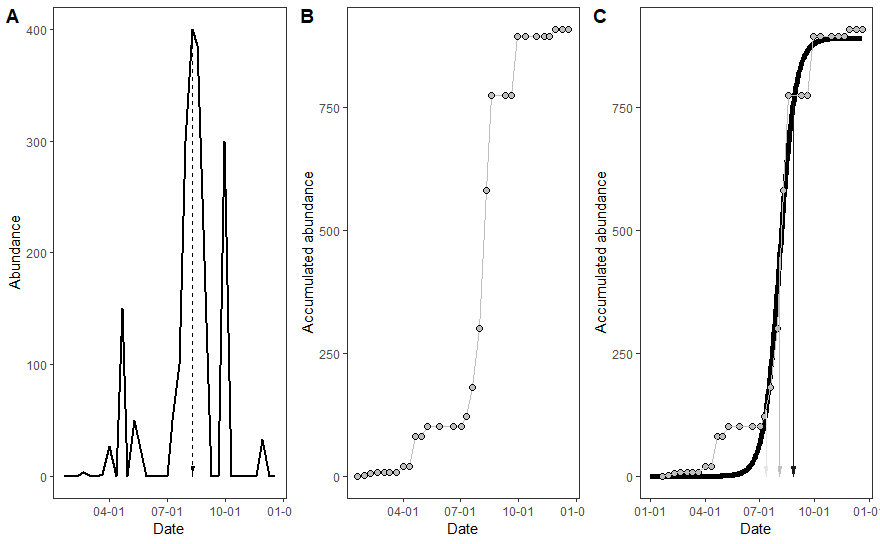
**Evaluation of validity of methods used**

**Goodness of fit of logistic regression used for phenological parameter evaluation**

The initial data, reflecting seasonal changes in abundance of particular species in each year (Fig SE1.1A) was converted into the the graph of cumulative curve (Fig SE1.1B). Then the date of cumulative curve was used for logistic regression fitting (Fig. SE1.1C). Dates of phenological events (Begin-of-season, Middl-of-season and End-of-season) was estimated from logistic regression (see the main text).



Figur SE1.1 Stages of determination of phenological parameters (Acartia spp. in 1964 was taken as an example). A. The graph reflecting the initial data, the seasonal changes in abundance. Dotted line represents the Peak of abundance. B. The cumulative curve. C. The logistic regression fitted. Vertical arrowed lines represents Begin-of-season , Middle-of-season and End-of-season (15%, 50% and 85% of asymptote of logistic curve respectively).

To assess the goodness of fit for logistic models we visually examined the the correspondence of observed cumulative curve and logistic regression line using the Residual Standard Error (RSE). Fig. ES1.2 illustrate the logistic lines with maximal RSE for each species (the worst approximation of cumulative curve by logistic regression). Fig. ES1.3 the same for models with minimal RSE (the best approximation). In all cases logistic curve approached close to initial data.

