California Technical Forum (Cal TF)
Technical Forum (TF) Meeting #10
July 23, 2015
Pacific Energy Center
San Francisco

I. Participants

Annette Beitel, Cal TF Facilitator
Jenny Roecks, Cal TF Staff
Alejandra Mejia, Cal TF Staff

Grant Brohard, TF Member
Doug Mahone, TF Member
Pierre Landry, TF Member
Bryan Warren, TF Member
Sherry Hu, TF Member
Mary Matteson Bryan, TF Member
Alina Zohrabian, TF Member
Spencer Lipp, TF Member
David Springer, TF Member
George Beeler, TF Member
Mark Modera, TF Member
Ryan Hoest, TF Member
Steven Long, TF Member
Owen Howlett, TF Member
Tom Eckhart, TF Member
Martin Vu, TF Member
Ed Reynoso, TF Member
Mike Casey, TF Member

Amir Roth, Department of Energy (DOE)
Kyle Benne, National Renewable Energy Laboratory (NREL)
Thianzen Hong, Lawrence Berkeley National Laboratory (LBNL)
Michael Blonsky, EMI Consulting
Beckie Menten, Cal TF PAC Member, MCE
Alice Stover, MCE
Mike Wilson, International Building Performance Simulation Association
Martha Garcia, Southern California Gas (SCG)
Marc Costa, The Energy Coalition
Josh Krauss, Pacific Gas & Electric (PG&E)

Christine Hanhart, UCONS
Alfredo Gutierrez, Southern California Edison (SCE)
Jia Huang, PG&E
Brian Smith, PG&E
Rick Ridge, Ridge & Associates
Rod Williams, Embertec
Steve Kromer
Todd Malinick

On the Phone
Armen Saiyan, TF Member
George Roemer, TF Member
Larry Kotewa, TF Member
Jon McHugh, TF Member
Bruce Harley, TF Member
Yeshpal Gupta, TF Member
Christopher Rogers, TF Member

Pete Ford, San Diego Gas & Electric (SDG&E)
Chan Paek, SCG
Roger Baker, ComEd
David Shallenberger, Synergy Companies
Ben Lipscomb
Celia Johnson, Future Energy Enterprises
Kyaewoon Chung
Matt Tyler, Clear Result
Michelle Friedrich, Sacramento Municipal Utility District
Nenita Plorin, SDG&E

II. Key Decisions and Action Items

Variable Refrigerant Flow Abstract

- ACT: Clarify if cost estimates come from retrofit or new construction program data.
- ACT: Workpaper should use different values for new construction and retrofit applications.
- ACT: Rud Judkoff can provide latest modeling engine comparison results to show how results of DOE 2.2 compare to EnegyPlus.
- **Abstract approved to proceed to workpaper development using EnergyPlus for modeling**

- Validating EnergyPlus modeling capacity at a component level would be too granular to provide added value given that EnergyPlus already models at the component level.
- Further validating EnergyPlus results against DOE-2.2 would not increase the accuracy of the measure.

Comparing the base and measure case estimates at the hourly level would increase measure complexity and add no additional value to workpaper RPP Workpaper Net to Gross

- **10 year forecasted NTG approved for use in workpaper and Program Implementation Plan until implementation data becomes available.**

Commission Staff Feedback on Cal TF Workpapers to Date

- Technical Forum suggestions for optimizing Energy Division involvement with Cal TF review process:
  - The workpaper development and review process would achieve significant benefits from having CPUC staff or consultant participation in TF meetings so that:
    - Staff/consultants can discuss/contribute their recommendations and requests with TF members and the workpaper developer, which could alter TF recommendations and outcomes.
    - Staff/consultants can hear TF member perspective, information and recommendations, which may alter staff/consultant input and requests.
    - Even if TF and staff/consultants do not always reach agreement on all issues, “real time” staff/consultant participation is likely to narrow issues, improve process efficiency, and reduce costs compared to the current ex ante team workpaper review process.
  - Energy Division staff should be empowered to make decisions during Cal TF meetings.
  - CPUC workpaper review process should consider and give weight to technical judgments rendered by the Technical Forum.
  - It will be important for Energy Division to review, comment on, and approve the final Cal TF ex ante development guidelines that will provide needed clear and written guidance to workpaper developers.

