January 16, 2012

Home Energy Performance Scores:

Efforts to Date with Modeling Tool Comparison and Summary of Key Issues

Prepared by:

MetaResource Group

Portland, Oregon

For:

Energy Trust of Oregon, Inc.

Table of Contents

Home Energy Performance Scores: Efforts to Date with Modeling Tool Comparison and Summary of Key Issues

Authors: Jennifer Stout and Steven Scott, MetaResource Group
January 16, 2011

INTRODUCTION

This report summarizes Energy Trust's efforts to date to establish an "asset-based" energy performance metric for existing homes, and to test and compare available modeling tools. The report describes yet unresolved issues with energy performance scoring of homes in general.

This report does not make recommendations on specific energy performance scoring tools, next steps, or policies. Its purpose is to help frame the issues and act as a useful reference document for various entities including Energy Trust, local utilities, Bonneville Power Administration, realtors, homebuilders, and city and state governments. Note that this report is a snapshot in time since tool developers continually revise and update their tools.

This report discusses energy performance scores that are "asset-based" rather than "operationally-based." "Asset-based" scores remove occupant behavior, and consider only the dwelling's physical structure, the applicable climate, and a standard set of operating parameters (e.g., thermostat settings). The region is moving strongly toward asset-based energy performance scoring tools. Other tools (e.g. OPOWER) are being tested as mechanisms to influence occupant behavior.

A useful analogy to an asset-based home energy score is the Miles per Gallon (MPG) rating on a new car. The Environmental Protection Agency (EPA) calculates the MPG ratings using a set of controlled conditions[1]. While individual car owners' actual driving behavior and circumstances rarely match these conditions exactly, the rating allows buyers to compare cars.

Dwellings are more complicated than cars; they differ substantially in vintage (age), square footage, climate, heating systems, architecture, and fuel types. To overcome these challenges, different entities in the Northwest and nationally have been developing and field-testing asset-based modeling tools over a number of years. The concept is that, like MPG, simple, consistent, and reasonably accurate home energy scores will tell homeowners and buyers how their

[1] www.fueleconomy.gov/feg/how_tested.shtml

existing or prospective dwelling performs relative to others, and that this information will motivate them to make decisions regarding efficiency improvements and purchases that result in regional energy savings.

MetaResource drew the information in this report primarily from the following Energy Trust reports. Key report content is summarized at the end of this report.

- *Energy Trust Residential Home Energy Review: Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data*, Jennifer Stout and Steven Scott, MetaResource Group, December 2011
- *Energy Performance Score Modeling Comparison*, Dave Robison, Stellar Processes, December 2011
- *Energy Performance Score 2008 Pilot – Findings and Recommendations Report*, Earth Advantage Institute, Conservation Services Group, and Energy Trust, August 2009
- *Engineering Review and Process Evaluation of the Energy Trust New Homes Program*, Research Into Action, January 2010
- *Billing Analysis of Energy Performance Score Homes*, Jason Christensen, Energy Trust, May 2011
- *Process Evaluation of the 2009-2010 New Homes Program*, Research Into Action, September 2011
- *Homeowner Communications Research,* DHM Research, October 2011
- *Existing Homes Energy Performance Score: 2011 Pilot Initiative Summary*, Energy Trust, January 2011

A BRIEF HISTORY OF HOME ENERGY SCORES

In 1997, Denmark was the first country to adopt legislation requiring that all homes and commercial properties obtain and disclose energy scores. Two years later, Australia adopted similar legislation, though only for the residential sector. In 2003, the European Union (EU) adopted legislation for both the residential and commercial sectors; those laws and the scoring systems took effect in 2009. Under the laws, nearly every EU property owner must obtain an energy audit and score prior to selling his or her home or commercial building and must disclose that score to potential buyers. Currently, a process is under way to improve the policy further.[2]

In the United States, over a dozen states and municipalities have considered disclosure and/or upgrade legislation, with seven new policies in place as of the summer of 2009.[3] In 2010, the

[2]Dunsky Energy Consulting and Northeast Energy Efficiency Partnerships, *Valuing Building Energy Efficiency through Disclosure and Upgrade Policies – A Roadmap for the Northeast US*, November 2009. See page 21 of the NEEP report for a full list of countries and U.S. cities and states who have adopted, or are considering, EPS policies. http://neep.org/uploads/policy/NEEP_BER_Report_12.14.09.pdf

[3]See Dunsky/ NEEA, page 21.

U.S. House and Senate debated legislation to encourage widespread adoption of upgrade policies and disclosure of the energy scores of homes.

In Oregon, a team of the Energy Trust, Earth Advantage, PECI, and Conservation Services Group developed the Energy Performance Score (EPS). The development team has had an EPS in place for new construction available since mid-2009, and piloted a system for existing homes in 2008.

In 2009, Senate Bill 79 created the 13-member *Oregon Task Force on Energy Performance Scores.* SB 79 charged the Task Force with [1] making recommendations to the Oregon Department of Energy (ODOE) regarding a voluntary scoring system, and [2] reporting to the Oregon state legislature regarding a mandatory scoring system.[4]

The Task Force's recommendations to ODOE regarding a voluntary building energy scoring system formed the basis of administrative rules that went into effect July 1, 2010; the rules spell out a consistent methodology for building energy scoring, the metrics and format for displaying the score, and software approval requirements. One key element was that the scoring system for both existing and new dwellings be asset-based.

The majority of Task Force members did not support mandatory building energy scores. Their recommendations to the state legislature were[5]:

- For residential scoring, physical inspection of the building should be required.
- Those engaged in the business of producing building energy scores should be required to have a certification from Building Performance Institute (BPI) or Residential Energy Services Network (RESNET).
- Either USDOE or ODOE should approve the software tools used for producing energy performance scores.

In 2010 in Seattle and Bellingham, Washington, Earth Advantage began piloting an energy performance score for existing homes, based on the analytical engine called SIMPLE. Conservation Services Group has also been developing and testing an enhanced version of its current modeling tool Energy Measure Home for forecasting savings in Energy Trust's established, existing homes program.

---

[4] http://www.oregon.gov/ENERGY/docs/reports/legislature/2011/oregon_taskforce_on_energy_performance_scores_2010.pdf?ga=t

[5] In making its recommendations, the Task Force drew on a number of reports, and particularly on *Valuing Building Energy Efficiency Through Disclosure and Upgrade Policies – A Roadmap for the Northeast US*,[5] published in 2009 by Dunsky Energy Consulting and the Northeast Energy Efficiency Partnerships (NEE). This report also recommended use of an asset-based tool for residential home energy performance scoring.

The US Department of Energy (DOE) created its own Home Energy Score (HES) tool that utilizes an asset-based version of Home Energy Saver with benchmarks based on Residential Energy Consumption Survey (RECS) data. The DOE announced ten tests in various U.S. states and cities in 2011, including one through the Energy Trust.

In the summer of 2011, Energy Trust ran a pilot residential home scoring effort delivered through its Home Energy Review program. Two of the three groups received home energy performance scores (one received the Energy Trust's Energy Performance Score (EPS), the other the DOE's "Home Energy Score" [HES]). The third group received a recommendations report but no score. Analysis of Internet surveys of these pilot participants is included in this report.

DISCUSSION OF KEY ISSUES

Below is a discussion of key issues regarding energy performance scoring tools and home energy performance scoring in general.

1. A number of findings suggest that home energy performance scores may not be a key driver for homeowners to make efficiency upgrades to their existing homes, either while living there, or in anticipation of selling them. However, other findings indicate that many homeowners do believe energy performance scores can be useful in a home sale or purchase. Key findings are described below:

    - The Stellar Processes report reconfirmed that numerical improvements in a home's energy performance score from efficiency upgrades are generally small relative to the overall score. See point #2 below for a more in-depth discussion of this point.

    - MetaResource Group analyzed the Internet survey responses of participants from the Home Energy Review (HER) pilot groups to assess if there were statistically significant differences among the groups' self-reported measure follow-through rates.[6] If there were differences, MRG sought to determine if the following key features of the pilots might be factors in those differences: one group included only the new Custom Home Energy Report (CHER); the second included the new CHER plus the EPS; the third included the CHER plus the HES. (All three groups received their site visits around June.) However, MRG found no

[6] In this report, follow-through is defined as the percentage of participants in each pilot and in the original HER program who installed one or more of nine measures incented by Energy Trust within three months of receiving their on-site visit

statistically significant differences among the follow-through rates reported by the three groups at about three months out.[7] (Participants in the original HER program were not surveyed.) That being said, readers should bear in mind two factors that could temper the strength of these results. The first factor is that only three months had elapsed between the site visits and the surveys; one year would be preferable (even though all three groups received their site visits and surveys during the same time period). The second factor is that there were only 14 survey respondents in one of the two groups. While MRG used the Comparison of Means Test (Student t-test) for small samples, the small sample should still be kept in mind.

- MRG did find statistically significant differences among the follow-through rates calculated by Energy Trust staff using the Fast Track data on measure incentives paid for the three *full populations* of pilot participants (not just Internet survey respondents), as well as for a recent population of participants in the original HER program.[8] However, discussions with the staff revealed that it is most likely that factors *other than* the addition of the new CHER and home energy scores account for these differences in follow-through rates. MetaResource Group discusses these factors in detail in their report *Energy Trust Residential Home Energy Review: Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data*, Jennifer Stout and Steven Scott, MetaResource Group, December 2011 (pages 8-10). To briefly summarize, one key issue is "contamination" of the site visit recruitment pool on the part of the contractor who conducted the site visits. Other factors include the timing of the site visits and follow-up analysis, and types of measures participants installed (expensive versus inexpensive).

- MetaResource Group's analysis also found that respondents from the two pilots groups who received a score were not statistically more or less satisfied with their program experience than those respondents who did not receive a score. The most prominent factor in satisfaction was the energy advisor's on-site visit (audit). Homeowner responses to surveys conducted by Earth Advantage for their report also reflected that homeowners value audits.

---

[7] The self-reported follow-through rates from the Internet surveys were 11 percent for the CHER-only pilot, 8 percent for the CHER + EPS pilot, and 15 percent for the CHER + HES pilot.

[8] The staff analysis found that Fast Track follow-through rates were 4 percent for the CHER Pilot, 6 percent for the EPS Pilot, 25 percent for the HES Pilot, and 13 percent for the original HER program.

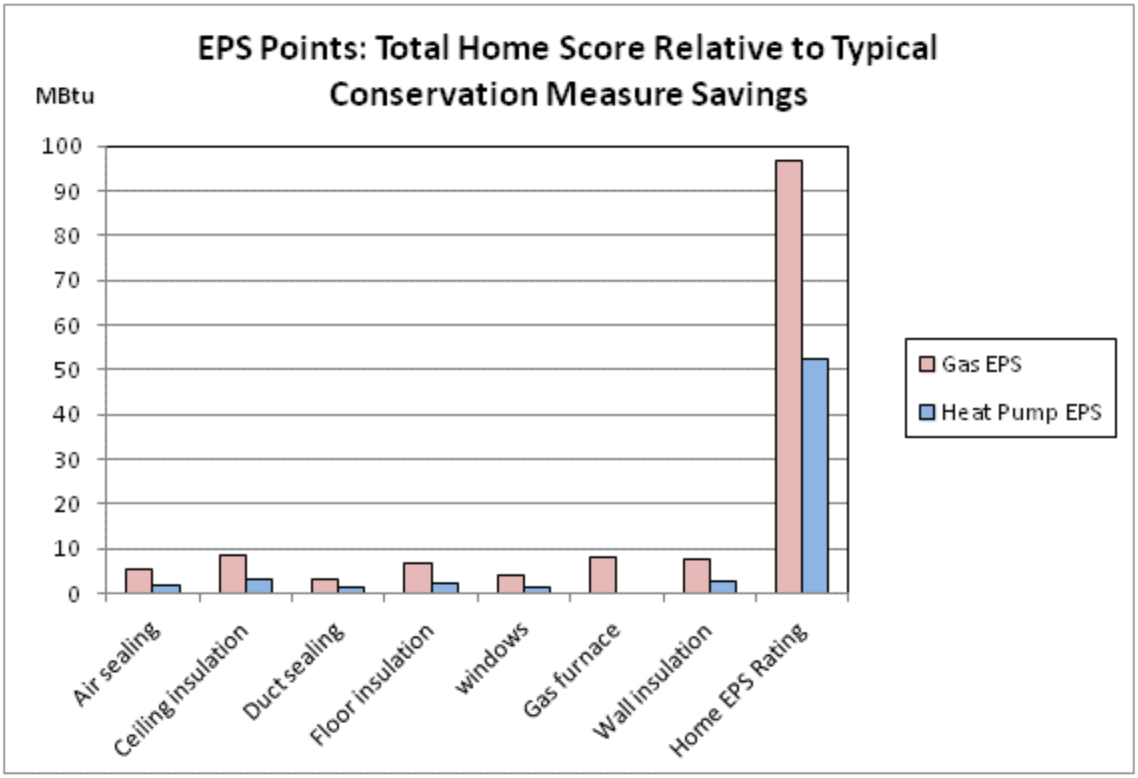
- MetaResource Group's analysis also found that 46 percent of the survey respondents who received a home energy score (a combination of the two pilot groups) said they thought an energy score would be useful if they sold their home, and 83 percent said they would want to see the score if they were buying a home. While these are positive results, they must be interpreted as speculative since energy performance scores are not yet a significant factor in the existing homes market – consumers do not know how useful they will be.

- In the open-ended survey questions, pilot participants expressed particular interest in information on energy costs and cost savings, and wanted Energy Trust to provide more information, suggesting that this is a key driver. Homeowner responses to surveys conducted by Earth Advantage for their report also emphasized the importance of cost.

2. The Stellar Processes report reconfirmed that numerical improvements in a home's energy performance score from efficiency upgrades are generally small relative to the overall score. The modesty of the potential change in score could diminish the impact of energy performance scores on consumers' decisions to upgrade, and could reduce the effectiveness of a score as a market signal to potential buyers of a previously-owned home, assuming the buyer is comparing the scores of two or more homes with similar characteristics (vintage, square footage, climate, heating system, architecture, and fuel type).

   Figure 1 below (Stellar Processes Report Figure 2) illustrates this issue. Note that the improvement in a home's score is even smaller for home with a heat pump than a home with a gas furnace.

Figure 1



On the other hand, a logical corollary to the above finding is that homes of markedly different vintages may have substantially different energy performance scores because of the building energy codes in place at the time of construction (all other things being equal). This could potentially encourage the purchase of newer homes. If this is a concern for any of the region's stakeholders, it could be addressed by normalizing the scores by using different "reference averages" for different vintages. However, normalizing might eliminate the potential for these differences in scores to encourage further efficiency improvements in older homes.

3. Findings from both the Stellar Processes and DHM Research Reports indicate there are some technical and communication challenges related to consumers' understanding of scores. For example, Stellar Processes points out consumers may not be familiar with "MBtu. If a normalized scale is used, such as 0 to 100 units, consumers may wonder what the units mean. A further challenge is that different tools may define their normalized scales differently in terms of whether a high or a low score is better. Indeed, DHM Research found that many of the participants in their focus groups did not understand that the lower scores presented to them – those closer to zero – were better than the higher scores. DHM also noted that not all participants grasped the

analogy of the home energy score to an automobile's MPG rating because they did not have any context for what the energy performance score meant. Some in the industry have suggested that a more meaningful approach for the consumer would be to normalize to a percentage scale, such that the consumer would see that one home's energy use is, say, "110 percent of average" and another's is "120 percent of average." Suffice to say that the choice of scale and units has been the subject of long, hard conversation.

4. Another complexity in the scoring of homes is the issue of establishing the appropriate "reference average" to which a home's individual score is compared so that the consumer can understand how their particular home performs relative to others. The comparison is only meaningful if the "reference average" is derived from homes whose combination of characteristics is similar to that of the individual home being scored. Relevant characteristics include vintage, square footage, climate, heating system, architecture, and fuel type. Many characteristics mean many possible combinations, and so many "reference averages" must be created. Energy Trust has created some of these reference averages, but this reality introduces complexity for both the person doing the energy performance scoring and those interpreting them.

5. Stellar Processes points out that if one compares energy performance *scores* and energy *costs* for homes with systems of different fuel types, the question arises of whether Energy Trust's message should be "save energy" or "save money." In some cases, a home's energy performance score favors one fuel or system type, while cost analysis favors another. For example, all other things being equal, heating a home with a gas furnace will require more BTUs than heating a home with an electric heat pump (if BTUs are measured at the site).[9] Thus, the home with the gas furnace will have a worse score than the home with an electric heat pump. However, the dwelling with the gas furnace will <u>cost</u> less to operate. Should Energy Trust's message to consumers should be "save energy" or "save money"? Table 1 below illustrates this issue.

Table 1: Scores and Operating Costs Compared for Heat Pump Versus Gas Furnace

[9] This assumes that one comparing the two homes based on their "site energy" use. *"Site energy"* is the amount of energy a dwelling consumes "on site," as reflected in its utility bills. "Site" and "source" energy are discussed further in point #5.

Portland Prototype Dwelling	SEEM Energy Score	Operating Cost
Prototype A: 2,200 SQFT Heat pump	69	$1,889
Prototype B: 2,200 SQFT Gas furnace	113	$1,607

6. Stellar Processes found challenges with energy performance scores related to fuel type. Below is a discussion of ways this issue might be approached.

    - Compute the score for a home of a particular fuel type *only* by comparing its energy usage to the reference average for that fuel type. This means one would score a home with a *heat pump* by comparing it to a reference average for homes with *heat pumps*, and one would score a home with a *gas furnace* by comparing it to a reference average for homes with *gas furnaces*.

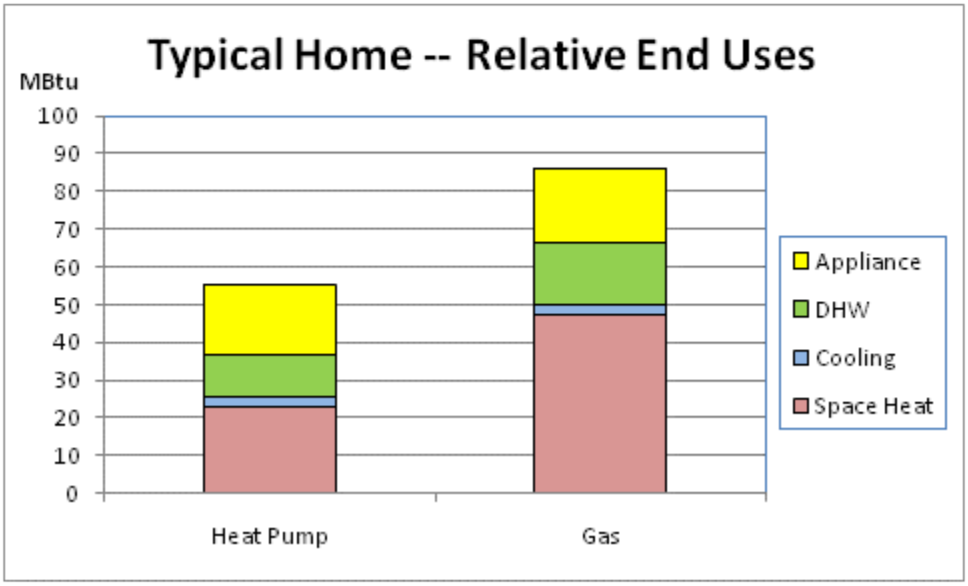
    - Normalize the scores to percentage scales. As noted in point #2 above, if a tool bases its score on MBtu, the meaning of the score's units may be unclear to the consumer. A more meaningful approach for the consumer may be to normalize to a percentage scale. Thus, a consumer considering buying a home with a gas furnace might learn that the home uses "110 percent of average *for homes that have gas furnaces.*" Similarly, a consumer considering buying a home with an electric heat pump might learn that the home uses "110 percent of average *for homes that have electric heat pumps."* This approach would help the consumer understand how efficient an individual home is – with a particular type of heating system, and a particular profile (vintage, square footage, climate, energy efficiency measures, etc.) – in comparison to homes that have the same type of heating system. However, such a scale would not encourage conversion of resistance-heated homes to whole house or ductless heat pumps, which is one of the most significant programmatic ways to change heat use in electrically heated homes.

    - Normalize scores by fuel. Alternatively, scores could be normalized by fuel. This would mean that if you scored two homes that were identical in every way *except* that one had a heat pump and one had a gas furnace, there scores would be the same. Normalizing by fuel could also work for homes with resistance heat, ductless heat pumps, and hearths, but doing so would create considerably more complexity in the energy performance scoring process.

