LEGACY TREE DATA: A NATIONAL DATABASE OF DETAILED TREE MEASUREMENTS FOR VOLUME, WEIGHT, AND PHYSICAL PROPERTIES

P.J. Radtke, D.M. Walker, A.R. Weiskittel, J. Frank, J.W. Coulston, J.A. Westfall[1]

Abstract—Forest mensurationists in the United States have expended considerable effort over the past century making detailed observations of trees' dimensions. In recent decades efforts have focused increasingly on weights and physical properties. Work is underway to compile original measurements from past volume, taper, and weight or biomass studies for North American tree species. To date, taper records have been recovered from over 150,000 trees, and biomass records from another 22,000. Upon completion the database will serve many purposes including the development and testing of taper, volume, and biomass estimators for about thirty U.S. tree species that comprise roughly two-thirds of the Nation's growing stock. The work is going very well, especially for eastern species that currently make up a majority of the collection to-date. Work will continue in the East, but a major emphasis going forward will be the collection of data sets from western species.

INTRODUCTION

Forest mensurationists in the United States have expended considerable effort over the past century making detailed observations of trees' dimensions and physical properties. Many studies involved felling trees to make dimensional measurements of main stem attributes including: scaling diameters and lengths of merchantable logs (Allen, 1902; Kenety, 1917); determination of cubic foot volumes, taper, bark thickness, and internal defects (Hornibrook, 1950; Pemberton, 1924); and tree-ring analysis for reconstructing height and diameter growth over time (Bishop et al., 1958). Stem wood and bark physical properties have been carefully studied in felled-tree studies, including extensive efforts carried out by the Forest Service to characterize wood specific gravity in commercially important species (Maeglin and Wahlgren, 1972; Mitchell, 1964).

Interest in estimating biomass yields and production have necessitated studies involving the green and dry weight contents of felled trees, including branch and foliage components together with stem wood and bark, and roots (Whittaker and Woodwell, 1968). The collection of felled tree data has continued to the present time, with measurement protocols varying to suit underlying research goals. Often the goals involve the measurement of stem dimensions along with weights and basic physical properties of aboveground components (Saucier and Clark, 1985). Studies primarily concerned with outside-bark stem dimensions have successfully relied on nondestructive techniques including the use of optical dendrometers or direct measurement with the use of ladders or climbing ropes (Reed, 1926; Westfall and Scott, 2010).

Despite shifting interests and research goals over time, researchers have recognized the value of incorporating information from past studies into new analyses and tools. Often summary results such as tables or equations are used in place of original measurements because the original data are not available (Jenkins et al., 2003). Challenges with standardizing

[1] Associate Professor (PR) Virgina Tech, Forest Resources & Environmental Conservation, Blacksburg, VA 24060. pradtke@vt.edu, 540-231-8863; Research Associate (DW) Virgina Tech, Forest Resources & Environmental Conservation; Associate Professor (AW) University of Maine, School of Forest Resources; Research Assistant (JF) Virgina Tech, Forest Resources & Environmental Conservation; Research Forester (JC) USDA Forest Service; Research Forester (JW) USDA Forest Service.

measurements collected using different protocols or instruments have also been noted to be problematic (Wiemann and Williamson, 2012). Despite these challenges, many original data sets listing detailed tree measurements are known to exist, and methods exist for standardizing or "harmonizing" data collected using different tools and methods (Stahl et al., 2012). The goal of this work is to compile a comprehensive set of detailed field and laboratory measurements for use in developing estimators of standing tree volume, weight, and carbon contents. A further goal is to develop a sharable repository to facilitate long-term archival and re-use of legacy tree data for a wide range of research or management applications.

APPROACH

Facilities and individuals affiliated with the U.S. Forest Service, universities, private companies, and various other institutions were identified as potential points of contact for acquiring legacy data. An initial review of published research articles, written reports, and theses and dissertations was conducted to find data sets directly incorporated into written materials. When only available as paper copies, print materials were scanned to digital image format, primarily as multi-page portable document format (PDF) files. Optical character recognition (OCR) software was used to extract tabular data from PDF files or other digital images into ASCII text. Other printed lists of tree-related measurements such as unpublished computer printouts were also digitized using OCR when original digital files could not be found.

