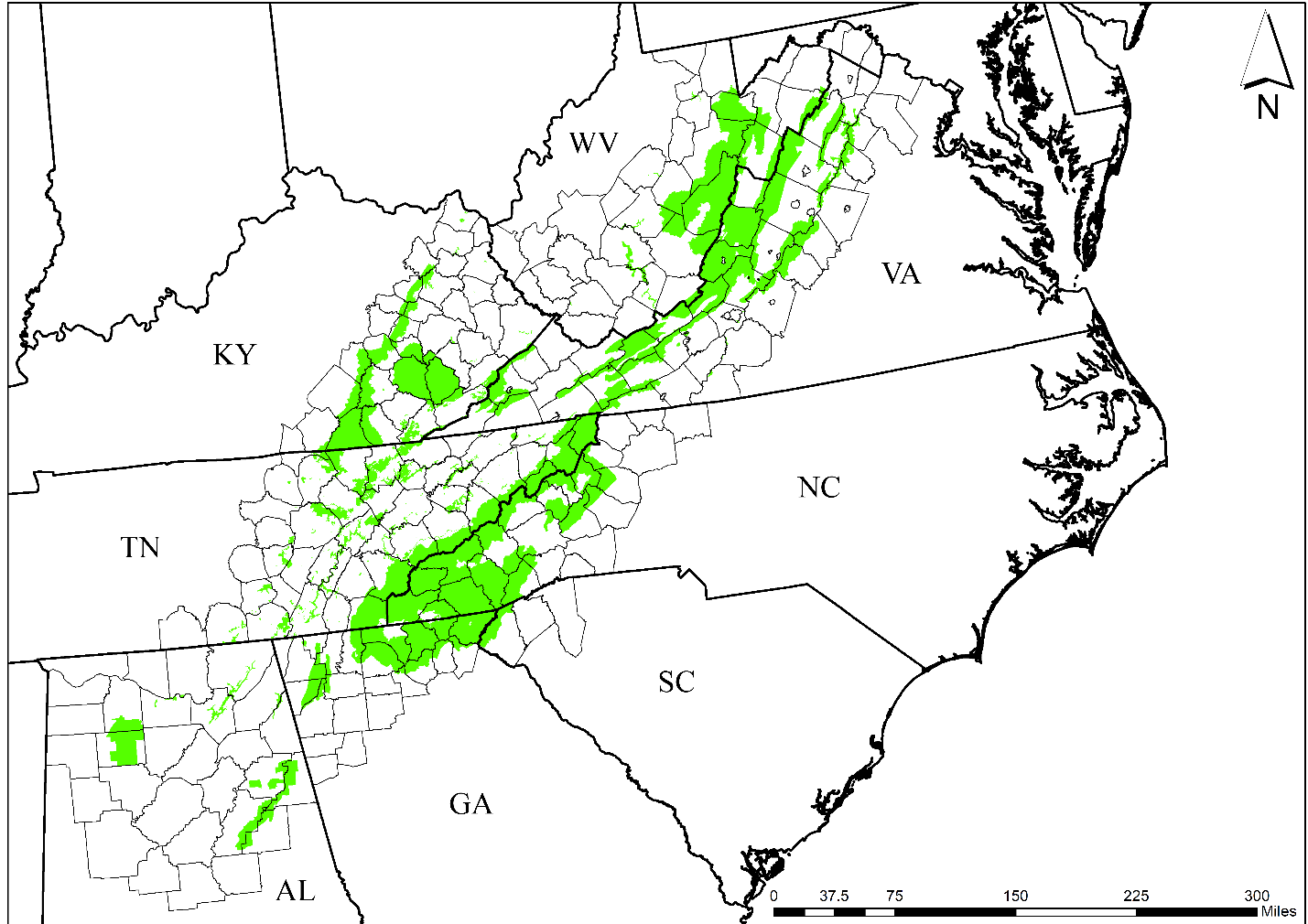
**Supplementary material**

**Text summary**

There are three figures and six texts in Supplementary material. Figure S1 shows the study area with 231 counties in the eight states of the Central and Southern Appalachian Region (USA). Figure S2 shows scatter plots between economic impact ($/$) and forest-dependent biodiversity ROIs (ha/$) with explicit cost or with relative opportunity cost. Figure S3 shows protected forestlands from the optimal budget distributions under assigned weight between forest-dependent biodiversity and economic impact of 100%–0% with explicit cost or relative opportunity cost. S1 contains brief description of the Maxent modeling and how aggregated suitable habitats are converted to accumulated species ranges. S2 contains the brief description of the IMPLAN. S3 contains brief description of how annual return from forestland is estimated. S4 contains the brief description of how annualized urban return is estimated. S5 contains brief description of the MINIMAX approach. S6 contains brief description of the sensitivity analysis of the main finding.



Permanently protected areas

**Figure S1** Study area with 231 counties in the eight states of the Central and Southern Appalachian Region of the United States

Note: Dotted line represent a fitted line

**Figure S2** Scatter plots between economic impact ($/$) and forest-dependent biodiversity ROIs (ha/$) with explicit cost or with relative opportunity cost

Explicit cost

Relative opportunity cost

Forest-dependent biodiversity: Economic impact = 100% : 0%

**Diagram, schematic

Description automatically generated**

**Diagram, schematic

Description automatically generatedText

Description automatically generated**

**Figure S3** Protected forestlands from the optimal budget distributions under the 100% assigned weight to forest-dependent biodiversity and 0% weight to economic impact with explicit cost or relative opportunity cost

**S1.