

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** Final presentations should be very exciting-- fruits of your labor over the entire semester in reality. The fundamentals in the first third, the technologies in the second, gearing up toward the cross cutting themes in the third. I understand that we've had an accelerated project schedule this semester. We've completed the entire projects over the duration of about a month and a half, so you are to be congratulated for your hard work in a very intense period of time McKenzie tiger team style.

For that, I reward you with doughnuts and coffee over there. I understand many of you were up late last night, so you're welcome to ingest some shortchange carbohydrates and some caffeine. If you would like to get some, get it now. We're going to have another minute of blah, blah, blah before we dive into the presentations and the real fun begins. I'd like to introduce our panelists up here in the front who will be the evaluation criteria a la *American Idol* style, except that you're, of course, a lot smarter and equally well dressed.

Starting from right to left in front of me we have Dr. Jasmin Hofstetter, who comes from IES, Spain. That's the Institute for Solar Energy in Spain. That's where she did her PhD with Antonio Luque. Those who have been studying intermediate band solar cell materials may know the name as one of the fathers of the field. She studied under Antonio Luque's organization with Carlos de Canizo and is the winner of presentation awards at scientific conferences, among others, so Jasmin is welcome here.

We have thought that Dr. Mark Winkler, as well, a PhD in Eric Mazur's laboratory at Harvard. Those who are familiar with femtosecond laser characterization may be familiar with Eric Mazur, also one of the fathers of the field. Mark started the

Harvard Journal Energy Club, which is Harvard's version of the MIT Energy Club-- a lot smaller and a lot less dynamic than MIT's version, but nevertheless to be congratulated. And of course, a very good organization. I kid. There's a little bit of MIT Harvard rivalry.

And of course, our very own Joe Sullivan, who has been with you the entire semester. For those who might not be familiar as much with this research as you are with his teaching, Joe is studying intermediate band solar cell materials here at MIT in the Media Lab and has been working for the last-- what is it now?

**JOE SULLIVAN:** Three and a half.

**PROFESSOR:** Three and a half years-- focused on intermediate band solar cell materials, coming from a very broad background in energy from climate science. So with that, I'd like to welcome our first team down, and the floor is yours.

**STUDENT 1:** Good morning, we are the PV smart retrofit group, and our project goal was to assess whether or not there was an electrical benefit or loss from retrofitting an old home with a PV system. Now, in less lofty terms, it essentially means from an on site energy perspective, does it make sense to put PV panels on my house? You'd think that that's a kind of an obvious question. We'd all say, well, yeah.

We produce energy for free, except that you have to consider other things, such as shading from trees that you'd have to cut down or the color of the roof that you might be changing by adding a black panel. These would both reduce the thermal load of your house in normal situation. So by adding a PV panel and increasing the thermal load, you actually add the energy to cool your house during the summer.

We considered multiple variables in this project. One was location, which has an effect on the amount of sunlight you're receiving. The PV panel's presence and its size. We had a couple of situations where there was no panel as a kind of a baseline. And then also different sizes to figure out if a bigger panel had a bigger effect. We looked at two different colors, black and white. Those are the ends of the spectrum, and they give us endpoints to look at, and the color matters because a

darker color will absorb more heat.

We looked at roof pitch. The reason for this-- well, first off, roof pitch is the angle of your roof. And we looked at this because we assumed that our panels were fixed and parallel to the roof, so this kind of controlled what angle your solar panel was facing towards the sun. And finally you had your house footprint, which is the area that the house covers, and when you combine that with the roof pitch, you get the area of the roof, which is really what we're concerned with.

We had five scenarios, and as you can see from our cute little diagrams, we have a black roof, a white roof, a white roof with the tree, a white roof with a solar panel, and a white roof with a solar panel and a cut tree so the solar panel is getting plenty of sunlight. In evaluating the five scenarios, we had three models of increasing complexity from left to right, which you can note from the fact that model one only covers three of the five, and these models all had common assumptions.

The first was that you had a single story house in a suburban locations, so you didn't have shading from other buildings, for example. We had a common house size of about 2000 square feet, or 186 square meters. The roof pitch, which in construction is set usually at 5/12, which means 5 inches of rise for 12 inches of travel. We had an unfinished attic space, which means that it's sealed to the outside, but you didn't make it livable, a five kilowatt PV system covering 36 square meters, and we chose reflectance values of 0.08 for black and 0.35 for white, and this, again, reflects the effect that color has on heat absorption. With that, I will turn you over to, Jordan.

**JORDAN:**

So the first one that we looked at was basically using most readily available and simple models there can be. So this is from the Department of Energy to measure-- well, to get a gauge of how the color of the roof and the roof properties affect the thermal loads in the house. This is a simple one dimensional model where you've got an inside cavity of 65 Fahrenheit, and you basically input the location, and from that it has a lookup table of the average insulation on the house, as well as the number of heating degree days and cooling degree days relative to that 65 inside

temperature.

The other parameters are just at the roof, so we insert the reflectance, and we're using values that represent real tiles for an average house, black and white, as well as the thermal resistance and the heat absorbance of the tiles-- from this model, the outputs and the thermal loads in terms of the heating you need to put in, as well as the cooling energy load.

So with the thermal model assessed, we can assess the photovoltaic output. And for this, we're just using a simple model PVWatts. This basically takes in the location and the angle of the panels, as well as a derate factor from converting to AC to DC. It's quite a simple model, and output from this is the amount of kilowatt hours per year that you get from the panels.

