Dynamic Generalised Additive Models (DGAM) for forecasting discrete ecological time series

Nicholas J Clark [1] and Konstans Wells [2]

[1] School of Veterinary Science, The University of Queensland, Gatton QLD 4343 Australia

[2] Department of Biosciences, Swansea University, Swansea SA2 8PP, UK

Appendix S4: Supplementary Figures

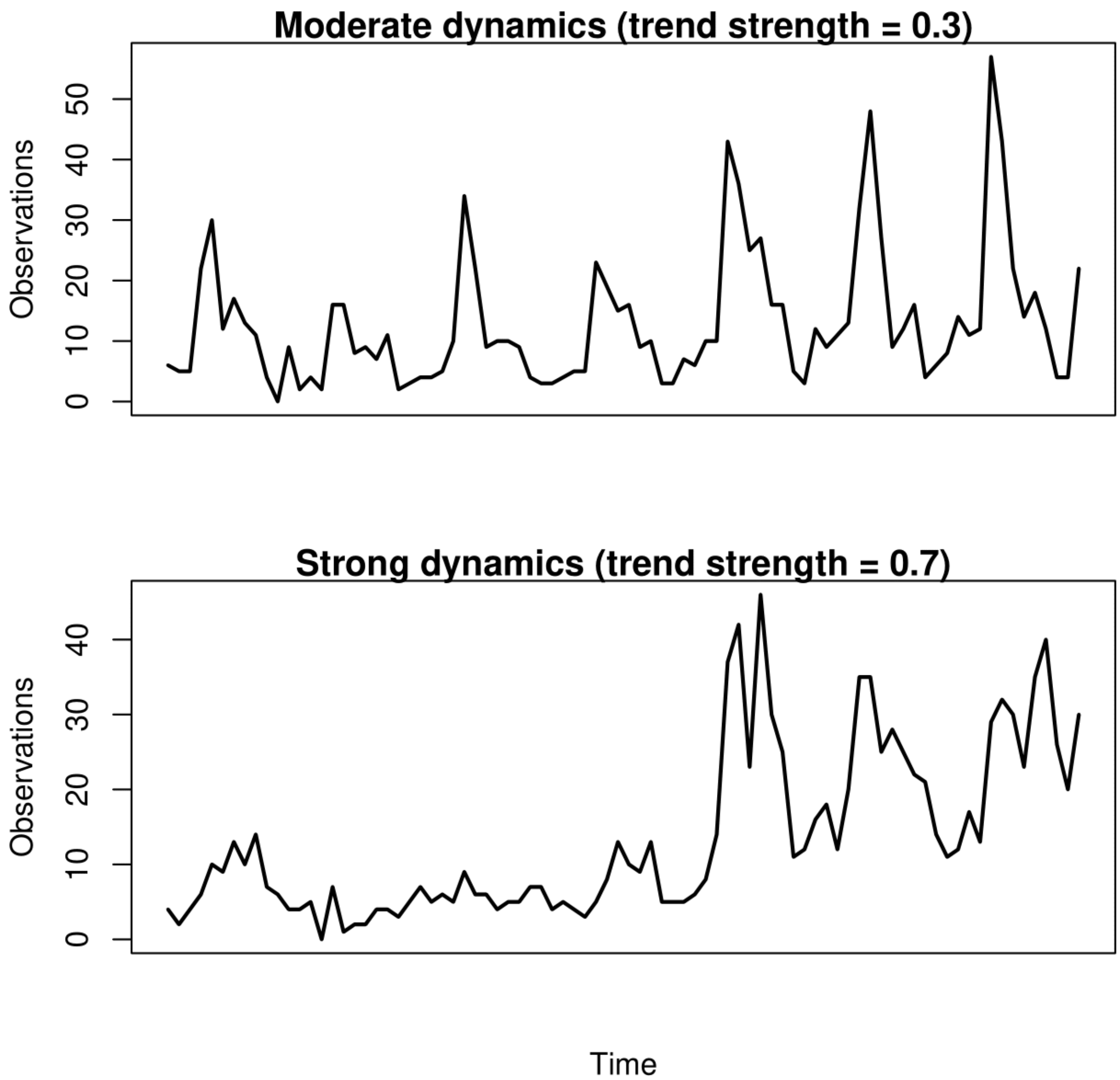

Figure S1: Two simulated discrete time series following the same seasonal pattern (with a seasonal frequency of 24 observation periods per year) but with different strengths of temporal dynamics. Both series were simulated using a Poisson observation process

