EPA-420-P-98-022

Calculation of Age Distributions in the Nonroad Model: Growth and Scrappage

Report No. NR-007

February 19, 1998

Nonroad Engine Emissions Modeling Team
Assessment and Modeling Division
US EPA, Office of Mobile Sources

Introduction

The NONROAD model calculates nonroad equipment populations by age (i.e., an age distribution) for given equipment types and scenario years. This calculation is necessary for the model to account for a number of factors which affect nonroad emissions over time, including emissions deterioration, new emissions standards, technology changes, changes in equipment sales and total equipment population, and scrappage programs. The NONROAD model calculates equipment age distributions for the base year (which in the beta version of the model is 1996) based on estimated 1996 engine populations and the scrappage function. The model calculates age distributions for future years by stepping through each year between the base and future years; for each year, the model projects equipment populations, scrappage for each model year of equipment still in service, and equipment sales.

The methods employed for these calculations, the manner in which these methods are encoded in the NONROAD model, and the default assumptions and inputs used by the NONROAD model are described in this report. The Methodology discussion is divided into seven sections. The first section describes the concept of base year populations and its use in NONROAD. The second section describes the model's use of population growth rates to project equipment populations in future years. The third section describes the methods used by the model to calculate scrappage. The fourth section describes the method used by NONROAD to calculate the age distribution in the base year. The fifth section describes the method used by the model to calculate annual sales based on estimated scrappage and projected population growth in each year. The sixth section brings the methods presented in earlier sections together to describe how NONROAD computes the age distribution for all future years. The seventh section discusses the implications of and potential drawbacks to the methods used by the NONROAD model to calculate age distributions and its relationship to the methods used by the California Air Resources Board's OFFROAD model.

The Calculations discussion describes the manner in which the methods presented in the Methodology discussion are encoded in the NONROAD model. This discussion is divided into three sections. The first section describes the calculations used to calculate the growth rate for a given equipment population. The second section describes the algorithm used to determine the age distribution, equipment population, and equipment sales for the base year. The third section describes the algorithm used to determine the age distribution, equipment population, and equipment sales for future years.

The Inputs discussion describes the information that must be provided to the NONROAD model to perform the calculations necessary to determine equipment populations, age distributions, and sales. This discussion includes separate sections on equipment populations, growth indicators, scrappage functions, and activity data.

The report ends with a glossary that defines the key terms used in this report. Those terms are italicized the first time they appear in this report subsequent to this introduction.

Methodology

The NONROAD model calculates equipment populations, the *age distribution* of those populations, *annual equipment sales*, and equipment scrappage. These calculations are performed for the *base year*, the specified *target year* for the model run, and for all years between the base year and target year. The methods used by the model to perform these calculations are described below. In general, the model uses the *population growth rate* to project equipment populations from a base year (in which the equipment population is known) through the target year. The model determines equipment sales by adding equipment scrappage to the change in equipment population. The model calculates annual equipment sales using the following equation:

$$Sales(X+1) = [Population(X+1) - Population(X)] + Scrappage(X+1)$$

where

Sales(X+1)	= Equipment sales in year X+1
Population(X+1)	= Equipment population at the end of year X+1
Population(X)	= Equipment population at the end of year X
Scrappage(X+1)	= Equipment scrapped during year X+1

The model uses a *scrappage curve* to determine the proportion of equipment of a given age that has been removed from service. This proportion can then be multiplied by the annual sales for the year in question to determine the *accumulated scrappage* for equipment of that age. Subtracting the accumulated scrappage from initial sales yields the population of equipment still in service. With this information, the age distribution of the population can be calculated by dividing the population at each age by the total population. The details of these calculations are discussed more fully below.

Base Year Populations

Estimates of the base year equipment populations are contained in the NONROAD model's input files. The basis for these estimates is discussed at length in the Population

Estimates Technical Report (NR-006). The beta version of the NONROAD model uses 1996 as its base year.

