

Weather and landscape morphology drive thermal regime variation among Mývatn ponds, and implications for resident Arctic charr

Grant E. Haines<sup>1</sup>

[grant@holar.is](mailto:grant@holar.is)

ORCID: [0000-0001-9085-0022](https://orcid.org/0000-0001-9085-0022)

Joseph S. Phillips<sup>2</sup>

ORCID: [0000-0003-2016-1306](https://orcid.org/0000-0003-2016-1306)

Bjarni K. Kristjánsson<sup>1</sup>

ORCID: [0000-0001-6984-5771](https://orcid.org/0000-0001-6984-5771)

Camille A.-L. Leblanc<sup>1</sup>

ORCID: [0000-0003-4861-9948](https://orcid.org/0000-0003-4861-9948)

1 - Hólar University, Dept. of Aquaculture and Fish Biology

2 - Creighton University, Dept. of Biology

## ABSTRACT

Thermal stratification is common in lentic freshwater systems, and has extensive effects on ecosystem function and the interactions between aquatic organisms and their surroundings. Although thermal regimes in temperate systems are well-characterized, small arctic and subarctic lakes and ponds can have irregular thermal regimes, and the conditions leading to stratification and mixing are less predictable. Nevertheless, they can be important habitats for northern fish species, and their ecological processes, including thermal stratification and dissolved oxygen dynamics, may determine whether they can support fish populations. Where such populations do persist, temperature and dissolved oxygen determine where and when fish are active, including foraging and spawning behavior, and consequently may shape trajectories of local adaptation. We studied differences in thermal stratification regimes in two adjacent ponds in a unique cave pond system in Iceland, where conditions shaped by thermal stratification may be acting as selective agents for phenotypic divergence of Arctic char. In particular, we characterize differences in thermal stratification regimes in relation to the orientation of cave openings and the highly irregular catchment topography, describe the association of the thermal regimes with dissolved oxygen dynamics, and describe potential consequences for the ponds' resident charr populations.

## INTRODUCTION

The Arctic charr (*Salvelinus alpinus*) is a culturally and economically significant species of salmonid with a holarctic distribution (Reist et al. 2013). Although it exhibits extremely high levels of intraspecific diversity and is able to occupy a large range of aquatic habitats, its habitat use is constrained by both temperature and dissolved oxygen concentrations (Elliott and Baroudy 1995; Kelly et al. 2020). By changing lake and pond thermal regimes, climate change is expected to dramatically alter the suitability of some habitats, forcing charr populations to migrate or adapt to these changes where migration is not possible (Lehtonen 1998; Kelly et al. 2020; Gillis et al. 2024). Understanding the current interactions between charr populations and the thermal regimes of their habitats is therefore critical for anticipating future adaptation and local extinction risk.

The mixing regimes of lakes and ponds are major sources of variation between lentic aquatic habitats with large effects on the ecological processes occurring within them. Stratification between water layers establishes gradients of not just temperature, but also dissolved oxygen (DO), salinity, CO<sub>2</sub>, CO<sub>4</sub>, and nutrients (Gibson 1999; Song et al. 2013; Rabaey and Cotner 2024). In most lakes wind drives mixing, and most large temperate lakes are dimictic, mixing in the spring and autumn (Boehrer and Schultze 2008), while some polar lakes may remain stratified throughout the year (Gibson 1999) or stratify along a chemocline generated by snow- and ice-melt (Cortés and MacIntyre 2020). Although some deep Icelandic lakes are dimictic, notably Þingvallavatn (Adalsteinsson et al. 1992) and Lagarfljót (Ramón et al. 2020), most Icelandic most lakes or ponds are much smaller and shallower. Small, shallow lakes often have more complex mixing regimes. Holgerson et al. (2022) found that temperate ponds in Ontario larger than 4 ha are usually isothermal with any stratification breaking down at night, but ponds with a smaller surface tend to mix intermittently if less than ~74 cm deep, or rarely or never mix if deeper than 74 cm. The extent to which these observations apply to ponds with complex or irregular basin topography, and in cooler subarctic regions is unclear. Nevertheless, such environments may be important drivers of ecological processes across a landscape, or inhabited by populations or species that are of conservation interest.

Temperature itself, along with some ecological parameters associated with thermal mixing regimes, particularly dissolved oxygen, affect fish physiology, development, and behavior (McDonald and McMahon 1977; Cote et al. 2020; Rodrigues et al. 2022), and in the long term can create selective pressures on freshwater fishes and may e.g., reinforce reproductive isolation between populations (Ohlberger et al. 2013; Kahilainen et al. 2014). Although Arctic charr are very tolerant of cold water, even to the extent that they can continue to grow at nearly freezing temperatures (Elliott and Baroudy 1995), their phenology and development is heavily influenced by temperature, DO, and their interaction. In hypoxic conditions, embryo weight and weight at hatching are reduced, and mortality of free-swimming embryos at relatively warmer temperatures dramatically increases (Gruber and Wieser 1983). Even in cold water, low DO can result in significant mortality of free-swimming embryos, and at 5°, Baroudy (1993) showed that 7-day survival was 100% at 75% oxygen saturation (9.6 mgL<sup>-1</sup>), but only 25% at 50% oxygen saturation (6.4 mgL<sup>-1</sup>).

We focus our investigation of thermal regimes on ponds that are situated within lava caves inhabited by small populations of small bodied Arctic charr in the Mývatn and Laxá Conservation Area in Northeast Iceland. Environmental microgeographic variation can therefore be expected not only to produce variation in ecological dynamics, but also to facilitate evolutionary divergence over very small spatial scales through selection on behavior, developmental, and physiological processes. Such small-scale divergence has been documented in

aquatic habitats spanning lotic-lentic gradients (Haenel et al. 2021), between reproductive habitats of different temperatures (Kavanagh et al. 2010; Arietta and Skelly 2021), and between distinct habitat patches within tidal marshes (Wagner et al. 2017). We also know that invertebrate communities in the cave ponds around Mývatn, on which the charr are dependent for food, are shaped by temperature and dissolved oxygen concentrations (Kristjánsson et al. 2024). Especially for metapopulations with small constituent populations vulnerable to local extinction, such adaptive variation is expected to be critical for enabling adaptation to climate change (Denney et al. 2020; Arietta and Skelly 2021).

To consider whether variation in thermal regimes between ponds could be an important driver of small-scale ecological variation and divergence in Arctic charr, which is widespread in Northern latitudes and exhibits particularly high diversity in phenotype (Klemetsen 2013), we analyzed depth profiles for two adjacent subarctic cave ponds. These ponds are partially subterranean and have highly irregular basin morphology, and we hypothesized that the deeper and more sheltered of these two ponds, C24, would stratify and the other, C23, would mix more regularly or continuously (see pond descriptions below). We then assessed meteorological variables for their potential contributions to the ponds' thermal mixing regimes. The relative contributions to pond stratification of localized landscape morphology and weather shared across the Mývatn lava fields is important for determining the potential heterogeneity of mixing regimes across the hundreds of ponds in this area, and determining whether temporal variation in thermal regimes is predictable. Finally, we tested for associations between changes in stratification intensity and dissolved oxygen, to determine if it is likely that stratification plays a role in modulating DO dynamics.

## METHODS

### *Site Description*

The ponds described here are located in a lava field directly northwest of the southern basin of Mývatn, a large shallow lake in Northeast Iceland. The lava field contains numerous caves formed by lava tubes, many of which are filled with groundwater that creates permanent ponds. Twenty ponds, in two locations west and northwest of the lake have been the subjects of a long-term monitoring project since 2012 because they contain small populations of small-bodied Arctic charr, described in Leblanc et al. (2024).

This study focuses on two cave ponds in particular, hereafter referred to as C23 and C24 (65.61722° N, 17.05426° W and 65.61718° N, 17.05376° W, respectively) which are located 14 m apart from one another. Despite their proximity, they contain Arctic charr populations with no significant gene flow from C24 to C23 and only low rates of gene flow from C23 to C24 (migration rate from C23 to C24 = 0.110 per generation,  $p = 0.04$ ; Judson et al. 2024;  $F_{st} = 0.093$ ; Leblanc et al. 2024). C23 has two openings separated by a low arch approximately 5m wide, with the passage between them oriented in an east-west direction, has a total of 2.2 m<sup>2</sup> of its surface open to the sky, and 5.8 m<sup>2</sup> of vertical area separating the surface of the water and the ceiling of the cave at the openings. C24 has one larger opening and an immediately adjacent smaller one, with a combined surface of 0.8m<sup>2</sup> exposed to the sky, and neither opening has a vertical area – meaning there is no space between the surface of the water and the ceiling of the cave chamber. Although we believe most of the pond area of C23 is visible, C24 likely includes a substantial area that is inaccessible without diving equipment. Both ponds have mostly rocky substrates with some fine organic mud in deeper areas, and lack macrophytes, but have attached algae and diatoms that serve as primary producers.

### *Temperature and Dissolved Oxygen Measurements*

We recorded temperature profiles by suspending HOBO data loggers (UA – 002-64 Onset Corporation, Bourne, MA) 20 cm apart along vertical transects from buoys at the deepest accessible points of C23 and C24. Because water levels were unusually high when temperature loggers were deployed due to late snowmelt, suspending them from anchored buoys instead of anchoring the loggers directly allowed them to maintain their depth relative to the surface, but likely resulted in some change in the distance from the benthos (which was <20 cm when deployed). Loggers recorded temperatures in the caves at 6-hour intervals from June 20, 2024 to September 4, 2024.

We obtained daily average air temperature, wind speed, and wind direction data for the Mývatn weather station (65.61933° N, -16.97684° W, 3.5 km east of the ponds) from Veðurstofa Íslands (the Icelandic Meteorological Office) for the period during which our water temperature loggers were deployed in C23 and C24. We also collected air temperature data at 6-hour intervals using a HOBO logger deployed at a site approximately 3.68 km south of the ponds (65.58413° N, -17.05642° W) to obtain data on diel fluctuations of air temperature and ensure that temperature data from different locations within the study area were consistent with each other.

Dissolved oxygen (DO) was recorded using a miniDOT logger (PME, Vista, CA) placed near the center of C23, which took measurements at 1-hour intervals from June 21, 2024 – September 18, 2024. This probe was positioned so that it was 18 cm above the bottom, and the temperatures it logged corresponded with the temperatures at 100 cm depth in our depth profile. The logger was removed for 10-30 min on August 21 so that the pond could be electrofished, and replaced at a slightly different location, and although there appears to have been little effect on the measured water temperature, the move of the probe and wading through the pond to electrofish may have influenced subsequent DO concentrations.

### *Limnological Calculations and Data Analysis*

All data processing and analysis was conducted in R (R Core Team 2023). Data from temperature profiles were synchronized to common six-hour intervals using linear interpolation (Gebhardt et al. 2024). We calculated the thermocline, when possible, using the *ts.thermo.depth* function in the R package *rLakeAnalyzer* (Winslow et al. 2019). We calculated Relative Thermal Resistance to Mixing (RTRM; Wetzel 2001; Song et al. 2013) – a metric of water density differential based on temperature – in both caves between the shallowest (20 cm) and deepest (100 cm in C23; 140 cm in C24) HOBO loggers. Because the chemical identities of solutes are unknown in the ponds, and because of the similarity of their water chemistry due to proximity and a common catchment, RTRM was calculated using densities for pure water.

We tested for the influence of daily average wind direction, wind speed, and air temperature on RTRM as a measure of the extent of stratification in C23, using generalized additive models (GAMs) as implemented in the R package *mgcv* (Wood 2011; Wood 2023). All tested models included main effects of air temperature and wind direction, and some also included wind speed and its interaction with wind direction. Because the variance of RTRM increases with its value, all GAMs used a Tweedie distribution – which is flexible enough to encompass both gamma and poisson distributions at differing parameter values – with a log link function. Additionally, because RTRM is not formally bounded by zero but is unstable at negative values, the absolute value of the minimum RTRM was added to all response values, so no values were negative and models could be fitted using the log link function. Because the thermal gradient is a response to previous conditions, models were run using both 1-day and 2-

day lags on independent variables. The smoothing splines for wind direction were implemented as cyclic cubic regression splines (Wood 2003, 2017).

While the non-seasonal models have the benefit of being able to predict RTRM using only a single day's meteorological data, their main drawback is that they cannot account for seasonal variation or temporal autocorrelation. To account for temporal variation, the best-fitting GAM model was rerun with a term for “day of the year”, fit with both a thin plate regression spline (TPRS) and an adaptive smooth (details in supplementary materials; Simpson 2018). The temporal and non-temporal models with the lowest AICs were then used to hindcast RTRMs of C23 from May 15 – September 30 (inclusive) from 2013 – 2024, or all but the first year of the Arctic charr monitoring program.