    - Another way to handle the issue of fuel type is to use "source energy" rather than "site energy" to compare the two homes' energy use, and to generate their energy performance scores. "Site energy" is the amount of energy a dwelling consumes "on site," as reflected in its utility bills. "Source energy" is the total primary (raw) fuel needed to delivery heat and electricity to the dwelling. Source energy incorporates transmission, delivery, and production losses. If one assumes that the source electricity is generated by a gas-fired power plant, these losses would mean the dwelling with the electric heat pump would actually use more BTUs of source energy, and therefore have a worse score, than the home heated with the gas furnace (the scores would reverse from those calculated using site energy). However, a confounding factor is that the generation of source energy varies with each utility's resource plan, and individual utility plans themselves change over time. Further, a tool user has a choice between inputting the source energy for the "marginal" electrical plant (i.e., the next unit of generation) or the "average" electric plant (all units combined). The "marginal" electric plant will use less source energy than the "average" electric plant because of increases in turbine efficiency. Thus, the choice would have a significant impact on a home's energy performance score.

In Figure 2 below, the relative BTU usage of end uses is shown using site energy. If source energy were used instead of site energy, the red block for space heat for the heat *pump* would actually be *larger* than the red block for space heat for the gas furnace.

Figure 2 – Relative End Uses Using "Site Energy"



7. Stellar Processes' tool testing found that home energy performance scoring tools are different from one another in their computational methodologies and assumptions, and the resulting scores are not consistent from one tool to another. For example, if one used Tool A to score one home, and Tool B to rate a second home, the two scores might be different from the two scores if one were to use just Tool A to score both homes. The two approaches might lead to different conclusions about which home is more efficient.

   Stellar Processes' study revealed this lack of consistency. Table 2 below shows how two similar homes ranked relative to one another, based on their energy scores; the scores were not consistent from tool to tool, thus the rankings "flip-flopped" from one tool to the next. That is, given that the lower ranking is preferable, Home #2 was preferred according to Ecotope's Simplified Energy Efficiency Model (SEEM) and Earth Advantage SIMPLE 2.0, but not according to EnergyMeasure™HOME (EMHome), a proprietary tool developed by Conservation Services Group (CSG), and Energy Savvy.

   It matters which model is chosen and whether the model is validated and upgraded over time. Vetting and designing models and defining what model is appropriate for a market is not a simple, fast, or inexpensive task. Whether one or a set of models is used in the market, and how this selection is done are significant issues not addressed here.

Table 2: Comparison of Rankings of Two Homes Based on Scores from Different Tools

	SEEM	EMHome	Earth Advantage	Energy Savvy
Home #2	7	4	3	8
Home #10	8	2	4	7

8. Three of the five tools tested by Stella*r Proc*esses (Energy Savvy, EMHome, and Earth Advantage SIMPLE 2.0) produced home energy performance scores that were reasonably similar to the scores generated by Ecotope's Simplified Energy Enthalpy Model (SEEM).[10] There are two key reasons that it is important that tools used in the Northwest be able to produce scores similar to those of SEEM. First, SEEM is a more sophisticated tool than the five tools tested by Stellar Processes (it accounts for a greater number of parameters for structure, systems, and end uses); therefore, if three of the tools' results are similar to SEEM's, it is an indication that these three are fairly sound technically. Second, Northwest planners use SEEM as their standard tool for detailed technical studies of dwelling energy use; it is important to have regional alignment between planning tools and tools used in the field. Regarding the other two tools, Stellar Processes found that Home Energy Score (HES) tool, developed by the U.S. Department of Energy, overestimates total consumption, and Recurve has discrepancies with regard to heating and cooling equipment performance. The report suggested that Energy Trust remove these from consideration in the form they were at the time of the study.

9. As an overall observation, the discussion above suggests that the application of home energy performance scores to existing homes presents a number of challenges. To re-cap, these challenges include: small differences in scores from efficiency upgrades; consumer understanding of scales and scores; the need for an appropriate "reference average"; the issue of fuel and system type; cross-fuel comparison; and inconsistency in scores from one tool to another. While owners of existing homes who responded to the Internet survey did say scores would be useful in the sale and purchase of an existing home, these responses must be interpreted as speculative since scores are not yet a significant factor in the real estate market – consumers do not know how useful they will be. The challenges make the scoring process complex, and require owners, buyers, and real estate agents to be judicious in how they use and interpret scores. If Energy Trust can overcome these challenges, it can then apply further criteria to tool selection

[10] The Northwest Power and Conservation Council commissioned Ecotope to develop. SEEM is a more detailed tool than the commercially available ones. It models sophisticated interactions within the structure, and allows explicit examination of such parameters as duct losses and air leakage. In addition, SEEM has been tested to confirm that results match actual cases.

such as cost of data collection, software, obtaining inputs and outputs, and doing upgrades as needed. Eventually, they will be able to weigh all these factors against each other, and make a decision on whether to implement home energy scores widely in the existing home market.

REPORT OVERVIEWS

Below is an overview of recent key reports on home energy performance scoring tools and related issues. Reports on existing home are summarized first, followed by reports on new homes.

REPORTS ON HOME ENERGY SCORING FOR EXISTING HOMES

Below is a summary list of the existing home reports that MetaResource Group has summarized:

- *Energy Performance Score Modeling Comparison*, Dave Robison, Stellar Processes, December 2011
- *Energy Trust Residential Home Energy Review: Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data*, MetaResource Group, December 2011
- *Energy Performance Score 2008 Pilot – Findings and Recommendations Report*, Earth Advantage Institute, Conservation Services Group, and Energy Trust, August 2009
- DHM Research, *Homeowner Communications Research,* October 2011; Coates Kokes, *Energy Trust Focus Groups*, October 2011 and *Implications of Focus Findings*, November 2011
- *Oregon Task Force on Energy Performance Scores: Report to the Oregon Legislative Assembly* September 2010
- Sentech, *Review of Selected Energy Auditing Tools in Support of a National Building Performance Assessment and Rating Program*, November 2010

***Energy Performance Score Modeling Comparison*, Dave Robison, Stellar Processes, December 2011**

The goals of the Stellar Processes report were:

- Determine if the different tools result in consistent and similar scores for existing homes.
- Determine if the scores are comparable with SEEM results.
- Recommend which tools Energy Trust should consider.

- Recommend tool improvements.

The tools analyzed were:

- EnergyMeasure™HOME (EMHome): This tool is a proprietary tool developed by Conservation Services Group (CSG).
- Home Energy Score Tool (HEST): this tool was developed by the United States Department of Energy (USDOE).
- Earth Advantage (EA): Simple 2.0 is a tool developed by Michael Blasnik and is used by Earth Advantage Institute in their programs in Washington and other states around the country.
- Energy Savvy: Is a tool intended for homeowner use, and collects substantially less information than the other tools surveyed. Energy Savvy is the current web tool that Energy Trust provides on its website for customers to develop home energy profiles.
- Recurve, a newly developed modeling tool.

MetaResource Group did not include a summary herein of the key findings of this report because we discuss them in the main report.

***Energy Trust Residential Home Energy Review: Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data*, MetaResource Group, December 2011.**

MetaResource Group analyzed Internet surveys that were conducted by Conservation Services Group (CSG) of participants in the existing home Energy Peformance Score pilot based on Energy Trust's"Home Energy Reviews" (HER).

Participants in two of the pilot groups received home energy scores. One group received a score generated by the Home Energy Score Tool (HEST), developed by US DOE. The other group received a score generated by EnergyMeasure™HOME (EMHome), developed by Conservation Services Group (CSG). In addition, MetaResource Group (MRG) reported results of Energy Trust staff's analysis of measure follow-through data from the "Fast Track" program-tracking database.

MetaResource Group did not include a summary herein of the key findings of this report because we discuss them in the main report.

***Energy Performance Score 2008 Pilot – Findings and Recommendations Report*, Earth Advantage Institute, Conservation Services Group, and Energy Trust, August 2009**

In 2008-2009, with the assistance of Earth Advantage Institute and Conservation Services Group, Energy Trust tested four home energy scoring tools: REM/Rate, Earth Advantage Simple 2.0, and two versions of Home Energy Saver Tool (HEST).

The consultants audited about 300 homes, and used 190 in the study. The study compared the tools' predictions of energy consumption for the homes, to the homes' weather-normalized utility billing data. Since utility billing data reflects occupant behavior, behavioral factors were also included in the modeled predictions, even though Energy Trust's intent was that the EPS would be an "asset-based" tool, excluding the impact of occupant behavior.

Below are key study conclusions followed by our recommendations.

Earth Advantage and Conservation Services Group's Conclusions on Tools

- The four tools did not predict actual consumption with a level of accuracy acceptable to Energy Trust.
- SIMPLE generally produced the most accurate results with a mean absolute percent error of 25.1 percent compared to HES-Full with 33.4 percent, REM/Rate with 43.7 percent, and HES-Mid with 96.6 percent. In other words, SIMPLE predicted energy use on average within plus or minus 25.1 percent of actual use.
- Comparisons to billing data are not an accurate test of a models' forecasting efficacy because billing data reflect homeowner behavior.
- A set of enhancements to the less-complex energy models might improve accuracy.

Earth Advantage and Conservation Services Group's Conclusions on Stakeholder Feedback

Earth Advantage Institute conducted a series of stakeholder surveys over the course of the pilot, the responses to which guided the conceptual development of the EPS. Key findings included the
following:

- The EPS concept appeals considerably to consumers, home improvement contractors, builders, and real estate professionals.
- Energy costs and cost savings are major issues and comprise the common language for understanding energy.
- The EPS must be clear and objective, and presented by a trustworthy source.
- Carbon emissions are relevant and very important to homeowners.
- Homeowners are most familiar with energy use in terms of Watts and kilowatt-hours.
- Homeowners want information on energy performance and where to make improvements.
- Homeowners perceive home energy audits as useful for a variety of reasons.
- Homeowners thought that their homes were more energy efficient than preliminary results indicated.

Based on the software analysis, research, and survey findings, Earth Advantage and Conservation Services Group made the following recommendations for the development of an Energy Performance Score (EPS). MRG has abbreviated each recommendation; the reader will find the full text in the report.

1. Develop the EPS along two tracks: the EPS with energy and carbon scores and a performance profile of energy related elements of the home (EPS *score only*), and the EPS that additionally includes recommendations for energy upgrades (EPS *w/ upgrades*). Have trained and certified third-party auditors conduct the auditing for both tracks.

2. In order to offer a credible level of accuracy, EPS certified software programs should be able to predict energy use within 25 percent for 70 percent of homes, and within 50 percent for 90 percent of homes when compared to actual use.

3. The EPS energy score should be expressed as the total annual energy required for the house under normal conditions and be expressed in kilowatt-hours per year.

4. The EPS should include a carbon score that reflects the greenhouse gas emissions associated with the home's energy use.

5. The EPS should include an energy analysis report with an accounting of the annual estimated energy use and fuel costs for major end-uses. For existing homes, the report may also include recommendations for energy upgrades and the associated costs and predicted savings.

6. Ideally, the EPS will be a coordinated effort to ensure consistency of the core tool elements and modeling protocols.

At the time the report was published, the consultants noted that Energy Trust and Earth Advantage Institute were moving ahead with the refinement of SIMPLE, as well as working with recent relevant legislation and efforts at home energy auditing in Oregon to develop and further test the EPS concept in different climate zones.

DHM Research, *Homeowner Communications Research,* October 2011; Coates Kokes, *Energy Trust Focus Groups*, October 2011 and *Implications of Focus Findings*, November 2011

DHM Research conducted two focus groups on September 26-27, 2011, in Portland. The focus groups were to support Coates Kokes' (CK) communication strategy work with Energy Trust. Eighteen people participated in the groups, and all were homeowners who lived in the Portland metro area.

The main objective of the focus groups were to assess homeowners' reactions to different branding approaches, and secondarily, to test the effectiveness of Energy Trust's key messages and the understandability of the Home Energy Recommendations Report and the Energy Performance Score (EPS).

The focus group participants were not participants in Energy Trust's energy performance score pilot Home Energy Review (HER) Program, and saw the sample EPS for the first time in the focus group. Below is a summary of the key focus group findings:

- The participants greatly valued the *Home Energy Recommendations Report.* They found that it was easy to understand, conveyed important information, and would help them make the most important energy saving improvements to their homes.
    - All the participants said that the report would be helpful.
    - They said that the charts were easy to understand and that the report clearly indicated which actions they should prioritize for the most benefit.
    - While the participants appreciated that the report indicated the incentives available and the long-term cost savings, some would have liked more information about the upfront costs of the recommendations.

- The participants valued the concept of the *Energy Performance Score*; however, they had difficulty interrupting the actual results in the report.
    - Nearly all the participants said that the *Energy Performance Score* would be valuable for making decisions about their homes, when selling their homes, and when buying new homes.
    - Several participants struggled to read the graphs in the report. They were confused that lower values meant better results – this seemed counterintuitive to them.
    - They did not all grasp the analogy of the *Energy Performance Score* to an automobile's miles per gallon rating. Unlike an MPG rating, they did not have any context for what a score of, say, 15, meant for carbon emissions.
    - Several participants said that comparing their results on an *Energy Performance Score* to the average Oregon home was not valuable. They would prefer more specific comparisons, such as to homes in their neighborhood or homes of similar sizes and ages.

- Saving energy was important to customers, primarily to save money but also to help the environment.
    - The messages and value statements that most resonated with the participants connected saving energy to saving money and benefiting the environment.
    - Secondary reasons related to economic development, home comfort, or the source of energy.

***Oregon Task Force on Energy Performance Scores: Report to the Oregon Legislative Assembly* September 2010**[11]

In 2009, Senate Bill 79 created the 13-member *Oregon Task Force on Energy Performance Scores.* SB 79 charged the Task Force with [1] making recommendations to the Oregon Department of Energy (ODOE) regarding a voluntary scoring system, and [2] reporting to the Oregon state legislature regarding a mandatory home energy disclosure system.[12]

The Task Force's recommendations to ODOE regarding a voluntary building energy performance scoring system formed the basis of administrative rules that went into effect July 1, 2010; the rules spell out a consistent methodology for building energy performance scoring, the metrics and format for displaying the score, and software approval requirements. One key element was that the scoring system for both existing and new dwellings be asset-based.

The majority of Task Force members did not support mandatory building energy scores. Their recommendations to the state legislature were[13]:

- For residential scores, physical inspection of the building should be required.
- Those engaged in the business of producing building energy performance scores should be required to have a certification from Building Performance Institute (BPI) or Residential Energy Services Network (RESNET).
- Either USDOE or ODOE should approve the software tools for producing energy performance scores.

Sentech, *Review of Selected Energy Auditing Tools in Support of a National Building Performance Assessment and Rating Program*, November 2010[14]

Sentech prepared this report for the U.S. Department of Energy (DOE) in support of its effort to develop a national program, the Home Energy Score (HES) program, to assess the energy performance of houses. The program will provide information to current and prospective homeowners about the energy performance of the house and potential areas of improvement, along with associated cost estimates. As a component of this program, DOE is interested in understanding the variety and characteristics of currently available audit tools that have national validity. Of particular interest is the ability of these tools to accurately analyze residential building performance--regardless of climate, fuel source, architecture, and building

[11] http://www.oregon.gov/ENERGY/docs/reports/legislature/2011/oregon_taskforce_on_energy_performance_scores_2010.pdf?ga=t

[12] http://www.oregon.gov/ENERGY/docs/reports/legislature/2011/oregon_taskforce_on_energy_performance_scores_2010.pdf?ga=t

[13] In making its recommendations, the Task Force drew on a number of reports, and particularly on *Valuing Building Energy Efficiency Through Disclosure and Upgrade Policies – A Roadmap for the Northeast US*,[13] published in 2009 by Dunsky Energy Consulting and the Northeast Energy Efficiency Partnerships (NEE). This report also recommended use of an asset-based tool for residential home energy performance scoring.

[14] http://apps1.eere.energy.gov/buildings/publications/pdfs/homescore/auditing_tool_review.pdf

system--with a reasonable level of tool inputs (i.e. asset-based scores). Additionally, DOE is interested in the ability of these tools to produce reports on estimated fuel consumption and lists of recommended building energy efficiency improvements.

The report describes the following tools:

- RESNET-accredited tools: REM/Rate®, EnergyGauge®, Energy*Insights*®.
- Tools commercially available and commonly used for energy audits and home performance programs: BEACON Home Energy Advisor®, Home Energy Tune-uP®, TREAT®, and RealHomeAnalyzer.
- Other tools, either government-produced or benchmarking applications: HES*Pro*, NEAT®, and Green Energy Compass.

The key report finding was that no one tool fully captures all the characteristics currently thought to be important to a national home performance assessment program: low cost, universal availability, ease of use with reasonable input requirements, conformance to a universally accepted accuracy standard, and the ability to generate improvement recommendations and associated costs. However, the report does conclude that the audit tools as a population appear to address the potential needs of a national program.

Readers might be particularly interested in Tables 1, 2, and 3 on pages 25, 29, and 33. Each table is a matrix of audit tool criteria and attributes for different groupings of tools.

Table 2: Strength of EPS in Predicting Energy Consumption by Fuel Type

	Correlation
Gas heating load	25%
Annual therms	43%
Annual kWh	24%
Total annual Btus	54%

The analysis also demonstrated that the EPS tool met its goal of predicting energy use within 25 percent for 70 percent of homes, and within 50 percent for 90 percent of homes when compared to actual use. (The analysis found that approximately 75 percent of the homes are within 30 percent of predicted usage, and about 94 percent were within 50 percent.)

<u>Calculation of Energy Savings</u>

The calculation of average savings was based on a comparison between the energy consumption of the 91 gas-heated, single family, EPS homes built in 2008-2009, and the energy consumption of 307 gas-heated, single family, ENERGY STAR® homes built in the Northwest in 2006-2007. (The latter analysis had already been done by KEMA.) At that time, the requirements for ENERGY STAR® homes were very similar to the most recent Oregon code. The consumption of the EPS Homes was determined based on billing analysis. Gas savings were nine percent and were statistically significant; electrical savings were 3 percent but were not statistically significant.

APPENDIX A

December 13, 2011

Residential Home Energy Review

Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data

Prepared by:

MetaResource Group
Portland, Oregon

For:

Energy Trust of Oregon, Inc.

Table of Contents

Energy Trust Residential Home Energy Review

Analysis of Pilot Group Internet Survey Results and Energy Trust Fast Track Data

December 13, 2011 MetaResource Group

Introduction

Energy Trust has offered the "Home Energy Review" (HER) program to residential customers since 2003. The program provides participants a free on-site home visit from an energy advisor who then prepares a written report describing the dwelling's energy performance and recommending measures to improve it.

In 2011, Energy Trust conducted three residential pilot programs that were based on the original HER program, but had distinct new features. Below are descriptions of the three pilot programs. The dwellings in all three pilots used natural gas as their primary heat source and were located within the Portland metropolitan area.

- New Custom Home Energy Report (CHER) ("CHER Pilot"): Participants in this pilot program received (1) the same on-site visit as provided in the original Energy Trust HER program, and (2) a newly designed written report called a "Custom Home Energy Report" (CHER).

- CHER plus "Energy Performance Score" (EPS) ("EPS Pilot"): Participants in this pilot program received (1) and (2) as described above for the CHER Pilot, plus an "Energy Performance Score" (EPS). Conservation Services Group (CSG) developed the EPS.

- CHER plus "Home Energy Score" (HES) ("HES Pilot"): Participants in this pilot program received (1) and (2) as described above for the CHER Pilot, plus a "Home Energy Score" (HES) for the dwelling. The US Department of Energy developed the HES.

Energy Trust hired Conservation Services Group (CSG) to conduct Internet surveys between June and August 2011 of pilot participants. The purpose of the surveys was to assess participants' opinions of various aspects of the pilots, and participants' self-reported "follow-through" on energy efficiency measures. Energy Trust defines follow-through as the installation in a dwelling of at least one incented energy efficiency measure after the program participant has received the on-site visit and written report. For the HER program, nine incented measures[15] were applicable to the types of dwellings that participated in the pilot programs, namely existing dwellings heated primarily with natural gas.

[15] Because the participating dwellings were existing ones heated primarily with natural gas, they were eligible for nine incented measures: air sealing, ceiling insulation, duct insulation, duct sealing, floor insulation, tanked water heater, tankless water heater, wall insulation, and windows.

In addition to the self-reported information on follow-through from survey respondents, Energy Trust staff collected and analyzed follow-through data from the "Fast Track" program-tracking database. The Fast-Track data were based on the records of when the Energy Trust paid measures incentives to customers. The same nine measures were analyzed as were asked about in the Internet surveys.

The following bullets show the number of participants in each of the three pilot programs, and the number of Internet survey respondents. Internet surveys were not administered to participants in the original HER Program.

