In this next section, we will examine estimating the emissions from a known location using measurement data. We're going to use the CAPTEX three hour measurements, those collected near the source to determine what emission rate would've been required from the Dayton, Ohio, release location to give us those values. So the unknown in this case is the emission rate, the known is the emission location. So let’s start by running the model for the base case with the actual emissions just to see what the results would be.

Let's do a reset to clear out the unwanted files and we will again load the CAPTEX surface configuration for the surface samples, and that was the CONTROL file that you had previously saved. It is also in the tutorial files directory. And the only difference, we can make that change and now. For here is the three hour samples, remember went through, started 18 on the 25th, and went through 12 UTC on the 26th, so we only need to do a calculation for 19 hours. And we need to retrieve the name list file for CAPTEX as well and then save. Now we can go ahead and run the model and this will give us the baseline performance for the model for those three hour samples only.

Now we will be doing the statistical analysis by going to the Convert to DATEM menu and we could select the three hour samples from this data file of all samples collected. And we called, these are the three hour samples versus six hour samples. The three hour samples were all collected close, relatively close in. To

avoid you having to edit this file to extract the three-hour samples, we have already done that for you in this file, the CAPTEX three hour sampling file. So we will now go to the Utilities, Convert to DATEM menu, and select the three hour measured data file, in the tutorial directory, that file.

We are using the picogram, convert gram to picogram, and we will create the DATEM file, and we can compute the statistics. And we have a correlation of .52 and the ratio of calculated to measured is almost unity. That is the model is not over- or under-predicting. And we can look at the scatter plot, and there are values around a factor of 10 over- and under-prediction, but there are many samples around the 1 to 1 line. So this is not an unusual type of model performance.

Now to do this calculation in the reverse, if you will, let's determine what source is required. We need to decide what we know, what we don't know. So we don't know the emission rate, we're going assume we don't know the emission rate, and we're going to assume we don't know when the emission occurred.

So we will rerun the model after making just one change. Let's go into the pollutant menu, and since we don't know the emission rate, we're going to assume a unit emission, and since we don't know the hours of emission, we're just going to assume it's emitting for the entire run duration. Now then, we started these calculations at 17 Z, because we know that that's when the emission started, but the meteorological file actually starts at 15 Z, so we should

actually do our calculations from 15 Z because we do not know when the emissions truly started. So instead of running for 19 hours, were going to run an additional two hours, which makes that 21.

And that means, we need to start here 15 and run for an additional two hours. What we're doing is, in the previous calculation, we showed that the model did not over- or under-predict. There was relatively no bias and we used the actual emission rate of 67,000. Had we run the model with an emission rate of let's say 30,000, we would've under-predicted concentrations by a factor of two. And this under-prediction or over-prediction ratio is what we will use to correct, to determine the correct emission rate. So effectively, the equation that we are solving here is that the downwind concentration and grams per meter, the measured concentration, or the model predicted concentration, is really equal to the dilution provided by the model or the atmosphere, in the real world, times emission rate.

And this can be rearranged so that we can solve for the emission rate, where we divide the measured concentration by the dilution factor or the model prediction. As I was saying in this hypothetical example, where if we set the emission rate to 30,000 and we're under-predicting by a factor of two, then this ratio would come out to be two, it would be, the model predicted values would be half of the measured values, telling us that we need to correct the emission factor by a factor of two from the one that we had assumed. However by running the model with unit

emission, then the answer that comes out will be directly the emission rate. So let's go ahead and save and run model.

And let's do the statistics for the results, that is the Convert to DATEM and we're using the three hour data only, and we will leave the convert here to pg. And create this file and run the statistics. And you can see now, as you would expect, because we ran with a unit of emission rather than 67,000, we are under-predicting by quite an amount. So the average calculated as 0.07, the average measured is 2304. So using this equation, if we make this division, we get an answer on the order of 33,000 g per hour. So it's not actually, it's within a factor of two of what we know the true emissions to be, but you have to remember that this, because the way the flow was, it's really hard to tell from these data, when the emissions actually started and stopped. There's not enough information here to determine that because concentration was being contributed to those sampling locations for emissions, for model emissions that occurred prior to 17 and after 20 hours, when the true emission stopped. The emissions from the model were still contributing to the sampling locations, partly because of the three hour sample duration that we're using, and also partly because of the dispersion in the atmosphere. So there is uncertainty that will be added to the result, but it does bring us within reasonable limits of an answer.

We should save these configurations so that we can use it in other examples later on. So I would suggest that you

save as src_fwrd and we'll save the name list as well, src_fwrd_setup. The statistical program does generate multiple files. One of which is the individual samples, so these are the unit release calculations paired up with each sampler, each measured sample, and from these you can infer an emission rate from each one of these samples, by just dividing the calculated into the measured value. And we just looked at the mean overall result and you would see there'd be some, quite a bit of variability, in the sample, samples, and the estimates, the resulting estimates.

So in subsequent sections on the source attribution methods, we will discuss other techniques that you can use to time differentiate when the emissions start and stop.

And this concludes the forward calculation for source attribution for unknown emissions.