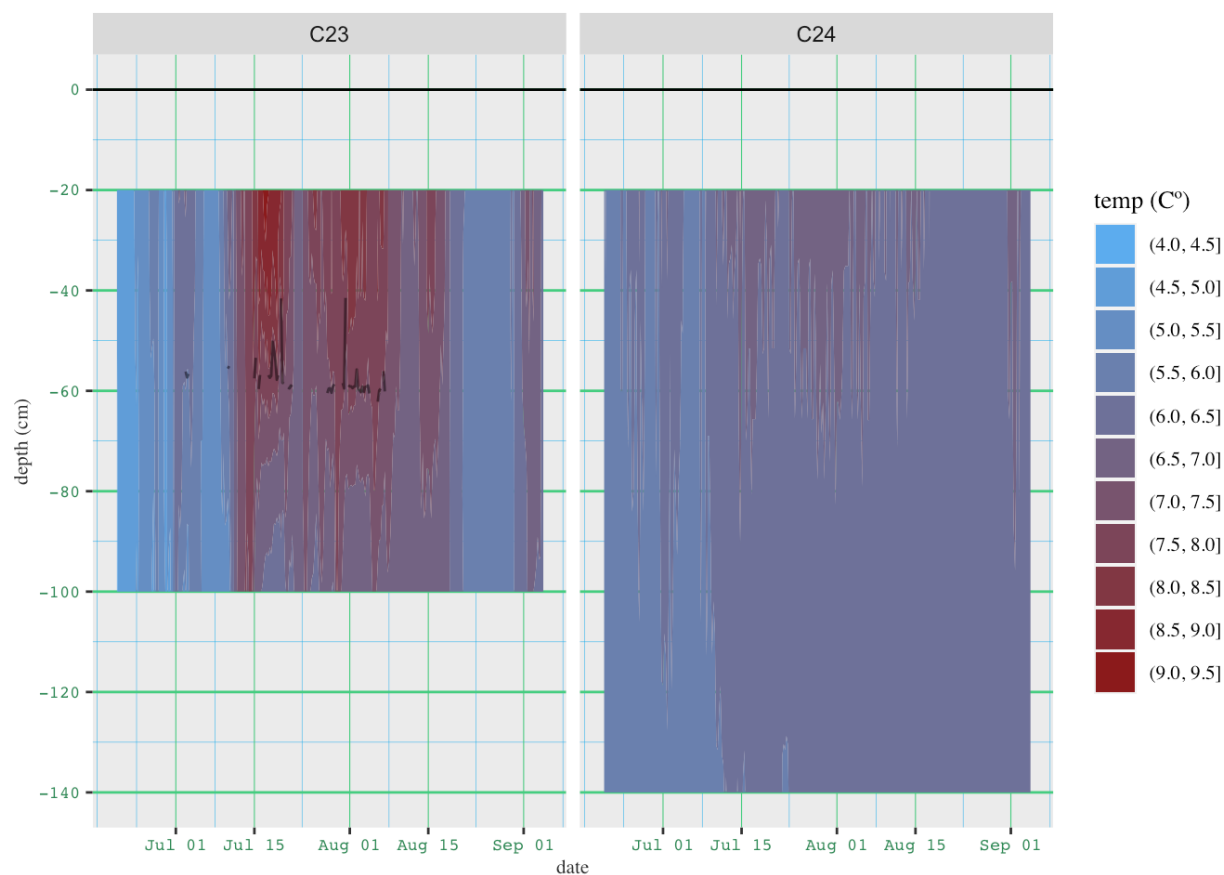


Figure 1.

Temperature profiles of both ponds throughout the study period. Temperatures in both panels are on the same color scale. The grey line in the C23 profile shows the calculated thermocline depth.

Temporal autocorrelation of model residuals was accounted for using a first-order autoregressive process using the *corARI* function in the nlme R package (Pinheiro et al. 2023), following procedures described in Simpson (2018) and Enevoldsen et al. (2022) that are further described in the supplementary materials.

Finally, although establishing a mechanistic model of DO variation in these ponds is outside the scope of this paper, we used a rolling correlation of DO and RTRM measured at 6-hr intervals to determine whether a mechanistic association between DO and thermal

stratification is likely. Because DO experienced dramatic diel fluctuations from mid-July through the 3<sup>rd</sup> week of August, DO and RTRM changes relative to the same time on the previous day were used to remove this effect. Correlations used a 3-day (12 time-point) lagging window. DO data following the August 21 removal and replacement of the DO probe are reported in supplementary materials (Figure S1), but were excluded from this analysis because of a subsequent abrupt decline in diel fluctuations that are assumed to be related to the disturbances associated with electrofishing or probe movement.

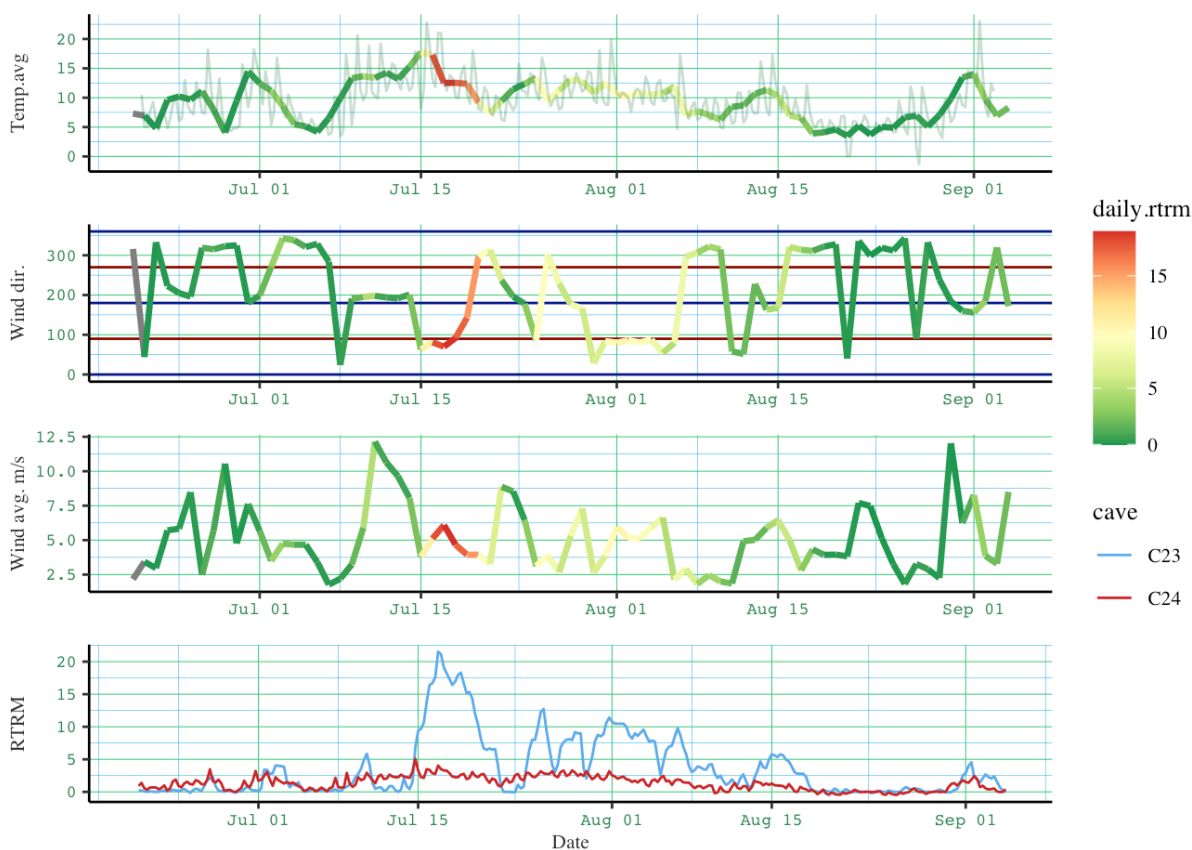


Figure 2.

Daily mean temperatures, wind direction, and wind speed at the Mývatn weather station, and RTRM at both caves throughout the study period. The pale green line in the first panel represents air temperature recorded from a HOBO logger at 6-hour intervals 5.28 km southwest of the weather station and 3.68 km south of the ponds to show that air temperature varies little across the lake, and therefore the weather conditions recorded at the weather station should be similar to those at the ponds. The diel fluctuations in the logger temperature data are a consequence of the 6-hour sampling resolution. The daily mean RTRM between the 20 cm and 100 cm depth loggers in C23 is illustrated in the first three panels by line color, RTRMs of C23 and C24 at 6-hour intervals are shown in the final panel.

## RESULTS

### *Thermal Gradient Formation*

Throughout the study period, C24 did not stratify, and temperature varied little during the study period (Figure 1, S2). The difference between temperatures recorded at 20 cm and 140 cm rarely differed by a maximum of 1.0°C, but rarely exceeded 0.5°C, irrespective of air temperature, indicating constant, thorough mixing. Loggers in C23, by contrast, recorded several distinct

periods of thermal stratification between mid-July and mid-August, the most prominent of which began on July 14 and lasted for approximately a week (Figure 2). C23 had a maximum temperature difference of  $3.0^{\circ}\text{C}$  between depths of 20 cm and 100 cm.

As might be expected for such a shallow pond, when a thermocline developed in C23, it was unstable and short-lived. After thermocline depth calculations of exactly 30 and 90 cm – which were assumed to be calculation artefacts – were removed, the most durable thermocline persisted for 4 days, from 18:00 on July 31 to 12:00 on Aug 4. Furthermore, as C23 became more stratified the temperatures recorded by the deeper loggers would decline, becoming colder than when the water column was thoroughly mixed, even as water temperatures near the surface increased (Figure S2). C24 did not form a thermocline throughout the study period.

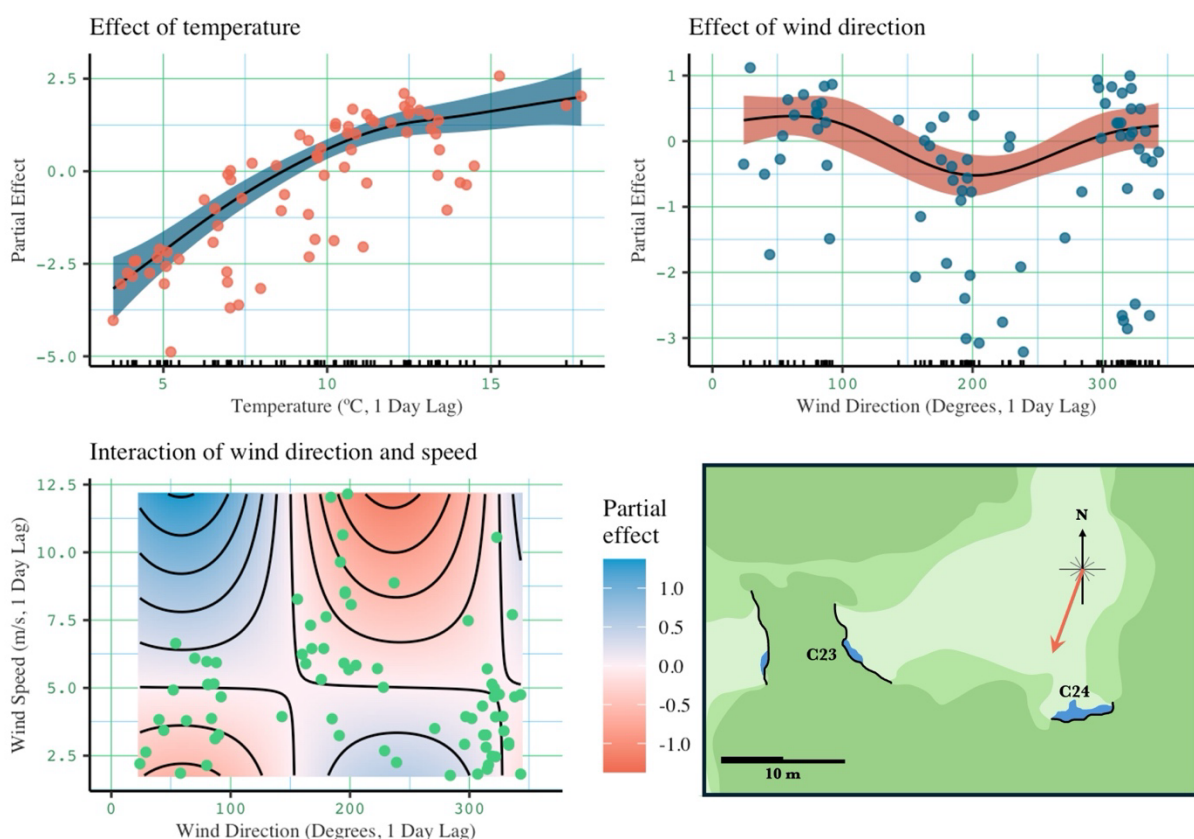


Figure 3. GAM model 2 partial effects on RTRM in C23. Negative partial effects result indicate increased mixing, while positive effects result in increased stratification. In the map, black lines indicate the cave openings, and shading indicates elevation and vegetation type. Pale green indicates grassy areas only slightly above the water table, while the darker greens indicate elevated areas with mostly low, shrublike vegetation. The orange arrow is oriented towards  $200^{\circ}\text{SSW}$ .

