**Information about the Standardizing State Datasets scripts**

James Duncan

Forest Ecosystem Monitoring Cooperative, femc@uvm.edu

Last Updated: 9/11/2018

The set of scripts in this zip file work with each state’s data as delivered by the state, walking the coding scheme forward to the DASM method. More information about the project in general can be found at <https://www.uvm.edu/femc/data/archive/project/northeastern_ads/>. The technical report found there includes details on the scripts, and this information is copied below.

## Maine

*Script name: ME\_Aggregate.py*

Prior to 1997, the name of the file indicated the damage causing agent. Severity data were only available for spruce budworm (in the “Defol” attribute) and browntail moth (in the “LEVEL” attribute). Surveyors provided damage ratings such as “medium to severe” or “moderate” or “light”. However, the coding could span from light to severe in a single polygon. Thus, we coded polygons with “light” or “low” in the text as a severity of 1), along with polygons with only “medium” or “moderate” rankings. If the severity text contained “heavy” or “severe” and did not contain the words “light” or “low”, we coded them as a severity of 2. We provide details of the severity recoding in Table 2.

Table 2. Recoding scheme used for translating severity attributes in Maine's historical data to the ADS severity standard.

|  |  |
| --- | --- |
| **Original severity rankings** | **Standardized severity code** |
| “Light”, “Very Light”, “Continuous”, “Scattered”, “Medium”, “Moderate”, “Negl to Light”, “Trace – Scattered”, “Light – Scattered”, “Light – continuous w/spots Med”, “Light – Continuous”, “Occasional Light – Medium”, “Light – Medium with Occasional Heavy”, “Light to Moderate”, “Light – Moderate”, “Light to Heavy”, “Light – Heavy”, “Medium – Scattered” | 1 |
| “Heavy”, “Severe”, “Medium – Heavy”, “Medium to Heavy”, “Medium and Heavy”, “Medium with Occasional Heavy”, “Medium with Spots of Heavy”, “Medium – Severe”, “Moderate – Heavy”, “Moderate – Severe”, “Heavy – Severe”, “Heavy to Severe” | 2 |

## Massachusetts

*Script name: MA\_Aggregate.py*

The text in the column “COMMONNAME” was used for all recoding work, as it contained some combination of agent, damage type, and damage severity. We provide details of the recoding in Table 3, Table 4, and Table 5.

Table 3. Recoding scheme used for translating severity attributes in Massachusetts’ historical data to the ADS severity standard.

|  |  |
| --- | --- |
| Original severity-related text in COMMONNAME attribute | **Standardized severity code** |
| Any text containing: “50%”, “30%”, “Light”, “Moderate”, “Modrate” | 1 |
| “Rusty Pine Hard Defoliation” | 2 |
| Any text containing “Heavy”, “75%” | 2 |

Table 4. Recoding scheme used for translating damage causing agents in Massachusetts’ historical data to the ADS damage causing agent standard.

