Логика построения модели была следующей.

The logic linear model construction was as follows

1. Зависимая переменная. Поскольку зависимая переменная представлена долями (величина распределена в пределах от 0 до 1) обычную линейную регрессию, основанную на нормальном распределении остатков, строить было бы некорректно. В связи с этим мы использовали регрессионный анализ, основанный на бета-распределении с “logit” link-function.

The logic of model constructing was as follows.

Dependent variable. Since the dependent variable was represented by proportions (the value is distributed between 0 and 1) it would be incorrect to build an ordinary linear regression based on normal distribution of residuals. Therefore, we used a regression analysis based on a beta distribution with a "logit" link-function.

1. Природа предикторов. Целью нашего исследования было изучение пространственного варьирования переменной отклика. То есть? мы хотели оценить есть ли какой-то неслучайный паттерн в пространственном распределении доли левосторонних особей. Самый простой способ - это построить модель, связывающую зависимую переменную с географическими координатами (Lat и Lon). Однако эти две величины оказываются в случае имеющегося материала коллинеарны. Использование коллинеарных предикторов нарушает условия применимости множественных линейных моделей. Альтернативой для описания положения точки в пространстве может быть измерение расстояния от любой произвольно взятой точки. В качестве точки отсчета мы выбрали геометрический центр европейской популяции. Однако, само по себе расстояние не задает, однозначно, положение точки в пространстве. Поэтому в модель мы включили одну из координат (Lon). Variance inflation factor для моделей, включающей эти два предиктора, был меньше 2. Стало быть коллинеарности этих двух предикторов выявлено не было.

Nature of predictors. The goal of our study was to examine the spatial variation of the response variable. We wanted to assess whether there is some non-random pattern (gradient) in the spatial distribution of the proportion of left-sided morph. The easiest way is to build a model linking the dependent variable to geographic coordinates (Lat and Lon). However, these two coordinates would be collinear in the case of material available. The use of collinear predictors violates the conditions of applicability of multiple linear models. An alternative way to describe the position of any point in space is to measure the distance from given pount to any arbitrarily taken point chusen as point of reference. We chose the geometric center of the European population as the reference point. However, the distance from the reference point per se does not specify uniquely the position of a point in space. Therefore, we included one of the geographic coordinates (Lon) in the model. To note, Lon is a sort of distance from other reference object - zero meridian. The variance inflation factor for the models including Distance and Lon as predictors was less than 2. Therefore, no collinearity between these two predictors was detected.

1. Две модели вместо одной. Теоретически вместо двух моделей можно было бы построить одну, включив в качестве предиктора фактор «Subpopulation” с двумя градациями: “Arctic” and “European”. Однако, нетрудно заметить, что этот предиктор был бы коллинеарен, как с долготой, так и с расстоянием от центра популяции. Что опять же нарушает условия применимости множественной регрессии. Можно было бы построить nested model, но поскольку нас не интересовало сравнение двух субпопуляций между собой, то это, на наш взгляд, было бы избыточным.

Two models instead of one. Theoretically, instead of two models, we could construct a singular model by including the factor "Subpopulation" with two levels: "Arctic" and "European." However, it is easy to see that this predictor would be collinear with both Lon and Distance. Which again violates the assumptions of multiple regression. We could construct a nested model, but since we were not interested in comparing the two subpopulations with each other, this would, in our opinion, be redundant.

1. Проверка на пространственную автокорреляцию остатков. Мы не включали в модель никаких иных предикторов, кроме тех, которые задают само положение точек в пространстве. В такой ситуации применение каких-то сложных геостатистических моделей, включающих модлеирование пространственных автокорреляций, тоже неоправдано.

Check for spatial autocorrelation of residuals. We did not include in the model any predictors other than those that set the position of points in space itself. In this situation application of any complex geostatistical models including modelling of spatial autocorrelations is also unreasonable.