Supplementary Figures for the article (Figs. S1–S25):
Simulating future land use and cover of a Mediterranean mountainous area:
the effect of socioeconomic demands and climatic changes

by

Diogenis A. Kiziridis, Anna Mastrogianni, Magdalini Pleniou, Spyros Tsiftsis, Fotios Xystrakis, and Ioannis Tsiripidis

Contents

Figure **S1**

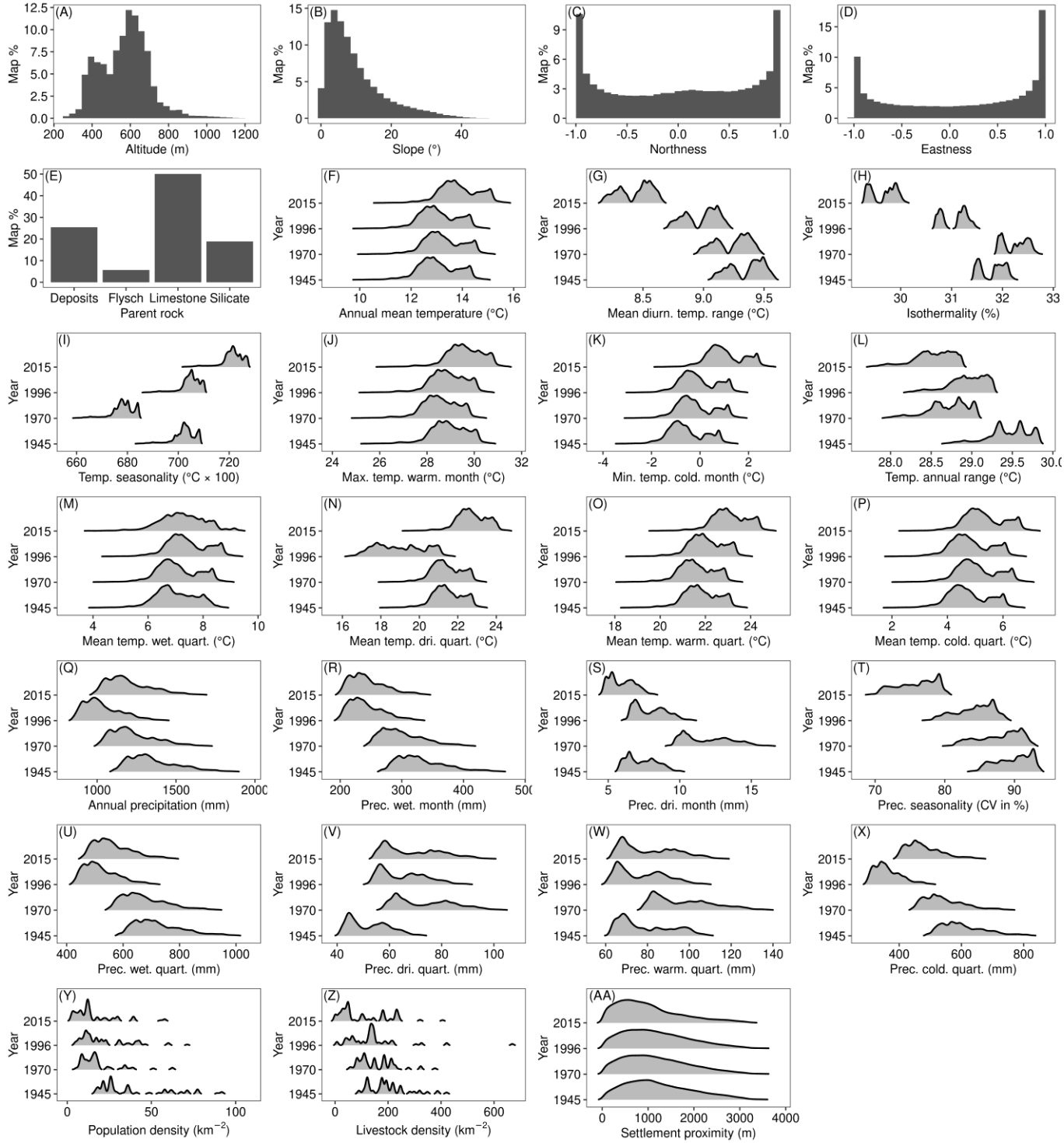

Fig. S1: Distribution of biophysical and socioeconomic predictors in the cells of our study area's maps at the historical reference years. Some variables were fixed in time, whereas others varied between years.

Figure S2

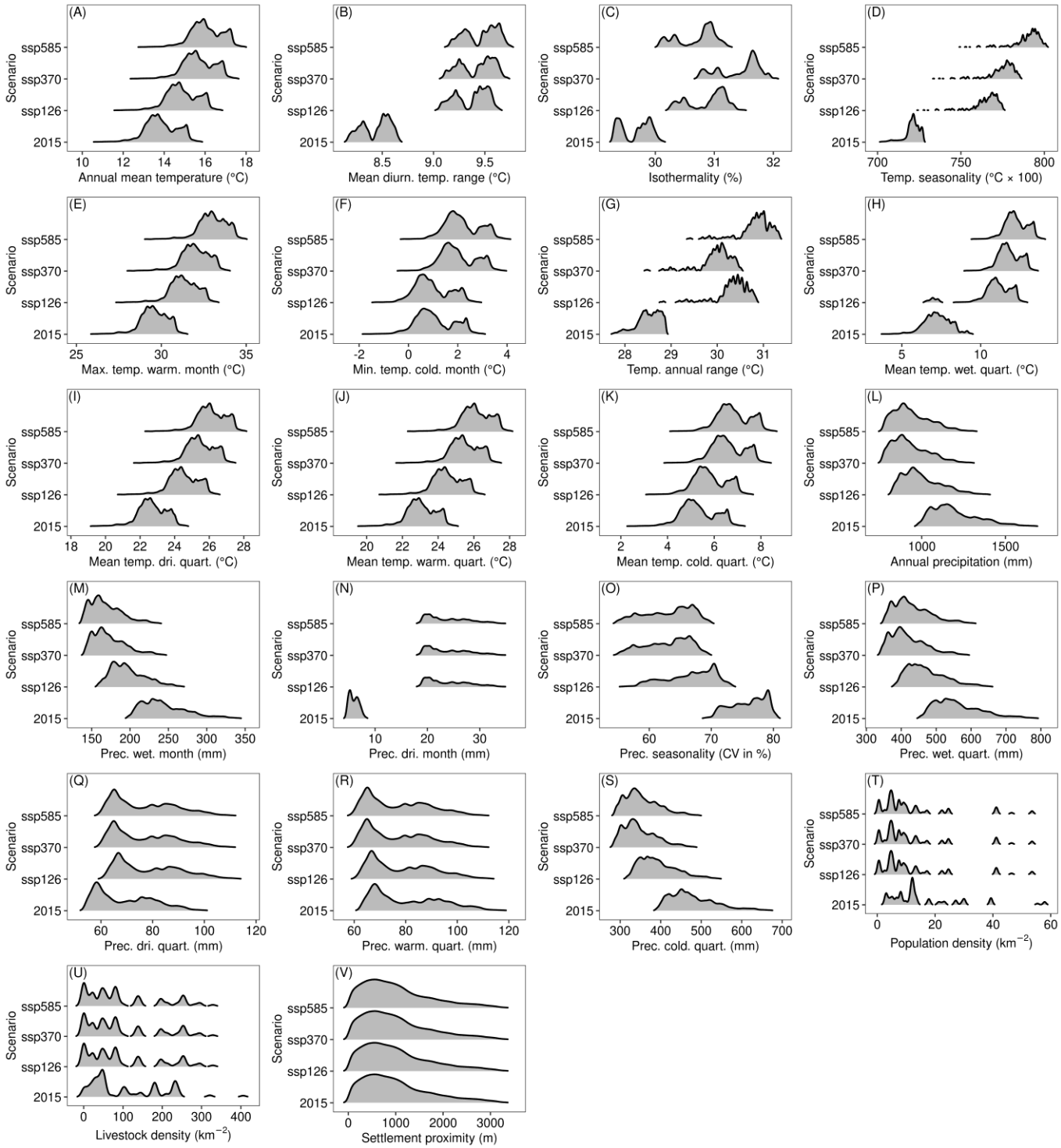

Fig. S2: Distribution of bioclimatic and socioeconomic predictors in the future at the cells of our study area's maps. For comparison, we additionally provide the distribution of the predictors in the latest reference year 2015. The three climatic scenarios concern only the 19 bioclimatic predictors (panels A–S), i.e. the socioeconomic predictors took the same values between these climatic scenarios (panels T–V). From "ssp126" to "ssp585", the climatic scenarios are stacked from more to less optimistic.