|  |  |
| --- | --- |
| Original damage causing agent related text in COMMONNAME attribute | **Standardized damage causing agent code** |
| “Gypsy Moth”, “Gypsy”, “Gyspy Moth” | 12089 |
| “Oak Leaf Skeletonizer”, “Oak Leaf Skeltonizer”, “Oak Skeletonizer” | 12029 |
| “Anthracnose” | 25025 |
| “Apple Scab” | 22089 |
| “Ash Rust” | 25048 |
| “Beech Bark Disease” | 22042 |
| “Birch Leaf Miner” | 12070 |
| “Black Locust Gall Midge” | 12119 |
| “Black Turpintine Beetle”, “Black Turpingtine Beetle” | 11011 |
| “Brown Tail” | 12068 |
| “Canker” | 22700 |
| “Cankerworm”, “Fall Canker Worm”, “Fall Cankerworm” | 12014 |
| “Change in Water Table” | 50004 |
| “Cherry Scallop Shell Moth” | 12081 |
| “Cieda” | 13006 |
| “Dead Trees (Flooded)” | 50004 |
| “Diplodia”, “Diplodia (Pine)”, “Dipoldia” | 25058 |
| “Drought”, “Drought (Beech)” | 50003 |
| “Fire”, “Fire / Chemical Dump Site” | 30000 |
| “Flooding” | 50004 |
| “Forest Tent” | 12096 |
| “Hemlock Looper” | 12083 |
| “Insect” | 10000 |
| “Linden Looper” | 12067 |
| “Locust Leafminer”, “Locust Leaf Roller”, “Locust Leaf Ruller, “Locust Twig Borer” | 12119 |
| “Logging” | 70007 |
| “Looper”, “Looper/Light” | 12083 |
| “Maple Saddled Prominet”, “Saddle Priminet”, “Saddle Prominent”, “Saddle Prominet”, “Saddled Prominent” | 12079 |
| “Nantucket Pine Month”, “Pine Tip” | 15065 |
| “Needle Miner” | 12004 |
| “Oak Leaf Tier” | 12055 |
| “Pear Thrips” | 14058 |
| “Pine Looper”, “Pine Pitch Looper” | 12174 |
| “Pine Needle Miner” | 12069 |
| “Red Spruce Winter Inj.” | 50014 |
| “Satin Moth” | 12086 |
| “Spruce Gall Adelgid” | 17002 |
| “Storm Damage” | 50000 |
| “Tier” | 12055 |
| “Turp Beetle” | 11011 |
| “Willow Leaf Beetle” | 12215 |
| “Beech Maple”, “Beech/Maple”, “Beech Mortality”, “Birch”, “Brown Dying Pine/Spruce”, “Browning”, “Brown Pine”, “Brown Pine White Spruce”, “Dead & Dying Pine - Spruce”, “Dead & Rotting Oaks”, “Dead Hemlock”, “Dead Spruce”, “Dead Timber”, “Dead Trees”, “Defoliated Softwoods”, “Dying”, “Evergreen Growth”, “Fuzzy Tops”, “Heavy Hemlock Mortality”, “Hemlock”, “Housing Development”, “Larch”, “Miner”, “No Cause”, “Not Defoliated”, “Oak”, “Oak Leaf”, “Oak Leaf Minor”, “Pine”, “Pine – Unknown”, “Pine/Unknown”, “Red Maple Thin Foliage”, “Red Maple Thin Foliage No Caus”, “Skeletonizer”, “Sprayed Area”, “Spruce”, “Swamp”, “Thin Crowns” | Unable to translate – 99999 |

Table 5. Recoding scheme used for translating damage type in Massachusetts’ historical data to the ADS damage type standard.

|  |  |
| --- | --- |
| Original damage type related text in COMMONNAME attribute | **Standardized damage type code** |
| “Off Color”, “Rusty Pine Hard Defoliation”, “Defoliation” | 1 |
| “Beech Mortality”, “Birch”, “Dead & Dying Pine - Spruce”, “Dead & Rotting Oaks”, “Dead Hemlock”, "Dead Oak”, “Dead Pine”, “Dead Spruce”, “Dead Timber”, “Dead Trees”, “Dead/Dying”, “Dead Trees (Flooded)”, “Defoliated Softwoods”, “Pine Heavy Mortality” | 2 |
| “Brown Dying Pine/Spruce”, “Browning”, “Brown Pine”, “Brown Pine White Spruce” | 3 |
| All Other Codes | Unable to translate: -1 |

## New Hampshire

*Script name: NH\_Aggregate.py*

The New Hampshire data prior to 1997 contained a column for each individual damage agent, with some sort of severity code when that agent was found in the polygon. All observations were defoliation observations, DMG\_TYPE was assigned to 1 for these polygons. Severity was recorded as the following groupings: low (“l” or “L”), low to medium (“LM” or “lm”), medium (“M” or “m”), high (“H” or “h”) or unknown (“UN”). We combined the groups low, low-medium, and medium to represent the lower ADS severity code (e.g., severity rating of 1), and high as the high-severity code (e.g., severity rating of 2) (Table 6). We provide details of the agent recoding in Table 7.

