**Explanation and Interpretation of Species Tables**

*This document describes the variables provided on the excel spreadsheet as well as tools to interpret the data provided. The material presented here is also presented, for the most part, within the supplement sheets on the excel file.*

*Louis Iverson, Anantha Prasad, Matthew Peters, Stephen Matthews. 27 September 2019*

Landscape Change Research Group, Northern Institute of Applied Climate Science, Northern Research Station, USDA Forest Service, 359 Main Road, Delaware, OH 43015. Louis.iverson@usda.gov

*What can this table tell you?*

1. *How many species are present in this area (according to FIA)?*
   1. *The column N counts the species on this table. Some are present and some are modeled to have suitable habitat appear by 2100. Those with %CellFIA >0 are known to be present. Some others may be present but were not found on FIA plots.*
2. *What species are present in this area (according to FIA)?*
   1. *Look at Common Name or Scientific Name, those with %CellFIA >0 have FIA indication of presence.*
3. *How common is each species, when it is found, i.e., is it common but in particular habitats (according to FIA)?* 
   1. *FIAiv shows the average importance of the species, when it is found.*
4. *How abundant is each species, taken across the region of interest (according to FIA)?*
   1. *FIAsum shows sum of importance values across the area (*this valuecan be compared across areas even though those areas are unequal in area)*.*
   2. *Abund classifies FIAsum into: Abundant, Common, Rare, Absent*
5. *How much confidence do we have in the models?*
   1. *ModRel gives classes of model reliability: Low, Medium, High*
6. *What do the models suggest may happen to habitat suitability for the species by the year 2100, according to a low [or high] emissions scenario?*
   1. *ChngCl45 [or ChngCl85] gives a class for potential change under low [or high] emissions: Large Increase, Small Increase, No Change, Small Decrease, Large Decrease, Very Large Decrease, Unknown (can’t model it), or New Habitat (new habitat gained by 2100). Currently, we are on the path of high emissions.*
7. *According to literature, how adaptable are the species to the direct and indirect impacts of a changing climate?*
   1. *Adapt shows a score (1.7-8.5) and color code (Low-pink, Medium-yellow, High-green) of a compilation of modification factors that estimate the adaptability of the species.*
8. *How capable might the species be for coping with the changing climate, within the region of interest, under a low [or high] emissions scenario?*
   1. *Capabil45 [or Capabil85] gives an estimate of the species capability to cope, based on its change in habitat suitability (ChngCl45 or ChngCl85), its adaptability, and its abundance in the region of interest. If the species is abundant locally, it is assumed to also be in refugia and niches somewhat buffered from the climate impacts.*
9. *What species might be considered most vulnerable to the changing climate in this area?*
   1. *Those species listed under Capabil45 [or Capabil85] as Poor, Very Poor, or Lost will be the most vulnerable according to our assessments.*
10. *Natural migration of the species has been modeled with SHIFT – it predicts where the species may migrate within the next 100 years that is currently unoccupied by the species. Assuming a generous migration rate of ~ 50 km/century, where is the species likely to colonize?*
    1. *%OccCol (long file) is the sum of areas currently occupied and those areas projected to get occupied at >50% probability within 100 years.*
11. *What portion of the area has at least a 2% chance [>0%] of getting colonized over 100 years?*
    1. *%2Col (long file) estimates the percentage of the region of interest with at least 2% chance of colonization, and is not already occupied according to FIA. One cutoff we use for potential planting guidelines is that the species requires at least 5% of the region of interest to have >2% chance of colonized (color-coded dark green in the table).*
    2. *%AnyCol (long file) provides the percentage of the area, not already occupied according to FIA, that has >0% chance of getting colonized.*
12. *When we consider both the new suitable habitat (Habitat Quality) predicted from the DISTRIB2 models, and the chance of that new area getting colonized (Colonization Likelihood), what species may have the highest (or at least some) chance of migrating and finding suitable habitat within 100 years (under low and high emissions)?* 
    1. *HQCL45 (long file, low emissions) and HQCL85 (long file, high emissions) are indices that use a weighted average calculation to assign a score for the potential to migrate into suitable habitat – the higher the value, the greater the chance, with values > 1 indicating the presence of suitable and colonizable habitats.*

**Interpretation of tables**

**Assumption: you have the ‘color-coded’ excel file (Sxx-Exx-short) open of the area of interest, be it national forest (buffered) or the 1x1 degree grids.**

