DOCKETED

Docket Number:	17-BSTD-02
Project Title:	2019 Title 24, Part 6, Building Energy Efficiency Standards Rulemaking
TN #:	223106-2
Document Title:	Obstacles to Commercialization of Variable Capacity Heat Pump Systems in the California Market
Description:	N/A
Filer:	Patty Paul
Organization:	Mitsubishi Electric US/Bruce Severance
Submitter Role:	Public
Submission Date:	4/2/2018 1:40:29 PM
Docketed Date:	4/2/2018

OBSTACLES TO COMMERCIALIZATION
OF VARIABLE CAPACITY HEAT PUMP SYSTEMS
IN THE CALIFORNIA MARKET

By Bruce Severance
Regulatory Compliance
Mitsubishi Electric US

Introduction

Variable capacity heat pumps (VCHP) and variable refrigerant flow systems (VRF) hold the promise of drastically reducing HVAC system carbon footprint in nearly every residential and commercial application. However, debate about whether AHRI test protocols and efficiency ratings are accurate and the comparatively low CEC "compliance credit" for these systems is greatly inhibiting market penetration in California. It is a common experience among architects and designers that they may need to integrate VRF or VCHP systems to reach the LEED accreditation or all-electric ZNE goals they are targeting, but the compliance credit offered for these high-efficiency HVAC systems shows up as a compliance penalty when the CBECC models are run. The discrepancies in system efficiency ratings between standard test protocols (AHRI Standards 210/240 and 1230) and the CBECC modeling software are widely recognized by CEA energy analysts and they are frequently the central topic of project cost discussions and compliance strategies with architects. Energy consultants in the state widely perceive the misallocation of compliance credit to be a barrier to compliance approval of all-electric ZNE projects. This comparatively low compliance credit is at least in part due to the inherent challenges of testing VCHP technology that continuously vary capacity as a means of optimizing efficiency. Efficiency in this system is not a fixed quantity and there are variables, controls and climate conditions that cause efficiency and capacity to increase or decrease as the system ramps up or down in response to loads and climate conditions. Like the V-8 that can shut down four cylinders, this variable capacity feature is in fact what optimizes efficiency. However, traditional test methods which take "snapshots" of performance, become a less appropriate lens for evaluating efficiency that varies over time.

In addition to this, it is evident that source energy calculations integrated into the CEC's algorithms for TDV (time dependent valuation of energy) assume 70% grid losses regardless of whether or not on-site distributed generation partially or entirely eliminates grid losses. Critics from both inside and outside the CEC have hinted at institutional resistance to revising these source energy inaccuracies. In the meantime there is roughly a three to one favor for gas heating systems and appliances that enjoy several prescriptive paths under the proposed 2019 Code, while all-electric ZNE options are portrayed as inherently less efficient due to grid losses. There is no apparent consideration or prospect of TDV calculations becoming more equitable for all-electric ZNE homes that may, for example, offset 100% of their cooling loads during peak PV output in a high-cooling load climate. Clearly, if a level playing field is to be created for all promising advanced technologies the 2019 Code must include an all-electric prescriptive path, and TDV must account for the extent that distributed generation reduces or eliminates grid losses given climate conditions, patterns of energy use and whether HVAC loads are being offset by solar or wind at the time HVAC equipment is running.

The performance and efficiency ratings of HVAC systems listed in the AHRI Directory (Air-conditioning, Heating, and Refrigeration Institute) have been well accepted as the national standard recognized by both industry and regulatory stakeholders for over 50 years. The testing standards developed by AHRI have evolved over many decades to perfect and improve control and definition of lab test variables so that all manufacturers can be assessed on a "level playing field" wherein bias toward one technology is eliminated through the repeatable and verifiable methodologies that attempt to closely approximate real world conditions in which these systems operate. Although field-installed HVAC system testing can provide valuable data and product development feedback loops, there are no nationally recognized standards or test protocols for field testing, and it would be beneficial to evolve such standards over time. One important focus of this paper is to evaluate the appropriateness and accuracy of field testing as a method of vetting, verifying or cross checking AHRI Standards 210/240 and 1230 lab testing protocols and ratings. If this protocol vetting process is to be accurate, it must assess and control numerous field test variables that can significantly impact the validity of data. Because this type of field verification appears to be given more weight than nameplate ratings in California, it is especially critical that we consider the complexity of variables that must be controlled in an effort to achieve a "fair and level playing field" for evaluating VRF and VCHP systems.

In general AHRI testing standards are revised every five years, however, AHRI Standard 210/240 -2017 had not had a major revision since 2008. It also should be noted from the history of the 1230 Standards, such field testing protocols are likely to take as much as a decade to develop, vet and build industry consensus. In the near term context of the zero energy requirements that are taking effect in a year, it is in the best interest of the public, designers, manufactures and regulators to move forward with existing testing and rating methodologies in order to assign appropriate and fair compliance credit to VRF and VCHP technologies while field testing and consensus building around field test protocols are ongoing. With the launch of the first zero energy building standard in the world only a year away, the importance of developing clearer definition of test protocols and fair and objective process and criteria for assigning compliance credit have never been more critical to the collective goal of achieving zero energy building methods at cost parity with current building practices. Creating such a "level playing field" for rating systems and assigning compliance credit is in fact the key to removing the only obstacles to commercialization of the most advanced energy-saving technologies that can be integrated into buildings in California.