- CHER Pilot: 180 participants; 53 survey respondents; 29% response rate[16]
- EPS Pilot: 227 participants; 76 survey respondents; 33% response rate[17]
- HES Pilot: 55 participants; 20 survey respondents; 36% response rate[18]

Energy Trust then hired MetaResource Group (MRG) to analyze both the results of the Internet surveys and the data from the Fast Track database. Energy Trust asked MRG to focus on the following four topics:

- Follow-Through: The percentage of participants in each pilot and in the original HER program who installed one or more of nine measures incented by Energy Trust within three months of receiving their on-site visit, and how these percentages compare (1) among the three groups of pilot participants, and (2) between the pilot participant groups and the participants in the original HER program
- Satisfaction: Participant satisfaction with the pilot programs
- Preference for EPS versus HES: Whether participants in the EPS Pilot liked their scoring system more than participants in the HES Pilot did
- Opinion of CHER: Pilot participant perceptions of the newly designed Custom Home Energy Report (CHER)
- Statistically significant differences, if any, among the three groups of pilot program participants related to the issues in the three bullets above

Key Takeaways

- Follow-Through: Based on analysis of the Internet surveys, the new CHER and the two different home energy scores (HES and EPS) do not appear to be driving differences in follow-through rates. MRG did not find statistically significant differences among the follow-through rates self-reported by the Internet survey respondents. MRG also looked at the results of analysis of

[16] In Appendix A, readers can find a copy of the newly designed CHER, embedded within the CHER Pilot Internet survey.
[17] In Appendix B, readers can find a copy of the newly designed CHER, and a copy of the EPS, both embedded within the EPS Pilot Internet survey.
[18] In Appendix C, readers can find a copy of the newly designed CHER, and a copy of the HES, both embedded within the HES Pilot Internet survey.

follow-through conducted by Energy Trust staff using Fast Track data for the full populations of pilot participants, and for a recent population of participants in the original HER program. While MRG did find statistically significant differences among these follow-through rates, discussions with the staff revealed that factors *other than* the differences in features of the pilots themselves most likely account for these differences. These factors are discussed in the detailed findings section.

- Satisfaction: MRG's analysis of the Internet survey responses indicated that respondents from all three pilot programs are very satisfied with their program experience. These results are similar to those of two additional past surveys of participants in the original HER program; one survey was conducted in 2010 and the other in the second quarter of 2011. Our analysis of Internet survey responses also indicated that respondents from the two pilots groups who received an EPS or HES score in their written report were not definitively more or less satisfied than those the survey respondents from the CHER Pilot who did not receive a score. In open-ended responses to the Internet surveys, respondents particularly praised the energy advisors. However, a number of respondents said the scheduling process was too complicated.

- Preference for EPS versus HES: The Internet surveys asked a series of six questions regarding participants' perceptions of the scores' clarity and usefulness. Respondents could answer "yes," "no," or "unsure." A majority of respondents from both the EPS Pilot and the HES Pilot responded "yes" to all six questions. While there were variations between the two pilot groups in how large each majority was, MRG's analysis found only one question that appeared to indicate a preference for the HES over the EPS. On whether participants perceived their score as "clear and visual," the mean of responses for the EPS was significantly lower than for the HES score at the 95 percent and 99 percent confidence levels.

- Opinion of CHER: Respondents indicated they found the new Custom Home Energy Report (CHER) useful, but gave lower ratings for the energy cost and savings information than for the other question elements. Further, in the open-ended responses, respondents asked that Energy Trust provide more information on energy costs and cost savings. MRG recommends that Energy Trust conduct further research on these issues, perhaps through focus groups. Although Energy Trust conducted focus groups previously, they did not cover some of the topics discussed in this report.

- Sample Sizes: Because only 20 HES Pilot participants responded to the Internet survey, Energy Trust might consider conducting additional focus groups with pilot participants, particularly from the EPS and HES Pilot groups, to obtain additional feedback on the scoring tools and other topics.

Detailed Findings

Detailed Findings on Follow-Through Actions: Energy Trust defines participant "follow-through" as the installation of at least one of nine measures for which Energy Trust offers an incentive.[19] Data analyzed for this report were collected three months after homeowners received their site visits. For the three pilot groups, participants received their site visits around June. For the original HER, participants received their site visits around March.

- MetaResource Group's analysis of the Internet survey results indicated that key differences in the features of the three pilots (most importantly, the addition of a new Custom Home Energy Report (CHER) and the two home energy scores [EPS and HES]) do not appear to be driving differences in follow-through rates. MRG found no statistically significant differences among the follow-through rates reported by the Internet survey respondents from the three different pilot participant populations. These rates were 11 percent for the CHER Pilot, 8 percent for the EPS Pilot, and 15 percent for the HES Pilot. (As noted earlier, CSG did not survey participants in the original HER program.) However, these findings carry two caveats. One is that only three months had elapsed between the site visits and the surveys; one year would be preferable. The other is that there were only 14 survey respondents in one of the two groups. While MRG used the Comparison of Means Test (Student t-test) for small samples, the small sample should still be kept in mind.

  MRG did find statistically significant differences among the follow-through rates calculated by Energy Trust staff using the Fast Track data for the three *full populations* of pilot participants (not just Internet survey responses), as well as for a recent population of participants in the original HER program. The staff analysis found that follow-through rates were 4 percent for the CHER Pilot, 6 percent for the EPS Pilot, 25 percent for the HES Pilot, and 13 percent for the original HER program. However, discussions with Energy Trust staff indicated that it is most likely that factors ***other than*** the addition of the new CHER and the home energy scores account for these differences. MRG describes these factors below:

  - *"Contamination" of Site Visit Recruitment Pool:* Conservation Services Group (CSG), which conducts the site visits for the standard HER program, and conducted them for each of the three pilots, was having difficulty fulfilling their site visit quotas for the CHER and EPS pilots because there was a limited pool of residential customers who had requested site visits. To meet their quotas, CSG resorted to recruiting participants from a pool consisting of customers who had requested only Energy Trust's free "Energy Saver Kit," (ESK), but who had not requested a free site visit. CSG persuaded these ESK customers to participate in the free site visit. Since these site visit invitations were an "upsell" by CSG, and these customers had originally only ordered the free ESK, Energy Trust staff believe these customer had not intended to get a site visit or install incented measures, and therefore their rates of follow-through were significantly lower than the HES Pilot and the HER group. The lower follow-through rates for those participants who had only requested the ESKs brought the follow-through rates for the entire CHER and

[19] Because the participating dwellings were existing ones heated primarily with natural gas, they were eligible for nine incented measures: air sealing, ceiling insulation, duct insulation, duct sealing, floor insulation, tanked water heater, tankless water heater, wall insulation, and windows.

EPS Pilot groups down. (See Table 1, section titled "Follow-Through Rates for Pilot Participant Populations."

- *Seasonality:* Seasonality may be another factor in the lower follow-through rates for the CHER and EPS Pilots (4 and 6 percent respectively). On average, the CHER and EPS groups received their site visits around June. Energy Trust staff have found, not surprisingly, that customers who receive a site visit in the late fall and winter when the weather is colder, follow through at higher rates than customers who receive a site visit in the late spring or summer; "seasonality" drives follow-through. However, the HES group received its site visits in June, and had a follow-through rate of 25 percent, so seasonality does not hold up in this case.

- *Measure Type/Cost:* Another finding was that there was a statistically significant difference in the follow-through rate for the population of HER participants (13 percent) versus the population of HES participants (25 percent). (This difference was statistically significant at the 95 percent confidence level although not the 99 percent confidence level.) One possible reason for the lower follow-through rate by HER participants as compared to the HES participants was that based on analysis of the Fast Track data, the HER participants installed relatively more expensive measures – new windows and insulation of various kinds – than did participants in the HES Pilot.

Table 1 below shows the follow-through rates and the results of the difference of means tests.

Table 1: Follow-Through by Participants: One or More Measures Incented

Pilot Name	Follow-Through Rates for Internet Survey Respondents
CHER Only (n = 52) (Site visits around June)	11%
EPS (n = 75) (Site visits around June)	8%
HES (n = 20) (Site visits around June)	15%
HER – not surveyed	N/A
	Follow-Through Rates for Pilot Participant Populations
CHER Only (N = 180)(Site visits around June)	4%
EPS (N = 227) (Site visits around June)	6%
HES (N = 55) (Site visits around June)	25%
HER (N = 485) (Site visits around March)	13%
	Comparison of Means Tests for Internet Survey Respondents
CHER Compared to EPS	No statistically significant difference found
CHER Compared to HES	No statistically significant difference found
EPS Compared to HES	No statistically significant difference found
	Comparison of Means Tests for Pilot Participant Populations and Original HER Program Participant Population
CHER Compared to EPS	No statistically significant difference found
CHER Compared to HES	Statistically sig. diff. at 95% and 99% confidence levels
EPS Compared to HES	Statistically sig. diff. at 95% and 99% confidence levels
Original HER Program Compared to CHER	Statistically sig. diff. at 95% and 99% confidence levels
Original HER Program Compared to EPS	Statistically sig. diff. at 95% and 99% confidence levels
Original HER Compared to HES	Statistically sig. diff. at 95% confidence level but not at the 99% confidence level

Detailed Findings on Program Satisfaction:

- Survey respondents from all three pilot groups expressed high levels of program satisfaction overall (a mean of 3.8 or above out of 5) on a number of different pilot program aspects. MRG selected the following aspects for analysis: “Overall experience,” “Energy Advisor’s knowledge of energy efficiency,” and “Packet of materials.” For “Overall experience,” 99 percent of all survey respondents said they were ”Satisfied,” “Very satisfied,” or “Extremely satisfied”; 88 percent said “Very” or “Extremely.” These results are similar to those of two additional past surveys of participants in the original HER program. In one conducted in 2010, 93 percent of HER participants said they were “Satisfied” or “Very satisfied.” In the other conducted the first quarter of 2011, the result was 90 percent.

- MRG converted the Likert scales to five-point numerical scales for the questions related to satisfaction. Tables 2, 3 and 4 below[20] show that all of the 24 mean ratings were 4.0 or above, except two of 3.8. One rating of 3.8 was the mean response of HES Pilot respondents on the "Packet of materials" (see Table 2), and the other rating of 3.8 was for this same respondent group on their perception of how much they learned about the "Estimated utility cost savings after energy improvements are made" (see Table 4).

- As shown in Table 2, the mean rating of the EPS Pilot respondents was higher than the mean rating of the HES Pilot respondents on "Overall experience" and "Packet of Materials." On "Overall experience," the difference was significant at the 95 percent confidence level but not at the 99 percent confidence level based on a Comparison of Means test (Student t-test). Thus, a definitive interpretation of this result was difficult. On "Packet of materials," the difference was significant at both the 95 percent confidence and the 99 percent confidence levels. This might appear to show a preference for the EPS. However, as we explain below, this apparent preference is contradicted by EPS and HES Pilot participants' responses to a later multi-part question that asks directly about participants' perceptions of the scoring systems.

- MRG also noted that the mean ratings shown in Table 2 for the CHER Pilot and the EPS Pilot were almost identical; this finding suggests that for these respondents including the EPS did not affect satisfaction. MRG believes that the results of both the past and current surveys indicate that the home visit, rather than the CHER, is the most important factor contributing to participants' satisfaction.

[20] The Internet survey had more questions than MRG analyzed in this report. MRG selected those questions we judged most pertinent to the issues of interest to Energy Trust and most likely to reveal differences among the pilot groups.

Table 2: Participant Satisfaction[21]

	How would you rate your level of satisfaction with your Home Energy Review? Please indicate on a scale of Extremely Satisfied to Extremely Dissatisfied. **(1 to 5 Scale)**		
	Overall Experience	***Energy Advisor's Knowledge of Energy Efficiency***	***Packet of Materials***
Pilot Group	**Average**	**Average**	**Average**
CHER Only (n = 52)	4.4	4.5	4.3
EPS (n = 75)	4.4[22]	4.5	4.4[23]
HES (n = 20)	4.0	4.2	3.8
	Statistically sig. diff. between EPS and HES at 95% confidence level but not at 99% confidence level	No statistically sig. diff. found	Statistically sig. diff. between EPS and HES at 95% and 99% confidence levels
All Pilot Participants n = 146	***Overall Experience***	***Energy Advisor's Knowledge of Energy Efficiency***	***Packet of Materials***
Extremely satisfied	48%	55%	48%
Very satisfied	40%	37%	34%
Satisfied	11%	8%	15%
Very Dissatisfied	1%	1%	2%
Extremely Dissatisfied	0%	0%	0%

[21] MRG did not select the following two sub-questions for analysis: "scheduling" and "cash-back incentives from Energy Trust."

[22] Statistically higher mean than HES at 95% confidence level (two-tail). Calculated t = 2.11 (df = 92); critical is about 2.010.

[23] Statistically higher mean than HES at 95% and 99% confidence levels (two-tail). Calculated t = 2.76 (df = 92); critical is about 2.010.

Table 3: Participant Perceptions of HER Recommendations[24]

	How would you describe the recommendations you received? Please indicate your thoughts on a scale of Completely Agree to Completely Disagree. **(1 to 5 Scale)**		
	The Recommendations I Received Are Clear	***The Recommendations I Received Are Accurate***	***The Recommendations I Received are Useful***
Pilot Group	**Average**	**Average**	**Average**
CHER Only (n = 43)	4.5	4.4	4.4
EPS (n = 62)	4.6	4.2	4.4
HES (n =16)	4.4	4.4	4.6
	No statistically sig. diff. found	No statistically sig. diff. found	No statistically sig. diff. found

In Table 4 below, respondents from the HES Pilot have a lower mean rating than the respondents from the other two pilots on "estimated utility cost savings…" However, the Student t-test did not indicate that the difference is statistically significant.

Table 4: Participant Perception of How Much They Learned[25]

	How would you rate how much you learned during your Home Energy Review? Please indicate on a scale of I Learned a Lot to I Did Not Learn Anything. (Scale of 1 to 5)	
	Estimated Utility Cost Savings After Energy Improvements Are Made	***The Best Energy Improvements for Your Home***
Pilot Group	**Average**	**Average**
CHER Only (n = 50)	4.1	4.3
EPS (n = 60)	4.3	4.6
HES (n = 17)	3.8	4.5
	No statistically sig. diff. found	No statistically sig. diff. found

- The survey asked respondents to select from closed-ended lists the aspects of their Home Energy Review with which they were most, and least, satisfied. Table 5 below shows the results. MRG calculated the frequencies for each HER aspect by dividing the total number of times

[24] MRG analyzed all sub-questions on this question.
[25] MRG did not select the following four sub-questions for analysis: "current conditions and energy performance of your home"; "comfort, health, and safety benefits after energy improvements are made"; "estimated costs of energy improvements"; and "available resources to help make energy improvements (e.g. trade ally contractors, financing)."

respondents from all three Pilot groups selected that aspect divided by the total number of respondents that gave one or more responses to either part of the question (n = 135). MRG removed from the total those survey respondents that left both parts of the question blank.

Table 5: Respondent Satisfaction: Specific Program Aspects

Program Aspect	***What aspects of your Home Energy Review are you most satisfied with?***	***What aspects of your Home Energy Review are you least satisfied with?***
-	**n = 135**	
Energy advisor's level of knowledge	67%	10%
Overall assessment of your home	54%	11%
Summary of your home's energy efficiency needs	63%	14%
Contractor recommendations	25%	13%
Information on next steps	36%	14%
Other	7%	7%

- As shown above in Table 5, respondents most often selected the energy advisors' level of knowledge as the aspect they were most satisfied with (67 percent). Next was the summary of their homes' energy efficiency needs (63 percent). At the same time, 14 percent of other respondents also cited the latter program aspect as the aspect with which they were the *least* satisfied. Fourteen percent of respondents also cited information on next steps as the aspect with which they were the *least* satisfied. It is notable that 14 percent found the summary of their homes' energy efficiency needs the least satisfactory program aspect.

- MRG also coded and analyzed open-ended question asking for feedback about the pilots (all pilot groups combined). Survey respondents made many positive comments about their experience with the pilot programs, particularly about the energy advisors. Of note were four comments in which respondents mentioned their advisors by name. For example, "My home advisor [Name] was not only knowledgeable but he took me everywhere with him, except under the house, to point stuff out. He was great."

- Respondents did make some negative comments in the open-ended question about the HER, mostly focused on the following:

    - Respondents' perception that setting up a HER appointment was too long and complicated (6 comments)
    - Their perception that they did not get a timely response by Energy Trust staff or contractors to inquiries following the HER (2 comments)
    - Their perception that the HER was not specific enough to their home (6 comments)
    - Their comments that they would like more information on energy costs and costs savings (6 comments)

- MRG coded the open-ended responses and provides the results below. Of note are the six comments on what participants perceived as a lengthy and/or difficult process for getting a HER scheduled.

Table 6: Open-Ended Survey Respondent Feedback on HER Program

Comment	Number of Times
Scheduling process lengthy and/or difficult	6
More on costs and cost savings	6
More on incentives	5
A report more customized to home's characteristics	5
Time of use rates	2
More/different analysis of heat loss	2
Lack of follow up on Qs following getting HER	2
More info on incentives for homeowners to do work themselves	2
Feasibility of solar	2
Information on local contractors	1
Other	6
Payment for upgrades through energy bill	
Better advertising	
Use Smart Meters to see impact of changes in real time	
Visual data on energy use/savings.	
Prioritization for improvements	
More on water heaters	

  - *Recommendations:* MRG recommends that Energy Trust explore the processes for scheduling appointments and for following up on specific customer requests.

Detailed Findings on Energy Performance Score (EPS) versus Home Energy Score (HES):

- One pilot group received the EPS and the other the HES. The survey asked participants to answer "yes," "no," or "unsure" to six questions related to whether the scores were clear and useful. MRG analyzed the percentages and also converted the "yes" responses to 1, the "no" responses to 0 and the "unsure" responses to 0.5, making it possible to conduct a Comparison of Means test (Student t-test) to see if there were any statistically significant differences that might indicate that respondents favored one scoring system over the other. While there were variations between the two pilot groups in how large each majority was, MRG's analysis found only one question that appeared to indicate a preference for the HES over the EPS. On whether participants perceived their score as "clear and visual," the mean of responses for the EPS was significantly lower than for the HES score at the 95 percent and 99 percent confidence levels.

- Over 70 percent of both EPS and HES Pilot respondents answered "yes" to nine out of the 12 questions. Seen in a positive light, a number of respondents appear to have found the scores meaningful and useful; on the other hand, a number appear to have not.
- Fifty-seven percent of HES Pilot respondents said "yes" to "The score is useful" while 85 percent of EPS Pilot respondents said "yes." However, as shown below, MRG's Comparison of Mean's Test did not show that the difference in mean ratings on this question was statistically significant.
- On whether participants perceived their score as "clear and visual," the mean of responses for the EPS was significantly lower than for the HES score at the 95 percent and 99 percent confidence levels.
- The EPS mean rating was higher than the HES mean score on "The score will be useful if I sell my home" but MRG only found significance at the 95 percent confidence level.

Table 7: EPS and HES Pilot Participant Perceptions of Scores

	Understand How Score Calculated	**Score is Appropriate**	**Score is Clear and Visual**	**Score is Useful**	**Score Will Be Useful if I Sell My Home**	**Would Want to See Score if Buying a Home**
EPS (n = 54)						
Yes	84%	91%	87%	85%	50%	87%
No	9%	7%	4%	7%	11%	6%
Unsure	7%	2%	9%	7%	39%	8%
Average	0.87	0.92	0.92	0.89	0.69	0.90
HES (n = 14)						
Yes	93%	71%	100%	57%	29%	71%
No	0%	7%	0%	14%	43%	14%
Unsure	7%	21%	0%	29%	29%	14%
Average	0.96	0.82	1.0	0.71	0.43	0.79
	No stat. sig. diff. found	No stat. sig. diff. found	Stat. sig diff. at 95% and 99% conf. levels	No stat. sig. diff. found	Stat. sig. diff. at 95% conf. level (but not 99%)	No stat. sig. diff. found

- Participants in the EPS program received both an energy score and a carbon score and were asked which they thought was most important. Forty-six percent said both, 39 percent said energy, and 2 percent said carbon. When asked why they chose both or only energy, the common responses were that an energy score lets one relate one's performance to financial savings (eleven responses), it is easier to understand (four responses), and it is more believable than a carbon score (four responses).

    - *Recommendation:* Because the sample for the HES group was only 14, MRG recommends that Energy Trust consider conducting additional focus groups with pilot participants, particularly from the EPS and HES Pilot groups, to obtain additional feedback on the clarity of the score format and other topics.

Detailed Findings on Perception of Custom Home Energy Report (CHER): Energy Trust introduced a new customized version of the HER, the CHER.