1. **Understanding the species and numbers of species present or potentially present:** The “**N**” variable counts the rows in the table, listing species within the area of interest, and includes (1)all native species currently occuring according to FIA, (2) non-native introduced species occuring according to FIA, (3) species not identified via FIA plots but modeled to occur, and (4) species modeled to provide suitable habitat for the area by 2100. As such, extraction and counts of subsets of species within these classes can easily be done via sorting on any particular column of interest. Species can also be presented and sorted by common name, scientific name, or FIA code.
2. **Identification of ranked order of current species:** when the **FIAsum** column is sorted largest to smallest, the table lists most dominant to least dominant species, according to the FIA inventory data. It is the sum, within the area of interest, of the importance value (IV) of the species (based equally on number of stems and basal area of those stems and standardized to 0-100) for each cell multiplied by the area of each cell occupied by that species. This number has also been calibrated to a standardized 10,000 km2 (2.47 million acres) area, the approximate area of a 1x1 degree cell, so that **FIAsum** values can be compared across areas even though those areas are unequal in area. This gives a current snapshot of the status of most of the species. However, with only 1 FIA plot per ~5000 acres, there will be occassional species missed by the plots; of course these are usually rare species. The **FIAsum** variable is also used to class the Abundance (**Abund**) into Abundant, Common, Rare, or Absent according to the cutoffs listed in the variable explanation document.
3. **Identification of species’ importance where present:** FIAi provides an indication of the average importance value (IV) of a species, on a 0-100 scale, for the cells that have the species. Therefore, certain species that may be uncommon across an area but abundant in certain locations may obtain higher scores than generally common species.
4. **Identification of species’ adaptabilty to cope with conditions likely under a changing climate:** The **Adapt** variable lists a score, ranging from 8.5 for red maple, the most adaptable species to 1.7 for black ash, the least adaptable species according to 21 modification factors. These scores were obtained via several literature sources, and are described in Matthews et al. 2011. It should be noted, however, that these scores were achieved via reviewing literature across the range for each species; there may be cases were these scores should be changed based on local knowledge.
5. **Idenfication of changes in suitable habitat according the the models:** the **ChngCl45** and **ChngCl85** variables provide, for RCP4.5 and RCP8.5 respectively, the potential change in suitable habitat for the species by year 2100. *It is important to note that a change in suitable habitat does not mean the species importance will actually change in that area by 2100, only that the habitat is expected to increase, decrease, or not in suitability for that species over time.*
6. **Identification of the capability of species with cope with the changing climate:** with the variables **Capabil45** (RCP4.5) and **Capabil85** (RCP8.5) we combine the influence of changes in suitability, adaptability, and abundance of the species within the area of interest. As such, this is our best estimate of the species’ ability to withstand the extremes and prolonged issues from climate change, *within the area of interest*. If the species is abundant in the area of interest, it has more resources to withstand disturbances and find refuge somewhere. In contrast, rarer species may have a higer probability to be pinched out under particular extreme events. We provide tables initally sorted by **Capabil45** so that species are ranked from Very Good to Lost. Following those classes on the table are species listed as ‘New Habitat’, ‘FIA only’, or ‘NNIS’, representing, respectively, those species that have New Habitat projected for the area of interest, those species that have statistically unacceptible models so we only report their current inventory information, and those species that are non-native invasive species (**NNIS**). Note: we also report, on the long form, an initial rating of capability that does not include the last factor of abundance in the area at present. For planting (see #1 below), it doesn’t really matter how abundant the species is presently.
7. **Identification of species with New Habitat but also some probability of colonization within 100 years:** With **Capabil45** or **Capabil85**, or **ChngCl45** or **ChngCl85**, we provide those species identified as New Habitat. With the SHIFT analysis, we provide the percent of the area with any chance of getting colonized (%AnyCol, long file), so, generously, we can say that the species has at least a small chance to naturally migrate to the New Habitat within 100 yrs. If these criteria are met, we suggest that the species has a relatively greater potential to succeed, as compared to those with 0 in %AnyCol, if the species were planted within the area of interest. Higher values of %AnyCol would also indicate a relatively higher chance to succeed. To subset these further, we suggest you especially consider those species with at least a 5% of the area with at least 2% probability of colonization (%2Col >5). If HQCL45 or HQCL85 (last 3 variables three long file only) are >0 of those species, there is even greater support for planting those species.
8. **Identification of species currently present but rare, and potentially suitable for expansion based on SHIFT model:** Some species are present but with a low percentage of occupany now or in the future across the area of interest. Given that most land managers have a preference for planting species in, or at least near, their natural range, these are species that may be candidates for planting to expand their position. We select those species with a %OccCol (long file) of less than 25% of the area. %OccCol is the sum of areas currently occupied (PerOcc) and those areas projected to get occupied at >50% probability within 100 years (Per50Col). We then can use the variables that combine habitat quality and colonization likelihood (HQCL45 and HQCL85) to assist in selecting potential species for planting. Sorting on %OccCol provides the list from >0.01 to 25.0; then those species with HQCL45 or HQCL85 showing values >0 will have both suitable habitat and some likelihood of colonization with higher values represented higher potentials for those species.