Many electronic files were obtained and processed using a range of hardware and software tools to compile measurement data into ASCII text files that could be easily shared and transferred to other computer software for management, analysis, and public distribution where allowed by data set owners. Non-print media obtained with viable data sets included punched cards, reel- and cartridge-type magnetic tape, floppy disks, CD-rom optical disks, hard disk drives, and portable media such as USB thumb drives or files transmitted over the Internet. A range of binary file formats were encountered including several from electronic spreadsheet, database management, and statistical analysis software. Data originated from various operating systems including VAX/VMS®, MS-DOS®, Microsoft Windows®, Unix®, Linux®, and Apple® O/S versions spanning over thirty years of change in mainstream computer technology.

Hand-written paper forms were also common, including field-tally sheets, Forest Service 558a forms, keypunch forms, and various notebooks containing field and laboratory records. These data were entered manually since OCR was ineffective at digitizing hand-written entries. It was also necessary to manually enter data from some printouts of fixed-width computer files, since OCR has only limited ability to determine the number of blank spaces strung together in a sequence. Scatterplot digitization software was used to efficiently extract stem taper diameter and height coordinate pairs from 558a forms (p. 71; Chapman and Demeritt, 1936).

Legacy Tree Database

A relational database schema was developed to accommodate the range of attributes and study designs represented in legacy data sets. The public view of the database contains roughly 575 attributes, including key fields, arranged into seven basic tables (Table 1—Legacy database relational tables and selected attributes in each table. Relational keys not shown.). With few exceptions, the tree was the basic unit of observation in the database. A few studies were added to the compilation that collected specimens or samples deemed useful for the development of biomass estimators even if no tree-level data were available. An example was a set of wood and bark moisture content measurements collected from quaking bolts upon delivery to a mill.A number of primary and secondary keys link relevant observations among tables and provide flexibility to add measurements to the database without restructuring the underlying database structure. Widely-used and freely-available software tools, such as PostgreSQL, Apache Server, and PHP were used as much as possible to promote accessibility and encourage adoption of the database structure or its contents over the WWW.

Figure 1—Examples of print and electronic media containing legacy tree data records.

Table 1—Legacy database relational tables and selected attributes in each table. Relational keys not shown.

Table	No. Attributes	Typical Attributes
LOCATION	30	Source; LAT/LON; Study design;
TREE	400	Size; Component; Weight; Volume; MC; SG; Protocol
STEM	10	Taper
SECTION	35	Dimensions; Weight
DISK	50	MC; SG; Bark; Wood
BRANCH	35	Position; Dimensions; Component; Weight
CORE	15	SG

STATUS AND DISCUSSION

To-date, 220 distinct data sets containing 174,115 individual trees have been fully incorporated into the database format, with many others not yet fully digitized or formatted to match the legacy database structure. Over 151,000 trees having stem taper measurements are included, along with about 22,000 and 11,000 trees having dry weights and green weights, respectively (Table 2—Number of Legacy trees having dry and green weight records and taper measurements for selected species.). Mainly just those species represented by 100 or more trees measured for stem taper are shown in Table 2—Number of Legacy trees having dry and green weight records and taper measurements for selected species., except for a few commercially important western species that are currently underrepresented in the collection and others that have stem forms not conducive to taper measurements.

Tree records from the Southern Region make up the largest subset of legacy data to-date, with 45% of all taper data being from the four major southern yellow pines, sweetgum, and yellow-poplar. The high proportion of data from southern species is a result of several factors: 1) that the work was aided by a small number of individuals who were able to assist us in obtaining large collections of tree biomass and taper data from the Southeast; 2) that our search began with southern and eastern locales, and only recently has been expanded to find comparable data sets further north and west; and 3) that many northern and western studies of stem volumes were conducted decades earlier than southern studies, especially in southern pines which have been intensively managed only since about 1950 (Fox et al., 2007).