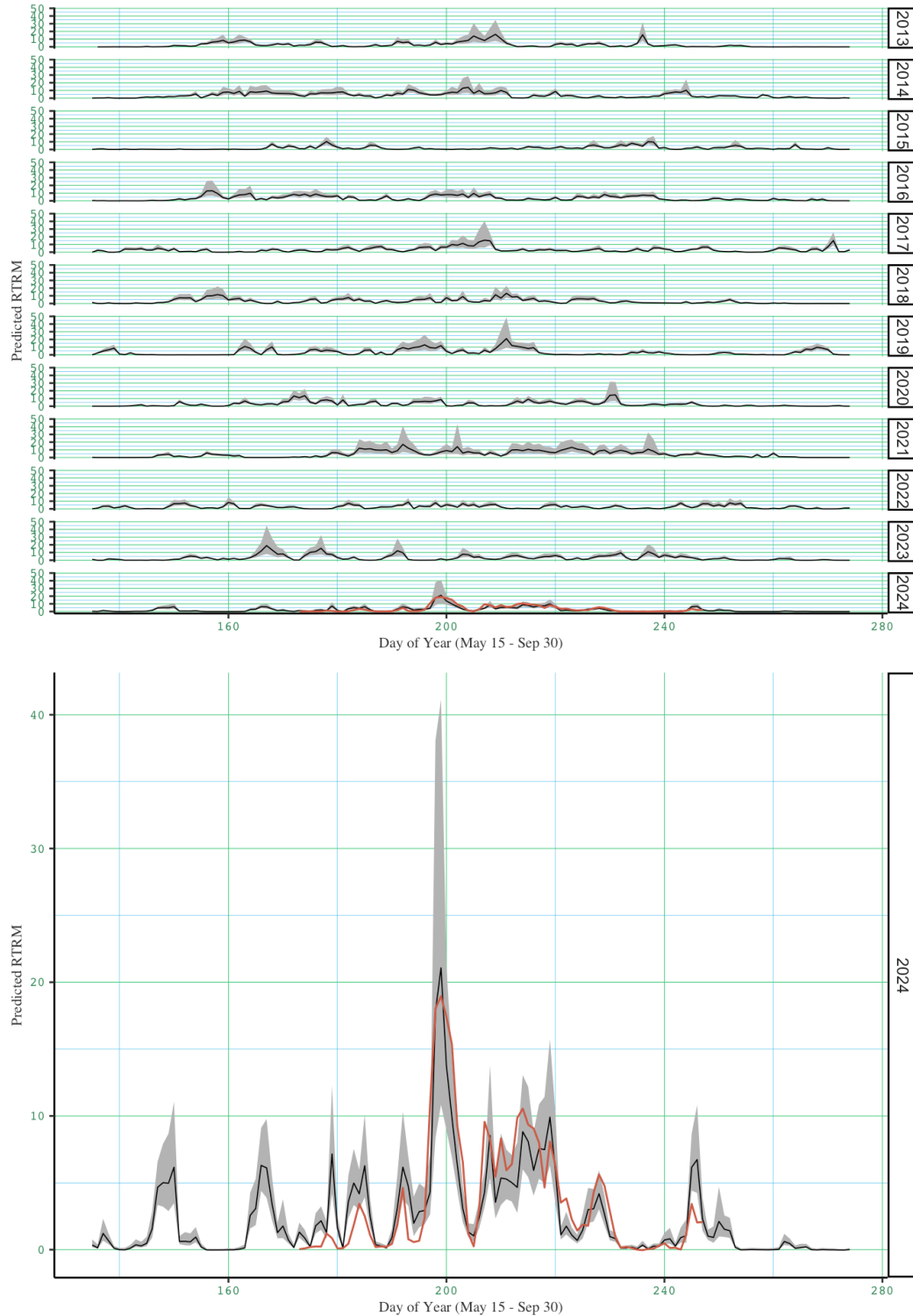


Figure 4. Daily mean RTRM empirically calculated (orange) and predicted RTRM from GAM Model 2 (black), with 95% credible interval shaded. This model does not contain a temporal term, so each point is predicted from only the previous day's meteorological data.



### *Stratification Models*

GAM models with a 1-day lag on meteorological variables yielded better fits than models with 2-day lags (Table 1). Model 2, the 1-day lagged model with an interaction between wind speed and direction, but without a main effect of wind speed, had the lowest AIC of tested models without a temporal component, so its results will be presented here (Table 2, Figure 3). Model 2 explained 69.3% of the deviance in RTRM, and all three model terms were significant. The negative effect of wind on RTRM was strongest when winds were blowing towards approximately 200° SSW. The effect of air temperature on RTRM increased steadily to approximately 12° C before the rate of increase in the effect of temperature slowed slightly. When wind speed and three-way interactions of temperature, wind direction, and wind speed were included in 1-day lagged models, the main effect of wind speed tended to be significant, while the interaction did not (Table S1, Figure S3).

Models incorporating a seasonal component (*day*), performed better than those without, explaining between 84.2% (Model 12) and 88.6% (Model 14) of deviance, and both adaptive spline models providing slightly better explanatory power than the models using TPRS for the *day* term. Unlike the non-seasonal models, in both Model 12 and Model 14, neither the main effect of wind direction nor the three-way interaction between temperature, wind speed, and wind direction were significant, but in Models 11 and 13, all effects were significant (Table 2, Figure S4). In Models 11-13, a strong seasonal effect of *day* was present, and peaked at the end of July or the beginning of August. Although a similar general pattern was observed in Model 14, this term was much more sensitive. After incorporating an autoregressive model component to account for residual autocorrelation, only the *temperature* term of Model 2-AR1 remained statistically significant, and only *temperature* and *day* remained significant in Model 11-AR1 (Table 2, Figures S5-6). However, the AIC of 2-AR1 remained higher than three of the four seasonal models excluding autoregression. Because the results of Models 11 and 13 were nearly identical, only Model 11 was fitted with an autoregressive component. Additionally, the autocorrelation coefficients used are conservative, possibly overestimating autocorrelation, and as a consequence the AR1 models should be interpreted with considerable caution (Figures S7-8, see Supplementary Methods).

Using Model 2 to model RTRM throughout the summer of 2024 (Figure 4) shows correspondence between measured and predicted values, albeit with a seasonal pattern in the difference between the two. The autoregressive version of this model predicts much lower RTRM in the middle of the summer, and higher RTRM in cooler weather than the empirical thermal gradients (Figure S9). Because Model 11 incorporates seasonality, its fitted values more closely match the measured thermal gradient at the beginning and end of the deployment period, but also predict complete mixing in the weeks before and after temperature logger retrieval (Figure S10). This is in contrast to Model 2 which predicts two periods with moderate gradients in May and June before logger deployment. This pattern in the differences between the seasonal and non-seasonal models' predictions is consistent for all years from 2013-2024, with Model 11 predicting very weak thermal gradients except in July and August, and Model 2 regularly predicting periods of at least moderate gradients well into May and September. Model 11-AR1 produces somewhat a less accurate model fit than Model 11, but this effect is less extreme than its non-seasonal counterpart (Figure S11).

### *Dissolved Oxygen*

For most of the study period, the change in strength of the thermal gradient, even when very weak, was negatively correlated with change in DO, sometimes very strongly (Figure 5).

However, this negative association weakened and briefly reversed following the period of strongest stratification in July, became reestablished for approximately a week, then again reversed during the weaker stratification at the beginning of August. The correlation coefficient between daily changes in DO and RTRM then gradually declined until it was once again strongly negative by the beginning of September.

Table 1. Model formulas and AIC fits for GAM models

Brief descriptions of all tested general additive models of pond thermal gradient for C23. Model terms are: T – temperature,  $W_d$  – wind direction in degrees,  $W_s$  – wind speed, D – day of the year. Model 2, the best-fitting non-seasonal model, was used as the basis for seasonal models 11 and 13. Model 4, the best-fitting non-seasonal model that incorporated a three way interaction between weather variables was used as the basis for seasonal models 12 and 14. Models 2, 11, and 13 are the focus of the remainder of results reported here. For these reasons, Models 2, 4, 11, and 13 are indicated in bold below. Full model results and a discussion of Model 14 are presented in the supplementary materials.

Model No.	Formula	Lag (days)	df	AIC
1	$\sim T + W_d$	1	11.37	273.87
<b>2</b>	<b><math>\sim T + W_d + W_d W_s</math></b>	<b>1</b>	<b>11.79</b>	<b>268.87</b>
3	$\sim T + W_d + W_s + W_d W_s$	1	13.21	269.44
<b>4</b>	<b><math>\sim T + W_d + W_s + T W_d W_s</math></b>	<b>1</b>	<b>9.89</b>	<b>270.23</b>
4.1	$\sim T + W_d + T W_d W_s$	1	9.86	271.81
5	$\sim T + W_d + W_s + W_d W_s + T W_d W_s$	1	16.23	271.62
6	$\sim T + W_d$	2	8.27	286.67
7	$\sim T + W_d + W_d W_s$	2	11.69	287.23
8	$\sim T + W_d + W_s + W_d W_s$	2	11.69	287.23
9	$\sim T + W_d + W_s + T W_d W_s$	2	9.28	283.88
9.1	$\sim T + W_d + T W_d W_s$	2	9.28	283.88
10	$\sim T + W_d + W_s + W_d W_s + T W_d W_s$	2	10.97	285.02
<b>11</b>	<b><math>\sim D + T + W_d + W_d W_s</math></b>	<b>1</b>	<b>15.89</b>	<b>224.96</b>
12	$\sim D + T + W_d + W_s + T W_d W_s$	1	16.48	229.33
<b>13</b>	<b><math>\sim D + T + W_d + W_d W_s</math></b>	<b>1</b>	<b>15.65</b>	<b>223.85</b>
14	$\sim D + T + W_d + W_s + T W_d W_s$	1	21.05	215.02

Table 2. GAM Model Results for C23

Summary results for models 2, 11, and 13. All GAMs were run using the Tweedie distribution with a log link function. Models 11 and 13 incorporate seasonal components, while model 2 requires only the weather data from the previous day. Thin plate regression splines are indicated by TPRS, and the Wind Speed term uses a TPRS spline with shrinkage to penalize large swings in responses. The wind direction spline is cyclic such that values of 0 and 360 are equivalent. Models 2 – AR1 and 11 – AR1 incorporate first-order autoregressive components with values of autocorrelation coefficient  $\rho$  reported (see supplementary materials for more on how this value was obtained). Reported REML scores for these two models are fREML scores.

Model	Terms	Spline/ Smooth Type	edf	F	p
-------	-------	------------------------	-----	---	---

2	T	TPRS	3.17	25.03	< 0.001	***
	W <sub>d</sub>	Cyclic	2.00	4.60	< 0.001	***
4	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	1.93	0.81	0.004	**
	T	TPRS (shrinkage)	2.64	22.52	< 0.001	***
	W <sub>d</sub>	Cyclic	2.29	3.77	0.004	**
	W <sub>s</sub>	TPRS (shrinkage)	0.86	1.11	0.022	*
11	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	0.00	0.00	0.576	
	D	TPRS	3.89	15.42	< 0.001	***
	T	TPRS	3.00	23.51	< 0.001	***
	W <sub>d</sub>	Cyclic	1.32	1.51	0.0243	*
	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.13	1.31	< 0.001	***
13	D	Adaptive	3.62	16.48	< 0.001	***
	T	TPRS	3.05	23.62	< 0.001	***
	W <sub>d</sub>	Cyclic	1.32	1.50	0.0244	*
	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.15	1.31	< 0.001	***
2 – AR1	T	TPRS	2.41	10.29	< 0.001	***
	W <sub>d</sub>	Cyclic	0.45	0.17	0.299	
	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	0.39	0.04	0.271	
11 – AR1	corAR1(~ D)	$\rho = 0.83072$				
	D	TPRS	3.33	6.74	< 0.001	***
	T	TPRS	2.37	13.46	< 0.001	***
	W <sub>d</sub>	Cyclic	0.00	0.00	0.013	*
	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	0.85	0.11	0.206	
	corAR1(~ D)	$\rho = 0.5926354$				
	Deviance Explained (%)	DF	AIC	-REML	Scale Estimate	
2	69.3	11.79	268.87	134.94	0.730	
11	84.9	15.89	224.96	114.70	0.435	
13	85.1	15.65	223.85	112.34	0.431	
2 – AR1	39.8	7.63	228.89	81.47	1.173	
11 – AR1	76.1	11.54	217.04	81.99	0.552	

## DISCUSSION

We showed that two cave ponds separated by only 14 m, and serving as habitats to separate populations of Arctic charr, experience divergent thermal regimes as a consequence of the interaction between weather and the irregular morphology and orientation of the caves themselves. Additionally, mixing in C23 appears to be associated with changes in DO, and stratification may reduce the relative role of physical processes – compared to biological processes – in governing oxygen dynamics. This suggests that variation in pond thermal regimes across this landscape plays an important role in shaping ecological processes at a very fine scale, and may be a direct and indirect agent of selection on the resident charr populations.

