ENERGY STAR® Connected Thermostat Products

Discussion Document – Description of Candidate Savings Methodology & Algorithms Implemented in Software

Rev. SEP 15, 2016

OVERVIEW

This discussion document describes the software implementation for assessing Connected Thermostat field savings. The software currently includes 3 candidate methods for modeling the unique relationship between thermal load and HVAC run time in a home. EPA intends to select one of these methods for use in the Version 1.0 ENERGY STAR Connected Thermostat (CT) Program. As of this writing, the specification and test method for this program are in draft and several Alpha iterations of the software have been released. This discussion document has been developed to foster understanding of the software implementation of these methods.

SCOPE

This document is applicable to Alpha release v0.3.4-alpha. EPA intends to issue revised versions of this document as the software is further developed.

CANDIDATE METHODS

An overview of the candidate methods and software implementation is provided below. These methods are used to model the relationship between thermal conditions (outside and inside temperatures) and HVAC run time. Once models are established, they are used by the software to assess baseline run times, that is, what HVAC run times would be if the home were kept at the occupant's preferred comfort temperatures for heating and cooling (detailed in Section 8.a.ii).

- *Linear fit ΔT method* – uses linear regression to characterize the relationships between ΔT and heating run time on core heating days; and ΔT and cooling run time on core cooling days. ΔT is the difference between indoor and outdoor temperature.
- *Daily HTD/CTD method* – a measure of thermal demand expressed as Heating Thermal Demand (HTD) and Cooling Thermal Demand (CTD); calculated for each core heating and core cooling day using daily average temperatures. The model assumes HVAC run time is directly proportional to heating (or cooling) demand.
- *Hourly HTD/CTD method* – a measure of thermal demand expressed as Cooling Thermal Demand (CTD) and Heating Thermal Demand (HTD), calculated for each core heating and core cooling day, where daily thermal demand is the sum of hourly thermal demand values calculated for each hour using hourly average temperatures. The model assumes HVAC run time is directly proportional to heating (or cooling) demand.

INPUT DATA

Data used to assess CT savings is necessarily limited by availability to the following:

1. Data reported by the CT

2. Installation ZIP Code
3. HVAC Type (as far as can be determined by the CT)
4. Inside Temperature
5. Set Temperature (heat & cool)
6. HVAC Run Time (primary heat, aux. heat, emerg. Heat, cool)
7. Publicly available outside temperature. Software uses values reported by the nearest weather station in the same climate zone as the CT, as determined from the ZIP code.

BASELINES

The software assesses savings against a baseline condition expressed as average conditioned space temperatures. This document details development of per home comfort baselines that separately assess resident's comfort preferences for heating and cooling and assume 24/7 use of these temperatures as the baseline condition.

Note, however, that these candidate methods may be used to assess savings relative to ANY average temperature baseline, including regional baselines that define median temperature preferences per climate zone or geographic region, or average temperatures assessed from self-reported thermal preferences.

ENERGY STAR CT SOFTWARE

The software consists of documentation plus two modules which assess CT savings expressed as run time reduction relative to the baseline.

- *Thermostat Module*

  The thermostat module separately assesses HVAC heating and cooling savings attributed to a CT.

  1. Input files consist of:
     - a metadata file that contains the unique ID, controlled HVAC type, ZIP code and interval data file name for each CT in the data set, and
     - an interval data file for each CT in the data set. Input file contents and format is detailed in the software documentation.
  2. The module will not assess savings for thermostat data files that are missing too much data, or if corresponding outdoor temperature data is not available. Indoor and outdoor temperature data gaps of up to 2 hours are interpolated. Days with longer gaps are not included in the analysis. If more than 5% of days are missing HVAC run time data, the thermostat is excluded.[1]
  3. The thermostat module output is a.csv file that includes a separate row for each CT in the data set. Columns in the output file include:
     a. Unique CT ID
     b. Controlled HVAC type
     c. ZIP code in which the thermostat is installed
     d. Heating savings

[1] Intent is for a subsequent software version to also filter out CTs where >5% of days are missing indoor/outdoor temperature data.

e. Cooling savings
f. Resistance Heating Utilization
g. Associated statistical data

A detailed description of the output data is included in the software documentation

- *Statistics Module*

1. A thermostat module output file comprises the input to the statistics module
2. The statistics module outputs a .csv file that includes:
   a. Regional heating and cooling savings as average savings of multiple instances of a CT model in each of the five Energy Information Administration (EIA) climate regions. The output file includes the climate zone, decile bins, means, and standard errors of the mean for all numerical outputs in the individual thermostat output files generated by the thermostat module.
   b. National heating and cooling savings assessed as a weighted average of regional savings. Weighting is by the proportion of national heating and cooling energy, respectively, used in each of the five EIA climate regions
   c. Resistance heat utilization (%) in 5°F bins from 0°F to 60°F (applicable to heating with heat pumps only)
   d. an estimate of statistical power, in order to guide partners toward a data set that provides sufficiently certain statistics

Source code and detailed documentation are available at the following locations:

Documentation: http://thermostat.readthedocs.org/en/latest/

Source code: https://github.com/impactlab/thermostat

OVERVIEW OF SOFTWARE USAGE

Step 1 – The thermostat module is used to assess savings for n CTs that comprise a sample set.