** Maxent modeling is widely used in species niche and distribution modeling (Peterson et al. 2011), and it works by using presence-only species records with environmental variables (i.e., temperature and precipitation as GIS layers) as input to estimate the probability of each species’ presence at a pixel under given environmental conditions (Elith et al. 2011). While the climatically suitable habitats for species accounted for temperature and precipitation, they did not consider the ecological condition of the landscape.

The maximum entropy algorithm found the probability distribution of a species at a given pixel derived from the presence-only species records by satisfying any certain sets of constraints from the environmental variables among every probability distribution (i.e., Maxent model; Phillips & Dudík 2008). Then, the probability distribution of each species from the maximum entropy algorithm was converted to a binary prediction, which equals 1 if the species exists and 0 if the species does not exist at the pixel by setting a threshold for existence of species (i.e., 10th training threshold; see detail in Zhu et al. 2021). The binary suitability variables of species, representing suitability of habitats for 258 terrestrial vertebrates (including mammals, birds, reptiles and amphibians), were aggregated at the county level. The model focused on species of policy concern in the region by the U.S. Fish and Wildlife Service (2020), Landscape Conservation Cooperative Network ([2020](about:blank)), and USGS Science Analytics and Synthesis ([2020](about:blank)).

Given evidence that the ecological condition of the landscape depends on whether or not the landscape is protected (Armsworth et al. 2020), we assumed the probability that a species persists is a function of the proportion of the species range that is covered by the aggregated suitable habitats that are protected forest areas. One important simplification is to assume that the accumulated species range is an additive sum of aggregated suitable habitats on protected and unprotected forest areas. Here, we assumed 1 hectare of fully protected area in a species’ range makes the highest contribution to assisting its persistence and assign a relative weight of 1 to that 1 hectare, while we assigned a relative weight, (0 < < 1) to 1 hectare of unprotected land following Armsworth et al. (2020).

