Supporting Information

A MODEL USING LOCAL WEATHER DATA TO DETERMINE THE EFFECTIVE SAMPLING VOLUME FOR PCB CONGENERS COLLECTED ON PASSIVE AIR SAMPLERS

NICHOLAS J. HERKERT, ANDRES MARTINEZ*, AND KERI C. HORNBUCKLE *

Department of Civil and Environmental Engineering IIHR-Hydroscience and Engineering

The University of Iowa. Iowa City, IA 52242 USA

* Phone (319) 335-5148 (K.C.H), (319) 335-5647 (A.M.); FAX: (319) 335-566; e-mail: Keri-hornbuckle@uiowa.edu (K.C.H.), andres-martinez@uiowa.edu (A.M.).

Contents

Integrated Surface Database (ISD) Lite Dataset Processing Script

For hourly weather parameter, the Model uses the Integrated Surface Database (ISD) Lite database accessed through NOAA's National Centers for Environmental Information (NCEI) website.[1] A Matlab script, developed and run in Matlab version R2015a, for processing the Raw ISD Lite datasets is attached as a Matlab file (**process_isd_met_data.m)**.

PUF-PAS Effective Volume Model

The Matlab script, developed and run in Matlab version R2015a, for calculating PUF-PAS Effective Volume is attached as a Matlab file (**PUF_PAS_Effective_Volume_Model.m**)

PCB Physical Chemical Properties (CSV File)

The model requires the physical-chemical properties for all PCB congeners, such as molecular weight (MW), octanol-air partitioning coefficient (K_{oa}) at 25 °C, and internal energies of octanol–air transfer (dU_{oa}). These values are given in an accompanying CSV (comma separated values) file that is critical to use of the PUF-PAS Effective Volume Model (**PCB_PROPERTIES_MW_DU_KOA.csv**). Ensure this file is in the same workspace as the PUF-PAS Effective Volume Model script.

PCB LFER Descriptors (CSV File)

If the user decides to use the linear free energy relationship (LFER) to predict K_{PUF} for PCBs partitioning to polyurethane foam disks, the model will require the accompanying CSV file containing the descriptors for PCBs (**PCB_LFER_descriptors.csv**). Ensure this file is in the same workspace as the PUF-PAS Effective Volume Model script.

Chicago Sampling Locations

Table S1: Summary of site locations, longitude, latitude, and number of samples collected at that site.

Site ID	Sampling Site Name	Longitude	Latitude	N	Mean Concentration (pg m-3)	Standard Error
AU	Aurora	-88.329374	41.784717	9	74	1.25
CN	Channahon Park	-88.187542	41.468465	9	102	1.29
CP	Chase Park	-87.669303	41.967150	3	241	1.35
GE	Graves Elementary School	-87.805740	41.782596	6	289	1.34
HC	Harrison Crib	-87.572256	41.916117	1	998	N/A
***IT**	Illinois Institute of Technology (IIT)	-87.624700	41.834400	6	571	1.05
JP	Jefferson Park	-87.762946	41.968148	14	198	1.22
JT	Joliet Township	-88.124903	41.529934	12	2507	1.18
JN	Jardine Water Plant - Over Intake	-87.606188	41.8959050	12	1907	1.19
***JS**	Jardine Water Plant - Southeast Corner	-87.602717	41.89355	12	716	1.17
KC	Kane County Health Department	-88.329528	41.784623	2	230	1.18
***LM**	Lemont	-87.990570	41.668120	11	277	1.12
NC	Naperville City Hall	-88.153036	41.770978	11	98	1.28
NP	Norwood Park	-87.794134	41.986593	7	154	1.23
PP	Portage Park	-87.762161	41.955091	12	313	1.19
SP	Sauganash Park	-87.737399	41.988401	12	286	1.19
SL	Schiller Park	-87.865005	41.954700	10	215	1.17
VM	Village of McCook	-87.832527	41.800371	11	131	1.20
WH	Waukegan Harbor	-87.822736	42.363092	6	391	1.24
WP	Winnemac Park	-87.684438	41.974165	14	165	1.11

*Denotes sites with samples in the subset of 180 total samples specifically examined in this study.