**Possible method to assist in selecting candidate species to plant (these are coded on ‘SSO, species selection option’ as 0-3 as follows.**

We emphasize that the following is one approach, among others, that decision makers must use to narrow the potential list of species to plant. We assume the manager would prefer to select species that are fairly well adapted to predicted future climatological conditions and that are already within, or at least close, to the range of the species. Our species models and information concerning adaptability were based on the species’ natural distributions across their entire range within the United States. Of course managers, especially those planting in parks or urban areas, will have additional criteria for selection based on factors such as form and safety.

1. **Select species present currently (often common or abundant) that are likely to cope with the changing climate:** for these species, we simply select among those with a Fair, Good, or Very Good capability to cope under low (Capability45, long file) or high (Capability85, long file) emissions. In this case, we use the initial Capability Class, prior to modification by the abundance in the area of interest. We assume that the current abundance of the species should not factor into the decisions to plant. Coded ‘1’ under **SSO (species selection option)**.
2. **Select species present currently, but less common (often rare and often with poor capability classes), yet are potentially in a position to expand over time:** for these species, we select those with an occupancy currently or with at least a 50% likelihood of colonization within 100 years (%OccCol, long file) of less than 25%, and then evaluate the combined habitat quality/colonization likelihood weighted score (either HQCL45 or HQCL85) for values >0. The higher the HQCL score, the higher chance the species is appropriate for planting. Keep also in mind the capability, adaptabiliy, change classes, and model reliability of these species as you narrow your search, as some may just have too many negative factors to consider planting. Coded ‘2’ under **SSO (species selection option)**. Most often coded "Infill + or Infill ++" under SHIFT45/SHIFT85.
3. **Select species not recorded (via FIA) currently, but with potential to migrate into the area of interest within 100 years:** these species are labeled New Habitat under **Capabil45** or **Capabil85**, and also show some chance (%0Col, long file) of colonization within 100 years. This is an optimistic view because SHIFT is using an optimistic rate of migration of 50 km/century and an optimistic long-distance potential of 500 km. To subset these further, we suggest you especially consider those species with at least a 5% of the area with at least 2% probability of colonization (%2Col >5). Coded ‘3’ under **SSO**, or even check out the hidden values for 5 or 20% colonization probability. Often coded "Migrate + or Migrate ++ under **SHIFT45/SHIFT85**.
4. **Select other species:** we emphasize that these analyses are only to be used as general guidelines for species selection. The models are built from FIA data across the eastern US, and for certain species, local influences (e.g., lake effects) will override the general tendencies across the entire eastern US. The models also necessarily are built from coarse-level data and are unable to zero in on special or rare habitats and may have low model reliability (ModRel). Therefore, please do not discard species from consideration if they do not show up of the three lists mentioned above, but use your local knowledge to select species that may be suited for particular niches in your project area. Coded ‘0’ under **SSO**.

**Variable Descriptions**

Following is a brief summary of each column on the species list pages. Each column on the excel spreadsheet can be sorted (click on header variable arrow) to gain maximum information extraction. General information regarding our modeling schemes can be found in Iverson et al. 2008, 2011; 2017; 2019; Matthews et al. 2011; Peters et al. 2015; 2019; Prasad et al. 2013; 2016. In the listing of column headers below, the **bolded** variables are visible on the short spreadsheet (Sxx\_Exx-short), while the non-bolded variables are only on long (Sxx\_Exx-long) spreadsheet. The order below follows the order of the headers on both spreadsheets, just that you won’t see the non-bolded variables on the short sheet. We advise staying with the short spreadsheet unless you really want to ‘deep-dive’ into the data providing information to the short spreadsheet.

**DISTRIB2 output variables**

**FIA**: a numeric code assigned according to US Forest Service Forest Inventory and Analysis (FIA) units, indicating the species is found or is projected to have suitable habitat under some climate change scenario by 2100, within at least 1 cella within in the area of interest.

