WILDFIRE RISK TO COMMUNITIES

Interactive maps, charts, and data to help communities understand, explore, and reduce wildfire risk.

www.fs.usda.gov/wildfirerisk

Wildfire Risk to Communities 2.0:

Updated methods for geospatial datasets for populated areas in the United States

A white paper included with:

Jaffe, Melissa R.; Scott, Joe H.; Callahan, Michael N.; Dillon, Gregory K.; Karau, Eva C.; Lazarz, Mitchell T. 2024. Wildfire Risk to Communities: Spatial datasets of wildfire risk for populated areas in the United States. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2020-0060-2

May 15, 2024

TABLE OF CONTENTS

INTRODUCTION

The Wildfire Risk to Communities project (WRC) was created in response to direction by the U.S. Congress in the 2018 Consolidated Appropriations Act (i.e., 2018 Omnibus Act, H.R. 1625, Section 210: Wildfire Hazard Severity Mapping). That legislation directed the USDA Forest Service to develop and publish, within two years, national geospatial products depicting wildfire hazard and risk for communities across the United States. The focus of the legislation was firmly on communities. The intent was to help U.S. communities understand components of their relative wildfire risk profile, the nature and effects of wildfire risk, and actions they can take to mitigate risk.

To meet the intent of the Omnibus Act, the Forest Service formed a team of experts to develop the necessary data and build a website for effective delivery of information to communities. The team consisted of wildfire analysts from the Fire Modeling Institute (FMI), part of the Forest Service's Rocky Mountain Research Station (RMRS), and wildfire modeling and geospatial data experts at Pyrologix, LLC. A non-profit partner, Headwaters Economics, also became a critical player in developing the public-facing website with interactive maps and charts, and clear communication targeted to local government officials and private citizens who could take actions to mitigate risks in their communities.

The result of those initial efforts was the Wildfire Risk to Communities website (www.wildfirerisk.org) that was launched in April 2020. The data products published in that initial rollout were built on the nationwide wildfire hazard data from Short et al. (2020), and they represented the first time wildfire risk to communities had been mapped nationally with consistent methodology down to the level of individual communities (Scott et al. 2020a, 2020b). The data provided foundational information for comparing the relative wildfire risk among populated communities in the United States.

Data Update

Wildfire hazard reflects the overall likelihood and probable intensity of wildfire, and it is not static through time. It is influenced by the condition of vegetation and other burnable materials (i.e., fuels) and by temperature, moisture, and other dynamic elements of the fire environment. When the first version of WRC data was released in 2020, the wildfire hazard data reflected ground conditions as of the end of 2014. This lag of six years reflects the enormous amount of work involved in updating vegetation and fuel maps across the United States and the running wildfire simulation modeling for such a large area. It was always our intent that those initial data products would be the first of many versions, with data updates as frequently as possible and practical.

This second release of WRC data (WRC 2.0) represents the first update to our initial data products. The methods used to generate the hazard and risk data for this update differ in many ways from those used in the initial 2020 release (WRC 1.0). In addition to the differences in methodology and new input data used for mapping wildfire hazard across all lands (detailed in the companion white paper with Scott et al. 2024), many changes and updated data sources were used to create the new populated areas data. Important differences in the populated areas data from version 1.0 to version 2.0 include:

- Building footprint data used for WRC 1.0 came primarily from the 2018 Microsoft Building Footprint dataset. For WRC 2.0, we used a custom building footprint dataset produced by combining the public USA Structures building footprints[1] with a commercially-available dataset from ONEGEO[2], both reflecting 2022 conditions.

- U.S. Census Bureau population and housing data used in WRC 1.0 came from the 2018 American Community Survey (ACS) and Populated Areas Program (PEP) datasets. For WRC 2.0 we used an updated 2021 vintage of both those datasets and we also leveraged the 2020 Decennial Census Redistricting data (described in more detail below).
- The populated areas data publication for WRC 2.0 includes three new raster datasets not included in WRC 1.0: Building Count, Housing Count, and Population Count.

The purpose of this white paper is to provide detailed descriptions of the methods used in updating the spatial datasets for populated areas for the WRC 2.0 data release. This includes methods for: 1) producing spatial datasets of populated areas, and 2) creating spatial datasets that represent wildfire exposure and risk to populated areas in the United States. There are two companion papers: one that describes methods for landscape-wide characteristics of wildfire hazard and risk (Scott et al. 2024), and one that details the process we used to summarize those data for communities in the United States and delineate Community Wildfire Risk Reduction Zones (Dillon et al. 2024).

DATA AND METHODS

Input Datasets

The input datasets used to produce the updated populated areas data products include datasets related to population, building sizes and locations, land cover, and wildfire hazard (probability and intensity). We describe the sources of those input datasets in the following sections.

Population – U.S. Census Bureau 2021 5-year American Community Survey

The American Community Survey (ACS) is an ongoing annual survey conducted by the U.S. Census Bureau to collect updated estimates of U.S. social, economic, housing, and demographic data (U.S. Census Bureau 2022). The Census Bureau contacts over 3.5 million households every year as part of the survey. From this sample of households, the Census Bureau then produces statistical estimates of a range of population and housing characteristics at geographic subdivisions down to the Census Block Group level. ACS creates rolling 5-year estimates from the previous 60 months of collected responses covering all areas. We used Block Group population estimates for 2021 published in the 2021 5-year ACS data products (U.S. Census Bureau 2022). They are based on data collected from January 1, 2016 to December 31, 2020.