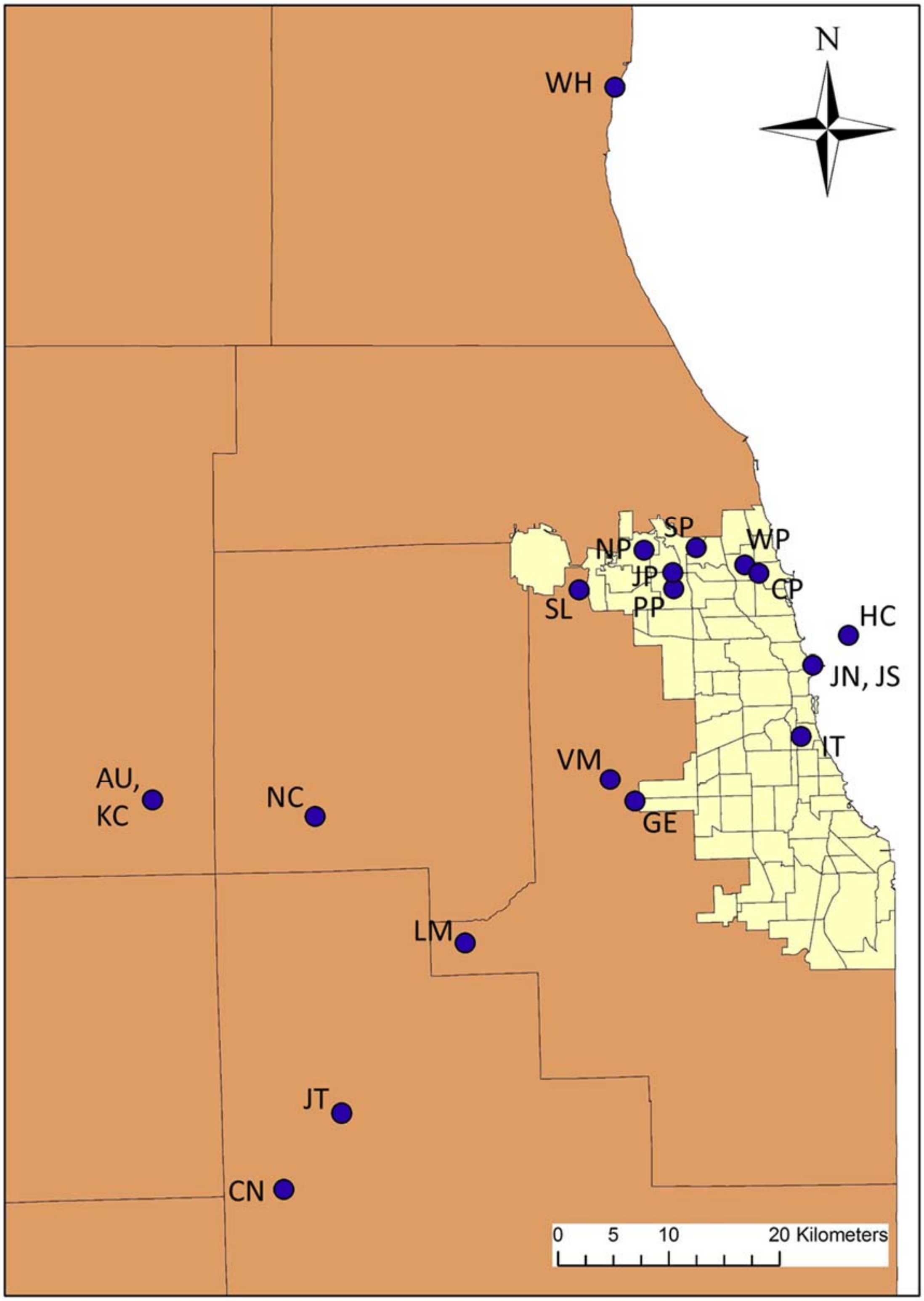


Figure S1: The 20 sampling sites used for the study in the Metropolitan Chicago area are denoted above on the map with their corresponding site IDs as noted in Table S1.

Sampler Design for Comparison Studies

Table S2: Description of PUF disk and sampler housing differences between the comparison studies.

