This document describes the main purpose, and data inputs and outputs, for each script

**AHS\_HSM\_dev**

* Script to develop parameters and data objects based on American Housing Survey (AHS) for use in the Housing Stock Model (HSM)
* Inputs:
  + AHS Sample Case History files
  + AHS microdata 1973-2019
  + Manufactured Housing Shipments annual, 1959-2018
  + Completed Single- and Multifamily housing units, annual, 1979-2019
  + Function: run\_sm
  + Function: vacrate\_fn
  + Function: demrate\_fn
* Outputs:
  + summaries.RData - Summaries of housing stock characteristics and evolution, 1985-2019, for US total and four Census Regions
  + ConLinModels.RData - Linear models to estimate construction as a function of occupied stock growth (OSG) for negative values of OSG, and to estimate growth factor as a function of change in vacancy rates
* Also many figures are created to validate the stock model developed for US total and four census regions

**run\_sm**

* Function to run stock model for national and census regions

**run\_sm\_cty**

* Function to run stock model for counties

**dem\_rate\_fn**

* Function to calculate loss rates by cohort based on loss rates by age ranges and model year

**vacrate\_fn**

* Function to calculate vacancy rates by cohort based on vacancy rates by age ranges and model year

**housing\_stock**

* Script to make summaries of total and occupied housing stock in 2019 from American Community Survey (ACS) data, using a RAS function to calculate total housing stock by type and cohort (combined) based on data of total housing stock by type, and cohort (separately) and occupied housing stock by type and cohort (combined)
* Inputs:
  + ACS Table DP04, 5yr and 1yr versions, county and state versions
  + ACS Table B25127, 5yr and 1yr versions, county and state versions
* Outputs:
  + OccHousing2019.RData – Occupied housing stock by type and cohort in 2019 for all counties
  + TotalHousing2019.RData – Total housing stock by type and cohort for each county in 2019

**pop\_housing**

* This script acquires data of total and occupied housing units, and population, by house type and county, in 2019. Then creates description of the housing stock in 2020 (with same details), and projections of population, population share by house type,, and household size by house type in each county for each year 2020-2060
* Inputs:
  + Population projections for US counties for 2020-2100, from Hauer (2019): 'SSP\_asrc.csv' file, accessible at <https://osf.io/uh5sj/>
  + US Census Bureau population projections <https://www2.census.gov/programs-surveys/popproj/datasets/2017/2017-popproj/np2017_d1_mid.csv>
  + household population by house type, from ACS Table B25033, 5yr table
  + census estimates of total resident population in 2019 <https://www2.census.gov/programs-surveys/popest/tables/2010-2019/counties/totals/co-est2019-annres.xlsx>
  + OccHousing2019.RData – Occupied housing stock by type and cohort in 2019 for all counties, produced by housing\_stock
  + TotalHousing2019.RData – Total housing stock by type and cohort for each county in 2019, produced by housing\_stock
* Outputs
  + InitStock20.RData – Housing stock in 2020 by county, with detailed breakdown by type, age cohort, and vacancy status. Also include population, vacancy factor, and household size by house type and county
  + CountyProjections.RData – population, share by house type, and household size by county, from 2020-2060

**hsm\_cty**

* Script to run housing stock model projections for all US counties, 2020-2060
* Inputs:
  + Function: run\_sm
  + Function: vacrate\_fn
  + Function: demrate\_fn
  + InitStock20.RData – describes the housing stock by county in 2020, created based on ACS data by pop\_housing.R.
  + CountyProjections.Rdata – county level population projections based on Hauer and US Census population projections, created by pop\_housing.R
  + ConLinModels.RData – created by AHS\_HSM\_dev
  + summaries.Rdata – created by AHS\_HSM\_dev
* Outputs:
  + County\_SMDF\_Scenarios.RData – stock model data frames (SMDF) for each county
  + County\_Scenario\_SM\_Results.RData – stock model results for each county

**sm\_result\_vis**

* Script to visualize results from the housing stock model
* Inputs:
  + County\_Scenario\_SM\_Results.RData - stock model results for each county, created by hsm\_cty
  + InitStock20.RData – describes the housing stock by county in 2020, created based on ACS data by pop\_housing.R.
* Outputs:
  + US\_smop\_scenarios.RData – describes stock model output (SMOP) for different housing stock scenarios at national level
  + HSM\_results/NewConEstimates.csv - new construction per scenario/year
  + Optional run of function at the end to save figures in Figures directory. Need to uncomment figures to plot and save.