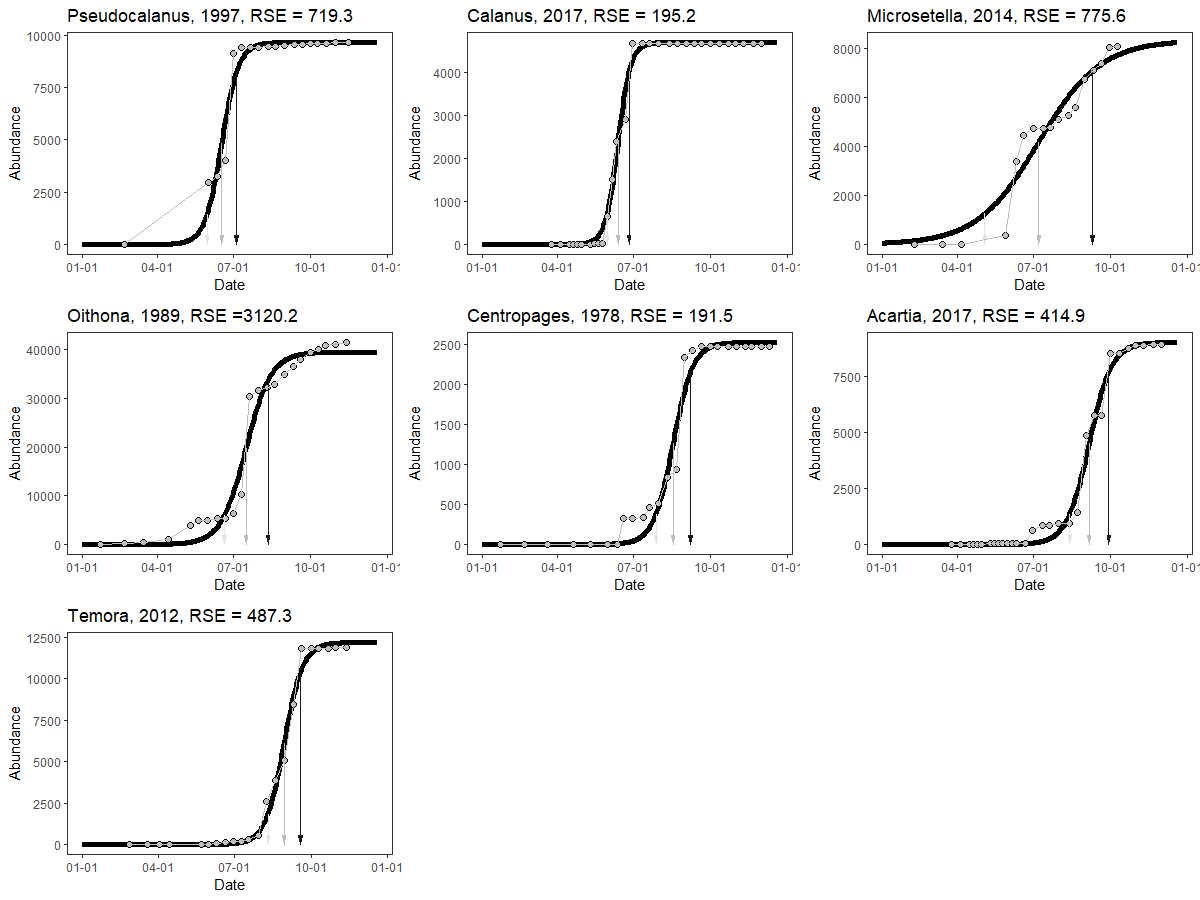


Fig. ES1.2. The abundance cumulate curve and fitted logistic curve for particular species in the case of worst approximation (maximal RSE).

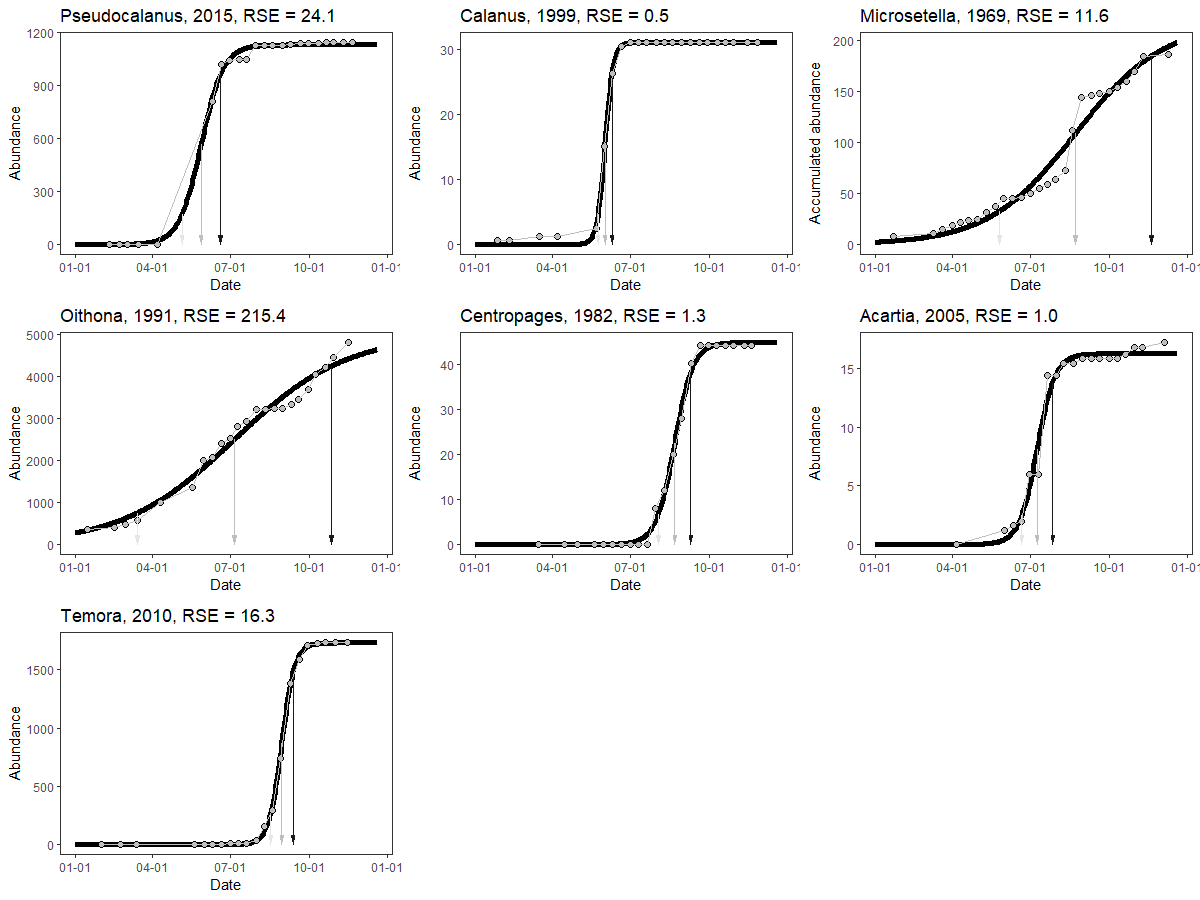


Fig. ES1.3. The abundance cumulate curve and fitted logistic curve for particular species in the case of best approximation (minimal RSE).