**STUDENT 2:**

So I'll be talking about a couple thermal electric model. This is the model that we built in our group, and we developed this model based on two sets of individual parameters and individual models. I would say that one comprises of the thermal model, and there is the electric model. So what we need to note from this model here is it's a further step in complexity when compared to the module one that Jordan just discussed. And it takes into account various input parameters that the model one doesn't take into account.

So the basic structure of this model is as follows. We have a thermal model which takes into account several input conditions, such as the insulation and shading, and it outputs a living space temperature, which is this temperature of the living room in our house. And then this temperature is fed into an electric model, which calculates the cost and energy values, and thus we can compare an energy production and energy consumption.

Going to the thermal model in detail, I've just shown a picture here. So it considers basically when we start from the top of the house, we use insulation and shading as the input parameters. We then calculate the temperature of the PV panel, and then the temperature of the roof using two different energy balance models. And these two energy balance models are pretty robust in the sense that they consider all

these physical phenomenon which are realistic, such as the convection, radiation, and PV electrical output. And based on these energy balance equations, we can calculate the PV panel temperature and then the roof temperature.

And once we get these two temperatures, we get the heat flux that goes into the roof and that enters the attic. And once you get this, we find out the attic temperature, which then determines what is the ceiling temperature. And then we consider the convection via ceiling, and then finally we end up with the living space temperature. And then the couple this living space temperature with another electric model, which I'll discuss now.

So the electric model is basically based on an ideal gas assumption. So it basically-- what it does is it calculates the energy needed by the AC, which could be the heating or cooling, in order to maintain the living space at a particular temperature. And we use this formula,  $m \cdot C_p \cdot \Delta T$ , which is an ideal gas formula, which gives out the energy needed by the cooler. And then so the electrical model uses the power consumption, and then we know the power production through PV output. So comparing these two, we can really assess whether PV installation is favorable or not.

And this is just the model in making. What we want to signify here is we actually made this model on our own, and this is the MATLAB code we wrote. And the model not only just predicts energy values, but it also can do a lot more things, such as predicting temperature. And what I've shown here is the PV output, and then the cooling load that is required, so it can do a lot of other things, as well. This is the temperature of the roof in terms of direct sunlight and diffuse sunlight. So if anyone is interested, I'd be happy to discuss with them more. Thank you. And I'll now pass it on to Heidi.

**HEIDI:**

So I take over from here talking about the third and most complex model that we used. For this model, we used two different softwares-- one called BEopt from NREL, and the second, called EnergyPlus, which you'll see later, developed by the Department of Energy. So what we did in this model was we took into account the

3D effects of these thermal and electric loads that have been talked about in the other two models.

The first thing we did was to actually model the 3D house, as you can see right over here. BEopt allows for a very nice interface, where you can easily model the house and easily put a whole bunch of input parameters for the house. And for these input parameters, we consulted with an experienced building inspector for the construction inputs and used BEopt default values for the rest of the inputs.

Once this model of the house was done, we had to actually export it into EnergyPlus because EnergyPlus gives us a much more detailed look into all these different parameters, I guess. And you can actually go in and modify different things. For the materials, we can modify every single material property-- conductivity, density, specific heat-- and we actually did that for the roof. And we also removed a whole bunch of other miscellaneous loads that BEopt had included.

So this is how we modeled the trees and the panels that we talked about in our scenarios before. For the trees, we modelled them as these really large 5 by 20 meter rectangles that act as shading for the house. So the trees are located on the south side of the house, and we modelled them as deciduous trees, so we set up a transmitting schedule so that they have a higher transmittance in the winter when the leaves have fallen and a low one in the summer. And for the case of scenario E in which the trees are cut, we modelled the trees as being five meters tall.

On the other hand, for the PV panel, we modeled it as being fixed on the roof. We actually had to completely change the model from BEopt, because they modeled it as being decoupled from the entire system, placed 30 meters away from the house. So we completely changed that, and we read extensively into EnergyPlus literature and found this particular object-- I guess you could call it-- called the integrated exterior vented cavity object.

And what this does is it models a surface as being, in our case, 0.5 meters away from the roof, and it models the convection and radiation between these two surfaces. We also considered the solar panel to have a solar absorbance of 0.92

and thermal emissivity of 0.9.

**STUDENT 3:**

And so we actually got a lot of results, as you might imagine, from all those different models, but just for the purposes of comparison for the presentation, we're just going to show you-- just summarize results. And basically what we're showing you here is specifically for Boston, and this figure that we're showing is the y-axis is the relative energy gain. So in order to compare them, we decided within each model compare it to a common situation. So we decided to say that, if we're in Boston, let's say we start with a white roof and a tree, so that's scenario C there, so that's why it's 0. So everything is in comparison to that.

And right off, we could see that, as we might expect, putting a PV on makes sense energetically. And specifically scenario D, which is where you completely remove the tree instead of just cutting it, in Boston at least, is what makes the most sense. And this is for both models-- actually, for all the models-- although model one can't really model a tree necessarily. So that's why it basically doesn't apply for that.

In terms of comparing the results between the models, model one actually does a pretty decent job in Boston in getting close to model three, which is impressive, because model one is significantly more crude, much simpler than model three, which required many, many inputs. And the reason we believe the discrepancy is between models two and three is that model two-- the way that we basically treated the solar insulation-- it doesn't treat diffuse sunlight differently, whereas basically the models PVWatts and EnergyPlus will take all that into account, so we think that's a larger reason for that.