Population – U.S. Census Bureau 2021 Population Estimates Program

The U.S. Census Bureau also produces estimates every year of the population for the nation, states, and counties as part of the Population Estimates Program (PEP)[3]. This program measures population

[1] Oak Ridge National Laboratory (ORNL) and U.S. Federal Emergency Management Agency (FEMA) Geospatial Response Center. 2022. USA Structures. https://gis-fema.hub.arcgis.com/pages/usa-structures. Downloaded December 2022.

[2] ONEGEO. 2022. 3D Building Footprints. https://onegeo.co/. Acquired January 2023.

[3] https://www.census.gov/programs-surveys/popest.html

changes through public records of births, deaths, and migration and updates population counts annually for a continuous time series between Decennial Census counts (U.S. Census Bureau 2021). Estimates are released every year on July 1 of the current year, referred to as the vintage year.

We used county level population estimates for the vintage year 2021 that capture changes from April 1, 2020 to July 1, 2021 (U.S. Census Bureau 2021)

Population – U.S. Census Bureau 2020 Decennial Census

We used population data from the Census Bureau from the 2020 Decennial Census at the Census Block level[4].

Population & Housing Units – U.S. Census Bureau 2020 Decennial Census Redistricting Data

In WRC 1.0, we estimated population and housing unit counts at the Census Block level through indirect methods. We first estimated population at the Census Block Group level from ACS and PEP data, then distributed that population evenly across building footprints to get numbers of population-per-building. To get numbers of housing units in a Census Block, we then divided the population in a Census Block (derived by summing the population-per-building values in each Block) by published County averages of persons per housing unit.

For WRC 2.0 we were able to use direct counts of housing units by Block, published in the Decennial Census Redistricting Dataset[5]. The Redistricting data include an Occupancy Status dataset that records the number of occupied, vacant, and total housing units by Census Block. We used the total number of housing units by Block (vacant and non-vacant) as our housing unit count. However, this dataset did not include group housing both institutional and non-institutional.

To ensure representation of populations living in group quarters we used group quarters population information in the redistricting data to supplement Block-level population from the 2020 Decennial Census. The Group Quarters Population by Major Population Type recorded the population in institutionalized populations (i.e. correctional facilities for adults, juvenile facilities, nursing facilities/skilled-nursing facilities) and noninstitutionalized populations (i.e. college/university student housing and military quarters). We used the total population of both institutionalized and noninstitutionalized populations from the redistricting data.

[4] U.S. Census Bureau. 2020. 2020 Census: P.L. 94-171 Redistricting Data Total Population, accessed March 23, 2023. https://www.nhgis.org/tabular-data-sources#2020

[5] H1OCCUPANCY STATUS 2020:DEC Redistricting Data (PL 94-171) https://data.census.gov/table/DECENNIALPL2020.H1?q=&y=2020&d=DEC%20Redistricting%20Data%20(PL%2094-171)

Census Dataset	2021 American Community Survey (ACS)	2021 Population Estimates Program (PEP)	2020 Decennial Census Redistricting Data			Current Population Survey (CPS)
Census Sub dataset	**5-yr estimates**	**2021 PEP**	**Total Population**	**Occupancy Status**	**Group Quarters Population by Major Population Type**	**Annual Social and Economic Supplement – Historical Households Table**
Summary Level used in WRC 2.0	Block Group	County	Block	Block	Block	National
Area Covered	Covers all areas	Covers all areas	Covers all areas	Covers all areas	Covers all areas	Covers all areas
Data Periodicity	5-yr*	1-yr	10 yrs	10 yrs	10 yrs	1-yr
Data Collection Period	1/1/2016-12/31/2020	4/1/2020-7/21/2021	2020	2020	2020	2022
How was the data used in WRC 2.0?	Improve currency of population estimates by adjusting Decennial data at the Block Group level	Improve currency of population estimates by adjusting Decennial data at the county level	Enable downscaling of ACS population from Block Group to Block level	Calculate Housing Unit count	Supplement Occupancy Status data to include group quarters in Housing Unit Count	Acquire estimate of people per household to compute population per housing unit for group quarters

*ACS 5-year estimates are released each year but are calculated as the average of the previous 5 years of data.

Building Footprints

3DBuildings—We acquired 3DBuildings from ONEGEO[6] to represent building footprints as of 2022. This is a commercially available building footprint dataset covering all 50 states and the District of Columbia.

USA Structures—We acquired a publicly available building footprint dataset for all 50 states and the District of Columbia from USA Structures[7], a collaboration of Oak Ridge National Laboratory (ORNL) and the Federal Emergency Management Agency (FEMA). This dataset is limited to buildings with a

[6] https://onegeo.co/. Dataset acquired December 14, 2022.
[7] https://gis-fema.hub.arcgis.com/pages/usa-structures. Dataset acquired January 12, 2023.

calculated footprint area of 450 ft^2 (41.8 m^2) or greater. Like the 3DBuildings dataset, the USA Structures dataset represents building footprints as of 2022, however a simple analysis showed that the USA Structures dataset contains buildings not present in the 3DBuildings dataset.