Figure S3

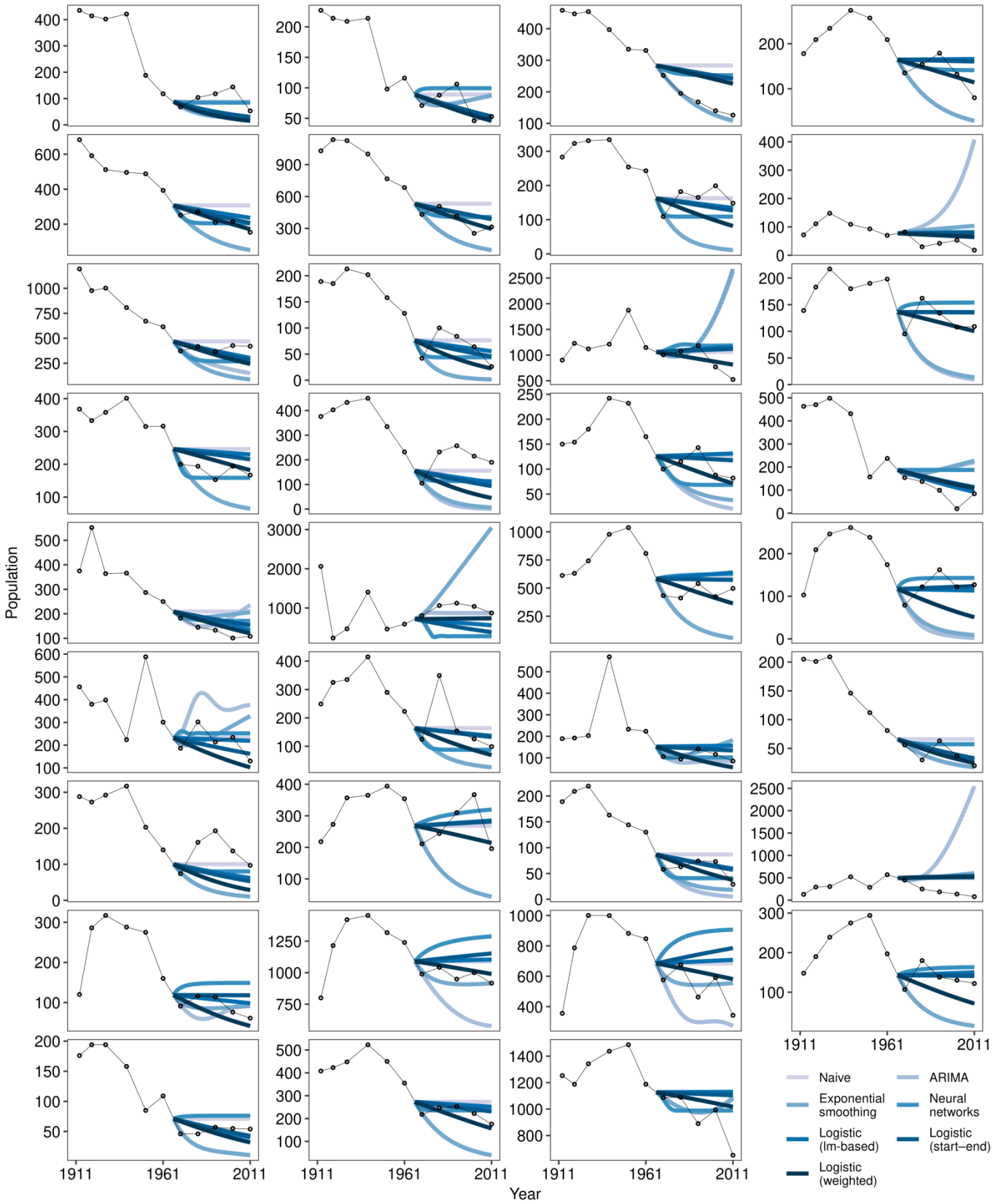

Fig. S3: Validating the prediction of population time series in the 35 municipal districts by seven models (bottom right). Each panel is for a different district. The thinner line segments just connect the points which are the census data observations. The thicker and colour-shaded curves are the predictions of the models.

Figure S4

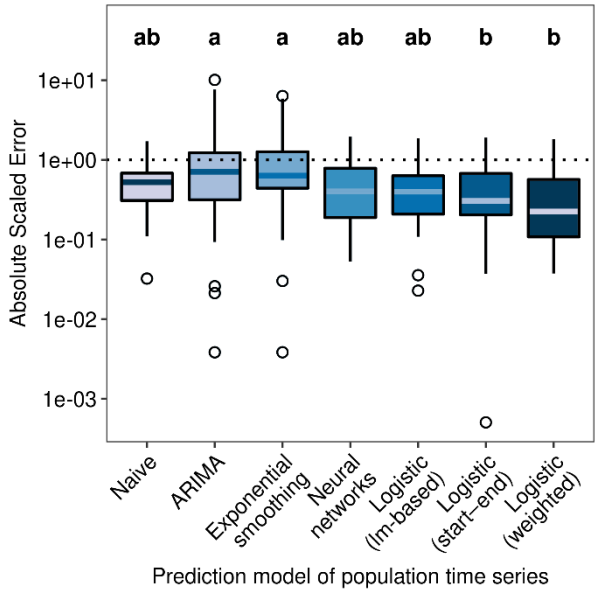

Fig. S4: Distribution of errors in the prediction of population time series during validation of the seven models. The 44-year window predictions from 1967 to 2011 were validated against the latest year 2011 of census data. Each boxplot summarises the distribution of the 35 error values from the 35 municipal districts. The labelling above the boxplots groups the models according to all-pairs comparisons with exact tests of Friedman-type ranked sums, from lower (a) to higher performance (b).

Figure S5

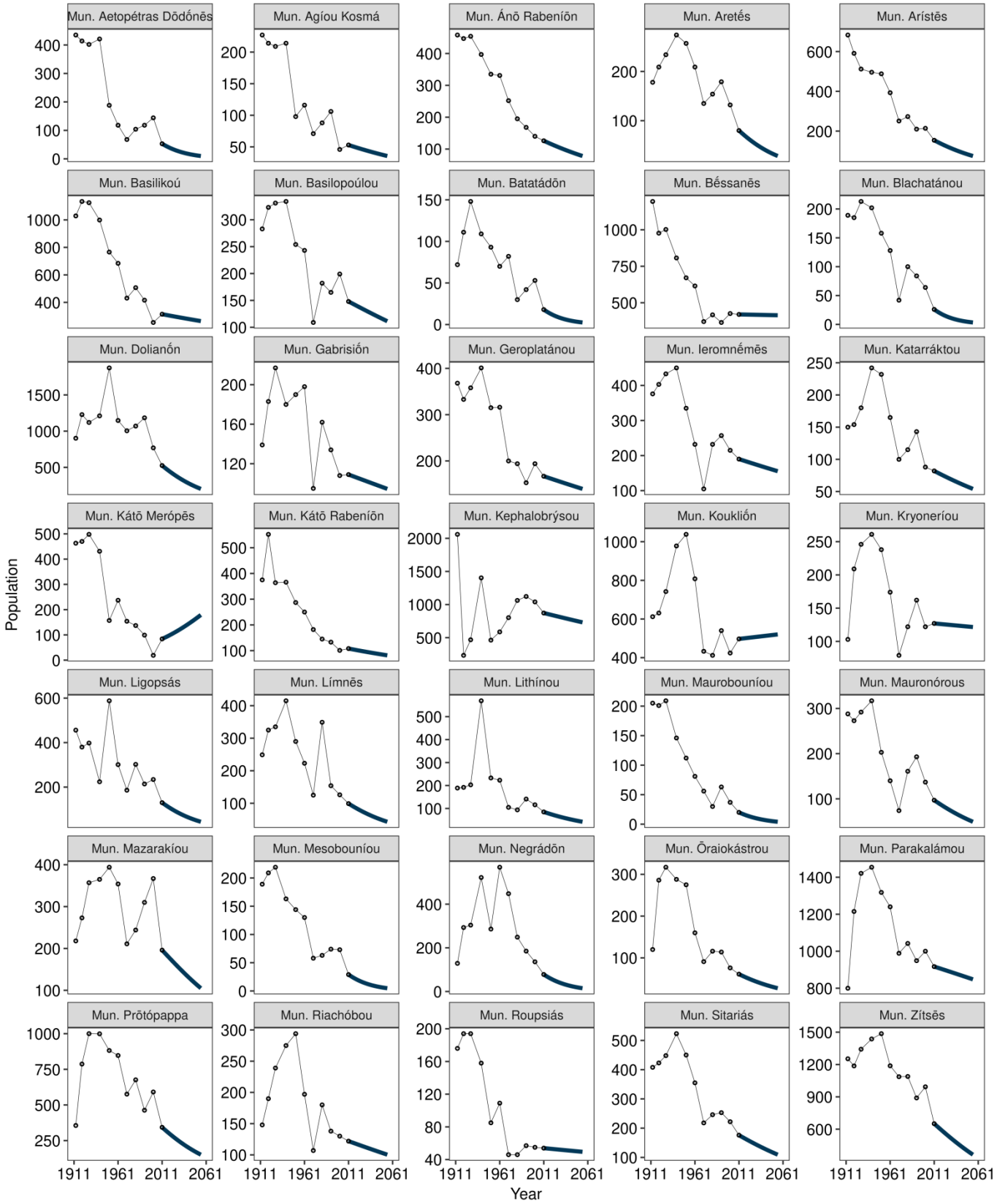

Fig. S5: Future predictions of population time series for the 35 municipal districts by the selected model with the lower prediction error during validation. The thinner segments just connect the points which are the census data observations. The thicker curves are the predictions of the model. The predictions are for 44 years since year 2011 for which we had the latest census data, i.e. until year 2055. Panels are ordered alphabetically by district name, and their order corresponds to the row-wise panel order of Fig. S3. The selected model was the Logistic (weighted), i.e. more recent observations had exponentially higher weight in the calculation of the mean of the growth/decay parameter *r* of the logistic model. The other parameter of the logistic model, the carrying capacity *K*, was equal to the maximum population size which was observed in the census data of each district.