Table 6. Recoding scheme used for translating severity attributes in New Hampshire’s historical data to the ADS severity standard.

|  |  |
| --- | --- |
| **Original severity-related text in the agent columns** | **Standardized severity code** |
| “L, “M”, m”, l”, “lm”, “LM” | 1 |
| “H”, “h” | 2 |

Table 7. Recoding scheme used for translating damage causing agents in New Hampshire’s historical data to the ADS damage causing agent standard.

|  |  |  |
| --- | --- | --- |
| **Column name of damage agent** | **Interpretation** | **Standardized damage causing agent code** |
| “DEFO\_GM” | Gypsy Moth | 12089 |
| “DEFO\_SP” | Saddled Prominent | 12079 |
| “DEFO\_SB” | Spruce Budworm | 12038 |
| “DEFO\_HL” | Hemlock Looper | 12083 |
| “DEFO\_BLM” | Birch Leaf Miner | 12070 |
| “DEFO\_CSM” | Cherry Scallop Moth | 12081 |
| “DEFO\_LLM” | Locust Leaf Miner | 12119 |
| “DEFO\_SMDIE” | Sugar Maple Dieback | 99999 |
| “DEFO\_ANTHR” | Anthracnose | 25025 |
| “DEFO\_HDIE” | Hemlock Dieback | 99999 |
| “DEFO\_RSB” |  | 99999 |
| “DEFO\_WI” | Winter Injury | 50014 |
| “DEFO\_ELB” | Elm Leaf Beetle | 12141 |
| “DEFO\_HAIL” | Hail | 50006 |
| “DEFO\_SDD” | Sulfur Dioxide Damage | 50001 |

## New York

No additional processing was required for New York data beyond the methods described in the Methods section.

## Vermont

*Script names: (VT\_defoliation\_crosscoding\_1985\_1994.py) and (VT\_MergeYearlyDatasets.py)*

Data in Vermont were collected on paper maps prior to digital collection, and Vermont had digitized data back to 1985. However, VTFPR indicated that the compilation prior to 1995 may not include all mapped disturbances (Barbara Schultz, personal communication, 2015). We projected data to the NAD 1983 Vermont State Plane Meters projection (EPSG code 32145).

Vermont utilized a unique coding system prior to the federal standardization of codes in 1999 (USFS 1999). We translated Vermont-specific codes into the ADS standardized coding schemes. First, if they did not exist, we added columns to each year of data for DCA1, DCA2, DMG\_TYPE1, FOR\_TYPE1, HOST1, STATE, and RPT\_YR. For data between 1985 and 2011, we used existing codes in VTDMGTYP, VTDMGAGT, and VTHOST to populate the standardized fields. In some cases, we needed either DMGNAME to further refine the coding, or a combination of VTDMGTYP, VTDMGAGT, and VTHOST to determine the correct codes. After 2012, the standard field were consistently populated. We provide additional details of these coding changes in Table 8, Table 9, and Table 10.

Table 8. Recoding scheme used for translating damage causing agent attributes in Vermont’s historical data to the ADS damage causing agent standard.

| **Original damage causing agent in VTDMGAGT** | **Standardized damage causing agent code** |
| --- | --- |
| "ALM" | 12026 |
| "Anth" | 25025 |
| "AY" | 24004 |
| "BBD" | 22042 |
| "BLF" | 25076 |
| "BLM" | 25001 |
| "BLS" | 12070 |
| "BSW" | 12120 |
| "Bvr" | 41002 |
| "BWA" | 14003 |
| "Cns" | 70006 |
| "CSSM" | 12081 |
| "DED" | 24022 |
| "Drt" | 50003 |
| "ELB" | 12141 |
| "ETC" | 12093 |
| "FCW" | 12014 |
| "Frst" | 50005 |
| "FTC" | 12096 |
| "FWW" | 12082 |
| "Fyr" | 30000 |
| "GM" | 12089 |
| "Hail" | 50006 |
| "HL" | 12083 |
| "Ice" | 50011 |
| "LAT" | 12037 |
| "LC" | 12047 |
| "LD" | 24011 |
| "LLM" | 12119 |
| "Log" | 70007 |
| "MLC" | 12127 |
| "MS" | 14051 |
| "Mse" | 41004 |
| "Ndc" | 25005 |
| "Oys" | 14028 |
| "PG"[[1]](#footnote-1) | 12130 |
| "PLF" | 25036 |
| "Porc" | 41006 |
| "SAS" | 12167 |
| "Sat" | 12086 |
| "SBW" | 12038 |
| "Scb" | 25004 |
| "See" | 50800 |
| "Sno" | 50011 |
| "SP" | 12079 |
| "SWI" | 50014 |
| "Thr" | 14058 |
| "Wet" | 50004 |
| "Wnd" | 50013 |
| "ZPM" | 15022 |

