 1265–1270
696 Seymour M, Altermatt F. 2014. Active colonization dynamics and diversity patterns are
697 influenced by dendritic network connectivity and species interactions. *Ecol Evol.*
698 4(8):1243-1254.

699 Shen L, Xu H, Guo X. 2012. Satellite remote sensing of harmful algal blooms (HABs) and a
700 potential synthesized framework. *Sensors*. 12(6):7778-7803.
701 Simberloff D, Stiling P. 1996. How risky is biological control? *Ecology.* 77(7):1965-1974.
702 Skidmore RE, Maberly SC, Whitton BA (1998) Patterns of spatial and temporal variation in
703 phytoplankton chlorophyll a in the River Trent and its tributaries. *Science of the Total*
704 *Environment*. 210-211, 357-365.
705 Smayda TJ. 1998. Patterns of variability characterizing marine phytoplankton, with examples
706 from Narragansett Bay. *ICES J Mar Sci.* 55(4):562-573.
707 Smith VH. 1983. Low nitrogen to phosphorus ratios favour dominance by blue-green-algae
708 in lake phytoplankton. *Science.* 221(4611):669-671.
709 Smith VH. 2003. Eutrophication of freshwater and coastal marine ecosystems - a global
710 problem. *Environ Sci Pollut Res.* 10(2):126-139.
711 Sommer U, Lengfellner K. 2008. Climate change and the timing, magnitude, and
712 composition of the phytoplankton spring bloom. *Glob Chang Biol.* 14(6):1199-1208.
713 Strayer DL. 2010. Alien species in fresh waters: ecological effects, interactions with other
714 stressors, and prospects for the future. *Freshw Biol.* 55:152-174.
715 Tipping E, Woof C, Rigg E, Harrison AF, Ineson P, Taylor K, Benham D, Poskitt J, Rowland
716 AP, Bol R, Harkness DD. 1999. Climatic influences on the leaching of dissolved
717 organic matter from upland UK Moorland soils, investigated by a field manipulation
718 experiment. *Environ Int.* 25(1):83-95.
719 Tyers M. 2017. riverdist: River network distance computation and applications [R package].
720 Version 0.13.1. https://cran.r-project.org/web/packages/riverdist/index.html.
721 Van Wichelen J, van Gremberghe I, Vanormelingen P, Debeer A-E, Leporcq B, Menzel D,
722 Codd GA, Descy J-P, Vyverman W. 2010. Strong effects of amoebae grazing on the
723 biomass and genetic structure of a *Microcystis* bloom (Cyanobacteria). *Environ*
724 *Microbiol.* 12(10):2797-2813.

725 VanDonk E, Lurling M, Hessen DO, Lokhorst GM. 1997. Altered cell wall morphology in
726 nutrient-deficient phytoplankton and its impact on grazers. *Limnol Oceanogr.*
727 42(2):357-364.
728 Ver Hoef J, Peterson E. 2016. SNN: Spatial modeling on stream networks [R package].
729 Version 1.1.8. https://cran.r-project.org/web/packages/SSN/index.html.
730 Watts G, Battarbee RW, Bloomfield JP, Crossman J, Daccache A, Durance I, Elliott JA,
731 Garner G, Hannaford J, Hannah DM, Hess T, Jackson CR, Kay AL, Kernan M, Knox
732 J, Mackay J, Monteith DT, Ormerod SJ, Rance J, Stuart ME, Wade AJ, Wade SD,
733 Weatherhead K, Whitehead PG, Wilby RL. 2015. Climate change and water in the
734 UK - past changes and future prospects. *Prog Phys Geogr.* 39(1):6-28.
735 Whitehead PG, Wilby RL, Battarbee RW, Kernan M, Wade AJ. 2009. A review of the
736 potential impacts of climate change on surface water quality. *Hydrol Sci J.* 54(1):101-
737 123.
738 Willis K, Garrod G. 1991. Valuing open access recreation in inland waterways - on-site
739 recreation surveys and selection effects. *Reg Stud.* 25(6):511-524.
740 Winder M, Cloern JE. 2010. The annual cycles of phytoplankton biomass. *Philos Trans R*
741 *Soc B*. 365(1555):3215-3226.
742 Yoshida T, Takashima Y, Tomaru Y, Shirai Y, Takao Y, Hiroishi S, Nagasaki K. 2006.
743 Isolation and characterization of a cyanophage infecting the toxic cyanobacterium
744 Microcystis aeruginosa. *Appl Environ Microbiol.* 72(2):1239-1247.
745 Zhang C, Zhang J. 2015. Current techniques for detecting and monitoring algal toxins and
746 causative harmful algal blooms. *J Environ Anal Chem*. 2(1). doi: 10.4172/2380-
747 2391.1000123.
748 Zhang Q, Zhu H, Hu Z, Liu G. 2016. Blooms of the woloszynskioid dinoflagellate *Tovellia*
749 *diexiensis* sp. nov. (Dinophyceae) in Baishihai Lake at the eastern edge of Tibetan
750 Plateau. *Algae.* 31(3):205-217.

