In this section we will look at another method for solving the coefficient matrix. We will use cost function minimization. If you're continuing from the previous section, you should still have the nine output files, the nine simulations for each release hour from 15 UTC until 23 UTC on the 25th. If you do not have those in the working directory, that would be these output files here, these are the HYSPLIT output files for each release hour. If you do not have those, then you need to go back to the 12.7 section and generate those outputs. And those outputs in this section, we used to create the transfer coefficient matrix or the coefficient matrix that we then solved using singular value decomposition. And this was the matrix that was generated using those nine files, one file for each release hour, that his columns A through I, and then rows two through the end here, represent each sample collected at a different location and possibly a different time. And these then represent the dispersion values, the dilution factors generated by HYSPLIT for that release hour to this particular sample location or time, end time. And the last column represents the measured observations at those sampling locations, at those receptor locations.

In this section we are going to use a different approach to solve the coefficient matrix, we will use cost function minimization. And the cost function is represented by this equation, and what we are doing is trying to minimize this cost function, and minimize these two terms. The first term represents the unknown, that is Q, the emission rate at various times and locations, source locations, in this case we really have one source location. And we're minimizing the difference between the unknown emissions and a first guess emission, and this is divided by the error variance of the emissions. This value is of course, estimated. And this second term represents the difference between the HYSPLIT model prediction, that is a H here minus the observation at the measurement point, and the HYSPLIT model prediction is really just composed of the product of the source term and the transfer coefficient matrix value, or the coefficient matrix value, or the dilution factor, if you will, for that source and receptor location. More detail on this approach can be found in the referenced paper.

To solve this equation, what we're going to do is go into the utilities menu and into the transfer coefficient tab and instead of SVD, which we did in the last section, we will go to the cost function menu. And this happens when you open up the GUI without populating the menu, so quit, set up, and save, and now back to utilities, transfer coefficient, cost function. And we need to create the wild card for the HYSPLIT output file names that would be TCM0925, and we need to define the measured data, and as you know we are using the three-hour measurements only for these calculations. We are doing conversion to pg and we will output the coefficient matrix in this c2array comma delimited file. And we also need to define a first guess for the emissions, so the equation does require a first guess, and for the purposes of this calculation, we will select 30,000, which is half the rate that we know it should be. And what happens in the minimization process is if the errors are so large or a minimization cannot be achieved, then the solution would be the first guess solution for the emissions at that time.

So the first thing we need to do is create the file name, list of files, and that creates an INFILE and to confirm that this is correct, we could go to the working directory and check INFILE, and we can see all the release time files have been found. And step two is we create the transfer coefficient array, and we have 16 rows and nine columns and that can also be found in the working directory. This is the same file we were looking at before. And the next step is we need to create the input file for the solver, for the cost function minimization, and at this point we're going to leave everything here as the defaults, the only thing that you really need to enter is the first guess. You would need to read more about the approach to change these, but we do not make assumptions about the errors in the source, that is the guess uncertainty and also in the observations.

And we’re going to select solve, but first we need to create, which we forgot to do. Now we click on solve and if you recall the hours 15, 17, 18, and 19, should be the emission hours of about 67,000 and this cost function minimization program also did not do very well, even though we did include some measure of the uncertainties in the computations or in the observations and the emissions. So this still was not sufficient and for the very same reasons that the singular value decomposition solution give us unrealistic results, is that there are still too many errors in the model estimates, the dilution factors generated by the model for each emission time to the observation, or receptor locations.   And the same limitations apply and we should be really doing this with more data, more observations.

If we were to use the synthetic observations which were generated from the model calculation, so that the model in that situation would be perfect because it really, there are no errors in the model because the observational data were generated from the model. If we were to select those data instead, just like we did for the hypothetical in the singular value decomposition section, select these measurement data and create the new coefficient matrix file, and a new input file, and solve, we now get answers for those three release hours that give us a much better match.

So the approach that we've outlined in this section and the previous section, using transfer coefficient matrices, and solving for the unknown, the emissions is a viable approach, but you do need to edit the model predictions and measurements to ensure that the data you have have few errors or as few errors as possible, and the more data you have, the better the results, the more robust results would be. These two examples were really presented as an example of how to do these calculations. Unfortunately we had to rely on hypothetical data to give you some realistic looking results.

In any event this concludes our discussion on this cost function minimization and it really concludes section 12, in terms of source attribution approaches. The next thing we will do is as simple exercise.