While the RTRM values calculated in C23 are lower than what is typical for lakes and ponds that are considered stratified because of their low year-round temperatures, both C23 and C24 are largely sheltered from most of the disturbances that would induce mixing in more typical ponds. Both have very small areas exposed to direct sunlight, meaning that radiative heating of the benthos is not a factor that would induce mixing, and while C23 has a larger surface exposed to the air, the air temperature in the chamber under the arch is often cooler than the outside air temperature, as is common among caves in the area. As a consequence, even relatively low RTRMs can resist mixing in C23, except when northeasterly winds are able to force air to flow under the arch and interact with the surface of the pond. Although C24 has even less exposure to wind, it also has a much smaller surface capable of interacting with the air. This likely prevented sufficient warming of its surface to induce stratification in the first place.

The lava fields around Mývatn include hundreds of ponds both within and outside of lava tube caves, which are inhabited by numerous small populations of charr (Leblanc et al. 2024) and stickleback (Seymour et al. 2013), with varying levels of physical connectedness and gene flow between ponds. While both of the ponds we studied here are very small, and support populations of only a few dozen adult charr (census  $N$ : C23 = 60, C24 = 42;  $N_e$ : C23 = 6, C24 = 22; Leblanc et al. 2024), the metapopulation they are a part of may consist of thousands or tens of thousands of individuals. Additionally, many of the caves are irregularly shaped, and often have several distinct basins and openings with varying orientations and degrees of exposure to fluctuations in air temperature and changes in wind speed and orientation. As a consequence, the morphological variation among caves may drive ecological dynamics in this system in much more complex ways than the methods employed in the present study would allow us to predict.

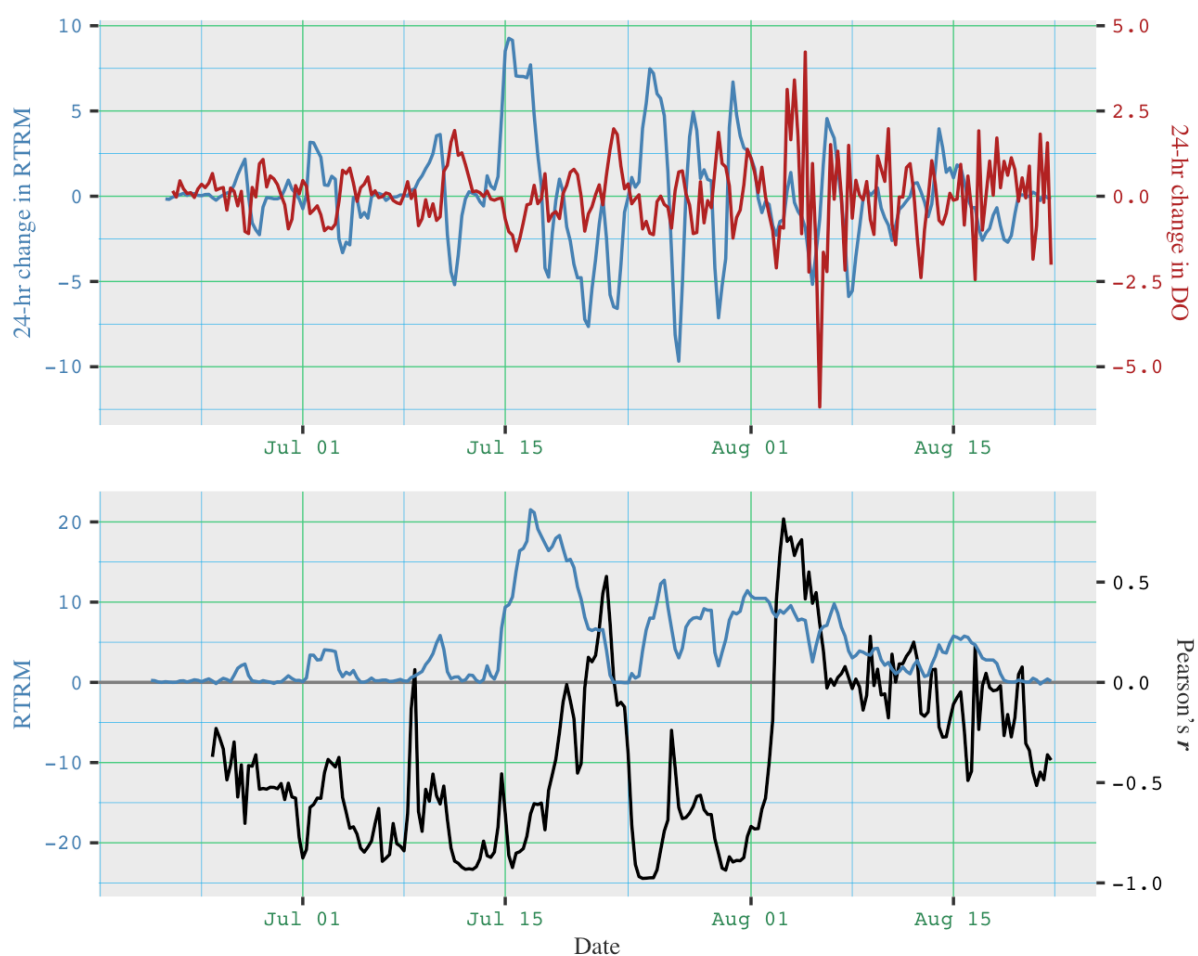


Figure 5.

The association between daily change in DO and thermal gradient intensity, calculated in 6-hr increments. The top panel shows 24-hr change in RTRM (blue) and DO (red) throughout the temperature profile logger deployment. The bottom panel shows a rolling correlation of daily change in DO and thermal gradient intensity (Pearson's  $r$ ; black), with a lagging 12-point, 36-hr window.

### *Limitations*

There are numerous caves of varying morphology across the lava fields around Mývatn, and this study only included two with relatively simple morphology compared to some caves in the area with multiple basins, openings, and cave branches. It is therefore unclear how prevalent thermal stratification is in ponds across this landscape, and how important its effects are to the area's ecology. Additionally, some ponds include multiple basins that are deep enough to stratify, and cave openings that are vertically, rather than laterally-facing, meaning less direct interaction with the climatic conditions that induce both stratification and mixing.

We have also not directly considered groundwater temperatures or flows in our models of stratification. The flow of groundwater in an approximately south-westerly direction, towards the lake, is likely responsible for maintaining cool water temperatures in the ponds during summer, and preventing them from freezing from top to bottom in the winter, but the rate of flow and baseline temperature of the groundwater is not known. The seasonal effect of groundwater temperature on stratification is likely captured by the *day* term in our temporal GAMs (see discussion of model 14 in supplementary methods and figure S12), but without direct measurements, we cannot confirm this. In addition, because large portions of these ponds are underground, some common limnological variables are not easily measurable, e.g. volume, so our models were designed with easily accessible meteorological data, but similar models for small ponds may be improved with the incorporation of such information.

Finally, because water is densest at 4° C, ponds can be warm- or cold-stratified, and RTRM has positive values during both warm- and cold-stratification. However, because our data were collected in the summer, only four days in our dataset had average temperatures < 4°, and none of those reached 3°. As a consequence, our models show only a positive effect of temperature on stratification, even though this relationship should reverse as temperatures cross and extend below the 4° line. This means that the current model fit should not be used to predict RTRM during cooler times of the year, although ongoing data collection should make this possible in the future.

### *Conclusion*

While the ponds surrounding Mývatn are of particularly complex morphology due to the volcanic history of the region, many lakes and ponds across northern latitudes have extremely complex and irregular morphologies. As we have shown, this variation can result in wildly divergent pond mixing patterns, even over very short distances. As important as the ecological dynamics of this divergence are when considered over a large scale, they may also drive variation across northern aquatic communities and populations at a very local scale, and be an engine of adaptation in the face of climate change. Adaptive evolution over contemporary time-scales is a common feature of wild populations (Sanderson et al. 2021), and across Arctic and subarctic environments, lakes and ponds are expected to stratify more frequently and for longer durations as the climate warms (Woolway et al. 2021). The ranges of coldwater fish populations are expected to shift as a consequence (Gillis et al. 2024), forcing populations to either adapt, migrate northward, or face local extinction, and the ability to predict ecological processes is required to understand the shape these responses will take at the local scale.

### *Acknowledgements*

We wish to thank Alison Derry for providing comments on an early version of this manuscript. We also thank Kári H. Árnason, Anthony R. Ives, and Emily Adler for their support in the field, as well as Sjoerd Dankers and Bart van der Zande for field assistance. Rannís (grant 228720-051)

provided funding support for this project, and we would also like to thank Creighton University for providing funds to JSP for travel costs and field equipment. Meteorological data were obtained from Veðurstofa Íslands (the Icelandic Meteorological Office) on December 4, 2024.

#### References

- Adalsteinsson, H., P. M. Jónasson, and S. Rist. 1992. Physical characteristics of Thingvallavatn, Iceland. *Oikos* 64:121-135.
- Arietta, A. Z. A. and D. K. Skelly. 2021. Rapid microgeographic evolution in response to climate change. *Evolution* 75:2930-2943.
- Baroudy, E. 1993. Some factors affecting survival and distribution of Arctic charr (*Salvelinus alpinus* (L.)) in Windermere. Department of Biological Sciences. University of Lancaster, Lancaster, UK.
- Boehrer, B. and M. Schultze. 2008. Stratification of lakes. *Rev. Geophys.* 46:RG2005.
- Cortés, A. and S. MacIntyre. 2020. Mixing processes in small arctic lakes during spring. *Limnol. Oceanogr.* 65:260-288.
- Cote, D., B. Tibble, R. A. Curry, S. Peake, B. K. Adams, K. D. Clarke, and R. Perry. 2020. Seasonal and diel patterns in activity and habitat use by brook trout (*Salvelinus fontinalis*) in a small Newfoundland lake. *Environ. Biol. Fishes* 103:31-47.
- Denney, D. A., M. I. Jameel, J. B. Bemmels, M. E. Rochford, and J. T. Anderson. 2020. Small spaces, big impacts: contributions of micro-environmental variation to population persistence under climate change. *AoB PLANTS* 12:plaa005.
- Elliott, J. M. and E. Baroudy. 1995. The ecology of Arctic charr, *Salvelinus alpinus*, and brown trout, *Salmo trutta*, in Windermere (northwest England). *Nord. J. Freshwat. Res.* 71:33-48.
- Enevoldsen, J., G. L. Simpson, and S. T. Vistisen. 2022. Using generalized additive models to decompose time series and waveforms, and dissect heart–lung interaction physiology. *Journal of Clinical Monitoring and Computing* 37:165-177.
- Gebhardt, A., R. Bivand, and D. Sinclair. 2024. *interp: Interpolation Methods*.
- Gibson, J. A. E. 1999. The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica. *Antarct. Sci.* 11:175-192.
- Gillis, D. P., C. K. Minns, S. E. Campana, and B. J. Shuter. 2024. Major changes in fish thermal habitat diversity in Canada's Arctic lakes due to climate change. *Commun Earth Environ* 5:89.
- Gruber, K. and W. Wieser. 1983. Energetics of development of the Alpine charr, *Salvelinus alpinus*, in relation to temperature and oxygen. *J. Comp. Physiol., B* 149:485-493.
- Haenel, Q., K. B. Oke, T. G. Laurentino, A. P. Hendry, and D. Berner. 2021. Clinal genomic analysis reveals strong reproductive isolation across a steep habitat transition in stickleback fish. *Nature Communications* 12:4850.
- Holgerson, M. A., D. C. Richardson, J. Roith, L. E. Bortolotti, K. Finlay, D. J. Hornbach, K. Gurung, A. Ness, M. R. Andersen, S. Bansal, J. C. Finlay, J. A. Cianci-Gaskill, S. Hahn, B. D. Janke, C. McDonald, J. P. Mesman, R. L. North, C. O. Roberts, J. N. Sweetman, and J. R. Webb. 2022. Classifying mixing regimes in ponds and shallow lakes. *Water Resources Research* 58:e2022WR032522.
- Judson, B. J., B. K. Kristjánsson, C. A.-L. Leblanc, and M. M. Ferguson. 2024. The role of neutral and adaptive evolutionary processes on patterns of genetic diversity across small cave-dwelling populations of Icelandic Arctic charr (*Salvelinus alpinus*). *Ecol. Evol.* 14:e11363.
- Kahilainen, K. K., W. P. Patterson, E. Sonninen, C. Harrod, and M. Kiljunen. 2014. Adaptive radiation along a thermal gradient: Preliminary results of habitat use and respiration rate divergence among whitefish morphs. *PLoS ONE* 9:e112085.
- Kavanagh, K. D., T. O. Haugen, F. Gregersen, J. Jernevall, and L. A. Vøllestad. 2010. Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evol. Biol.* 10:350.
- Kelly, S., T. N. Moore, A. d. Eyto, M. Dillane, C. Goulon, J. Guillard, E. Lasne, P. McGinnity, R. Poole, I. J. Winfield, R. I. Woolway, and E. Jennings. 2020. Warming winters threaten peripheral Arctic charr populations of Europe. *Clim. Change* 163:599-618.
- Klemetsen, A. 2013. The most variable vertebrate on Earth. *J. Ichthyol./Vopr. Ikhtiol.* 53:781-791.

