Appendix S2: Simulation methods and results using average temperature as an alternative biological predictor.

Methods

Simulation 1 (Accuracy of species distribution maps from ZI models) was repeated using an alternative biological predictor to altitude - average temperature in ºC across the study area between 1970-2000 obtained from WorldClim (WorldClim, accessed 10/05/18) at a 30-second resolution, and then converted to a 1km^2 resolution. Following the protocol of Simulation 1, a species with 5000 occurrence points was simulated across the study area based on the temperature layer converted to a probability layer using a logarithmic scale; the species was simulated to prefer higher average temperatures (Figure S2.1). The same bias predictor of distance to nearest town centre was used, and the simulation was again repeated 10 times for each set of town centres. All of the model structures in Tab. 3 were used. Model predictive power was assessed using 'deviation from the best model' ('*D*').

Results

Results from the alternative run of Simulation 1 with the species preferring high temperatures echo those using altitude, in that the count abundance predictions provide the most accurate estimates (according to the metric '*D*') of true species abundance (Figure S2.2). Again, the GLMs and the ZI sampling abundance predictions perform poorly in comparison and are unable to capture the effect of sampling bias or model successfully the excess zeros. Of the ZI models, all with the exception of ZI2 (where the bias predictor is omitted from the zero component but included in the count component), are able to provide good estimates of true species abundance. The zero component of the ZI models is again able effectively to identify and model the sampling bias (Figure S2.1). Although the correlations between distance from nearest town centre and average temperature are higher than for altitude (which is reflected in the zero component of the ZI6 models which include the bias predictor in this component), the ZI models are still able to produce accurate abundance maps using the count abundance predictions.