And then similar thing in Phoenix. And basically we see a slightly different case here. Actually scenario E, where you just-- you install the PV, but you cut the tree instead of completely removing it-- gives you a slightly better increase in net energy gain. In terms of-- and this is, again, just in terms of energies. There could be rounding errors. This is ignoring the fact that, if you cut a tree down, the net greenhouse gas emissions would be changed and altered, and this is ignoring all those other effects.

This is just in terms of net energy gain of the house. And just to point out the discrepancy between model one here, we think that model one is actually, again-- that's using PVWatts to get your energy output. And that's basically assuming peak solar insolation, whereas our models basically use more empirical formulas to get the estimates for your PV output.

And then the last thing we did to do the sensitivity-- or to compare the models, was do a sort of sensitivity analysis. So basically on your y-axis, what you have is your percent change in your relative energy gain-- relative, again, to that scenario C-- divided by the percent change in your parameters. So we just considered four parameters. The x-axis is a rough estimate of how difficult it would be to actually change that parameter in your house.

So the far left one is the PV size. We just assumed if you went-- instead of a five kilowatt to a 5.5 kilowatt, so that's why the price is roughly \$3,000. We assumed about a \$5 to \$6 per watt installation cost for that. So that's basically the easiest one to do, and you get various significant change in your thermal energy gains because of that.

And then just because of time I'm going to rush through these, but basically a lot of the models follow similar trends in terms of the sensitivities. The last one is the roof pitch. You obviously wouldn't really want to change that. It's very expensive to do. Luckily, for most of the models, it's not actually that sensitive to it, at least within a close amount to where you start with. So in conclusion, in all three models, it makes sense to install a PV system. It's kind of what we expected from the beginning, but it's nice to get that sort of conclusion.

For models one and two, we actually were able to get pretty reasonable results, but they're limited in terms of what you can consider within those models. The advantage of looking at this basically is that, if you have a user that isn't as familiar with EnergyPlus in model three, which is very sophisticated-- it requires a lot of inputs-- they can still get a rough estimate, which is relatively close, using these much simpler models. And model three-- again, we were taking that to be the more

realistic case, but you'd have to compare it to real life data and do an empirical analysis to see how close it actually does correlate.

And then just, again, to summarize the results we got from model three-- in Boston, it makes sense to install the PV, but completely remove the tree. And your payback period is about 24 years. In Phoenix, the best scenario is to install the PV, cut the tree, and it's about 51 years. And interestingly, the maximum of that relative energy gain was essentially the same in both, even though the scenarios were different. So we'd just like to acknowledge Professor Buonassisi for helping assist us and guiding the direction of the project, and Bryan Urban at Fraunhofer and other members at Fraunhofer for giving us guidance. And with that, we would like to ask you for questions.

[APPLAUSE]

**AUDIENCE:** My question is I grew up in a neighborhood that has a lot of trees, and so cutting down all the trees wouldn't be very practical, but do you at all consider PV systems that could handle shading at different times of the day? So somehow decoupling different parts of it knowing that some of them will be in sunlight for part of the day, some of them will be shaded, and that will change?

**STUDENT 3:** So you could make the model-- especially in EnergyPlus, you could make it as complex as you want. We just did this for simplicity, just to put boundaries around what are problem is that we were considering, but you could definitely-- yeah, you could definitely add that to the model if you wanted to.

**STUDENT 1:** Cutting the tree was not worrying about the output of the actual cell in terms of whether or not some shading was going to bring down the rest of the cell. The reason we would cut the trees is to increase the maximum amount of sunlight per day hitting the panel. And it's because we were looking mostly at endpoints, trying to get the spectrum ends. That was why we went to such extremes. Cutting down selective trees and just parts of trees would be kind of in between. It's a little bit more difficult to assess.

**STUDENT 2:** Cutting down trees is actually [INAUDIBLE]. It's not [INAUDIBLE] because we found that the shading factor doesn't play a much bigger role if you look at the relative [INAUDIBLE]. So you would as well have increased the-- there will be objectively small loss in the energy gain, but that shouldn't matter much.

**AUDIENCE:** Sort of a philosophical question. If the payback period in Phoenix is 51 years, is it worth it? That's a long time period for-- I guess economically you could say that you could do other things with that capital instead that would have a shorter payback period.

**JORDAN:** Well, 51 years-- then the answer is probably not if you're looking at the benefit of cost money-wise. We did analysis about the energy. So we found an estimate for the embedded energy of the panel. This is from-- I can't remember the source, but they change quite a bit. But this example says it's 1,500 kilowatt hours per meter squared, so that equivalent to nine years payback. Pretty much nine years in Boston. I think it turned out to be eight years in Phoenix. So there is a benefit energy-wise, but in this example, perhaps not cost-wise-- perhaps not the most advantageous to do per dollar.

**MARK WINKLER:** A related question, actually. Can you back to your two slides back maybe?

**JORDAN:** Yeah.

**MARK WINKLER:** Your look at the net energy gain was quite similar. So why the large difference in yearly savings and payback period.

**STUDENT 3:** Just the cost of electricity in each location. We estimated it as about \$0.07 per kilowatt hour in Phoenix and \$0.17 in Boston for residential.

**STUDENT 1:** Which is why you would get a shorter payback period for Boston-- is because the cost of the electricity that you're [INAUDIBLE].

**MARK WINKLER:** So I would have assumed that the generation mix is sort of similar. Is that regulatory, or-- I would assume they're coal/gas centric generation mixes.

**STUDENT 3:** Yeah. You mean in terms of how the houses are--

**MARK WINKLER:** This is a little outside the scope of what you guys did. I was just curious if you guys had any sense of why the big difference in wholesale electricity price is between Boston--

**STUDENT 3:** I think part of it is just how plentiful energy is. I guess Boston is at the very end in the corner of the US. It's more difficult to get fuel, oil, gas shipped over here. I think Arizona-- I think they're relatively close to a nuclear power plant over there. Oil-- I think it's just location--wise.