- Kristjánsson, B. K., D. Combot, A. Reilent, J. S. Phillips, and C. A.-L. Leblanc. 2024. Invertebrate diversity in groundwater-filled lava caves is influenced by both neutral- and niche-based processes. *Ecol. Evol.* 14:e11560.
- Leblanc, C. A., K. Räsänen, M. Morrissey, S. Skúlason, M. Ferguson, and B. K. Kristjánsson. 2024. Fine scale diversity in the lava: genetic and phenotypic diversity in small populations of Arctic charr *Salvelinus alpinus*. *BMC Ecology and Evolution* 24:45.
- Lehtonen, H. 1998. Does global warming threaten the existence of Arctic charr, *Salvelinus alpinus* (Salmonidae), in northern Finland? *Ital. J. Zool.* 65:471-474.
- McDonald, D. G. and B. R. McMahon. 1977. Respiratory development in Arctic charr *Salvelinus alpinus* under conditions of normoxia and chronic hypoxia. *Can. J. Zool.* 55:1461-1467.
- Ohlberger, J., Å. Brännström, and U. Dieckmann. 2013. Adaptive phenotypic diversification along a temperature-depth gradient. *The American Naturalist* 182:359-373.
- Pinheiro, J., D. Bates, and R Core Team. 2023. nlme: Linear and nonlinear mixed effects models.
- R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rabaey, J. S. and J. B. Cotner. 2024. The influence of mixing on seasonal carbon dioxide and methane fluxes in ponds. *Biogeochemistry* 167:1297-1314.
- Ramón, C. L., M. C. Priet-Mahéo, F. J. Rueda, and H. Andradóttir. 2020. Inflow dynamics in weakly stratified lakes subject to large Isopycnal displacements. *Water Resources Research* 56:e2019WR026578.
- Reist, J. D., M. Power, and J. B. Dempson. 2013. Arctic charr (*Salvelinus alpinus*): a case study of the importance of understanding biodiversity and taxonomic issues in northern fishes. *Biodiversity* 14:45-56.
- Rodrigues, T. H., A. J. Chapelsky, L. E. Hrenchuk, G. R. Mushet, L. J. Chapman, and P. J. Blanchfield. 2022. Behavioural responses of a cold-water benthivore to loss of oxythermal habitat. *Environ. Biol. Fishes* 105:1489-1507.
- Sanderson, S., M.-O. Beausoleil, R. E. O'Dea, Z. T. Wood, C. Correa, V. Frankel, L. D. Gorné, G. E. Haines, M. T. Kinnison, K. B. Oke, F. Pelletier, F. Pérez-Jvostov, W. D. Reyes-Corral, Y. Ritchot, F. Sorbara, K. M. Gotanda, and A. P. Hendry. 2021. The pace of modern life, revisited. *Mol. Ecol.* 31:1028-1043.
- Seymour, M., K. Räsänen, R. Holderegger, and B. K. Kristjánsson. 2013. Connectivity in a pond system influences migration and genetic structure in threespine stickleback. *Ecol. Evol.* 3:492-502.
- Simpson, G. L. 2018. Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution* 6:149.
- Song, K., M. A. Xenopoulos, J. M. Buttle, J. Marsalek, N. D. Wagner, F. R. Pick, and P. C. Frost. 2013. Thermal stratification patterns in urban ponds and their relationships with vertical nutrient gradients. *J. Environ. Manage.* 127:317-323.
- Wagner, D. N., T. Z. Baris, D. I. Dayan, X. Du, M. F. Oleksiak, and D. L. Crawford. 2017. Fine-scale genetic structure due to adaptive divergence among microhabitats. *Heredity* 118:594-604.
- Wetzel, R. G. 2001. Fate of Heat. Pp. 71-92. *Limnology: Lake and river ecosystems*. Academic Press, San Diego, Ca.
- Winslow, L., J. Read, R. Woolway, J. Brentrup, T. Leach, J. Zwart, S. Albers, and D. Collinge. 2019. rLakeAnalyzer: Lake physics tools.
- Wood, S. 2023. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation *in* S. Wood, ed, CRAN.
- Wood, S. N. 2003. Thin Plate Regression Splines. *Journal of the Royal Statistical Society: Series B (Methodological)* 65:95-114.
- Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Methodological)* 73:3-36.
- Wood, S. N. 2017. *Generalized Additive Models*. Chapman and Hall/CRC, New York.
- Woolway, R. I., S. Sharma, G. A. Weyhenmeyer, A. Debolskiy, M. Golub, D. Mercado-Bettín, M. Perroud, V. Stepanenko, Z. Tan, L. Grant, R. Ladwig, J. Mesman, T. N. Moore, T. Shatwell, I. Vanderkelen, J. A. Austin, C. L. DeGasperi, M. Dokulil, S. La Fuente, E. B. Mackay, S. G. Schladow, S. Watanabe, R. Marcé, D. C. Pierson, W. Thiery, and E. Jennings. 2021. Phenological shifts in lake stratification under climate change. *Nature Communications* 12:2318.

## Supplementary Tables

Supplementary Table 1. GAM Model Results for C23

Summary results for models except 2, 11, 13, 2 – AR1, and 11 – AR1, which are reported in Table 2. All GAMs were run using the Tweedie distribution with a log link function. None of the below models incorporate autoregressive components. As described in Table 1, models 1-5, 12, and 14 have a 1-day lag on all independent variables, while models 6-10 use a 2-day lag. The wind direction spline is cyclic such that values of 0 and 360 are equivalent.

Model	Terms	Spline/ Smooth Type	edf	F	p	
1	T	TPRS	3.39	22.03	<0.001	***
	W <sub>d</sub>	Cyclic	3.25	2.38	<0.001	***
3	T	TPRS	2.96	27.74	<0.001	***
	W <sub>d</sub>	Cyclic	1.63	1.93	0.022	*
	W <sub>s</sub>	TPRS (shrinkage)	0.95	2.12	0.001	**
4.1	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.36	0.67	0.012	*
	T	TPRS (shrinkage)	2.71	22.23	<0.001	***
	W <sub>d</sub>	Cyclic	2.50	6.61	<0.001	***
	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	0.44	0.01	0.301	
5	T	TPRS (shrinkage)	2.42	25.43	<0.001	***
	W <sub>d</sub>	Cyclic	1.46	1.55	0.031	*
	W <sub>s</sub>	TPRS (shrinkage)	0.94	2.09	<0.001	***
	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.57	0.88	0.004	**
6	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	2.18	0.06	0.200	
	T	TPRS	2.93	22.88	<0.001	***
7	W <sub>d</sub>	Cyclic	1.06	0.25	0.128	
	T	TPRS	2.92	24.90	<0.001	***
	W <sub>d</sub>	Cyclic	0.78	0.45	0.152	
8	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.44	0.37	0.095	
	T	TPRS	2.92	24.90	<0.001	***
	W <sub>d</sub>	Cyclic	0.78	0.45	0.152	
	W <sub>s</sub>	TPRS (shrinkage)	0.00	0.00	0.549	
9	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.44	0.37	0.095	
	T	TPRS (shrinkage)	2.23	28.49	<0.001	***
	W <sub>d</sub>	Cyclic	0.35	0.15	0.262	
	W <sub>s</sub>	TPRS (shrinkage)	3.2 x10 <sup>-5</sup>	0.00	0.501	
9.1	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	2.08	0.15	0.015	*
	T	TPRS (shrinkage)	2.23	28.49	<0.001	***
	W <sub>d</sub>	Cyclic	0.35	0.15	0.262	
10	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	2.08	0.15	0.014	*
	T	TPRS (shrinkage)	2.16	28.65	<0.001	***
	W <sub>d</sub>	Cyclic	0.36	0.15	0.252	
	W <sub>s</sub>	TPRS (shrinkage)	1.6 x10 <sup>-5</sup>	0.00	0.399	
12	W <sub>d</sub> W <sub>s</sub>	Cyclic:TPRS (shrinkage)	2.00	0.13	0.025	*
	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	1.14	0.14	0.196	
	D	TPRS	3.99	16.29	<0.001	***
	T	TPRS (shrinkage)	2.35	27.11	<0.001	***
14	W <sub>d</sub>	Cyclic	2.2x10 <sup>-4</sup>	0.00	0.365	
	W <sub>s</sub>	TPRS (shrinkage)	1.93	3.13	<0.001	***
	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	2.16	0.07	0.116	
	D	TPRS	9.90	10.71	<0.001	***
14	T	TPRS (shrinkage)	2.33	21.77	<0.001	***
	W <sub>d</sub>	Cyclic	2.1x10 <sup>-5</sup>	0.00	0.878	
	W <sub>s</sub>	TPRS (shrinkage)	0.94	2.10	0.002	**
	TW <sub>d</sub> W <sub>s</sub>	TPRS: Cyclic:TPRS (shrinkage)	1.31	0.03	0.244	



## Supplementary Figures

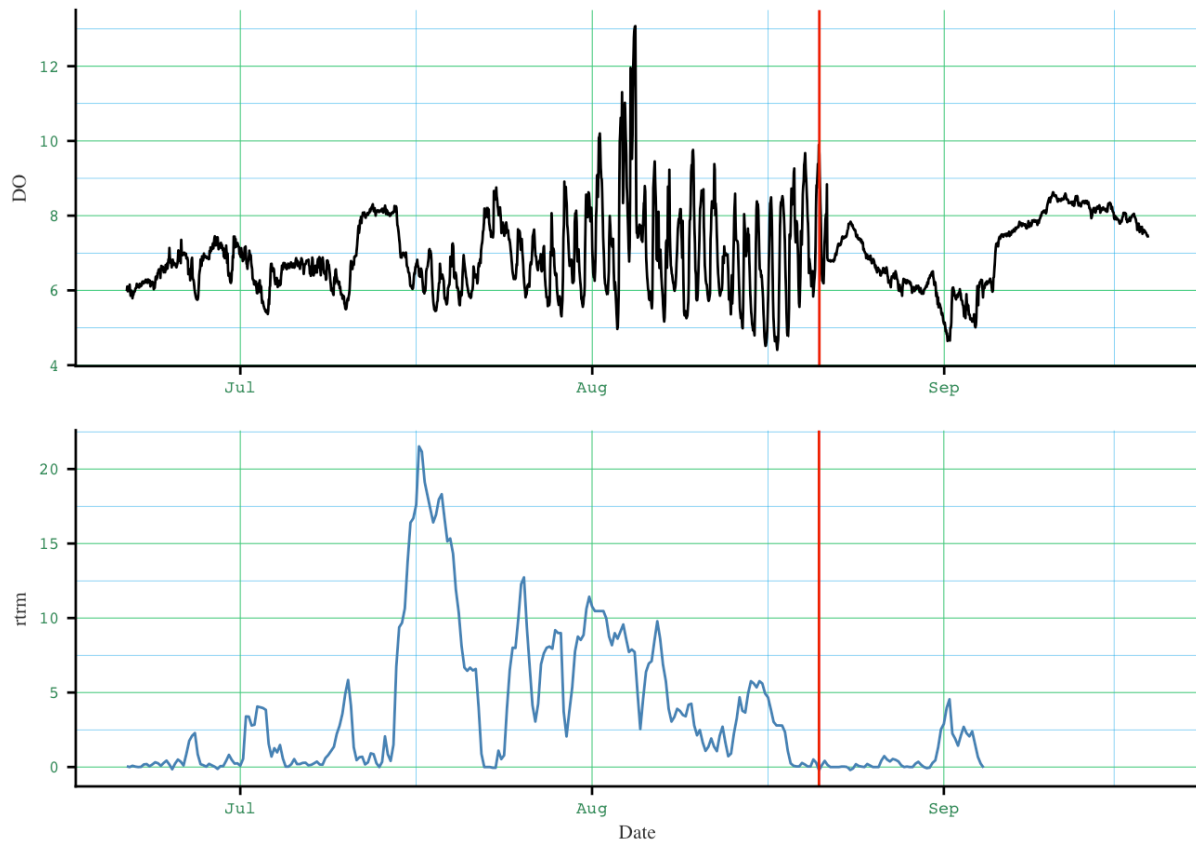


Figure S1. DO Concentrations (mgL<sup>-1</sup>) and RTRM in C23  
Full DO concentration recordings from MiniDOT logger, and RTRM. Vertical line indicates when the logger was removed and then replaced so the pond could be electrofished.