Figure S6

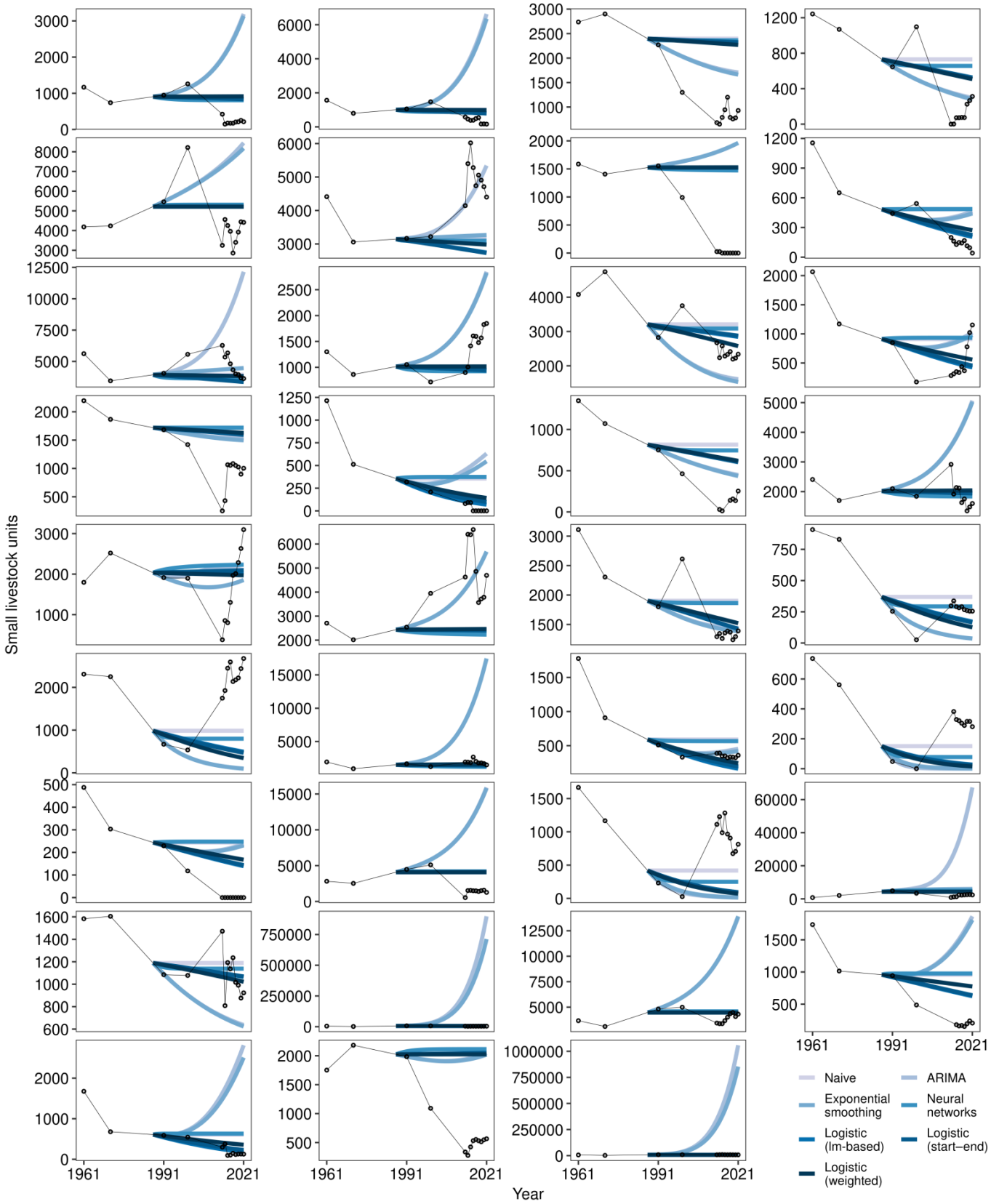

Fig. S6: Validating the prediction of livestock time series in the 35 municipal districts by seven models (bottom right). Each panel is for a different district. The thinner line segments just connect the points which are the census data observations. The thicker and colour-shaded curves are the predictions of the models.

Figure S7

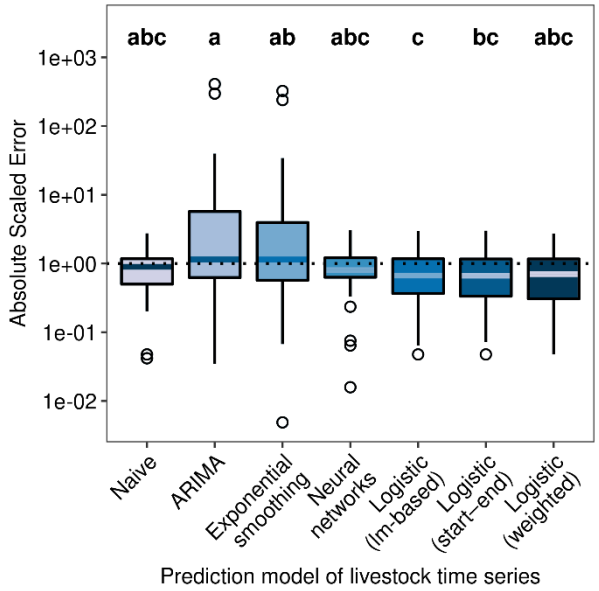

Fig. S7: Distribution of errors in the prediction of livestock time series during validation of the seven models. The 34-year window predictions from 1987 to 2021 were validated against the latest year 2021 of census data. Each boxplot summarises the distribution of the 35 error values from the 35 municipal districts. The labelling above the boxplots groups the models according to all-pairs comparisons with exact tests of Friedman-type ranked sums, from lower (a) to higher performance (c).

Figure S8

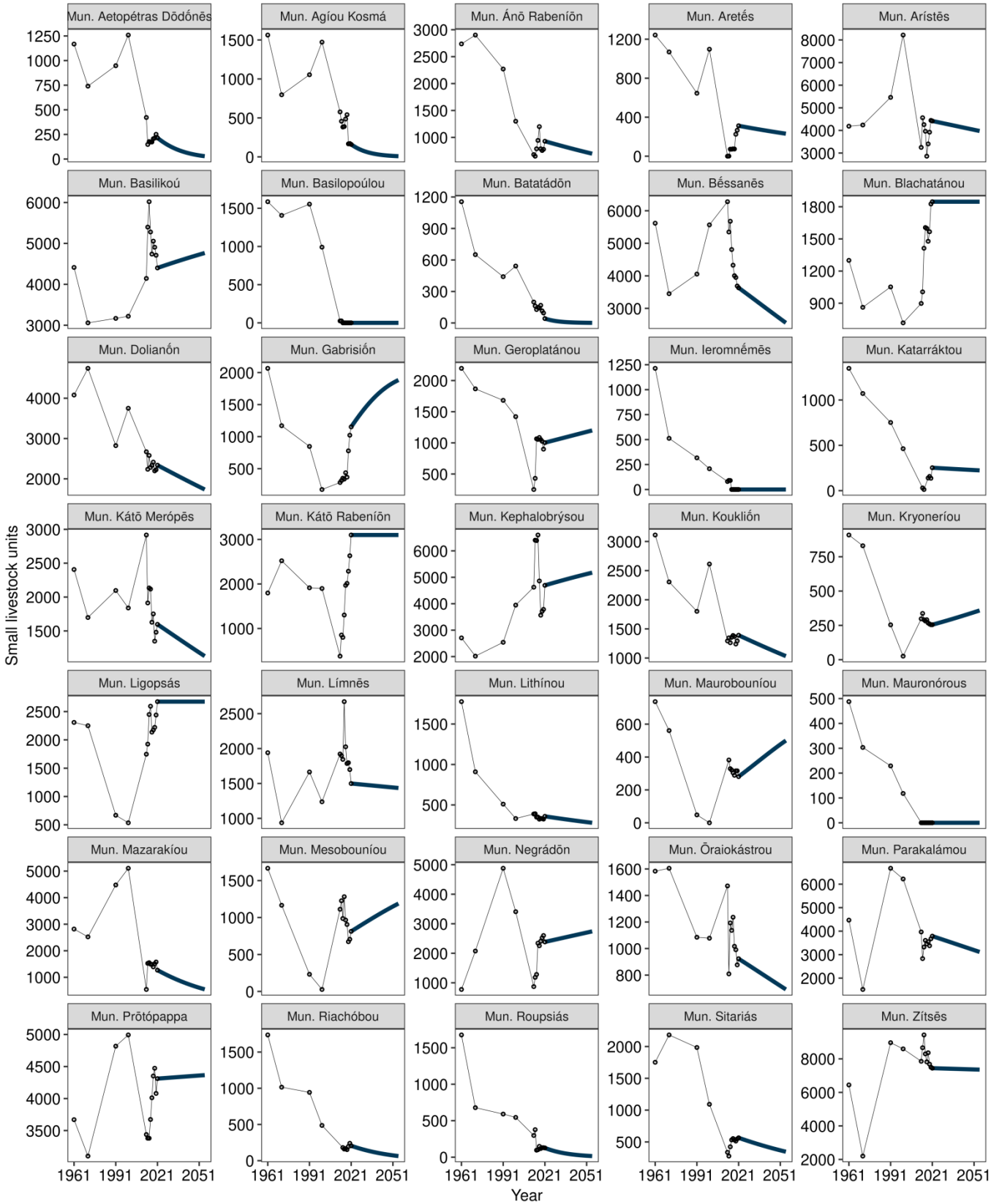

Fig. S8: Future predictions of livestock time series for the 35 municipal districts by the selected model with the lower prediction error during validation. The thinner line segments just connect the points which are the census data observations. The thicker curves are the predictions of the model. The predictions are for 34 years since year 2021 for which we had the latest census data, i.e. until year 2055. Panels are ordered alphabetically by district name, and their order corresponds to the row-wise panel order of Fig. S6. The selected model was the Logistic (weighted), i.e. more recent observations had exponentially higher weight in the calculation of the mean of the growth/decay parameter *r* of the logistic model. The other parameter of the logistic model, the carrying capacity *K*, was equal to the maximum livestock size which was observed in the census data of each district.