Smart Thermostats

- ACT: IOUs to provide the Technical Forum with pilot results before interim workpapers can be developed.

III. EnergyPlus as a DOE-2.2 Alternative

Amir Roth, Department of Energy—

PowerPoint Presentation

Pierre Landry—Has anybody done a DOE-2.2 vs. EnergyPlus comparison?

Amir Roth—Yes, I believe Cadmus is doing one right now.

Steven Long—Edison has also done one too. I think the big difference was that heating was off by a little bit.

Sherry Hu—Yes, there was roughly a 20% difference in heating energy consumption.

Kyle Benne—It is very unlikely that either program is getting the basic physics wrong. Assuming the loads and boundary conditions are all input equivalently, this difference likely amounts to an inconsistency in how the mechanical systems are controlled. There could be differences in how the two programs are responding to the thermostat set point and particularly the dead band between the high and low cutoff.

Annette Beitel—The modeling engine is one piece of the Cal TF's broader "DEER Alternatives" work. The question for the Technical Forum today is if DOE's proposal meets the various criteria we're using to evaluate our eventual recommendations for fixing the ex ante framework. Remember that the criteria are both policy considerations—including the Commission's long-standing goals for the framework—and key technical objectives—including ease of use and reduced costs.

Doug Mahone—For year the CEC resisted updating their modeling tools because they knew the new software would produce a different set of values compared to what Energy Division is using for ex ante value development. We should be cognizant that this would take some getting used to. Does DOE have a plan for dealing with the changes in modeled values as the open source tool continues to get updated?

Amir Roth—EnergyPlus *is* an open source good. However, that does not mean that anybody can change it. What it means is that anyone can look at the code and create applications or interfaces for the tool. Only very few people—all at LBNL right now—can actually make the changes.

Kyle Benne—And the reviewers who do have the ability to introduce changes to the code must go through a thorough testing system before any update can be made. This is all publically catalogued.

Tom Eckhart—I think this is a great improvement, but it does not address all of the problems with the current ex ante framework.

Annette Beitel—Absolutely. This is only one piece of the solution being considered by the DEER Alternatives subcommittee.

Kyle Benne—OpenStudio actually facilitates calibration with field data, because it is so much more automated that California's current tool.

Steven Long—One of the problems we have in California is that DEER and READI don't use the actual code for results. The approaches are interpretations of the code. Can EnergyPlus help us address that problem?

Kyle Benne—As part of our 50 state project we *are* developing building models that incorporate ASHRAE 90.1, so that should help address that concern.

Annette Beitel—We have Mike Wilson in the room. He is the Executive Director of the International Building Performance Simulation Association. Do you have anything to add to this discussion Mike?

Mike Wilson—I'd just like to make it clear that this is a crucially important discussion. We need policy makers to understand and opine on it, because modeling tools have important effects through the entire energy efficiency industry.

Mark Modera—I remember from years ago that calibrating models was a very complicated, manual process. How does your proposal help with that?

Kyle Benne—Having a lot of knobs to turn isn't as much of a problem now that we have such improved computing capabilities.

Mark Modera—I was more worried about making sure the right knobs are being turned—that the calibration process is well documented and you're getting the right answer for the right reasons.

Amir Roth—There have been several improvements in how we address that exact concern. One of the more promising approaches is Bayesian calibration, where you actually get probability distributions. There is also already a model calibration standard and our tool is automating this step too.

David Springer—I really like this proposal. I believe this can help California close the gap between Code CASE studies and measure development.