**Common Name**: species common name used by FIA.

**Scientific Name**: species scientific name used by FIA.

**Rang:** Range code for Distribution (Wide vs. Narrow, first letter), Commonness (Dense vs. Sparse, second letter), and Importance (High IV vs. Low IV, third letter). Distribution was based on the percent of the Eastern US occupied by the species (Wide =>10% occupied, Narrow <10%). Commonness was based on the percent of 10x10 km cells with the species detected by FIA (Dense=>40%, Sparse <40%). Importance was based on the mean Importance Value where present (High=>5, Low <5). This code thus gives a quick evaluation of the nature of the species’ distribution throughout the eastern U.S. If there is an ‘X’ in the fourth position of the code, the species was so rare as to be unreliably modeled.

**MR**: the model reliability of the species’ model predicting current and future suitable habitat (High, Medium, Low), based on several statistical parametersb. If coded ‘FIA only’, the model is unacceptable for predicting into future, thus the change classes and capability classes are unknown.

**%Cell**: the percentage of cellsa within the 1x1 degree of Lat/Lon (or other area) zone that currently have the species present according to the FIA data. It does not mean the species actually covers that amount of ground within the sample area. Also, if the 1x1 is only a partial rectangle because of coast or water, or because of the region of interest, there may be pieces without FIA plots or missing environmental predictors, therefore small fractions of the 1x1 without certain common species. Also, the region could be selecting only a small fraction of a cell with the species present.

act\_ncell (long file): number of cellsa occupied according to FIA data. Because FIA doesn’t sample everywhere (1 plot per ~5000 acres), this number is usually less than the modeled current value (mod\_ncell).

mod\_ncell (long file): number of cellsa predicted currently according to modeled output.

ccsm45\_ncell (long file): number of cellsa predicted by 2100 according to modeled output using CCSM4 model and RCP4.5.

had85\_ncell (long file): number of cellsa predicted by 2100 according to modeled output using Hadley model and RCP8.5.

gcm45\_ncell (long file): number of cellsa predicted by 2100 according to modeled output using combined CCSM4, Hadley, and GFDL models and RCP4.5.

gcm85\_ncell (long file): number of cellsa predicted by 2100 according to modeled output using combined CCSM4, Hadley, and GFDL models and RCP8.5.

act\_sumIV (long file): the area-weighted sum of the importance value (IV) according to FIA records for the species within the area of interest. These values have not been corrected for a standardized area, as is FIA sum.

**FIAsum**: the area-weighted sum of the importance value (IV) according to FIA records for the species within cellsa within the area of interest. These values have been corrected for partial 1x1 degree zones (to 10,000 km2), and for varying sizes north to south (curvature of earth makes zones narrower towards the poles) according their proportion of a full 1x1 degree zone at mid latitudes.

**FIAiv**: the average importance value (IV) according to FIA records for the species, where it occurs, within the 10x10 and/or 20x20 km cellsa of the area of interest. The 0-100 score is based on number of stems and basal area of tree species recorded during the most recent FIA inventory cycle (FIA data used from across the eastern US range from 2001-2016). IVs are averaged at the cella level to indicate the average abundance of a species. Note that this IV is for the species only within the cells where it now occurs, not averaged over the entire region of interest.

MODi (long file): the area-weighted average importance value (IV) according to a Random Forest model for the species within cellsa. This includes cellsa that have modeled suitable habitat for the species, so that FIA plots need not be present for the model to predict species presence in the smaller cellsa.

CCSM45i (long file): the area-weighted average importance value (IV) according to a Random Forest model for the species within cellsa, under the Representative Concentration Pathway (RCP) 4.5 (relatively low emission future) for the CCSM4 general circulation modelc by 2100. This model scenario represents a relatively mild change in climate conditions as compared to current.

HAD85i (long file): the area-weighted average importance value (IV) according to a Random Forest model for the species within cellsa, under the Representative Concentration Pathway (RCP) 8.5 (relatively high emission future) for the Hadley general circulation modelc by 2100. This model scenario represents a relatively large change in climate conditions as compared to current.

G45i or G85i (long file): the area-weighted average importance value (IV) according to a Random Forest model for the species within cellsa, under the Representative Concentration Pathway (RCP) 4.5 (relatively low emission future) or 8.5 (high emission pathway) average of 3 general circulation modelsc (GCMs) by 2100. The 0-100 score is based on number of stems and basal area.

CCSM45r (long file): the ratio of future (2070-2099) suitable habitat (=CCSM45i) to current (1981-2010) habitat (=MODi), so that a ratio of 1 indicates no change in suitable habitat, <1 indicates a potential loss in habitat, and >1 indicates a potential gain in habitat by 2100 according to the mildest model scenario, CCSM45.