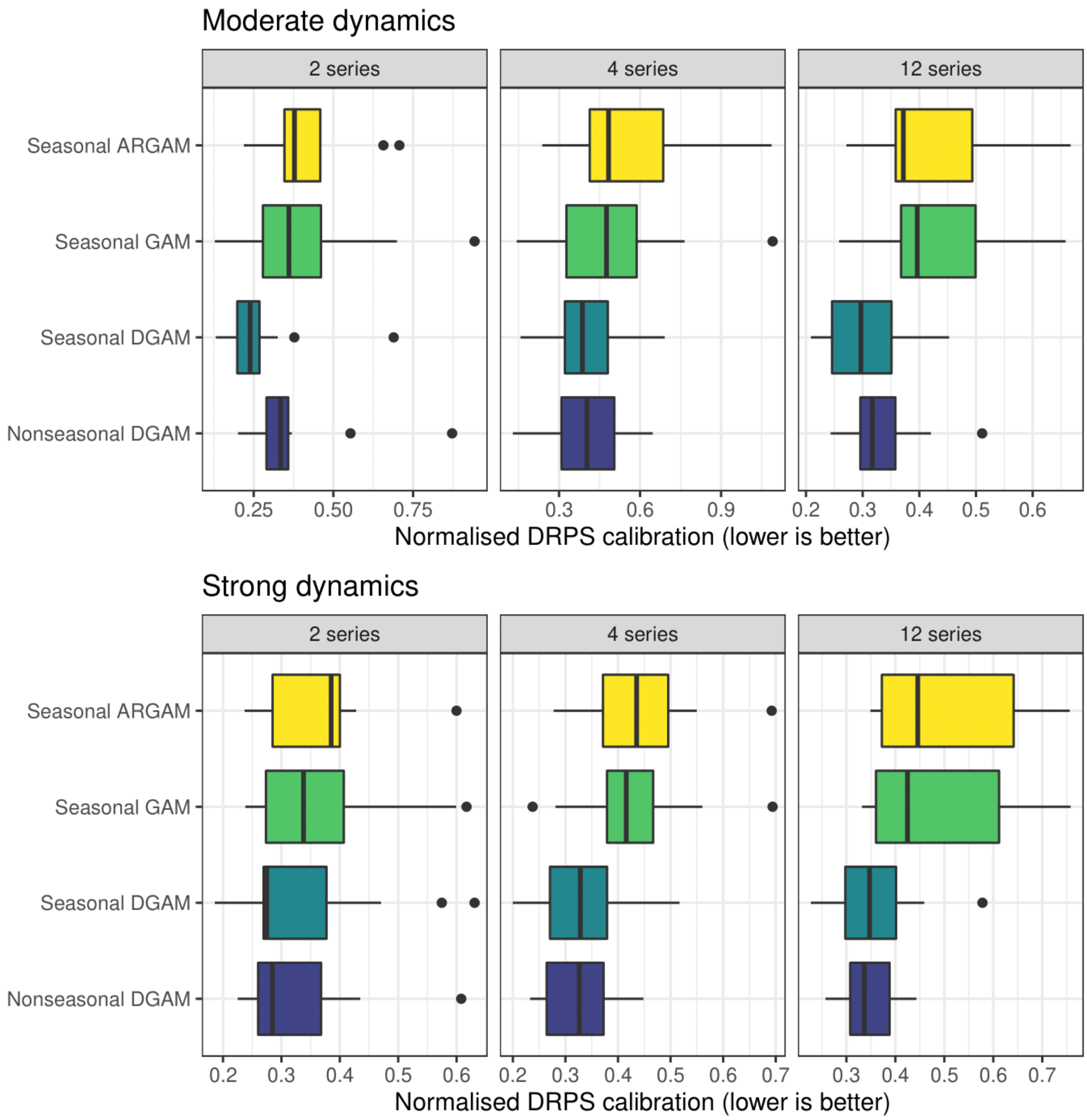

Figure S2: Normalised Discrete Rank Probability Score (DRPS) performance for out of sample forecasts from competing models fitted to sets of simulated discrete time series. Panels depict models fitted with different dimensionality (number of series) and temporal dynamics strength. The Seasonal GAM and Seasonal ARGAM were fitted using R package *mgcv*, while the Seasonal and Nonseasonal DGAMs were fitted using the *mvgam* package. Lower scores indicate better model performance.

