



RESEARCH NOTES

MICHIGAN TECHNOLOGICAL UNIVERSITY

FORD FORESTRY CENTER

L'ANSE, MICHIGAN 49946

Number 4

November, 1967

COMPUTER CHECKING FOREST INVENTORY DATA FOR ERRORS

By J. W. Meteer, J. J. Jewell, and C. B. Stott¹

INTRODUCTION

A standard segment for error checking individual tree data is an integral part of the Michigan Technological University (MTU) Standard Program for Continuous Forest Inventory (CFI).

The Standard Program consists of separate program parts or segments for error checking; tree computation; plot totals and summarization; and for tables construction of Area and Combination Breaks (1) (3).

By tradition, the first important task faced in the computation of CFI data has been the correction of errors. Procedures have progressed through various stages from the ocular scanning of trial listings, to formalized step by step procedures on a card sorter, to the present routines prepared for the stored-program electronic computer (7) (8).

No logical or mechanical procedure will find many of the errors which can occur in measurement and coding. However, the traditional checks for improbable relationships, for impossible codes, and for missing data will detect those serious errors that may stop computer operations, create meaningless data, or print out impossible and unacceptable answers.

PROGRAM FEATURES

When work was started on the MTU Standard Program, there was some doubt

¹Respectively, Assistant Professor and Research Forester, Ford Forestry Center; Instructor, Mathematics Dept.; Consultant; Michigan Technological Univ.

as to whether a standard error check segment would be practicable because the problem appeared to be so strongly "job-oriented." An important objective of the standard system was to keep programming techniques simple to understand and use by the average user. Thus it was not deemed appropriate to use a complex set of "parameter input" cards to set data limits and identify code meanings.

The program developed, therefore has the following features and limitations:

1. A standard card format is required for data input (4). Therefore a separate "conversion routine" is used to prepare data for error checking. This routine is a special, relatively simple program (5).
2. Some error checking is done in the conversion routine.
3. A basic core of standard codes is used. These are constructed in the conversion routine.
4. The standard error checks are limited to those applicable to the standard coding scheme and to the specific measurement and classification procedures for which designed.
5. Relatively general data limits are specified. However, these limits can be easily changed by those who wish to alter the program for their own special use.

The resultant program is written in Fortran II. It takes card input (in the form of the "standard cards" developed in the converting routine). It is presently written for the IBM 1620 (II) computer with 40,000 positions of core storage. It is being modified for use on newer high speed computers.

This program segment is basically well suited for the card-oriented machine such as the 1620. These machines are available to many users. This is an important advantage, for error checking should be done concurrently with field work. Small amounts of data can be batch processed while field work is in progress; card files can be corrected at once and results fed back to the field crews. Operation is not expensive and it is recommended that after such piecemeal processing, one last run be made of the total block of cards to prepare a final correct data file for input to the tree computing segment.

Actual CFI case data from 6 different sources has been run on the IBM 1620 computer at Michigan Technological University with excellent results at modest cost.

THE ERROR SYSTEM

The error check system is designed to check individual tree data from permanent CFI or other numbered inventory plots on which repeat measurements are taken. The specific features are derived from a study of CFI project requirements in the North Central Region of the U.S. However, the basic structure contains principles which suit the program to wider application, at least in the Eastern U.S.

Specific checks made are:

1. For impossible data; exceeding prescribed standards.
2. For improbable relationships between data and classification codes for a single measurement (or for each of 2 single measurements) on a given tree.
3. For improbable or impossible relationships between the 2 measurements on a given tree.
4. Tree and plot numbers sequence within district or total job.
5. Missing tree numbers and duplicate tree numbers.

Several program options are available:

1. For a single measurement only.
2. For two measurements on the same trees.
3. Where cordwood usable lengths are recorded on all trees (including sawtimber trees which would also have sawlog U.L.)
4. Where cordwood usable lengths are taken on cordwood, chemical-wood and cull trees only.
5. Output is error messages only.
6. Output is error messages plus "correct" data file. (This would be important in a tape version of the program.)