The Obstacles to Variable Capacity Heat Pump Marketability

There are few barriers to the commercialization of variable capacity heat pump (VCHP) systems other than questions surrounding rating methods and test protocols which pose a significant obstacle to widespread market penetration in California*. The various manifestations of VCHP and VRF technologies for both residential and commercial applications are mature and have achieved economies of scale that make them significantly more competitive with more conventional HVAC systems than they were a

*AHRI refers to VCHPs as variable-speed mini-splits and multi-splits (VSMS) for capacities less than 65,000 Btu/H and Variable Refrigerant Flow (VRF) multi-split systems for greater than 65,000 Btu/H decade ago. Recent cost comparison research for tract developments in Austin, Texas indicate that the cost premium for VCHP systems compared to condensing furnaces with split-DX AC systems was less than $1000 or about 10% including installation. However, VCHP systems offer significant product advantages in terms of both long term energy savings, indoor air quality and the comfort of true zonal control. It is clear from numerous case studies that VCHP residential systems operate at significantly higher efficiencies than conventional systems provided they are properly sized and they are installed in high performance shells.

While VCHP systems may have higher installation and integration costs in retrofit scenarios, their commercialization maturity has brought them close to cost parity with more conventional gas systems and their product advantages would allow them to compete effectively in the near-term market were it not for the CEC's reservations about accepting AHRI test data that would provide the basis for issuing them proportional compliance credit. Withholding this compliance credit assumes that architects applying for permits under the performance method of compliance (97% of all applications) will have to integrate other more expensive features in order to compensate for the artificially low compliance credit awarded to VCHP systems. On average, the compliance penalty awarded to VCHP systems imposes an artificial "tax" or cost premium in the order of $3000 to $10,000 depending on project scope, making it nearly impossible for architects to integrate the most efficient all-electric HVAC technology on the market in favor of high carbon-footprint .80 AFUE gas furnaces that have been the mainstay of the industry since the late 1970s. The CEC's own protocols and policies are dictating the use of these antiquated and archaic systems.

From a brief review of the last 15 years of test protocol development it is clear there have been some real challenges building industry and regulatory consensus at both the national and state levels. A great deal of analysis went into the development of inherently more complex test protocols that were developed by industry working groups between 2003 when the first variable refrigerant flow (VRF) systems were introduced in North America and AHRI's eventual acceptance of the Standard 1230 in 2009. The testing was based on the concept of having a defined "tested combinations" for both ducted and non-ducted system combinations (pairings of indoor and outdoor units). This iterative working group development process established that the sum of nominal capacity of the indoor units would be within 5% of the nominal capacity of the outdoor unit. This enabled multi-split VRF systems to be compared to traditional unitary systems on a level playing field for all manufacturers and models.

Through the working group test development process it was immediately clear to industry stakeholders that VCHP and VRF systems needed to evolve a rating metric that could more appropriately evaluate the efficiency of continuously modulated variable capacity systems. Traditional test protocols that take "snapshots" of capacity could no longer serve as an appropriate lens into performance. More test-points of performance at various levels of capacity would need to be taken and mathematical models to fill in performance curves under varying climate conditions would need to be developed. In the years that followed VRF technology introduction in North America, the new metric for the VRF systems evolved: the IEER, Integrated Energy Efficiency Ratio was adopted.

The IEER system efficiency testing is conducted at 100%, 75%, 50% and 25% capacity levels while accounting for numerous changing variables. Weighted averages were developed to calculate variable capacity systems in a way that accounted for the numerous variables that could not be assessed through traditional test protocols that took "snapshots" of performance at specific speed and ambient temperatures. The four weighted averages were based on simulations run using fifteen US cities in different climate types and across three building types. It took until 2012 for the DOE to vet and accept this test protocol and adopt it for use under federal standards. Today, AHRI selects about 20% to 25% of the basic model groups (BMGs) of each manufacturer to test every year so that over the course of 5 years all of the BMGs are tested. If any models do not test within 95% of their EER rating or within 90% of their IEER rating, they do not meet the compliance certification testing requirements and will be required to go through a mandatory rerating process. Through a rigorous, decade long protocol development period through which industry and stakeholder consensus was built, the IEER metric emerged as a much improved lens through which to evaluate variable capacity systems.

Even so, the IEER metric is not without its margin of error as it applies conglomerate weighted averages to predict and integrate variable demand conditions which will invariably match some climate conditions more closely than others. Without a doubt, it has greatly improved accuracy of VRF system ratings. In contrast, the HVAC Industry and AHRI, as its preeminent industry trade association, have not as yet evolved field test criteria that would serve as an appropriate cross reference or field verification process for the AHRI Standard test protocols. Given the complexity of variables that exist in the field, it is in fact a daunting task to define the many variables that may exist, and the prospect of building industry consensus on such a proposition is questionable. Many industry stakeholders do not believe the development of a field test standard can control the myriad of variables that exist under field conditions. But we shall nevertheless here discuss how such a standard would take shape if it were to be developed.