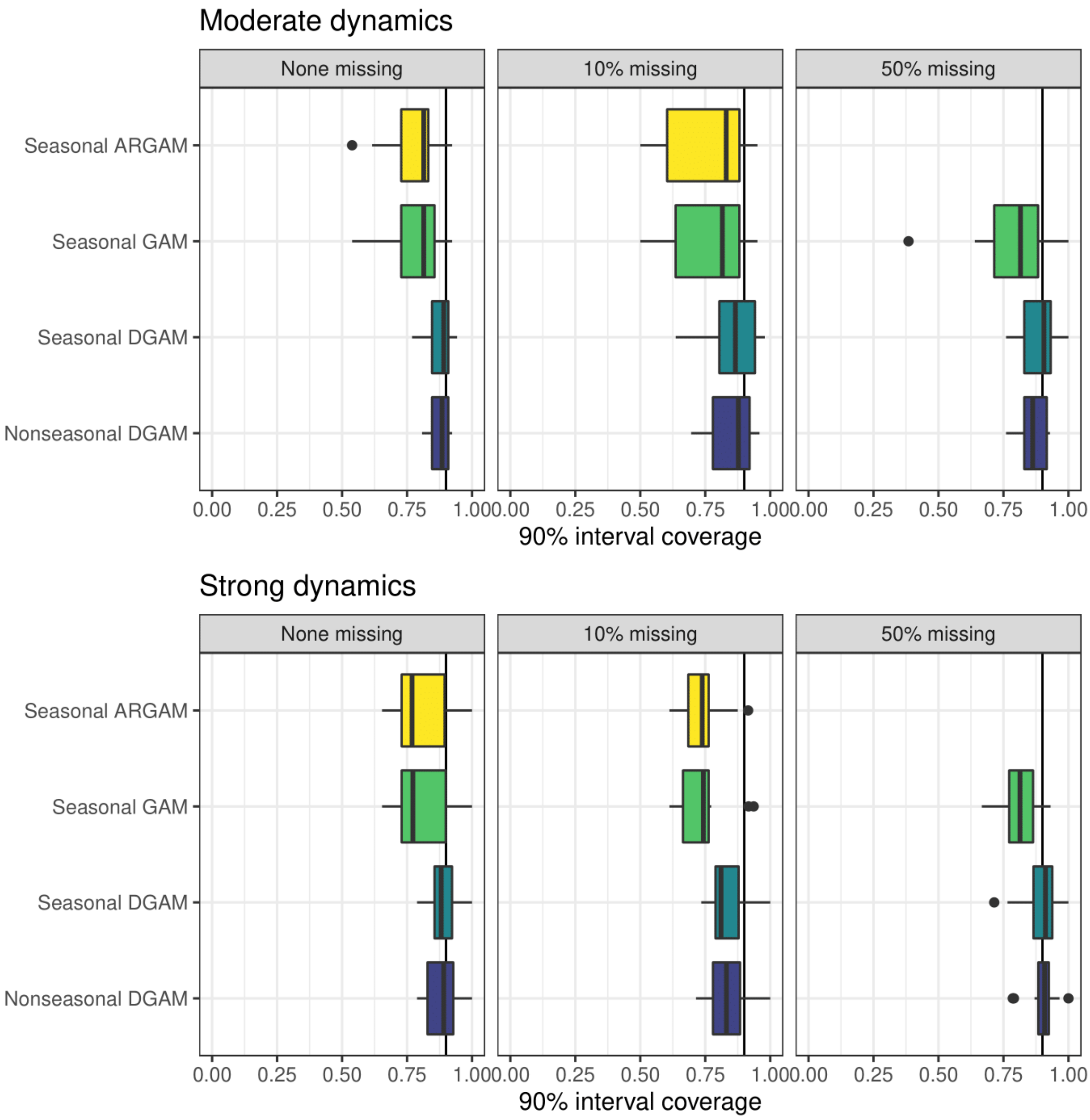

Figure S3: 90% interval coverage for out of sample forecasts from competing models fitted to sets of simulated discrete time series, plotted as a function of missingness (proportion of observations with NAs) and dynamics strength. The vertical line in each plot marks a coverage of 0.9. The Seasonal GAM and Seasonal ARGAM were fitted using R package *mgcv*, while the Seasonal and Nonseasonal DGAMs were fitted using the *mvgam* package. Scores closer to 0.9 are better.