751 Zhu WJ, Pan YD, You QM, Pang WT, Wang YF, Wang QX. 2015. Phytoplankton
752 assemblages in a newly man-made shallow lake and surrounding canals, Shanghai,
753 China. *Aquatic Ecol.* 49(2):147-157.

754 **Table Legends**

755 **Table 1** The minimum, median and maximum values of each environmental variable
756 calculated for each 5 km canal stretch buffer. The median values were calculated due to the
757 non-normal distribution of each variable (Shapiro-Wilk normality tests: $W > 0.915$, $P <$
758 0.001). Agricultural land, woodland and grassland denote the untransformed proportion of
759 each land cover type.

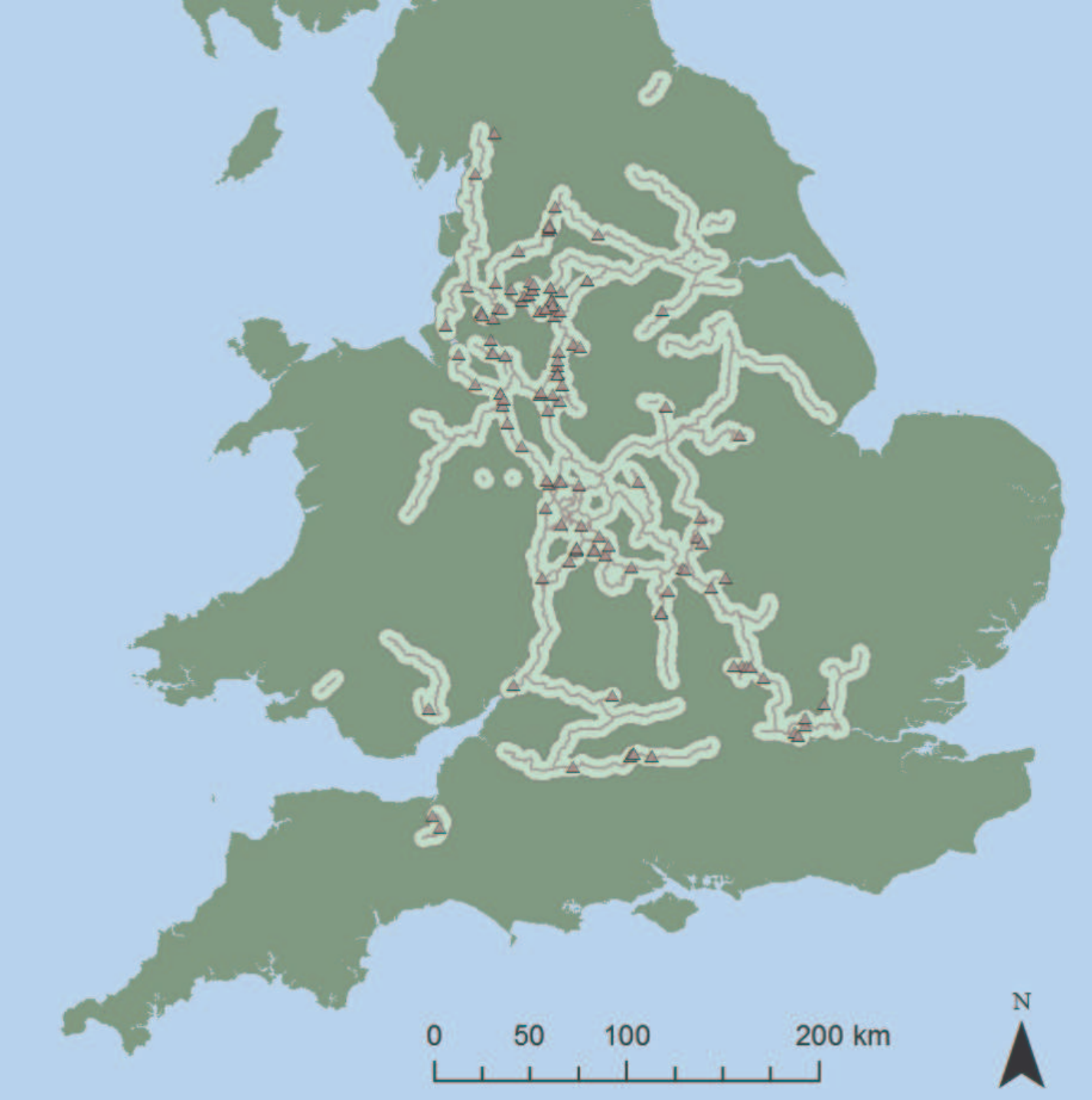
760 **Table 2** The generalised linear models with binomial errors output for the six environmental
761 predictor variables following model averaging with shrinkage. “Model presence” denotes the
762 number of models each variable was present in. Significant terms are marked in **bold.** See
763 text for details.