**JOE SULLIVAN:** There's a lot of coal there. They actually ship a lot of the electricity to California because they can't [INAUDIBLE] in California. One quick question, though. The relative energy gains are the same for both. Do you have different sized panels, or is heating that much? That's a big deal.

**STUDENT 3:** Heating, yeah.

**JOE SULLIVAN:** OK. So if you--

**HEIDI:** Also for Phoenix. For Phoenix, you can see this is the cooling over here and heating on the right over there, and for Phoenix, you can just look at the values and see that there's a lot more cooling than heating compared to the Boston case, where there's a lot more heating by many orders of magnitude more. And so that kind of balances it out.

**JOE SULLIVAN:** And so that's only in a cutting down a tree case if you were already well shaded or not shaded at all?

**HEIDI:** These are actually all the curves for all the scenarios, and--

**JOE SULLIVAN:** I see. OK.

**HEIDI:** So it does matter just because of location. So if you're not cutting down a tree, then there's no decrease in shading. Or is it just the panel itself that's heating up more?

**STUDENT 3:** Well, actually if you go back to--

**JOE SULLIVAN:** Sorry, I think I missed something.

**STUDENT 3:** Actually, in Boston it actually makes more sense to have a black roof than a white roof. So you actually want-- shading isn't necessarily good in Boston, just because there's so much heating that you need in the winter. It seems to be the dominant effect in Boston, and in Phoenix, it's the opposite. The cooling is the dominant effect.

**PROFESSOR:** Did you consider the possibility that snow also insulates the house once it falls on the roof?

**STUDENT 3:** No. We did not. I don't know if-- is that built into-- I don't know what would happen. No, I don't think we did, but that's a good point.

**AUDIENCE:** Along those lines, do you have any intuition as to why in one case it's better to cut down a tree other than remove it, and then the other is better to remove it rather than just cut it down? Are you expecting it to grow back and then have to incur more costs because you're going to have to cut it down again, or-- what's going on?

**STUDENT 3:** Well, just in terms of pure energy, it was very, very slightly better in this case to have the tree just cut just in terms of the balance between heating and cooling [INAUDIBLE].

**STUDENT 1:** I may be able to help clarify that. The idea is in Boston we're relatively cold most of the year, so the more sunlight that hits your house is going to add more heat to your house, and that's less energy that you have to pay for. So the reason that it's beneficial to cut down the tree completely in Boston is because it allows more sunlight to hit your house, whereas in Phoenix, you don't want the sunlight to hit your house. If you cut down the tree completely, that's more you have to pay for AC in the summer. So the--

**AUDIENCE:** [INAUDIBLE] the difference between cutting down the tree and removing it?

**STUDENT 1:** So--

**STUDENT 3:** So-- go ahead.

**STUDENT 1:** Cutting the tree is assuming that you're going to maintain it at that certain level. So by cutting the tree, you keep a certain amount of shading on the lower part of your house, but you still allow sunlight to hit your solar panel. Cutting down the tree completely means there's no shading on your house at all.

**AUDIENCE:** I guess I have a philosophical question. So I think there are a lot of people-- motivation for the solar panels is not just the [INAUDIBLE], but rather the desire to do something good for the environment, and to lower carbon emissions, et cetera. But when you cut down trees, that increases your carbon emission because you're reducing the plant, that reduces your carbon output. So given that you won't have a tree there for like 50 years, does that offset the carbon emission gains that you get by--

**STUDENT 1:** Just my two cents. If you really want to go in depth, you can look at how much carbon is going to be produced by the coal power plant to give you the energy that you're going to be using for 50 years and compare that to the amount of carbon that one tree was going to save you, or you could ask yourself, am I planning to have a child during those 50 years, which will produce so much more CO<sub>2</sub> than that tree will take out? Either way, it's relatively a small value. However, we do acknowledge that there was a lot of philosophical questions that we argued amongst ourselves, but we realized we didn't have the time to try to evaluate, or the materials, and scope.

**PROFESSOR:** One more question, and then we're going to have to switch groups. Jasmin?

**JASMIN**  
**HOFSTETTER:** Do you have any real data to compare your model results to. From your results, it seems that it doesn't make any sense to install solar panels ins Phoenix. Is that right? That's the impression?

**STUDENT 3:** Financially. Just purely financially, yeah. In terms of the PV output, it seemed to be pretty close in comparison to what we got from other sources. So it seems like the

net gain in energy is roughly right, but obviously people still install panels there. So either, I'm guessing, subsidies, or larger installations, or something else, or just the desire to install it just for installing it-- not necessarily for financial reason-- in Phoenix.

**JASMIN** What was the temperature that you assumed? I suppose you assumed a constant  
**HOFSTETTER:** temperature in the house that was like the--

**STUDENT 3:** Yes, that would also change it.

**JASMIN** What was this temperature?  
**HOFSTETTER:**

**STUDENT 3:** The set points for our model was-- the cooling set point was 71.

**HEIDI:** 76. Yeah, 76. And then the heating set point was 71.

**STUDENT 3:** Yes.

**JASMIN** Can you say that again, please?  
**HOFSTETTER:**

**HEIDI:** The cooling set point was 76 degrees Fahrenheit, and the heating was 71.

**STUDENT 3:** For both locations.

**HEIDI:** For both locations.

**JASMIN** Thank you.  
**HOFSTETTER:**

**AUDIENCE:** [INAUDIBLE].

**JOE SULLIVAN:** Are we out of--

**PROFESSOR:** No, that's it. You guys are done. Congratulations.

[APPLAUSE]

[INAUDIBLE] coming up, PV grid. What happens when you install loads, and oodles, and oodles, and oodles of solar onto the grid? We're going to hear all about. And take it away. Knock it out of park, guys.