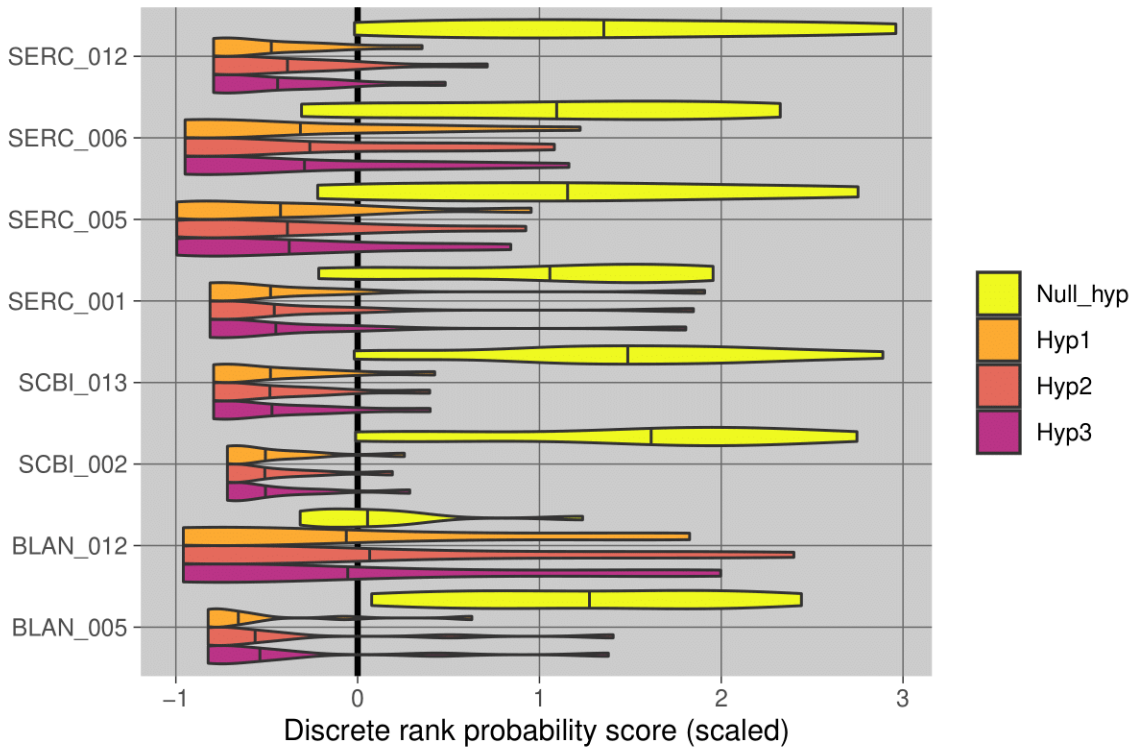

Figure S4: Discrete Rank Probability Score (DRPS) performance distributions for out of sample forecasts from competing models fitted to NEON's *Ixodes scapularis* abundance series, plotted as a function of site ID. Hypothesis definitions are outlined in main text section **CASE STUDY: FORECASTING TICK ABUNDANCES**.

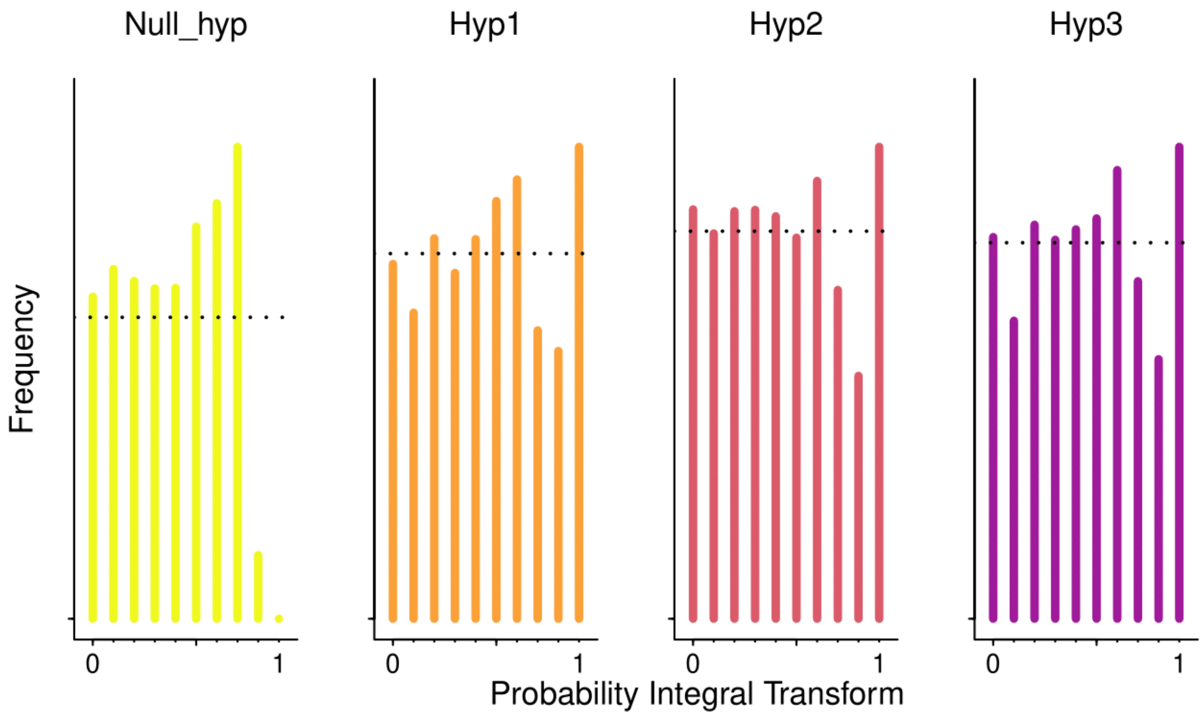

Figure S5: Probability Integral Transform (PIT) histograms for out of sample forecasts from competing models fitted to NEON's *Ixodes scapularis* abundance series. If probabilistic forecasts are well-calibrated, the histogram should be approximately uniform (indicating no systematic under-prediction or over-prediction). A left skew indicates the forecast tends to underpredict, while a right skew indicates the model tends to overpredict. Hypothesis definitions are outlined in main text section **CASE STUDY: FORECASTING TICK ABUNDANCES**.