Integrated Building Footprints—Each of the source building footprint datasets had benefits and drawbacks; neither dataset on its own appeared to fully represent the full extent of building locations across the country. We therefore made an integrated building footprint (IBF) dataset designed to reduce the number of buildings that would be missed if either source were used on its own. To make the IBF dataset, we started with all buildings in the 3DBuildings dataset and added buildings from the USA Structures dataset whose footprints did not intersect with a 3DBuildings footprint.

Qualifying Buildings—Another drawback to building footprint datasets is the presence of false-positives – footprint polygons that do not actually represent buildings. Often these are very small polygons that may reflect rocks, shadows, and other imagery features mistaken for buildings. To reduce the prevalence of these in our data, and to also keep our focus on primary residential, commercial, and industrial buildings (as opposed to sheds and other small outbuildings), we eliminated buildings with a footprint area smaller than 40 m^2 (430 ft^2). As this was below the minimum size for the USA Structures data (41.8 m^2 or 450 ft^2), this only eliminated small building footprints coming from the 3DBuildings dataset. We also eliminated a building footprint if its centroid fell on a pixel of uninhabitable land cover (i.e., open water and permanent snow/ice; see below). We refer to the resulting, filtered version of the IBF dataset as the qualifying building footprints (QBF) dataset.

Protected Areas

In addition to the filtering of buildings for our QBF dataset, we also used a protected areas mask to try to remove more false positives from our population and housing unit products. In doing so, we assumed that very few people live inside of legally designated Wilderness or Roadless areas and created our protected areas mask from the following datasets: Department of Interior wilderness areas (Craig Thompson, DOI Office of Wildland Fire, personal communication), and USDA Forest Service Wilderness Area boundaries, Other National Designated-Area boundaries, and Roadless Area boundaries. We acquired all datasets in September 2023.

We considered Wilderness and Wilderness Study Areas as a first step in filtering out false positives. We merged all Wilderness and Wilderness Study Area polygons into a combined dataset. We visually inspected these polygons with aerial imagery available in ArcGIS and found that some valid housing units exist directly on the edge of designated areas. To prevent losing those buildings when converting the polygons to a raster mask, we first did an inverse buffer of 200m (i.e. shrunk the polygons inwards by 200m). We then removed any “buildings” from inside the shrunken mask.

Next, we visually inspected Roadless Areas with aerial imagery and found that some states appeared to have valid buildings (possibly housing units) inside Roadless areas. Therefore, we subjectively chose specific states where we felt also removing buildings from inside Roadless Areas would be beneficial. The states where we removed “buildings” from Roadless Areas were: Arizona, California, Colorado, Florida, Idaho, Kentucky, Montana, Nevada, New Mexico, North Dakota, Oregon, Pennsylvania, South Dakota, Texas, Utah, Washington, and Wyoming.

We used the protected areas mask to filter the building locations used to create the population and housing unit data products. However, because real buildings that are not housing units may exist in protected areas, we did not use the mask to filter buildings used in the building rasters (Building Count, Building Density, and Building Cover).

Land Cover – LANDFIRE Fire Behavior Fuel Models

We used the LANDFIRE 2.2.0 (LANDFIRE 2020) Scott and Burgan Fire Behavior Fuel Model (FBFM40) raster dataset to identify burnable vs non-burnable and habitable vs uninhabitable land covers. The FBFM40 dataset represents the primary vegetation layer likely to carry fire, which is different than a typical land cover map that represents the dominant overstory. It is used as a primary input to the fire behavior modeling that generates burn probability and flame-length probability rasters, so using it to also represent land cover maintains logical consistency between the WRC datasets. The data cover three different extents; each extent used a different spatial reference (see table below).

Spatial domain	LANDFIRE version		Projection
Conterminous U.S. (CONUS)	2.2.0	LF 2020	Albers CONUS
Alaska (AK)	2.2.0	LF 2020	Albers AK
Hawaii (HI)	2.2.0	LF 2020	Albers HI

Non-burnable land cover was defined as areas mapped by LANDFIRE as any of the non-burnable fuel models in the FBFM40 raster: urban (91), permanent snow/ice (92), non-burnable agriculture (93), open water (98) and bare ground (99) (Scott and Burgan 2005). We considered everything else burnable land cover. We used both 30-m and 270-m resolution versions of the burnable land cover raster at different stages of data processing for Wildfire Risk to Communities.

We defined habitable land cover as all land cover types except open water and permanent snow/ice.

Burn Probability

The burn probability (BP) dataset used as input to the products described here is a 30-m raster representing the *circa* 2021 annual likelihood of wildfire occurrence in a given location. This 30-m BP raster was generated specifically for the Wildfire Risk to Communities project using a multi-stage raster geoprocessing-based resampling and smoothing process (Scott et al. 2024) applied to the most recent nationwide 270-m BP results from Dillon et al. (2023). That process involved spatially "upsampling" the 270-m data to 30-m resolution, then "oozing" burn probability values 1530 m (approximately 1 mile) into developed areas mapped as non-burnable in LANDFIRE fuels data. Details on the process are in the companion white paper included with the WRC 2.0 landscape-wide data products (Scott et al. 2024).

