In the previous section we compared the forward and backward dispersion calculations under ideal conditions. In this section will expand the backward dispersion calculation to try to estimate a source location and a source strength using the measured data from multiple samplers.

We're going to start by using the CAPTEX configuration for the basic surface tracer release. That is you want to load this CONTROL file and namelist file that we have generated early on when we configured the simulation for CAPTEX. So go to the set up menu and … well, let's, before we retrieve, let's do a reset and now retrieve, and that will be the CONTROL file for CAPTEX. Remember that was a 68 hour duration, and the same for the name list, retrieving the setup file.

So for this example we are again going to use the three-hour measurement data which ends on the 26th at 12 UTC. Which means we only need to do a simulation for 19 hours and we are starting on the 25th at 17, when the tracer release occurred, from Dayton, Ohio. Now because we're only running for 19 hours, we do not really need to release 50,000 particles, and since you are going to be doing multiple simulations, one for each of the three hour samples, go ahead and reduce the particle number, and if you remember that is menu number four. So instead of 50000, let's just release 5000 particles for each simulation.

So the next step will be to do what we call a special run.

And this special run will be called, or is called geolocation. This menu essentially populates the variables in a TCL script that is run within the graphical user interface, where we first have to define all the measured input data files that we will be using. In this case as you remember, we have already created a subset of the measurement data for the three hour samples and that would be in the tutorial CATPEX directory, the captex2_3hr.txt. And we can open that. We will not be doing any conversions and we will leave the numerator flag set. And what's going to happen is that the script will examine each three hour sample and configure a CONTROL file to do a backwards simulation, using the information that was populated into the graphical user interface for the forward calculation, for the CAPTEX surface experiment, surface release. And it will place the measured data in the numerator of the emission rate field. So that is we're going to generate output fields that are proportional to the strength of the measured concentration.

If you go ahead and click the execute button, what it had done at this step is to create all the CONTROL files for the backward calculation. I'm going to click on continue and open up the working directory, and here you can see we've created 48 CONTROL files. In each CONTROL file, for instance, if I were to find, let me put them in numerical order here. Let's open up number 23. And you can see that it takes information from the forward run but it also takes information from the measured data file. So in this case, this particular sample, if I were to go to the measured data file, and look for 23. Now this is, let me,

too many windows here. Okay, so this happens to be station 318, so that's how it's identified in terms of the pollutant station 318. Now if look for station 318 in the measured data file, you can see it has values. Let's move this over a little bit. We need to find station 318 on the 26th at 12 and that would be the sample here I believe. So this goes from 9 to 12 on the 26th and, well sorry about that. What I'm trying to say is that's not the right sample, this is the right sample. This is the sample 318, here we go, so 318, this is 3556 from 0 to 0300. So this is the 3556 measurement, starting at the 26th at 0300, going backwards for three hours. That would take it to 00, so we go from 0300 to 00. And this is the release location, at that location 41, that gets translated here into the starting location in the CONTROL file. The end time of the sampling needs to correspond with the end time of all the three our sampling, so that all the samples are lined up. So we will be running from the 15th going back 19 hours, right. We started at 15Z, we go back 19 hours, the release starts at 03Z and continues to 00Z, the three hour release, and the calculations continue until it terminates on 17Z on the 25th.

And so there is this preprocessor program that was invoked here in the graphical user interface, that generated these CONTROL files. Now we're going to run the dispersion simulations next and this is just a script that goes through each CONTROL file and does those calculations. And as each one is completed, there is a message, some runs are faster than others, and if we go to the working directory, as this is running, you can see

the output files are being created here, the CONC.001 and so on.

Because of the fewer particles this particular simulation runs very quickly. And I should add that this information, the way this script runs is also shown in the batch file, or the shell script and this is exactly what we're doing here; configuring a forward simulation. And for instance, here is the preprocessor program that uses the three-hour measured data to generate the CONTROL files from the basic CAPTEX forward simulation CONTROL file, using the start and stop information, the meteorological data file, sampling duration and so on. All that information comes from the forward file. And that's merged with the data in the measurements to create the CONTROL file. Simply there is a script here that runs through each one of these CONTROL files that has a dot three digit suffix. And then there is a probability program that puts together all the three digit output files, the CONC.001 through 048, the last simulation that was done. And when it's all finished, you get the message calculations complete.