Formalizing these assumptions, we converted the aggregated suitable habitat for species *j* in county *i*, , to an accumulated species range for county *i*, *Hi* = , where , , and denote the total county area, protected forest area, and unprotected forest area (i.e., eligible forest area) of county *i*, respectively. The relative-weight parameter was assigned a value of 0.25 following Armsworth et al. (2020). In this equation, the additive sum of aggregated suitable habitats on protected and unprotected forest area was multiplied by the ratio of the aggregated suitable habitat to the total county area under the assumption that each hectare of protected forest area and one-quarter hectare of unprotected forest area sample the aggregated suitable habitat in proportion to the amount of the total county area. This sampling approach allows the suitable habitats for each species not to exceeding the forest area in the county.

**S2.** The IMPLAN is an Input–Output (I–O) model. The I–O model developed by Wassily Leontief plays an important role in quantitative economic techniques. It is used to estimate the interdependencies between various parts of a national or regional economy, and its main contribution is to show a consistent picture of the economic system (Ten Raa 2010). The economic impact generated by the I–O model measures changes of output, employment, labor income, and total value added. In addition, the economic impact is the sum of direct, indirect, and induced effects, which are summarized on a social accounting matrix. The direct impact is the direct infusion of money into the economy caused by economic events; the indirect impact represents the sum of inter-business sectors’ spending in the local economy (e.g., spending for buying intermediate goods), and the induced impact indicates household spending caused by the direct and indirect impacts (e.g., spending of households’ wages) (RVARC 2015).

**S3.** We obtained the area-weighted average of soil expectation value (SEV) per hectare between deciduous and evergreen forests at the county level based on stumpage price per hectare, timber volume per hectare, harvest rotation for each forest type, and a discount rate (Johnston & Williamson 2005). Stumpage prices are available in two parts of each state. We collected the data for the part in which the county is located from Timber Mart-South (2015) for AL, GA, NC, SC, TN, and VA, from Kentucky Division of Forestry (2015) for KY, and from West Virginia Division of Forestry (2015) for WV. We estimated timber volume per hectare for deciduous and evergreen forests at the county level by dividing standing timber volume from the Forest Inventory and Analysis (USDA Forest Service 2018) by county-respective area of forestland estimated from NLCD data for each forest type (USGS 2016). We used harvest rotations of 75 and 50 years for hardwood and softwood trees for deciduous and evergreen forests, respectively, from Smith et al. (2006).

The SEV for forest type *f* (i.e., deciduous and evergreen forest) in county *i* was annualized with a 5% discount rate and a 100-year horizon using where is the stumpage price and is the amount of harvested timber volume per hectare. Then, was multiplied by the ratio of forest type *f* relative to total forest area in county *i* to calculate the area-weighted that reflects the spatial ratios between the deciduous and evergreen forests within the county. Finally, the area-weighted *ASEV* for the two types of forests were summed at the county level to be used as the annual return from forestland for each county.

**S4.** First, we calculated land value ratios by dividing assessed land value by total assessed value at the parcel-level using data from tax assessors’ offices for a sample of counties (1-2 counties for each of the states where parcel-level data are available in our study area). The selection of the counties for the ratio data is simply based on data availability. While we have no control over which counties we choose, our sample data shows that the average land value ratio per hectare for each census-block group has wide disparity in the degree of urbanization with mean of 0.36, standard deviation of 0.86. We think because the ratio data collected is at the census-block group level, which is finer-grained data than the county-level data, the disparity in the degree of urbanization is well reflected in the ratio data. Second, we obtained land value ratios per hectare by dividing parcels’ land value ratios by their corresponding lot sizes. Third, we predicted the average land value ratio per hectare for each census-block group (CBG) by regressing land value ratios per hectare on respective CBG population densities and vectors of relevant distance variables (see a brief description of the regression results in the following paragraph). Fourth, we calculated the median assessed land value per hectare for each CBG as the product of its average land value ratio per hectare and median housing price. Fifth, we annualized the median assessed land values per hectare assuming a 100-year horizon and a 5% discount rate as annual return from urban use *ui*.