Population Growth: Growth Rates and Projected Populations

The NONROAD model projects equipment populations in future (post-base year) years by applying a growth rate to the base year equipment population. This growth rate is calculated by determining the *compound annual growth rate* (CAGR) in equipment population between 1989 and 1996, as reported by PSR. This approach is discussed more fully in the Growth Technical Report (NR-008). The model assumes that the population growth rate remains constant during the years between the base and target years. EPA's Nonroad Engine Emissions Modeling Team (NEEMT) has chosen this approach, instead of relying on economic indicator variables, for the reasons discussed in NR-008. Once the CAGR has been calculated, the model uses this growth rate to project the *total equipment population* in each future year up to and including the target year using the following equation:

Total population in year X+1 = [1+(CAGR/100)] * [Total population in year X]

Scrappage

The NONROAD model uses a *scrappage curve* to determine the proportion of equipment that has been scrapped as a function of equipment age. The model uses the scrappage curve originally developed for the EPA NEVES project as its default scrappage curve for all equipment types; the curve is scaled to the *average lifetime* of the equipment such that half of the units sold in a given year will be scrapped by the time those units reach the average expected lifetime and all units will be scrapped at twice the average lifetime. The average lifetime (in years) is calculated as the *average life* (in hours at full load) divided by the *activity level* (hours/year) and the *load factor*. The default scrappage curve is shown in Figure 1 (the data plotted in Figure 1 are also provided in Table 1, located on p. 16 of this document). In Figure 1, age has been transformed to a dimensionless quantity by dividing by the average lifetime for the type of equipment under consideration. Note that in Figure 1, the proportion of equipment that has been scrapped represents the *accumulated scrappage* since the equipment was placed in service.

The default scrappage curve used in the NONROAD model is based on a normal distribution of accumulated scrappage versus age. Other distributions can be used; for example, EPA's proposed standards for small spark-ignited engines uses a Weibull distribution to project fleet turnover. This distribution results in reduced scrappage in the first few years and increased scrappage in later years than is the case for a normal distribution. The NONROAD model allows a user-specified curve to be substituted for the default curve for any or all equipment types. A user-specified curve can vary the rate of scrappage with age (the shape of the curve) but must conform to the assumption that all units are scrapped within twice the average lifetime.

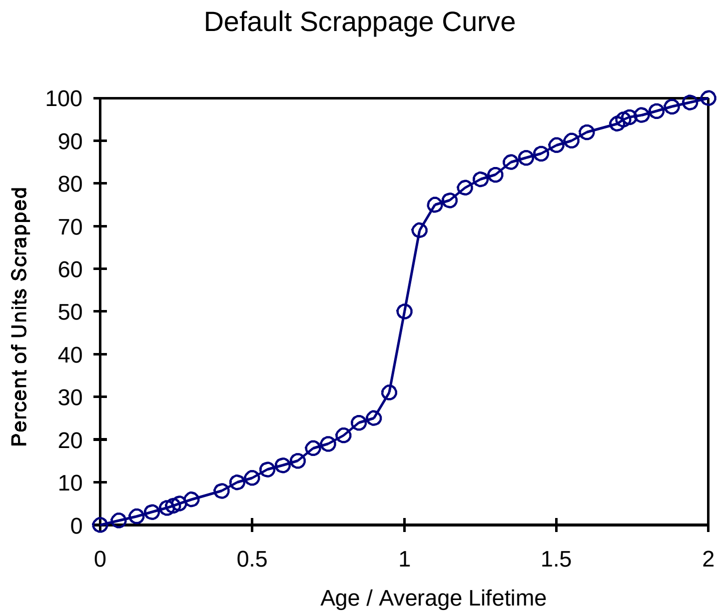

Figure 1. The default scrappage curve relating percent of units removed from service to age.

The *scrappage rate* is defined as the percentage of equipment of a given age removed from service in a given year. The scrappage rate can be derived from the scrappage curve by determining the slope of the scrappage curve at the age in question. Note that for the default scrappage curve, the scrappage rate changes as the equipment age changes: the scrappage rate is low when units are new, reaches a maximum when unit age is equal to the average lifetime, and then declines again for units that are older than the average lifetime.

Age Distribution in the Base Year

To determine the equipment population's actual age distribution as of the base year, one would need to know past equipment sales and scrappage. However, estimates of past equipment sales suffer from missing or incomplete data, are of poor quality, or are based on estimation methods that are incompatible with the methods used to estimate the base year equipment populations used in the NONROAD model. Given these uncertainties, the NEEMT has elected to determine the age distribution as of the base year by assuming constant equipment sales over the *(N)* years preceding the base year, where *N* is two times the average lifetime for the equipment type (in other words, the model does not backcast the effects of growth). Under this assumption, the base year age distribution can be determined directly from the scrappage curve.