			This Study	Melymuk et al (2011)[2]	Chaemfa et al (2008)[3] -Lancaster Design
PUF Disk	PUF Disk Manufacturer		Tisch Environmental	PacWill Environmental	Klaus Ziemer GmbH
	Length (cm)		14	14	14
	Thickness (cm)		1.35	1.35	1.2
	Density (g cm^{-3})		0.0213	0.0213	0.035
	Surface Area (cm^2)		365	365	365
Housing Design	Top	Diameter (cm)	24.0	23.0	26.0
		Depth (cm)	8.0	8.0	10.5
	Bottom	Diameter (cm)	20.0	19.0	19.5
		Depth (cm)	7.0	7.0	7.0
	Gap depth (cm)		1.5	0.75	1.5
	Gap width (cm)		2.0	2.0	2.5

Effects of Temperature Changes

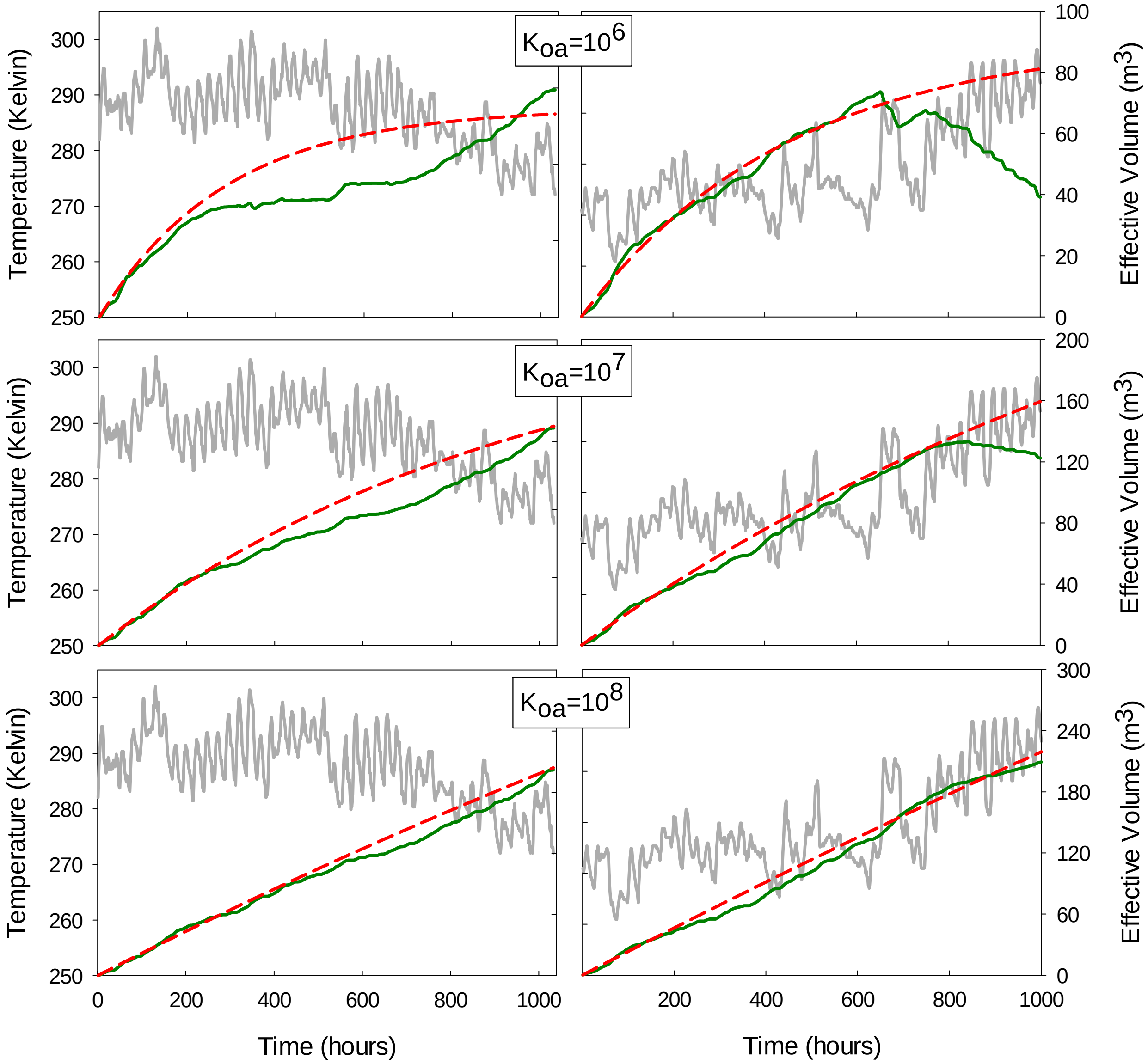


Figure S2: Example effective sampling volume curves for deployments with decreasing temperature (left) and increasing temperature (right). The solid grey line represents hourly air temperature. The solid green line represents the effective sampling volume calculated using the method of this study, including hourly adjustments on k_v and K_{PUF}. The dashed red line, represent effective sampling volume calculated using average values for k_v and K_{PUF}.