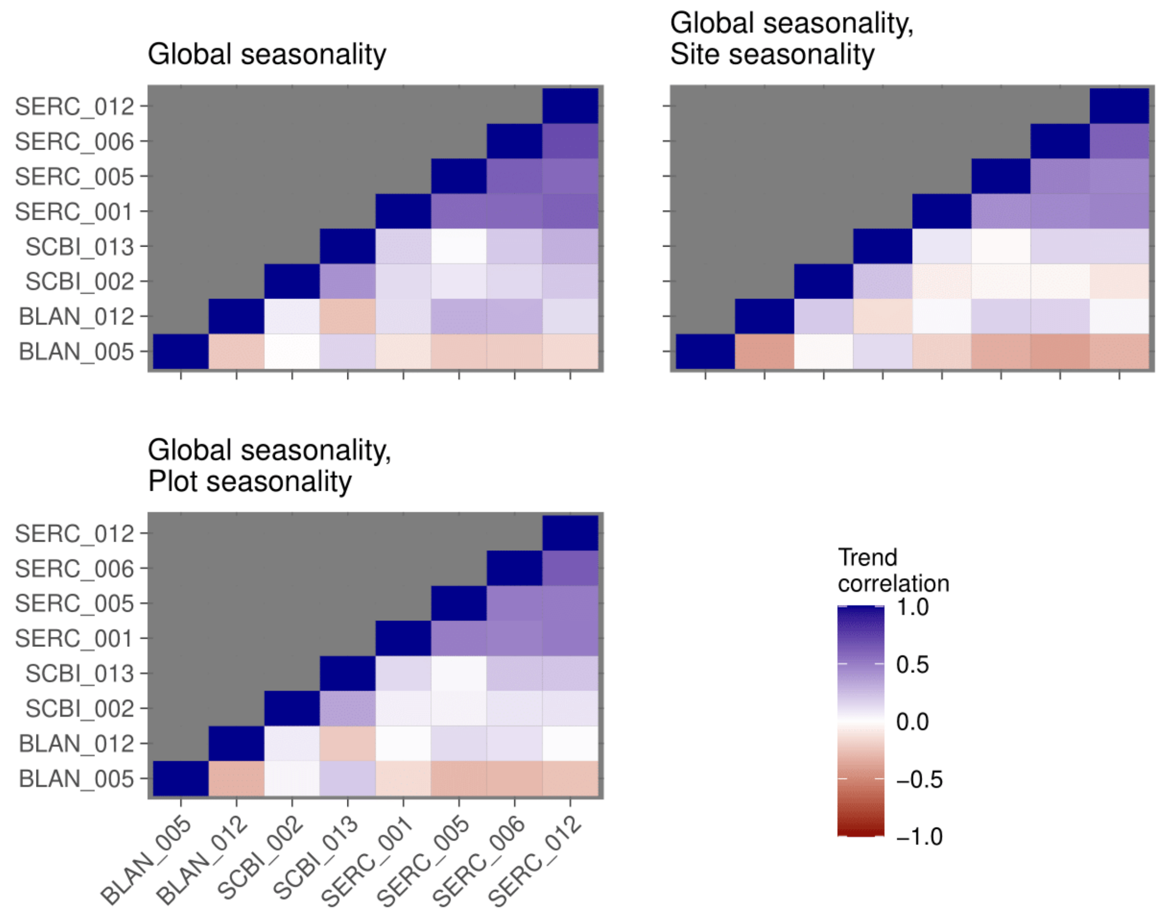

Figure S6: Posterior mean pairwise correlations among estimated latent dynamic trends from competing models fitted to NEON's *Ixodes scapularis* abundance series. Model definitions are outlined in main text section **CASE STUDY: FORECASTING TICK ABUNDANCES**.