Efforts to recover legacy data will continue under the auspices of this work for some time, with increased attention given to the recovery of data sets in all geographic regions of the United States. Significant collections of Central American and Canadian legacy data have been identified that may be suitable to include with the legacy tree database as well (Navar et al., 2013; Ung et al., 2008). Including trees outside of U.S. borders may aid in developing models for species whose ranges are not limited to the contiguous 48 states. The database design includes tools for continuously adding legacy records so that ongoing work or newly recovered data may be added to the compilation; further, as field researchers and practitioners gain familiarity with standardized sampling protocols, measurement attributes, and the existence of public data repositories like this one, the number of contributions is expected to grow.

Table 2—Number of Legacy trees having dry and green weight records and taper measurements for selected species.

		No. Legacy Trees		
Common_Name	SPCD	Dry wt.	Green wt.	Taper
balsam fir	12	271	23	1,459
white fir	15	12	30	941
subalpine fir	19	169	-	217
Alaska yellow-cedar	42	4	-	737
Atlantic white-cedar	43	-	-	259
eastern redcedar	68	655	17	894
tamarack (native)	71	26	-	167
western larch	73	77	-	15
Engelmann spruce	93	107	-	600
white spruce	94	340	57	1,277
black spruce	95	415	9	1,148
red spruce	97	155	48	450
Sitka spruce	98	-	-	224
jack pine	105	224	-	3,080
sand pine	107	138	-	652
lodgepole pine	108	505	19	204
shortleaf pine	110	481	397	6,470
slash pine	111	991	833	14,208
spruce pine	115	75	1	182
sugar pine	117	4	-	217
western white pine	119	96	-	36
longleaf pine	121	751	787	6,088
ponderosa pine	122	543	254	2,890
Table Mountain pine	123	100	-	151
red pine	125	129	-	2,665
pitch pine	126	117	-	553
pond pine	128	118	18	1,157
eastern white pine	129	236	77	2,642
loblolly pine	131	3,450	2,950	31,146
Virginia pine	132	190	216	3,121
singleleaf pinyon	133	102	76	-
Austrian pine	136	-	-	285
Douglas-fir	202	548	18	1,349
baldcypress	221	28	28	290
pondcypress	222	93	83	696
northern white-cedar	241	20	-	273
western redcedar	242	66	-	540
hemlock spp.	260	-	-	216
eastern hemlock	261	58	32	623
western hemlock	263	91	-	615
red maple	316	709	391	3,864
silver maple	317	14	14	211
sugar maple	318	568	92	1,861
buckeye spp.	330	-	-	113
birch spp.	370	95	69	469
yellow birch	371	466	20	773
sweet birch	372	38	32	145
paper birch	375	304	-	1,562
hickory spp.	400	281	199	2,993
pecan	404	8	-	101
shagbark hickory	407	-	-	128

		No. Legacy Trees		
Common_Name	SPCD	Dry wt.	Green wt.	Taper
hackberry spp.	460	16	14	242
flowering dogwood	491	140	60	290
common persimmon	521	11	6	104
American beech	531	339	10	1,077
ash spp.	540	467	279	686
white ash	541	161	71	388
black ash	543	-	-	163
green ash	544	43	43	164
loblolly-bay	555	-	-	134
American holly	591	22	13	129
black walnut	602	1	1	299
Arizona walnut	606	190	-	-
sweetgum	611	780	771	5,810
yellow-poplar	621	440	416	4,767
magnolia spp.	650	175	-	37
cucumbertree	651	1	1	157
sweetbay Magnolia	653	17	8	417
water tupelo	691	202	203	348
blackgum	693	137	32	1,189
swamp tupelo	694	184	203	1,655
bay spp.	720	188	-	-
American sycamore	731	41	40	430
poplar spp.	740	170	142	286
balsam poplar	741	21	16	215
eastern cottonwood	742	72	-	176
bigtooth aspen	743	85	-	566
quaking aspen	746	535	12	2,517
black cherry	762	145	78	825
oak spp.	800	113	6	707
white oak	802	491	380	5,284
scarlet oak	806	159	142	1,700
northern pin oak	809	-	-	127
southern red oak	812	92	84	1,900
cherrybark oak	813	19	19	414
laurel oak	820	48	48	1,187
overcup oak	822	4	4	183
swamp chestnut oak	825	1	1	244
water oak	827	218	234	2,126
Texas red oak	828	-	-	107
willow oak	831	72	72	520
chestnut oak	832	156	150	2,512
northern red oak	833	332	162	2,371
post oak	835	37	30	1,163
black oak	837	128	139	1,923
live oak	838	-	-	238
black locust	901	29	21	514
basswood spp.	950	32	32	360
American basswood	951	31	-	510
elm spp.	970	68	56	820
American elm	972	133	-	241
unknown/other tree	999	573	46	855
All species		22,373	11,164	151,139