An important feature of the system is the use of the "tree change" code. Perhaps a misnomer, this name has developed through custom and is used to approve as valid those data which nevertheless fall outside the prescribed limits.

DATA LIMITS AND RELATIONSHIPS

Basic limits are set forth in Table 1. Also prescribed are standard codes for important tree-classification variables which are common to many CFI projects in the North Central and Northeastern U.S. (5). An entry is required for each item of tree-measurement data in the appropriate data fields. Codes of zero are allowed for tree-classification variables to provide acceptance of data from those projects in which not all items are classified. Otherwise the variables of vigor, sawtimber

quality, and mortality are checked against tree status classes. Originally, management potential and product codes were also checked here. Since experience with several cases showed a lack of standardization in coding, these checks were dropped from the system. It is obvious that some standardization of coding is required; and that tree status classification is one of the important keys to the error check program. If error checking is desired on variables not listed in Table 1, special routines must be written. The logical place to perform these checks is in the data conversion process (5).

Table 1. Basic Data Limits and Relationships for Error Checking each Measurement

Field	Standard Tree Status Codes						
	0 Ingrowth	5 Mort	6 Cut	2 Cordwood	9 Chem. Wd.	3 Sawtimber	4 Cull
DBH	0	0	0	5.0-39.9	5.0-39.9	11.0-39.9	5.0-39.9
Saw.U.L.	0	0	0	0	0	8-66	0
Saw.So.	0	0	0	0	0	50-99%	0
Cord.U.L.	0	0	0	4-98	4-98	4-98**	4-98
Cord.So.	0	0	0	50-99%	50-99%	50-99%**	0
Vigor	0	0	0	0, 1, 2, 3	0, 1, 2, 3	0, 1, 2, 3	4
Quality*	0	0	0	0	0	Hdwds 0, 1, 2, 3 (Softwds 0)	0
Mortality	0	Not 0	Not 0	0	0	0	0

* Sawtimber Quality for Hardwood Species

** A program option requires a zero for Cord. U.L. and Soundness in projects for which cord length is not measured on sawtimber trees.

Certain important relationships peculiar to many CFI projects are listed in Table 2. Usable Length-DBH relationships are an important check here, as is tree quality for hardwood sawtimber trees vs DBH. Standards for the latter apply to either the Forest Products Lab or Northern Hardwood and Pine Association rules (6) as applied to a butt-log tree grade. Vigor is checked against tree soundness. This is a feature of the vigor grades advocated by Stott (9). It is these two latter tests that may cause problems in the use of this program for special data from research plots or for CFI projects from other regions where these tree quality and vigor grading systems are not used. It should be remembered that changes can be made in Fortran statements to create separate but parallel program options; with the basic program structure being unchanged.

Table 2. Special Data Relationships Subject to Checking for Errors.
Both M1 and M2 data are checked individually as follows:

<u>Vigor:</u>		<u>Soundness:*</u>	
Must be 3	when	50-74%	
Cannot be 1	when	85% or less	
Must be 4 or 0	when	00	
<u>Quality (Hardwood Sawtimber Trees):</u>		<u>DBH:</u>	
1	limited to	15.0"	or greater
2	limited to	12.0"	or greater
3	limited to	11.0"	or greater
<u>Maximum Cordwood U.L. = (DBH X 5) + 10 (use whole number DBH).</u>			
<u>DBH</u>	<u>U.L.</u>	<u>DBH</u>	<u>U.L.</u>
5	- 35	10	- 60
6	- 40	11	- 65
7	- 45	12	- 70
8	- 50	13	- 75
9	- 55	14	- 80
		15	- 85
		16	- 90
		17	- 95
		18+	- 98
<u>Maximum Sawlog U.L. (for whole number DBH).</u>			
<u>DBH</u>	<u>U.L.</u>	<u>DBH</u>	<u>U.L.</u>
11	- 28	14	- 52
12	- 38	15	- 56
13	- 46	16	- 58
		17	- 62
		18+	- 66
Cordwood U.L. should equal or exceed Sawlog U.L. (if T.S.** is 3).			