Figure S9

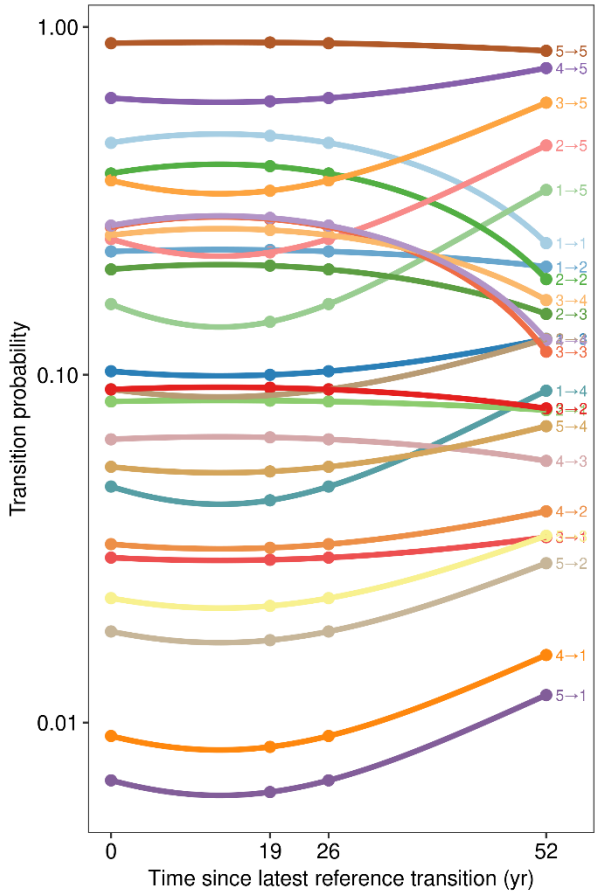

Fig. S9: The quadratic fits in the regression-based procedure for Markov transition probability estimation of the validation period's business-as-usual demand scenario of the trans-CLUE-S model. Each quadratic curve is for a different transition, as labelled on the far right. LUC type indices: (1) farmland; (2) grassland; (3) open-scrub; (4) closed-scrub; and (5) forest. For example, "5→1" is for the forest-to-farmland transition. The calibration period was the 26-year period of 1970–1996 upon which the Markov-based extrapolation was done for: the 1996–2022 period (26 years since the calibration period on the x-axis); and the 1996–2048 period (2 × 26 = 52 years). The quadratic fits facilitated the interpolation of the transition probabilities for the desired 1996–2015 validation period (19 years until the validation year 2015).

Figure S10

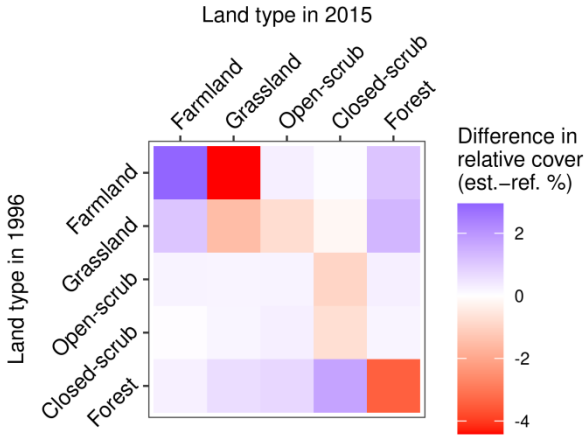

Fig. S10: Differences in the transitioned relative cover of the estimated by the quadratic-based procedure (est.) minus of the reference time period 1996–2015 (ref.).

Figure S11

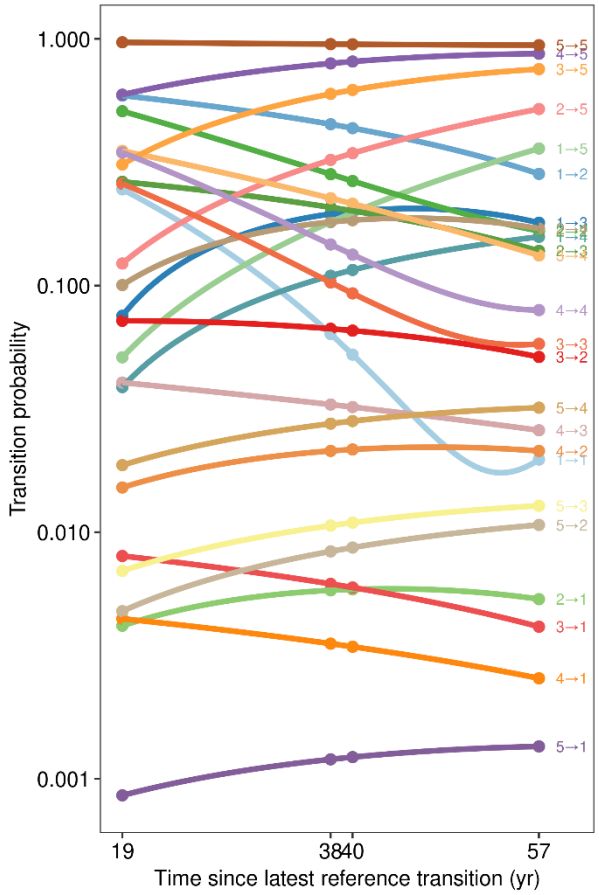

Fig. S11: The quadratic fits in the regression-based procedure for Markov transition probability estimation of the 2015–2055 future period's business-as-usual demand scenario of the final trans-CLUE-S model. Each quadratic curve is for a different transition, as labelled on the far right. LUC type indices: (1) farmland; (2) grassland; (3) open-scrub; (4) closed-scrub; and (5) forest. For example, "5→1" is for the forest-to-farmland transition. The calibration period was the 19-year period of 1996–2015 upon which the Markov-based extrapolation was done for: the 2015–2034 period (19 years since the calibration period on the x-axis); the 2015–2053 period (2 × 19 = 38 years); and the 2015–2072 period (3 × 19 = 57 years). The quadratic fits facilitated the interpolation of the transition probabilities for the desired 2015–2055 period (40 years until the prediction year 2055).

Figure S12

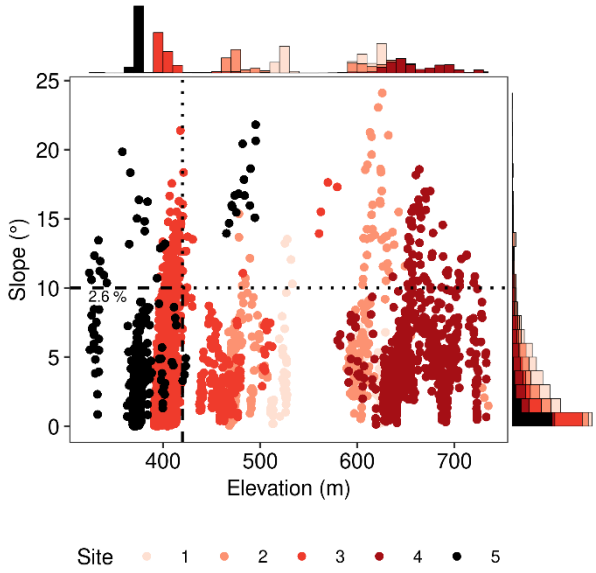

Fig. S12: Selecting farmland of intensive farming which will be assumed to additionally persist, for providing an additional demand scenario. Data points correspond to the map cells of farmland in 2015, and are coloured by the site they occur, according to the bottom legend. The marginal histograms provide the stacked-by-site frequency distributions of the cells' elevation (top x-axis) and slope (right y-axis). The dotted lines indicate the threshold value of these two predictors below of which the cells were selected. These selected cells inside the dashed-lined rectangle at the bottom left comprised the indicated relative cover of persisting farmland in year 2055.

Figure S13

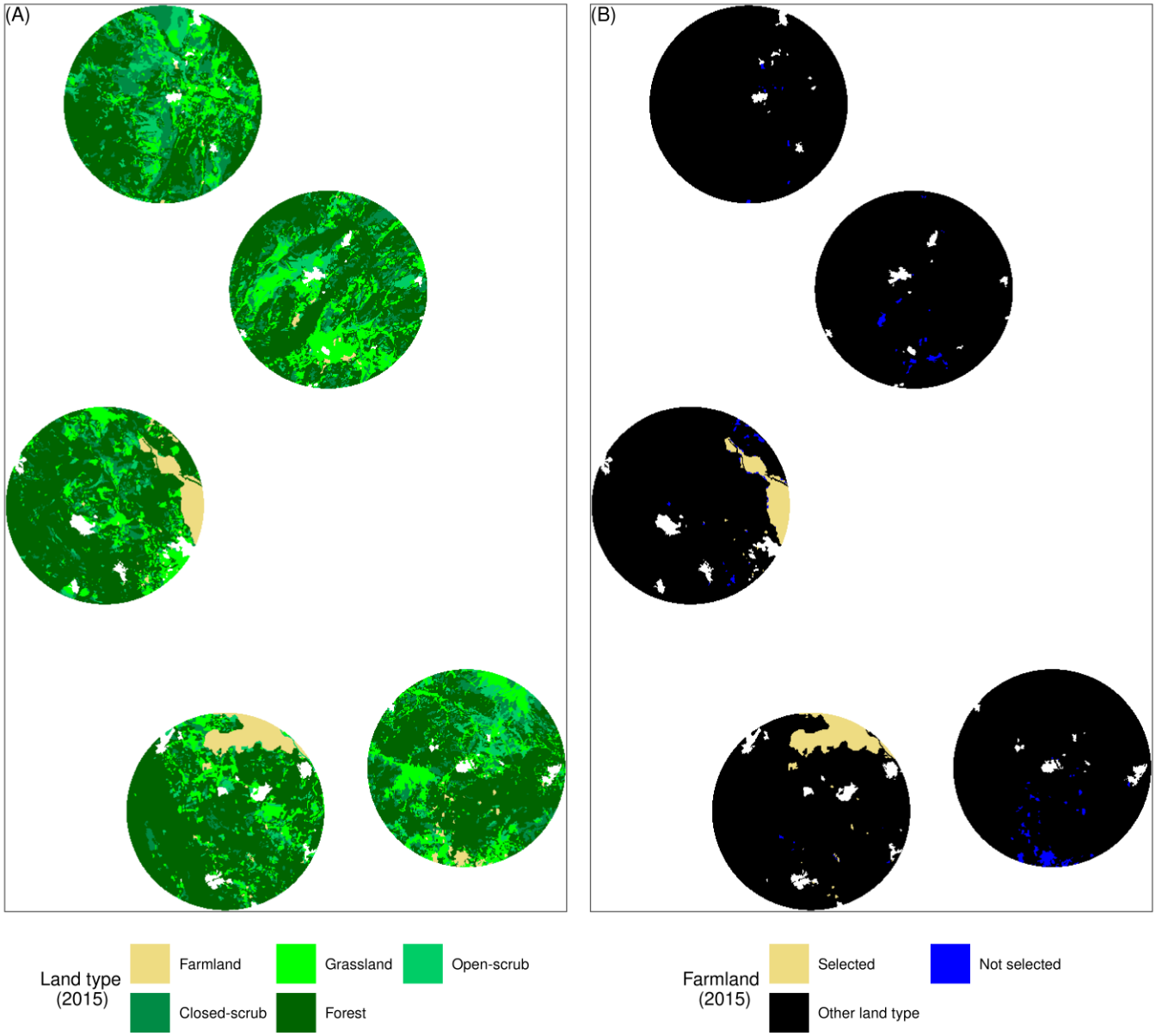

Fig. S13: Spatial distribution of the additional farmland in 2015 which was assumed to persist due to being of intensive farming, according to the selection illustrated in Fig. S12. The two maps show the land cover of 2015 (A), and the corresponding positions of the 2015 farmland which was selected and not selected to persist (B).