The NEEMT recognizes that the assumption of constant sales for the years up to and including the base year may introduce inaccuracies in estimating the rate at which existing equipment is scrapped. For example, to the extent that annual equipment sales increased during the years preceding the base year, the model would overestimate the population of equipment produced between *N* and *N*/2 years prior to the base year and would underestimate the population of equipment produced between the base year and *N*/2 years prior to the base year. As a result, the model would tend to overestimate the rate of scrappage of existing engines for the first *N*/2 years after the base year and then underestimate the rate of scrappage for the subsequent *N*/2 years. Furthermore, overestimating scrappage rates during the first *N*/2 years would tend to result in overestimates of sales and the benefits of lower emission standards during those years. The NEEMT plans to minimize these problems by updating its base year population estimates periodically as new information becomes available, thereby helping to compensate for any discrepancies between modeled and actual scrappage. These updates would take the form of new default population input files. The NEEMT continues to investigate these issues and welcomes suggestions regarding alternative sources of information about the age distribution in the base year, scrappage rates, and other model inputs.

Annual Sales for Future Years

As the model steps through each year from the base year to the target year, it estimates the sales expected to take place during each year. As discussed above, annual sales are determined by adding the scrappage projected to occur during the year to the increase in equipment population projected for that year:

Sales(X+1) = [Population(X+1) - Population(X)] + Scrappage(X+1)

where

Sales(X+1)	= Equipment sales in year X+1
Population(X+1)	= Equipment population at the end of year X+1
Population(X)	= Equipment population at the end of year X
Scrappage(X+1)	= Equipment scrapped during year X+1

The scrappage projected for equipment sold in the target year depends on the type of scenario being modeled. For *end-of-year* scenarios, the projected scrappage of newly-sold equipment is based on an average equipment age of 6 months. For *mid-year* scenarios, the projected scrappage of newly-sold equipment is based on an average equipment age of 3 months.

For a mid-year scenario, the target year sales estimated using the equation presented above represent only one-half year of sales, including one-half year of equipment population growth and one-half year of scrappage replacement. To determine projected full-year sales during the target year, the NONROAD model multiplies the calculated sales by two.

Age Distribution for Future Years

To calculate age distributions for future years, the NONROAD model steps through each year between the base and target years, calculating the effects of growth and scrappage during each step. As the model steps through each year, it estimates the equipment population of each age expected to remain in service using the scrappage curve mentioned above. The age distribution is determined by dividing the equipment population of each age by the total equipment population.

Discussion

The methods used in the NONROAD model have considerable similarity to the methods used in the CARB OFFROAD model. Both models use the same default scrappage curve and similar methods for calculating an age distribution; both models also assume that equipment durability remains constant over time. At the present time, the NEEMT is aware of two major differences between the models. First, the CARB model is only capable of calculating year-end age distributions, whereas NONROAD can calculate either a year-end or mid-year distribution. Second, the CARB model appears to use a different method to adjust age distributions for growth. This method is not fully understood by the NEEMT at this time but will be described in a future technical report.

Calculations

Calculation of Equipment Population Growth Rate

The NONROAD model uses an exponential growth model to project equipment sales over time, a model that is analogous to compound interest. The subroutine *GRWFAC* computes the *Compound Annual Growth Rate* (*CAGR*) from default or user-specified indicator data. CAGR is expressed as percent growth in total equipment population per year. It is calculated by fitting an exponential function to indicator data for 2 years that best match the time period of the base and scenario years. Thus, at least 2 years of data are required for every growth indicator used by the model.

Calculation of the Age Distribution and Equipment Population for the Base Year and Prior Years

The calculation of the age distribution of engines for the base year and any previous year is performed in the *MODYR* subroutine. The steps employed are:

1. Calculate the hours accumulated in service for units of each *model year* using activity data for each year. NONROAD assumes that all engines of a given application, size, and engine type which are still in service operate for the same number of hours during any given time period. Note that the hours accumulated by units sold in the base year must be adjusted downward to reflect the units'

average age at the time of evaluation. For an end-of-year scenario, base year units were on average sold in June and therefore have accumulated only half a year's activity. For a mid-year scenario, base year units have accumulated only a quarter of a year's activity, on average. A mid-year scenario is defined as spring or summer season scenario, or a monthly scenario earlier than October. All other scenarios are defined as end-of-year scenarios.