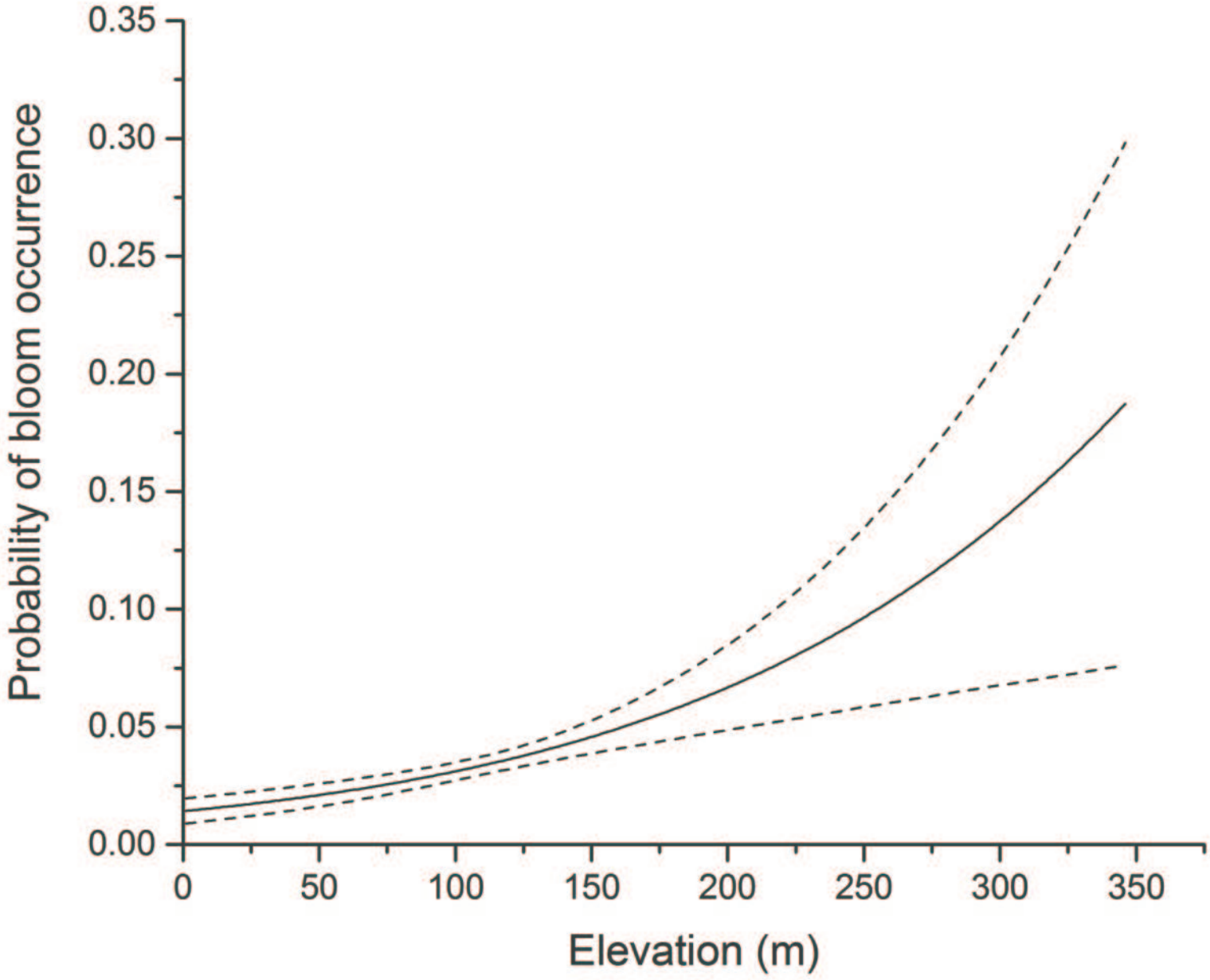
764 **Figure Legends**

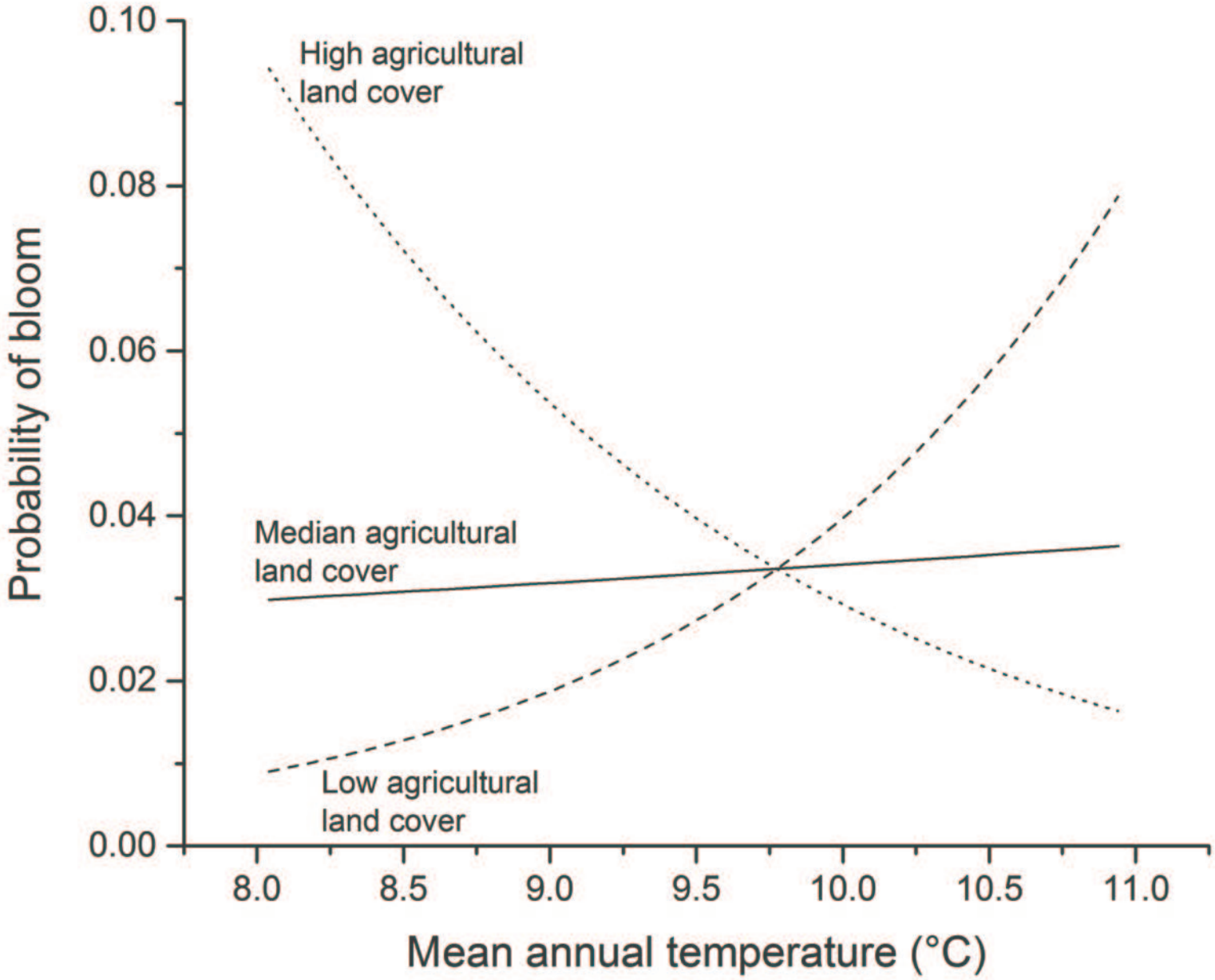
765 **Figure 1** Locations of phytoplankton blooms (marked with triangles) within the UK canal
766 network (5km buffer shown around each canal stretch).

767 **Figure 2** The predicted probability of phytoplankton bloom presence at different levels of
768 elevation (solid line), with the standard errors displayed (dotted lines).

769 **Figure 3** The predicted probability of phytoplankton bloom presence at differing transformed