Figure S14

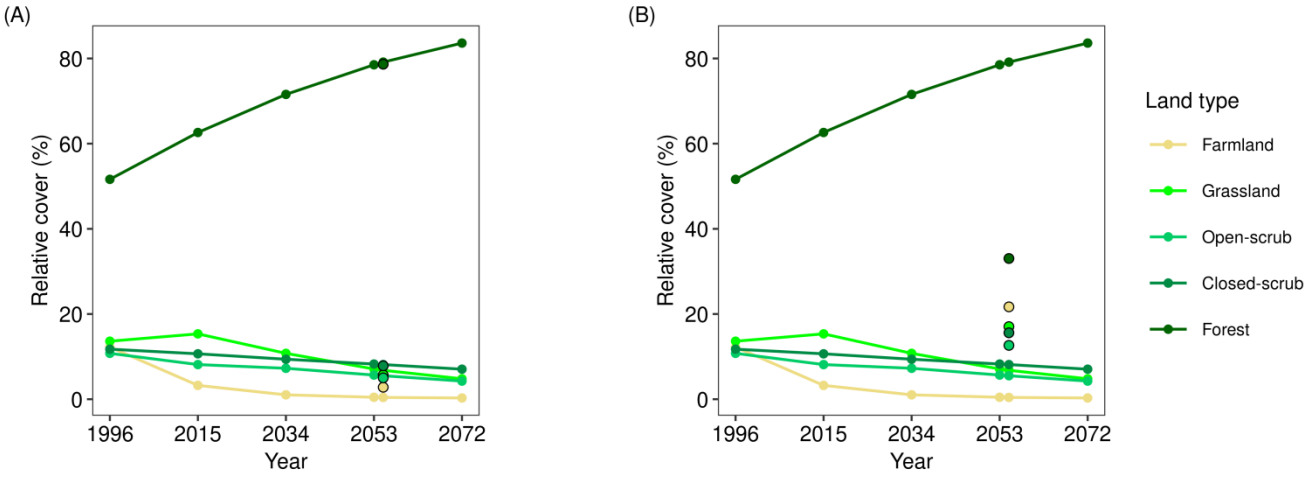

Fig. S14: Relative cover of the LUC types under the three demand scenarios of trans-CLUE-S for predicting land cover in year 2055. The business-as-usual scenario is indicated with the line segments which connect the data points. The data points with the black outline are for: (A) the third scenario which assumes that additional farmland to the business-as-usual will persist due to persistence of intensive farming; and (B) for the fourth scenario which assumes that demand in 2055 will be equal to the relative cover of the LUC types in 1970 due to extensification of farming.

Figure S15

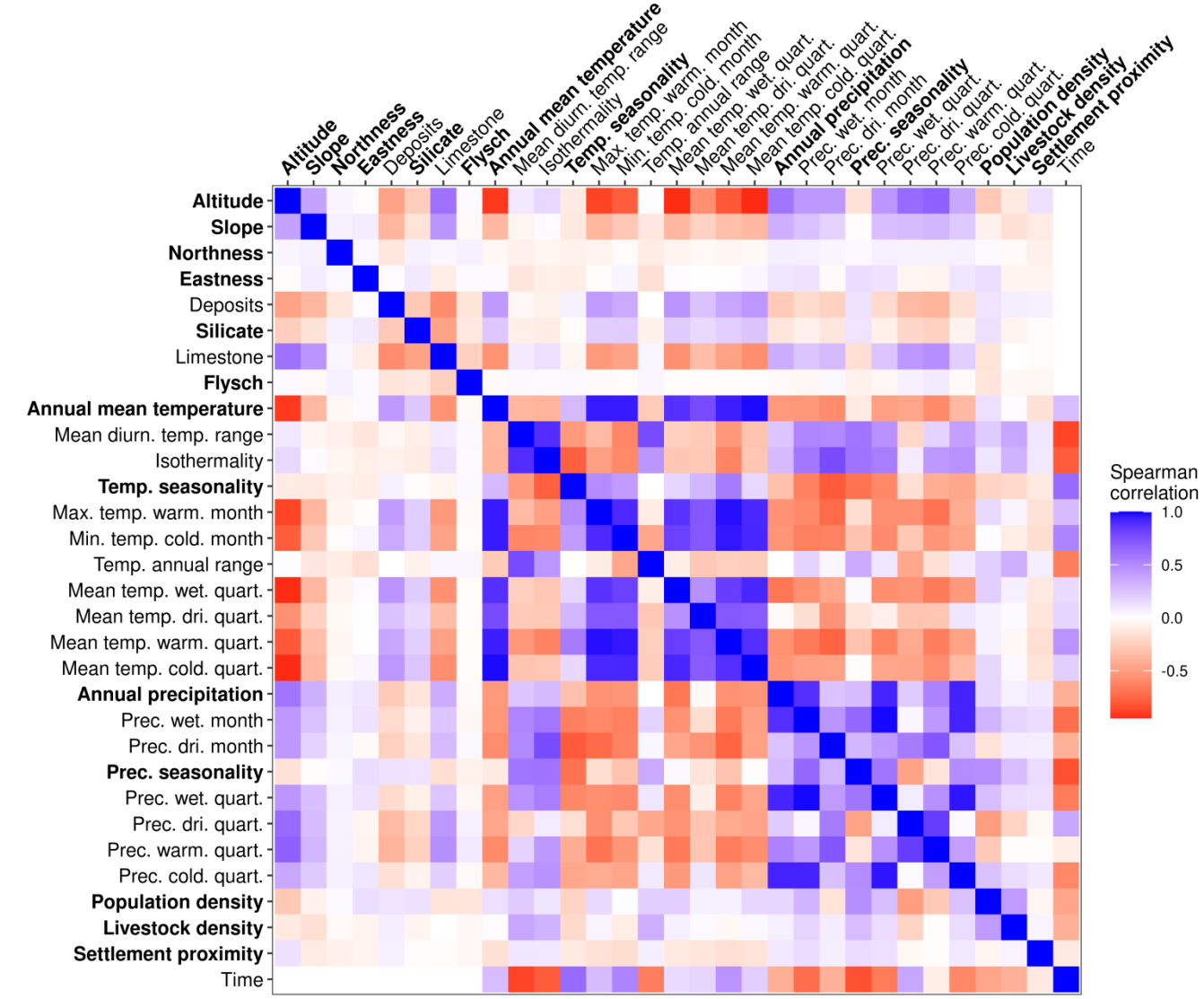

Fig. S15: Spearman correlation coefficient values between candidate predictors for the demand allocation suitability component of the trans-CLUE-S model. We have also appended a “time” variable (years since the earliest reference year 1945, for years 1945, 1970, 1996 and 2015). The four categories of parent rock could be more conveniently introduced to the Random Forest suitability models as presence–absence, binary variables, i.e. one variable for the presence of each category of parent rock. Aiming for lower inter-correlations but also higher interpretability, we selected the variables in bold for these multinomial classification models which return the suitability of each cell to be covered by each LUC type.

Figure S16

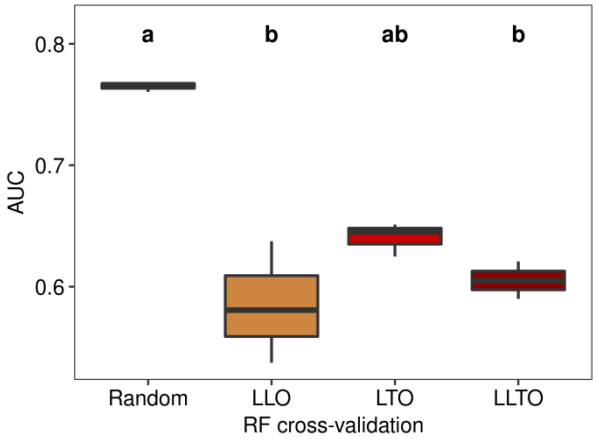

Fig. S16: Distribution summaries of the Area Under the ROC Curve (AUC) in the cross-validation partitions of four Random Forest models (RF). These four multinomial classification models were trained and cross-validated with data from the reference maps of 1945, 1970 and 1996. The cross-validation types were the following: Random (partitioning all observations randomly to training and validation sets); Leave-Location-Out, LLO (partitioning observations according to the study sites they occurred), Leave-Time-Out, LTO (partitioning observations according to the study year), and the Leave-Location-and-Time-Out, LLTO (partitioning observations according to both study site and year). The letters above the distributions indicate their grouping based on the post-hoc Dunn's test of equality between all pairs of cross-validation types using rank sums.