Long Deployment Periods

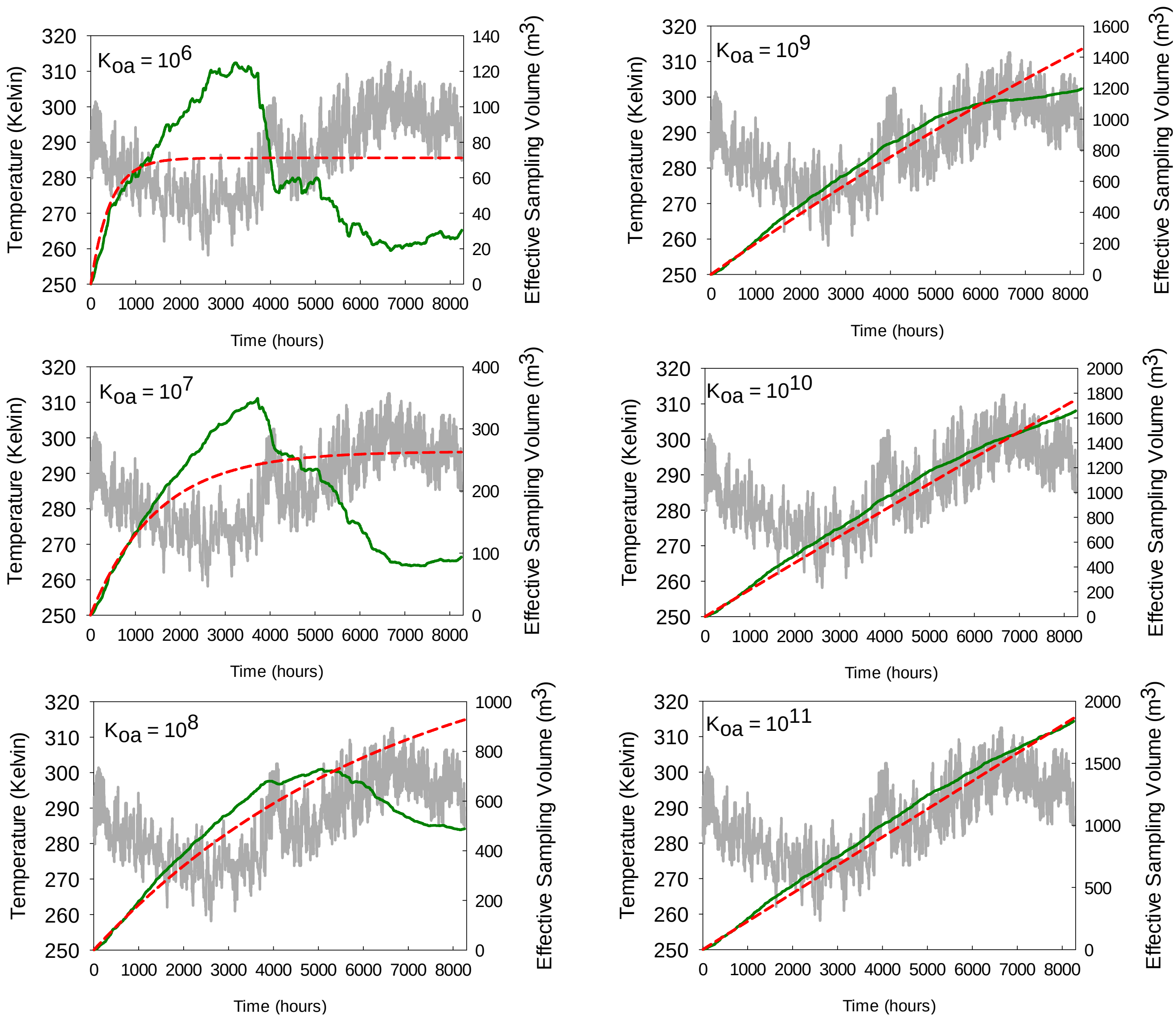


Figure S3: One sample involved in our study was deployed for 344 days from October 3th, 2011 to November 11th, 2012. The solid grey line represents hourly air temperature. The solid green line represents the effective sampling volume calculated using the method of this study, including hourly adjustments on k_v and K_{PUF}. The dashed red line, represent effective sampling volume calculated using average values for k_v and K_{PUF}.