Conditional Risk to Potential Structures (cRPS)

Some populated-area WRC outputs are based on the Conditional Risk to Potential Structures (cRPS) raster, which represents the potential consequences of fire to a home or other structure at a given location if a fire were to occur there and if a home were located there. It is a measure that integrates the expected range of wildfire intensities with generalized consequences to a structure on every pixel but does not account for the annual probability of fire occurrence. It is analogous to conditional Net Value Change (cNVC) described by Scott and Thompson (2015). cRPS is referred to as Wildfire Consequence in the Wildfire Risk to Communities web application.

Scott et al. (2024) calculated the initial cRPS raster at 30-m by applying one of three response functions representing the relative effect of wildfire on structures (i.e., relative degree of damage or loss) at different intensities to 30-m flame-length probability (FLP) rasters. Response functions were developed for three lifeforms separately: grass/herbaceous, shrub, and tree, and reflect the assumption that consequence is greatest in tree fuels, lower in shrubs, and lowest in grass fuels, across all intensity levels. A value of 0 means no damage to a structure, and a value of -100 means complete loss. Scott et al. (2024) applied the response function to all pixels across the landscape, even if no structures are present.

The response function values used were:

Lifeform	FLP1	FLP2	FLP3	FLP4	FLP5	FLP6
Tree	-25	-40	-55	-70	-85	-100
Shrub	-20	-35	-50	-65	-80	-95
Grass	-10	-25	-40	-55	-70	-85

The final cRPS raster was created by using a modified version of the oozing approach mentioned above for BP that oozes the same distance but does not decay the cRPS values.

The 30-m FLP rasters used in creating the cRPS were produced using different input fuels data and a different modeling process than what was used for the BP raster. For generating the FLP rasters, Scott et al. (2024) used fuels data representing conditions *circa* 2023 and a FlamMap-based modeling process. Pairing more current, and finer-resolution, intensity data (*circa* 2023) with slightly older (*circa* 2021) burn probability provided a viable solution for producing the most current national risk maps possible with available data and technology.

Exposure Type

One populated-area WRC output required use of the Exposure Type raster from Scott et al. (2024). Exposure Type characterizes the way that a structure could be exposed to wildfire with values ranging from 0 to 1. Where the underlying land cover is considered burnable in the LANDFIRE fuels data, the value of the Exposure Type raster is 1 indicating pixels where a home would be "directly exposed" to wildfire. Where land cover is non-burnable developed, agricultural, or bare ground and the upsampled and oozed BP is non-zero (i.e., within approximately 1 mile of a 500-ha contiguous area of burnable vegetation), homes would be "indirectly exposed" to wildfire. The value of Exposure Type in these areas is between 1 and 0, varying by distance to burnable fuels, with pixel values decreasing toward 0 as they get further from burnable fuel. Finally, where the land cover is non-burnable and the upsampled and smoothed BP is zero, the value of the Exposure Type raster is 0 indicating pixels where a home would have little-to-no exposure to wildfire due to its distance from a large contiguous patch of burnable vegetation.

Methods for Specific Raster Datasets

For this release of the Wildfire Risk to Communities data, we produced a series of eleven nationwide raster datasets covering CONUS, Alaska, and Hawaii. All rasters have 30-m spatial resolution. Specifics for each dataset are described in the sections that follow.

Building Count

Building Count is a 30-m raster representing the count of qualifying building centroids in the QBF dataset located within each 30-m pixel.

The Building Count dataset is the raster equivalent of a building-point feature class. It is produced to facilitate data analysis (summarizing building count for any geography) rather than map display (Building Density or Building Cover is preferred for map display). Building Count raster values are integers representing the number of qualifying building centroids within a 30-m pixel.

Building Density

Building Density is a 30-m raster representing the density of qualifying building centroids in the QBF dataset (buildings/km^2). We generated the Building Density raster using the Building Count raster and land cover data from LANDFIRE. Our approach to generating the Building Density raster generally follows the methods used by the Westwide Wildfire Risk Assessment to produce the Where People Live raster (WPL; Sanborn Map Company 2016), although Building Density represents all qualifying buildings as opposed to just housing units in the WPL dataset.

We generated the smoothed Building Density raster using a four-step process:

1. Calculate a 200-m radius moving-window sum of the 30-m Building Count raster;
2. Calculate a 200-m radius moving-window sum of habitable land cover (in km^2);
3. Divide the sum-of-Building Count raster produced in step 1 by the sum-of-habitable-land-cover raster produced in step 2 to generate building density in buildings/km^2.
4. Convert the raster values to integers. Values less than 0.25 buildings/km^2 were set to zero. Values between 0.25 and 1 were set to a value of 1 building/km^2. All other values were rounded to the nearest integer.

Building Coverage

Building Coverage is a 30-m raster depicting the percentage of habitable land area covered by qualified building footprints (QBF). We generated the Building Coverage raster using the QBF dataset and land cover data from LANDFIRE. The primary steps in developing our Building Coverage raster are described below.

1. *Estimate building footprint area*

   We first divided building footprints in the IBF dataset into two categories: small buildings and large buildings. The intention behind this was to reduce computational demands stemming from the sheer number of small buildings while not losing spatial distribution of larger building footprints across multiple pixels.

