Modeling Wood Fuel Production

Update to the Team of Specialists on Forest Product and Wood Energy Statistics
28 April 2022
Why FAO? FAO is the custodian of a data series on wood fuel production that dates back to 1961. The data are reported by countries to FAO and our partners through the Joint Forest Products Questionnaire (JFSQ) and provided publicly via the FAOSTAT database.

Why estimate? The data series should not have gaps for countries so that totals can be estimated, for example, global production of wood fuel.

When do we need an estimate? If countries do not report data in a given year, FAO uses external information sources and/or a model to estimate wood fuel production for that year.

How will the model be implemented? There is a relatively short time window between when the country data are submitted (or not submitted) and when FAO must release an estimate for that year.

Why “now”? The model currently in use by FAO is a collection of linear and log linear equations based on work completed from approximately 1998-2002. With new information available and new statistical tools available, we would like to update the model used by FAO.
What we have*: FAOSTAT

- FAO’s wood fuel data are based on figures reported by country officers.

- Non-repeating official data for ≥ 10 years since 1960.

- Official data < 10 years since 1960.
  - Often, these countries have a high reliance on woodfuel

- FAO uses a model to estimate national wood fuel production

*in 2019
Problem Statement

We need a method for estimating wood fuel (and wood charcoal) in the “current” year for countries that do not submit data in that year.
Step by Step

- **Design and initiate a systematic country-by-country data search**
  - Early 2020

- **Complete 145 countries and initiate an advisory committee**
  - Summer 2020

- **Assess additional information and previous work**
  - Fall 2020

- **Convert all data (where possible) to national estimates of production**
  - Fall 2020

- **Preliminary identification of groups of countries that should share a modeling framework**
  - Fall 2020

- **Data preparation automation procedures (cleaning)**
  - 2021

- **Create new models for each country group (fuel, wood, charcoal, HH, industry)**
  - 2022

- **Develop the underlying conceptual model**

- **Data Management System**

- **Data Visualization**

- **Information Model**

- **Assemble covariates**

- **Imputation for covariates**

- **Final clustering**
## Data found as of 30 June 2021

### Number of countries search and number of data points found

<table>
<thead>
<tr>
<th>Continent</th>
<th>Number of countries with &gt;10 years FAOSTAT data</th>
<th>Number of countries searched</th>
<th>Countries for which we “found” data</th>
<th>Countries with no data</th>
<th>Number of data points found</th>
<th>Number of Countries with AFREC data</th>
<th>Number of AFREC data points</th>
<th>Total number of additional data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>8</td>
<td>49</td>
<td>46</td>
<td>3</td>
<td>943</td>
<td>44</td>
<td>1139</td>
<td>2082</td>
</tr>
<tr>
<td>America</td>
<td>8</td>
<td>33</td>
<td>15</td>
<td>18</td>
<td>771</td>
<td>0</td>
<td>0</td>
<td>771</td>
</tr>
<tr>
<td>Asia</td>
<td>9</td>
<td>41</td>
<td>23</td>
<td>18</td>
<td>490</td>
<td>0</td>
<td>0</td>
<td>490</td>
</tr>
<tr>
<td>Europe</td>
<td>37</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>69</td>
<td>0</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>Oceania</td>
<td>1</td>
<td>14</td>
<td>2</td>
<td>12</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>145</td>
<td>91</td>
<td>54</td>
<td>2309</td>
<td>44</td>
<td>1139</td>
<td>3448</td>
</tr>
</tbody>
</table>

### Energy Balance Field Survey Other

- Africa: 61%
- America: 19%
- Asia: 28%
- Europe: 53%
- Oceania: 14%

- IO: 38%
- NO: 54%
- RO: 0%
- T: 7%
- L: 0%
How are data distributed in time?
Conversions needed to scale up to national estimates of total production (m$^3$) ...

DRAFT RESULTS – PLEASE DO NOT CITE OR SHARE
The Simple Model

1) Calculate per capita WF demand (in volume)

\[
\text{Per Capita Consumption WF} = \frac{\text{Production WF} + \text{Import WF} - \text{Export WF}}{\text{Total Population}}
\]

2) Model per capita WF demand (conceptual model)

\[
\text{Per Capita Consumption WF} \sim \text{Poverty Indicators} + \text{Forest Indicators} + \text{Climate Indicators} + \text{Landform Indicators} + \text{Social Indicators} + \text{Industry Indicators}
\]

3) Calculate national charcoal demand (in weight)

\[
\text{Consumption C} = \text{Production C} + \text{Import C} - \text{Export C}
\]

4) Model the proportion of WF demand met with charcoal.

\[
\text{Proportion of WF Demand met with Charcoal} \sim \frac{\text{Prop of Population Urban} + \text{GDP/Person} + \text{Other wood energy production}}{}
\]

5) Convert back to production of WF (volume) and charcoal (weight).

