

October 1, 2001

Listing A Complete Set of Kincaid SF6 Concentration Values
Subtitle: Why you should question your sanity for trying the above!

By

John S. Irwin

When I first thought about providing a complete listing of both the Development and Evaluation data from the 1980-1981 EPRI Kincaid experiments, I thought I was doing the 'right thing.' Now, I am not so sure. You want to be helpful to the community of modelers trying to evaluate their models, but you also want to provide some continuity with past efforts and schemes used.

I am finding that trying to allow use of the QA codes previously defined is problematic. I could do such, by publishing the Developmental SF6 data using the receptor assignment to arcs that was previously published. This can be done for both the Developmental and Evaluation data sets. However, if anyone tries to use the raw concentration values, they will soon find what I found out. See below.

1. Defining receptors to arcs:

You would think this would be rather simple. I have several choices: 1) use the designations in the archive; 2) use the files that list concentrations by arcs (KB-5.dat and KB-6.dat), or 3) devise my own scheme.

Use #1:

I first tried choice #1, using the receptor assignments in KB-50.dat and KB-51.dat. When I then tried to link up the QA codes that had been arrived at by Steve Hanna and Joe Chang, I found that 79 arcs did not even appear. After studying the problem, I noticed that the assignments in the archive were very coarse. Some receptors designated to be on the '3-km arc' were closer to 2km or 5km, and this was not an isolated incident. There were many such occasions.

So, I abandoned this approach.

Use#2:

This is easier to say than to do. I found that a receptor might be selected to be on the 2-km arc for one hour, and then placed on the 3-km arc for another. To implement this in code would require arc-by-arc instructions for each hour (painful). Furthermore, I found that many nonzero concentration values, that were never labeled as 'suspicious' (I assumed someone had assigned arcs as starting with "99" had a suspicious or redundant concentration values.), were dropped from consideration, and never reported in the concentration by arc listings in KB-5.dat and KB-6.dat. When I incorporated these '99' concentration values back on arcs they fell

near to, they seemed every bit in accord with the values already there. So, I could not confirm why these values were designated to arcs starting with '99.'

In a few cases, I might agree with dropping from consideration a value. Such cases were those where the receptors were nearly in identical positions (within 0.1km for one another), but with vastly different concentration values reported. My problem with this is who knows whether that is questionable. My view of plume transport and diffusion is that material ejected for the stack is caught in passing eddies (whirls) that have differing strengths, durations and periods of rotation. This view of plume transport and diffusion means that receptors nearly adjacent might 'see' very different concentration versus time results. In a not very good analogy, tornadoes can destroy a house and leave an adjacent house untouched. Why can not such a thing happen in plume transport and diffusion?

So, I abandoned this approach.

Use#3:

In this approach, I assign each receptor to each receptor arc using objective criteria. Such assignments are fixed, and never change from one hour to the next. The subroutine below is what I devised. Essentially, it defines the midpoint between each receptor arc (minus a fudge factor of 0.000001 km for what I call the 'top', which is the upper bound for each arc) in receptor assignments.

```
subroutine DefineArc(Dist4real,iarcout)
! Define receptor to arcs using objective criteria
! The routine below should perform the task
dimension Arcs(12)
data      Arcs/0.5,1.0,2.0,3.0,5.0,7.0,10.0,15.0,20.0,30.0,40.0,50.0/

iarcout = 0

do ia=1,12

if(ia.eq.1) then
bottom = 0.01
top     = 1.0 - 0.250001
elseif(ia.eq.12) then
bottom = 50.0 - 5.00
top     = 80.0
else
bottom = arcs(ia) - (arcs(ia)-arcs(ia-1))/2.0
top    = arcs(ia) + (arcs(ia+1)-arcs(ia))/2.0 - 0.000001
end if

if( Dist4real.gt.bottom .and. Dist4real.lt.top ) then
iarcout = ia
end if

end do

return
end
```

The result from using the above assignment procedure was encouraging. When I linked up the QA codes to the arcs I only had three arcs not identified, and these were easily seen to be arcs where receptors were moved to the 0.5km arc, when none were there to be found to begin with. Furthermore, for these three arcs, all the concentration values were zero, and the QA code was 0.

That was the good news. The bad news was that I found arcs labeled as having a QA code of 3, with only 0, 1, or 2 nonzero concentration values. Further investigation revealed that typically a QA code was assigned to arcs having at least 5 nonzero values, and these were rare. The more general trend was that in order to have a QA code of 3, you needed at least 8 nonzero concentration values.