Annette Beitel—Thank you very much for your input on this. We will take the discussion back to the DEER Alternatives subcommittee, where we will propose concrete implementation recommendations for this as well as the other pieces of our solution to the problem. Please let us know if you are not a member of the subcommittee but would like to participate in the modeling engine portion of the work

IV. Variable Refrigerant Flow Abstract

Sherry Hu, PG&E—

PowerPoint Presentation

Bryan Warren—I actually just installed this technology in my home and am really happy with it. I actually find it more comfortable than the system I replaced.

Owen Howlett—I'm a little confused by the Measure Cost table. Can you please explain it?

Sherry Hu—The column on the left is the incremental cost from the two baselines on the right.

Jon McHugh—How many heads and how many tons are you including in your measure system?

Ben Lipscom—The heads tend to be 3-4 tons. The commercial application varies from 20 to 100 tons.

Spencer Lipp—Is the cost for the retrofit or new construction program?

Sherry Hu—I believe it is a weighted average of the program data.

Mark Modera—Your cost estimates seem fairly high.

Sherry Hu—Yes, the cost is high, which is probably related to the fact that it came from estimates from only two manufacturers. Energy Solutions is working on a survey to improve that estimate.

Mark Modera—On your question for the group, you're not going to get reduced ducts if this is a retrofit. I would suggest splitting out retrofit and new construction values.

- ACT: Clarify if cost estimates come from retrofit or new construction program data.
- ACT: Workpaper should use different values for new construction and retrofit applications.

Steven Long—Did you try to calibrate the CBECC Com. models against DEER?

Sherry Hu—Yes. This is described later in the presentation.

Steven Long—And how old was the data you used?

Sherry Hu—We are using the most recent program data.

Proposed Response to CPUC Energy Division (ED) Feedback on Modeling Tool

Spencer Lipp—What was provided before that ED found inadequate?

Sherry Hu—This is not entirely new information to them. I believe that what ED really wants from us is a validation of EnergyPlus with Doe-2.2.

Steven Long—I think the proposed approach is more than reasonable. However, I've seen some early data that showed the manufacturers were significantly off. I would just recommend you make sure you are using the most recent data.

Sherry Hu—Yes, slide 19 shows that we are using the latest data.

Martin Vu—I think it is unreasonable to request that DOE-2.2 be used to validate modeling for a measure it itself doesn't have the capability to model.

Amir Roth—The latest validation project shows less than 2% difference between the EnergyPlus and DOE-2.2.

- ACT: Rud Judkoff can provide latest modeling engine comparison results.

Mike Casey—In terms of the range of installations, aren't some of these units installed for more than one room so you do have some amount of ducts?

Doug Mahone—Yes, but that is still not enough to affect savings.

Steven Long—Seems like your response addresses the ED request appropriately.

Spencer Lipp—To the extent that it can be answered. The request is very vague and seems to go beyond existing CPUC requirements.

Mark Modera—Safeguarding for every single possible efficiency erosion in each different application would be an enormous task. Separating new construction from retrofit and commercial from light commercial may help easy ED's anxieties about the measure. .

Sherry Hu—We do have different measure codes for retrofits and new construction.

Annette Beitel—It seems like the group does support the use of EnergyPlus. What does the group think about the level of granularity being requested by Energy Division?

Mark Modera—I'm not even sure that separating between new construction and retrofit would affect the choice or accuracy of the modeling tool.

Group—EnergyPlus is already component-based, so ED's component request seems way too granular.

Steven Long—They are asking for a higher level of verification and documentation than exists for DEER or DOE-2.2. The ex ante team is trying to set a higher bar for use of EnergyPlus than exists for DOE-2.2 and DEER.

David Springer—I am more than satisfied by the evidence PG&E has already compiled.

Annette Beitel—So, to conclude, is the request that EnergyPlus be validated by DOE-2.2 reasonable?

Group—No.