- The Internet survey asked respondents how useful they thought eight different sections of their CHER were on a scale of "Extremely Useful to Extremely Unuseful" (sic). MRG selected six for analysis. MRG converted the Likert scales to five-point numerical scales for the questions. Mean ratings were highest for the information on (1) existing conditions, (2) where the home loses energy, and (3) recommended improvements. It appears that respondents found the information on energy costs and savings relatively less useful. Table 8 shows MRG's analysis of this question.

- MRG conducted a Comparison of Means test (Student t-test) and found no statistically significant differences among the three pilot groups in their perception of the usefulness of the CHER.

Table 8: Participant Perception of Usefulness of Custom Home Energy Report

	How useful are the following sections of your custom recommendations report? Please indicate on a scale of Extremely Useful to Extremely Unuseful (1 to 5 Scale)		
	Existing Conditions	***Recommended Improvements***	***Where Your Home Loses Heat (Graph***
Pilot Group	**Average**	**Average**	**Average**
CHER Only (n = 43)	4.0	4.0	4.1
EPS (n = 60)	4.1	4.2	4.1
HES (n =15)	4.0	4.0	4.0
	No statistically sig. diff. found	No statistically sig. diff. found	No statistically sig. diff. found
	Annual Estimated Energy Cost Savings (Graph)	***Current Annual Estimated Energy Costs (Graph)***	***Annual Savings for Recommended Improvements***

CHER Only (n = 43)	3.9	3.9	3.9
EPS (n = 60)	3.9	4.0	3.9
HES (n =15)	3.7	3.9	3.7
	No statistically sig. diff. found	No statistically sig. diff. found	No statistically sig. diff. found

- *Recommendation:* Energy Trust might consider revisiting how the CHER presents energy costs and savings to see if improvements might be possible.

Survey Response Rates and Sample Sizes:

- MRG's analysis was impacted by the fact that fewer customers participated in the HES Pilot than in the other two pilots, and their survey response rate was only 40 percent compared to about 90 percent for the other two. Only 20 participants responded, and on many of the individual questions, the removal of non-responses reduced the sample size even further. This sample size falls below 30, the minimum for optimal statistical analysis. Even using statistical methods designed for small samples wherever possible, on some of the topics of interest to Energy Trust and other stakeholders MRG was not able to reach definitive conclusions.
    - *Recommendation:* Energy Trust might consider conducting additional focus groups with pilot participants, particularly from the EPS and HES Pilot groups, to obtain additional feedback on the scoring tools and other topics.

APPENDIX B

Energy Performance Score

Modeling Comparison

For

Energy Trust of Oregon

By

David Robison, P. E.

Stellar Processes

Stellar Processes, Inc.

18490 Vista Drive

West Linn, OR 97068

(503) 827-8336

January 18, 2012

Table of Contents

List of Figures

List of Tables

Chapter 1 Executive Summary

Energy Trust defines the Energy Performance Score (EPS) as the estimated annual energy consumption in units of MBtu. The customer will also be provided with a reference baseline that represents average residential consumption for a similar house. This report examines the efficacy and accuracy of several modeling tools relative to a widely accepted benchmarking tool. The goals of this project are to determine if the different tools result in consistent and similar scores comparable with regional results.

The estimating tools reviewed in this study are:

- EnergyMeasure Home™ (EMHome): This tool is a proprietary tool developed by Conservation Services Group (CSG).
- Home Energy Scoring Tool (HEST): this tool was developed by the United States Department of Energy (USDOE).
- Earth Advantage (EA): Simple 2.0 is a tool developed by Michael Blasnik and is used by Earth Advantage Institute in their programs in WA and other states around the country.
- Energy Savvy: Is the current web tool that Energy Trust provides on its website for customers to develop home energy profiles.
- Recurve, a newly developed modeling tool.

Key Findings

1. Overall, the desired goal of a scoring method is achievable. Recurve and Home Energy Scoring Tool (HEST) were eliminated from consideration. Rankings were as follows:

Ordinal Ranking	Overall Rank
Earth Advantage	1
EMHome	2
Energy Savvy	3

2. Consistency is a major issue. That is, if Home A is compared to Home B, the relative choice should not be affected by the choice of scoring tool. All of the tools demonstrated some “flip flop” of relative ranking position compared to their SEEM ranking.
3. To accurately reflect efficiency improvement, one hopes that the tools are unbiased and reasonably capture end-uses. All of the tools suffer problems when one looks at end-uses. However, the differences are small relative to total consumption. Changes in the EPS score due to end-use efficiency improvements may not be accurately computed but any errors in the total rating number will be small.
4. The tools include different embedded assumptions regarding appliance energy use. These appliance differences can mask a difference in other end-uses. It would be helpful to develop a standard set of assumptions regarding appliance usage.

5. The reference benchmark will need further development. The Energy Trust reference was developed for gas-heated homes in the Willamette Valley. To extend use of the EPS scoring, reference benchmarks will be needed for other climates and for new construction[26].
6. The use of a gas-heated reference for heat pump cases is potentially an issue. A fuel-blind rating would change the reference baseline depending on the fuel type.

[26] Energy Trust Note: Stellar Processes was only given a reference benchmark for gas heated homes in the Willamette Valley because all of the homes in the field pilot were gas heated and in the Willamette Valley. The appropriate benchmarks for other fuel types and climate zones have been developed.

Chapter 2 Methodology

Introduction

The Energy Trust Existing Homes Program (Program) is in the process of introducing an energy performance score (EPS). This report examines the efficacy and accuracy of several modeling tools relative to a widely accepted benchmarking tool. The final decision regarding EPS policy will also take into account such factors as the relative cost of gathering inputs and simplicity of use. Such a decision is beyond the scope of this report. Instead, this report focuses solely on the issue of comparing the benchmarking results to a standard simulation tool.

The rating focuses on an asset-based performance score rather than trying to estimate operational energy consumption, given that the latter is highly dependent on occupant behavior.

The goals of this project are to:

- Determine if the different tools result in consistent and similar scores.
- Determine if the scores are comparable with SEEM results.
- Provide a recommendation of which tools should be considered by the Program.
- Provide recommendations for how the tools could be improved.

Energy Performance Score (EPS)

First, it is important to define the methodology behind providing a home rating score. The goal is to provide the consumer with a rating number that is easy to understand – similar to the miles-per-gallon rating on autos. This means that the rating should be "asset-based", depending upon the characteristics of the structure and independent of behavioral choices made by the owners. Of course, behavioral choices have a large impact on energy consumption but other owners may not have the same lifestyle. Thus, the "asset-based" EPS rating needs to focus on the physical structure, equipment and improvements, assuming "typical" inhabitants.

Energy Trust proposes to use the home's total energy consumption in units of MBtu to be the home's rating number. To facilitate comparison, customers will also be provided with a benchmark reference for average energy consumption.

To define the average, Energy Trust reviewed the billing records of customers. Currently, reference average consumption is defined for gas-heated homes in the Willamette Valley[27]. Eventually, a reference will need to be developed for homes in more extreme weather zones and also for homes constructed to a new code minimum. Electrically-heated homes are compared to the same reference which implies that heat-pump homes can be expected to achieve a lower rating score. The customer average annual consumption was found to be 86.5 MBtu.

[27] Energy Trust, Memorandum - Benchmark Methodology 4.25, 4/26/11.

There has been discussion that the benchmark reference could be adjusted for house size. That is, the customer will receive the score for their own house and also a comparison score for an average house of similar size. Methodology for the size-adjusted benchmark has not yet been finalized. For this report, there are a few instances where we present the size-adjusted benchmark for reference. We assumed a reference benchmark based on the ETO Study reference above.

$$Base\ Gas\ (MBtu) = 27.44 + 0.0228 * Square\ Footage$$

$$Base\ Electric\ (KWh) = 4816 + 2.39 * Square\ Footage$$

$$Average\ (MBtu) = Base\ Gas + Base\ Electric * 3.413/1000$$

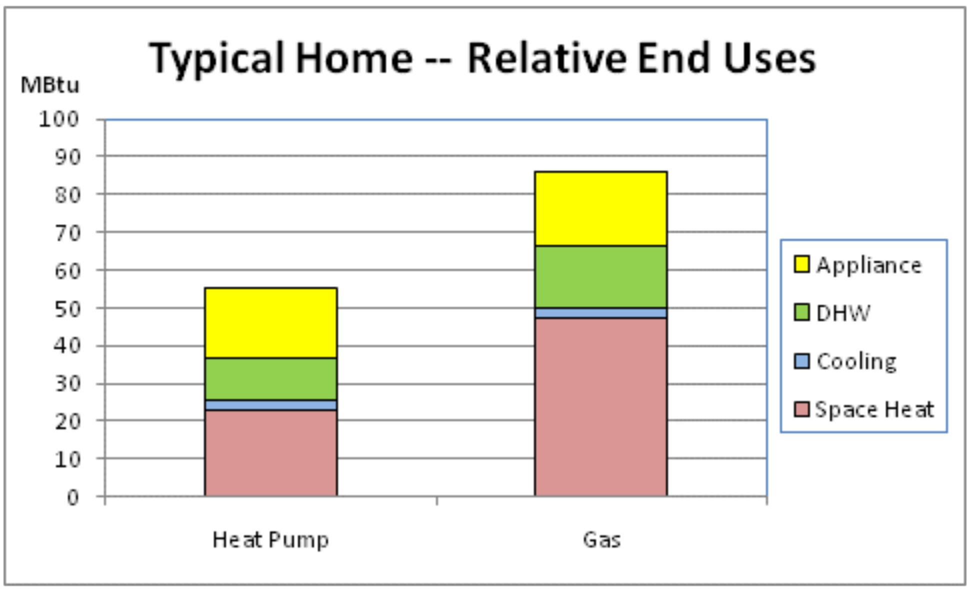


Figure 1. Relative End Uses

As shown in Figure 1, space heating accounts for about two thirds of the total MBtu consumption in a gas-heated home. The fraction is somewhat smaller in a heat pump home. Although we will examine the other end-uses, these are a smaller fraction of total consumption. Errors in parameters that relate to electrical consumption, such as cooling or residential lighting, will not have a major effect on total consumption. Parameters that affect gas space heating, such as insulation levels, duct losses and burner efficiency, will have an effect and these should be carefully considered.

EPS Score of Conservation Measures

The relative change in EPS score is likely to be low for conservation measures. Table 1, Table 2 and Figure 2 present some typical conservation savings from recent program evaluations[28] and the corresponding change in EPS score. (Score is computed as annual MBtu). As shown in Figure 2, relative impact of the conservation measures is small relative to the total house EPS score – and even smaller for heat pump homes. The change in EPS score is about the same size as the variances and discrepancies that occur between different tools. The fact that gas and heat pump homes are so different further complicates the desire to apply the same tool consistently for all customers. If the EPS score is to provide useful consumer feedback, accuracy will be important.

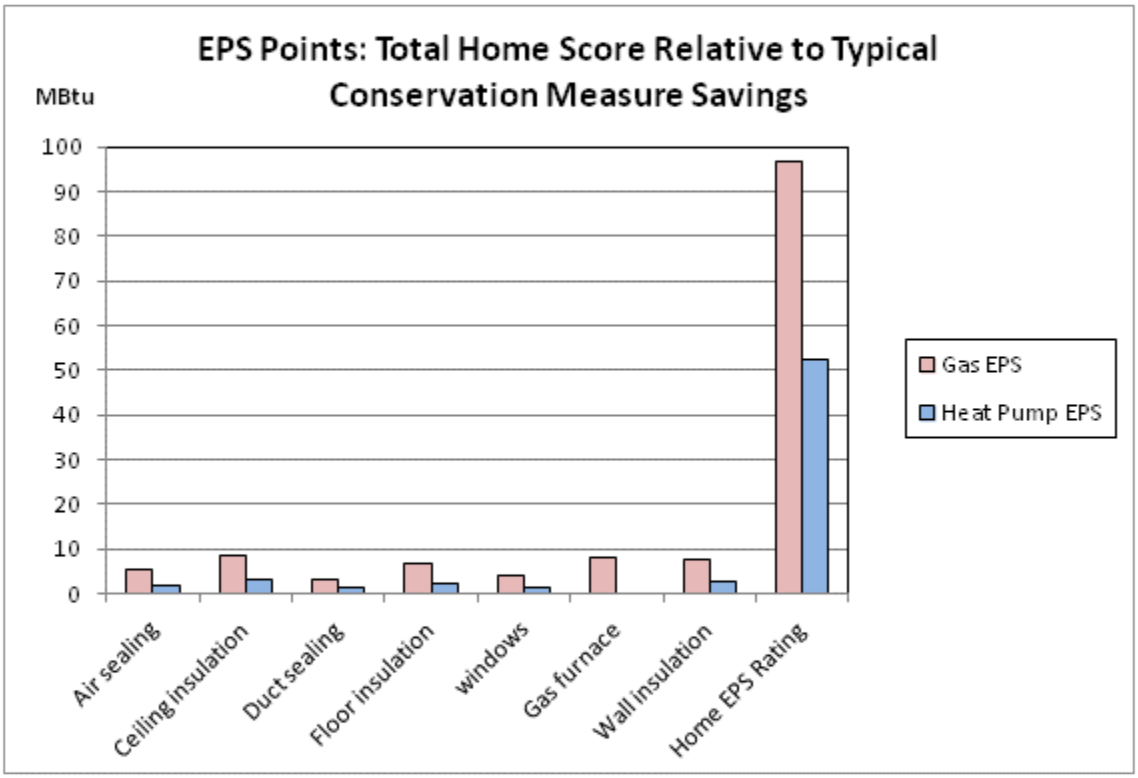


Figure 2. EPS Scores for Conservation Measures

[28] Source: Brien Sipe, 2009 Energy Trust gas impact evaluation estimates of savings for measures, Energy Trust of Oregon, 9/30/11.

Table 1. Typical EPS Score for Gas Conservation Measures

Measure	% of Space Heat	Therm Savings	Gas EPS Score Change
Air sealing	9%	53	5
Ceiling insulation	14%	86	9
Duct sealing	6%	33	3
Floor insulation	11%	66	7
windows	7%	41	4
Gas furnace	13%	79	8
Wall insulation	13%	75	8
Base Home EPS Rating -- Gas			97

Table 2. Typical EPS Score for Heat Pump Conservation Measures

Measure	% of Space Heat	Heat Pump EPS Score Change
Air sealing	9%	2
Ceiling insulation	14%	3
Duct sealing	6%	1
Floor insulation	11%	2
windows	7%	2
Wall insulation	13%	3
Base Home EPS Rating -- Heat Pump		53

Simulation Tools Compared to Standard

A variety of tools have been proposed to do the energy simulation. Technically, there are issues concerning how such tools assess such factors as duct leakage or equipment performance. SEEM is a tool used in the Northwest to explicitly simulate these factors. Accordingly, one part of this project was to compare the other tools to SEEM.

The energy performance score (EPS) tools examined are:

- **EnergyMeasure Home™ (EMH):** This tool is a proprietary tool developed by Conservation Services Group (CSG). This version of EMHome is designed to support a non-diagnostic visual audit in Oregon. Alternate configurations to support diagnostic testing and more detailed audit methodologies are in use in the Oregon Home Performance with ENERGY STAR® program, Clean Energy Works Oregon and elsewhere around the country.

- **Home Energy Scoring Tool (HEST):** this tool was developed by the United States Department of Energy (USDOE). It is currently being piloted around the country. Energy Trust is one of ten USDOE pilot partners.

- **Earth Advantage (EA):** Simple 2.0 is a tool developed by Michael Blasnik and is used by Earth Advantage Institute in their programs in WA and other states around the country.

- **Energy Savvy:** Is the current web tool that Energy Trust provides on its website for customers to develop home energy profiles. It is also being considered for use by the Energy Advisors.

- **Recurve**, a newly developed proprietary modeling tool without a previous program history in Oregon.[29]

- **SEEM:** The generally accepted modeling tool of the Pacific Northwest. This tool is not being considered for use by the Program, but provides a generally accepted baseline to compare the tools listed above.

SEEM is a more detailed modeling tool that was developed to model sophisticated interactions within the structure. SEEM allows explicit examination of such parameters as duct losses and air leakage. In addition, SEEM has been tested to confirm that results match actual cases. As a result, SEEM has been the standard tool within the Northwest region for detailed technical studies. The appendix includes a more detailed description.

The critical part of this analysis is defining some standard prototypical homes that can be compared across the various tools. We defined three prototypical buildings that are roughly representative of older, moderate age and new construction.

- 1,344 square feet. Small single story over crawl space. Insulation levels typical of older homes.
- 2,200 square feet. Medium split-level w/ some second floor space over garage. First floor over crawl space. Higher insulation level typical of 1980’s vintage homes.
- 2,688 square feet. Medium one story house over heated basement. Insulation level meets latest Oregon code.

It must be noted that these prototypes were arbitrarily defined – for example, there was no attempt to develop average component description from the overall customer population. In general, the simulation tools do not attempt to incorporate all possible construction details. That is, the tools work from a simple and less detailed description of the structure and equipment. Thus, part of the test is to see if there is a difference between these simple tools and the more detailed SEEM simulation.

[29] Recurve, Inc, 220 Montgomery St. Suite 820, San Francisco, CA 94104, www.recurve.com

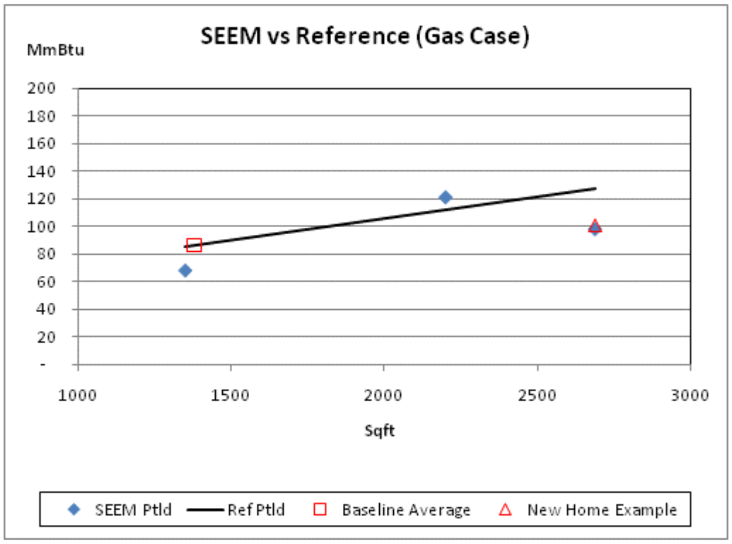


Figure 3. Prototypes Compared to Reference Baseline

Figure 3 shows how the SEEM prototypes compare to the reference baseline. The reference baseline is represented by a line since it is size-dependent. The three prototypes are shown as blue triangles. Note that these triangles are not exactly on the line because the structural components assumed are not necessarily "average". The Energy Trust study identified the overall average baseline (shown as a red square) and the new home baseline (shown as a red triangle). Note that the new home baseline agrees closely with the "new home" prototype. The two smaller prototypes match the ETO benchmark line fairly well.

We also modeled four different equipment types as follows:

- Zonal electric heat. Represents the actual thermal loads of the structure without duct losses or system inefficiencies.
- Electric furnace. Above with duct losses added.
- Gas furnace. Above with burner inefficiencies added.
- Heat pump. Furnace system with heat pump. Comparison with furnace allows for assessment of how heat pump performance is modeled.

Of course, only the gas and heat pump systems can be considered typical of what consumers install. The value of the intermediate system types is that it allows an assessment of how the various models deal with system inefficiencies.

We also modeled all the prototypes in three cities (Portland, Redmond and Medford) in order to have a diversity of climates. Thus, models were run were run for each home size, heating system and location, leading to 36 cases.

Note that the comparison is based on the overall modeled results. We did not attempt to verify the algorithms used within the programs to compute energy usage. That is, we did not require the modelers to use the same set of inputs. Instead, we asked them to use the standard defaults and assumptions that would apply where there are not on-site measurements. All modelers were provided with the same site information describing building components at the level expected from a simple on-site audit.

Test Group Implementation

Another part of this study was to compare how the various tools performed on field data collected from a test set of actual homes. For this purpose, we reviewed a group of 35 gas-heated homes and compared simulation results. This group of gas-heated homes was selected as representative of the current building stock but does not attempt to be a statistically selected sample.

The test group is comprised of recently constructed homes. They are generally large in size – 3 to 5 bedrooms. Most have conventional gas furnaces and water heaters. Two cases utilized a demand water heater and one case utilized a gas boiler and hydronic heating system. All were modeled assuming Portland climate.

Chapter 3 Overall Results

Summary of Model Comparisons

In general, the tools provide similar ratings for an EPS score. The issue is that one would like the consumers to be assured that efficiency improvements (such as weatherization, duct sealing or efficient equipment) will result in a reasonably accurate score improvement. As demonstrated, the tools vary in their ability to incorporate such component features. However, the differences are relatively small.