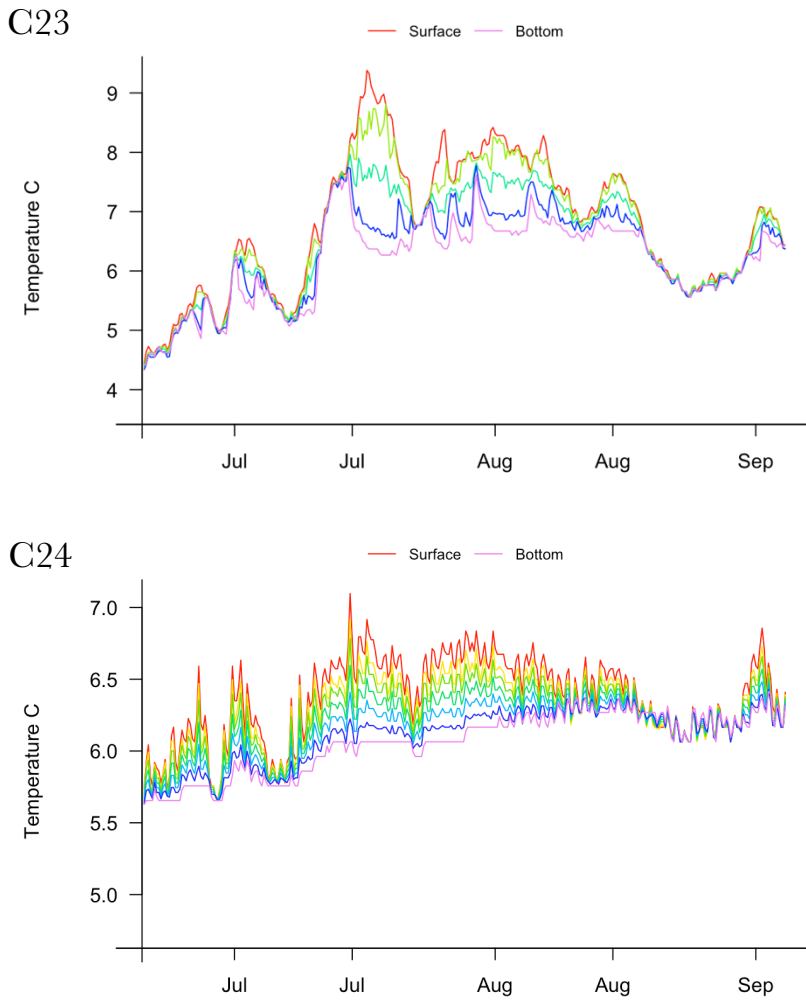


Figure S2. Thermal Profiles by Logger

Thermal profiles are shown with color gradient indicating logger depth, with loggers separated by 20 cm. The data presented here is the same as the data presented in Figure 1, but visualized in a way that emphasizes temperature differences between depths.

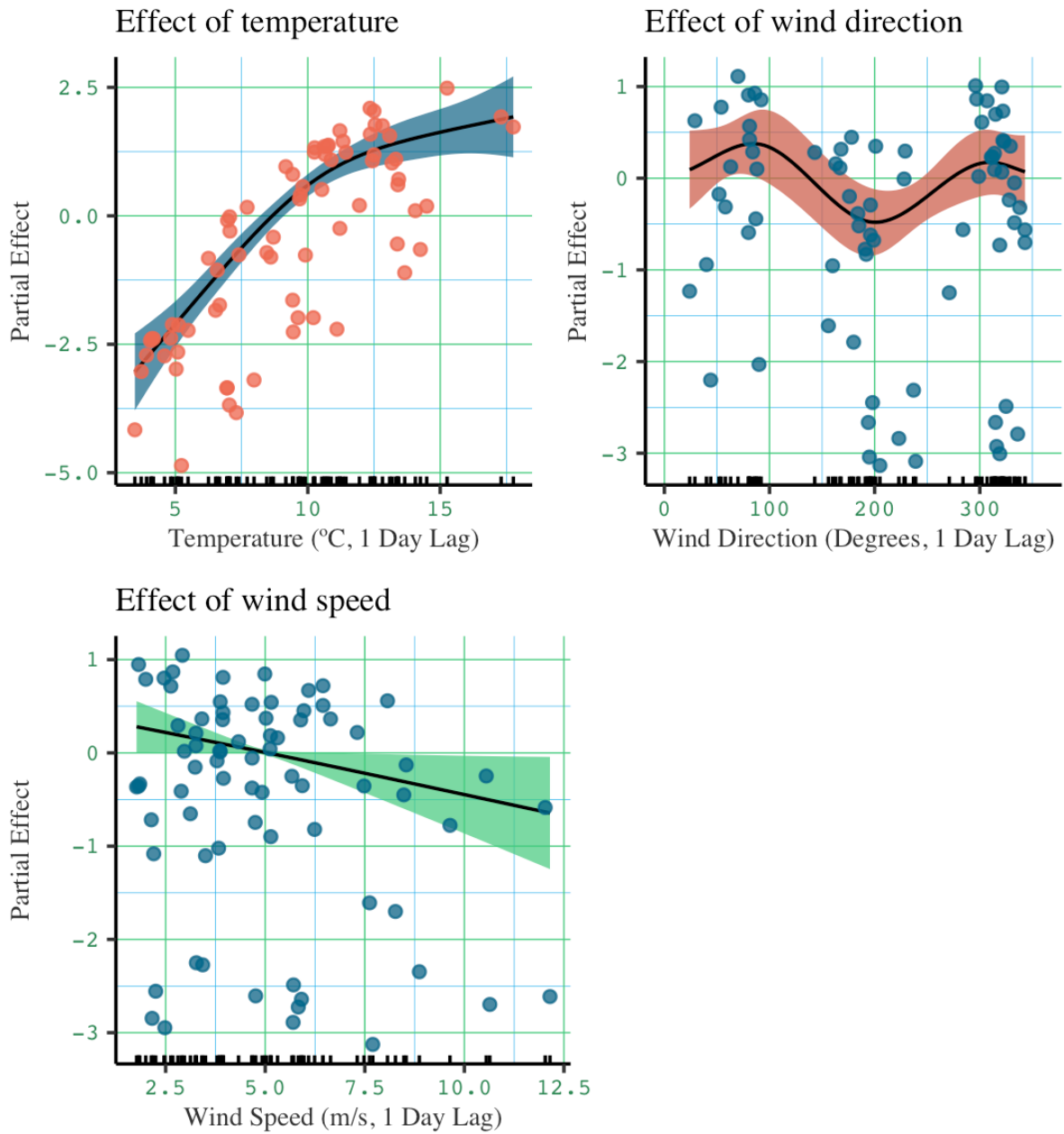


Figure S3. Partial Effects of GAM No.4

Partial effects of GAM model 4. The three-way interaction between temperature, wind direction and wind speed was not significant and is not shown.

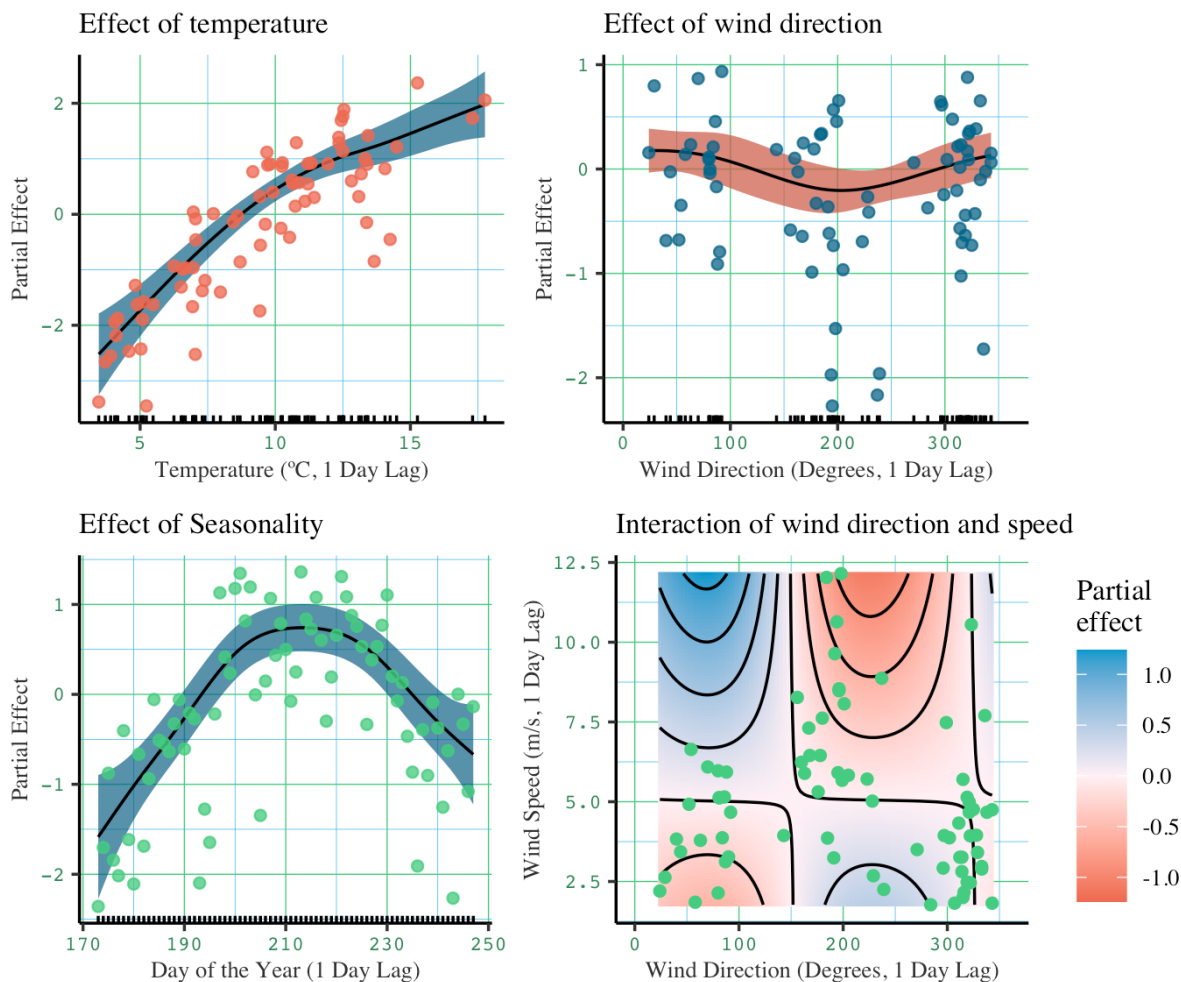


Figure S4. Partial Effects of GAM No.11

Partial effects of GAM model 11. All effects included in this model are shown, and are statistically significant. Model 13 is not shown because effects were nearly identical to those in model 11. Note that at higher wind speeds, winds blowing in a south-westerly direction had a negative, de-stratifying effect.

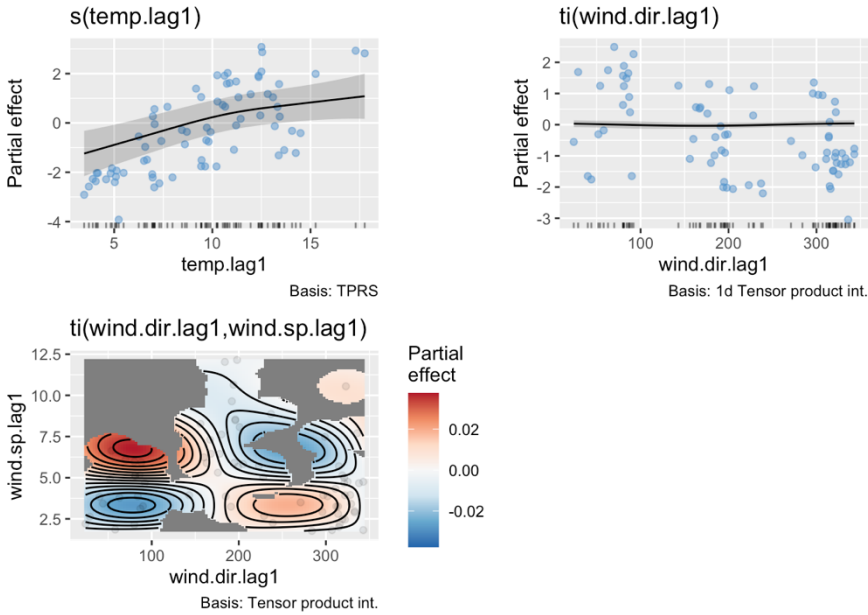


Figure S5. Partial Effects of GAM No.2-AR1

Partial effects of GAM model 2-AR1. Of these effects, only *temperature* is significant, when the AR1 correlation structure is included in the model.