1. *Assess Savings for CT1:*

   a. *Develop Thermal/HVAC Models* – using each of the candidate methods, the software constructs separate models of the relationship between heating and cooling HVAC run time, outside temperature, and temperature choices for the CT;

   b. *Determine Baseline Comfort Temperatures* – the software parses the CT setpoint history[2] to determine occupants' preferred set temperatures for heating and cooling;

   c. *Assess Baseline Run Times* – the software uses each model to assess baseline heating and cooling run times for the CT; i.e. what run times would have been under 24/7 use of baseline comfort temperatures;

   d. *Assess CT Savings* – the software generates an output file that includes CT savings; expressed as percent run time reduction, relative to the baseline; for each of the methods. For heat pumps, the software also outputs Resistance Heating Utilization

[2] Intent is for a subsequent software version to parse the indoor temperature history reported by the CT to determine occupants' preferred set temperatures for heating and cooling.

(RHU), the proportion of total heating run time that includes auxiliary and emergency heating.

e. *Repeat for each of the remaining CTs in the sample set.* Once evaluation of the sample set is complete, the output file will include *n* rows of data, one row for each of the *n* CTs in the sample set.

Step 2 – The statistics module is used to assess aggregate national and regional savings

1. *Regional CT savings* – With a thermostat module output file serving as the input, the statistics module outputs average regional HVAC savings for each candidate method in the (5) U.S. climate zones defined by the EIA[3].
2. *National CT savings* – for each of the candidate methods, the software also outputs national heating and cooling savings as a weighted average of regional savings. Regional savings are weighted by the proportion of national heating or cooling energy consumed in each EIA climate zone.
3. *Resistance heating utilization (RHU)* – for heat pumps, the statistics module also outputs average regional and national RHU.

THERMOSTAT MODULE – ASSESSMENT OF SAVINGS FOR A CT

As noted, the thermostat module currently assesses savings for an individual CT using three candidate methods, as detailed below.

Common Terms and Calculations

The following terms and calculations are common to all (3) candidate methods:

- *Core Days*
  - Core heating days – days where daily heating run time ≥ 1 h with no cooling, excluding days that exceed the missing data thresholds, as previously noted.
  - Core cooling days – days where daily cooling ≥ 1 h with no heating, excluding days that exceed the missing data thresholds, as previously noted.

  Note: Only primary heating source run times are assessed to determine core heating days. That is, auxiliary and emergency heat run times associated with heat pumps are not included.
- *Baseline Assessment* – occupant's preferred comfort temperatures for heating and cooling[4]:
  - Using data reported by the CT for core heating days, determine the preferred comfort set point for heating:

[3] https://www.eia.gov/consumption/residential/reports/images/climatezone_eere-lg.jpg

[4] Method described in paragraph 3.1.2 of "A DATA-DRIVEN FRAMEWORK FOR COMPARING RESIDENTIAL THERMOSTAT ENERGY PERFORMANCE," Bryan Urban and Kurt Roth, Fraunhofer USA, July 2014.

$T\ base\ heat\ (°F)$ *= 90th percentile of user heating set point (CT target temperature) history for core heating days*.[5] Not a Number (NaN) values are ignored.

- Using data reported by the CT for core cooling days, determine the preferred comfort set point for cooling:

  $T\ base\ cool\ (°F)$ *= 10th percentile of user cooling set point (CT target temperature) history for core cooling days*.[5] NaN values are ignored.

METRIC 1 – Linear fit ΔT method

- Assess Heating Savings
  - Develop the CT/home's unique thermal/HVAC heating run time model (Linear fit ΔT method)
    - Calculate the average daily indoor minus outdoor temperature difference ($daily\ \Delta T_d$) for each core heating day. Note that $daily\ \Delta T_d$ is $demand\ \Delta T$ in the thermostat module.

      $$daily\ \Delta T_d\ (°F) = \frac{\sum_{n=1}^{24}(hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n})}{24}$$, where

      d is the core heating day ($d = 001, 002, 003 \ldots x$), and
      n is the hour; ($n = 01, 02, 03, \ldots 24$).

    - For the set of all core heating days, perform a linear regression to determine the relationship between ΔT and heating run time. The Thermostat Module uses a least squares optimization to determine and record the values of τ_h and α_h which best fit the data.[6]

      $$Daily\ RT\ heat_d\ (minutes) = \alpha_h * (daily\ \Delta T_d - \tau_h)$$ where

      d is the core heating day ($d = 001, 002, 003 \ldots x$),
      α_h is the slope of the line (expected to be positive), and
      τ_h is the ΔT associated with zero run time (expected to be positive).

  - Calculate the cumulative baseline heating run time for the set of all core heating days in the CT interval data file (i.e. cumulative baseline heating run time is what would have occurred over the set of all core heating days in the reporting period, had the home been held constant at $T\ base\ heat$, the preferred heating comfort temperature):
    - For each core heating day in the CT interval data file, calculate the average daily indoor baseline minus outdoor temperature difference ($daily\ \Delta T\ base\ heat_d$) that would have occurred had the home been held constant at $T\ base\ heat$:

      $$daily\ \Delta T\ base\ heat_d\ (°F) = \frac{\sum_{n=1}^{24}(T\ base\ heat - hourly\ outdoor\ T_{d.n})}{24}$$, where

      d is the core heating day ($d = 001, 002, 003 \ldots x$),
      n is the hour; ($n = 01, 02, 03, \ldots 24$), and
      $T\ base\ heat$ is the occupant's preferred comfort temperature for heating.

    - Calculate baseline run time as the sum of daily baseline run times for the set of core

---

[5] Intent is for a subsequent software version to parse the indoor temperature history reported by the CT to determine occupants' preferred set temperatures for heating and cooling.