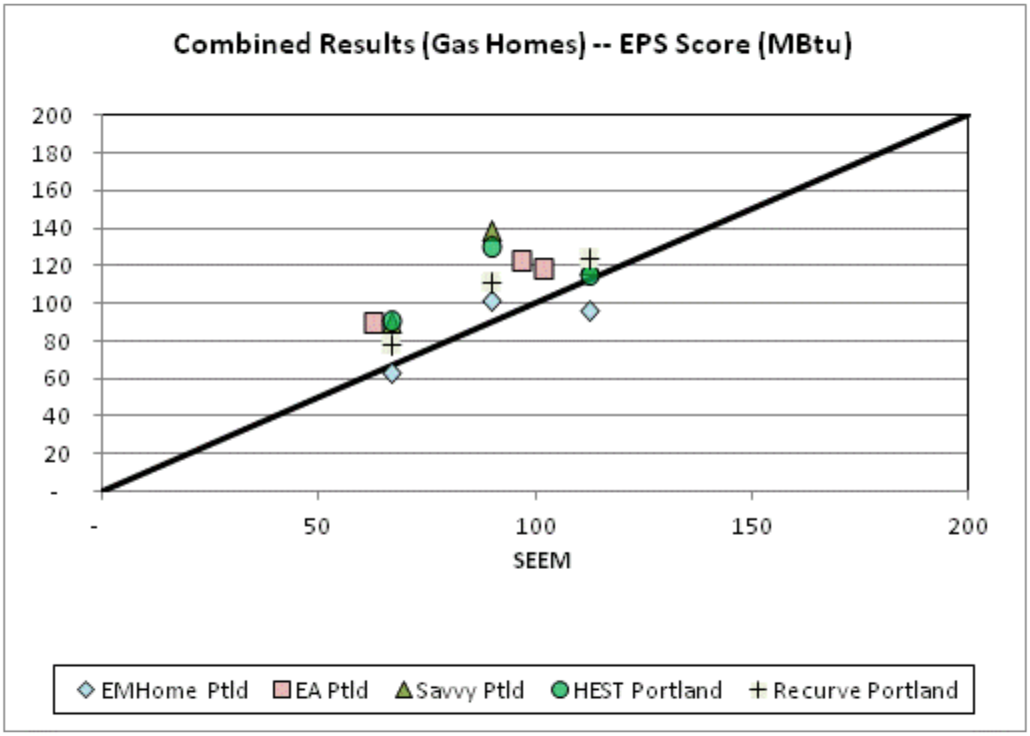


Figure 4. EPS Scores

In Figure 4 and subsequent charts, the x-axis represents the SEEM simulation result which is assumed to be the "correct" standard. The y-axis represents the result from the specific modeling tool. The 45-degree line is a reference line -- not a regression line. If the tools agree with SEEM, they would align along the 45-degree line. Points above the line are over-estimating and points below the line are under-estimating relative to SEEM.

Figure 4 shows the resulting EPS scores that would be computed for customers. Since the baseline reference was specified for the Willamette Valley, we are only looking at the Portland cases. And since gas homes are the current program focus, we limited discussion to that system type. Generally, the tools have a tendency to over-estimate. EMHome appears to provide the most reasonable estimates.

Figure 5 shows how a "typical" case might compare with the various tools. This typical case represents an average gas home with consumption of 120 MBtu. Looking at the gas cases, most of the tools are close to the SEEM score; Recurve has the largest discrepancy. The same homes treated with a heat pump system generally result a lower score relative to the reference.

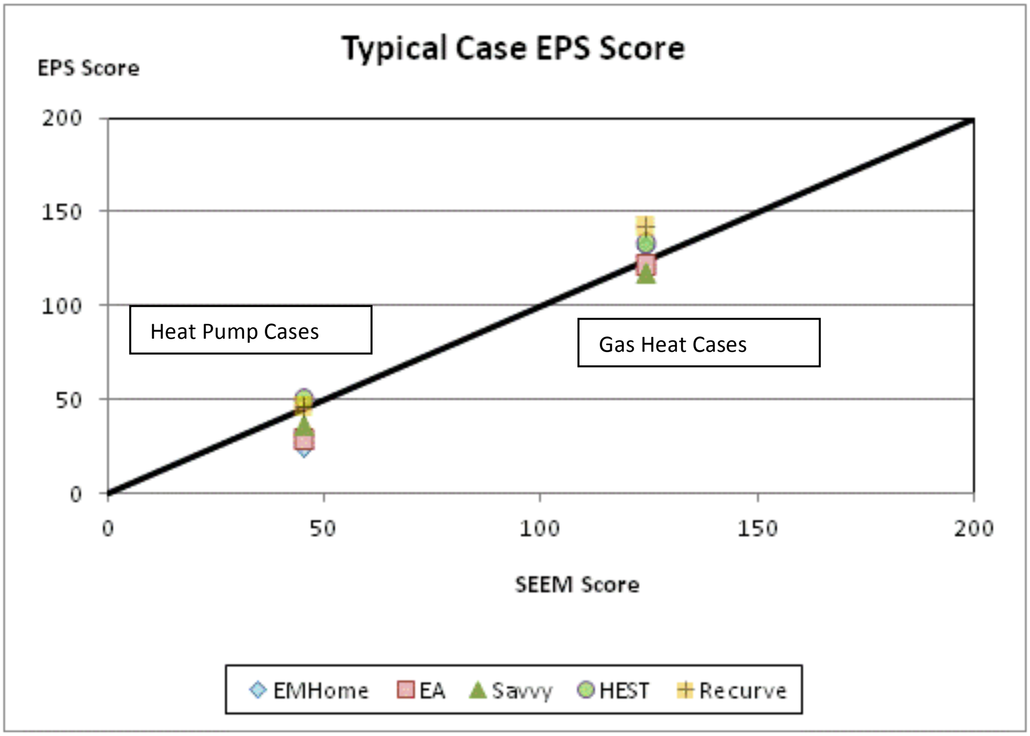


Figure 5. Typical Case Compared

What accounts for the differences between the modeling tools? For that, we have to look in more detail at the results for discrete end-uses. The EPS score is computed as the sum of all the energy consumption. One observation is that the different tools might actually compute different consumption for the end-uses but differences are counter each other when summed.

Appliance Loads

For example, we observe that the tools apply different assumptions regarding the energy usage of household appliances. Figure 6 shows assumed electric consumption for lights and appliances.

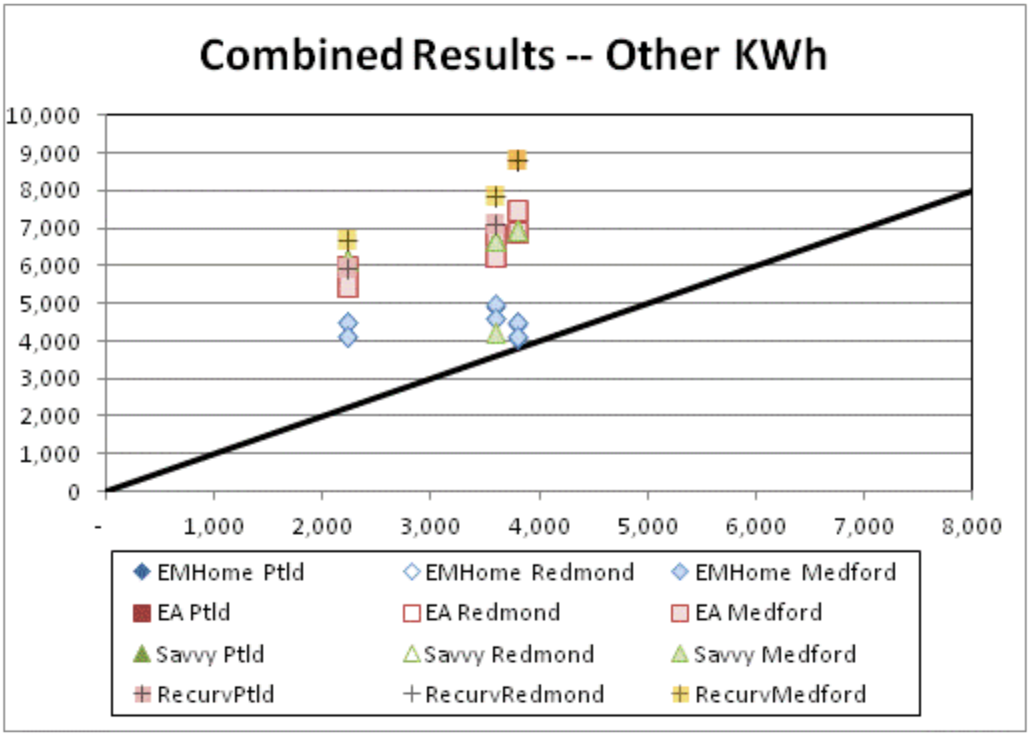


Figure 6. All Tools Compared (Other kWh)

SEEM has applied a standard amount of "other consumption" to all the homes. The other tools have varying assumptions for appliance loads but generally are much larger than SEEM assumptions. (To some extent, these loads affect space heating requirements since they are partly useful internal gains.) However, the impact is small – a difference of 4,000 kWh amounts to14 MBtu – say 10% of total score.

Table 3. Summary of SEEM Appliance Loads

Internal Gain Assumptions Standard Home (2200 sqft)			Applied to Prototypes, Annual kWh Loads		
			1344 sqft	2200 sqft	2688 sqft
Lighting	1608	Btu/hr	1,050	1,719	2,100
Clothes Washer/Dryer	315	Btu/hr	808	808	808
Refrigerator	223	Btu/hr	572	572	572
Dishwasher	140	Btu/hr	359	359	359
Misc. Cooking and Plug Loads	600	Btu/hr	941	1,540	1,882
Ventilation			117	117	117
Total			3,731	4,999	5,722

To verify the amount of assumed appliance loads required further investigation. The modeling assumptions were not provided for the proprietary models. For SEEM, appliance loads are based on regional studies and are tied to internal gains that influence the modeling results. As shown in Table 3, plug loads and lighting vary with prototype home size but the other appliances are fixed loads. SEEM is unique in adding a ventilation load consistent with the most recent Oregon building code. (Code now requires controlled ventilation with an air-to-air heat exchanger although this would not be likely in older homes.) SEEM modeling was repeated using the appliance loads specified in Table 3 which resulted in slightly lower heating consumption compared to preliminary modeling. Further definition of appliance loads, such as specification of cooking load, would be helpful but would not have a significant impact on the whole-house EPS score.

Equipment Performance Assumptions

As previously mentioned, assumptions regarding heating system performance play a small but important part in computing total energy consumption. Examination of the different model results allows one to generalize about the underlying assumptions embedded in the models.

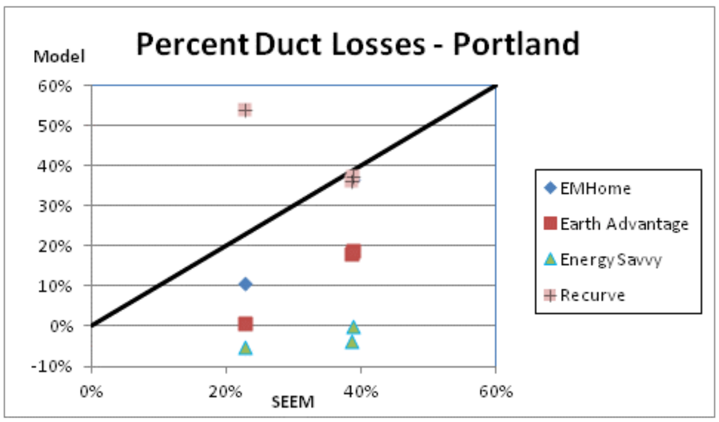


Figure 7. Duct Losses Compared -- Portland

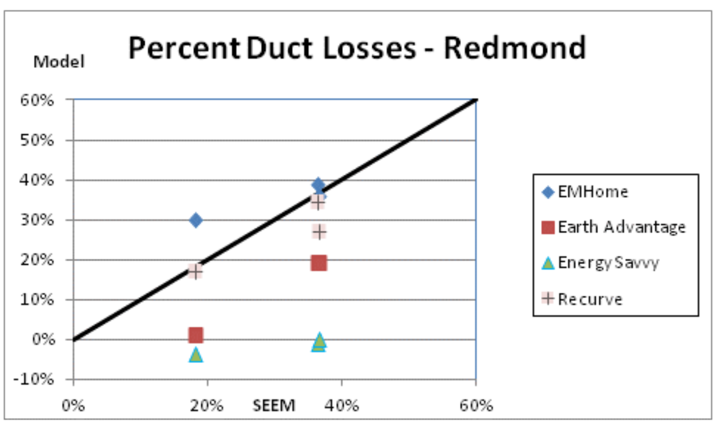


Figure 8. Duct Losses Compared – Redmond

Figure 7 compares duct losses by comparing the electric baseboard model to the electric furnace model. The difference in space heating consumption is considered to be a loss in heating efficiency. SEEM includes an explicit calculation of ducts and buffer spaces so it is presumed to be a more precise estimate. The estimate depends on choices made by the user is setting up the model – which accounts for some of the variation between prototypes. All of the tools allow for inputs representing duct variables. However, Figure 7 and Figure 8 demonstrate very different net results. Energy Savvy appears to ignore duct losses entirely and apparently includes an efficiency loss for electric resistance heating. Earth Advantage (EA) shows no change due to climate and Recurve even shows a counter-intuitive decrease in losses. It is not known if these results are due to an underlying difference in the calculation methods or due to "operator error" in how the modelers used the tools.

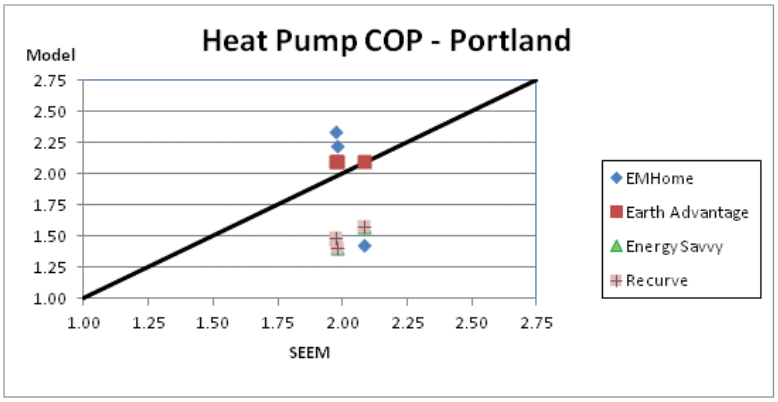


Figure 9. Heat Pump COP Compared -- Portland

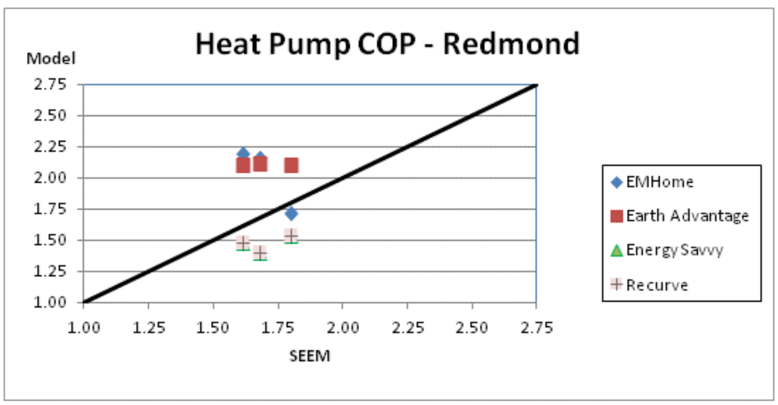


Figure 10. Heat Pump COP Compared -- Redmond

Figure 9 and Figure 10 show a similar comparison for heat pump seasonal Coefficient of Performance (COP) in two climates. In these charts, COP is computed as the ratio of space heating consumption for the heat pump model compared to the electric furnace model. SEEM includes an assumption of effective demand-limiting controls as required in Oregon. Such controls increase seasonal COP, but the improvement depends both on climate and the building structure's short-term hourly demand. SEEM then applies an hourly modeling method to rigorously compute climate and demand impacts.

Earth Advantage and Energy Savvy do not demonstrate an effect of climate on performance. (Earth Advantage has since modified their tool to fix this problem.) Energy Savvy tends to underestimate COP, perhaps because it does not utilize demand-limiting controls. As would be expected, Energy Measure Home (EMH) and Recurve compute some performance decrease in the more extreme Redmond climate, but only a small one. Tools that lack the capacity to correctly model equipment performance would be of limited use in program operations. In the following charts, we examine the how the specific tools treat sub-components of consumption.

Simulation Tools Compared

EnergyMeasure Home (EMHome)

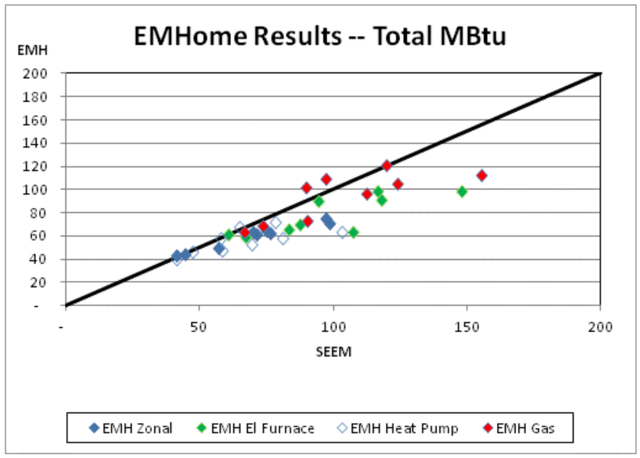


Figure 11. EMHome MBtu by System Type

Figure 11 shows results for EnergyMeasure Home (EMHome) by system type. The blue diamonds (Zonal) show a result close to just the thermal loads of the structure, without consideration of equipment performance. The green diamonds (El Furnace) show what happens when duct losses are added. The red diamonds (Gas) show a further addition of burner inefficiencies. The white diamonds (Heat Pump) show how heat pump performance is estimated – it will depend to some extent on climate. Of course, only the gas and heat pump systems are representative of consumer choices but the inclusion of other system types allows one to assess how the different system components influence the modeling tool. One observes that this tool tends to have a bias – that is, it under-predicts more for homes with large consumption.

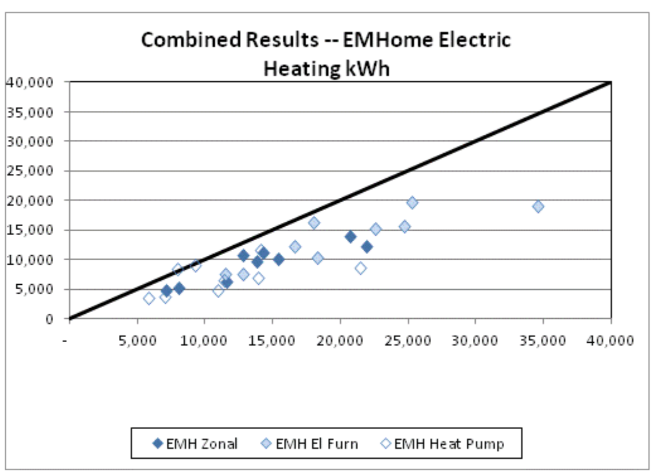


Figure 12. EMHome KWh by System Type

Earth Advantage (EA) (Simple)

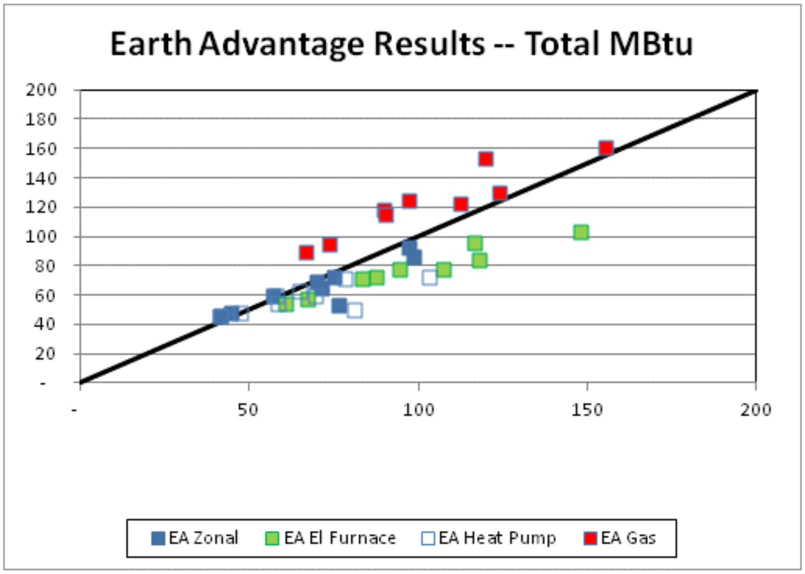


Figure 13. EA MBtu by System Type

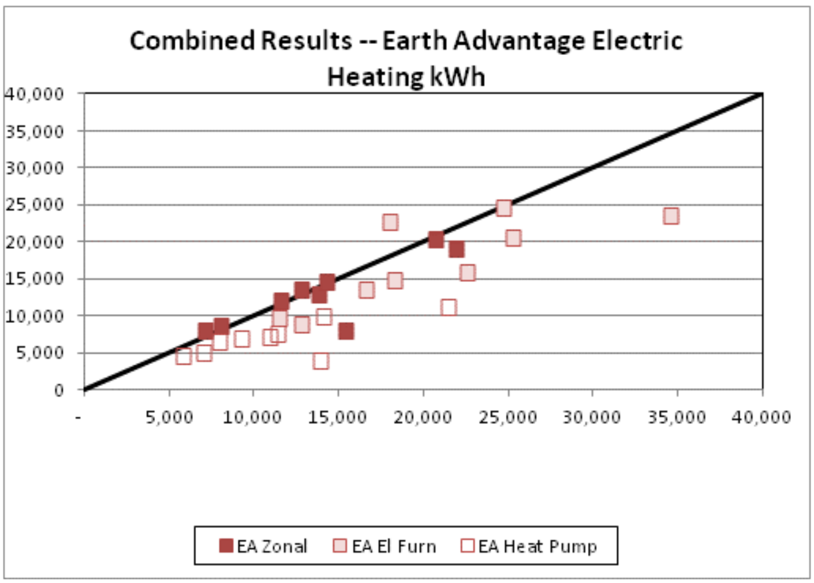


Figure 14. EA KWh by System Type

Figure 13 demonstrates relatively good and unbiased agreement with SEEM. The total MBtu chart hides the fact that heating performance is less accurate -- duct losses tend to be under-estimated.