2. Calculate the average equipment lifetime (in years) as the number of years at which the accumulated hours of service equal the average equipment life (in hours), adjusted for the average load that the equipment's engine experiences, using the following equation:

$$Average\ Lifetime\ (years) = \frac{Average\ Life\ (hrs)}{Activity\ (hrs/yr) * Load\ Factor}$$

where the average life is expressed in terms of hours of operation at full load and the load factor is expressed in terms of the average fraction of available power that is used by the equipment during operation.

3. Use the average equipment lifetime to transform the dimensionless scrappage curve (Figure 1) to an *annualized scrappage curve.*

4. Use the annualized scrappage curve to estimate the *in-service fraction* for each model year, which is the fraction of units that were produced in each year from the base year back to twice the average equipment lifetime and which are still in service as of the base year. The NONROAD model calls this fraction *modfrc*; the fraction for a specific year *M* is called *modfrc(M)*. Figure 2 shows an example for equipment with a ten-year life.

5. Calculate the base year age distribution. Assume that the sales in each year were the same so that the population of units still in service during the base year for a given model year (*M*) is proportional to *modfrc(M).* The *model year fraction* is the proportion of in-service equipment of a specific model year *M* and is called *modscp(M)*; as of the base year, it is given by:

$$\text{modscp}(M) = \text{modfrc}(M) / \sum \text{modfrc}(L)$$

where $\sum modfrc(L)$ is summed over the *N* years between the base year and (*N*-1) years prior to the base year (*N* = twice the average equipment lifetime).

6. To calculate a given model year's equipment population as of the base year, the model year fraction is multiplied by the application's total base year equipment population.

Note that step 5 calculates an age distribution based on the assumption that sales did not vary over the *N* years prior to and including the base year. This assumption is equivalent to assuming that total equipment populations did not vary over those years. The age distribution for later model years can be adjusted to account for changes in population, as discussed below.

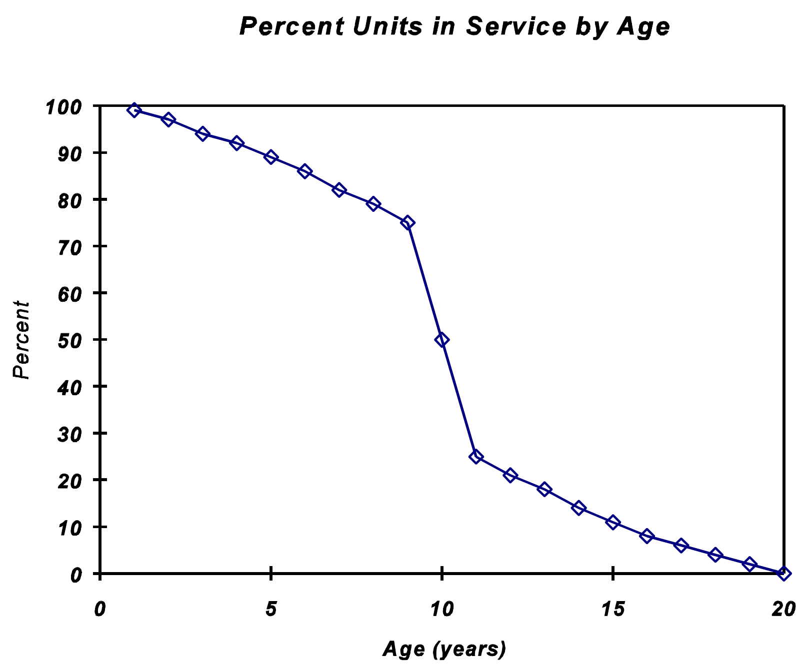

Figure 2. Fraction of equipment still in service versus equipment age, assuming an average equipment lifetime of ten years.