[6] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

heating days in the CT interval data file:

$RT\ base\ heat\ (minutes) = \sum_{d=1}^{x}[\alpha_h * (daily\ \Delta T\ base\ heat_d - \tau_h)]_+$, where

d is the core heating day ($d = 001, 002, 003 \ldots x$), and

α_h, and τ_h , determined above, are constants that define the CT/home's thermal/HVAC heating run time model.

$[\ ]_+$ indicates that the term is zero if its value would be negative.

- CT heating savings is the percent heating run time reduction relative to the baseline, for the reporting period in the CT interval data file

$$HS\ (\%\ RT\ reduction) = 100 * \frac{(RT\ base\ heat - RT\ actual\ heat)}{RT\ base\ heat}$$, where

$RT\ base\ heat$ is the sum of baseline heating run times for all core heating days in the CT interval data file, as detailed for each method, and

$RT\ actual\ heat$ is the sum of heating run times for all core heating days in the CT interval data file.

▪ Assess Cooling Savings

- Develop the CT/home's unique thermal/HVAC cooling run time model (Linear fit ΔT method)
  - Calculate the average daily indoor minus outdoor temperature difference ($daily\ \Delta T_d$) for each core cooling day. Note that $daily\ \Delta T_d$ is $-\ demand\ \Delta T$ in the thermostat module.

    $$daily\ \Delta T_d\ (°F) = \frac{\sum_{n=1}^{24}(hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n})}{24}$$, where

    d is the core cooling day ($d = 001, 002, 003 \ldots x$), and

    n is the hour; ($n = 01, 02, 03, \ldots 24$).
  - For the set of all core cooling days, perform a linear regression to determine the relationship between ΔT and cooling run time. The Thermostat Module uses a least squares optimization to determine and record the values of τ_c and α_c which best fit the data.[7]

    $Daily\ RT\ cool_d\ (minutes) = \alpha_c * (\tau_c - daily\ \Delta T_d)$ where

    d is the core cooling day ($d = 001, 002, 003 \ldots x$),

    α_c is the slope of the line (expected to be positive),

    τ_c is the ΔT associated with zero run time (expected to be positive), and

- Calculate the cumulative baseline cooling run time for the set of all core cooling days in the CT interval data file (i.e. cumulative baseline cooling run time is what would have occurred over the set of all core cooling days in the reporting period, had the home been held constant at $T\ base\ cool$, the preferred cooling comfort temperature):
  - For each core cooling day in the CT interval data file, calculate the average daily indoor baseline minus outdoor temperature difference ($daily\ \Delta T\ base\ cool_d$) that would have occurred had the home been held constant at $T\ base\ cool$:

---

[7] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

$$daily\ \Delta T\ base\ cool_d\ (°F) = \frac{\sum_{n=1}^{24}(T\ base\ cool - hourly\ outdoor\ T_{d.n})}{24}$$, where
d is the core cooling day ($d = 001, 002, 003 \ldots x$),
n is the hour; ($n = 01, 02, 03, \ldots 24$), and
$T\ base\ cool$ is the occupant's preferred comfort temperature for cooling.

- Calculate baseline run time as the sum of daily baseline run times for the set of core cooling days in the CT interval data file:
$$RT\ base\ cool\ (minutes) = \sum_{d=1}^{x}[\alpha_c * (\tau_c - daily\ \Delta T\ base\ cool_d)]_+$$, where
d is the core cooling day ($d = 001, 002, 003 \ldots x$), and
α_c, and τ_c , determined above, are constants that define the CT/home's thermal/HVAC cooling run time model.
$[\ ]_+$ indicates that the term is zero if its value would be negative.

- CT cooling savings is the percent cooling run time reduction relative to the baseline, for the reporting period in the CT interval data file

$$CS\ (\%\ RT\ reduction) = 100 * \frac{(RT\ base\ cool - RT\ actual\ cool)}{RT\ base\ cool}$$, where

$RT\ base\ cool$ is the sum of baseline cooling run times for all core cooling days in the CT interval data file, as detailed for each method, and

$RT\ actual\ cool$ is the sum of cooling run times for all core cooling days in the CT interval data file.

METRIC 2 – Daily HTD/CTD method

- Assess Heating Savings

- Develop the CT/home's unique thermal/HVAC heating run time model (daily HTD method)

- Calculate daily Heating Thermal Demand (HTD_d) for each core heating day in the CT interval data file:

- Calculate the average daily indoor minus outdoor temperature difference ($daily\ \Delta T_d$) for each core heating day
$$daily\ \Delta T_d\ (°F) = \frac{\sum_{n=1}^{24}(hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n})}{24}$$, where
d is the core heating day ($d = 001, 002, 003 \ldots x$), and
n is the hour; ($n = 01, 02, 03, \ldots 24$).

- Starting with an assumed value of zero for Tau ($\tau_h = 0$), calculate the daily Heating Thermal Demand (HTD_d), as follows:
$$daily\ HTD_d\ (°F) = [daily\ \Delta T_d - \tau_h]_+$$, where
$daily\ HTD_d$ is the daily Heating Thermal Demand (°F),
d is the core heating day; ($d = 001, 002, 003 \ldots x$),
τ_h is the ΔT associated with $HTD = 0$ (zero heating run time), and
$[\ ]_+$ indicates that the term is zero if its value would be negative.