Figure S17

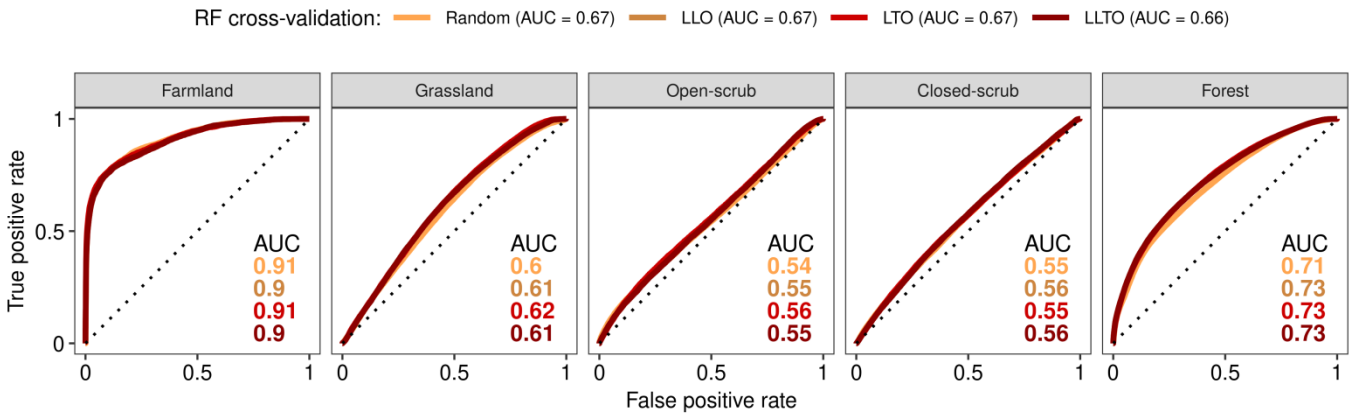

Fig. S17: The Area Under the ROC Curve (AUC) of the four types of cross-validation for the Random Forest (RF) models which were candidate suitability models for the trans-CLUE-S model. These four multinomial classification models were trained with data from the reference maps of 1945, 1970 and 1996. Each panel is from validating the model predictions of a different LUC type in the reference map of 2015, i.e. without the trans-CLUE-S model. The ROC curves are coloured according to the legend. The legend additionally provides each model's average AUC among the LUC types, whereas the AUC for each LUC type is given in each respective panel. The cross-validation types were the following: Random (partitioning all observations randomly to training and validation sets); Leave-Location-Out, LLO (partitioning observations according to the study sites they occurred), Leave-Time-Out, LTO (partitioning observations according to the study year), and the Leave-Location-and-Time-Out, LLTO (partitioning observations according to both study site and year).

Figure S18

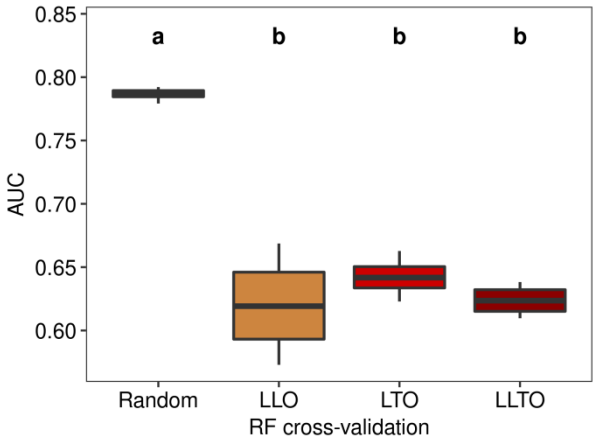

Fig. S18: Distribution summaries of the Area Under the ROC Curve (AUC) in the cross-validation partitions of four candidate Random Forest models (RF) for use as suitability models of the final trans-CLUE-S model for predicting land cover in 2055. These four multinomial classification models were trained and validated with data from the reference maps of 1945, 1970, 1996 and 2015. The cross-validation types were the following: Random (partitioning all observations randomly to training and validation sets); Leave-Location-Out, LLO (partitioning observations according to the study sites they occurred), Leave-Time-Out, LTO (partitioning observations according to the study year), and the Leave-Location-and-Time-Out, LLTO (partitioning observations according to both study site and year). The letters above the distributions indicate their grouping based on the post-hoc Dunn's test of equality between all pairs of cross-validation types using rank sums.

Figure S19

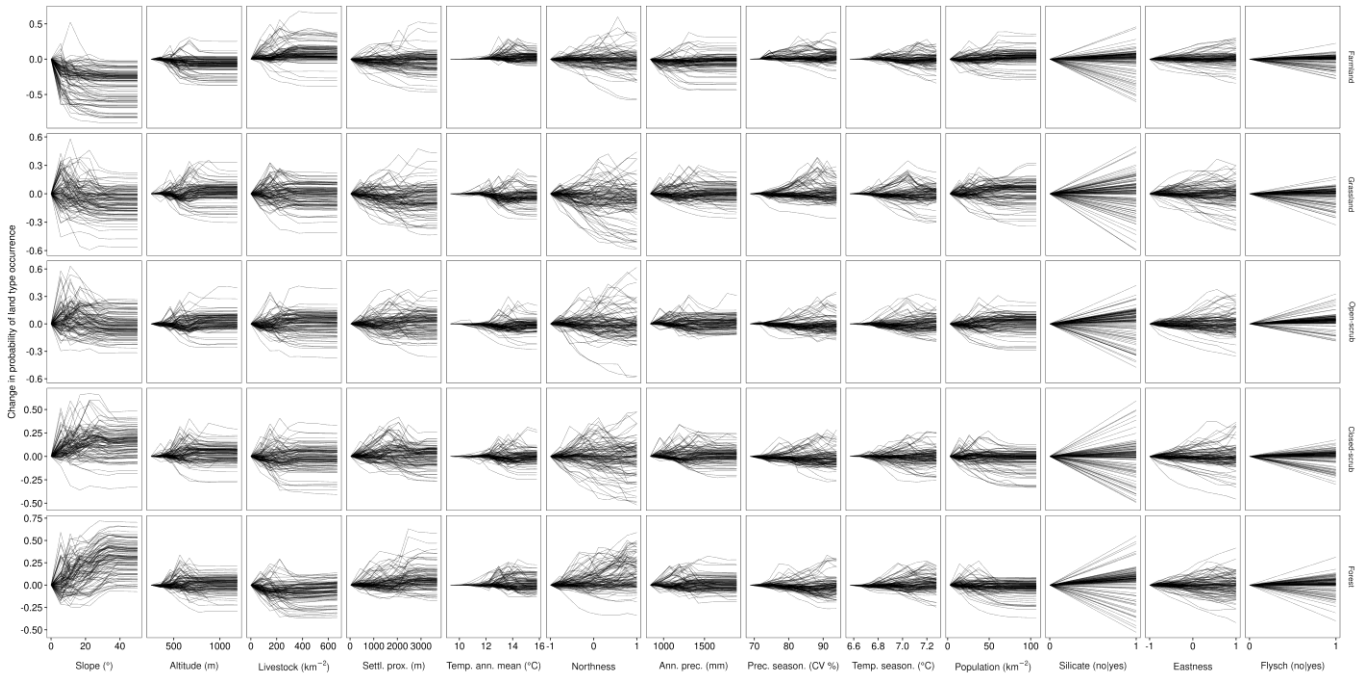

Fig. S19: Individual marginal effect curves of the predictor variables (columns) in the Random Forest model of suitability for predicting the cover of each LUC type (rows) in 2055 with the trans-CLUE-S model. Only a random sample of 100 curves is shown with the dark and thin curves. For comparability, the curves are centred to the mean prediction for the left-most, minimum value of each predictor. The predictors (columns) are ordered from left to right in decreasing relative importance for the classification of LUC type by the Random Forest model.

Figure S20

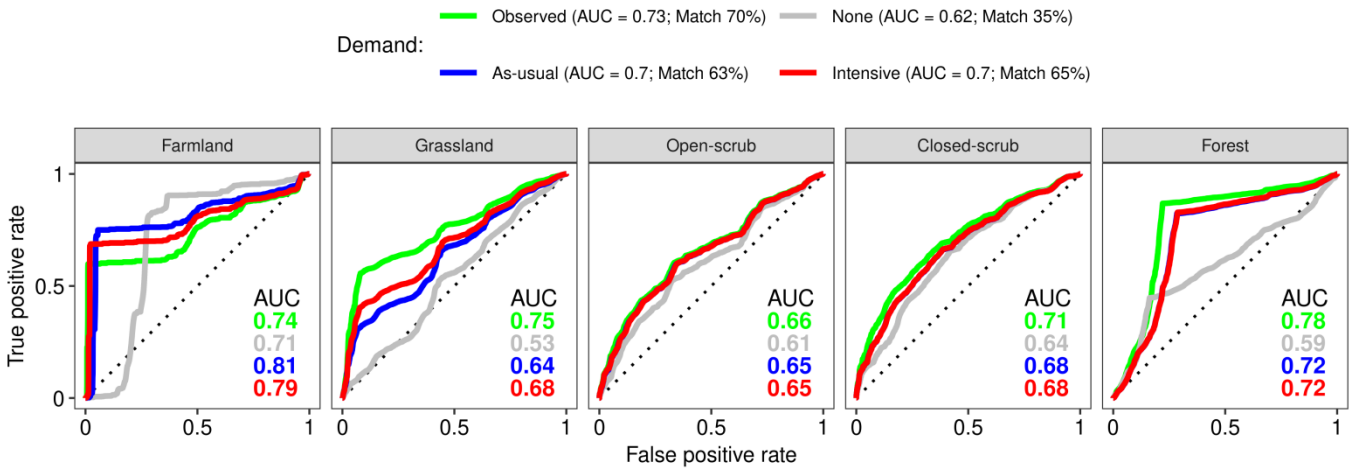

Fig. S20: Validating the trans-CLUE-S model against the observed map of 2015. The model was tested against four demand scenarios, according to the legend at the top. The case of observed demand is about using the actual demand from the 1996–2015 period. The match inside the brackets of the legend refers to the percent of the predicted 2015 map which had the same LUC as the observed LUC map of year 2015.