**IBRAHIM:** So as [? Tony ?] mentioned, we're the PV grid project. I'm Ibrahim.

**MARY:** I'm Mary.

**RITA:** I'm Rita.

**ASHLEY:** I'm Ashley.

**JARED:** I'm Jared.

**IBRAHIM:** All right, so I'm just going to start with the motivation behind our project. So as we discussed in class, PV installations have witnessed very significant growth rates over the last few years. Last year alone PV installation growth rates were around 17%. Around 18 gigawatts globally were installed. As the cost of PV approaches grid parity, more investors and consumers are going to want to adopt PV systems.

However, one lingering or major obstacle preventing the further or high penetration levels of PV systems is intermittency. So as we discussed in class, there's variability in terms of the solar resource, both on a long-term scale and a short-term scale seconds to minutes. So on a long-term scale, we're talking about the position of the sun relative to the Earth and so on. So in that respect, that's predictable and can be planned for. When we define or talk about intermittency, it's the short-term unpredictable effects that change the power output significantly.

So what we have here is the fractional change in power output over the course of one day. So as you can see, between the two consecutive seconds, the power output can almost double, and it can at other times drop by half. So from a system operator perspective, that's obviously a major challenge because demand should match supply at all times. So again, these effects, or these intermittency issues, arise due to regional weather patterns that can be predicted and also due to local weather patterns that are less predictable.

So in our project, what we tried to address is, can the weather report be used to predict the power output from an ensemble of smaller distributive PV systems? That is, can we average out these local less predictable intermittency effects? I'll give it to Mary to discuss our approach.

**MARY:**

So our goal of this project was to design a model that could quantitatively analyze a PV grid and determine its robustness in terms of variability. And our main components were meantime between failure-- which is the average time between two system failures, which Rita will define and discuss later-- number of systems in the grid, and geographic dispersion, which we measured through geometric mean distance.

Our data set was from the Oahu airport, which is part of the National Renewable Energy Laboratory. There are 17 systems all within about a kilometer of each other, so it's a very small, very dense system, but there was second interval data for a year, which we used. So there's a fair amount of data to give us an estimate of how intermittency varies over the course of a system and the number of systems and density.

**RITA:**

So our first step was to define what was a PV system failure. In order to do so, we accessed the CAISO website-- that is, the California Independent System Operator-- and we took that data from one week of the actual demand and hour hand demand forecast. They give this value for every hour, so we took the value for every hour of the week, and then we plotted in this graph that we have a line for each day of the week, and can see that both the magnitudes and the shapes throughout the week are almost the same.

We can also see that the values are almost all positive. This means that they usually underestimate. They usually think that the demand is going to be under what it really happens. And so what we defined was that, if this estimation is OK for CAISO, if they can manage that the grid with this variation, then they could also manage the grid with this variation in a PV output.

And so we looked at 5:00 PM. That is the hour that we have the biggest variation between the two, and we averaged the value, and we got to 6%. So this means that, if our intermittency is above 6%, we are going to have a PV system failure. If the variation is below 6%, then the intermittency is not going to be a failure. Then we could define the mean time between failures-- that is, the mean time between two intermittencies higher than 6%.

**JARED:**

OK, now that we have some context of what the problem is, and we have an idea of what variability is, and we have a data set to work with, I'm going to talk about how we actually solve the problem. We use coding in MATLAB to handle this huge data set. NREL had 17 systems out there for every second of an entire year. And so we took all the files from NREL and put them all into one huge matrix. You can imagine it was-- it ended up being about 23 fields by several million, and it's about 677 megabytes.

So actually handling the data was an issue in itself. I don't recommend it with an old computer. And we also the GPS coordinates for each of those locations and the variability from the California ISO, so with that data we could begin to build our code to figure out a quantitative description of mean time between failure and our idea of density.

So once everything was loaded into one big matrix that we could work with, we moved on to use the GPS coordinates. And of those 17 systems, we found every single combination of 17 choose 2, 17 choose 3-- every possible way that you could connect these systems-- and came up with something like 60,000 different ways of connecting these, and then for each possible connection, we had a function that would calculate this geometric mean distance that would give you an idea of the density of that particular connection.

And so to compare our mean time between failure for these systems while holding the density constant, we then searched through those possible combinations and found this magic number that kind of existed for each of those possible combinations of two, of three, or four, all the way through. And it kind of lined up for

geometric mean distance of 400 meters. So using that set, we could then go on and see how increasing the number of systems helped the mean time between failure. And then for a given set, we ended up using eight. 17 choose 8 gave us like 24,000 possible ways to connect them.

We searched through and found varying densities for one set number. Then finally we wrote a function that calculated the mean time between failure that went through our data from NREL and said-- looked at the fractional difference and said, OK, each time it's about 6%, that's a failure, and then measured that distance, took the average of that, and that was our mean time between failure. And then finally, once we had all that together, we crunched all the numbers, took a long time on my computer. We were able to plot it together and get some very nice trends.

One of the great things, I think, about our code is that it was only 525 lines, and if you've ever built a programmer, a big application, that's really small. It's very easy just to go in and see exactly what's going on. So it's very flexible. We could hand it off to another company, to another research group, and they go in and adapt it to just about any data set. If you are able to get data in California, or from Germany, or from somewhere else, and bring it into our format, it's very easy just to plug it in and run the data. Very, very minimal changes within our code.