770 proportions of agricultural land, with increasing mean annual temperature (℃).


0
50
100
200 km
N


Probability of bloom occurrence
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
50
100
150
200
250
300
350
Elevation (m)


0.10
0.08
0.06
0.04
0.02
0.00
Probability of bloom
High agricultural
land cover
Median agricultural
land cover
Low agricultural
land cover
8.0
8.5
9.0
9.5
10.0
10.5
11.0
Mean annual temperature (°C)

Variable	Minimum	Median	Maximum
Elevation (m)	2.57	101.83	346.08
Annual temperature (°C)	8.04	9.33	10.97
Agricultural land	0.04	0.64	0.78
Woodland	0.01	0.05	0.08
Grassland	0.00	0.04	0.29

Variable	Coefficient	SE	Z value	Df	P value	Model presence
Agricultural land	**35.2**	**9.3**	**3.754**	**1**	**< 0.001**	**3/3**
Mean elevation	**2.9**	**0.9**	**3.011**	**1**	**< 0.001**	**3/3**
Mean annual temperature	**5.2**	**1.3**	**3.939**	**1**	**< 0.001**	**3/3**
Agricultural land × temperature	**-33.5**	**9.3**	**3.619**	**1**	**< 0.001**	**3/3**
Grassland	0.8	0.4	0.065	1	0.869	1/3
Woodland	0.9	0.3	0.023	1	0.931	1/3