Energy Savvy

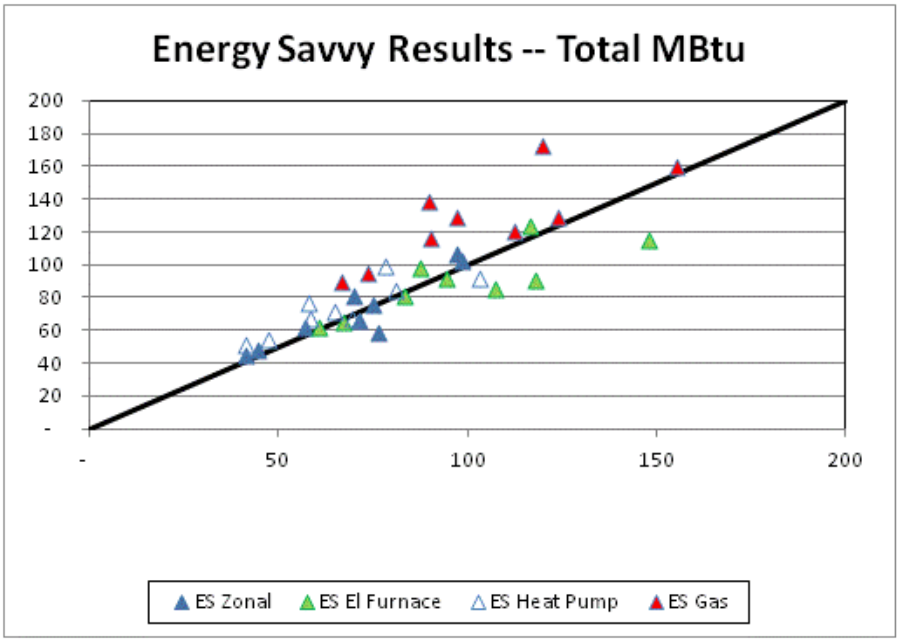


Figure 15. Energy Savvy MBtu by System Type

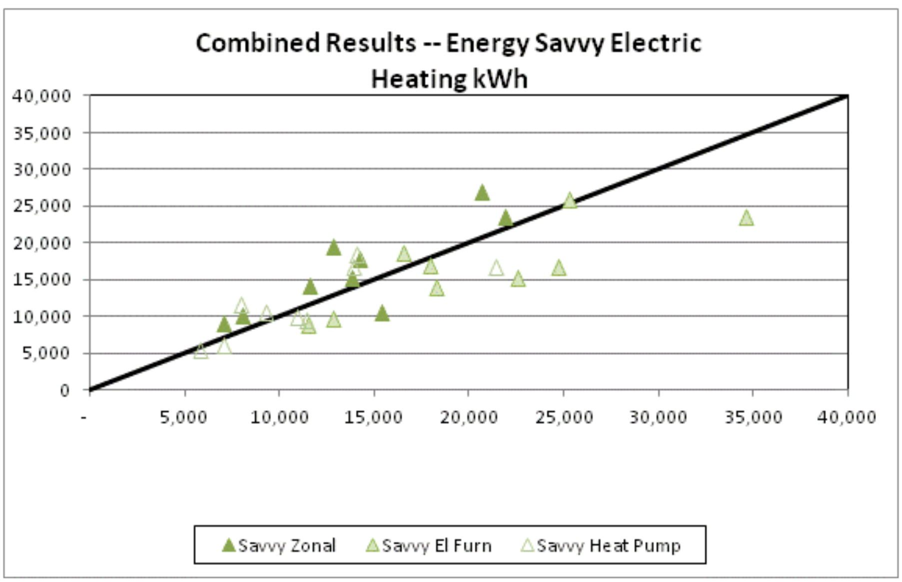


Figure 16. Energy Savvy KWh by System Type

Figure 15 and Figure 16 show relatively more scatter compared with SEEM.

Recurve

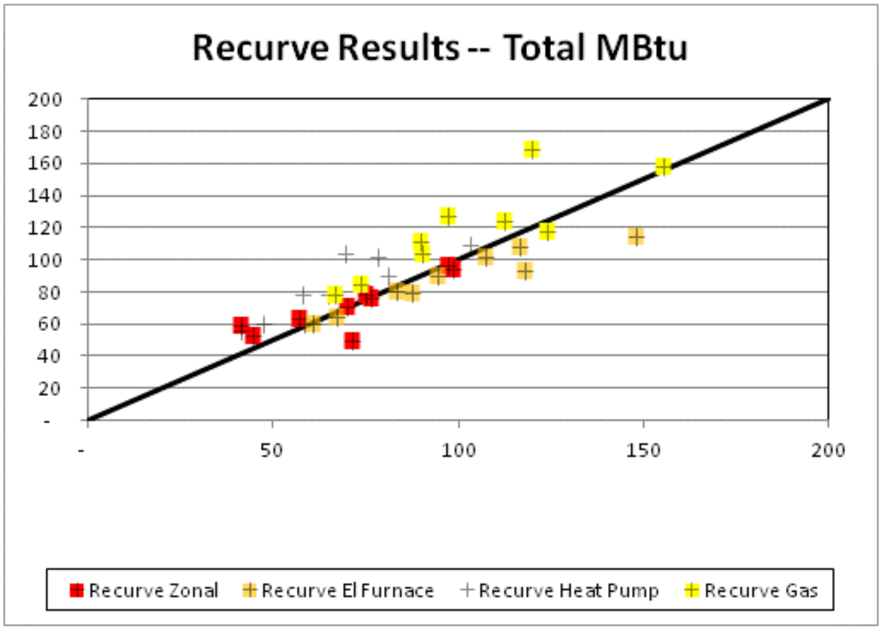


Figure 17. Recurve MBtu by System Type

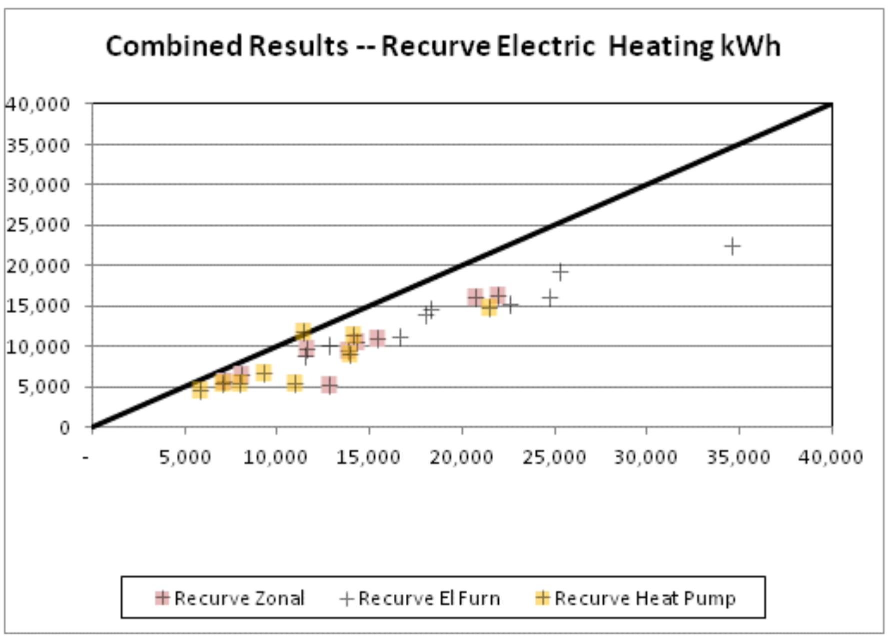


Figure 18. Recurve kWh by System Type

Figure 18 shows some bias on estimating electric heating. The overall MBtu estimates in Figure 17 hide variation due to major differences in other end-uses. This is discussed in more detail in the appendix.

Home Energy Scoring Tool (HEST)

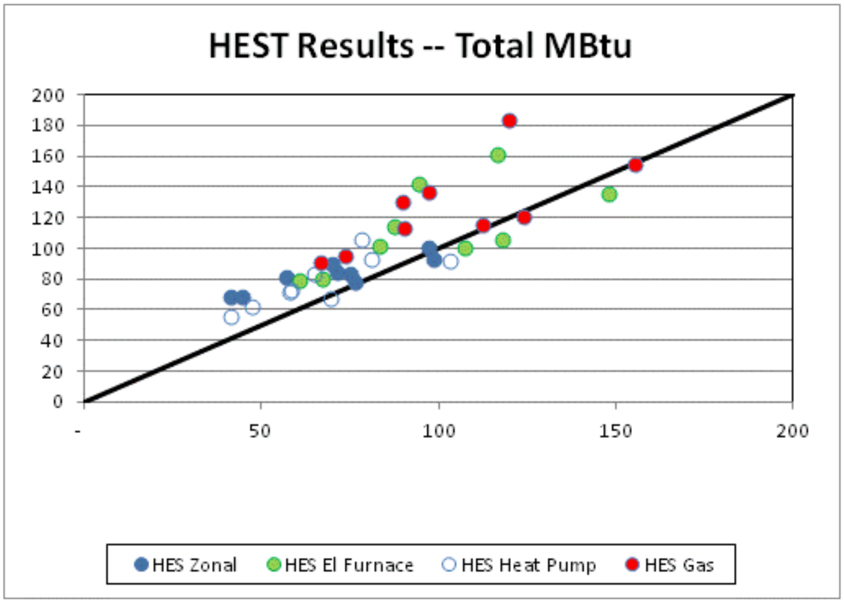


Figure 19. HEST MBtu by System Type

HEST tends to over-predict consumption as shown in Figure 19. Consumption for the basic structure tends to be high.

EPS Scores for Test Group

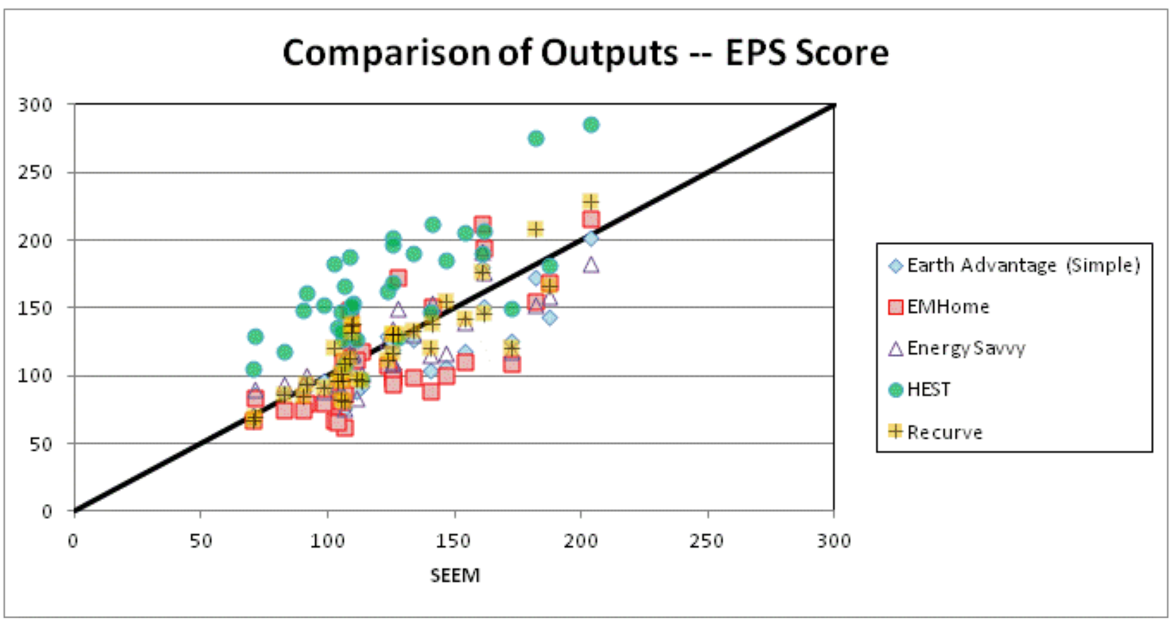


Figure 20. Test Group EPS Score

The EPS scores shown in Figure 25 are difficult to interpret due to the number of data points. Details are discussed in the appendix. Qualitatively, one observes that HEST distinctly over-predicts. The other tools are similar to each other. A more quantitative analysis follows.

Ranking of Modeling Tools

Table 4. Tool Features and Omissions

Feature	EMH	Earth Advantage	Energy Savvy	Recurve	Importance
Space Heating	Poor	Yes	Yes	Yes	High
Gas System Performance	No information	Poor	Yes	Yes	High
Adjust HP for Controls	Yes	No	No	No	High
Adjust HP for Climate	Yes	No	Poor	No	Moderate
Appliance Loads	Marginal	Poor	Poor	Poor	Low
Furnace Ducts	Poor	Yes	Poor	Yes	Low
Cooling	Poor	Acceptable	Acceptable	Poor	Very Low

Table 4 presents some conclusions regarding features of the various programs based on results for the prototypes. HEST is not included because it does not provide a breakdown by enduse. Being able to model space heating and equipment performance are high importance. The ability to model heat pump performance may become more important as the program expands to other fuel types. If the models do not already include such components, it will be difficult to add that capability.

Models based on historical regression coefficients may not be flexible when dealing with new circumstances or operating conditions. For example, EMHome is one program that allows for changes to input parameters. Earth Advantage is an example of a program that relies on previously developed regression coefficients.

Ability to model appliance loads and cooling is less important. Ability to model duct losses affects scores applied to conservation measures. Likely the models can incorporate changes for these end-uses.

To quantify model performance, we proposed two metrics. The first metric is based on accuracy compared to SEEM results. The Root Mean Standard Error (RMSE) of Gas Homes is computed as a percent of the total and provides a measure of accuracy[30]. The metric is computed for both the prototypes and the test group. For each group, one can assign a relative rank with "1" being the best.

A more important question is the consistency of results. That is, if one compares one house to another, do the different tools provide the same conclusions regarding relative scores? Or does the choice of modeling tool result in "flip flops" between the relative rankings of houses? Table 5 demonstrates an example from the test group of homes. This table shows the relative ranking with each tool for two homes. Under each tool, the homes are ranked in order of increasing energy use, with number "1" having the lowest energy use. A lower ranking number means lower energy use -- so that home would be preferred. When the homes are tightly clustered, small differences in the modeling results lead to different comparative rankings. In this example, one observes that Home #2 is preferred under SEEM and Earth Advantage but not under EMHome or Energy Savvy.

[30] Standard error is a statistical measure of "goodness of fit" and directly relates to the confidence limit around the average result.

Table 5. Example of Relative Ranking by Tools

Home	SEEM	EMHome	Earth Advantage	Energy Savvy
Home #2	7	4	3	8
Home #10	8	2	4	7

For a consistency metric, we computed the relative ranking of each case based on SEEM runs and compared that to the relative ranking using the alternative tool. For each tool, we then computed the average deviation in relative rankings. This average deviation quantifies the consistency of each tool. For the prototypes, the deviations were similar but the test group showed more differences between the tools.

Note that scores were computed for only three tools. Recurve and HEST were eliminated from the ranking process based on concerns noted in Table 4. While Recurve has some good features, it falters on equipment performance and cooling. HEST is the least satisfactory modeling tool.

In Table 6 and Table 7, we report quantitative metrics for the test cases. Note that the ordinal rankings for the two data sets yield different results. Further detail of metrics for both Gas and Heat Pump homes and the Test Group cases are shown in the appendix.

Table 6. Quantitative Statistics Comparing Prototype Models

EPS Score	Consistency	Consistency Ranking	Relative RMSE	RMSE Rank
Earth Advantage	1.6	1	22%	2
Energy Savvy	2.1	3	29%	3
EMHome	1.8	2	18%	1

Table 7. Quantitative Statistics for Test Case Models

EPS Score	Consistency	Consistency Ranking	Relative RMSE	RMSE Rank
Earth Advantage	5.0	2	16%	2
Energy Savvy	4.9	1	16%	1
EMHome	5.9	3	24%	3

Using the quantitative metrics of Table 6 and Table 7, one can compute an ordinal ranking of the various tools as shown in Table 8. The overall rank is computed using the specified weights. Consistency and accuracy for both study groups are weighted equally. The option of applying different weights was explored but did not greatly change the overall ranking. As is apparent, a high rank in one metric does not necessarily mean a high rank in others.

Table 8. Ordinal Ranking of Tools

Ordinal Ranking	Prototype Zonal Space Heat Consistency	Prototype EPS Accuracy	Test Case Space Heat Consistency	Test Case EPS Accuracy	Overall Rank
Weights	25%	25%	25%	25%	
Energy Savvy	3	3	1	1	3
Earth Advantage	1	2	2	2	1
EMHome	2	1	3	3	2

Chapter 4 Recommendations

Overall, the desired goal of a scoring method is achievable. Two tools were removed from consideration. Home Energy Rating (HEST) appears to seriously over-estimate total consumption. Recurve has discrepancies with regard to heating and cooling equipment performance.

Consistency is a major issue. If Home A is compared to Home B, the relative choice should not be affected by the choice of scoring tool. All of the tools demonstrated some “flip flop” of relative ranking position compared to their SEEM ranking.

All of the tools suffer problems when one looks at end-uses. In particular, it would be useful to develop a standard assumption for appliance energy use so this end-use is not a source of discrepancy.

The reference benchmark will need further development. The Energy Trust reference was developed for gas-heated homes in the Willamette Valley. To extend use of the EPS scoring, reference benchmarks will be needed for other climates and for new construction.

The use of a gas-heated reference for heat pump cases may also be an issue. A fuel-blind rating would change the reference baseline depending on the fuel type.

Chapter 5 Technical Appendix

SEEM Modeling Standard

The SEEM program is designed to model residential building energy use. The program consists of an hourly thermal simulation and an hourly moisture (humidity) simulation that interacts with duct specifications, equipment, and weather parameters to calculate the annual heating and cooling energy requirements of the building. There are current efforts to add infiltration and ventilation to the simulation. It is based on algorithms consistent with current American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE), American Heating and Refrigeration Institute (AHRI), and International Organization for Standards (ISO) calculation standards.

SEEM, written at Ecotope, was developed by and for the Northwest Power and Conservation Council (NPCC) and the Northwest Energy Efficiency Alliance (NEEA). SEEM is used extensively in the Northwest to estimate conservation measure savings for regional energy utility policy planners. It is the simulation engine used to provide heating and cooling energy savings estimates for the residential sector in the Northwest Power Plan, for the Performance Tested Comfort System (PTCS) incentive program, the Northwest EnergyStar for Homes program, as well as numerous other utility program offerings. SEEM is also used to support state building energy code revisions including the Washington and Oregon state energy codes.

SEEM has been calibrated against several residential data sources. This step is required by the Regional Technical Forum (RTF) as part of the vetting of savings and performance estimates used for regional conservation planning. These data sources include sub metered data from the RSDP and RCDP programs as well as billing analysis from the NEEM program and from the PTCS heat pump program. This set of data represents most of the long term end use data currently available in the region.