HAD85r (long file): the ratio of future (2070-2099) suitable habitat (=Had85i) to current (1981-2010) habitat (=MODi), so that a ratio of 1 indicates no change in suitable habitat, <1 indicates a potential loss in habitat, and >1 indicates a potential gain in habitat by 2100 according to the harshest model scenario, HAD85.

G45r or G85r (long file): the ratio of future (2070-2099) suitable habitat (=G45i/G85i) to current (1981-2010) habitat (=MODi), so that a ratio of 1 indicates no change in suitable habitat, <1 indicates a potential loss in habitat, and >1 indicates a potential gain in habitat by 2100 according to the lower (or higher) emissions scenario, average of 3 GCMs.

**ChngCl45** or **ChngCl85**: class of potential change in habitat suitability by 2100 according to the ratios of future (2070-2099) suitable habitat for an average of 3 GCMs to current (1981-2010) modeled habitat, using either lower emissions (RCP 4.5) or higher emissions (RCP 8.5). Classes follow ratios as shown in the table below, for common species. When MODi was >0 for <10% of the cellsa within the 1x1 degree zone, the species was categorized as ‘rare’ and the change classes were expanded, so that NoChange=ratio 0.6-4.0; Sm. inc.=ratio 4.0-8.0; Lg. inc.=ratio >8.0; Sm. dec.=ratio 0.2-0.6; Lg. dec.=ratio <0.2.

|  |  |  |
| --- | --- | --- |
| **Class** | **Criteria** | **Description** |
| NoChange | Ratio of 0.8 – 1.2 | No Change to very little change in habitat suitability |
| Sm. inc. | Ratio of 1.2 – 2.0 | Small increase in habitat suitability |
| Lg. inc. | Ratio > 2.0 | Large increase in habitat suitability |
| Sm. dec. | Ratio of 0.5 – 0.8 | Small decrease in habitat suitability |
| Lg. dec. | Ratio < 0.5 | Large decrease in habitat suitability |
| Very Lg. dec. | FIAi > 0.5 and GCM45i = 0 | Very large decrease in habitat suitability |
| Unknown | ModRel = unacceptable | Not enough information to assign a class due to very low model reliability or species is extremely rare either now or potentially in future. |
| New Habitat | MODi = 0 and GCM45i > 0 | Species is not currently reported by FIA and has not been modeled under current conditions to have suitable habitat, but could gain suitable habitat under future climate conditions. |

**Adap**: Adaptability score for the species, according to a literature review of 12 disturbance and 9 biological characteristics, or modification factors (Iverson et al. 2011; Matthews et al. 2011). Scores range from 1.7 (very low adaptability to a changing climate) to 8.5 (very high adaptability), with color class breaks at 3.4 and 5.2. Thus low adaptability is 1.7-3.4, medium is 3.5-5.2, and high is >5.2.

**Abund:** based on FIAsum within a 1x1 degree (or other) zone, the abundance of the species based on last FIA inventory cycle, corrected for 1x1 degree zone size as described above. We realize these are subjective cutoffs, but does have merit based on multiple evaluations. Users could change cutoffs to meet their needs.

|  |  |
| --- | --- |
| **Class** | **Criteria** |
| Abundant | FIAsum > 75 |
| Common | 5 ≤ FIAsum ≤ 75 |
| Rare | 0 < FIAsum < 5 |
| Modeled | FIAsum = 0 and MODi > 0 |
| Absent | FIAsum = 0 and MODi = 0 |

Capability45 or Capability85 (long file): See next variable for explanation. Used for coding as a potential planting species, if the classes are coded as Very Good, Good, or Fair.

**Capabil45** or **Capabil85**: the overall estimate of capability for the species to cope with the changing climate within the study area. Capability class is based on the Change Classes (potential changes in suitable habitat by end of century), which in turn are based on the ratio of future to current (see above) across three GCMs at RCP 4.5 or 8.5 and the Adaptability of the species to the added disturbances likely under climate change. Classes range from Very Good to Good to Fair to Poor to Very Poor to Lost (no habitat under GCM85i) to Unknown (Unknown ChngCl45/ChngCl85, see ChngCl45 description above) to New Habitat (see ChngCl45 description above) to NNIS (if the species is not native), to FIA only (if the species only is recorded in FIA inventories but not sufficient for modeling. First, the ChngCl45/ChngCl85 is considered, followed by Adapt to arrive at an initial capability. First, the ChngCl45/ChngCl85 is considered, followed by Adapt to arrive at an initial capability (long file as variable Capability45 or Capability85. This is the value used for coding ‘1’ under the species selection option (**SSO**). Then Abundance is used to modify classes so that if the species is Abundant, we increase capability by one class (e.g., poor to fair, or good to very good); if species is Rare, we decrease capability by one class (e.g., poor to very poor); if species is Common, there is no change in capability. The idea is that common species are more likely to find refugia into the future, and rare species are less likely.