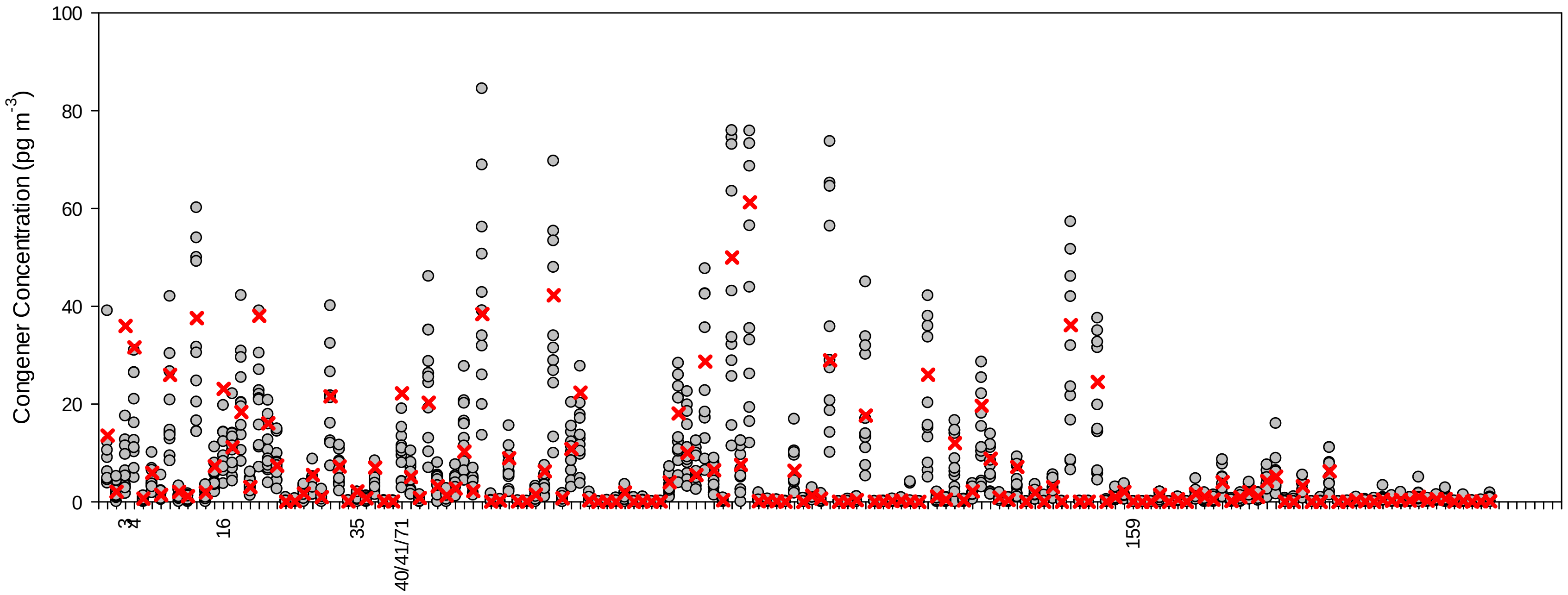


Figure S4: Profile comparison of long deployment sample (denoted with red x) at Jardine Water Plant - Southeast Corner (JS) site and all other samples (grey circles) for JS site. The six PCB congeners (3, 4, 16, 35, 40+41+71, and 159) that fell out of the range of what was observed at the JS site in all other samples collected at the JS site, are denoted on the x-axis.