Annette Beitel—And lastly, would there be value to comparing the base and measure cases at an *hourly* level?

Group—It is very unreasonable.

Mike Casey—And it is also beyond any of the capabilities of either of the modeling tools.

Martin Vu—It would be helpful if you could estimate the time and cost it would take to comply with these requests. I think it would go a long way to showing that the added precision and complexity is not cost effective.

- **Abstract approved to proceed to workpaper development using EnergyPlus for modeling**
    - Validating EnergyPlus modeling capacity at a component level would be too granular to provide added value given that EnergyPlus already models at the component level.
    - Further validating EnergyPlus results against DOE-2.2 would not increase the accuracy of the measure.
    - Comparing the base and measure case estimates at the hourly level would increase measure complexity and add no additional value to workpaper accuracy.

V. RPP Workpaper Net to Gross

Brian Smith, PG&E; Rick Ridge, Ridge & Associates—

PowerPoint Presentation

Annette Beitel—So, in the ten year Program Implementation Plans you will be reporting ten year estimates for all parameters?

Brian Smith—Yes, but that does not mean that we can't update those estimates as we get new data.

Annette Beitel – In the E3 calculators can you enter different NTG values for each year of the program such that the program TRC would based on expected results on a yearly basis?

Rick Ridge – No, the E3 calculator only allows one NTG value to be added for the entire ten-year period.

Annette Beitel – How difficult would it be to modify the E3 calculator to allow for NTG values to vary each year over the ten year period.

Rick Ridge – That would be difficult.

Pierre Landry—What is the reason why you're choosing just one NTG?

Rick Ridge—The Prahl market transformation white paper recommends this approach in order to avoid unnecessary complexity.

Pierre Landry—Will the Program Implementation Plan (PIP) for the RPP program be clear about your modeling method and assumptions?

Rick Ridge—Yes.

Pierre Landry—And you will also update all values contained in the PIP when you update the NTG based on actual program data?

Rick Ridge—Yes, and that could be as early as the end of the first year.

Doug Mahone—Is the forecasted NTG greater than 1 really possible?

Rick Ridge—It certainly is if you have enough market effects and multipliers. We've found NTGs greater than 2 for some NYSERDA market transformation programs.

Steven Long—You're talking about updating your ex ante estimates with implementation data. How will that be different from what you will be reporting?

Rick Ridge—We will report a ten-year forecast value using as much up to date data as possible.

Brian Smith—We will definitely have to revisit these estimates since we're expecting to see a hokey-stick type pattern in effects.

Grant Brohard—How much savings are you expecting the first year?

Rick Ridge—Statewide, including LADWP, we would expect somewhere around 7 Gigawatt hours in the first year.

Brian Smith—These products were specifically chosen for not having enough savings on a per unit basis to justify individual programs on their own.

Grant Brohard—And you are requesting this for a pilot?

Brian Smith—Yes.

Jon McHugh—How does the 5% initial Partial Leading Indicator (PLI) NTG on slide 14 relate to the NTGs you are now forecasting?

Rick Ridge—That PLI NTG only applies to the very limited activities we piloted last year. As we discussed last meeting, that pilot mostly tested key operational metrics and should not be used to extrapolate to the full program design.

Jon McHugh—Are you anticipating that the baseline and measure efficiencies will change over time?

Rick Ridge—Certainly. Codes, standards, and what gets incented will change over time.

Annette Beitel—Are there any other questions or concerns from the group?

Group—None.

Annette Beitel—So, does the Technical Forum approve the use of these ex ante NTG values for the next two to three years?

Group—Yes.

Annette Beitel—Is anyone opposed?

Group—No opposition.

- **10 year forecasted NTG approved for use in workpaper and Program Implementation Plan until implementation data becomes available.**

VI. Commission Staff Feedback on Cal TF Workpapers to Date

Annette Beitel, Cal TF Staff—

PowerPoint Presentation

Jon McHugh—I'd like to point out that the CEC doesn't always support the DLC. The CEC has a substantially higher standard available.