   Small buildings

   We defined a small building as any building footprint in the QBF covering an area less than or equal to 900 m^2 (the size of one 30-m pixel) and mapped its building footprint area (m^2) to the 30-m pixel its centroid falls within. We then summed building footprint area for all remaining small building centroids within a pixel to obtain the footprint contribution from small buildings.

   Large buildings

   We defined a large building as a building footprint covering more than 900 m^2 (i.e., more than one pixel). We wanted to allocate building footprint area appropriately for any given building shape over 900 m^2, so we rasterized the building polygons to 30-m resolution, snapped to the LANDFIRE FBFM40 raster, and removed any building points located on water or ice as determined by the LANDFIRE FBFM40 dataset.

2. *Generate the Building Coverage raster*

To generate the Building Coverage raster, we first calculated the sum of the small and large building footprint contributions for a given 30-m pixel. We then generated the smoothed Building Cover raster using a three-step process:

1. Calculate a 75-m radius moving-window sum of the small-and-large-combined 30-m building footprint area raster on habitable land cover (m^2);
2. Calculate a 75-m radius moving-window sum of habitable land area (m^2);
3. Divide the smoothed building-area raster produced in step 1 by the smoothed habitable land cover raster produced in step 2 to generate proportion of building coverage.

The moving window consists of 21, 30-m pixels (a 5x5 matrix with the four corner cells omitted), which covers approximately 4.7 acres (1.9 ha). The values in the raster from step 3 represent the proportion of habitable land cover in the approximately 5-acre area surrounding each pixel covered by a building footprint.

Finally, we created the Building Coverage raster by multiplying the building cover proportions (0 to 1) by 100 and rounding to the nearest integer. This converted proportions to percentages with whole number values. Values less than or equal to 1 percent were set to zero.

Building Exposure

For WRC 1.0, we produced a Building Exposure Type raster to characterize the spatial coincidence of wildfire hazard with where buildings are located. That raster delineated whether buildings at each pixel would be directly exposed to wildfire from adjacent wildland vegetation, indirectly exposed to wildfire from indirect sources such as embers and home-to-home ignition, or unlikely to be exposed to wildfire due to distance from direct and indirect ignition sources.

While developing the framework and methodology for WRC 2.0 datasets, we decided to build on the concept of building exposure and create a product that enables communities to spatially identify where different types of risk mitigation activities are likely to be most effective. That product, the Community Wildfire Risk Reduction Zones (CWiRRZ) raster, is available as a stand-alone dataset, detailed in a companion white paper (Dillon et al. 2024).

Population Count (PopCount)

Population Count (PopCount) is a 30-m raster with pixel values representing residential population count (persons) in each pixel. It is produced to facilitate data analysis (summarizing population count for any geography) rather than map display (Population Density is preferred for map display). We generated the PopCount raster using both the PEP and ACS population estimates for 2021, population data from the 2020 Decennial Census, and building footprint data from Qualifying Building Footprint dataset filtered by protected areas. The primary steps are defined below.

1. *Estimate 2021 Census Block population count*

   We estimated 2021 population count for a Census Block using two approaches—one based on 2021 5-year ACS data and another based on 2021 PEP data. The final estimated population count for a given Block was taken to be the larger of the two estimates. PEP and ACS estimate population at different times of the year and with different definitions of residency—using the larger of ACS and PEP avoids the potential for undercounting population due to these differences.

   2021 5-year ACS

The 2021 5-year ACS population data are estimates resolved only to the Block Group, not to the Block. With initial testing, we found this was inadequate to allocate residential population spatially across the landscape because Blocks within a Block Group can have buildings but not residential population (for example, Census Blocks that consist entirely of industrial and commercial). We downscaled the Block Group population to the Block level by allocating the Block Group population to its Blocks in proportion to the Decennial Census proportions. For example, if a particular Block accounted for 10% of the Decennial Census population of the Block Group to which it belongs, then we allocated 10% of the ACS Block Group population to that Block. This approach maintains the ACS population within a Block Group but also allocates that population to individual Blocks. This approach also prevents allocating population to commercial or industrial Blocks with little or no residential population.

2021 PEP

In a second approach, we leveraged 2021 PEP data (by county) to increase (or decrease) the Decennial Census Block population by the fractional population change for the county containing the Block.

Final 2021 Census Block population

We estimated the final 2021 population for a Census Block as the larger of the 2021 5-year ACS and 2021 PEP-adjusted Decennial Census population estimates. This approach was used to capture seasonal residences present in in the ACS data, and group quarters (prisons, dormitories, etc.) which are captured in the Decennial Census.

2. *Allocate population to building points*

We calculated area of each building footprint polygon (in square meters) from the QBF dataset filtered by protected areas and converted to point features (footprint centroids).

For each Census Block nationwide, we divided the 2021 adjusted Census Block population count obtained in the first step by the count of remaining building points in the Block. This ratio is the effective population count per building point—the population represented by each building point. This became a Block level population-per-building-centroid attribute in the building points dataset. As such, every building centroid in a Block has the same value for this attribute.