SEEM offers a number of advantages over other simulation programs. The step-by-step hourly calculations accurately model both air temperature and mean radiant temperature using a state of the art algorithm. Next, heat pumps and air-conditioners are modeled on real performance data from manufactures' catalogues. SEEM also provides the capability to use multiple control strategies and thermostat setups for the equipment. Further, SEEM closely tracks duct losses to user specified zones (inside, outside, crawl, attic) and accurately models their impacts. Additionally, SEEM contains a comprehensive below-grade heat loss algorithm to model building ground contact through slabs, crawl spaces, and basements. Lastly, weather data for the simulation comes from the widely used Typical Meteorological Year (TMY) datasets.

Similarities with All Tools

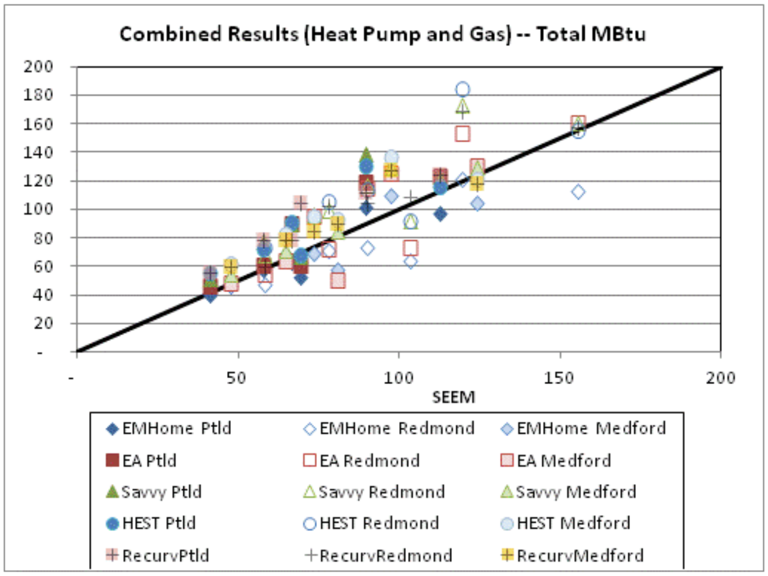


Figure 21. All Tools Compared (Total MBtu)

Figure 21 shows heat pump and gas cases for all the modeling tools – which makes it difficult to decipher. It is apparent that all the tools "work" (i.e. agree with SEEM) to some extent. However, there is a lot of scatter.

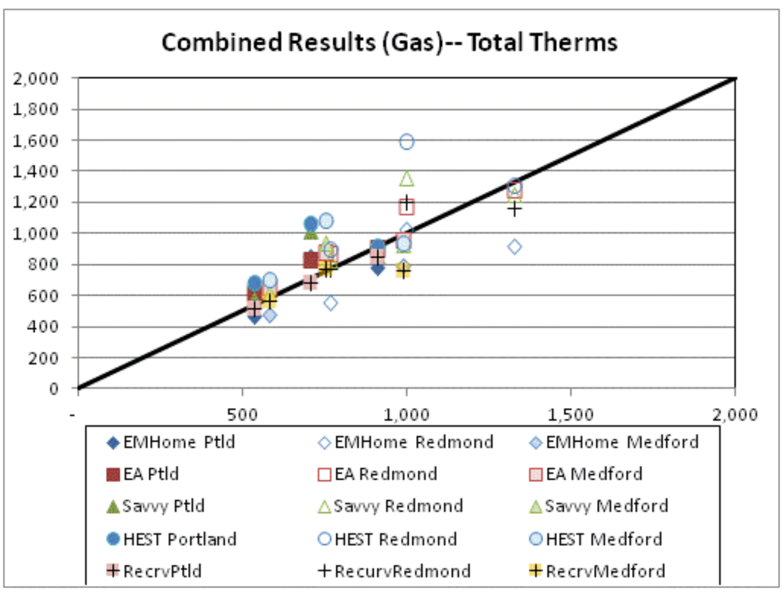


Figure 22. All Tools Compared (Therms)

It is instructive to review how these tools treat the end-use components within total consumption. Figure 22 shows results for only gas consumption. It is more apparent here that the tools often over-estimate relative to SEEM.

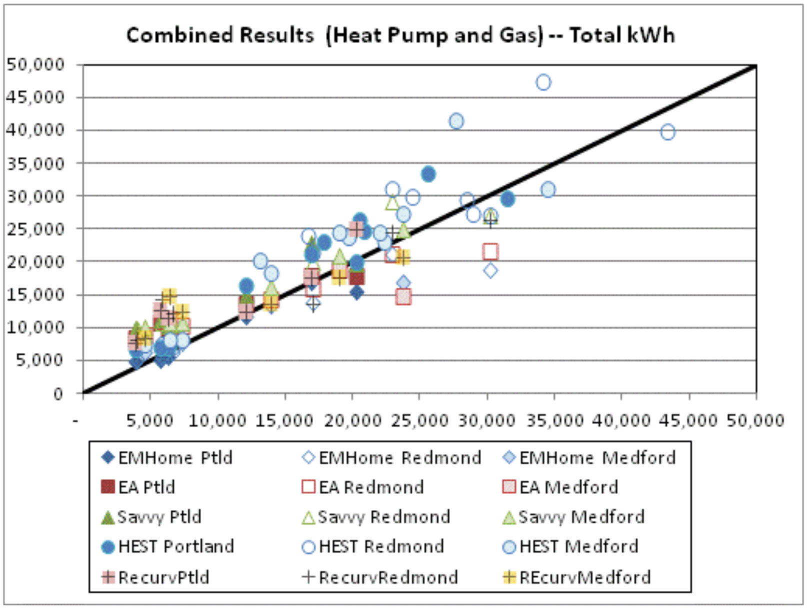


Figure 23. All Tools Compared (kWh)

Figure 23 shows estimated electric consumption.

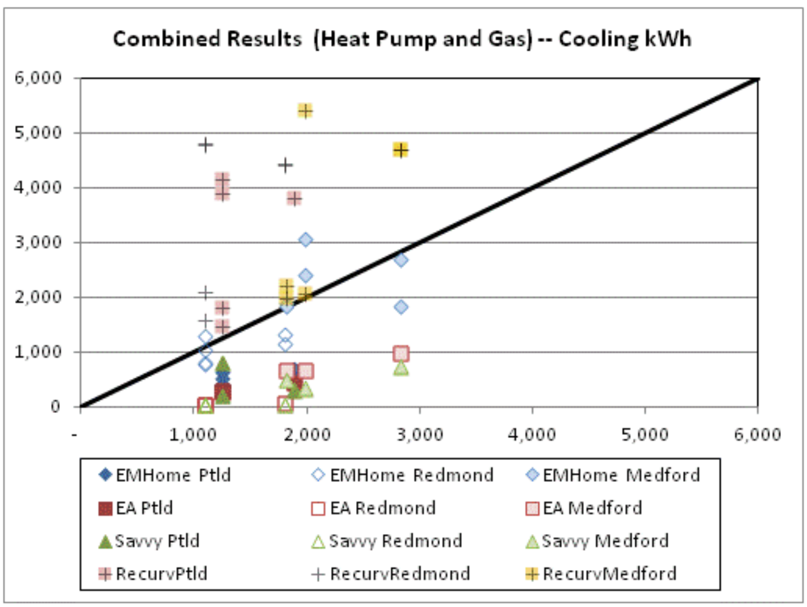


Figure 24. All Tools Compared (Cooling kWh)

Figure 24 shows the cooling end-use for the various tools. National modeling tools often over-estimate cooling loads in Pacific Northwest climates. This appears to be the case for Recurve. In contrast, Earth Advantage (EA) tends to under-estimate cooling relative to SEEM. (Enduse breakdown was not available for HEST.)

Test Group Implementation

Figure 25 compares the results of alternative modeling tools on the test group of 35 gas-heated homes. There is a fairly large amount of scatter, indicating that the tools differ in their modeling approach. Once again, SEEM is treated as the reference on the x-axis. For these cases, the SEEM models attempted to include both appliance consumption based on family size. HEST estimates predict significantly more consumption than the other estimates. Figure 26 shows that EMH estimates lower therms than the other tools. (HEST end-use breakdowns were not available.)

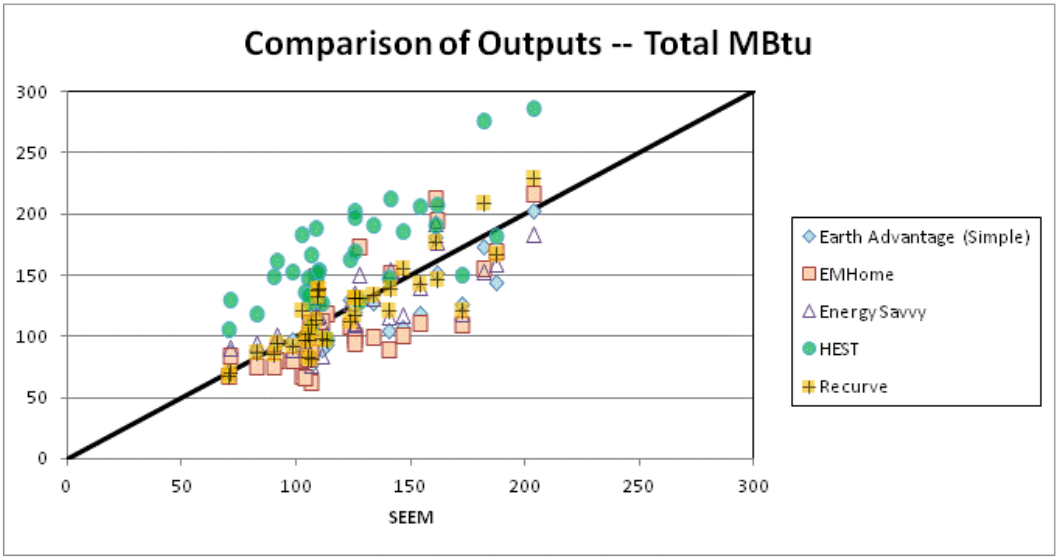


Figure 25. Test Group Total MBtu

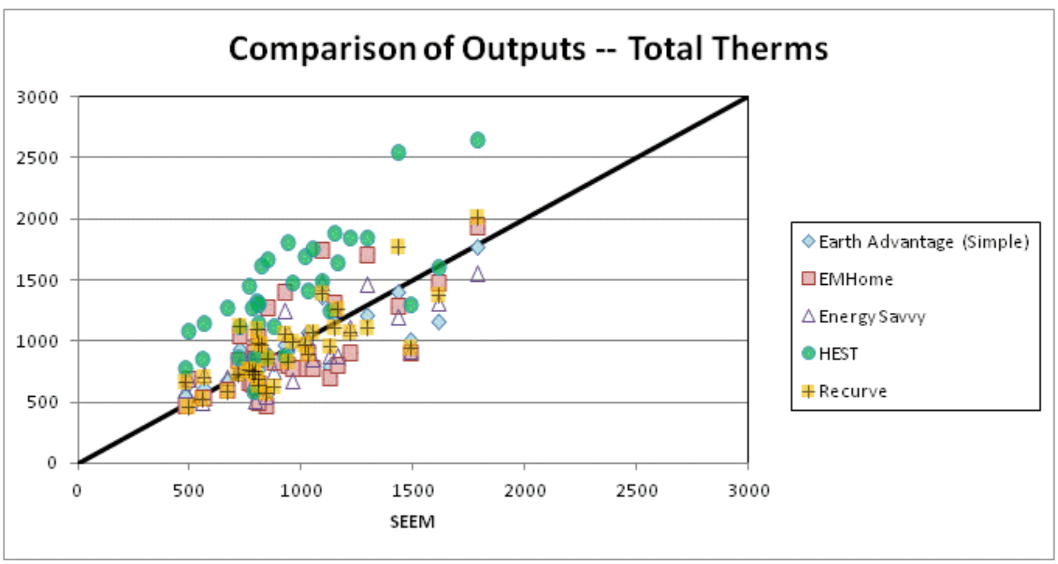


Figure 26. Test Group Therm Estimates

Review of the end-uses shows more differences between the tools. Figure 27 shows wide scatter in predictions of space heating. Figure 28 shows electricity for space heating. The amount of electricity is small for gas-heated homes, representing energy used by the furnace fan. Earth Advantage and Energy Savvy totally ignored fan energy. Recurve includes fan energy but tends to under-estimate.

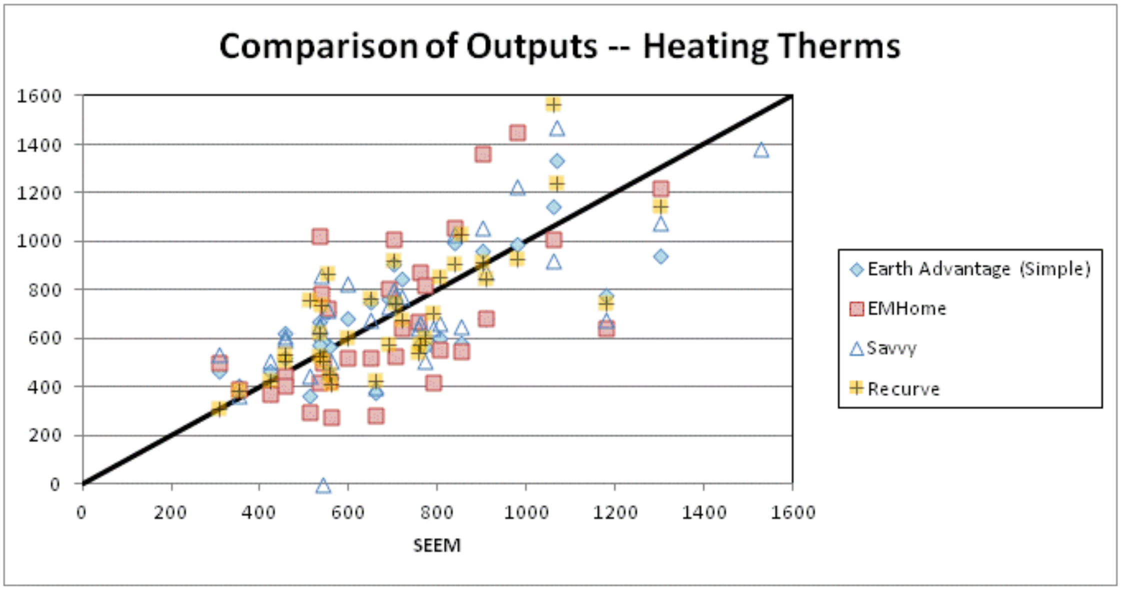


Figure 27. Test Group Heating Therms

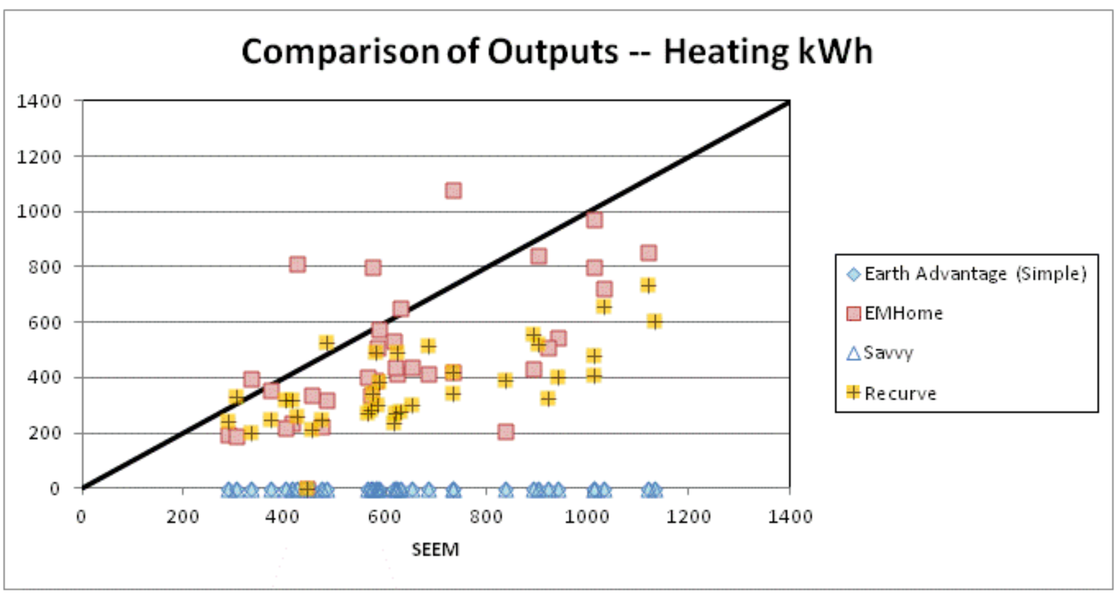


Figure 28. Test Group Heating KWh

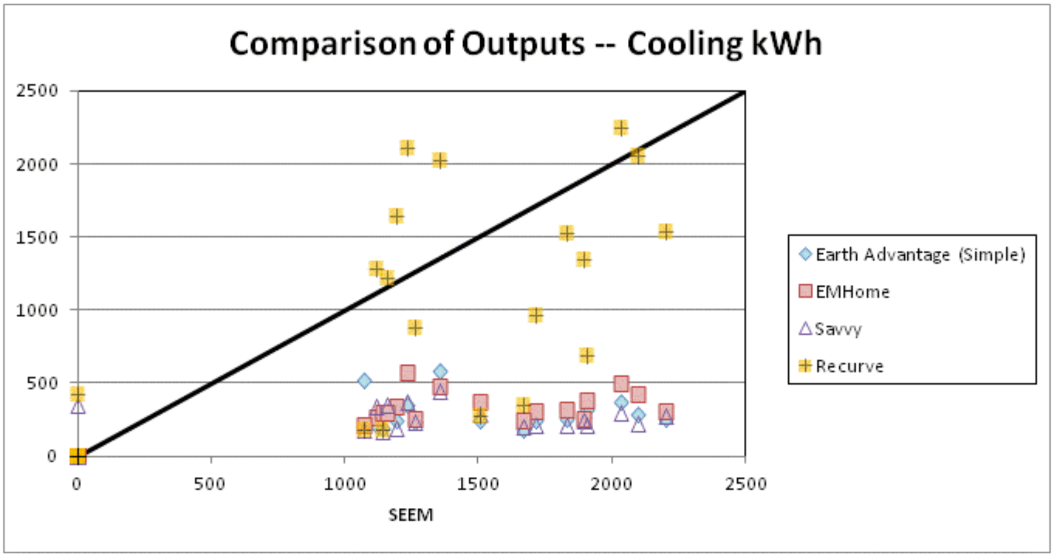


Figure 29. Test Group Cooling KWh

Similar results are shown for cooling consumption in Figure 29. Most of the tools under-estimate cooling consumption -- although this is a small end-use in our climate. Recurve shows wide discrepancy in results compared to the other tools.

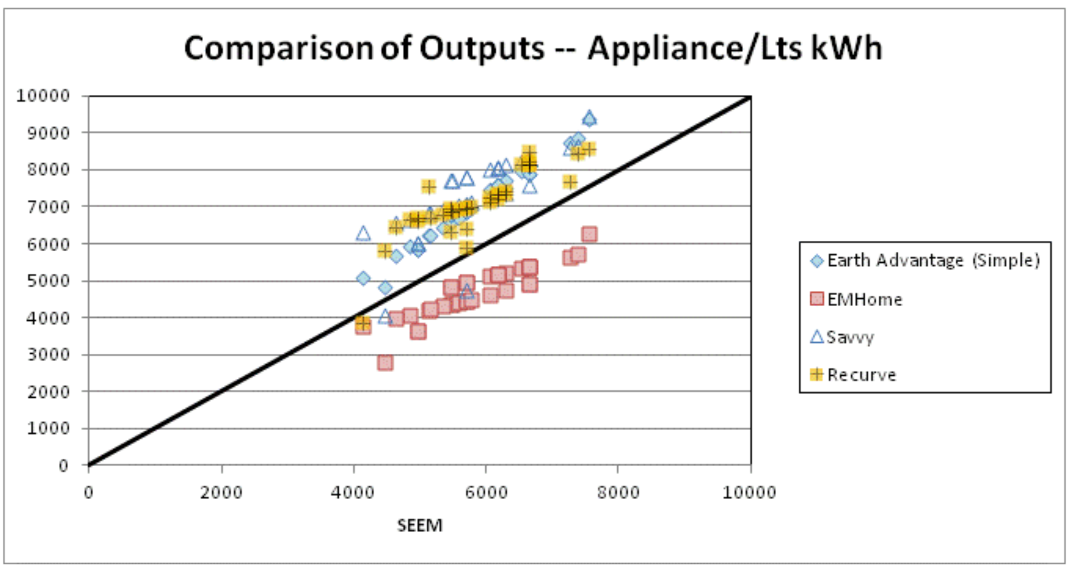


Figure 30. Test Group Appliance/Lights KWh

Differences in the modeling assumptions appear in appliance and lighting loads. Most of the tools assume higher loads than SEEM. EMHome assumes less kWh consumption.