Annette Beitel—That CEC standard is for residential, and of course it makes sense to use the CEC standards for the residential application; however, the CEC does not have a standard for commercial LED yet.

Doug Mahone—Did Commission staff say what they would accept if not the DLC?

Annette Beitel—No.

Mike Casey—Do you know if the APS workpaper went to a different reviewer than the person who gave the initial early feedback?

Annette Beitel—It was the same reviewer.

Spencer Lipp—Another benefit to earlier involvement from CPUC staff is that they would get to hear the discussion for each of the parameters. We thoroughly vet every estimate and they don't see that by just getting the final end product.

Pierre Landry—Yes, but whoever comes to the table needs to have the requisite decision-making authority.

- Energy Division staff should be empowered to make decisions during Cal TF meetings.

Doug Mahone—I think we need to be pushing for a solution that clearly says that in cases of professional opinion, they should accept our collective decision.

Annette Beitel—That is a very important point. We have a group of over 30 highly qualified technical experts, with extensive graduate degrees, and decades of energy efficiency experience; the value of that experience should not be disregarded.

- CPUC workpaper review process should consider and give weight to technical judgments rendered by the Technical Forum.

Doug Mahone—I would say that the implementation of standards like “best available information” is also an exercise of professional judgment. We spend hours making sure that each measure’s data satisfies that standard and it’s ridiculous for Commission staff to just waving that away.

David Springer—I think it would be valuable to compare the time it take new measures to be approved by the ex ante team and the time it takes to update Title 24.

Owen Howlett—It is telling how entire code updates have been introduced in the time it has taken some workpaper to be approved—and the Title 24 process includes a public input process.

Mary Matteson Bryan—Who is developing the ex ante development guidelines?

Annette Beitel—SCE is developing written guidelines for ex ante value development. Cal TF is supplementing this work through a subcommittee. .

Mary Matteson Bryan—We need to keep in mind that Energy Division will have to accept and recognize them for the guidelines to be useful.

- It will be important for Energy Division to review, comment on, and approve the final Cal TF ex ante development guidelines that will provide needed clear and written guidance to workpaper developers.

VII. Smart Thermostats

Oriana Tiell, PG&E—

PowerPoint Slides

Annette Beitel—Do we know if the baseline is a properly programmed thermostat or the more likely improperly programmed ones?

Oriana Tiell—No, we don’t.

Mike Casey—And these are thermostats that the utility can control directly?

Oriana Tiell—Yes.

Doug Mahone—I thought the CEC stopped requiring programmable thermostats because they actually used more energy?

Group—ENERGY STAR did that.

Doug Mahone—So why does it make sense to use those as the baseline?

Grant Brohard—Because they are still code in California.

Mark Modera—Isn't EPRI studying these?

Lucy Morris—I believe we are expecting results from that study by next year. The individual studies were intended to be RTC, but some of them had to be modified.

David Springer—Is the data reliably sourced? Can we see it? Is it statistically significant? Can models replicate the savings?

Oriana Tiell—Many of these white papers are published by the manufacturers. I haven't seen a breakdown of the specific features that lead to the savings. The idea of this study is to better understand the important features.

Alfredo Gutierrez, SCE—

PowerPoint Slides

Jenny Roecks—Randomized controlled trial *are* the gold standard for data accuracy, but since that type of data is not available yet, SCE would like to create an interim workpaper that allows them to offer the measure and collect data in the meantime. What we would like to hear from the group is if you think the data that SCE does have is enough to create an interim workpaper.

David Springer—Have you tried to analyze the magnitude of the savings?

Alfredo Gutierrez—I haven't seen the data so right now we are using the estimates you heard from PG&E.

David Springer—Do you think the savings will get lost in the noise?

Alfredo Gutierrez—That is definitely a possibility, but we won't know until we try.