Table below. Initial capability class ratings defined as colored classes, depending on the change class and adaptability of the species, and other classes for selected species. For final Capability rating (Very Good to Very Poor): if abundance was ‘abundant’, move up one class; if ‘rare’ move down one class; if ‘common’ stay in class.



**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**SHIFT output variables**

SHIFT applies to natural colonization into previously unoccupied habitats for each species. First of all, DISTRIB habitat model predicts future abundances as though there is no limit to migration. We therefore call them "habitats" instead of "distribution" or "abundances". One of the criticisms of species distribution models is that they assume no migration limit. Therefore, the role of SHIFT is to provide migration limits to DISTRIB (we call it adding a dose of reality to the predicted future habitats) - i.e., what is the likelihood that the future habitats predicted by DISTRIB will be colonized?

SHIFT calculates the colonization likelihood based on the FIA abundances and generation times of the species (among other things) to match the future predicted habitats of DISTRIB. Here is where Habitat Quality and Colonization Likelihood (HQ-CL) becomes relevant. HQ-CL represents areas that combine habitat quality with colonization likelihoods - i.e., where in the forested eastern US landscape will suitable combinations of HQ-CL be found (obviously we want high Habitat Quality and high Colonization Likelihood - but these combinations may be limited. However, we can identify the best HQ-CL combinations available for any species and target them for management (relocation etc.) if needed since they in turn can contribute propagules to assist in further migration.

Another point is the value of SHIFT in selecting potential species for planting. Because DISTRIB delineates only suitable habitat, it could show unlikely delineations for some species due to model related artifacts. We believe it prudent to plant those species that are relatively closer to their locations of potential migration naturally within 100 years (e.g., 1 seed zone away). We have taken a very generous approach to selecting species, first noting those with even a fraction (>0) of probability of colonization within 100 years, then noting that it preferable to select those with at least 5% of the area with at least 2% probability of colonization. This process narrows the list, and discourages the selection of species that are way, way out of current range limits.

**SHIFT45 or SHIFT85:** A combined classification, for RCP4.5 or 8.5, separating out those species that may (1) infill, where a species is currently recorded to be present via FIA and likely to spread out within the zone; (2) be likely, where a species was not reported by FIA but models suggest it is likely to be present; and (3) migrate, where a species is not reported or modeled to be in the zone, but has some potential to migrate there within 100 years - and could be considered as a high candidate for translocation. See long definition for details. Each label has a '+' = lower or '++' = higher modifier based on the strength of the signal.

**SSO:** Species Selection Options code. These codes may help screen species for potential planting or otherwise promoting in the region. If the SSO code is 1, the species is currently present in the region and has Fair, Good, or Very Good Capability to cope under low or high emissions; thus the models suggest it should do ok into the future. If SSO code is 2, the species is rare or close by in the region but does have a good chance of spreading within the region especially if it has a high HQCL score (see long definition). If SSO code is 3, the species is not recorded from the region via FIA plots, but does have some chance of getting colonized naturally within 100 years. If SSO code is 0, though not recommended according to our models, there will be many other reasons to evaluate for potential to plant or otherwise encourage in the area, based on local ecology and desires. More details for each code are found above under the section “**Possible method to assist in selecting candidate species to plant (these are coded on ‘SSO, species selection option’ as 0-3 as follows”.**

**N**: the count of species considered in the study area.

TotPerPotCol (long file): Percent of total study area that is potentially colonizable (i.e., the percent of the area that is not water or is not currently occupied, nor will it have any chance of colonization in next 100 years (PerNever), plus the percentage that has at least some (>0%) chance of getting colonized (%0Col)).

PerNever (long file): Percent of study area that is not currently occupied, nor will it have any chance of colonization in next 100 years.

%OcCol (long file): Percent of study area that is occupied added to areas that have some chance (>0%) of getting colonized.

PerOcc (long file): Percent of study area that is currently occupied according to FIA data.

Per50Col (long file): Percent of study area that has at least 50% chance of getting colonized within 100 years.

%2Col (long file): Percent of study area that has at least 2% chance of getting colonized within 100 years.