A recent study by Cadmus Group for the Vermont Public Service Department entitled "Evaluation of Cold Climate Heat Pumps in Vermont" indicates that heat pumps generally performed within 90% of their rated capacities under extreme cold conditions unless installed in older building shells with unforeseen heating and cooling load issues (leaks and thermal bridging). Field rated capacities dropped significantly (17%-22% deviation) when building shells were lower performing. Given the indications from this study that field-rated capacities vary significantly with building shell efficiency, it appears that field testing of VCHPs in owner-occupied new tract housing designed to low-load standards is needed to verify these correlations.

Similarly, the Ecotope Study reveals a high correlation between AHRI ratings and field tests with the added measure of cross-referencing air-side and refrigerant side capacity to confirm these correlations. To their credit, their conclusions also point to the need to develop and design field test protocols based on an iterative process of protocol refinement as correlations between lab and field tests increases over time. In turn field tests may guide or refocus future revisions of AHRI protocols: "The early analysis of the field data compares well to the lab measurements. Both measurements provide a useful cross reference for each other. The lab data is collected in a stable, repeatable, and highly controllable situation which provides a "reference set" for the field measurements of similar DHPs (Ductless Heat Pumps). Likewise, the field metering of COP shows which equipment operating modes are most common and therefore the most important parameters to measure in the lab. At the outset of the project, Ecotope did not anticipate the amount of synergy between the two data sets. It yields more confidence in both, while simultaneously demonstrating the benefits of an integrated evaluation approach to ductless heat pumps."

Insights revealed in these studies offers us a launch point. They reveal there is a long list of variables and parameters that need to be controlled if field test results are to accurately corroborate AHRI ratings. Because we are discussing the corroboration of two test protocols, a great deal of industry stakeholder discussion is needed to vet methodology and build consensus. We are talking about using one test to test the veracity of another (a test of a test). So how can we actually claim that one shows greater resilience or accuracy than another without such discussion? The results from each offers its fair share of insight. One would be prudent to ask: “Is the goal of lab versus field testing the same?” They are in fact just different lenses through which to evaluate performance, and epistemologically speaking, there is no “master protocol” that provides the acid test of all others.

It is important to note that the correlations alluded to in the Cadmus Study are largely theoretical. Because complete HERS ratings were not performed on any of the sample homes and all samples were existing structures that were not built to the low-load standards prescribed by the proposed 2019 California Energy Code, it is difficult to interpolate the capacity or efficiency values of the equipment, as they are directly affected by the installation variables themselves. Funds were not available to test building leakage and heat loss through thermal bridging, so it was not possible to accurately calculate room by room heat loads. The Cadmus Study also appropriately admits to margins of uncertainty. Because the research is specific to cold climate testing and the study acknowledges that there is insufficient data for cooling loads to accurately vet the accuracy of cooling capacities under these climate conditions. With too few data points and test samples, the uncertainty of some of their findings increases. Similar tests would need to be conducted in both humid and dry cooling load climates with complete energy audits of the shells to confirm this high correlation for these climate and load conditions. There does appear to be a significant enough body of data in the Cadmus Study to indicate that low-load building shells are less likely to have an impact on field-verified heat pump capacity tests, and this hint of a correlation should offer guidance for designing future field test protocols. Low load buildings are simply less likely to introduce uncontrolled variables, thereby increasing the accuracy and repeatability of the data.

In their current state of evolution, field test protocols are not useful as a substitute for lab ratings precisely because they are susceptible to too many uncontrolled variables. However, they provide some opportunity for a field cross reference of nameplate ratings within a broader margin of uncertainty. Insufficient field testing has been done to vet field test protocols, isolate the variables or even define the range of uncertainty. Given that the CEC has its concerns about a discrepancy between AHRI ratings and their field test results, it is more critical than ever to promote discussion and consensus around a process of iterative, continuous improvement of field test protocols while noting the absence of broadly recognized process controls and margins of uncertainty. Because there is a high probability that field test protocols will take a decade to perfect, compromise is needed to accept some intermediary process for either accepting AHRI ratings in the interim, or evolving an interpolation of them specific to higher load climate zones based on the limited data we have to date, such as the Cadmus Study, the CVRH Study and the Ecotope Study. Bringing together the very knowledgeable researchers that were involved with these studies to have a candid roundtable discussion on lessons learned would be extremely constructive, and perhaps a significant milestone in the iterative protocol development process.

In the paragraphs that follow, we shall discuss some variables that enter into field verification testing that should be controlled in order to assure a “level playing field” for evaluation of a field test methodology. It is unlikely that field test protocols will evolve to a level of acceptable accuracy and certainty in the near term, although it is imaginable that they may at some point call current AHRI test protocols into question or identify test and calculation methods that are in need of improvement. In this case new lab test protocols may be required which examine a greater range of load conditions (more test-point temperatures) or a different weighted average formula (as in the current IEER equation).