- For the set of all core heating days in the CT interval data file, use ratio estimation to calculate α_h, the homes responsiveness to heating, which should be positive
$$\alpha_h(^{minutes}/_{°F}) = {}^{RT\ actual\ heat}/_{\sum_{d=1}^{x} daily\ HTD_d}$$, where
d is the core heating day; ($d = 001, 002, 003 \ldots x$), and

$RT\ actual\ heat$ is the sum of heating run times for all core heating days in the CT interval data file.[8]

- For the set of all core heating days in the CT interval data file, optimize τ_h that results in minimization of the sum of squares of the difference between daily run times reported by the CT, and calculated daily heating run times. Next recalculate α_h (iaw the above step) and record the model's parameters (τ_h , α_h)[9]:

  $\sum_{d=1}^{x}(actual\ RT\ heat_d - daily\ RT\ heat_d)^2$ is minimized, where

  $daily\ RT\ heat_d = \alpha_h * daily\ HTD_d$

  $actual\ RT\ heat_d$ is the total daily heating run time reported by the CT for that core heating day

  d is the core heating day ($d = 001, 002, 003 \ldots x$)

  Note that α_h characterizes the response of the home to heating and τ_h is the difference between inside and outside temperatures when heating run time = 0

- Calculate the cumulative baseline run time for the collection of all core heating days in the CT interval data file (i.e. what would have occurred over the set of all core heating days in the reporting period, had the home been held constant at the preferred heating comfort temperature).

  - Calculate the Heating Thermal Demand for each core heating day in the CT interval data file

    - Calculate the average daily indoor baseline minus outdoor temperature difference ($daily\ \Delta T\ base\ heat_d$) that would have occurred had the home been held constant at $T\ base\ heat$:

      $daily\ \Delta T\ base\ heat_d\ (°F) = \frac{\sum_{n=1}^{24}(T\ base\ heat - hourly\ outdoor\ T_{d.n})}{24}$, where

      d is the core heating day ($d = 001, 002, 003 \ldots x$),

      n is the hour; ($n = 01, 02, 03, \ldots 24$),

      $T\ base\ heat$ is the occupant's preferred comfort temperature for heating, and

      $daily\ outdoor\ T_d$ is the average daily outdoor temperature calculated above.

    - Calculate the baseline Heating Thermal Demand ($HTD\ base_d$) for that day

      $HTD\ base_d\ (°F) = [daily\ \Delta T\ base\ heat_d - \tau_h]_+$, where

      τ_h, determined above, is a constant that is part of CT/home's thermal/HVAC heating run time model,

      d is the core heating day ($d = 001, 002, 003 \ldots x$), and

      $[\ ]_+$ indicates that the term is zero if its value would be negative

  - Calculate baseline run time as the sum of daily baseline run times for the set of core heating days in the CT interval data file

---

[8] For CTs that control heat pumps, $RT\ actual\ heat$ includes only compressor-based heating, i.e. emergency and auxiliary heating run times are not included

[9] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

$RT\ base\ heat\ (minutes) = \sum_{d=1}^{x} \alpha_h * HTD\ base_d\ (°F)$, where

α_h , determined above, is a constant that is part of CT/home's thermal/HVAC heating run time model

$HTD\ base_d$ is the baseline daily Heating Thermal Demand

d is the core heating day ($d = 001, 002, 003 \ldots x$)

- CT heating savings is the percent heating run time reduction relative to the baseline, for the reporting period in the CT interval data file.

  $HS\ (\%\ RT\ reduction) = 100 * \frac{(RT\ base\ heat - RT\ actual\ heat)}{RT\ base\ heat}$, where

  $RT\ base\ heat$ is the sum of baseline heating run times for all core heating days in the CT interval data file, as detailed for each method, and

  $RT\ actual\ heat$ is the sum of heating run times for all core heating days in the CT interval data file.

- Assess Cooling Savings
  - Develop the CT/home's unique thermal/HVAC cooling run time model (daily CTD method)
    - Calculate daily Cooling Thermal Demand (CTD_d) for the each core cooling day in the CT interval data file:
      - Calculate the average daily indoor temperature

        $daily\ indoor\ T_d\ (°F) = \frac{\sum_{n=1}^{24} hourly\ indoor\ T_{d.n}}{24}$, where

        d is the core cooling day ($d = 001, 002, 003 \ldots x$),

        n is the hour; ($n = 01, 02, 03, \ldots 24$), and

        $hourly\ indoor\ T_{d.n}$ are the conditioned space temperatures reported by the CT.
      - Calculate average daily outdoor temperatures for each core cooling day. Using the ZIP code included in the metadata file, look up hourly temperatures from the closest NOAA weather station, within same climate zone.

        $daily\ outdoor\ T_d\ (°F) = \frac{\sum_{n=1}^{24} hourly\ outdoor\ T_{d.n}}{24}$, where

        d is the core cooling day ($d = 001, 002, 003 \ldots x$),

        n is the hour; ($n = 01, 02, 03, \ldots 24$),

        $hourly\ outdoor\ T_{d.n}$ are hourly outdoor temperatures.
      - Calculate the average daily indoor minus outdoor temperature difference ($daily\ \Delta T_d$) for each core cooling day

        $daily\ \Delta T_d\ (°F) = \frac{\sum_{n=1}^{24}(hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n})}{24}$, where

        d is the core cooling day ($d = 001, 002, 003 \ldots x$), and

        n is the hour; ($n = 01, 02, 03, \ldots 24$).
      - Starting with an assumed value of zero for Tau ($\tau_c = 0$), calculate the daily Cooling Thermal Demand (CTD_d), as follows:

        $daily\ CTD_d\ (°F) = [\tau_c - daily\ \Delta T_d]_+$, where

        ($daily\ CTD_d$ is the daily Cooling Thermal Demand (°F),

        d is the core cooling day; ($d = 001, 002, 003 \ldots x$),

τ_c is the ΔT associated with $CTD = 0$ (zero cooling run time), and
$[\ ]_+$ indicates that the term is zero if its value would be negative.