* Use sawlog soundness for test when T.S. is 3.

** T.S. is tree status

Growth and change between two respective measurements are checked as indicated in Table 3.

Table 3. Correlations Between M1 and M2 Data. *

VARIABLE	LIMITS																																									
DBH	No Negative Growth Maximum Growth Maximum Growth	(NEW must at least equal OLD). 2.5" in 5 yrs. or MI x 0.5"*** On Ingrowth trees = 5.0" + (MI x 0.5")																																								
Sawlog U.L.	No Negative Growth Maximum Growth	(NEW must at least equal OLD). 16' in 5 yrs. or 6' + (MI x 2').																																								
Cordwood U.L.	No Negative Growth Maximum Growth	(NEW must at least equal OLD). 22' in 5 yrs. or 12' + (MI x 2').																																								
Tree Status	Unusual Codes	4-2, 4-3, 4-6, 4-9 not permitted																																								
Soundness (Cord. or Sawl.)	<p align="center"><u>Table of Allowable Cull-Class Increase Changes</u> (maximum 13%)</p> <table> <thead> <tr> <th><u>OLD</u></th> <th><u>NEW</u></th> <th><u>OLD</u></th> <th><u>NEW</u></th> </tr> </thead> <tbody> <tr> <td>97</td> <td>- 97</td> <td>78</td> <td>- 86 (1 standard grade)</td> </tr> <tr> <td>93</td> <td>- 97 (1 standard grade)</td> <td>65</td> <td>- 78 (1 standard grade)</td> </tr> <tr> <td>86</td> <td>- 97 (2 standard grades)</td> <td>50</td> <td>- 50</td> </tr> <tr> <td></td> <td></td> <td>00</td> <td>- 50 if a T.S. change</td> </tr> </tbody> </table> <p align="center"><u>Table of Allowable Cull-Class Decrease Changes</u> (maximum 18%)</p> <table> <thead> <tr> <th><u>OLD</u></th> <th><u>NEW</u></th> <th><u>OLD</u></th> <th><u>NEW</u></th> </tr> </thead> <tbody> <tr> <td>97</td> <td>- 86 (2 standard grades)</td> <td>78</td> <td>- 65 (1 standard grade)</td> </tr> <tr> <td>93</td> <td>- 78 (2 standard grades)</td> <td>65</td> <td>- 50 (1 standard grade)</td> </tr> <tr> <td>86</td> <td>- 78 (1 standard grade)</td> <td>50</td> <td>- 50</td> </tr> <tr> <td></td> <td></td> <td>65 & 50</td> <td>- 00 if a T.S. change</td> </tr> </tbody> </table>		<u>OLD</u>	<u>NEW</u>	<u>OLD</u>	<u>NEW</u>	97	- 97	78	- 86 (1 standard grade)	93	- 97 (1 standard grade)	65	- 78 (1 standard grade)	86	- 97 (2 standard grades)	50	- 50			00	- 50 if a T.S. change	<u>OLD</u>	<u>NEW</u>	<u>OLD</u>	<u>NEW</u>	97	- 86 (2 standard grades)	78	- 65 (1 standard grade)	93	- 78 (2 standard grades)	65	- 50 (1 standard grade)	86	- 78 (1 standard grade)	50	- 50			65 & 50	- 00 if a T.S. change
<u>OLD</u>	<u>NEW</u>	<u>OLD</u>	<u>NEW</u>																																							
97	- 97	78	- 86 (1 standard grade)																																							
93	- 97 (1 standard grade)	65	- 78 (1 standard grade)																																							
86	- 97 (2 standard grades)	50	- 50																																							
		00	- 50 if a T.S. change																																							
<u>OLD</u>	<u>NEW</u>	<u>OLD</u>	<u>NEW</u>																																							
97	- 86 (2 standard grades)	78	- 65 (1 standard grade)																																							
93	- 78 (2 standard grades)	65	- 50 (1 standard grade)																																							
86	- 78 (1 standard grade)	50	- 50																																							
		65 & 50	- 00 if a T.S. change																																							

* Apply to "repeater" trees (in the case of Sawtimber specifications, T.S. 3-3).