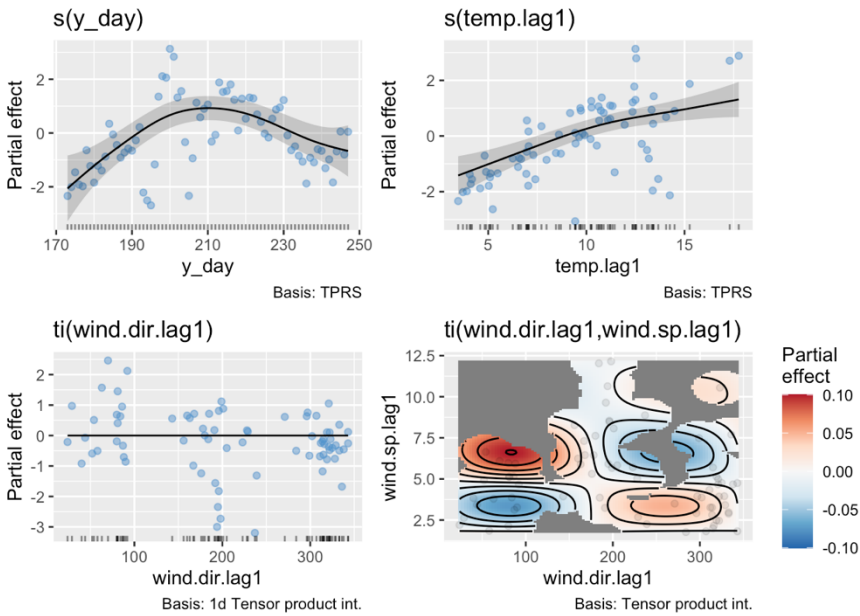


Figure S6. Partial Effects of GAM No.11-AR1

Partial effects of GAM model 11-AR1. Of these effects, only *day* and *temperature* are significant, when the AR1 correlation structure is included in the model.

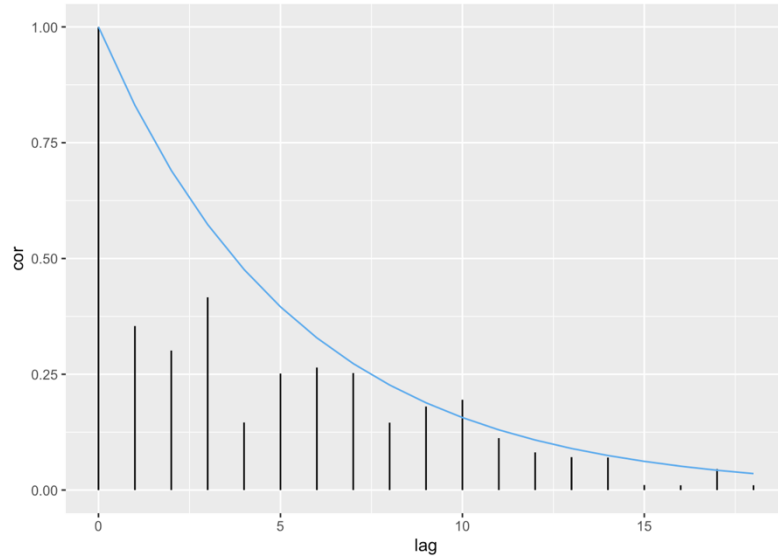


Figure S7. Model 2 Residual Autocorrelation

Vertical bars are an ACF plot of GAM model 2 residuals, showing clear autocorrelation. Blue curve shows exponential decay in temporal autocorrelation, using autocorrelation coefficient obtained by *gamm()* model with first-order autoregressive correlation structure, as described in supplementary methods.

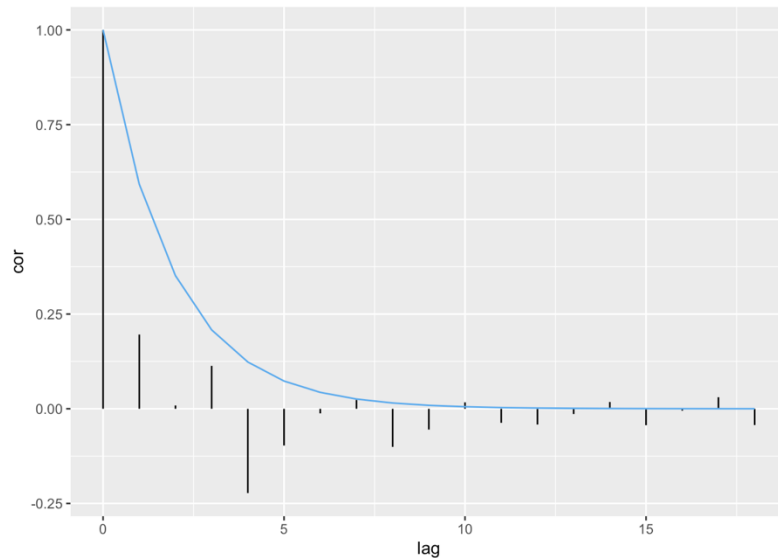


Figure S8. Model 11 Residual Autocorrelation

Vertical bars are an ACF plot of GAM model 11 residuals, showing clear autocorrelation. Blue curve shows exponential decay in temporal autocorrelation, using autocorrelation coefficient obtained by *gamm()* model with first-order autoregressive correlation structure, as described in supplementary methods.

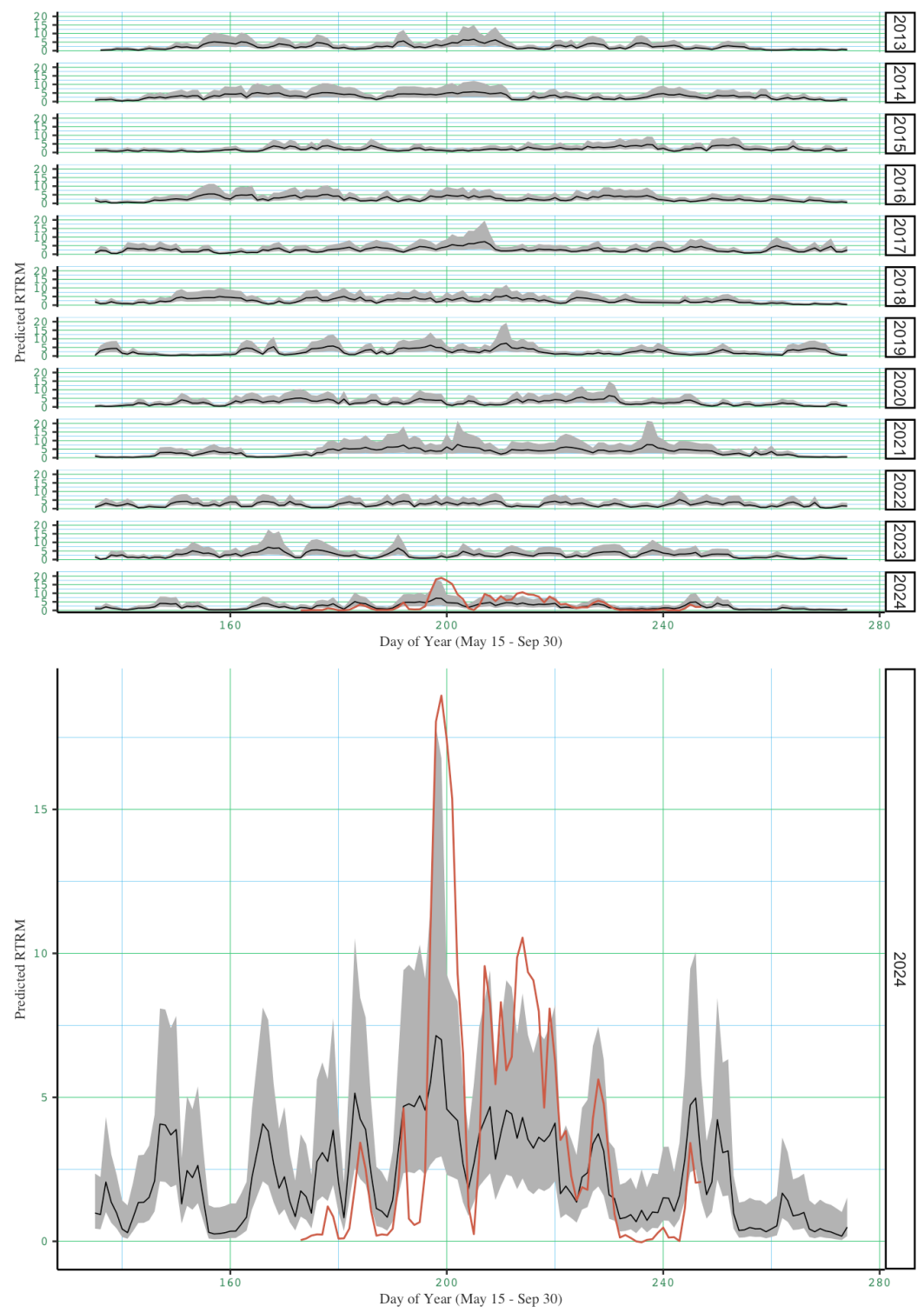


Figure S9. Hindcast Using Model 2-AR1 and Meteorological Data  
Figure shows Model 2-AR1 fitted values of RTRM using historical meteorological data for May 15-September 30 in 2013-2024 (top panel) and 2024 only (bottom panel). Orange lines in 2024 represent measured daily average RTRM values. Grey bands represent 95% credible interval

around fitted values. This model includes an AR1 correlation structure, and thus incorporates autoregressive effects.

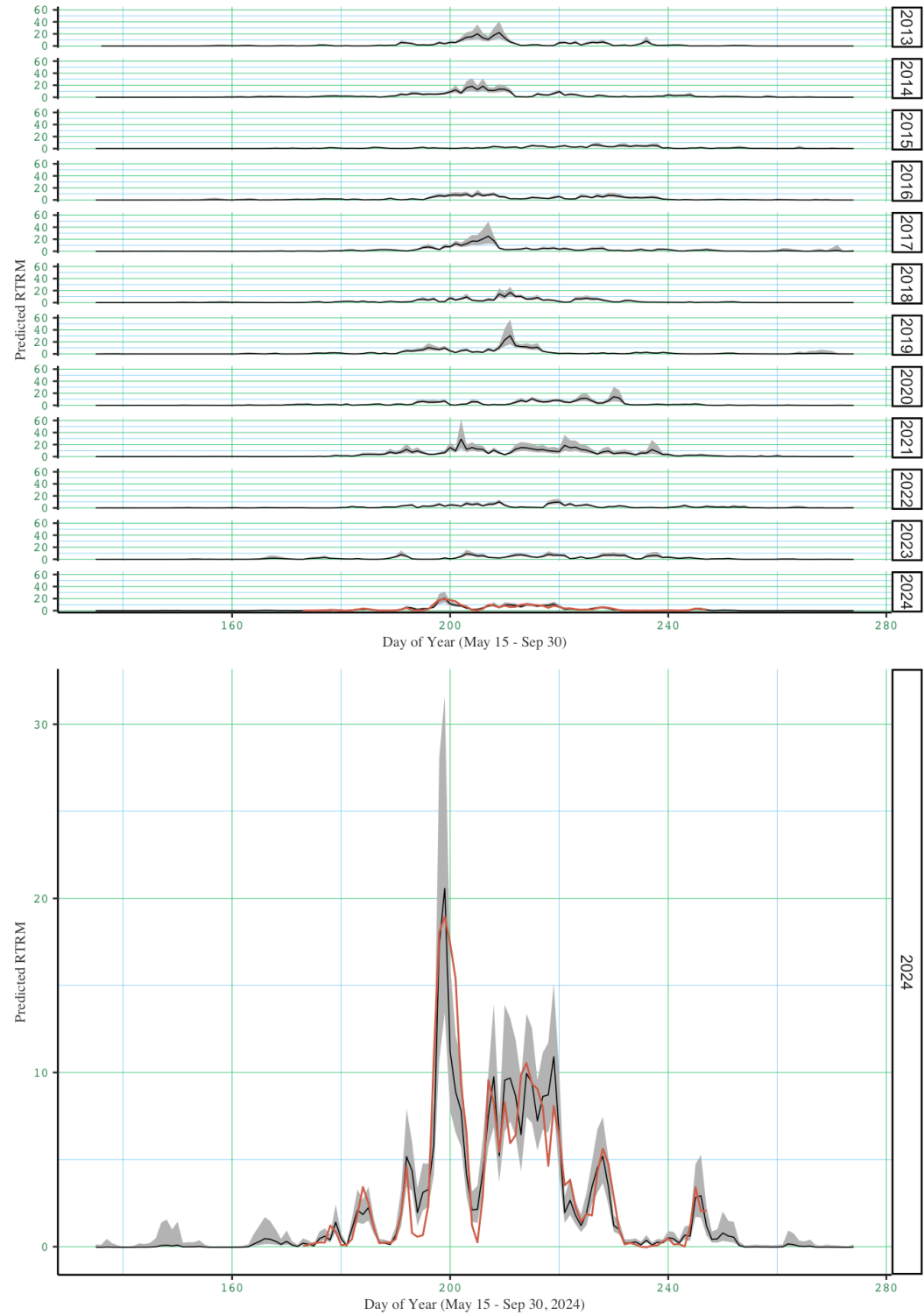


Figure S10. Hindcast Using Model 11 and Meteorological Data



Figure shows model 11 fitted values of RTRM using historical meteorological data for May 15-September 30 in 2013-2024 (top panel) and 2024 only (bottom panel). Orange lines in 2024 represent measured daily average RTRM values. Grey bands represent 95% credible interval around fitted values. This model includes a *day* term, and thus incorporates seasonal effects.