Figure S21

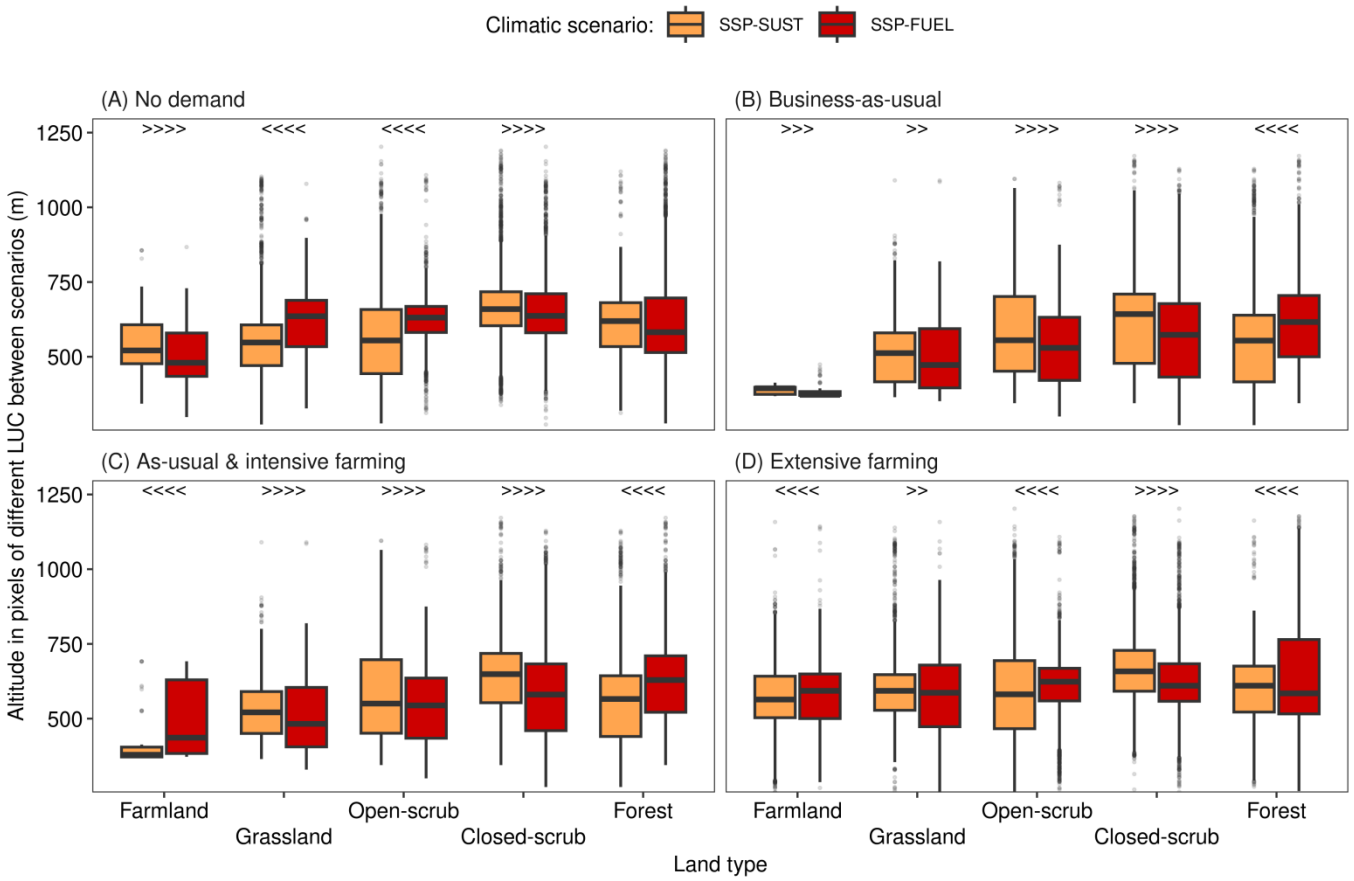

Fig. S21: LUC type elevation at the map cells which differed in the prediction of LUC between the most optimistic (SSP126) and most pessimistic (SSP585) climatic scenario, under each of the following four scenarios of demand in 2015–2055 LUC transitions: (A) no pre-specified demand; (B) as-usual from 1996–2015 to 2015–2055; (C) as-usual but with the 2015 intensive farming preserved; and (D) inverse transitioning to 1970 which had characteristics of extensive farming. Inequality symbols above the boxplots indicate statistically significant difference in the median slope by Wilcoxon test between scenarios. In specific, four inequality symbols were used for p-value ≤ 10^{-4}, three symbols for $p \leq 10^{-3}$, two for $p \leq 0.01$, one for $p \leq 0.05$, and no symbol for $p > 0.05$ level of statistical significance.

Figure S22

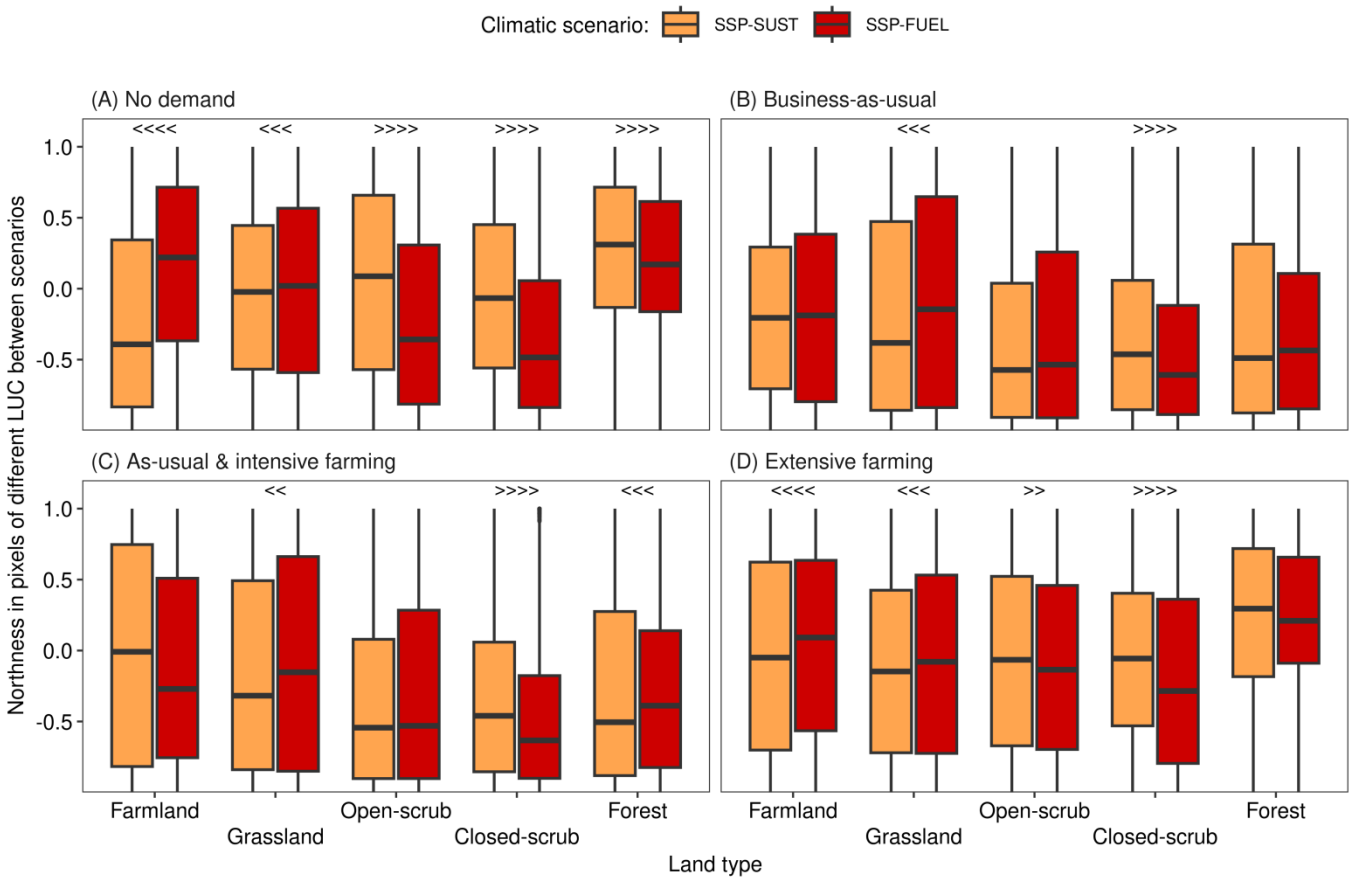

Fig. S22: Same as Fig. S21, but for northness.