- For the set of all core cooling days in the CT interval data file, use ratio estimation to calculate α_c, the homes responsiveness to cooling, which should be positive
  $\alpha_c(\text{minutes}/{}^\circ F) = RT\ actual\ cool / \sum_{d=1}^{x} daily\ CTD_d$, where
  d is the core cooling day; ($d = 001, 002, 003 \ldots x$and
  $RT\ actual\ cool$ is the sum of cooling run times for all core cooling days in the CT interval data file.

- For the set of all core cooling days in the CT interval data file, optimize τ_c that results in minimization of the sum of squares of the difference between daily run times reported by the CT, and calculated daily cooling run times. Next recalculate α_c (in accordance with the above step) and record the model's parameters (τ_c, α_c)[10]:
  where $\sum_{d=1}^{x}(actual\ RT\ cool_d - daily\ RT\ cool_d)^2$ is minimized, where
  $daily\ RT\ cool_d = \alpha_c * daily\ CTD_d$
  $actual\ RT\ cool_d$ is the total daily cooling run time reported by the CT for that core cooling day
  d is the core cooling day ($d = 001, 002, 003 \ldots x$)

  Note that α_c characterizes the response of the home to cooling and τ_c is the difference between inside and outside temperatures when cooling run time = 0

- Calculate the cumulative baseline run time for the collection of all core cooling days in the CT interval data file (i.e. what would have occurred over the set of all core cooling days in the reporting period, had the home been held constant at the preferred cooling comfort temperature).
  - Calculate the Cooling Thermal Demand for each core cooling day in the CT interval data file
    - Calculate the average daily indoor baseline minus outdoor temperature difference ($daily\ \Delta T\ base\ cool_d$) that would have occurred had the home been held constant at $T\ base\ cool$:
      $daily\ \Delta T\ base\ cool_d\ ({}^\circ F) = \frac{\sum_{n=1}^{24}(T\ base\ cool - hourly\ outdoor\ T_{d.n})}{24}$, where
      d is the core cooling day ($d = 001, 002, 003 \ldots x$),
      n is the hour; ($n = 01, 02, 03, \ldots 24$)
      $T\ base\ cool$ is the occupant's preferred comfort temperature for cooling, and
      $daily\ outdoor\ T_d$ is the average daily outdoor temperature calculated above.
    - Calculate the baseline Cooling Thermal Demand ($CTD\ base_d$) for that day
      $CTD\ base_d\ ({}^\circ F) = [\tau_c - daily\ \Delta T\ base\ cool_d]_+$, where
      τ_c, determined above, is a constant that is part of CT/home's thermal/HVAC cooling run time model,

[10] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

d is the core cooling day ($d = 001, 002, 003 \ldots x$), and
$[\ ]_+$ indicates that the term is zero if its value would be negative

- Calculate baseline run time as the sum of daily baseline run times for the set of core cooling days in the CT interval data file
  $RT\ base\ cool\ (minutes) = \sum_{d=1}^{x} \alpha_c * CTD\ base_d\ (°F)$, where
  α_c , determined above, is a constant that is part of CT/home's thermal/HVAC cooling run time model
  $CTD\ base_d$ is the baseline daily Cooling Thermal Demand
  d is the core cooling day ($d = 001, 002, 003 \ldots x$)

- CT cooling savings is the percent cooling run time reduction relative to the baseline, for the reporting period in the CT interval data file.

  $CS\ (\%\ RT\ reduction) = 100 * \frac{(RT\ base\ cool - RT\ actual\ cool)}{RT\ base\ cool}$, where

  $RT\ base\ cool$ is the sum of baseline cooling run times for all core cooling days in the CT interval data file, as detailed for each method, and

  $RT\ actual\ cool$ is the sum of cooling run times for all core cooling days in the CT interval data file.

METRIC 3 – Hourly HTD/CTD method

- Assess Heating Savings
  - Develop the CT/home's unique thermal/HVAC heating run time model (hourly HTD method)
    - Calculate daily Heating Thermal Demand ($daily\ HTD_d$) for each core heating day in the interval data file:
      - Calculate the average hourly indoor minus outdoor temperature difference ($hourly\ \Delta T_{d.n}$) for each core heating day:
        - Using the ZIP code included in the metadata file, determine the closest NOAA weather station that is in the same climate zone as the CT and look up hourly outdoor temperatures:
          $hourly\ outdoor\ T_{d.n}\ (°F)$, where
          d is the core cooling day ($d = 001, 002, 003 \ldots x$),
          n is the hour; ($n = 01, 02, 03, \ldots 24$)
        - $hourly\ \Delta T_{d.n}\ (°F) = hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n}$, where
          d is the core heating day ($d = 001, 002, 003 \ldots x$), and
          n is the hour; ($n = 01, 02, 03, \ldots 24$).
      - Starting with an assumed value of zero for Tau ($\tau_h = 0$), calculate the daily Heating Thermal Demand ($daily\ HTD_d$), as follows:
        $daily\ HTD_d(°F) = \frac{\sum_{n=1}^{24}[hourly\ \Delta T_{d.n} - \tau_h]_+}{24}$, where
        d is the core heating day ($d = 001, 002, 003 \ldots x$),
        n is the hour; ($n = 01, 02, 03, \ldots 24$),
        τ_h is the ΔT associated with $HTD = 0$ (zero heating run time), and
        $[\ ]_+$ indicates that the term is zero if its value would be negative

- For the set of all core heating days in the CT interval data file, use ratio estimation to calculate α_h, the homes responsiveness to heating, which should be positive
  $\alpha_h(^{minutes}/_{\circ F}) = {}^{RT\ actual\ heat}/_{\sum_{d=1}^{x} daily\ HTD_d}$, where
  $RT\ actual\ heat$ is the sum of heating run times for all core heating days in the CT interval data file.