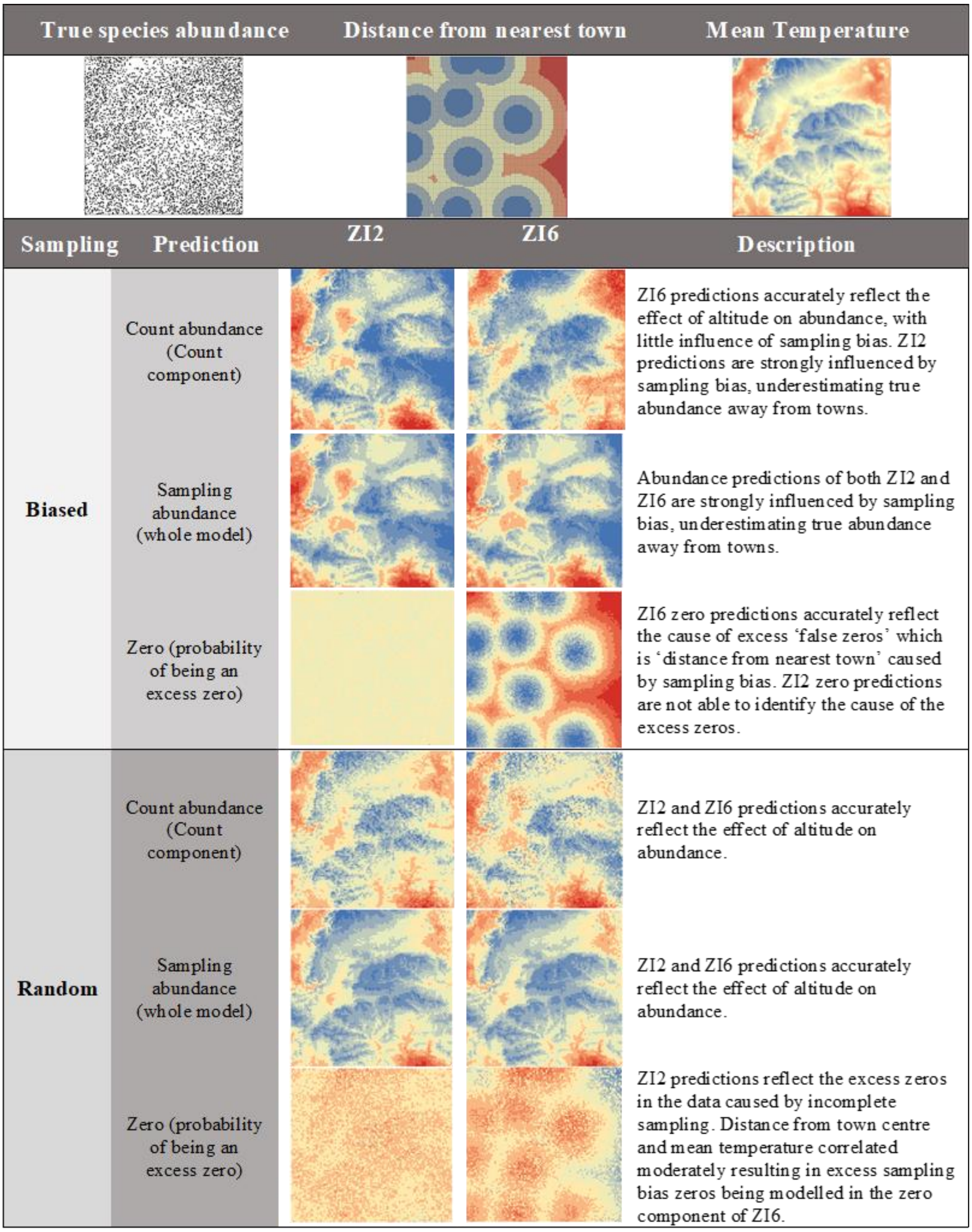

Figure S2.1. Example maps showing predicted abundance (count abundance and sampling abundance– see main text) and excess zeros (zero) for a hypothetical species whose occurrence is positively influenced by mean annual temperature, from two zero-inflated models (ZI2 and ZI6). Both models include a biological predictor (mean temperature) of both abundance and excess zeros, and a bias predictor (distance from the nearest town) as a predictor of abundance. ZI6 also includes distance from the nearest town as a predictor of excess zeros. Models were built with either data collected by randomly sampling grid cells (random) or with sampling bias (biased). Individual cells are colour coded based on abundance for the abundance predictions or on probability of being an excess zero for the zero predictions (high = red, low = blue).

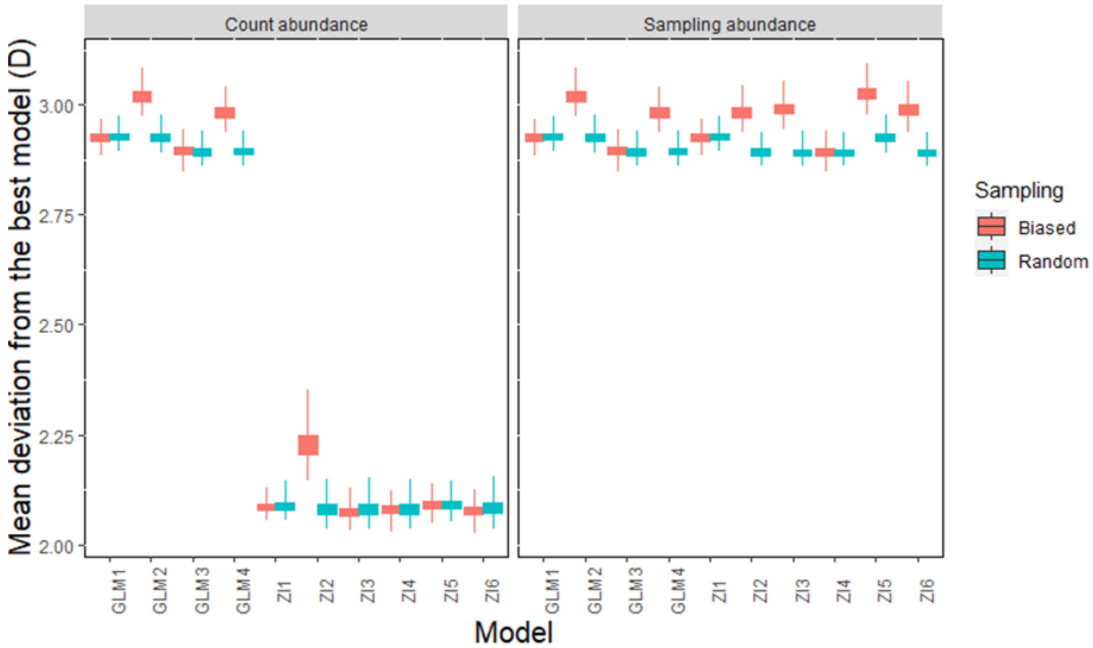

Figure S2.2. Evaluation of model predictions of abundance (based on D = 'deviation from the best model') for a hypothetical organism with a biological preference for warm temperatures. Mean D (± SE and data range) is shown for each sampling strategy (random or biased) across 10 different sets of hypothetical 'town centres' for each model. There are four non-zero-inflated generalised linear models, and six zero-inflated (ZI) models. For explanations of the structure of each model, see Tab. 3. Two types of prediction were evaluated: the count abundance predictions from the count component of the ZI models and the sampling abundance predictions from the whole of the ZI models or from the GLMs.