** MI is Measurement Interval.

The so-called "tree change" is an important item of field data. Although some CFI projects use a more elaborate coding scheme, the Standard Error Check Program is designed to accept the simple "zero" and "one" codes. Their use is outlined in Table 4. It has been observed that field crews are slow to learn the concept of the tree change code and it is therefore advisable to keep this coding scheme as simple as possible.

Table 4. Change Codes for use with MTU Standard Error Check Program

Code 0. Data are within limits. Code 1. Data are outside limits, but data are normal and there is <u>no error</u> .		
VARIABLE	TEST NORMAL LIMITS Use Change Code 0:	INHIBIT ERROR TESTS Use Change Code 1:
DBH*	Max. 39.9"	Over 39.9"
Cordwood U.L. Sawlog U.L.	(DBH x 5) + 10' to a 98' maximum DBH/SLUL** Maximum Table <u>DBH</u> <u>U.L.</u> <u>DBH</u> <u>U.L.</u> 11 - 28 15 - 56 12 - 38 16 - 58 13 - 46 17 - 62 14 - 52 18 - 66	Exceeds 98' or DBH/CWUL** formula Exceeds 66' or DBH/SLUL Table
DBH Growth	Positive up to 0.5"/yr.	Negative or Exceeds 0.5"/yr.
<u>U.L. Growth</u> Sawlog U.L. Growth Cord. U.L. Growth	Positive 6' + (MI x 2) max. 12' + (MI x 2) max.	Negative or Exceeds 6' in 1 yr. to 16' in 5 yrs. Exceeds 12' in 1 yr. to 22' in 5 yrs.
Soundness (Cord. or Sawl.)	Increase 13% max. Decrease 18% max.	Exceeds + 13% Exceeds - 18%
Tree Status Codes	All but 4-2, 4-3, 4-6, 4-9	4-2, 4-3, 4-6, 4-9

Note: Change Code will affect no error tests of limits not listed above (such as minimum SLUL of 8' or minimum CWUL of 4')

* DBH is Diameter breast high

** CWUL is Cordwood useable length and SLUL is sawlog useable length.

In addition to checking data limits and relationships, the program tests plot and tree numbers for sequence, for duplicate tree numbers and for missing trees. In this way, a correct data file can be prepared for input to the computing segments of the program with assurance that data cards are in the proper order.

Output from the program consists of error messages describing the nature of the error (Table 5). An alternate output is a correct data file. This option has not been used much because the card punch on the IBM 1620 is so slow. However, this important feature has been included in the basic system to provide the needed structure for fitting the program to a tape system.

Table 5. Sample of Error Messages from Output of Error Check Program.

GOGEBIC COUNTY TREE DETAILS APRIL 16, 1967		
MTU STANDARD CFI PROGRAM		
Plot	Tree	Tree Detail Error Check Phase 2 Category 1
Field(s)	in Error	
1002	8	CWSO GROWTH
1002	9	M2 CWUL/DBH
1002	39	M1 CWUL/DBH
2002	TREE 1	MISSING
1003	PLOT	OUT OF SEQUENCE
1003	46	DUPLICATE TREE NO.
1004	7	M2 CWUL/DBH
1004	47	TREE MISSING
1009	32	M1 VIGOR
1009	33	CWUL GROWTH
1010	11	M2 CWUL/DBH
1010	20	M1 CWUL/DBH
1010	38	M1 CWUL/DBH
1010	38	M2 CWUL/DBH
1014	TREE 1	MISSING
1017	12	M1 SPEC/QUAL
1068	TREE 1	MISSING
1068	19	M1 SPEC/QUAL
1071	6	DBH GROWTH
1071	16	CWSO GROWTH
1073	14	M1 T.S.
1073	14	M2 T.S.
1073	18	TREE MISSING
1073	17	TREE OUT OF SEQUENCE
1073	19	TREE MISSING
1073	26	CWSO GROWTH
1076	39	M1 T.S.