LITERATURE CITED

Allen ET, 1902. The western hemlock. Bulletin No. 33, USDA Bureau of Forestry, Washington, D. C. 55 p.

Bishop DM, Johnson FA, Staebler G, 1958. Site curves for red alder. Research Note No. 162, USDA Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, OR. 7 p.

Chapman HH, Demeritt DB. 1936. Elements of Forest Mensuration. Second ed.Albany, NY: Williams Press, Inc., Publishers. 451 p.

Fox TR, Jokela EJ, Allen HL, 2007. The development of pine plantation silviculture in the southern United States. Journal of Forestry 105:337-347.

Hornibrook EM, 1950. Estimating defect in mature and overmature stands of three Rocky Mountain conifers. Journal of Forestry 48:408-417.

Jenkins JC, Chojnacky DC, Heath LS, Birdsey RA, 2003. National-scale biomass estimators for United States tree species. Forest Science 49:12-35.

Kenety WH, 1917. Preliminary study of white spruce in Minnesota. Bulletin 168, University of Minnesota Agricultural Experiment Station, St. Paul, MN. 30 p.

Maeglin RR, Wahlgren HE, 1972. Western wood density survey: report no. 2. Research Paper FPL-183, USDA Forest Service, Forest Products Laboratory, Madison, WI. 24 p.

Mitchell HL, 1964. Patterns of variation in specific gravity of southern pines and other coniferous species. TAPPI 47:276-283.

Navar J, Rios-Saucedo J, Perez-Verdin G, de Jesus Rodriguez-Flores F, Dominguez-Calleros PA, 2013. Regional aboveground biomass equations for North American arid and semi-arid forests. Journal of Arid Environments 97:127-135.

Pemberton JE, 1924. A relation of bark to diameter and volume in redwood. Journal of Forestry 22:44-48.

Reed PM, 1926. Red pine in central New England: a preliminary study with volume and yield tables. 9, Harvard Forest, Petersham, Mass. 23 p.

Saucier JR, Clark A. 1985. Tables for estimating total tree and product weight and volume of major southern tree species and species groups. Washington, D. C.: Southwide Energy Committee, American Pulpwood Association. 59 p.

Stahl G, Cienciala E, Chirici G, Lanz A, Vidal C, Winter S, McRoberts RE, Rondeux J, Schadauer K, Tomppo E, 2012. Bridging national and reference definitions for harmonizing forest statistics. Forest Science 58:214-223.

Ung CH, Bernier P, Guo X-J, 2008. Canadian national biomass equations: new parameter estimates that include British Columbia data. Canadian Journal of Forest Research 38:1123-1132.

Westfall JA, Scott CT, 2010. Taper models for commercial tree species in the Northeastern United States. Forest Science 56:515-528.

Whittaker RH, Woodwell GM, 1968. Dimension and production relations of trees and shrubs in the Brookhaven Forest, New York. Journal of Ecology 56:1-25.

Wiemann MC, Williamson GB, 2012. Density and specific gravity metrics in biomass research. FPL-GTR-208, USDA Forest Service, Madison, WI. 6 p.