%0Col (long file): Percent of study area that has more than 0% chance of getting colonized within 100 years.

SumHQCL45 or SumHQCL85 (long file): the count of 1km x 1km cells (from SHIFT) that have at least some chance of colonization within 100 years. These counts are used as multipliers for the weighted scores described next.

HQCL45 orHQCL85: Habitat quality and colonization likelihood weighted score. This index is derived by combining the outputs of DISTRIB2 and SHIFT to get at desirable combinations of both habitat quality and colonization likelihood. The score is based on a cells’ habitat quality class of low (IV=1-5), medium (IV=6-15), high (16-100) as well as colonization likelihood probability of low (1-10%), medium (11-50%), and high (51-100%). The weighted score provides more value to locations with both higher levels of suitable habitat and higher levels of colonization likelihood according to the following table:

|  |  |  |
| --- | --- | --- |
| Habitat Quality (HQ) | Colonization Likelihood (CL) | Multiplier |
| Low (1) | Low (1) | 2 |
| Low (1) | Medium (2) | 3 |
| Low (1) | High (3) | 4 |
| Medium (2) | Low (1) | 5 |
| High (3) | Low (1) | 6 |
| Medium (2) | Medium (2) | 7 |
| Medium (2) | High (3) | 8 |
| High (3) | Medium (2) | 9 |
| High (3) | High (3) | 10 |

HQCL45\_11 or HQCL85\_11 (long file): Cell counts summed from the following combination: (HQ of ‘low’ and CL of ‘low’), reclassified as described under HQCL45. This score, and the four additional variables provided next provide information if one wants to probe further into the contributions of HQ and CL in the HQCL45 or HQCL85 scores.

HQCL45\_12 or HQCL85\_12 (long file): Cell counts summed from the following combination: (HQ of ‘low’ and CL of ‘medium’) plus (HQ of ‘low’ and CL of ‘high’), and reclassified as described under HQCL45.

HQCL45\_21 or HQCL85\_21 (long file): Cell counts summed from the following combination: (HQ of ‘medium’ and CL of ‘low’) plus (HQ of ‘high and CL of ‘low’), and reclassified as described under HQCL45.

HQCL45\_22 or HQCL85\_22 (long file): Cell counts summed from the following combination: (HQ of ‘medium’ and CL of ‘medium’) plus (HQ of ‘medium’ and CL of ‘high’), and reclassified as described under HQCL45.

HQCL45\_33 or HQCL85\_33 (long file): Cell counts summed from the following combination: (HQ of ‘high’ and CL of ‘medium’) plus (HQ of ‘high’ and CL of ‘high’), and reclassified as described under HQCL45.

SppNum (long file): FIA code.

aThe eastern US was carved into a hybrid grid of 10x10 or 20x20 km cells, based on the quantity of FIA inventory plots. In the western portion of the eastern US, the original vegetation was primarily prairie (now cropland), and forests are much less common. Therefore, these areas need larger cells in order to capture forest trends sufficiently for modeling, e.g., two or more FIA plots within a cell.

bMultiple factors were used to generate the ModRel classification, including the pseudo R2 (scaled) of the Random Forest imputed model, the Fuzzy Kappa, the percent deviance explained by the top 5 variables, the stability of the top 5 variables using 30 regression tree analysis outputs, and the total sum of squares (TSS) after removing records with very high coefficient of variation. Then if score = .23-.54 (low ModRel); .55-.69 (medium ModRel); >.7 (high ModRel).

cGeneral Circulation Models considered included the CCSM4 (Gent et al. 2011), GFDL CM3 (Donner et al. 2011), and HadGEM2-ES (Jones et al. 2011) models under the representative concentration pathways (RCP) (Moss et al. 2008) 4.5 and 8.5 for a historical period of 1981 – 2010 (PRISM) and modeled projections for the period 2011 – 2099. The following table presents characteristics of these scenarios (PANN=annual precipitation; Pgrow=precipitation during April-Sept; TANN=mean annual temperature; Tgrow=mean temperature April-Sept; TWIN=mean temperature Dec-Feb; TSUM=mean temperature June-August; Aridity Index=Aridity; TMIN=absolute minimum temperature; TMAX=absolute maximum temperature.