Blocks with buildings but no population

The QBF dataset includes buildings within Census Blocks with a population count of zero. These footprints usually represent commercial or industrial buildings, but they can also represent false positives in the building footprint dataset. Because there is no population in these Blocks, their population-per-building values are all zero and they do not contribute to the final Population Count raster. There are also cases where housing units exist in the Housing Unit rasters, but where Population raster values are zero. One explanation for these situations is that although housing units were detected in the Decennial Census Block counts, the American Community Survey received no response in those blocks (perhaps because they are occupied only seasonally or newly built), so no population was tallied.

3. *Generate the Population Count raster from the building points*

To produce the PopCount raster, we converted the building points to a 30-m raster where the raster value is the sum of the population-per-centroid attribute of all building centroids within each raster grid cell. The final Population Count raster is the effective population per building point multiplied by the number of building points falling within the pixel.

Population Density (PopDen)

Population Density (PopDen) is a 30-m raster of residential population density (people/km^2) as of 2021. We generated the PopDen raster using the Population Count raster and land cover data from LANDFIRE. Our approach to generating the population density raster generally follows the methods used by the Westwide Wildfire Risk Assessment to produce the Where People Live raster (WPL; Sanborn Map Company 2016), although PopDen represents population as opposed to housing units in the WPL dataset.

We generated the smoothed PopDen raster using a four-step process:

1. Calculate a 200-m radius moving-window sum of the 30-m PopCount raster;
2. Calculate a 200-m radius moving-window sum of habitable land cover (in km^2);
3. Divide the smoothed PopCount raster produced in step 1 by the smoothed habitable land cover raster produced in step 2 to generate population density in people/km^2.
4. Convert the PopDen raster values to integers. Values less than 0.25 people/km^2 were set to zero. Values between 0.25 and 1 were given a value of 1 person/km^2. All other values were rounded to the nearest integer.

Blocks with population but no buildings

There are 1,235 cases where a Census Block has a population count but no buildings. We refer to these as “missing Blocks” because we are missing buildings to which we can allocate the population to. This can occur for several reasons. In the Decennial Census, people experiencing homelessness are counted in the Census Block where they reside at the time of the Census, even if that is not within a building. Sometimes that occurs in Census Blocks with no buildings. In other cases, the geography of a Census Block does not contain any buildings, but it may yet be assigned the population of a nearby building, erroneously.

To distribute population within a missing Block, we first estimated the mean population density across the Block as the Census Block population count divided by the habitable land cover within the Block (all but water and permanent snow/ice). If the mean population density was greater than 1 person per 100 acres, we allocated the population count uniformly across the habitable land cover in the Block before generating the final PopDen raster. If the mean population density was less than or equal to 1 person per 100 acres, we reset the population density to zero. Setting this minimum mean density was necessary to avoid including large areas of very low population density in the final PopDen raster.

Housing Unit Count (HUCount)

Housing Unit Count (HUCount) is a 30-m raster representing the number of housing units in each pixel. It is produced to facilitate data analysis (summarizing housing unit count for any geography) rather than map display (Housing Unit Density is preferred for map display). We generated HUCount from the U.S. Census Bureau Redistricting data from 2020 and the building footprint data from the Qualifying Building Footprint Dataset filtered by protected areas. The primary steps are defined below:

1. *Allocate Housing Unit data to building points*

   We acquired the number of housing units in each census block from the Occupancy Status table of the 2020 Redistricting dataset. We then divided the reported housing units in each Block by the total number of buildings in the filtered QBF dataset to calculate the housing unit count per building centroid attribute.

2. *Add in Group Populations*

   The Occupancy Status table does not include populations that live in group quarters, both institutional (i.e. prisons, juvenile detention centers, and skilled nursing facilities) or noninstitutional (i.e. military housing, college/university student housing). To represent group populations, we used the Group Quarters Population by Major Population Type table of the 2020 Redistricting dataset. We converted to housing units by dividing by 2.5 persons per housing unit, which was the average population per housing unit as listed in the Annual Social and Economic Supplement to the Current Population Survey[8]. We then summed total group housing units in a Block and divided by the total number of buildings in the Block.

3. *Generate Housing Unit Count Raster*

   To produce the HUCount raster, we converted the building points to a 30-m raster where the raster value is the sum of the housing-unit-per-centroid attribute of all building centroids within each raster grid cell. We did this individually for both the redistricted housing unit data and the group population housing unit data. Then we summed both raster layers to create the HUCount raster.

Cell values in the HUCount raster were not converted to integers and represent fractions of housing units. For example, there are cases where a HUCount value is a fraction less than one. Though a fractional housing unit count is illogical, we maintained the continuous format of the values to avoid eliminating buildings in Blocks with very low population.

Housing Unit Density (HUDen)

Housing Unit Density (HUDen) is a 30-m raster of housing-unit density (housing units/km^2) created from the Housing Unit Count raster.

We created the smoothed HUDen raster using a four-step process:

1. Calculate a 200-m radius moving-window sum of the 30-m housing-unit count raster;
2. Calculate a 200-m radius moving-window sum of habitable land cover (in km^2);
3. Divide the smoothed housing-unit count raster produced in step 1 by the smoothed habitable land cover raster produced in step 2 to generate housing-unit density in housing units/km^2.
4. Convert the HUDen raster values to integers. Values less than or equal to 0.1 HU/km^2 were set to zero. Values between 0.1 and 1 were given a value of 1 HU/km^2. All other values were rounded to the nearest integer.