Calculation of the Age Distribution and Equipment Population for Future Years

The key difference between base year and future year age distributions, equipment population, and equipment sales is that future year calculations are adjusted to account for growth in equipment populations and sales. The adjustment of age distributions to account for population and sales growth is performed in the *GRWCLC* subroutine. The steps employed are:

1. Calculate populations of each model year back to twice the average equipment lifetime prior to the base year by multiplying the base year population by the model year fractions as of the base year (i.e., by the base year age distribution). Save these populations for use in future calculations.

2. Step from the base year out to the target year in one year increments and perform the following steps at each year until the target year is reached:

   A. Grow the total population by multiplying the previous year's population by (1 + CAGR/100), where CAGR is the compound annual growth rate.

   B. For all model years prior to the current year, recalculate the model year populations. Conceptually, this calculation is performed by multiplying the sales for each model year by the fraction of equipment projected to remain in use in the current year based on the scrappage curve. For model years prior to or including the base year, the NONROAD model uses a computational shortcut to calculate populations: the model re-uses the set of base year populations by model year shifted by a year. This approach is mathematically equivalent to direct calculation of model year populations for these years because the model assumes that sales were constant for all years up to and including the base year.

   C. Sum the populations for all model years up to the current year minus one. Subtract this sum from the total population to get the current model year population.

   D. Calculate the number of engines scrapped during the current year by summing the scrappage for each model year. For model years prior to the current year, scrappage amounts to one full year of scrappage. For the current model year, scrappage equals 1/2 year's worth of scrappage for an end-of-year scenario and 1/4 year's worth of scrappage for a mid-year scenario. Scrappage for a given model year is calculated by subtracting the model year's accumulated scrappage as of the year prior to the current year from the model year's accumulated scrappage as of the current year.

   E. Calculate the the current model year sales and save this value to use in subsequent calculations. For end-of-year scenarios, sales for the current model year are calculated by adding the total scrappage calculated in step 2D to the current model year population calculated in step 2C. For mid-year scenarios, the sum of total scrappage and current model year population must be multiplied by two in order to place the current model year sales on a full-year basis.

3. Step 2 gives populations by model year for the future year scenario. To convert to model year fractions, divide by the total population. Note that in *GRWCLC,* the base year population is used as the divisor to give model year fractions relative to the base year population (i.e., the fractions do not sum to one).

An example calculation using this algorithm is shown in Figure 3. Figure 3 shows the age distribution in the base year for equipment with a life of ten years. Also shown are the age distributions incorporating 2 percent per annum growth in population for 5, 10 and 20 years after the base year. Since the model year fractions in Figure 3 are shown relative to the base year population, the divisor in each curve is the same and the curves are effectively a direct measure of population by model year, making it easier to compare curves. Note that after 5 years of growth, the populations for the first 5 years have risen above the base year curve, but for years 6-20 the populations remain the same as for the base year. This makes clear the point that the growth algorithm does not perform any backcasting; thus, populations for model years prior to the base year are always assumed to be based on sales at the base year level. After 10 years of growth, effects have propagated into model years 1-10. After 20 years of growth, effects have propagated through the entire distribution.