**Literature Cited**

Donner, L. J., B. L. Wyman, R. S. Hemler, L. W. Horowitz, Y. Ming, M. Zhao, J.-C. Golaz, P. Ginoux, S. J. Lin, M. D. Schwarzkopf, J. Austin, G. Alaka, W. F. Cooke, T. L. Delworth, S. M. Freidenreich, C. T. Gordon, S. M. Griffies, I. M. Held, W. J. Hurlin, S. A. Klein, T. R. Knutson, A. R. Langenhorst, H.-C. Lee, Y. Lin, B. I. Magi, S. L. Malyshev, P. C. D. Milly, V. Naik, M. J. Nath, R. Pincus, J. J. Ploshay, V. Ramaswamy, C. J. Seman, E. Shevliakova, J. J. Sirutis, W. F. Stern, R. J. Stouffer, R. J. Wilson, M. Winton, A. T. Wittenberg, and F. Zeng. 2011. The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. Journal of Climate **24**:3484-3519.

Gent, P. R., G. Danabasoglu, L. J. Donner, M. M. Holland, E. C. Hunke, S. R. Jayne, D. M. Lawrence, R. B. Neale, P. J. Rasch, M. Vertenstein, P. H. Worley, Z.-L. Yang, and M. Zhang. 2011. The Community Climate System Model Version 4. Journal of Climate **24**:4973-4991.

Iverson, L. R., A. M. Prasad, S. N. Matthews, and M. Peters. 2008. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. Forest Ecology and Management 254:390-406.

Iverson, L., A. M. Prasad, S. Matthews, and M. Peters. 2011. Lessons learned while integrating habitat, dispersal, disturbance, and life-history traits into species habitat models under climate change Ecosystems 14:1005-1020.

Iverson, L. R., M. P. Peters, A. M. Prasad, and S. N. Matthews. 2019. Analysis of climate change impacts on tree species of the eastern US: Results of DISTRIB-II modeling. Forests 10:302.

Iverson, L. R., F. R. Thompson, S. Matthews, M. Peters, A. Prasad, W. D. Dijak, J. Fraser, W. J. Wang, B. Hanberry, H. He, M. Janowiak, P. Butler, L. Brandt, and C. Swanston. 2017. Multi-model comparison on the effects of climate change on tree species in the eastern U.S.: results from an enhanced niche model and process-based ecosystem and landscape models. Landscape Ecology 32:1327–1346.

Jones, C. D., J. K. Hughes, N. Bellouin, S. C. Hardiman, G. S. Jones, J. Knight, S. Liddicoat, F. M. O'Connor, R. J. Andres, C. Bell, K. O. Boo, A. Bozzo, N. Butchart, P. Cadule, K. D. Corbin, M. Doutriaux-Boucher, P. Friedlingstein, J. Gornall, L. Gray, P. R. Halloran, G. Hurtt, W. J. Ingram, J. F. Lamarque, R. M. Law, M. Meinshausen, S. Osprey, E. J. Palin, L. Parsons Chini, T. Raddatz, M. G. Sanderson, A. A. Sellar, A. Schurer, P. Valdes, N. Wood, S. Woodward, M. Yoshioka, and M. Zerroukat. 2011. The HadGEM2-ES implementation of CMIP5 centennial simulations. Geosci. Model Dev. **4**:543-570.

Matthews, S. N., L. R. Iverson, A. M. Prasad, M. P. Peters, and P. G. Rodewald. 2011. Modifying climate change habitat models using tree species-specific assessments of model uncertainty and life history factors. Forest Ecology and Management 262:1460-1472.

Moss, R., W. Babiker, S. Brinkman, E. Calvo, T. Carter, J. Edmonds, I. Elgizouli, S. Emori, L. Erda, and K. Hibbard. 2008. Towards New Scenarios for the Analysis of Emissions, Climate Change, Impacts, and Response Strategies. Technical Summary 978-92-9169-124-1, Intergovernmental Panel on Climate Change, Geneva.

Peters, M. P., L. R. Iverson, and S. N. Matthews. 2015. Long-term droughtiness and drought tolerance of eastern US forests over five decades. Forest Ecology and Management 345:56-64.

Peters, M., L. Iverson, A. Prasad, and S. Matthews. 2019. Utilizing the density of inventory samples to define a hybrid lattice for a macro-level species distribution model. Ecology and Evolution 9:8876-8899.

Prasad, A. M., J. Gardiner, L. Iverson, S. Matthews, and M. Peters. 2013. Exploring tree species colonization potentials using a spatially explicit simulation model: implications for four oaks under climate change. Global Change Biology 19: 2196–2208.

Prasad, A. M., L. R. Iverson, S. N. Matthews, and M. P. Peters. 2016. A multistage decision support framework to guide tree species management under climate change via habitat suitability and colonization models, and a knowledge-based scoring system. Landscape Ecology 31:2187-2204.