Housing Unit Exposure (HUExposure)

Housing Unit Exposure (HUExposure) is a 30-m raster that depicts the expected annual number of housing units within a pixel potentially exposed to wildfire. This is a long-term annual average and

[8] https://www2.census.gov/programs-surveys/demo/tables/families/time-series/households/hh6.xls

does not represent the actual number of housing units exposed in any specific year. It is calculated as the product of wildfire likelihood and housing unit count.

We calculated HUExposure using the 30-m burn probability (BP) raster from Scott et al. (2024) and the HUDen raster. It is necessary to convert HUDen to housing unit count by multiplying density (HU/km^2) by 0.0009 (km^2/pixel) to get housing units per pixel. To enhance map display, we used this HUDen calculation rather than the HUCount raster described above because it includes the 200-m radius smoothed footprint around the HUCount pixel. The equation for housing unit exposure, then, is simply:

$$HUExposure = (HUDen * 0.0009) * BP$$

Values of HUExposure are floating point decimal numbers that represent the annual average housing units exposed per pixel, over the 20,000+ annual simulation iterations (Dillon et al. 2023). They provide a relative measure of the mean annual number of homes that could be exposed to wildfire. Pixels where the Housing Unit Density raster is zero are NoData in the HUExposure raster.

Housing Unit Impact (HUImpact)

HUImpact (HUImpact) is a 30-m raster that represents the relative potential impact of fire to housing units at any pixel, if a fire were to occur. It is an index that incorporates the general consequences of fire on a home as a function of fire intensity and uses flame length probabilities from wildfire modeling to capture likely intensity of fire. HUImpact does not include the likelihood of fire occurring, and it does not reflect mitigations done to individual structures that would influence susceptibility. It is conceptually similar to Conditional Risk to Potential Structures (cRPS) in Scott et al. (2024), but also incorporates housing unit count and exposure type.

To create HUImpact, we first converted HUDen to housing unit count using the same method and logic applied for HUExposure (see above). This involved multiplying HUDen by 0.0009 to get the housing unit count per pixel while maintaining the smoothed spatial footprint of the HUDen raster. We then multiplied housing unit count per pixel by the cRPS raster produced by Scott et al. (2024). This raster captures the conditional probability of different fire intensities if a fire were to occur, and the relative consequences to a structure at those intensities (see details under Input Datasets above). Next, we multiplied by the Exposure Type raster produced in Scott et al. (2024) to mimic the reduction of potential losses with distance from burnable land cover. Where buildings are directly exposed to wildfire, the value of Exposure Type is 1 and HUImpact values are unmodified. Where buildings are indirectly exposed to wildfire, the value of Building Exposure is between 0 and 1, so HUImpact is reduced with increasing distance from areas of direct exposure. We calculate HUImpact with the following equation:

$$HUImpact = (HUDen * 0.0009) * cRPS * ExposureType$$

Calculated values of HUImpact can be very small decimal numbers. Because HUImpact is a unitless index, we chose to convert those to integers while preserving the relative values and as much precision as possible. To produce the final integer version of the HUImpact raster, we multiplied the initial values by 1,000,000 and rounded to the nearest integer. Pixels where the HUDen raster is zero are NoData in the HUImpact raster.

It is important to note that by using a consistent response function for all homes, we assume that all homes are equally susceptible to wildfire. In reality, an individual home's ability to survive wildfire is driven largely by local conditions that can be highly affected by a resident's or community's efforts toward mitigating wildfire susceptibility. The condition of vegetation in the immediate area around a home (known as the “Home Ignition Zone”) and the construction materials used in building a home (Quarles et al. 2010) could result in very different response function values for individual homes (Cohen 2019). Consideration of this local variation in susceptibility is well beyond the scope of the Wildfire Risk to Communities project, and HUImpact should be considered a landscape metric rather than specific to any one home.

Housing Unit Risk (HURisk)

HURisk is a 30-m raster that integrates all four primary elements of wildfire risk - likelihood, intensity, susceptibility, and exposure - on pixels where housing unit density is greater than zero. It is an index with similarities to Risk to Potential Structures (RPS) in Scott et al. (2024), but it also incorporates housing unit count.

The calculation of HURisk is very similar to HUImpact. We again used the housing unit count as derived from HUDen (representing exposure) and the cRPS raster from Scott et al. (2024) (representing intensity and susceptibility). We multiplied those two datasets by the annual burn probability (BP) raster from Scott et al. (2024)(representing wildfire likelihood). HURisk is analogous to the expected Net Value Change (eNVC) metric presented by Scott et al. (2013), if the only highly-valued resource or asset is housing units. The equation for HURisk is simply:

$$HURisk = (HUDen * 0.0009) * cRPS * BP$$

To produce the final integer version of the HURisk raster, we multiplied by 1,000,000 and rounded to the nearest integer, just as we did for HUImpact. Pixels where the HUDen raster is zero are NoData in the HURisk raster.