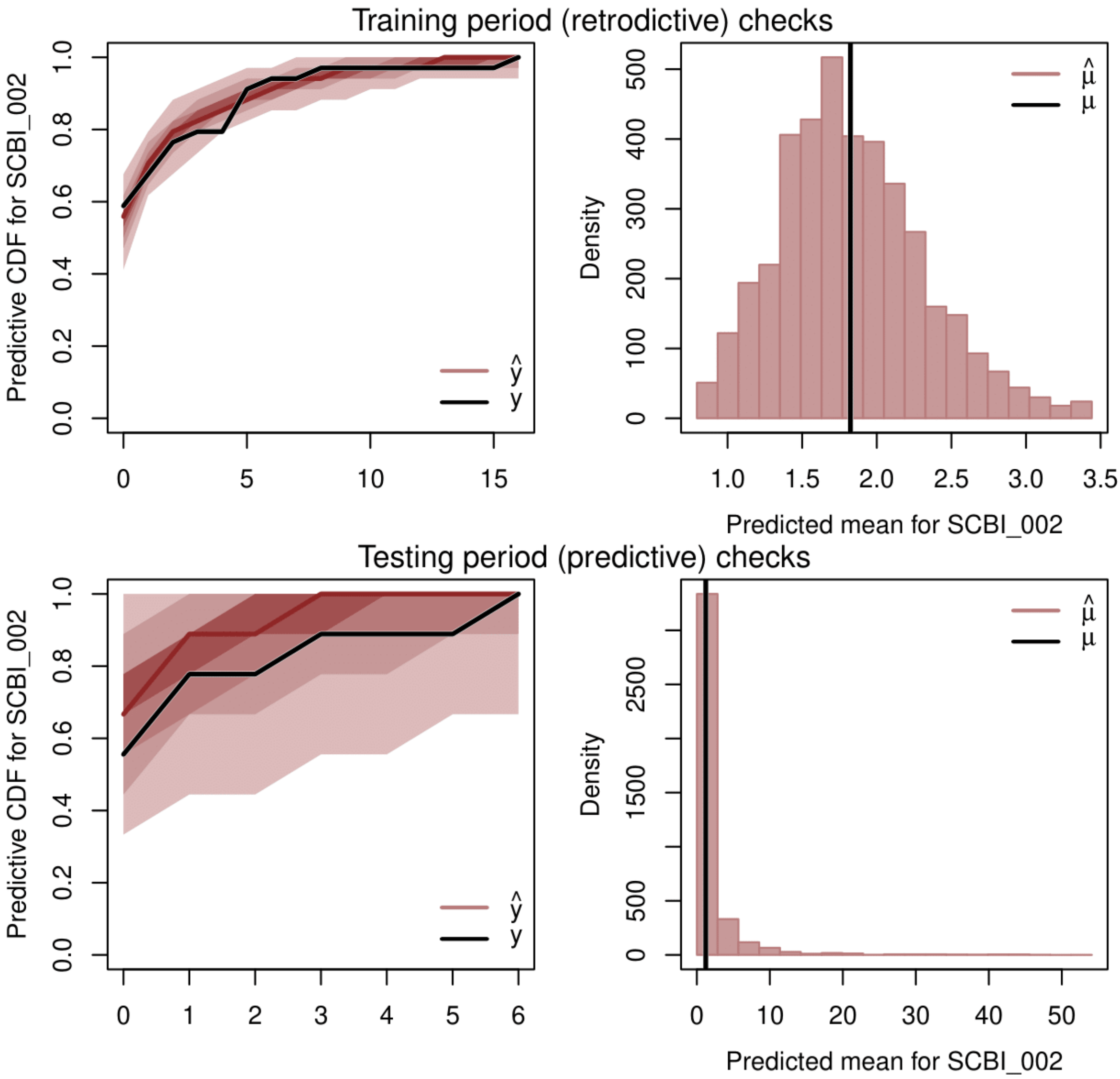

Figure S7: Example posterior retrodictive and predictive check visualisations (performed using the *mvgam* function *ppc*) from the best-performing *mvgam* model (Hyp3) for a single *Ixodes scapularis* plot (SCBI_002). Left-hand panels show the observed and estimated cumulative distribution functions for in-sample training data (top) and out-of-sample test data (bottom). Right-hand panels show observed and estimated means for training (top) and testing data (bottom). Shading of CDF plots shows posterior empirical quantiles. Hypothesis definitions are outlined in main text section **CASE STUDY: FORECASTING TICK ABUNDANCES**.