**bs\_combine**

* Script to analyze and combine buildstock (bs) files describing stock evolution in different scenarios, add floor area estimates to individual buildings, based on their ResStock floor area bin, for base projection and RFA projection, and calculate mean floor area by type, cohort, and geo (div, state, cty), for base and RFA. For RFA add a stipulation that county mean average floor areas by type are never lower than what they are in base. Calculate m2/cap per house type in each county for each scenarios.
* Inputs
  + Bs\_2020\_180k.RData – housing characteristics of the housing stock in 2020, created by buildstock batch using the NREL ResStock housing characteristics data (<https://github.com/NREL/resstock>). This file is too large to share on github. Users may contact the author ([peter.berrill@yale.edu](mailto:peter.berrill@yale.edu)) to request this data file.
  + bs\_baseRFA.RData – housing characteristics of new construction under the assumption of reduced floor area. Created by bs\_adjust in the residential energy projection repository (https://github.com/NREL/resstock/tree/feature/projections/projection\_scripts)
  + InitStock20.RData – describes the housing stock by county in 2020, created based on ACS data by pop\_housing.R.
  + US\_smop\_scenarios.RData – stock model output (SMOP) for different housing stock scenarios, created by sm\_result\_vis.R
  + County\_Scenario\_SM\_Results.RData - stock model results for each county, created by hsm\_cty
* Outputs
  + County\_FloorArea.RData – summary data frames of stock model outputs for each county including floor area characteristics
* rename the smop files as \*\_FA. Save as ‘HSM\_results/County\_FloorArea.RData’ and ‘resstock\_projections/ExtData/County\_FloorArea.RData’

**mgInt**

* Script to read in the Bill of Material excel files and calculate material and GHG intensities for each archetype by material category.
* Inputs:
  + framing.RData - share of new single-family built with wood/masonry framing by Division, based on Survey of Construction microdata
  + Bill of Materials files created in Athena Impact Estimator for 51 housing archetypes
* Outputs:
  + Arch\_intensities.RData – summary of material and GHG intensities (2020 and 2060) of 30 housing archetypes by 9 census division.
  + Full\_arch\_intensities (RData and csv files) – summary of material and GHG intensities (2020) for 51 housing archetypes, without aggregating by division-specific share of wood and masonry framed single-family homes

**mgInt\_apply**

* Script to apply calculated material and GHG intensities to counties and calculate material flows and GHG emissions from housing stock evolution. Create Figures 4, 5, 7, and the data needed for Figure 8.
* Inputs:
  + bs\_base.RData – characteristics of new construction by county 2020-2060, generated by buildstock batch sampling using the NREL ResStock housing characteristics data, with inputs of housing stock evolution from the housing stock model results
  + bs\_baseRFA.RData – characteristics of new construction by county 2020-2060, generated by buildstock batch sampling using the NREL ResStock housing characteristics data, with inputs of housing stock evolution from the housing stock model results, assuming no new housing is larger than 3,000 sqft per unit.
  + Arch\_intensities.RData - summary of material and GHG intensities (2020 and 2060) of 30 housing archetypes by 9 census division, created by mgInt
  + County\_FloorArea.RData - summary data frames of stock model outputs for each county including floor area characteristics, created by bs\_combine.
* Outputs:
  + County\_FloorArea\_Mat.RData - summary data frames of stock model outputs for each county including material flows
  + US\_FA\_GHG\_summaries.RData – national summaries of annual housing stock evolution including material flows, floor area, and GHG

**map\_results**

* Script to generate map results by county of different metrics, including ratio of material outflows/inflows per county (Fig 8 in main manuscript) as well m2/cap, vacancy rates, etc.
* Inputs:
  + County\_FloorArea\_Mat.RData - summary data frames of stock model outputs for each county including material flows