ACKNOWLEDGEMENTS

The authors would like to thank the following individuals who all contributed review comments and feedback on prototype datasets for this project: April Brough, Julie Gilbertson-Day, and Chris Moran currently or formerly Pyrologix; Alan Ager, Dave Calkin, Mark Finney, Jessica Haas, Don Helmbrecht, Karin Riley, Karen Short, Rick Stratton, Matt Thompson, Kit O’Connor, Nicole Vaillant, Erin Noonan-Wright currently or formerly with the USDA Forest Service; Michele Crist with the Bureau of Land Management; and Henry Bastian and Russ Johnson with the US Department of Interior Office of Wildland Fire. We are also grateful for the support and leadership from Wildfire Risk to Communities project lead, Jim Menakis with USDA Forest Service Fire and Aviation Management. Likewise, our partners in this project at Headwaters Economics, Kimi Barrett, Bill Daigle, Patty Hernandez Gude, Kelly Pohl, Brent Powell, Tara Preston, and Scott Story, have helped us to refine our communication about the datasets. This project was funded by the USDA Forest Service, Fire and Aviation Management. Funding was also provided by the USDA Forest Service Fire Modeling Institute, which is part of the Rocky Mountain Research Station. Some salary was provided by FMI through an Oak Ridge Institute for Science and Education (ORISE) agreement under the U.S. Department of Energy (DE-SC0014664). Work on dataset development was primarily completed by Pyrologix LLC under contract with the USDA Forest Service, Fire Modeling Institute.

REFERENCES

Dillon, Gregory K.; Lazarz, Mitchell T.; Karau, Eva C.; Story, Scott; Pohl, Kelly A. 2024. Community Wildfire Risk Reduction Zones for the United States. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2024-0030

Dillon, Gregory K.; Scott, Joe H.; Jaffe, Melissa R.; Olszewski, Julia H.; Vogler, Kevin C.; Finney, Mark A.; Short, Karen C.; Riley, Karin L.; Grenfell, Isaac C.; Jolly, W. Matthew; Brittain, Stuart. 2023. Spatial datasets of probabilistic wildfire risk components for the United States (270m). 3rd Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2016-0034-3

LANDFIRE, U.S. Department of the Interior, Geological Survey. 2020. LANDFIRE 2.2.0—Scott and Burgan Fire Behavior Fuel Models (FBFM 40) https://www.landfire.gov/

Scott, Joe H.; Burgan, Robert E. 2005. Standard fire behavior fuel models: a comprehensive set for use with Rothermel's surface fire spread model. Gen. Tech. Rep. RMRS-GTR-153. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 72 p. https://doi.org/10.2737/rmrs-gtr-153

Scott, Joe H.; Thompson, Matthew P.; Calkin, David E. 2013. A wildfire risk assessment framework for land and resource management. Gen. Tech. Rep. RMRS-GTR-315. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 83 p. https://doi.org/10.2737/rmrs-gtr-315

Scott, Joe H.; Brough, April M.; Gilbertson-Day, Julie W.; Dillon, Gregory K.; Moran, Christopher. 2020a. Wildfire Risk to Communities: Spatial datasets of wildfire risk for populated areas in the United States. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2020-0060

Scott, Joe H.; Gilbertson-Day, Julie W.; Moran, Christopher; Dillon, Gregory K.; Short, Karen C.; Vogler, Kevin C. 2020b. Wildfire Risk to Communities: Spatial datasets of landscape-wide wildfire risk components for the United States. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2020-0016

Scott, Joe H.; Dillon, Gregory K.; Callahan, Michael N.; Jaffe, Melissa R.; Vogler, Kevin C.; Olszewski, Julia H.; Karau, Eva C.; Lazarz, Mitchell T.; Short, Karen C.; Riley, Karin L.; Finney, Mark A.; Grenfell, Isaac C. 2024. Wildfire Risk to Communities: Spatial datasets of landscape-wide wildfire risk components for the United States, Second Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2020-0016-2

Short, Karen C.; Finney, Mark A.; Vogler, Kevin C.; Scott, Joe H.; Gilbertson-Day, Julie W.; Grenfell, Isaac C. 2020. Spatial datasets of probabilistic wildfire risk components for the United States (270m). 2nd Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2016-0034-2

Sanborn Map Company. 2016. West wide wildfire risk assessment: Final report. Prepared for the Oregon Department of Forestry, Western Forestry Leadership Coalition, Council of Western State Foresters, and funded by the USDA Forest Service. 106 p. https://www.thewflc.org/sites/default/files/WWA_FinalReport_3-6-2016-1.pdf

U.S. Census Bureau. 2021. American Community Survey Information Guide. https://www.census.gov/programs-surveys/acs/about/information-guide.html

U.S. Census Bureau. 2021. Population and Housing Unit Estimates Datasets: Vintage 2021. July 1, 2021. https://www.census.gov/programs-surveys/popest/data/data-sets.html

U.S. Census Bureau. 2021. 2017-2021 ACS 5-year Estimates. Released Dec 08, 2022. https://www.census.gov/programs-surveys/acs/news/data-releases.2021.html#list-tab-1133175109

U.S. Census Bureau. 2020. Methodology for the United States population estimates: Vintage 2021. Version 2, December 2021. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2020-2021/methods-statement-v2021.pdf