GAPS Template Comparison

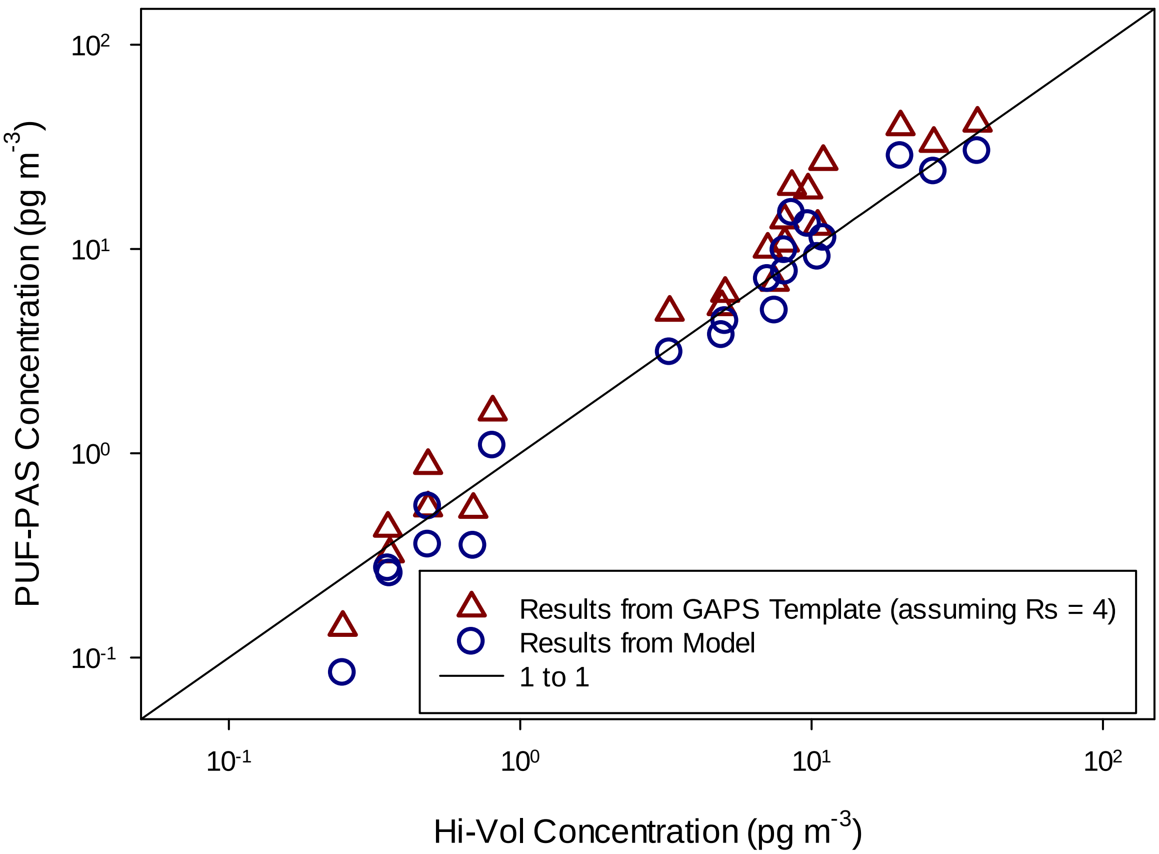


Figure S5: The model results for the Chicago comparison was also compared with results obtained using the commonly used GAPS template.[4] When using the GAPS template it was assumed that the sampling rate was 4 (the default) and particle phase sampling rate as a fraction of gas-phase rate was 1 (i.e. equivalent). From only a subset of congeners compatible between the datasets, the method ratio using the GAPS template calculation was 0.75, while the method rate using the model was 1.09.