An OLS regression was implemented by regressing the natural log of the parcel-level land value ratio per hectare on the natural log of population density and a vector of distance variables. The regression model hypothesizes that a parcel’s land value per unit of land is higher relative to its single-family house’s value in more densely populated areas that are closer to the city center with its associated facilities. The adjusted R2 for the log-log OLS model is 0.61, reflecting reliable goodness-of-fit for the model. Results for the model are reported in the table below, which shows that parcels in neighborhoods with higher population density, that are closer to the nearest city center, park, hospital, high school, and highway, and farther from the nearest golf course have higher land value ratios per hectare. These findings are generally consistent with other studies that suggest land value ratios per unit of land are higher in more urbanized areas than in less urbanized areas (Albouy & Ehrlich 2012). The regression coefficients from the table and the respective census-block group data are used to estimate the average land value ratio per hectare for each census-block group.

|  |  |  |
| --- | --- | --- |
| Variable | Definition | Coefficient  (Std Err) |
| ln (Population density) | Population density for census-block group (population per square mile) | 0.121\*  (0.054) |
| ln (City center) | Distance to the center of the nearest city center with population more than 10,000 (mile) | -0.518\*  (0.053) |
| ln (Park) | Distance to the nearest park (mile) | -0.047\*  (0.012) |
| ln (Hospital) | Distance to the nearest hospital (mile) | -0.524\*  (0.055) |
| ln (Golf course) | Distance to the nearest golf course (mile) | 0.233\*  (0.050) |
| ln (High school) | Distance to the nearest high school (mile) | -0.290\*  (0.048) |
| ln (Highway) | Distance to the nearest interstate highway (mile) | -0.171\*  (0.030) |
| Intercept |  | -0.699  (0.386) |

\* indicates significance at the 5%. ln(·) is the natural log of the value in parenthesis and Std Err is the standard error.

**S5.** Our multi-objective optimization framework used the estimated accumulated species range *Hi* from Maxent modeling, the total value-added from IMPLAN, and the budget required to protect all eligible forestland within each county. Two optimization problems were solved, one assuming the total explicit cost of annual return from forestland and the other assuming the total relative opportunity cost as the conservation investment.

The MINIMAX approach finds the minimum percentage deviation among maximum percentage deviations that stand for how far the actual value differs from the target value for an assigned hypothetical weight. The model is expressed in equations (1)-(4).

(1)

subject to

(2)

(3)

(4)

where is the decision variable that denotes the ratio of forestland enrolled in the PES program relative to total eligible area in each county *i,* is the hypothetical weight for objective *k* (i.e., maximizing accumulated species range or maximizing total value-added), is the target value of objective *k* that indicates the sum of maximum attainable values of the accumulated species range *Hi* or total value-added *TVAi* from each county *i* under the single-objective optimization for one of the two objectives. The coefficient of objective means the maximum values of accumulated species range *Hi* for county *i* or total value-added *TVAi* for county *i* when the eligible forestland for each county is fully enrolled in the PES program. represents the required budget for enrolling all eligible forestland in each county *i*, and *B* indicates the hypothetical budget constraint (i.e., US$ 10 million). Equation (2) indicates that the sum of hypothetical weights between the objectives must equal 1, equation (3) constrains between 0 and 1, and equation (4) is the budget constraint.

The coefficient of variation (CV) of maximum attainable values of accumulated species range *Hi* (i.e., standard deviation of 0.71 million hectares divided by its mean of 1.36 million hectares) is 0.52. The CV of maximum attainable values of total value-added *TVAi* (i.e., standard deviation of US$ 0.23 million divided by its mean of US$ 0.23 million) is 0.10 using the explicit cost, whereas the same value (i.e., standard deviation of US$ 32.24 million divided by its mean of US$ 14.93 million) is 2.16 using the relative opportunity cost. The CV of *TVAi* is different from when the PES program uses the explicit cost to when it uses the relative opportunity cost as *TVAi* is obtained by multiplying economic multiplier and required budget, which is a direct function of two cost measures. Pearson’s correlation coefficient between the two maximum attainable values is 0.55 and 0.33, using the explicit cost and the relative opportunity cost, respectively.