- For the set of all core heating days in the CT interval data file, optimize τ_h that results in minimization of the sum of squares of the difference between daily run times reported by the CT, and calculated daily heating run times. Next recalculate α_h (iaw the above step) and record the model's parameters (τ_h , α_h)[11]:
  $\sum_{d=1}^{x}(actual\ RT\ heat_d - daily\ RT\ heat_d)^2$ is minimized, where
  $daily\ RT\ heat_d = \alpha_h * daily\ HTD_d$
  $actual\ RT\ heat_d$ is the total daily heating run time reported by the CT for that core heating day
  d is the core heating day ($d = 001, 002, 003 \dots x$)

  Note that α_h characterizes the response of the home to heating and τ_h is the difference between inside and outside temperatures when heating run time = 0

- Calculate the cumulative baseline run time for the collection of all core heating days in the CT interval data file (i.e. what would have occurred over the set of all core heating days in the reporting period, had the home been held constant at the preferred heating comfort temperature).
  - Calculate the baseline daily Heating Thermal Demand ($daily\ HTD\ base_d$)for each core heating day in the CT interval data file
    - Calculate the difference between the occupant's preferred comfort temperature for heating and the average outside temperature for each hour of each core heating day ($hourly\ \Delta T\ base\ heat_{d.n}$):
      $hourly\ \Delta T\ base\ heat_{d.n}\ (°F) = T\ base\ heat - hourly\ outdoor\ T_{d.n}$, where
      d is the core heating day ($d = 001, 002, 003 \dots x$),
      n is the hour ($n = 01, 02, 03, \dots 24$),
      $T\ base\ heat$ is the occupant's preferred comfort temperature for heating, and
      $hourly\ outdoor\ T_{d.n}$ is the average outdoor temperature for that hour reported by the nearest NOAA weather station in the same climate zone as the CT.
    - Calculate baseline daily Heating Thermal Demand ($daily\ HTD\ base_d$)
      $daily\ HTD\ base_d = \frac{\sum_{n=1}^{24}[hourly\ \Delta T\ base\ heat_{d.n} - \tau_h]_+}{24}$, where
      d is the core heating day ($d = 001, 002, 003 \dots x$),
      n is the hour ($n = 01, 02, 03, \dots 24$),
      τ_h, determined above, is a constant that is part of CT/home's thermal/HVAC heating run time model, and

[11] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

$[\ ]_+$ indicates that the term is zero if its value would be negative.

- Calculate baseline run time as the sum of daily baseline run times for the set of core heating days in the CT interval data file

  $RT\ base\ heat\ (minutes) = \sum_{d=1}^{x} \alpha_h * daily\ HTD\ base_d\ (°F)$, where

  α_h , determined above, is a constant that is part of CT/home's thermal/HVAC heating run time model

  $daily\ HTD\ base_d$ is the baseline daily Heating Thermal Demand

  d is the core heating day ($d = 001, 002, 003 \ldots x$)

- CT heating savings is the percent heating run time reduction, for the reporting period in the CT input file.

  $$HS\ (\%\ RT\ reduction) = 100 * \frac{(RT\ base\ heat - RT\ actual\ heat)}{RT\ base\ heat}$$, where

  $RT\ base\ heat$ is the sum of baseline heating run times for all core heating days in the CT interval data file, and

  $RT\ actual\ heat$ is the sum of heating run times for all core heating days in the CT interval data file.

▪ Assess Cooling Savings

- Develop the CT/home's unique thermal/HVAC cooling run time model (hourly CTD method)

  - Calculate daily Cooling Thermal Demand ($daily\ CTD_d$) for each core cooling day in the interval data file:

    ▪ Calculate the average hourly indoor minus outdoor temperature difference ($hourly\ \Delta T_{d.n}$) for each core heating day:

      - Using the ZIP code included in the metadata file, determine the closest NOAA weather station that is in the same climate zone as the CT and look up hourly outdoor temperatures:

        $hourly\ outdoor\ T_{d.n}$, where

        d is the core cooling day ($d = 001, 002, 003 \ldots x$),

        n is the hour; ($n = 01, 02, 03, \ldots 24$)

      - $hourly\ \Delta T_n\ (°F) = hourly\ indoor\ T_{d.n} - hourly\ outdoor\ T_{d.n}$, where

        d is the core heating day ($d = 001, 002, 003 \ldots x$), and

        n is the hour; ($n = 01, 02, 03, \ldots 24$)

    ▪ Starting with an assumed value of zero for Tau ($\tau_c = 0$), calculate the daily Cooling Thermal Demand ($daily\ CTD_d$), as follows:

      $$daily\ CTD_d(°F) = \frac{\sum_{n=1}^{24}[\tau_c - hourly\ \Delta T_{d.n}]_+}{24}$$, where

      d is the core cooling day ($d = 001, 002, 003 \ldots x$),

      n is the hour; ($n = 01, 02, 03, \ldots 24$)

      τ_c is the ΔT associated with $CTD = 0$ (zero cooling run time), and

      $[\ ]_+$ indicates that the term is zero if its value would be negative

  - For the set of all core cooling days in the CT interval data file, use ratio estimation to calculate α_c, the homes responsiveness to cooling, which should be positive

$\alpha_c\left({}^{minutes}/_{^\circ F}\right) = {}^{RT\ actual\ cool}/_{\sum_{d=1}^{x} daily\ CTD_d}$, where

$RT\ actual\ cool$ is the sum of cooling run times for all core cooling days in the CT interval data file.