And then you could also build on our code to look at other problems. So we have the change in the-- we have the variability data as a function of time. We also have the solar output as a function of time. So you could conceivably go in and figure out how your meantime between failure changes based on the time of day and change your critical percentage based on the time of day, and there are several other problems that you could go, and launch off of our code, and continue on.

And if you're interested at all, I actually put the code of my public space. There's the link there. Check it out. It's pretty cool. And then Ashley is going to talk about our results.

**ASHLEY:**

Cool, so the first thing that we did in order to try to see the trends in these huge fields of data was just to plot the data. It was actually a much bigger task than I

thought it was going to be. The plot on the left is for one day's worth of data, and the plot on the right is one week's worth of data. The y-axis is power density in watts per square meter, and the x-axis is the time in seconds. The blue is all 17 of our systems together, and the red is just for one system.

So as you would assume, the power output for all 17 together is clearly a lot greater than the output from just one system, but this give us a sense of being able to see fluctuations within one day, and also were able to see when the sun rose, and peaked, and also fell each day.

And in order to quantify all those different fluctuations, we did the fractional change in power density versus time, once again, for one day, and then for one week. And red is the one system. Blue is all 17 systems together, and we can already see just from plotting the data that having all 17 systems together does start to average out the fluctuations of individual systems by a significant amount.

So then Rita earlier mentioned that we use 6% as our cut off for failure. We actually went ahead and did 6%, 12%, and 18% just to see how sensitive our analysis was to that threshold value. So here we have plotted on the left the meantime between failure versus the number of systems, and on the right, meantime between failure versus the geometric mean distance.

I also calculated these values for using a week's worth of data, a month's worth of data, and a year's worth of data. So the week would give you more fluctuations, but the year would give you the more long-term overall system behavior. Relationship between the mean time between failure and number of systems is quadratic, and we found a linear relationship between the mean time between failure and the GMD.

So this is four 6% cut off. This is for 12% cut off, and this is for 18% cut off. And the mean time between failure increases dramatically as you go from the 6% cut off to the 18% cut off. So a lot of this makes sense, but it was really cool to quantify that.

**RITA:**

So after applying those graphs we could take our conclusions and answer our

question. And so the first thing that we noticed, but we were expecting, is that a big data sample should be used if conclusions are going to be used as a design tool. As Ashley said, we used for a week, a month, and a year. And so we know that the bigger the data set, it's going to be-- it's not going to be influenced by abnormal things that can happen in a given day.

And we also saw that there is a linear relation between mean time between failure and GMD. When GMD increases-- that is, when density decreases-- we are going to have an increase in mean time between failure. This was also what we were expecting because the local effects will not affect systems that are further apart. We also saw that there was a quadratic relation between mean time between failure and number of systems. Number of systems increased. Mean time between failures also increased. This was also according to what we expected because we know that the percentage and the total output is going to be lower.

We also saw that the mean time between failure is very low, even when we can see the 17 systems together, we have about 900 seconds between failures. This means that some backup systems should be used in order to take over the load when we have a failure. And so now we're running conditions to ask our first question. And so we conclude that localized predictable intermittency do average out and that this effect decreases as the number of systems and the GMD increase.

The data that we used was for 17 systems, and the biggest distance between them was one kilometer. So we believe that it's important to run our code for a bigger set of data, because only in this way we can confirm our conclusions and guidelines for the design of PV systems can be defined. Thank you, and we'll be happy to answer your questions.

[APPLAUSE]

**JOE SULLIVAN:** So a couple things. First of all, you ended at exactly 15 minutes. I find that remarkable. Additionally, just-- sorry. Can you repeat what exactly a failure mode is defined as? Are you looking at 6% intermittency varying from second to second? So

if you look at the output from one second to the next, does that change by over 6%?

**RITA:** Mm-hm.

**JOE SULLIVAN:** It wasn't average out over an hour.

**RITA:** No, no. It was second by second.

**JOE SULLIVAN:** You got the 6% from Cal ISO.

**RITA:** We said that if there-- in a given hour, we measured-- let me just--

**JARED:** They only had an hour of data [INAUDIBLE].

**RITA:** Yeah, they only gave hourly data. So the difference between the actual demand and the hour-ahead demand forecast. So this is what they are expecting, but the difference between what they are expecting and what the grid is really asking them. So if they can manage this difference on a second base, they can also manage this difference on the PV grid.

**JOE SULLIVAN:** So you took the worst case. Is that how you got 6%?

**RITA:** Yeah, we took the average of the worst case. It's the 5:00 PM. The 5:00 PM is always the worst hour. It's always when they have that peak. And in fact, all of the base-- almost all of the base were around 6%. Our peak was like 6.8%, and we averaged, and it was 6%.

**MARK WINKLER:** So that's essentially their peaking capacity?

**RITA:** Yeah.

**ASHLEY:** Also, so I actually wrote down the numbers for 6% function, 18%-- like our mean time between failure. For 6%, we had up to 15 minutes between failure. So it's a pretty low amount of time between failures. And if you allow 12% as your intermittency, you can get up to about half a day. And then for 18% as your cut off, you get about nine days.

So it is still very intermittent, and you would pretty often have to have backup systems if you had the small of a system. So if you had a much wider spread system and a lot more systems in your grid, then you could definitely significantly increase the mean time between failure. Yeah, Joe?