But the reality of field testing is that it is highly suspect because there are inevitably so many variables to isolate and control that can affect test parameters, thereby challenging repeatability or coloring the data or the “lens” or paradigm through which we examine the results:

a) If the home is not occupied, how do you mimic the humidity and temperature impacts of occupants: breathing, showering and operating exhaust fans and dryers? Do you assume an accumulation of three quarts of water vapor per person per day? How could a standard protocol factor this variable?

b) If the home is occupied how is the behavioral variability controlled when there are significant differences in HVAC control use, showering and operation of windows? Should occupied homes be tested?

c) What is the test standard for how well the building should be air-sealed for the purposes of the test? Should in-field tests be performed on older structures? Is it adequate to perform energy upgrades on older homes (improve attics) if orientation and wall R-value can’t be upgraded or factored? Should all standardized field testing target optimized new construction only?

d) What should be the test standard for R-value of the walls and how should variation in insulation quality be accounted for? Should field testing be limited to new tract homes with photographic record of QII inspections?

e) What about differences in the structure’s framing, the extent of advanced framing measures and thermal bridging through the framing? How should such thermal bridging be quantified or standardized? Which shell performance measures have the most impact on system performance?

f) What about system and zone sizing, specific room by room load calculations and the system design schematics being controlled and following precise protocols? Should room by room load calculations be required for the system sizing for all equipment tests?

g) Should differences in the HVAC system integration philosophies be considered or prescribed? For example, “Old School” or traditional design protocols locate supply registers at the outside of the structure near the windows and prescribe hanging all ducts from the ceiling in the attic where temperatures and distribution losses are greatest. In contrast, “New School” thinking on ducted systems is to make ducts as short as possible, locating the air handler near the center of the structure and placing ducts low in the attic with registers at inside walls throwing the air to the outside walls. (Rick Chitwood School of thought)? Typically, deeply burying ducts and shortening them by positioning registers closer to the system can reduce cooling loads by 40% to 50%. Does it make sense to even test systems in field if they follow an inherently less efficient install protocol?

h) Should external static pressure (ESP) be verified on every ducted system that is tested? Is it sufficient to follow HERS testing protocols that rely on watt-draw as a measure of system flow but do not measure ESP directly? Should ESP be measured on mini-splits?

i) Should “throw” at the grills be tested to verify proper room mixing? Should IR imaging of room mixing be required to verify this design parameter meets a minimum requirement? Should temperature sensors be dispersed through each room to test for variations in mixing and stratification effects?

j) Should field tests performed to verify the accuracy of standard lab tests have stringent field install QC measures to verify that system installers are following manufacturer recommendations for air sealing, refrigerant line sealing and refrigerant charge verification?

k) Should test equipment be installed in the same precise way that it would be permanently installed in the home? Does it impact test results if ducts are temporarily suspended from the ceiling or running across the floor inside the conditioned space rather than inside chases, soffits or attics as they would be if permanently installed? Is there greater distribution losses in one test scenario versus the other?

This is just a preliminary list. There are far more issues to discuss than we can fit in a ten-page paper. Yet from this limited preliminary list of variables it is apparent that all of them should be controlled under field verification scenarios. Compounding the epistemological questions is that there is little to no consensus on how these parameters should be controlled, but it would be beneficial to the objective comparison of results and conclusions if standardized field test controls were recognized and adopted by researchers nationwide. From this point, how do field tests proceed in the face of such broad margins of uncertainty? Is it not necessary to build consensus regarding the objectives and methodology of the test at the outset and then revise the process iteratively as data is gathered? For example, there are those who have conducted field test research who have argued that installation performed by HVAC contractors should not be monitored or verified by engineers or tech support because “that’s not how it happens in the real world”. But wouldn’t the results under this assumption be more of a test of install quality than equipment efficiency? Which are we actually trying to test? Which is going to best attempt to verify rated capacities? What are the defined objectives of the test? Are we rating the manufactured appliance, or are we evaluating the work of the HVAC contractor that the field test team happened to hire for a particular field research site? The challenge of such field tests is that they are attempting to make the real world fit into a test tube, but the world doesn’t really want to fit neatly into that structure. It’s an epistemological challenge that cuts at the very roots of the scientific method which is generally more comfortably controlled and repeatable in laboratory settings.

The question it seems is not whether we are mimicking the real world, but whether an effort to crosscheck and verify the statistical and calculated accuracy of a lab test procedure can be an accurate measure if all such variables are not very strictly controlled. The fact that there is still not consensus on such field testing criteria points to both the inherent and inevitable complexity of such a paradigm. Clearly, there is a need for a great deal more discussion, perhaps a decade’s worth, to build the required consensus in the industry and define the necessary protocols and process controls. Those familiar with the math would not dispute that any one of these field test variables falling outside of a range of tolerance can bring the veracity of the methodology into question. Clearly, the credibility and accuracy of field research and the value of the data and conclusions will be enhanced by an effort to facilitate such an industry and regulatory consensus on field test methodology. Such an effort should be taken up as the subject of a research grant that analyzes these important epistemological questions and tries to come up with some clear directives, goals and interim test methodologies which may be eventually modified in a few years in light of new findings that help isolate the most significant test variables.

The difficulty of accurately field testing VCHPs and VRF systems is clearly not limited to one agency. Field test researchers invariably face challenges attempting to correlate data across radically different case study conditions. The Cadmus Study points to margins of “uncertainty” and variability in data, particularly on capacity testing in the context of older high-load homes that do not have optimized shells. There is also a degree of uncertainty inherent in studies which gathers data on homes that have had systems previously installed by your “average HVAC contractor” who may have not have had “new school” training, probably did not obtain permits and is highly unlikely to have performed shell leakage tests (blower door and duct leakage), and zone by zone heat load calculations to properly size equipment. Leaky building shells with high thermal bridging, combined with poorly sized equipment could in many instances reduce tested system efficiency to half of AHRI ratings. Leaks and improper refrigerant charge verification could take efficiencies even lower. Tests performed without controlling these variables by their nature will yield data and results with higher margins of uncertainty.