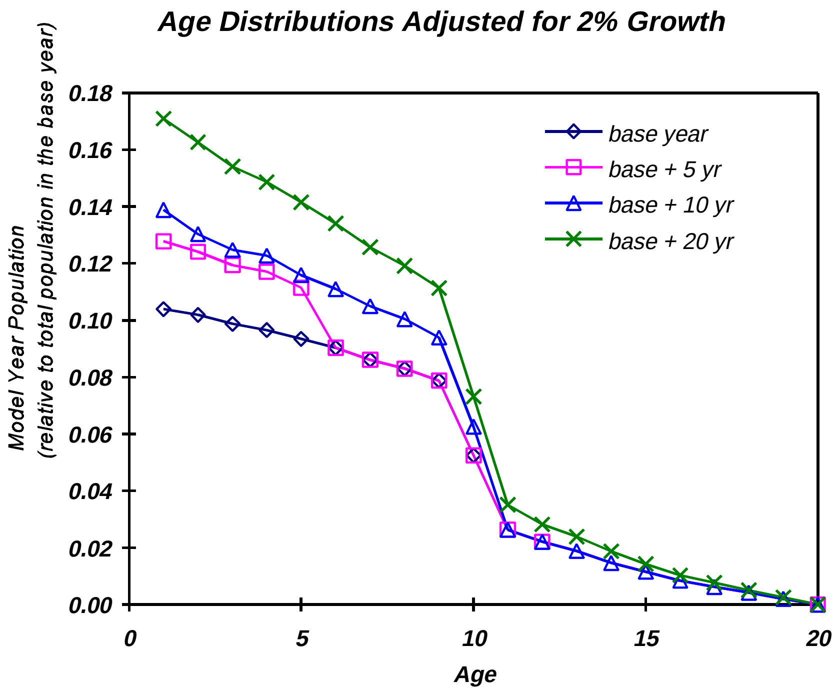

Figure 3. Model year fractions after adjustment for 2% annual growth beginning in the base year, shown for the base year, base + 5 years, and base + 10 years. Model year populations are expressed as fractions relative to the total equipment population in the base year.

Growth may be negative. Equipment populations may decrease due to economic factors or because of changes in technology (e.g., shift from 2-cycle to 4-cycle engines). Figure 4 shows the effects of negative 2% per annum growth for the same example as used in Figure 3. Note that the 5, 10 and 20 year curves drop below the base year curve.

If population data predict a strongly negative growth rate, the population could drop faster than the rate of attrition, resulting in a prediction of negative sales and therefore negative model year fractions. Figure 5 shows that such a situation occurs if negative 15% growth is assumed for the scenario used in Figures 3 and 4. The *GRWCLC* subroutine traps negative sales to zero which effectively limits the rate of population decrease to be no greater than the rate of attrition.

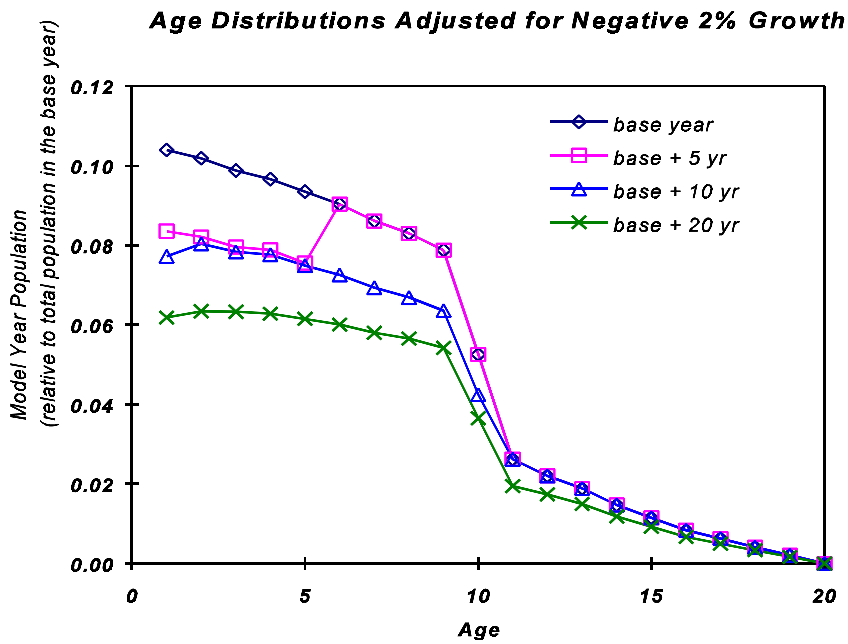

Figure 4. Model year fractions after adjustment for negative 2% annual growth beginning in the base year, for the base year, base + 5 years, and base + 10 years. Populations are expressed in terms of fractions of the total equipment population in the base year.