And we can now display the results. And that would be the last step 4 here, and the other options here permit you to put multiple time periods together, but we will look at each of the time periods individually. And this executes, as I mentioned the probability program, and what we are looking at here is the output from all the simulations. This is the mean output from the 48 simulations, and therefore the largest values are associated with the highest measured weighted concentrations. And for now we are

calling this as the source sensitivity function. It can have different names and can be defined different ways.

But let's just go on to the time period that's associated with the release. So this is the first three hour sampling period from 1800 to 2100, but it occurred just after the release started. And you can see that the highest weights of the backward calculation does correspond to the location near Dayton, the Dayton, Ohio, release location was approximately in this region right here: 40 N 84 W. So in this first step what we did was, we used the multiple measurement data, and a backward dispersion calculation to give us an area over which we might reasonably expect the source to have been. In this backward approach, where the source term was weighted by the measurements, zero measurements are given a very small nominal value, so that would tend to reduce the overall mean. So if you have a zero measurement going upwind, obviously a zero measurement means that there could not have been a source contributing to that station, so it needs to reduce the overall weight, the mean of the mean of upwind source estimates, when you merge all the measurements together.

Now there is a part two to the calculation. And the part two is that we select inverse as an option. What this does is so the model, when we are running the calculation, computes a dilution factor times a source term and it outputs then the calculated concentration. And what we did was, we used in lieu of a source term, we used instead a measured value, to give more weight to those

measurements that had the highest values. More weight to the upwind regions. But let's say we instead of putting in the measurement data, we put in the inverse of the measurement data, that is one over M, as a replacement for Q. Instead of D Q, we put, I should say D M, we put in D over M, or one over M rather, as the value for the emission, which means that the model will be outputting values of D over M.

Now if you were to take this simple equation and rearrange the terms, and we did this in an earlier section, right. We know that we can get an estimate of the emission, if we divide the measurement by the dilution factor, that is M over D. But if we are doing a calculation, where the output is going to be D over M, because we put in here one over M for Q, that is the output will be D over M, that's essentially equal to one over Q. So by putting in the inverse for the measured data value, the output of the dispersion calculations will be one over Q. We can't get the model to output Q directly because the model will always compute D, the dilution factor, in the numerator. So the only way to make this work is to put in one over M as the emission in the calculation, and then the model will output one over Q.

So what does that mean? We set this for inverse, we re-execute the CONTROL files. We've created new CONTROL files. Well I should say, before we do that, there is one thing that needs to be done, and that is we need to clean up a little bit, and because we created a bunch of files with a three digit suffix, so there is, and we

call those an ensemble by the way. We have ensemble cleanup, which gets rid of these three digit files, so let's do that first. And now we go ahead and re-compute the values, the CONTROL files that is, and if we look in the working directory, and if we are look back, for instance, at that same, we have actually fewer CONTROL files now, and these are now given by the inverse of the measurement. And I believe in this particular case station, CONTROL file 4, there are actually fewer of them now, is station 318, and but now we're putting in here very small number, which is the again the inverse of the of the measured value, on the 26 between the 26th zero and 03. So if you were to take that measured value, take the inverse, that's the number that's written in here. So the next step is simply to execute the simulations.

And also I should say there are fewer files in this approach because we're only looking at the nonzero measurements. Where is in the previous approach, the numerator approach, the nonzero, the zero measurements played an important role, whereas in this particular case they are not needed or they would not be useful. And then we can now execute, to see the final results, and this is the first time period, second, so on, and so this is the last time period, which corresponds with that first three hour sample after the release. And you can see that this particular approach is not necessarily useful for identifying a source location, but given what we learned in the first step, we know that the source is approximately in this region. And you can go to one of the utility programs like Convert to Station, at the release location and extract the exact value

here, and it's going to be something on the order of 3×10^{-17} and that would be of course in units of picograms, since the measurement data were picograms and if you were to take the inverse, you would get on the order of, an emission of about 30 kg, which is consistent with what we have learned from some of the previous steps.

Now the example that I showed you here is not particularly ideal and that the flow was southwest to northeast and fairly consistent for all the samples that were collected. Clearly if we had multiple events and that this was not a single three-hour emission, but we had a source that was a emitting for many days, under many different meteorological regimes, and we had samples collected in in many different downwind directions, putting together the backward calculations for something like that, we would have a much better triangulation in terms of identifying the source location, and even estimating the emissions. But the approach here was just to show you some of the methodologies that can be applied for using a backward dispersion calculation.

In subsequent sections we will look at some other approaches that we can use to refine the temporal source, can we identify when the source stops, and starts and stops.

For now this concludes our discussion of backward dispersion from multiple samplers.