Table 9. Recoding scheme used for translating damage type attributes in Vermont’s historical data to the ADS damage types standard.

|  |  |
| --- | --- |
| **Original damage type in VTDMGTYP** | **Standardized damage type code** |
| "bld" | 7 |
| "brk" | 6 |
| "brn" | 3 |
| "chl" | 3 |
| "col" | 3 |
| "def" | 1 |
| "dk" | 4 |
| "nDd" | 2 |
| "oDd" | 11 |
| "un" | 9 |
| "unk" | -1 |
| “thn” | The final coding depends on what is in other fields (Barbara Schultz, Personal communication, 2013). These were applied by hand after initial processing and do not appear in the script.  **For polygons in 1991**  If DMGNAME is “Drought Damage” then DMG\_TYPE1 is 1  **For polygons in 1992**  If DMGNNAME is “Heavy Seed” then DMG\_TYPE1 is 1  **For polygons in 1993**  If DMGNAME is “Maple Leaf Cutter” then DMG\_TYPE1 is 1  If DMGNAME is “Hardwood Decline” then DMG\_TYPE1 is 4  **For polygons in 1994**  If DMGNAME is “Heavy Seed” then DMG\_TYPE1 is 1  If DMGNAME is “Hardwood Decline” or “Birch Decline” or “Ash Decline” or  “Wet Site Related Decline” then DMG\_TYPE1 is 4  If VTHOST is “Cdr” and DMGNAME is “Heavy Seed” and VTDMGTYP is “thn” then HOST1 is 241  If VTHOST is “Cdr” and DMGNAME is “Arborvitae Mortality” and VTDMGTYP is nDd” then HOST1 is 241  IF VTHOST is “Cdr” and DMGNAME is “Unknown” and VTDMGTYP is “brn” then HOST1 is 241 |

Table 10. Recoding scheme used for translating host attributes in Vermont’s historical data to the ADS host standard.

| **Original damage type in VTHOST** | **Standardized host code** |
| --- | --- |
| "Ald" | 350 |
| "Apl" | 660 |
| "Ash" | 541 |
| "Be" | 531 |
| "BL" | 901 |
| "Bnt" | 601 |
| "Brc" | 375 |
| “Cdr” | Depends on values in other fields, if VTDMGAGT was “ALM” then the host code was set to 241, otherwise it was set to 99999 |
| "Chy" | 760 |
| "Elm" | 970 |
| "F" | 012 |
| "Hk" | 261 |
| "Hwd" | 001 |
| "Lch" | 071 |
| "Mr" | 316 |
| "Ms" | 318 |
| "Mxd" | 003 |
| "Oak" | 833 |
| "Pin" | 100 |
| "Pop" | 740 |
| "Pr" | 125 |
| "Pw" | 129 |
| "SF"[[2]](#footnote-2) | 002 |
| "Spr" | 090 |
| "Sr" | 097 |
| "Sw" | 094 |
| "Swd" | 002 |
| "Wlw" | 920 |

### Python Scripts

We developed two scripts, one for processing the older coding for data from 1985 to 1994 (VT\_defoliation\_crosscoding\_1985\_1994.py), and the other for combining all years into a single dataset (VT\_MergeYearlyDatasets.py). Note that the processing of damage type “thn” in VTDMGTYP was dependent on other fields for certain years. Refer to Table 9 for more information.

1. The code of “PG” corresponded to combined phigalia/gypsy moth damage, so the phigalia damage was recorded as 12130 in DCA1 and the gypsy moth damage was recorded as 12089 in DCA2 [↑](#footnote-ref-1)
2. When this host code was encountered, the FOR\_TYPE1 attribute was set to 100 [↑](#footnote-ref-2)