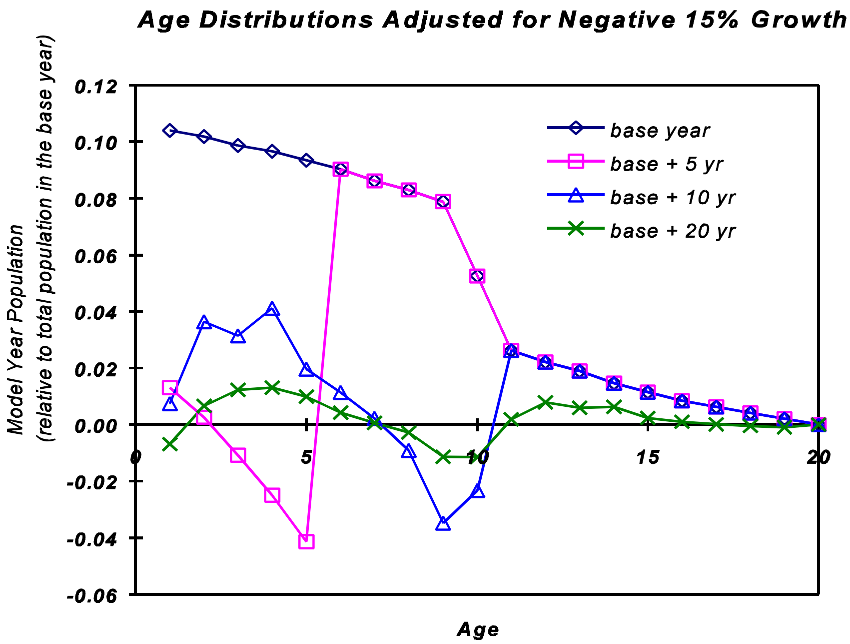

Figure 5. Model year fractions after adjustment for negative 15% annual growth beginning in the base year, for the base year, base + 5 years, and base + 10 years. Populations are expressed as fractions of the total equipment population in the base year.

Inputs

The calculation of age distributions requires the following data to be supplied to the NONROAD model:

- Equipment Populations
- Growth Indicators
- Scrappage Curve
- Activity Data

These data are input to the model using data files. The format of the data files is described below, examples of data are given, and the location of the data is specified.

Equipment Populations

Equipment population data for each state are provided in the directory "data/pop." If estimates are provided for more than one year, the NONROAD model will use the closest year which comes before the target year.

The format of the data in the /POPULATION/ packet is as follows:

1 - 5	FIPS code
7 - 11	Subregion code (used for subcounty estimates)
13 - 16	Year of population estimates
18 - 27	SCC code (no globals accepted)
29 - 68	Equipment description (ignored)
70 - 74	Minimum HP range
76 - 80	Maximum HP range (ranges must match those internal to model)
82 - 86	Expected average life (in hours of use)
87 - 96	Flag for scrappage distribution curve (DEFAULT = standard curve)
100 - 116	Population estimate

An example of data from a /POPULATION/ packet is given below (note that lines are wrapped to fit the page).

```
02000       1990 2260001010 2-Stroke  Motorcycles: Off-Road
16    25   750  DEFAULT                     7
02000       1990 2260001010 2-Stroke  Motorcycles: Off-Road
25    40  1500 DEFAULT                    525
02000       1990 2260001010 2-Stroke  Motorcycles: Off-Road
40    50  1500 DEFAULT                    131
02000       1990 2260001010 2-Stroke  Motorcycles: Off-Road
50   100  3000 DEFAULT                    482
```

Growth Indicators

The default data are based on national growth estimates for the various source category groups. Data are located in the "data/growth" directory in a single file that contains both data packets:

/INDICATORS/	Cross reference between SCC code and growth indicator code.
/GROWTH/	Numerical values for different growth indicator codes.

An indicator code is an alphanumerical code used to identify an actual growth indicator, such as population or employment growth. The growth indicator codes used in the /INDICATORS/ packet must match one of the codes provided in the /GROWTH/ packet. Cross referencing between the /INDICATORS/ and /GROWTH/ packets is based on FIPS code, SCC code, horsepower (HP) range, and technology type. The model uses the best match to the codes provided, falling back on global values if a unique match is not found.

The format of the data in the /INDICATORS/ packet is as follows:

1- 5	FIPS code	(00000 = applies to entire nation) (ss000 = applies to all of state ss)
7-10	Indicator code	(arbitrary alphanumeric code)
12-21	SCC code	(2260004000 = applies to all 2-stroke lawn and garden) (2600000000 = applies to all 2-stroke)
23-27	Beginning of HP range	
28-32	Ending of HP range	
34-43	Technology type (ALL = applies to all tech types)	

An example data record from the /INDICATORS/ packet is given below.

```
00000 001  2260001000     0 9999 ALL        2-Stroke Recreational Vehicles
```

The format of the data in the /GROWTH/ packet is as follows:

1- 5	FIPS code	(00000 = applies to entire nation) (ss000 = applies to all of state ss)
6-10	Subregion code (blank = applies to all subregions)	
11-15	Year of estimate (4-digit year)	
17-20	Indicator code (arbitrary alphanumeric code)	
26-45	Indicator value	

An example data record from the /GROWTH/ packet is given below.

```
02000      1992  001                       15.8
02000      2000  001                       20.4
02000      2005  001                       22.7
02000      2010  001                       24.7
```

Scrappage Curve

A single default scrappage curve is used by the NONROAD model. The scrappage curve is read in from the same data file as the growth data in the "data/growth" directory.

