Где-то надо сказать, что в анализах мы не учитываем данные 1961 и 1962 года.

К математическим методам

Вся обработка материала осуществлялась с помощью функций языка статистического программирования R 3.5.3 (R Core Team, 2019).

*R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.*

Выявление фенологических событий в популяциях отдельных видов

Мы выделили четыре ключевых события, даты проявления которых было необходимо оценить, исходя из имеющихся наблюдений: начало пребывания вида в планктоне (Begin), середина временного промежутка, когда вид представлен в планктоне (Middle), дата пика численности вида (Peak) и дата окончания пребывания вида в планктоне (End).

Поскольку в нашем распоряжении были сезонные наблюдения, организованные по декадной схеме, то даты ключевых событий могли приходиться на промежутки между наблюлениями и, стало быть, пропущены. В связи с этим мы воспользовались следующим методом руконструкции дат ключевых событий.

В качестве маркеров пребывания вида в планктоне мы использовали обилие копеподитов, как наиболее обильной и, стало быть, репрезентативной стадии. (Для Calanus glacialis была рассмотрена суммарная численность копеподитов I, II и III). Обилие особей этих стадий рассматривалось, как отражение общего обилия популяции.

Для каждого вида в каждый из календарных годов была построена кумулята обилия -- числовой ряд, отражающий накопленное число особей, отмеченных в пробах на данный момент наблюдений. Эта кумулята была аппроксимирована с помощью логистической кривой, которая описывала зависимость кумуляты от количества Юлианских дней, прошедших от 1 января данного года (эта дата была взята за начало отсчета в каждый отдельный год).

Подбор параметров логистической модели производился методом наименьших квадратов с помощью функции nls() (Bates, Chambers, 1992) из пакета Stats (R Core Team, 2019).

*Bates, D. M. and Chambers, J. M. (1992) Nonlinear models. Chapter 10 of Statistical Models in Seds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.*

После подбора параметров модели, связывающей кумуляту и время, прошедшее с начала года, мы оценивали три величины:

1. Дата (количество Юлианских дней, прошедших с начала года), на которую приходилось 15% от значения асимптоты логистической кривой подобранной для данного вида в данном году. Это значение рассматривалось, как дата начала пребывания вида в планктоне (Begin).
2. Дата, на которую приходилась точка перегиба логистической кривой. Это значение рассматривалась, как характеристика середины пребывания вида в планктоне (Middle).
3. Дата, на которую приходилось 85% от значения асимптоты. Это значение рассматривалось, как характеристика окончания пребывания вида в планктоне (End).

В качестве даты пика численности вида (Peak) рассматривалась дата прямого наблюдения (без учета аппроксимирующей логистической кривой), когда была отмечена максимальная численность за весь период наблюдения в данном году.

Методика заполнения пропущенных значений

В 1963, 1972 и 1990 гг имеющая временная сеть наблюдений не позволила адекватно описать кумуляту для C.glacialis (очень краткое пребывание вида в планктоне пришлось на промежутки между наблюдениями). В этих случаях подобрать логистическую кривую не удалось. В связи с этим пропущенные значения ключевых событий были реконструированы с помощью сингулярного спектрального анализа временных рядов (singular spectral analysis of time series), предложенного в качестве инструмента заполнения пропусков во временных рядах (Golyadina, Osipov, 2007; Golyadina, Korobeynikov, 2013). Для данного анализа была применена функция gapfill() из пакета Rssa (Golyadina, Korobeynikov, 2014).

*N. Golyandina, E. Osipov (2007): The "Caterpillar"-SSA method for analysis of time series with missing values. Journal of Statistical Planning and Inference, Vol. 137, No. 8, Pp 2642–2653*

*Golyandina, N., & Zhigljavsky, A. (2013). Singular Spectrum Analysis for time series. Springer Science & Business Media.*

*Nina Golyandina and Anton Korobeynikov (2014) Basic Singular Spectrum Analysis and Forecasting with R. Computational Statistics and Data Analysis, Vol. 71, 934-954*

Аналогичный подход бы применен для заполнения пропущенных значений во временных рядах факторов среды (см. ниже).

Параметры среды

В качестве парамтеров среды, потенциально оказывающих вличние на даты фенологических событий были использованы следующие показатели:

- TPD -- Julian day when temperature maximum was observed

- SpSD -- Julian day when hydrological spring started (water temperature reached 3 Celsium degree)

- SuSD -- Julian day when hydrological summer started (water temperature reached 3 Celsium degree)

- SuFD -- Julian day when hydrological summer finnished (????????)