Field testing *has* been regarded as extremely beneficial for product development efforts, iterative design and “continuous improvement”. The CEC field tests have created very productive feedback loops that facilitate these iterative improvements in design. For example, some models examined in the CVRH Study were found to have control algorithm issues, and others comparatively low efficiencies when systems are running continuously at maximum capacity. The latter appears to be more of a sizing issue which may have resulted from poor communication between the manufacturer teams and/or installation errors. Generally, VCHP and VRF systems are most efficient when they are frequently running at 60% to 70% of their rated capacity. Most manufacturers would recommend sizing to optimize efficiency in this “sweet spot”. Any potential “failure modes” or design shortcomings should be reported in writing back to the manufacturers so their development teams can engage in “continuous improvement” and process controls to address the issues. Such feedback loops foster a collaborative environment wherein field research can cast light on product development opportunities, and manufacturers can continue to improve efficiency and competitiveness.

Evaluation of Traditional Test and Rating Models

The traditional test protocols that were originally developed to rate single-stage furnaces and AC systems, take a “snapshot” of system performance to rate the efficiency of the equipment capacity and efficiency at 17°, 47°, 70°, 80° and 95°. It is logical and cost effective to evaluate capacity at these outside temperature test points, however, these snapshots may not give the full picture of performance in more extreme climates. Typically, one of these temperature points is used for modeling heating and cooling loads relative to “design temperatures” that are typical for different climate zones. In reality, even a single-stage system will not perform at the same efficiency when it is 80° as when it is 110° outside or at high and low humidity rates. So even for single-stage equipment these test methodologies are generalized and do not precisely match all install and climate conditions. Models and tested capacity ratings can only approximate reality. Snapshots are not motion pictures.

The situation is far more complex if an air conditioner operates at 20 speeds, or 200 speeds (continuously variable) rather than one fixed capacity. Capacity is varying with outside temperature and load conditions as well as with speed. Although the typical VCHP system varies capacity and efficiency continuously, system efficiency is generalized based on heating capacity rating at specific outside temperatures, say heating capacity at 47° in a mild climate, or cooling capacity at 95°. This “snapshot” of variable capacity is tantamount to rating the performance of an Olympic figure skater based on their facial expression in two photographs during the performance. Because the VRF system performance is continuously variable, we just aren’t going to capture the true performance of the system in our evaluation without a methodology that is more like a movie camera, and that type of clear model that perfectly captures the full continuum of performance doesn’t yet exist.

The Integrated Energy Efficiency Ratio (IEER) metric is an attempt at a more integrated approach to performance testing using a mathematical model based on well-researched weighted averages. Through an extensive and well-documented survey of heating and cooling load demand averaged across many cities in 15 US climate zones, weighted averages prorate the approximate periods of time that a system typically operates at four capacity levels: 25%, 50%, 75% and 100%. This IEER is a better lens than the traditional test protocols, and perhaps may be the most accurate acid test the industry will have over the next twenty years. It nevertheless falls short of perfect accuracy:

a) IEER rating is a conglomerate figure based on weighted averages of 15 climate zones, so it may not accurately reflect the field measured capacity under local climate condition;

b) the ambient (outside temp) test points are 65°, 68°, 81.5° and 95°, again a range of temperatures that may not reflect what is happening in Stockton or in Palm Desert in August;

c) IEER is a metric for cooling loads only and is not representative of a range of heating load scenarios and capacities that are normally evaluated in terms of HSPF (Heating Season Performance Factor) and COP (Coefficient of Performance).

The same complex method of weighted averages has not yet been applied to heating mode testing. Clearly, protocols will evolve over time to better mimic real-world conditions across a broader range of climate zones. Industry stakeholders come together regularly in AHRI working groups for Standards 210/240 and 1230 and are committed to continuously improve protocols. As we have seen from the history of the 1230 Standard, it is clear this evolutionary process takes time.

Federal DOE Efficiency Metrics Don’t Mesh with Utility Efficiency Needs

While ASHRAE Standards provided the basis for AHRI test protocol development, the actual federal regulations are mandated by the Department of Energy who view efficiency through the lens of system and test lab ratings. DOE does not exert jurisdiction over installation quality which is relegated to state-controlled agencies such as the CEC. In parallel with these unitary system ratings, the utilities are focused on metrics that allow them to save “nega-watts”, energy savings that allow them to use efficiency measures as a means to avoid investment in new generation capacity. They tend to focus on kWh’s saved and often interpolate EER and SEER ratings into projected kWh savings, return on investment (ROI) and years to payback that allow them to sell efficiency as an intelligent investment to consumers.

An integral part of these metrics and savings projections is the implicit assumption that an air-handler or mini-split’s rated efficiency and associated kWh savings are as predictable as the kWh per year savings of a refrigerator or a microwave. It is true that a given 10c.f. refrigerator will perform in about the same way in ten radically different homes. However a unitary appliance that behaves predictability is inherently different from HVAC systems that are integrated into and inseparable from the shell of the house. The presumption is that the SEER rating of an indoor and outdoor heat pump system should be roughly the same when tested in the field in ten different houses of radically different age, size, building type, and shell characteristics.