Table 1. Default Scrappage Curve.

Age/ Average Life	Percent Scrapped	Age/ Average Life	Percent Scrapped
0	0	1.05	69
0.06	1	1.1	75
0.12	2	1.15	76
0.17	3	1.2	79
0.22	4	1.25	81
0.24	4.5	1.3	82
0.26	5	1.35	85
0.3	6	1.4	86
0.4	8	1.45	87
0.45	10	1.5	89
0.5	11	1.55	90
0.55	13	1.6	92
0.6	14	1.7	94
0.65	15	1.72	95
0.7	18	1.74	95.5
0.75	19	1.78	96
0.8	21	1.83	97
0.85	24	1.88	98
0.9	25	1.94	99
0.95	31	2	100
1	50		

Activity

The /ACTIVITY/ packet defines how often a piece of equipment is used in a year. The file also contains some other information about the equipment, such as average load factor and tank volume.

The format of the /ACTIVITY/ packet is as follows:

1-10 SCC code
12-51 Equipment description (not used)

52-56	Region code
57-66	Technology type
67-71	Minimum HP
72-76	Maximum HP
77-81	Load factor
82-86	Tank volume (gallons)
87-96	Activity level units
97-106	Activity level

An example data record from the /ACTIVITY/ packet is (note that lines are wrapped to fit on page):

```
2260001010 2-Stroke Motorcycles: Off-Road
ALL    0 9999   76  0.5    Hrs/Yr       120       0.0
```

For more information on equipment populations and equipment activity, load factors and average life, see EPA's technical reports on these subjects, NR-006 and NR-005.

Glossary of Terms

Term	Definition
Accumulated scrappage	The total amount of scrappage that has occurred for equipment of a given model year since its introduction into service
Activity level	The number of hours per year that the equipment in question operates
Age distribution	The function that describes the proportion of in-service equipment by age; consists of the full set of model year fractions for a given year
Annual equipment sales	Total sales during a given calendar year
Annualized scrappage curve	The result of scaling the scrappage curve by the average lifetime of the equipment in question
Average life	The average number of hours that a given type of equipment operates at full load
Average lifetime	The age (in years) at which half of the equipment will have been removed from service; one-half the age at which all of the equipment will have been scrapped
Base year	The year for which the population of equipment in service is known
Compound Annual Growth Rate (CAGR)	The rate at which the population is projected to grow each year in order to reach a specified level
Current year	The year for which the model is currently calculating quantities such as population, sales, scrappage, and emissions
End-of-year	Refers to model runs for which the target date is October 1 or later

Glossary of Terms

Term	Definition
In-service fraction	The fraction of the engines originally sold in a given year which are still in service; the inverse of accumulated scrappage
Load factor	The average power level at which the engine operates divided by the maximum available power
Mid-year	Refers to model runs for which the target date is September 30 or earlier
Model year	Refers to the year in which equipment was produced. Equipment of the same model year was produced in the same year. To clarify the relationship between age and model year, consider the following example: 1990-model year equipment is (on average) six years old in 1996 and ten years old in 2000.
Model year fraction	The fraction of the total equipment population represented by a given model year at a given point in time
Population growth rate	The rate at which the equipment population increases each year
Scrappage curve	A graphical representation of the scrappage function
Scrappage function	The relationship between equipment age (expressed in terms of the fraction of average lifetime) and the proportion of equipment that has been removed from service, i.e., scrapped
Scrappage rate	The percentage of equipment of a given age removed from service in a given year
Target year	The year for which the NONROAD model's user wishes to estimate emissions and other quantities

Glossary of Terms

Term	Definition
Total equipment population	The total number of pieces of equipment in service at a given point in time; the sum of the populations of each model year still in service