* CWSO is cordwood soundness and SLSO is Sawlog Soundness

** M1 and M2 refer to measurement numbers 1 and 2

SUMMARY AND CONCLUSIONS

Although the error check program would not be considered "standard" by those whose tree vigor or quality grading rules differ substantially from those recommended by Stott (9) (10), it nevertheless has widespread application to CFI cases.

There are two alternatives to consider when designing a standard system for error checking:

1. Writing a basic generalized program with data limits, code meanings, and relationships set by special "parameter" or control cards.
2. Developing a more restrictive program with specified data limits and relationships, requiring greater standardization of coding.

The latter procedure was chosen in the design of this program. It has worked well on data from several CFI cases.

This program is not as restrictive as it might at first appear. Considerable flexibility can be developed in the converting routines. Furthermore, data limits can easily be changed by simple alteration of the Fortran statements. For example, the minimum D.B.H. can be changed from 5.0" to 3.6" or minimum DBH of saw-timber trees can be changed from 11.0" to 9.0". However, even the simplest change in the program should be accompanied by complete testing with pre-computed test data.

An important adjunct to the error check program is the conversion routine (5). In this preliminary program, codes are converted to the required standard coding schemes, thus obviating any need for changing field coding procedures. These converting programs can be quite simple. However, it is recommended that the simple double punch and blank column check, as well as more sophisticated checking of special items of data be done in the converting routine.

If CFI users could adopt standard coding systems, the error checking job would be much easier to standardize and a complete broad-spectrum error check system could be developed. Standard checks could be devised for most of the tree-classification variables in common use. Management potential and product are examples. Although more codes could be converted to a standard system in the conversion programs, this is really not the proper place to do this job. It should start with the field work.

LITERATURE CITED

1. Jones, Thomas A. (1965), Standardization of Computer Programs and Data Coding for CFI. Proceedings, Conference on Continuous Forest Inventory, pp. 51-57. Michigan Technological University, May 1965.
2. Meter, J. W. (1966), A Standard Computer Program for CFI. Ford Forestry Center Mimeo. Paper presented to Lake States Council of Industrial Foresters. March, 1966.
3. Meter, J. W. (1967), Computer Programming for CFI -- A progress report on the MTU standard program. Proceedings, Lake States Technical Division, American Pulpwood Association. June, 1967.
4. Meter, J. W., C. B. Stott, and J. F. Jewell (1966), Standard Input for Computer Processing of Forest Inventory Data. Research Note No. 1, Ford Forestry Center, Michigan Technological University.
5. Meter, J. W., C. B. Stott, and J. F. Jewell (1967), Conversion Programs for the MTU CFI standard program, Research Note No. 3, Ford Forestry Center, Michigan Technological University.
6. Northern Hardwood & Pine Manufacturers Association. (1959) Official grading rules for northern hardwood and softwood logs. N.H. & P. Association, Green Bay, Wisconsin.
7. Smith, R. H. (1965), Programming for error detection. Proceedings Conference on Continuous Forest Inventory, pp. 151-165, Michigan Technological University, May, 1967.
8. Stott, C. B. (1965), A brief on the error problem in CFI. Proceedings, Conference on Continuous Forest Inventory, pp. 140-142, Michigan Technological University, May 1965.
9. Stott, C. B. (1963), Tree vigor grading rules commonly used in CFI cases, North Central Region. Forest Control by Continuous Inventory. No. 108, U.S.F.S., Region 9.
10. Stott, C. B. (1961), Log and tree grading for wood quality; Parts I through V. In Forest Control by Continuous Inventory. No's. 84, 85, 86, 87, 91. U.S.F.S., Region 9.
11. U. S. Forest Service (1965), A guide to hardwood log grading. Northeastern Forest Experiment Station, U. S. Forest Service.