However, this is no more true than if we took the same compressor and coil system from two 10c.f. refrigerators and installed one of them in a 30 c.f. refrigerator with no seals. Now imagine variable capacity is part of the design and the 30c.f. refrigerator with no seals has a compressor running at three times the speed and 200% more of the time trying to catch up to the leakage at the refrigerator door. It is unreasonable to imagine that the high-efficiency compressor would operate at the same efficiency in both of these refrigerator boxes. Yet smart people may imagine HVAC system components should operate at maximum efficiency regardless of building shell performance. The coexistence of these radically different metrics and approaches to evaluating efficiency leads to the appearance that they are at odds, which they are not. They simply have different foci that appear in conflict if false assumptions are made about how they may be correlated.

"Minimally Efficient" VCHP System Rating by the CEC Versus AHRI Lab Ratings

From the examination of VCHP and VRF case studies it is clear that variable capacity systems clearly outperform more traditional HVAC solutions in a wide range of applications from single-family to high-rise office. The only significant obstacles to their broad acceptance in the California market is the CEC's requirement that they be "minimally" rated until appropriate modeling tools are developed. Take for example the following instruction in the 2016 compliance manual:

> *"These (VCHP and VRF) systems must be modeled as though they were minimally efficient units. The Energy Commission expects that the manufacturers will apply for a compliance option in the near future that will allow for the development of appropriate modeling rules to be included in the performance calculation approach."*

Heat pump and mini-split manufacturers nationwide puzzle over how the most efficient types of systems are required by the California Residential Compliance Manual to be "modeled as though they are minimally efficient". In many or most instances, 22 SEER systems are rated by the CEC at 14 SEER and additionally penalized by TDV calculations that give an additional 3 to 1 advantage to gas systems. Why are gas systems given such a huge rating and TDV advantage? How are the most efficient systems on the market so minimally rated? Although the modeling tools are now in place, sponsored by the manufacturers who paid for the modeling tool development by CEC's contract consultants, the AHRI rating of the equipment continues to be called into question by the CEC. It appears that the 2019 Code is on a path to continue the compliance penalty against clean technologies and favoring gas appliances: the proposed 2019 Code language offers no prescriptive path for all-electric ZNE homes. The CEC equipment rating directory still excludes VRF and VCHP systems and the TDV source energy calculations continue to favoring petroleum interests. How is this getting us closer to truly zero-carbon homes? It is time to get real about truly objective metrics that level the playing field.

Although the CVRH Study has provided a rigorous and valuable side by side comparison of different types of traditional and HP systems, it is questionable whether there has been a "level playing field" and control of all possible variables. This calls into question at least some of the study's conclusions. For example, the dehumidification and temperature control performance of a one-ton multi-split is compared to the dehumidification performance of a two-ton ducted heat pump system in an existing retrofitted (upgraded) house (Mayfair) but one whose shell performance is less than optimal despite the energy upgrades. The shell creates margins of uncertainty in the data as does the performance comparison between very different system sizes. This difference in equipment sizing may have resulted from differences in manufacturer recommendations derived during the test set-up, but it also appears that room by room load calculations were not completed and that differing zone loads may not have been factored. Although there were thorough energy upgrades completed, it is hard to imagine that the systems were sized based upon room by room, or zone by zone load calculations, a variable which is likely to have caused the mini-splits that were tested to run at maximum capacity and near minimum efficiency.

It is in the interest of all the stakeholders to advance high-efficiency technologies by developing level playing field metrics and protocols. Without fair cost-benefit comparisons between competing technologies the real societal benefits will not be realized: saved energy, more affordable ZNE prescriptive paths, smaller, less expensive solar panel arrays, lower demand on the grid, and a flattening of the grid demand curve (greater grid stability). Better communication is needed among all stakeholders so that the research is as atheistic and non-biased as possible. Without this collaboration and free and fair competition, end users and ratepayers will lose the benefits of clean technologies and all the societal benefits that they may bring. A great deal more discussion and perhaps field testing is needed to resolve disparate views on the validity of AHRI test protocols and to create industry consensus on field verification methods to vet lab test ratings. Clearly resolving these differences of opinion is in the interest of both manufacturers as well as the California building and regulatory stakeholders who must within two years find practical and affordable methods to build ZNE homes at or near cost parity with current building conventions.

Other Obstacles to Cost Parity in ZNE Structures

Over the past five years, some in the CEC have argued that VCHP systems should not be given their due compliance credit because they will offer a means by which designers can lower R-value and building shell performance. This concern seems to assume the worst of designers and this assumption does not connect with prevalent sentiments in the design field. This “slothful architect” argument seems to be based in the fear that designers, if given the option, will not invest in thermal performance of the shell which is an “infrastructure” item that is more difficult and more expensive to upgrade later. There are two reasons why this fear seems irrational:

1) Given the choice, no designer is going to spend \$20/s.f. to reach zero energy by upsizing the HVAC and solar systems to offset shell inefficiencies when they can spend \$2/s.f. to achieve ZNE through high-performance shell measures including passive house features and downsizing the mechanical systems. It is widely understood among architects that building shell improvements are more cost effective than compensating for shell inefficiencies with bigger mechanical and generation systems and virtually all of them feel and execute a fiduciary duty to build in the highest overall efficiency at the lowest possible cost. The will and integrity is there. They simply need more direction on how to get there in a cost effective manner.