**JOE SULLIVAN:** So you have this awesome graph. So if you go back to the time between failure number of systems. The interesting takeaway is how large of an area do you have average over, right? So 300 seconds on a grid perspective is unacceptable for widespread PV developed point. We need to be on the order of years. And so do you have an idea of what that distance is?

**ASHLEY:** If we just extrapolate it out?

**JOE SULLIVAN:** If you extrapolate-- this is obviously like we're taking the very, very edge of that function and then extrapolating [INAUDIBLE].

**ASHLEY:** So looking at the numbers--

**JOE SULLIVAN:** But it looks like it's going up exponentially, or do you have an idea of what that trend is?

**ASHLEY:** For 6% for the one year, it was almost exactly  $x$  squared. It was like  $x$  squared plus 50, or 100, or whatever that would be. So if you want, you could say, mean time how many seconds are in a year equals number squared. So the square root of however many number of seconds there are in a year would give you your number of systems required for a year between failures.

**IBRAHIM:** But this is for a given geometric mean distance, so you have two factors. If you sort of spread them out more, probably going to require less systems.

**JARED:** And if you looked that map, that's all at the end of a runway at the Honolulu airport. So if you have a huge field in Arizona, thousands of systems, your mean time between failure is going to be a lot better.

**MARK WINKLER:** So I'm actually really surprised that there's such a huge effect from adding systems, just because it seems as though the relevant length scales for weather should be

very large. Do you guys say anything about that?

**JARED:** I think the idea was that, for long-term weather, you can predict that. So if you know there's going to be a storm front coming through, you can add natural gas. You can add coal to the system. Back up--

**ASHLEY:** And that would cover the entire system.

**JARED:** Our kind of variability we're talking about is say, if one cloud goes over, or a flock of birds, or something. So we were thinking that would be on a few seconds for a single module for a cloud just go over shade it for a short distance. So if you add thousands of modules, the other modules wouldn't be shaded while that one is.

**MARK WINKLER:** But these fields-- I mean, 100 meters on the scale of cloud cover, this still seems like a somewhat small length scale. Let me rephrase the question. Do you think that the graph on the right would be a smooth function of distance, or do you think there's some length scale at which the behavior on that plot changes significantly?

**JARED:** That would be interesting if we could find--

**ASHLEY:** The assumption is definitely evenly dispersed in an area.

**JARED:** That would be something that, if we had another data set that had wider distances, it would be very easy to plug it in. I think our code's really flexible. It would show us that relationship.

**MARK WINKLER:** What do you guys think, though?

**JARED:** It's a good question.

**ASHLEY:** I wouldn't be surprised if it was linear still. I guess another complexity we could do would be you would have-- right now we just have one big field of systems, but if you had one set of systems that was spaced  $x$  distances apart, and then you had some number of kilometers away from another one space-- I'm not sure how exactly would model that, but I think that at that point I'm not sure what the curve would look like, but a continuing linear trend seems reasonable to me.

**IBRAHIM:** So I guess another thing to keep in mind is we did not take into account transmission costs, so I guess you'd have to weigh the cost of failure versus, I guess, the added incurred cost for transmission lines and so on, so there's sort of an optimum point where you want to space them and have a certain number of systems where I guess, after a certain point, your returns diminish and are not equal to, I guess, the cost of failure. So that's something where, I guess, future people can come in and expand on.

**AUDIENCE:** So all this is data from Hawaii, which has a very notable climate and weather. I've never been there, but--

[LAUGHTER]

Do you think that this is really-- your code is flexible, so I understand that, but do you think the conclusions are really extensible to other parts of the world with different weather patterns or climate?

**RITA:** That's why we think that the future mark is really to do it for a different place and for a bigger set of data because we really want to be sure that the conclusions are going to be applicable, because we had that same question. We were talking just about a small place. We said that it's one kilometer apart for the distance that we have. So we also want to run for a bigger set of data and for another place just to be sure that our conclusions are applicable everywhere.

**JARED:** And I would say the relationship would probably hold because if you have-- say, if your regional weather is very different, that wouldn't show up in the fractional second to second difference that we had. And so the timescale that we measured it on I think would be, say, small clouds or intermittent events that would occur over a wide range of different climates. The general regional weather is predictable, and it isn't investigated in our study at all. So I would say I think the relation would hold.

**ASHLEY:** I think that the big change between different regions would just be the total output power that you can get, but I think-- I wouldn't be surprised if the fluctuation is still the same or is similar. And I think certainly that, as you increase a number systems

and as you decrease the density with which they're packed, you should be able to have a more robust grid. I would be very surprised if that weren't the case.

**AUDIENCE:** Do you see shading for planes at the airport?

**ASHLEY:** There's no way for us to determine what causes the shading. The raw data we have is just output.

**JOE SULLIVAN:** Can you see how they move?

[LAUGHTER]

**ASHLEY:** It's like there is this line--

**IBRAHIM:** We actually did that for one plot. You could see the cloud moving around the plot.

**JOE SULLIVAN:** That's cool.

**IBRAHIM:** And you see the power output for [INAUDIBLE].

**ASHLEY:** Yeah, it was on the order of-- we had like two billion data points, I think, which was overwhelming. But yeah, it was really cool. Any more questions? Yeah?

**AUDIENCE:** Can you describe a little more what these PV system failures entail? And what happens, and how long does it take to get them back up and running? What has to be done to do that?

**ASHLEY:** You wanna get that one?

**JARED:** Sure. So basically there is a certain capacity that the grid would have. Say, you can compensate for a 6% drop in this case, or a 20% drop, or something like that. So if your system is completely powered by PV, which is not realistic, and you have, say, a 20% drop and nothing to compensate that, you have a blackout.