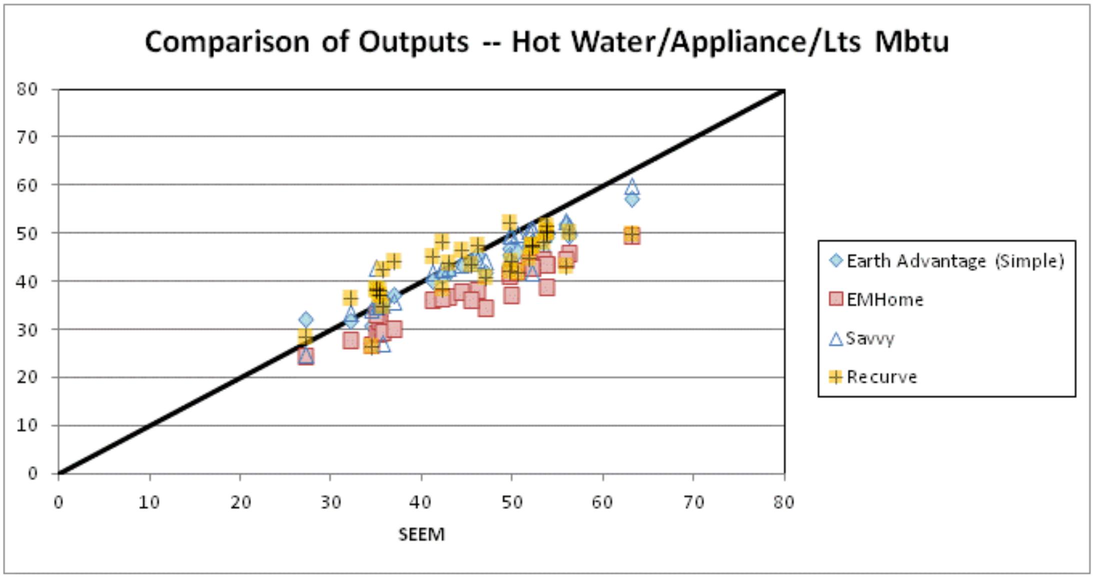


Figure 31. Test Group Appliance MBtu

Note that appliances are a relatively small part of total consumption. As an additional note, Energy Savvy is the only tool that assumed no gas used for cooking or appliances. This is apparently an error by the modeler. As shown in Figure 31, differences in appliance assumptions (including hot water) amount to perhaps 10 MBtu. Thus, it provides little explanation for the scatter observed in total MBtu (Figure 25).

Actual Energy Consumption

One might wonder how well any of these tools compare to the real world. The goal of this project was to investigate an "asset-based" energy score that is not influenced by customer behavior. However, we have actual consumption for the test homes. Figure 34 compares estimates by all the tools to actual consumption.

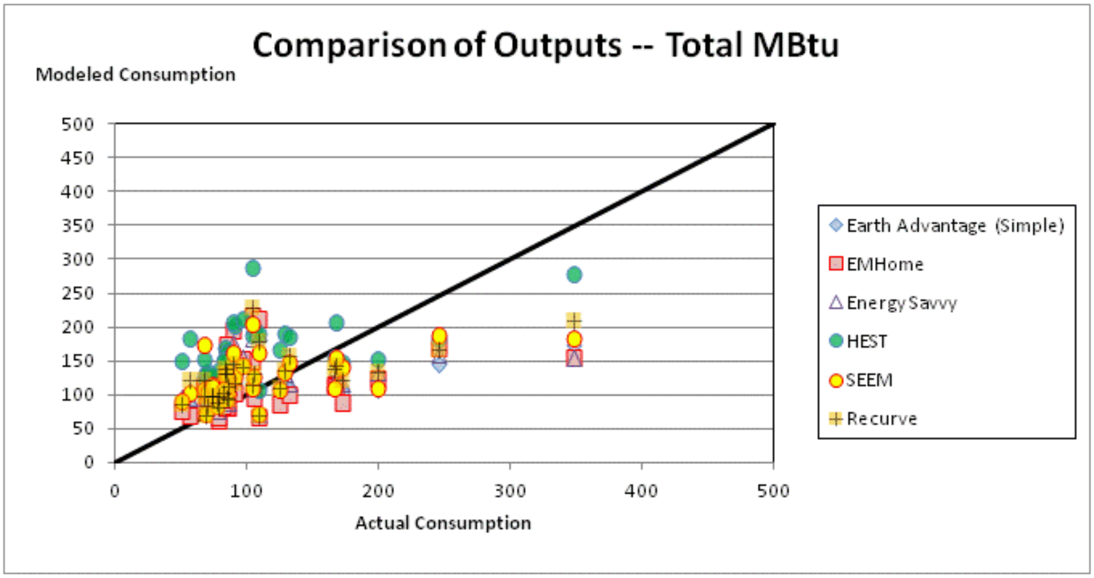


Figure 32. Modeled Compared to Actual Consumption

All the tools tend to under-estimate for large homes. Apparently, the standard assumptions regarding human occupancy do not scale up in a linear fashion. That is, large homes still have a similar family size so do not have as many persons per square foot. For that reason, the ranking methodology does not place a high premium on agreement with actual consumption.

As shown in Figure 33, consumption is not linear with respect to house size. The larger homes tend to be more expensive custom homes and the occupants have a higher lifestyle. The trendline shows a non-linear response to house size. For comparison, Figure also shows the reference benchmark (red line) suggested by the ETO study.

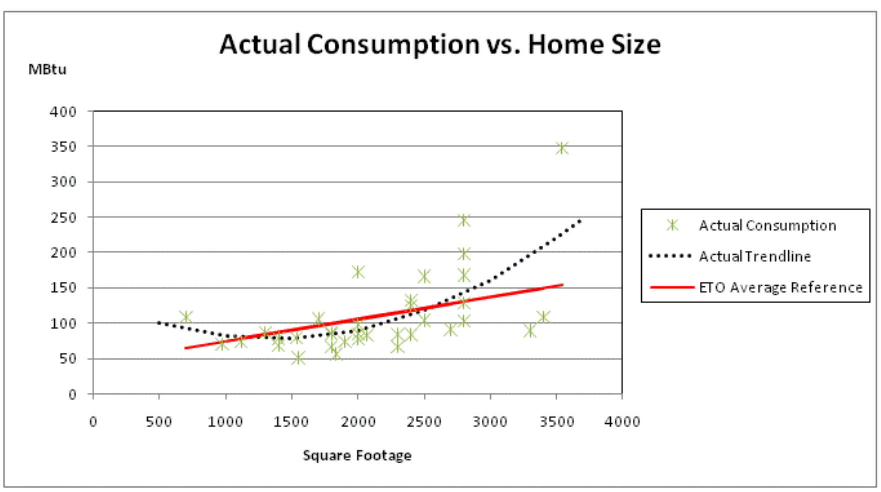


Figure 33. Actual Consumption versus Home Size

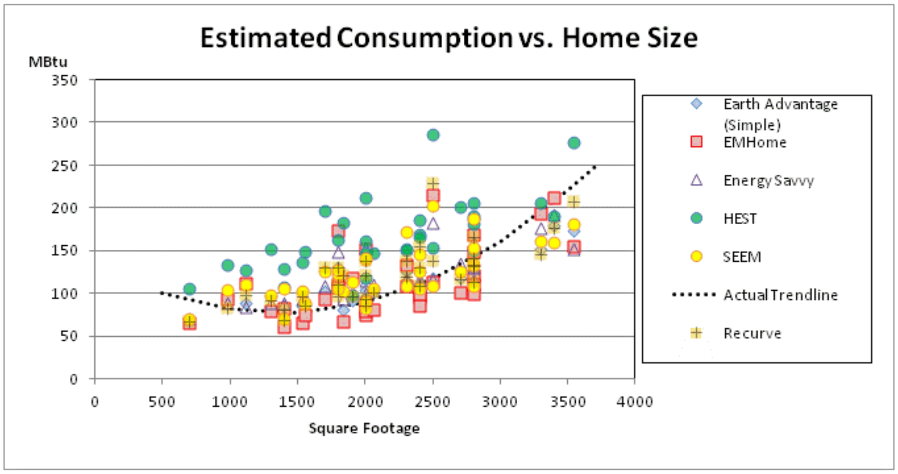


Figure 34. Estimated Consumption versus Home Size

Figure 34 shows the estimated consumption as a function of home size. For comparison, the chart also includes the actual trend line from Figure 33. With the exception of HEST, all the models assume typical occupancy parameters and under-predict the larger homes.

Quantitative Metrics

In review of results, we determined that a major issue is consistency. That is, if House A is compared to House B, the relative choice should not be affected by the choice of modeling tool. We determined this by comparing the relative ranking of the cases by SEEM and then by alternative tools. Figure 35 shows the relative ranking for prototypes. If the ranking agreed with SEEM, the data points would be arranged along the 45-degree line. Instead, there is a fair amount of scatter, although all the tools are similar. Figure 36 shows the same for the test group. This group shows more scatter and a larger range of differences between the different tools.

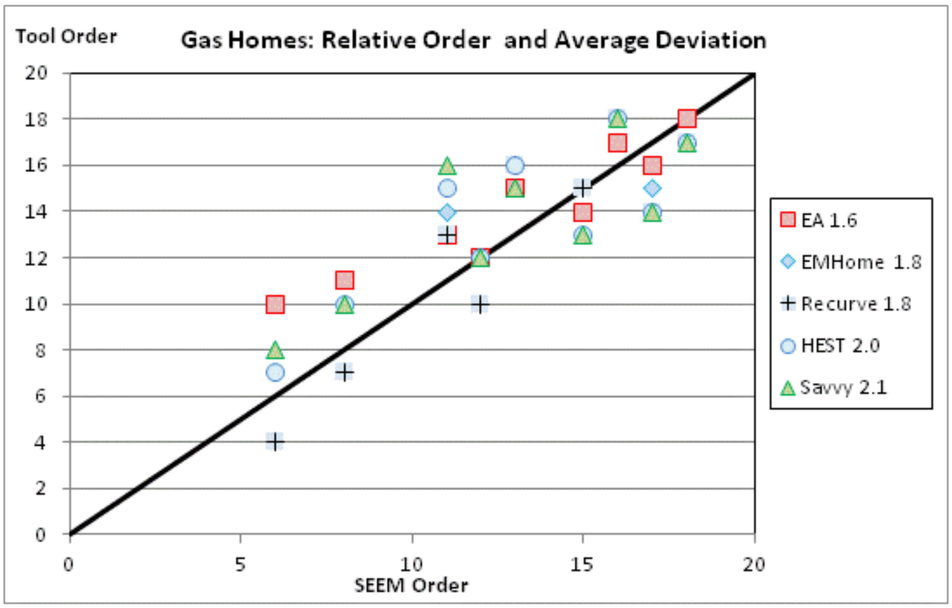


Figure 35. Deviation in Ranking for Prototypes

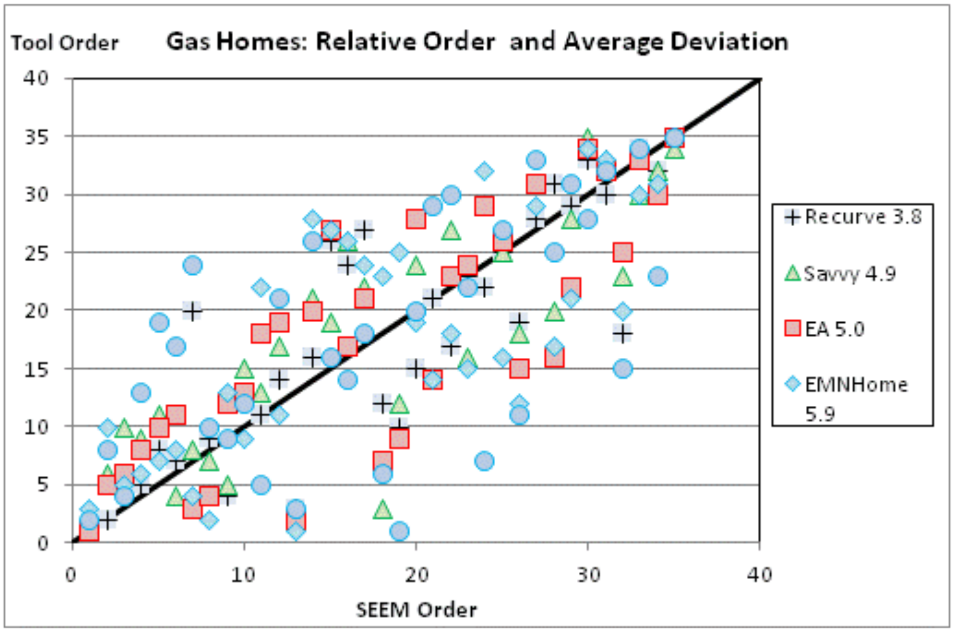


Figure 36. Deviation in Ranking for Test Group

As previously described, the average deviation in rankings represents an index regarding consistency with SEEM rankings. RMSE is an measure of overeall accuracy compared with SEEM. Table 9 and Table 10 report the computed indices and relative rank for both study groups.

Table 9. Quantitative Metrics -- Prototypes

Total MBtu	Consistency	Consistency Ranking	Relative RMSE	RMSE Rank
Earth Advantage	1.6	1	22%	2
Energy Savvy	2.1	3	29%	3
EMH	1.8	2	18%	1
Recurve	1.8		21%	
HEST	2.0		30%	

Table 10. Quantitative Metrics – Test Group Cases

EPS Score	Consistency	Consistency Ranking	Relative RMSE	RMSE Rank

Earth Advantage	5.0	2	16%	2
Energy Savvy	4.9	1	16%	1
EMHome	5.9	3	24%	3
Recurve	3.8		14%	
HEST	6.5		41%	

Table 11. Overall Rankings

Ordinal Ranking	Prototype Consistency	Prototye EPS Accuracy	Test Case Consistency	Test Case EPS Accuracy	Accurate Cooling	Include HP Performance	Overall Rank
Weights	25%	25%	25%	25%			
Energy Savvy	3	3	1	1	3	3	3
Earth Advantage	1	2	2	2	2	2	1
EMHome	2	1	3	3	1	1	2

As shown, we assigned ranking for a number of other categories that were not included in the overall ranking. They are shown here for general information. At some point, more focus could be applied to issue of equipment performance and the capacity to model heat pumps. In the end, the rank was based on accuracy and consistency with equal rank applied to the prototype and test group cases.

APPENDIX C

Existing Homes Energy Performance Score
2011 Pilot Initiative Summary

Energy Trust

Energy performance scoring (EPS) has been used to good effect in supporting energy efficiency efforts in new homes. Over the last several years there has been a strong desire to have an equivalent for existing homes. Because of the variability of existing homes, questions about the efficacy of the underlying modeling and the numerous interested parties, extending the EPS has complexity.

Energy Trust has advocated a systematic, structured approach to extending an EPS to existing homes. We have proposed incremental tests. The pilot approach proposed for 2011 is an extension and expansion of the work started in 2008 that resulted in the 2009 report.

Residential energy efficiency programs at Energy Trust are in the midst of revamping the approach to customers to provide better and more motivating information and services. The EPS has the potential to provide a tool that could be a compelling companion with these other changes or be more universal and more compelling in some circumstances. It may even be an appropriate tool for defining the value of energy efficiency in homes at re-sale.

Background

In 2008-2009, through Conservation Services Group and Earth Advantage Institute, Energy Trust undertook a test of the various modeling options behind an EPS. The study found:

1. Complicated models were no better at forecasting energy savings than less complex models.
2. The best performing non-complex model had an apparent error band of plus or minus 30%.
3. Comparisons to billing data are not the accurate test of a model's forecasting efficacy, due to homeowner behavior.
4. A set of enhancements to the less-complex energy models might improve accuracy.
5. More tests of improved models in comparison to a standardized baseline (non-bill) should be conducted.

In 2009 the Oregon Legislature created a taskforce to examine an EPS in Oregon. The taskforce noted the series of issues related to launching a state-wide EPS and recommended a voluntary, testing approach.

In 2010 in the cities of Seattle and Bellingham Washington, Earth Advantage began testing their approach to an EPS on existing homes, which includes multiple elements. Along with the score, it has an alternative delivery and contracting platform to make for a non-complex delivery from the consumers perspective. The enhanced approach utilizes an updated version of SIMPLE, the first version of which was tested in the study sponsored by Energy Trust, as the underlying model for rating the home.

Conservation Services Group has also been developing and testing an enhanced version of its current modeling tool RealHomeAnalyzer.2 for forecasting savings for use in Energy Trust's established, existing homes program.

The US Department of Energy (DOE) created its own Home Energy Score (HES) that utilizes an asset based version of Home Energy Saver with benchmarks based on Residential Energy Consumption Survey(RECS) data. The DOE announced ten tests in various states and cities in 2011, including one through the Energy Trust.

Energy Trust Pilot

The EPS is viewed by Energy Trust as a possible tool for strategic engagement of customers and markets, but only if it can be a reliable, accepted and effective tool. It could be used as a resource within the existing home program to give visibility and awareness to consumers on the performance of their home's shell and mechanical systems under standard operating conditions and how:

1. this compares to benchmarks with respect to energy consumption and carbon impacts, and
2. home performance would improve with investments in recommended energy efficiency actions.

The EPS has strong ties to the effort with real estate professionals to educate and train them to champion energy efficient homes. The EPS is a way for these market actors to educate customers on the relative merit of an energy efficient home and differentiate themselves in the market. EPS values may also play a role in mortgages, insurance products and financial initiatives.

Finally, if successful, Energy Trust contemplates an EPS as being deployed as part of the services that a BPI Certified Home Performance contractor or certified RESNET provider, could use as the basis to more simply estimate home-specific savings from recommended actions.

<u>Approach</u>
Given the conclusions of the first study, the development of new, less-complex models and the need to have further analytic review of alternative approaches, Energy Trust is interested in a next stage of testing an EPS for the general, existing-homes market. We propose a further set of testing that will compare the look and feel of the HES to one we utilize for new homes in Oregon and the impact these have on follow-through rates by home owners.

Additionally, we will examine three underlying analytic models: DOE's HES, an enhanced SIMPLE and a revised RealHomeAnalyzer (RHA2) from Conservation Services Group. We will not be testing the alternative contractor and delivery mechanisms being piloted by Earth Advantage in Washington as part of their EPS.

Goals

The goals of the proposed pilot on the next stage of an EPS are:

1. Does an EPS motivate a customer to act, in terms of sooner, deeper or both?
    a. Is one format of an EPS more motivating than another:
        i. DOE vs. ETO format, in terms of:
            1. Follow through rates for:
                1. Types of measures
                2. Quantity of measures
    b. Which information is most useful
        i. Score or no score
        ii. Energy usage
        iii. Savings by measure/action
        iv. Carbon footprint
    c. What are the customers' visual preferences
2. Does an EPS have efficacy in:
    a. Directing customers to do the right sorts of things
    b. Assigning a comparative ranking without diagnostic testing
    c. Predicting energy saved without diagnostic testing
    d. Which of three models perform better and under in terms of the above: (HES, SIMPLE and RHA2)
        i. What is ease of use in the field for the three models

The next stage of testing will involve 400 customers. The customers in the tests will be chosen randomly from those requesting home energy reviews. About 200 will get the US DOE version and another 200 will get the Energy Trust version. Both reports will have recommendations of energy savings actions but presentation and type of information will different between the two. Customers will be asked to complete surveys at one week and six months.

At present we expect the RHA2 model to generate the score in the Energy Trust version. The Energy Trust format of the EPS will have the same design, which is currently being used for Energy Performance Scores on new homes but will include an improved recommendations report from the one that is used for Home Energy Reviews.

We will compare actions taken at the one, three, six month and one year marks to see if the EPS drives noticeably higher levels of action compared to benchmark levels of action as. We will also see whether the presentation format (DOE or ETO) makes a difference. .

Behind the scenes each home will be modeled by the other two models (HES and SIMPLE) or (RHA2 and SIMPLE). The resulting energy usage and savings analysis generated using all three models will be compared to the regional reference tool in the Northwest (SEEM).

The numerical analysis will be conducted by an independent third party and help define if any of the models performs better and under which circumstance and with what error band. We expect to supplement data collected through the EPS pilot with additional home scores using program information in order to compile the appropriate dataset for the analysis. This part has already started.