2) Field research indicates that VCHP systems installed in higher performance building shells perform closer to their rated capacity by significant margins. Although more conclusive research is needed in this category, these preliminary findings are consistent with how variable capacity systems work; if loads are high due to air leaks and thermal losses through the shell, the variable speed fans and compressors are going to be running at higher speeds with predictably lower efficiencies. The benefits of integrating both high efficiency shell measures and VCHP systems produces cost-effective synergies. This synergistic relationship needs to be fully researched and vetted by the CEC and deserves the full focus of future field testing of VCHP systems in-low load, preferably passive solar homes.

In essence, it can and should be argued that with regulation, as with physics, there should not be a "free lunch". High building shell efficiencies should be required as a mandatory measures by the CEC with clear wall assembly direction including diagrams and checklists detailing wall assembly and air-sealing measures that should be clearly explained in the compliance manuals. The CEC's should overcome prior reluctance to define approximate costs and cost trade-offs of wall assembly materials and methods. Because cost-effectiveness is critical to winning contractor and market support, cost-trade off data should be approximated for all mandatory measures. For example, advanced framing (15 to 17% framing ratio) saves about $1.25 to $1.50/sf and spend that savings on blown-in compressible-type wall insulation systems to achieve a 35% increase in wall assembly R-value for about the same invested dollar. Building and lot orientation with attention to sun angles and insulated thermal mass such as R-10 full slab insulation clearly offer a 40% to 60% reduction in heating and cooling loads in many climate zones for a cost of $2.25/s.f. Such measures taken together can offer an ROI of about 15% to 30%. Independent cost analysis have indicated that all-electric 2019-compliant ZNE structures *can* achieve cost parity with structures built under current code requirements if such measures are mandatory, source energy calculations and TDV integrate the grid efficiency impacts of distributed generation and VCHP systems as well as HP water heaters are properly rated.

Most home owners are aware that they can't make a 15% return on investment in the best-case stock investment scenario, and they recognize the long-term value of the efficiency investment in the shell. Developers need to be trained or incentivized to worry less about incremental costs and focus on added value to customers, product advantages and ROI. Most architects and their client would embrace such mandatory measures. Since these efficiencies represent the most low-hanging fruit, the biggest returns for the lowest investment, the CEC has a fiduciary duty to make such measures mandatory. In light of the reality of climate science and AB 32 requirements, they have both a legal and moral duty to make such measures mandatory and create a level playing field through an all-electric prescriptive path. The synergies of combining high shell and HVAC performance with passive features is clear. Both HVAC and solar system costs are significantly reduced, but the greatest societal benefit may be the increased thermal storage capacity of the shell and its massive reduction of peak load demand on the grid, thereby helping to straighten out the "Duck Curve", reduce grid stresses and societal carbon footprint. The benefits of making shell performance measures such as passive solar features, highly insulated thermal mass and walls and high-performance attics mandatory is overwhelming. There is simply no reason to use the "slothful architect" reasoning as an argument against giving proper compliance credit to VCHP and VRF systems when making the CEC has the power to make cost-effective shell performance mandatory.

A Summary of System Rating Issues and Field Test Conditions

The particular advantage and testing challenge with VCHP systems is that they are variable capacity and their compressor and fans speeds are continually modulated so that their performance is optimized relative to cooling, or heating demand conditions. This both optimizes performance and makes them harder to test and rate because they don't just operate at one or three speeds.

In terms of performance, VCHP technology is an innovation that is as revolutionary as adding 20 gear ratios to a mountain bike (rather than one or three), which allows a minimal peak exertion for the rider to climb the hill. This performance optimization by definition means that variable capacity heat pumps and commercial VRF systems (variable refrigerant flow) are varying their efficiency continuously in response to load demands. Their optimum performance range is generally not at the higher speed range, because heat transfer efficiencies go down and static pressure and watt-draw of the equipment goes up at higher speeds. Like the bicyclist on level ground who doesn't want to be in either the highest or lowest gears, VCHPs generally operate happily in the 60% to 70% range of their capacity. Unlike conventional single-stage heat pumps, if VCHPs are a bit oversized, they modulate to the appropriate capacity range to handle the load efficiently and tend to operate less efficiently if undersized.

By comparison, the bike with one fixed gear, the single-stage ducted heat pump, can operate inefficiently if the system is oversized because it will draw higher watts, turn on and off too frequently and lose efficiency through these start-up cycles - like the engine that needs to warm up to reach its peak efficiency. Unlike single-stage HPs, it is better to oversize a VCHP systems a bit than undersize them because their peak performance is negatively impacted if they are running continuously at high speed. System sizing is therefore critical, and should be based on precise heat load calculations with room by room or zone by zone loads rather than overall building load calculations. If calculations are precise, based on a room-by-room load calculation method, the demand for conditioning is matched precisely by the supply without exceeding optimal capacity range of the selected system. A program such as WrightSoft that allows for both load calculations as well as room plans and system diagrams become critical for the execution of a properly matched load scenario.