- ICD -- Julian day when ice clearing was observed

- SpT -- Mean spring temperature

- SuT -- Mean summer temperature

- NAO -- North Atlantic oscillation index

- AOI -- Arctic oscillation index ?????????????

Поскольку жизненный цикл изученных видов охватывает как минимум два года, то состояние популяции в данном году может зависеть от событий произошедших в предыдущий год. В связи с этим, в качестве отдельных факторов, вовлеченных в дальнейший анализ рассматривались значения показателей, отмченные в предыдущем году.

Анализ многолетней динамики изученных величин

Для выявления многолетних линейных трендов в динамике фенологических показателей видов и их обилия, а также факторов среды для каждого из них были подобраны линейные модели, связывающие значение той или иной величины со временем (годы наблюдения). Подбор параметров линейных моделей осуществлялся методом наименьших квадратов с помощью функции lm() из пакета stats (R Core Team, 2019).

Вместе с тем, из-за очевидного присутствия в данных временных автокорреляций, мы не использовали стандартные оценки статистической значимости параметров моделей, которые требуют независимости наблюдений. Вместо этого мы спользовали метод модельных матриц [‘‘model matrix’’ approach] (Clarke & Gorley, 2006; Legendre & Legendre, 2012). Для этого мы вычисляли матрицу попарных эвклидовых расстояний между годами, на основе временного ряда, представляющего многолетние изменения той или иной величины. Вторая, «градиентная» модельная матрица отражала попарные эвклидовы расстояния между числами натурального ряда от 1963 до 2018. Далее проводилось вычисление Мантеловской корреляции между этими двумя матрицами. Оценка статистической значимости осуществлялась пермутационным методом (здесь и в дальнейших случаях пермутаионных оценок значимости использовалось 9999 пермутаций). В дальнейшем мы использовали уровни значимости, полученные в данном анализе. Поскольку во всех случаях приходилось иметь дело со множественными гипотезами (несколько видов, несколько параметров среды), то все значения Мантеловских уровней значимости подвергались корректировке в соответствии с процедурой контроля FDR [False Discovery Rate] (Benjamini, Hochberg,1995)

*Clarke, K. R. & R. N. Gorley, 2006. PRIMER v6: User Manual/Tutorial. Primer-E Ltd., Plymouth.*

*Legendre, P. & L. Legendre, 2012. Numerical Ecology. Elsevier, Third English edition.*

*Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.*

Если скорректированный уровень значимости был ниже критического уровня (в данном анализе за критический уровень принималось значение p=0.05), то это рассматривалось как свидетельство присутствия в многолетней динамике данной величины некоторого направленного (восходящего или нисходящего) тренда. Направленность тренда оценивалась по значению углового коэффициента подобранной линейной модели.

Фенологическая таблица и ее анализ

Для выявления факторов, оказывающих влияние на фенологические события видов, был применен канонический корреспондентный анализ [canonical correspondece analysis, CCA] (Ter Braak, 1986; Legendre, Legendre, 2012).

*Ter Braak, C. J. F. (1986) Canonical Correspondence Analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67, 1167-1179.*

*Legendre, P. and Legendre, L. (2012) Numerical Ecology. 3rd English ed. Elsevier.*

В качестве зависимой матрицы в этом анализе была использована «фенологическая матрица», колонки котрой были сформированы ключевыми событиями каждого из видов (4 х 7 = 21 колонка), а строками - годы. В ячейках фенологической матрицы были приведены даты ключевых событий (количество Юлианских дней от начала года).

В качестве матрицы-пердиктора выступала матрица, в котрой строками были годы, а колонками были параметры среды (см. выше). Вместе с тем, поскольку фенологические события в жизни планктона могут регулироваться не только абиотическими параметрами среды, но также и биотическими взаимодействиями с другими членами планктонного сообщества, в матрицу-предиктор мы также включили и значения обилия видов (величины были подвергнуты логарифмированию). Анализ был проведен с помощью функции cca() из пакета vegan (Oksanen et al., 2019).

*Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner (2019). vegan: Community Ecology Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan*

После того как была подобрана полная модель, включающая все возможные переменные из матрицы-предиктора, с помощью функции ordistep() из пакета vegan, был осуществлен подбор оптимальной модели в соответствии с протоколом пошагового обратного отбора [stepwise backward selection].

Для оценки статистической значимости финальной модели в целом, отдельных ограниченных осей (constrained axis) и оставшихся в модели переменных-предикторов был применен пермутационный метод, реализованный в функции anova.cca() из пакета vegan. Статистически значимыми считались оценки при критическом уровне значимости p = 0.1.