- For the set of all core cooling days in the CT interval data file, optimize τ_c that results in minimization of the sum of squares of the difference between daily run times reported by the CT, and calculated daily cooling run times. Next recalculate α_c (iaw the above step) and record the model's parameters (τ_c , α_c)[12]:

  $\sum_{d=1}^{x}(actual\ RT\ cool_d - daily\ RT\ cool_d)^2$ is minimized, where

  $daily\ RT\ cool_d = \alpha_c * daily\ CTD_d$

  $actual\ RT\ cool_d$ is the total daily cooling run time reported by the CT for that core cooling day

  d is the core cooling day ($d = 001, 002, 003 \ldots x$)

  Note that α_c characterizes the response of the home to cooling and τ_c is the difference between inside and outside temperatures when cooling run time = 0

- Calculate the cumulative baseline run time for the collection of all core cooling days in the CT interval data file (i.e. what would have occurred over the set of all core cooling days in the reporting period, had the home been held constant at the preferred cooling comfort temperature).
  - Calculate the baseline daily Cooling Thermal Demand ($daily\ CTD\ base_d$) for each core cooling day in the CT interval data file
    - Calculate the difference between the occupant's preferred comfort temperature for cooling and the average outside temperature for each hour of each core cooling day ($hourly\ \Delta T\ base\ cool_{d.n}$):

      $hourly\ \Delta T\ base_n\ (^\circ F) = T\ base\ cool - hourly\ outdoor\ T_n$, where

      d is the core heating day ($d = 001, 002, 003 \ldots x$),

      n is the hour ($n = 01, 02, 03, \ldots 24$),

      $T\ base\ cool$ is the occupant's preferred comfort temperature for cooling, and

      $hourly\ outdoor\ T_n$ is the average outdoor temperature for that hour reported by the nearest NOAA weather station in the same climate zone as the CT.
    - Calculate baseline daily Cooling Thermal Demand ($daily\ CTD\ base_d$)

      $daily\ CTD\ base_d = \frac{\sum_{n=1}^{24}[\tau_c - hourly\ \Delta T\ base\ cool_{d.n}]_+}{24}$, where

      d is the core cooling day ($d = 001, 002, 003 \ldots x$),

      n is the hour ($n = 01, 02, 03, \ldots 24$),

      τ_c , determined above, is a constant that is part of CT/home's thermal/HVAC cooling run time model, and

      $[\ ]_+$ indicates that the term is zero if its value would be negative.
  - Calculate baseline run time as the sum of daily baseline run times for the set of core

---

[12] Thermostat module calls Scipy Function leastsq, which uses Levenberg-Marquardt as implemented in MINPACK: LMDIF 1980

cooling days in the CT interval data file

$RT\ base\ heat\ (minutes) = \sum_{d=1}^{x} \alpha_c * daily\ CTD\ base_d\ (°F)$, where

α_c , determined above, is a constant that is part of CT/home's thermal/HVAC cooling run time model

$daily\ CTD\ base_d$ is the baseline daily Cooling Thermal Demand

d is the core cooling day ($d = 001, 002, 003 \ldots x$)

- CT cooling savings is the percent cooling run time reduction for the reporting period in the CT input file.

$$CS(\%\ RT\ reduction) = 100 * \frac{(RT\ base\ cool - RT\ actual\ cool)}{RT\ base\ cool}$$, where

$RT\ base\ cool$ is the sum of baseline cooling run times for all core cooling days in the CT interval data file, as detailed for each method, and

$RT\ actual\ cool$ is the sum of cooling run times for all core cooling days in the CT interval data file.

- Resistance Heating Utilization (RU) – for heat pump systems only, calculate RU in twelve outdoor temperature bins (0≤T<5°F, 5≤T<10°F,…, 55≤T≤60°F):

$$RU = \frac{(t_{emerg} + t_{aux})}{(t_{emerg} + t_{comp})}$$

where,

t_{emerg} = total emergency resistance heating run time in the interval data file, t_{emerg} is in lieu of compressor operation, e.g. use of electric resistance strip-heat under fault conditions or when the outside temperature is very low,

t_{aux} = total annual auxiliary resistance heating run time in the interval data file, t_{aux} is supplemental to compressor operation, e.g. use of electric resistance strip-heat to increase heating capacity, and

t_{comp} = total compressor heating run time in the interval data file. A proportion of t_{comp} may also include auxiliary resistance heating.

ASSESSMENT OF AGGREGATE NATIONAL AND REGIONAL SAVINGS

After the thermostat module generate an output file for the sample set, the statistics module leverages these files to assess average savings, expressed as percent run time reduction.