This is actually easier to accomplish in a newly constructed building where the builder and subs can collaborate on critical measures such as duct design, air-leakage targets, and install quality. On retrofit projects, however, clients call an HVAC contractor when the furnace or AC fails, usually noticed during cold weather, and the client and the contractor are in a hurry to just fix it, without much thought about the fact that the system could be cut to half the size if energy upgrades are part of the scope. In the crisis management of the moment, no one is thinking about fixing the larger air-sealing and duct leakage problems in the attic, and the HVAC contractor is just focused generally on completing another job by the end of the day and keep the work flow moving. CSLB has confirmed that 95% or more of the furnace replacement jobs in the state are completed without pulling permits. Under such circumstances there is no third party quality control, and HVAC contractors are installing new equipment and hooking up to existing ducts without testing them for leakage. If they are not pulling permits, they rarely if ever perform heat load calculations, so eye-ball engineering and system sizing prevail. Invariably, even those HVAC contractors that understand the building shell synergies never inform the client that there are energy upgrades that can be performed because this complicates their tasks and involves other contractors and slows their job turn-around.

Tragically, most of these potential deep retrofits are never realized due to inadequate HVAC and general contractor training, poor CSLB enforcement strategies, and the lack of severe penalties for installing systems without permits. According to NCI data is that the average system installed in the United States only achieves 54% of its rated capacity due to installation quality variables, and far more should be done to reap this low-lying fruit and their cost-effective efficiency benefits. For this and other reasons, testing equipment already installed in older homes may not be a reliable cross reference of nameplate capacity.

There are is an abundance of case studies that provide clear evidence of high installed efficiencies for both VRF and VCHP systems. The highest efficiency improvements can be realized in multi-zone VRF systems in low-rise and high-rise residential or office applications where heat gains on one side of the structure may call for cooling while heat loss on the other side calls for heating. Under such load scenarios, line-set branch circuit controls (BCC heat recovery devices) use waste refrigerant heat from one zone to heat another. This is why the tallest passive house building in the world (built to PH standards), the 24-story Cornell Tech Dormitory incorporates VRF systems while achieving a 60% to 70% reduction in HVAC loads despite conditions where multiple zones are calling for simultaneous heating and cooling. Similarly, the Cadmus Study which includes a survey of 42 homes with retrofit installations of VCHP systems found that field tested capacity was within 10% of manufacturer rated capacities in homes that had relatively low heating loads (few leaks and lower than average thermal losses). Numerous other case study projects indicate that if properly sized, VCHPs (including variable speed mini-splits and multi-split ductless systems) can reduce HVAC energy consumption from 15% to 60% depending on the application and climate zone, especially when combined with low-load, high performance shells. This is true even in low-rise and high-rise structures, greatly reducing solar system size and cost on projects where rooftop space is limited. Obviously, the efficiency afforded by integrating VCHPs should never be used as a justification to build lower quality building shells, as the research also indicates that poor thermal performance of the shell may significantly impact HVAC performance. For this reason, it seems advisable that all jurisdictions and national and state building codes should move swiftly to make high efficiency shell measures mandatory regardless of HVAC system specifications, types or sizing. Low-load shells are simply the most cost-effective way to increase overall efficiency at a relatively affordable cost especially in new construction, and this is an infrastructure investment that is difficult and expensive to upgrade later.

A Forward Glance

From looking at other case study and field test data, it is clear that the CVRH Study's conclusions about VCHP system performance are in part related to system integration issues relative to appropriate sizing in retrofit applications. Comprehensive multi-home field studies with interim field test protocols need to be conducted on optimized passive solar low-load tract homes where differing systems can be compared in homes with identical performance characteristics including optimized orientation. Interim field test protocols need to be developed through multi-lateral stakeholder working groups and eventually more formalized field test protocols can be vetted over time. Prior field research such as the Ecotope Study have hinted at this clearer vision of a field test protocol. But such a key test criteria should not be dictated without stakeholder input and a high level of industry consensus. Given the fast approaching implementation of the 2019 CEC Standards, it is critical that industry and regulators move quickly and in a spirit of collaboration to bring about the needed protocol clarification. In the meantime, compromise on AHRI ratings is essential to creating a level playing field for clean technologies. CEC's inequitable assessment of VCHP systems as well as the bias built into source energy factors and TDV algorithms are in fact creating artificial market barriers. These barriers are tipping the scale in favor of gas appliances and petroleum interests and make VCHP systems and other clean technologies less competitive.

The comparatively low compliance credit offered to VCHP systems makes them artificially more expensive to end users and far more difficult to integrate due to the additional cost of compensatory measures required under the performance method. Industry players and HP manufacturers stand poised to partner with regulators to create a truly fair and level playing field for this technology which promises to pave the way to affordable ZNE strategies. It is therefore critical to move rapidly to evolve this level playing field by adopting interim measures to either partially of fully recognize AHRI directory ratings, AHRI Standard 210/240 and 1230 test protocols and do so in a manner that will neither be biased against VRF systems nor show them undue favor. The evolution of industry-sanctioned field test protocols will likely take at least a decade and should proceed without impeding or delaying the needed interim measures. In this way near-term goals and long term goals can proceed in parallel and unimpeded. Invariably, all major HVAC manufacturers have embraced a culture dedicated to "continuous improvement" as a fundamental principle of quality and process control. They welcome specific engineering, process and test protocol feedback, and are committed to ongoing collaboration with regulatory agencies in perfecting systems and processes in the quest to minimize environmental impacts.

QUESTIONS? Contact: Bruce Severance, Regulatory Compliance, Mitsubishi Electric US: 805-574-3207-bseverance@hvac.mea.com