Additionally, to assess the quality of approximation, we measured the difference between Middle-of-season predicted by the logistic regression and date of Peak observed. The median of distribution of this value was close to zero for all species (Fig. ES1.4).

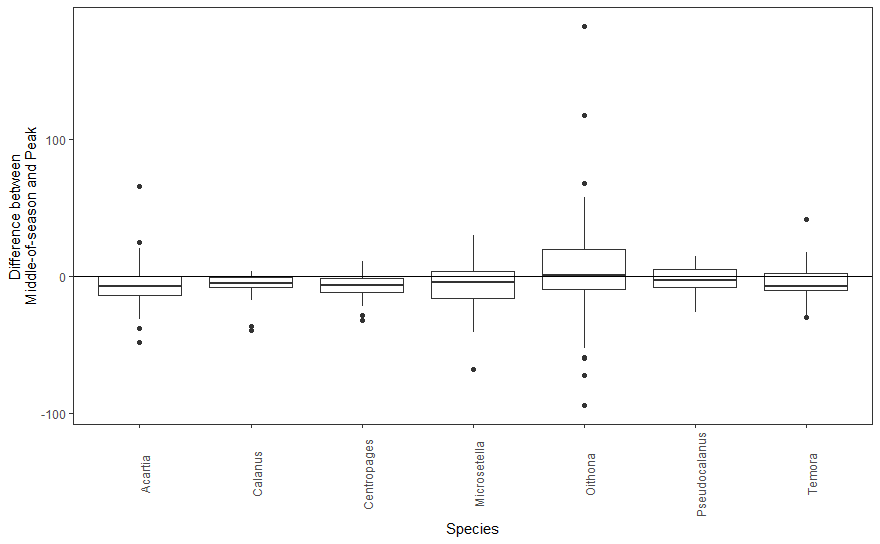


Fig. ES1.4. Boxplots representing the difference between Middle-of-season date and date of Peak. The median of all distributions (horizontal thick line inside the box) is close to zero.

**Gap filling using Singular Spectral Analysis (SSA)**

Для оценки возможного смещения оценок пропущенных значений, восстановленных с помощью SSA, мы использовали данные средовых параметров, где было пропущено 6 наблюдений (2% от общего количества наблюдений). Мы взяли матрицу с уже восстановленными данными, которая была использована в основном анализе. В этом массиве мы случайным образом выбрали 6 ячеек и обозначили их, как пропущенные значения. С помощью функции gapfill() из пакета Rssa (Golyadina et al., 2014) эти значения были восстановлены и была найдена разница между реконструированным и исходным значением. Такая симуляция была проведена 1000 раз. Распределение разностей реконструированных и восстановленных значений демонстрирует рисунок Fig. ES1.5. Медианы разностей для всех параметров оказались близки к нулю. Это свидетельствует об отсутствии систематического смещения оценок, связанных с применением метода SSA. Единственным параметром, в котором наблюдалось заметное смещение, было значение Summer Finish Day (SuFD), в среднем восстановленные значения были на 5 дней больше, чем исходные. Это единственный из параметров, в котором у нашем материале не было пропусков.

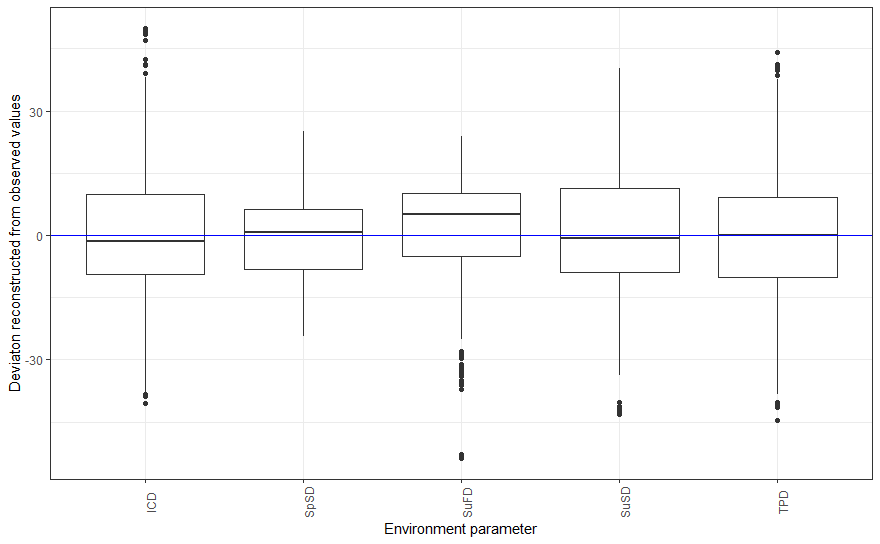


Fig. ES1.5 Boxplots representing the difference between values reconstructed by SSA and observed data in 1000 simulations. The median of all distributions (horizontal thick line inside the box) is close to zero.