DRAFT PLAN – PLEASE DO NOT CITE OR SHARE
First level clustering

Time series clustering using tsclust in the package dtwclust (Sarda, Espinosa, 2019)
Second level clustering
Cluster 2

Average Variability of the imputed means vs. Covariate

- **Group**
  - Climate
  - Country size
  - Development/Poverty
  - Forest Industry
  - Fuel Use

- **Missingness (%)**
  - 25
  - 50
  - 75

**MiceRanger for imputation**

DRAFT RESULTS – PLEASE DO NOT CITE OR SHARE
**Per Capita WF Demand 4 Modeling**

**BASE FOUND**

**PART I -> PART IV**

- **Found4Modeling.csv**
  - Widen data set: change name of Area.Code, remove excess columns, create separate columns for WF and C, fill flag columns with “DC”
  - UserProducerType and then with Element and then average remaining
  - Filter out high values of per capita consumption
  - Create per capita consumption dividing by population from covariates.csv

- **FinalFoundData.csv**
  - Remove 25 rows (non-countries and AFREC errors)
  - WOODFUEL: Scale data up to full year, all users in a country, from regional to national, and from households to people
  - CHARCOAL: Scale data up to full year, all users in a country, from regional to national, and from households to people
  - Conversion Factors for Found Data Updated 16AUG21.xlsx
  - Include AFREC data added if no other data for thatCountryxYear (new countries added!)
  - Country Code added for new countries
  - Associated indicators filled in for every new AFREC point

**COMPLETE BASE COVARIATES**

**FULLCOVARIATES**

- **FullCovariates.csv**
  - Create totals for wood production and flag for total
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create 5 new variables (sums or ratios)
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create totals for wood production and flag for total

- **Covariates.csv**
  - Go to sources as in CovariateDataDescription.do cx and bring in data
  - Additional data from FAOSTAT on wood production with flag Annual_2020_10_15.csv (FRA DATA) or direct from FRA
  - Merge all the country codes from all agencies to enable efficient download from non-FAO sites
  - Go to sources as in CovariateDataDescription.do cx and bring in data
  - Additional data from FAOSTAT on wood production with flag Annual_2020_10_15.csv (FRA DATA) or direct from FRA
  - Create 5 new variables (sums or ratios)
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create totals for wood production and flag for total

**BASE FAO**

**PART III -> PART V**

- **FAOstrong.csv**
  - Download data
  - Remove China
  - Merge coniferous and non-coniferous data and create correct flag for aggregate
  - Ensure encoding is UTF-8
  - Download data
  - Remove China
  - Merge coniferous and non-coniferous data and create correct flag for aggregate
  - Ensure encoding is UTF-8

**BASE CLUSTER**

- **clusters.csv**
  - ImputeCovariates.R — to be created

**Wood Charcoal % 4 Modeling**

**COMPLETE BASE COVARIATES**

**WOODCOAL**

- **FullCovariates.csv**
  - Create totals for wood production and flag for total
  - Fill in population data for countries missing from World Bank
  - Ensure encoding is UTF-8

**BASE COVARIATES**

**PART II**

- **Covariates.csv**
  - Create 5 new variables (sums or ratios)
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China

- **Conversion Factors for Found Data Updated 16AUG21.xlsx**
  - Created 5 new variables (sums or ratios)
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create totals for wood production and flag for total

- **LandAreatoFill.csv**
  - Merge all the country codes from all agencies to enable efficient download from non-FAO sites
  - Go to sources as in CovariateDataDescription.do cx and bring in data
  - Additional data from FAOSTAT on wood production with flag Annual_2020_10_15.csv (FRA DATA) or direct from FRA
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create totals for wood production and flag for total

- **FAOSTAT Strong.csv**
  - Created 5 new variables (sums or ratios)
  - Fill in population data for countries missing from World Bank
  - Fill in land areas for China
  - Create totals for wood production and flag for total

**DRAFT PLAN – PLEASE DO NOT CITE OR SHARE**
"There is nothing new except what has been forgotten."
~ Marie Antoinette