Figure S23

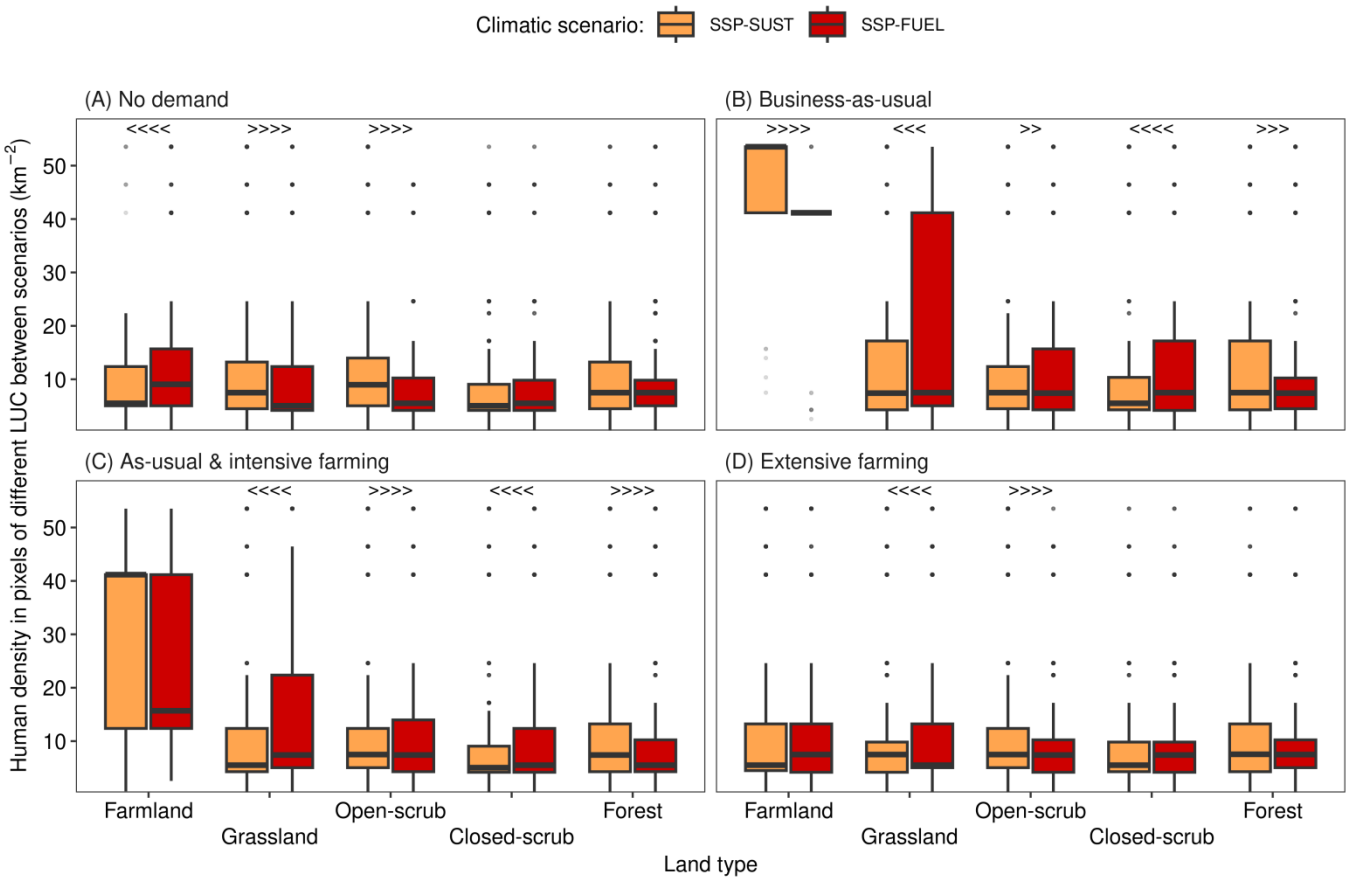

Fig. S23: Same as Fig. S21, but for population density.

Figure S24

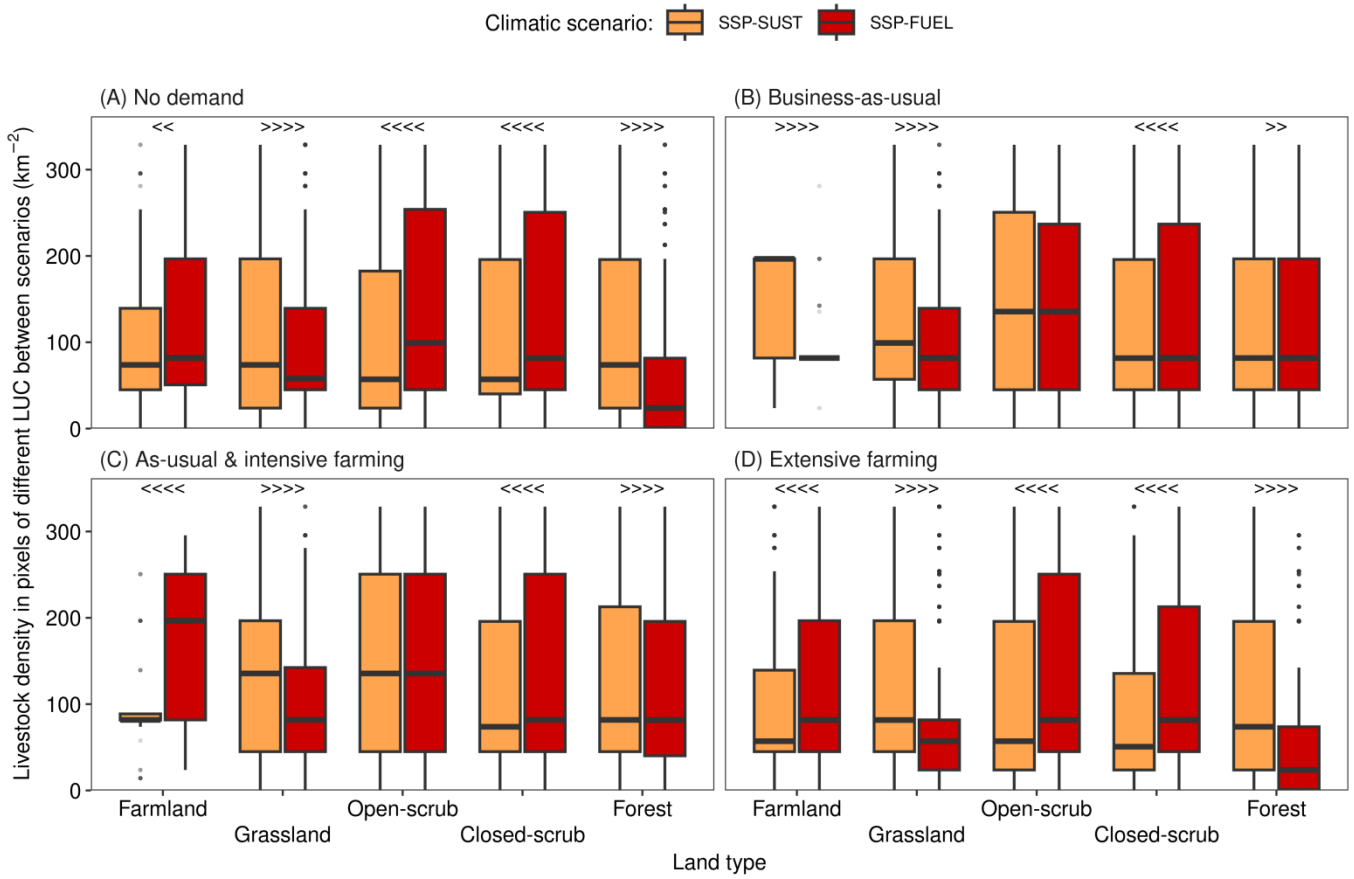

Fig. S24: Same as Fig. S21, but for livestock density.

Figure S25

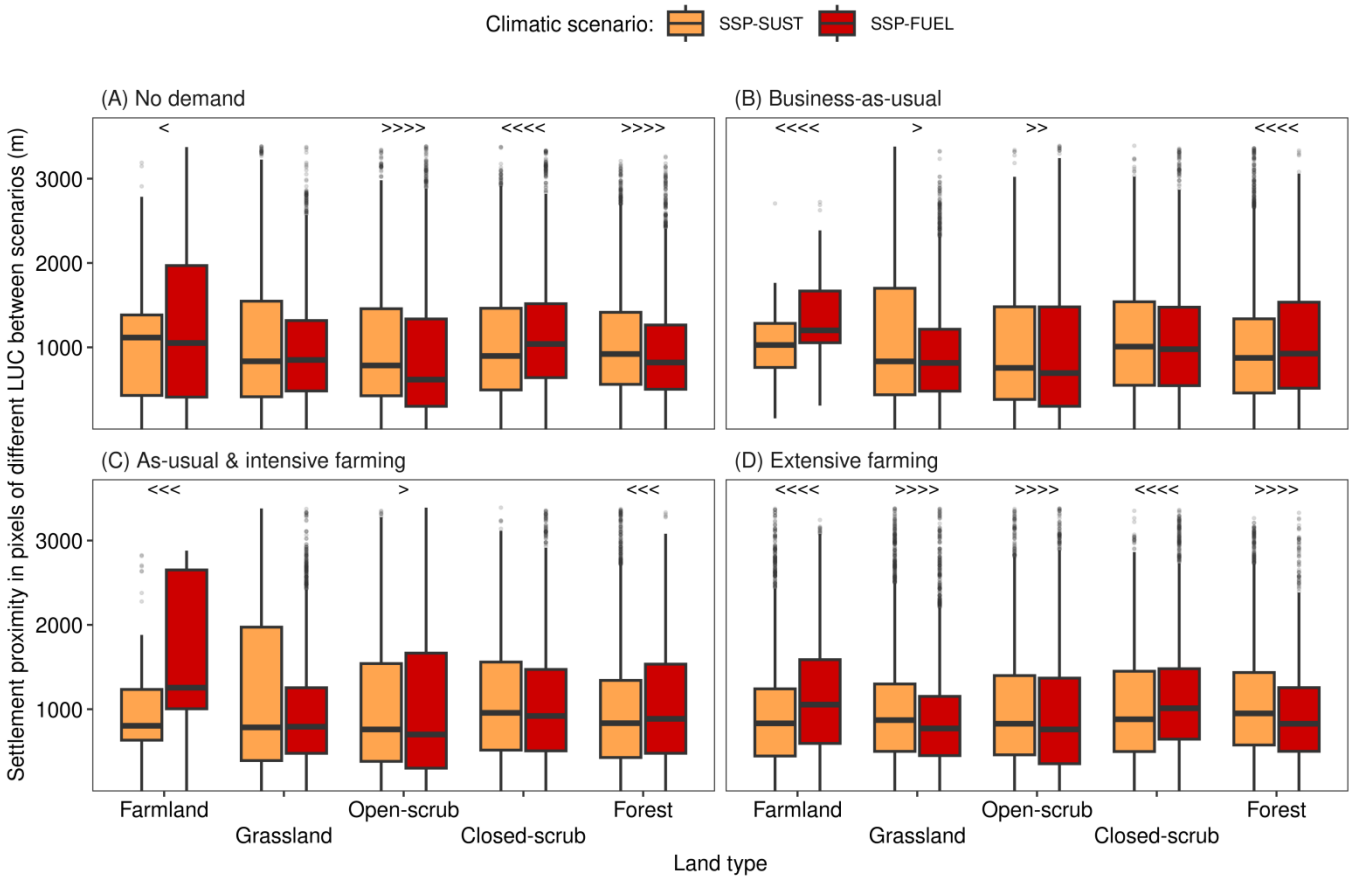

Fig. S25: Same as Fig. S21, but for settlement proximity.