And so we investigated 18%, for instance. So that would be, say, if your grid is a certain percentage of PV and then has natural gas, or coal, or something that you can bring online quickly to compensate a drop in PV. That would be an idea of what

a failure is-- if you aren't able to compensate that fluctuation

**AUDIENCE:** And how long [INAUDIBLE]?

**JARED:** How long would a failure last? It depends. If you can't meet the demand--

**AUDIENCE:** [INAUDIBLE].

**JARED:** For a PV system, I think the problem is the PV system would come back up right after the cloud was over, but if you can't meet power demand, you've got all kind of protection systems that would trip off, and it would be mess. So I don't think would come back very quickly.

**ASHLEY:** That's a good question, though.

**JARED:** That's a good question.

**PROFESSOR:** It's relevant because you can envision back up power that could kick in really quick, but exhaust itself within the period of the delta t necessary.

**AUDIENCE:** For your definition of intermittency, did you look at the absolute value or just the drop? Because the grid can't deal with excess power as well, and so I was just wondering if you had insight on that. Like if you dumped 60% more power in the demand, there's no way for you to--

**JARED:** We did the absolute value. So 6% more, 6% less.

**JASMIN**  
**HOFSTETTER:** So I'm going to ask you for real data. So do you know where more or less data points would lie for, let's say, PV systems on houses that are like-- with a typical distance in some kind of neighborhood.

**JARED:** I think that would just be you would adjust your geometric mean distance to whatever the distance from the houses are. I don't think our data has to be a solar farm, for instance. I think it could be houses in a neighborhood, for instance. So if they're perfectly connected to the grid, I think that our code would account for that.

**ASHLEY:** This was for eight, right?

**JARED:** Uh-huh.

**ASHLEY:** The right-hand graphic held the number of systems at eight. And so if you had eight houses spread apart by an average of 150 meters, then you would-- and if you considered a year's worth of data-- is it like 250? I just can't see it. So you'd have meantime between failure of 250 seconds, which is four minutes? Doing math under pressure.

**JARED:** Right, but if you have a grid to back that up, it's not big of a deal.

**AUDIENCE:** I'm confused about the plot on the right here. What it's suggesting is that one week you picked was significantly below the year average [INAUDIBLE], and you could have equally picked another week that was significantly above.

**JARED:** Right. This was, I think, just to give the trend. The relationship between the day, and a week, and a year is just the day that we-- I'm sorry. A week, and a month, and a year is just the week we picked, the month we picked. I think you see on some of the other plots that the week and the month actually shift. It's just the year was kind of the average of those.

**MARK WINKLER:** I would assume that areas, or specifically countries, that made large investments in solar would have studied this question in a detailed fashion. Do you know if, for example, Germany or Spain have looked at this problem when it's spread across hundreds of kilometers.

**ASHLEY:** Ibrahim, do you know that one? I think you might be--

**IBRAHIM:** I was actually very-- we didn't find a lot of literature actually. For wind, there was a lot of data out there, I guess, because the high penetration levels with PV. There were very few studies. Most of them actually were addressing the US. I didn't find any, actually, on Germany or Spain. Probably maybe they're in Spanish or German, so I don't know.

**JOE SULLIVAN:** So what I find really startling is that, for a given system, the time between failure of the 6% intermittency is on the order of a minute. Do you have any the idea-- is that

vastly different for wind and what that number is? And this is outside of your-- I'm just wondering if in your literature searching.

**JARED:** You probably know the most about it.

**ASHLEY:** You would know from it.

**JOE SULLIVAN:** It seems like you have this big rotor. There's some momentum, and that to slow that thing down requires more time, but I don't-- as opposed to electrons.

**JARED:** I would say wind would definitely have a much longer time scale than solar, I think. There's a lot of momentum there.

**RITA:** But when wind stops, the times that you have intermittency is going to be much bigger. And there'd be backup systems you need to have to take over for a long period of time.

**JARED:** And maybe in high winds you would have more of an issue, because if the turbine is spinning too fast, you actually have to stop it. So maybe there you'd run into issues of variability on the order of minutes.

**AUDIENCE:** So I think-- and this is kind of going back to location data set-- comes from Hawaii, which I would imagine has mostly direct sunlight. For locations such as Boston, would the data set change for, say, diffuse light and would that generally bring in panels closer together or require more panels at the same geometric distance to get the same results?

**ASHLEY:** Well, the raw data that we have doesn't separate direct and diffuse, so I think that the first thing would be we'd want to look at a data set and from whatever other location you wanted to know about and look at how diffuse and direct differs. I don't think we have a sense here of that effect. Does that answer your question to some degree?

**AUDIENCE:** Some degree. I don't know if someone else wants to add more.

**JOE SULLIVAN:** [INAUDIBLE] after you respond.

**JARED:** I think it's interesting-- I was just thinking about this. Something that might be interesting to investigate is concentrated solar. If it's easier to shade, it would look like a denser system. So maybe that would be-- maybe a concentrated solar farm might be a bad idea if you have lots of little clouds. So that's something I think that you could expand into from this project.

**IBRAHIM:** And another thing, I guess, to add to your point, if you look at, I guess, solar thermal systems probably because of the diffuse sunlight, the intermittency I would expect is going to be probably less. You're going to have less, or the mean time to failure is going to be longer, so you could maybe add a solar thermal system, sort of balance the power output, and decrease your intermittency even further.

**JOE SULLIVAN:** Any last questions? No? All right, let's thank our group.

[APPLAUSE]