- Assess average heating savings, cooling savings, and RU for each of the five EIA climate regions, for example

CT Savings $_{\text{heat Mixed-Humid}}$ = ∑ CT$_i$ Savings $_{\text{heat Mixed-Humid}}$ / *n*, where

i = 1 thru n, and

n = the number of CTs in the EIA Mixed-Humid climate region (min. 250)

- Nationally – from regional savings, national savings are calculated as weighted averages for heating and cooling. Weighting is by the proportion of national energy used in each of the five EIA climate regions, as follows:

Heating	
Climate Region	%BTU vs National
Very Cold/Cold	54.9%
Mixed-Humid	31.2%
Mixed-Dry/Hot-Dry	5.4%
Hot-Humid	4.9%
Marine	3.6%

Cooling	
Climate Region	%BTU vs National
Very Cold/Cold	9.6%
Mixed-Humid	33.8%
Mixed-Dry/Hot-Dry	14.3%
Hot-Humid	41.8%
Marine	0.5%

- All other National outputs, including National RHU, are calculated as an unweighted average over all CTs in the sample set.

- **SOFTWARE DEVELOPMENT ROADMAP**

Beta – EPA is developing a Beta version of the software which will be used for a data call to inform the Draft 3 Version 1.0 ENERGY STAR Connected Thermostat specification. Some or all of the following features will be included:

1. The baseline heating comfort temperature will be the 90th percentile of the indoor temperature history reported by the CT on core heating days.
2. Similarly, the baseline cooling comfort temperature will be the 10th percentile of the indoor temperature history reported by the CT on core cooling days.

EPA believes that T_{indoor} is more indicative of comfort preferences and the change from using $T_{setpoint}$ will further establish consistency with T_{indoor} being used to both assess comfort preferences and to develop home's thermal/HVAC run time models. This change is expected to improve the savings assessment as the majority of HVAC load is driven by the difference between T_{indoor} and $T_{outdoor}$ rather than between $T_{setpoint}$ and $T_{outdoor}$.

3. CT savings will continue to be assessed using the 3 candidate methods described in this document.
4. CT savings will be assessed against two baselines, as follows[13]:

[13] This feature will be implemented if it can be accomplished without undue schedule impact

a. Baseline run times calculated from 90th (10th) percentiles of heating (cooling) indoor temperature history.

b. Baseline run times calculated from regional average comfort temperatures[14].

EIA Climate Zone	Very Cold/Cold	Mixed Humid	Mixed Dry/Hot Dry	Hot Humid	Marine
Baseline heating temp (°F)	68	69	69	70	67
Baseline cooling temp (°F)	73	73	75	75	–

Note: The 2009 RECS does not include statistically significant data to inform the average baseline cooling temperature for the Marine climate zone.

Inclusion of this additional baselining method is informational. EPA intends to launch the version 1.0 ENERGY STAR CT program that assesses savings against baseline run times calculated from 90th (10th) percentiles of heating (cooling) indoor temperature history.

5. Mean regional and weighted average national savings will be assessed both with and without data filtering. Data filtering will be used to exclude CTs from assessing mean savings. Savings data will be separately output with:

   a. No filtering,

   b. Exclusion of CTs where Tau < 0°F or >25°F,

   c. In addition to filtering on Tau, CTs will also be excluded where CV(RSME) >0.6; that is, the error associated with the thermal / HVAC run time model exceeds 0.6

   d. In addition to filtering on Tau and CV(RSME), CTs will also be excluded where % savings is:

      i. in the top or bottom 1% of all CTs in that EIA climate zone,

      ii. in the top or bottom 2% of all CTs in that EIA climate zone, and

      iii. in the top or bottom 5% of all CTs in that EIA climate zone

   EPA collaboration with CT stakeholders has identified the need to filter out CTs where the thermal / HVAC run time models are poor. EPA reserves the right to adjust the above filtering parameters based on additional feedback from CT stakeholders.

6. Additional outputs:

   a. Percentage of CTs removed by each filter (by region and nationally)

   b. Output lower bound of 95% confidence interval for each % RT reduction output

[14] Average comfort temperatures for heating and cooling were developed from the 2009 EIA Residential Energy Consumption Survey (RECS)

V1.0 – EPA is planning a version 1.0 release in the November 2016 timeframe. Version 1.0 will include a single CT savings methodology and is intended to be used by CT Service Providers for qualification of Connected Thermostats as ENERGY STAR.

V1.x – EPA may develop version 1.x releases. "Dot" releases after launch of the ENERGY STAR CT program will not impact assessment of savings for qualification as ENERGY STAR, but may include minor user experience bug fixes, enhancements or additional features to inform future development of CT field energy savings assessment methods.

- **PROGRAM IMPLEMENTATION**

As EPA currently envisions the program implementation, ENERGY STAR CT partners will use the software to assess savings from their CT products for initial product certification, for periodic reporting and optionally for internal use.

Submission of the software output file will be required for both initial product certification and periodic reporting. Initial product certification will be through a Certification Body (CB) while periodic reporting will be with EPA.

Partners must strictly adhere to input data requirement for both initial ENERGY STAR certification and for ongoing annual reporting. These requirements will be included in the ENERGY STAR CT Method to Demonstrate Savings and in the software documentation. At this time, EPA anticipates input data requirements will include:

INITIAL PRODUCT CERTIFICATION

1. Submitted output file uses ≥ 150 CTs in each of the five EIA climate regions, (or more if needed to provide sufficient certainty of the statistical results)
2. Same start and stop dates for each CT interval data file
3. ≥ 6 months duration between start and stop dates

PERIODIC REPORTING

1. Submitted output file uses ≥ 150 CTs in each of the five EIA climate regions, (or more if needed to provide sufficient certainty of the statistical results)
2. Same start and stop dates for each CT interval data file, as specified in the Partner Requirements
3. 12 month duration between start and stop dates