Spencer Lipp—But even if you do something, the possibility that those savings may be from something you haven't accounted for is still very real. You need to know what the thermostats are doing to really know what's happening.

Alfredo Gutierrez—Nest used to be unwilling to share their data with us, but now they are providing us with HVAC run data.

Annette Beitel—Would this be addressed by a statistically significant confidence interval?

Spencer Lipp—That would of course be needed, but even then you may just be seeing noise.

Mike Casey—IP MVP recommends that you need at least 10% savings to detect from whole building utility billing data.

Sherry Hu—I think we need to figure this out before we even access the data.

Owen Howlett—But these aren't questions easily answered with statistics. All statistical models are based on normalized distributions. The problem is we don't know the underlying distribution.

Pierre Landry—What we need here is a good program model. We need focus groups to try and understand what is underlying the usage data. I'd suggest analyzing the data before you limit yourself to draconian statistical bounds.

Owen Howlett—I'd be much more convinced if we had better information about the thermostat set points.

Oriana Tiell—From my extensive program experience, it seems like your target sample size is way over optimistic. What is your strategy for reaching that target?

Alfredo Gutierrez—I believe we will be leveraging the customers that have already signed up.

Oriana Tiell—Ok, good. You also need to take into account the enrolled customers you will loose because of things like unavailable billing data, etc.

Jon McHugh—Can we trust Nest to essentially put their equipment on "dumb" mode to do a pre-post study?

From the phone—That would need to be a different study with different customer approval protocols.

Martin Vu—Wasn't there a representative from Nest at the last meeting who was saying they were working on data availability?

Lucy Morris—Yes, they are working on several "packages" of data availability.

Mark Modera—Also, sample sizes may help with getting rid of random error, but not bias.

Pierre Landry—Or small magnitudes. One way to better understand that is to chase down the numbers—ask people what they were doing during the program period.

David Springer—This seems like a technology much better suited for DR.

Sherry Hu—We also need to consider the cost effectiveness. For instance, we get a similar volume of savings from Opower, but the program is much cheaper.

Lucy Morris—PG&E had originally planned a combined DR and EE trial. However, the EE potential is the real unknown so that is why we are focusing on that aspect now. The vision *is* to combine the two.

Mark Modera—That's dangerous because the EE and DR savings may come from opposing approaches.

Oriana TIell—That may very well be true. However, we think this particular technology is a great opportunity to delve into the broader world of Integrated Demand Side Management, so this is why we are investigating all of the possibilities. This is a starting point for a much broader conversation and goal.

Martha Garcia, SCG—

<u>PowerPoint Slides</u>

Chan Paek—A preliminary TRC on the gas savings was around 0.2.

Oriana TIel—PG&E believes that adding the electric savings may make the measure cost effective, but this is based on very preliminary analysis.

Ed Reynoso, SDG&E—

PowerPoint Slides

Oriana Tiell—The beauty of this study is that Ecobee has data for each individual customer.

Ed Reynoso—The data does reside on the Ecobee server, but we are working with Itron for independent verification.

Annette Beitel—What has been Staff's rationale for rejecting propensity score matching in the short term?

Lucy Morris—I think it is a valid argument about the contentiousness of EE savings from this technology and wanting to use the gold standard RCT the first time around.

Spencer Lipp—I'm confused by the variation in savings between July and August. Those months have similar temperatures. That makes me question the level of noise you're picking up.

Nanita Plorin—We did have an issue with one of the July events where the entire population did not receive our signal.

Lucy Morris—But aren't DR savings removed from that table?

Spencer Lipp—If the daytime distribution of the temperature affects the savings, then maybe average temperature is the wrong metric to be using.

Jenny Roeck—So, it seems like the group is not satisfied with the preliminary information to support an interim workpaper. It seems like the IOUs will need to come ask with more analysis to support their proposal.

- ACT: IOUs to provide the Technical Forum with more analysis before interim workpapers can be developed.