**S6.** The optimal spatial budget distribution patterns between the two cost measures are likely sensitive to the relative weight assigned to unprotected forestland in the estimation of accumulated species range (i.e., = 0.25). Holding other things constant, a higher relative weight on unprotected area would be associated with a larger accumulated species range. Because unprotected forestland is significantly more positively correlated with the explicit cost (0.83) than with the relative opportunity cost (0.45), the increased accumulated species range associated with a higher alpha value would be larger for counties with higher explicit cost than those with higher relative opportunity cost. As a result, the additional optimal spatial budget distribution due to the higher alpha value would be greater for counties with higher explicit costs than for those with higher relative opportunity costs. Thus, the budget would be distributed more narrowly among counties when the cost measure is explicit cost than when it is relative opportunity cost. Consequently, the distributions of target counties would become less (more) divergent with higher (lower) alpha values.

**References**

Albouy D, Ehrlich G (2012) Metropolitan Land Values and Housing Productivity. National

Bureau of Economic Research Working Paper No. 18110. URL https://www.nber.org/system/files/working\_papers/w18110/revisions/w18110.rev1.pdf

Armsworth PR, Benefield AE, Dilkina B, Fovargue R, Jackson HB, Le Bouille D et al. (2020) Allocating resources for land protection using continuous optimization: biodiversity conservation in the United States. *Ecological Applications*.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical

explanation of MaxEnt for ecologists. *Diversity and distributions* 17: 43-57.

Johnston M, Williamson T (2005) Climate change implications for stand yields and soil expectation values: a northern Saskatchewan case study. *The Forestry Chronicle* 81:683- 690.

Kentucky Division of Forestry (2015) Kentucky Delivered Log Price Report. URL https://www.forestry.ky.gov.

Landscape Conservation Cooperative Network (2020) Appalachian. URL https://lccnetwork.org/lcc/Appalachian

Roanoke Valley-Alleghany Regional Commission (RVARC) (2015) Economic impact of the hotel Roanoke and conference center. URL https://rvarc.org/wp-content/uploads/2015/04/hotel-final-v6.pdf

Smith JE, Heath LS, Skog KE, Birdsey RA (2006) Methods for calculating forest ecosystem and

harvested carbon with standard estimates for forest types of the United States. General Technical Report 343. U.S. Department of Agriculture, Forest Service, Northeastern Research Station. URL https://www.fs.fed.us/ecosystemservices/pdf/estimates-forest-types.pdf

Ten Raa T (2010) Commodity and Sector Classifications in Linked Systems of National

Accounts. In Input–Output Economics: Theory And Applications: Featuring Asian Economies (pp. 17-24).

Timber Mart-South (2015) Timber Mart-South Supplemental Report. URL https://www.TimberMart-South.com

U.S. Department of Agriculture, Forest Service (USDA FS) (2018) Forest Inventory and

Analysis Databases (FIADB). URL https://apps.fs.usda.gov/fia/datamart/datamart\_excel.html

U.S. Fish and Wildlife Service (2020) Endangered Species Act. URL https://www.fws.gov/endangered/laws-policies/

U.S. Geological Survey (USGS) (2016) Multi-Resolution Land Characteristics Consortium

(MRLC): National Land Cover Database (NLCD). URL http://www.mrlc.gov/data

U.S. Geological Survey, Science Analytics and Synthesis (SAS) (2020) State Wildlife Action

Plans (SWAP). URL www1.usgs.gov/csas/swap

West Virginia Division of Forestry (2015) West Virginia Timber Price Report. URL https://www.wvforestry.com

Zhu G, Papeş M, Giam X, Cho SH, Armsworth PR (2021) Are protected areas well-

sited to support species in the future in a major climate refuge and corridor in the United States? *Biological Conservation* 255: 108982.