Of the three treatments I tried for assigning receptors to arcs, this scheme was (in my opinion) the least problematic. I could explain how it happened. But I had to face the reality that I could no longer use the QA codes originally devised.

2. Defining a QA coding scheme:

The process of accessing whether a concentration value is 'suspicious' reflects one's own preconceived notions of plume transport and diffusion. Here are mine that I tried to incorporate into my assignment scheme:

1. When concentration values are small (say under 100 PPT), I might view differences between adjacent values as being significant only if they differ by quite a bit (say by a factor of 3 or 4), especially for arcs close to the stack.
2. When concentration values are large (say over 500 PPT), I might view differences between adjacent values as being significant if they differ by a factor of 2 or so.

When concentration values are generally small for all the values along an arc, that one value is 3x or 4x larger than the rest seems possible, especially if the maximum value is 8 PPT or 20 PPT. This consideration is behind my thoughts for conditions #1 and #2 listed above.

I think the same when looking at the maximum concentration from one arc to the next, for a given hour. I would judge the maximum on each arc, looking to see if similar values are happening on the arcs in front and behind.

Before I did any of the testing of ratio values, I required a minimum number of nonzero values along an arc. If the number of nonzero values was 2 or less, I assigned a QA code of 0. If the number of nonzero values was 3 to 4, I assigned a QA code of 1. With 5 or more nonzero values I started testing ratio values as described.

With these thoughts in mind I devised the following subroutine:

```
subroutine judgeRatio(ArcMax,ratio,jsuspect)

jsuspect = 1

do i=1,10

itop      = i*50
ibottom   = 50*(i-1)
test      = 3.5 - i*0.125

if( ArcMax.lt.itop .and. ArcMax.ge.ibottom .and. ratio.gt.test ) then
jsuspect = -1
go to 900
end if

end do

900  continue
     return
     end
```

The subroutine is used both to judge concentrations along an arc, and to judge concentration maxima from one arc to the next for a given hour. The logic will not satisfy all, but here it is.

1. To test concentration values along an arc, I define the ratio as the maxima on an arc divided by the 2nd high maxima on the same arc. You will notice that as the ArcMax increases, I require the ratio to be closer to a factor of 2. For ArcMax below 50 PPT, the ratio has to be greater than 3.5 to be labeled 'suspicious.'

2. If in testing the concentration values along an arc, I find that it looks like I have a suspicious value. I then call up all the concentration values for the hour, and look to see what the ratio is between the ArcMax for my current arc, to those one arc closer and further from the stack, than the current arc. I use the same routine to test the concentration ratios formed using the adjacent arc's maximum concentration values. There are some details in doing this as follows:

- a. For arcs at 0.5km there is no 'nearer arc', so we can only test for the next further out arc. Likewise, for arcs at 50.0 km there is no 'further out arc', so we can only test for the next nearer arc.

- b. If I have missing (or no) values at the next nearer arc, and valid data for the next further out arc, I only use the ratio test results from the next further out arc. The same is done, if I have missing (or no) values at the next further out arc, but good data at the next nearer arc.

- c. If I have good data for both the nearer arc and further arc, both have to say the value is 'suspicious', or I will accept the current ArcMax as being reasonable.

The above scheme was used. I found 10 arcs having an initial QA code of 3, which were then deemed suspicious when tested against the next nearest arcs. I found three arcs having an initial QA code of 2, which were then deemed suspicious

when tested against the next nearest arcs. For these 13 arcs, I set the initial QA code to a negative value, so you will see 10 -3 values, and 3 -2 values.

A summary of the QA codes is provided in the following table.

Summary of QA codes assigned using objective criteria outlined above		Hanna and Chang
QA Code of -3	10	
QA Code of -2	3	
QA Code of 0	1363	442
QA Code of 1	446	256
QA Code of 2	307	248
QA Code of 3	679	338
Total	2808	1284

I have shown to the right in the above table a summary of the QA codes that were arrived at by Steve Hanna and Joe Chang for the developmental data set. I think my QA assignment scheme assigns more zeros, fewer 1's and 2's, and about the same number of 3's. What I did to reach this conclusion was to take the ratio of 2808/1284 and multiply this to each of the Hanna and Chang QA summary results. The logic can be questioned, but in essence I am saying given all things being equal, what would we anticipate the number would be for each QA code given an increase in the number of arcs by a factor of 2.19.

Since my assignments of receptors to arcs differs from that used by Steve Hanna and Joe Chang, and since I included all nonzero concentration values on each arc, which Steve Hanna and Joe Chang did not do; we can not directly compare the assignment schemes.

When I look at the QA codes I have arrived at, they seem reasonable. Now I need some help from other sets of eyes and opinions.