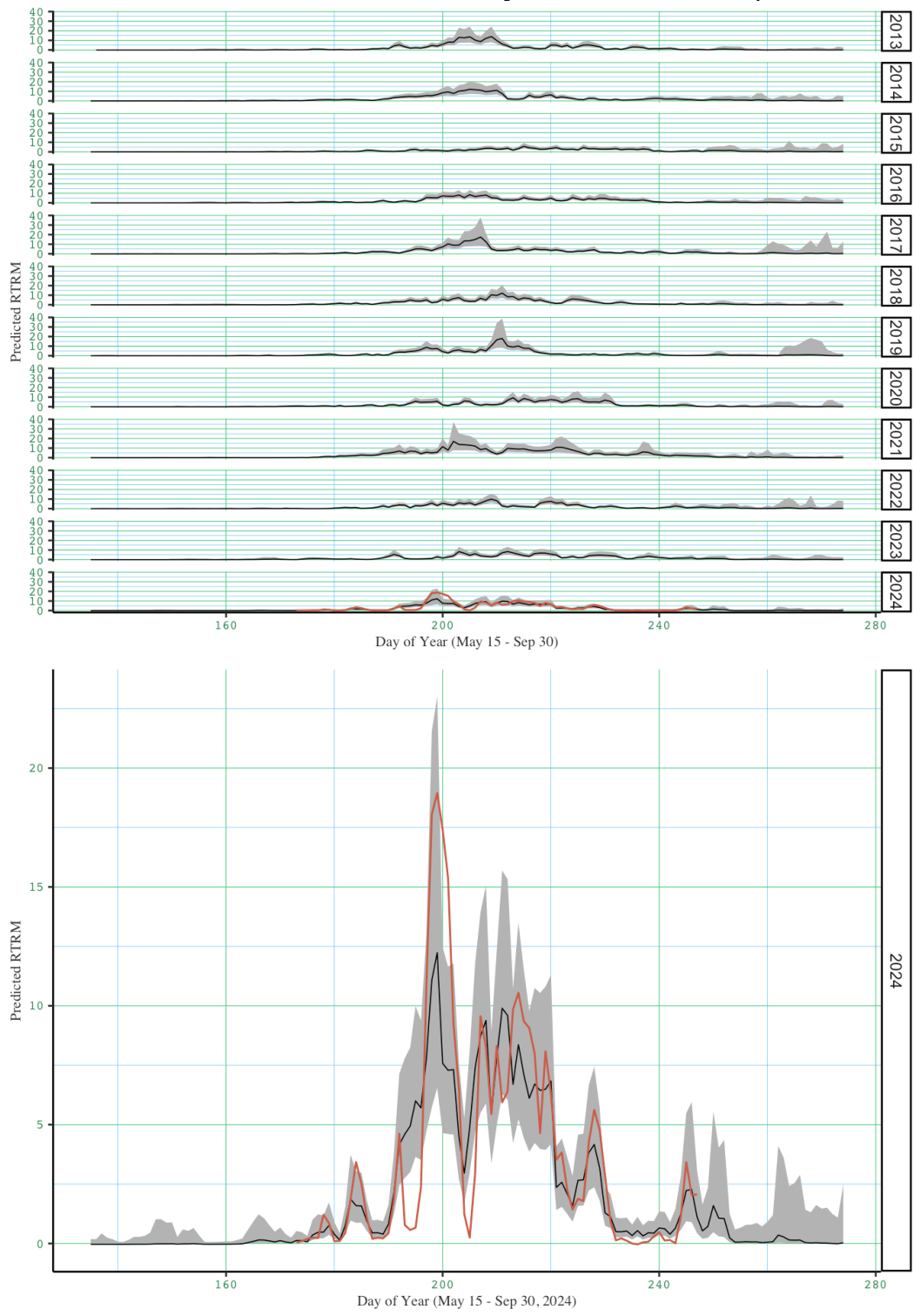


Figure S11. Hindcast Using Model 11-AR1 and Meteorological Data

Figure shows model 11–AR1 fitted values of RTRM using historical meteorological data for May 15–September 30 in 2013–2024 (top panel) and 2024 only (bottom panel). Orange lines in 2024 represent measured daily average RTRM values. Grey bands represent 95% credible interval around fitted values. This model includes a *day* term and AR1 correlation structure, and thus incorporates seasonal and autoregressive effects.

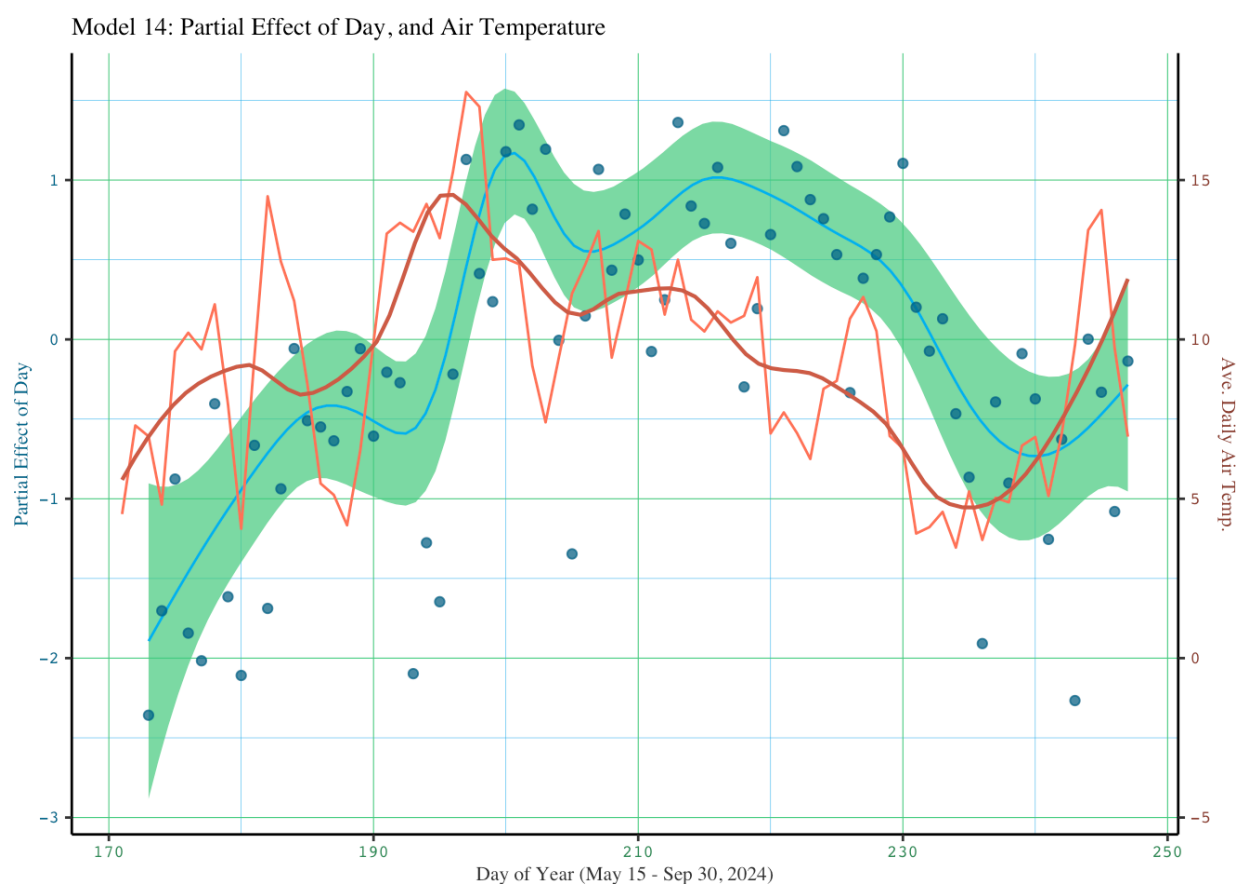


Figure S12. Model 14 Effect of *day* and Air Temperature

The partial effect of the day term in model 14 is shown in blue with a 95% CI and residuals. The average daily temperature recorded at the Mývatn weather station is shown in light orange with the darker orange curve showing a loess smooth with the  $\alpha$  value (span) set to 0.35. This shows a close association between air temperature and the effect of *day* with the *day* effect lagging 3–8 day behind temperature. Without further data to confirm this hypothesis, this could be understood as the *day* smooth absorbing the effect of groundwater temperature. Judging by AIC, this model fits the data better than every other model tested, despite not including an autoregressive correlation structure.

## Supplementary Methods

The appropriate distribution and link function for the GAMs used to model RTRM in C23 is somewhat difficult to determine because of the unusual characteristics of RTRM as a variable. Because the variance of RTRM increases with its value, it is heteroscedastic, but it is also not bounded by zero. Instead, RTRM values below 0 are possible, but unstable, because they indicate upper layers of water that are more dense than lower layers. This means that RTRM has a functional limit slightly below zero, and can dip into negative numbers on cool summer days (or more rarely on winter days when air temperatures are between 0° and 4°). We resolved this issue by modeling GAMs using Tweedie distributions with a log link function, and added the absolute value of the minimum RTRM to make all response values  $\geq 0$ . Where possible, the Tweedie distribution family was modeled using *mgcv*'s *tw()* function, which fits its own  $p$  parameter. However, while *tw()* is available for *gam()* and *bam()* models, it is not available in *gamm()*, which was used to determine autocorrelation coefficients as described below. In *gamm()* models we used *Tweedie()* for the model family and supplied the value for  $p$  returned in the output of non-autoregressive models.

Although estimates of autocorrelation coefficients can be obtained from autocorrelation functions (ACF) on model residuals, as in Enevoldsen et al. (2022), they can also be identified using general additive mixed models (GAMMs) with autoregressive correlation structures, as in Simpson (2018). In our models, the autocorrelation coefficients returned by GAMMs were substantially higher than those implied by ACFs (see Figures S7 and S8), presumably because the AR1 correlation structures absorbed some of the variance that was partitioned to the independent variables in non-autoregressive GAMs. We then used the correlation coefficients returned by the GAMMs as the input rho value in an analogous *bam()* model, which can use the more flexible *tw()* function to model the distribution family.

Because the autocorrelation coefficients are so high, the autoregressive models we report should be taken extremely cautiously, particularly because our independent variables are meteorological in nature, and therefore also temporally autocorrelated. Further, model 14, our best-fitting model (as judged by AIC), appears to suggest that the seasonal effects shown in models 11-13 may be detecting, but underfitting, what is actually an effect of groundwater temperature (Figure S12). An ACF plot of model 14's residuals show no significant autocorrelation, and fitting it as an autoregressive model is not possible due to the adaptive P-spline on the *day* effect.

### References

- Enevoldsen, J., G. L. Simpson, and S. T. Vistisen. 2022. Using generalized additive models to decompose time series and waveforms, and dissect heart–lung interaction physiology. *Journal of Clinical Monitoring and Computing* 37:165-177.
- Simpson, G. L. 2018. Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution* 6:149.