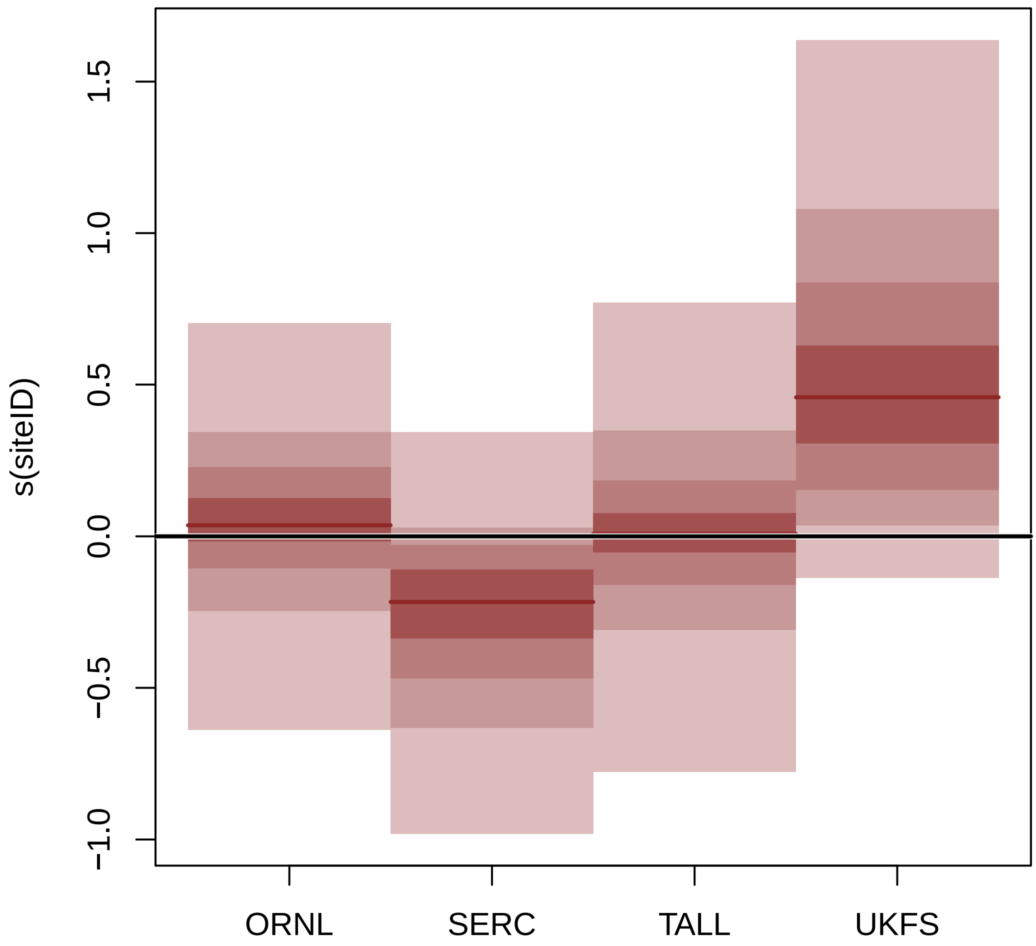

Figure S8: Output from the *plot_mvgam_randomeffects* function in *mvgam* showing posterior distributions for random intercept terms (on the log scale) for four *Amblyomma americanum* sites estimated from a dynamic GAM. Shading shows posterior empirical quantiles.