K_{PUF} Calculation Method Comparison

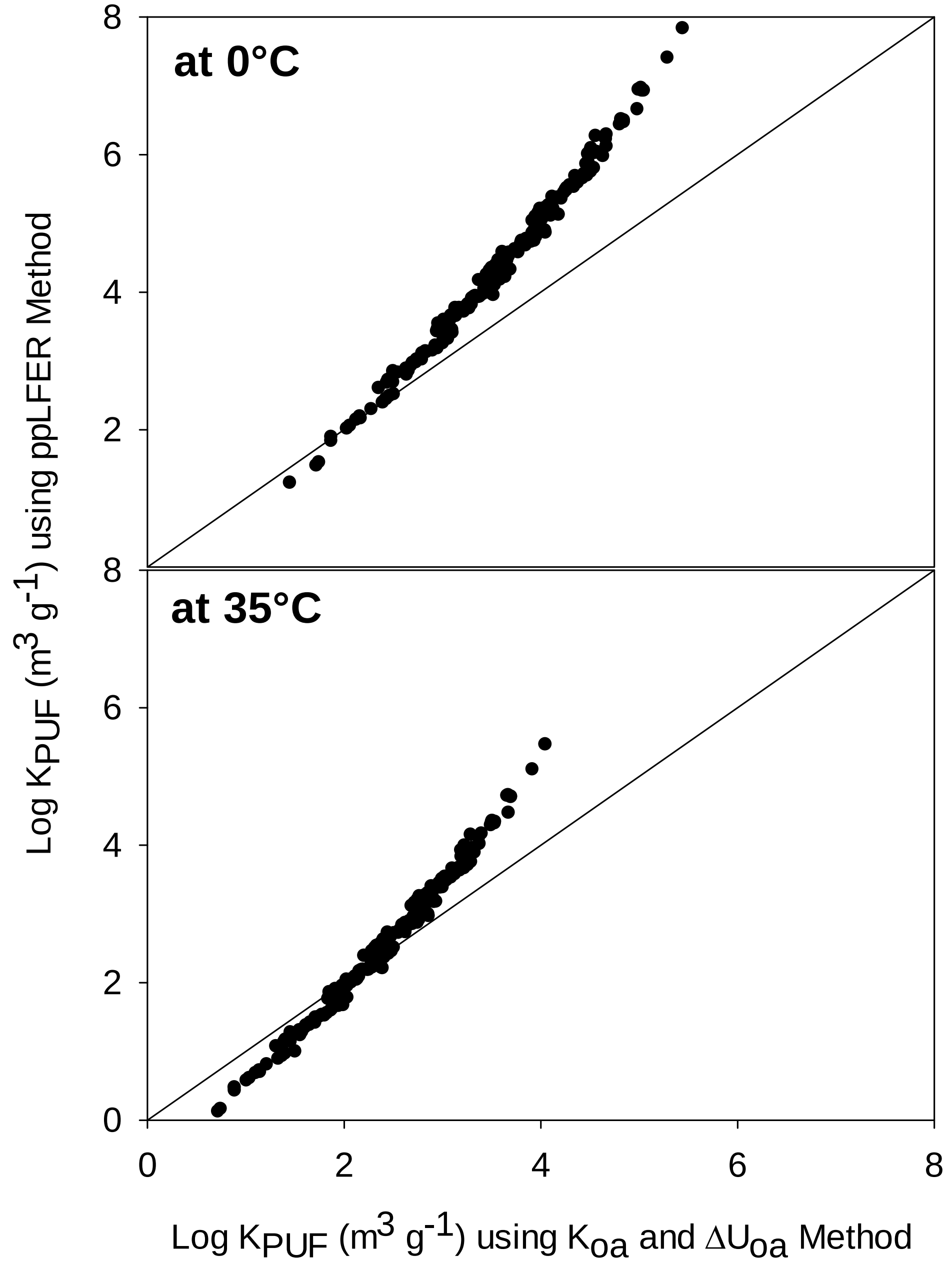


Figure S6: Comparison of the K_{PUF} values for all 209 PCB congeners using the temperature dependent LFER method for polyurethane foam given be Sprunger et al.[5] modified from Kamprad and Goss.[6] and empirical relationship with K_{oa} proposed by Shoeib and Harner.[7]

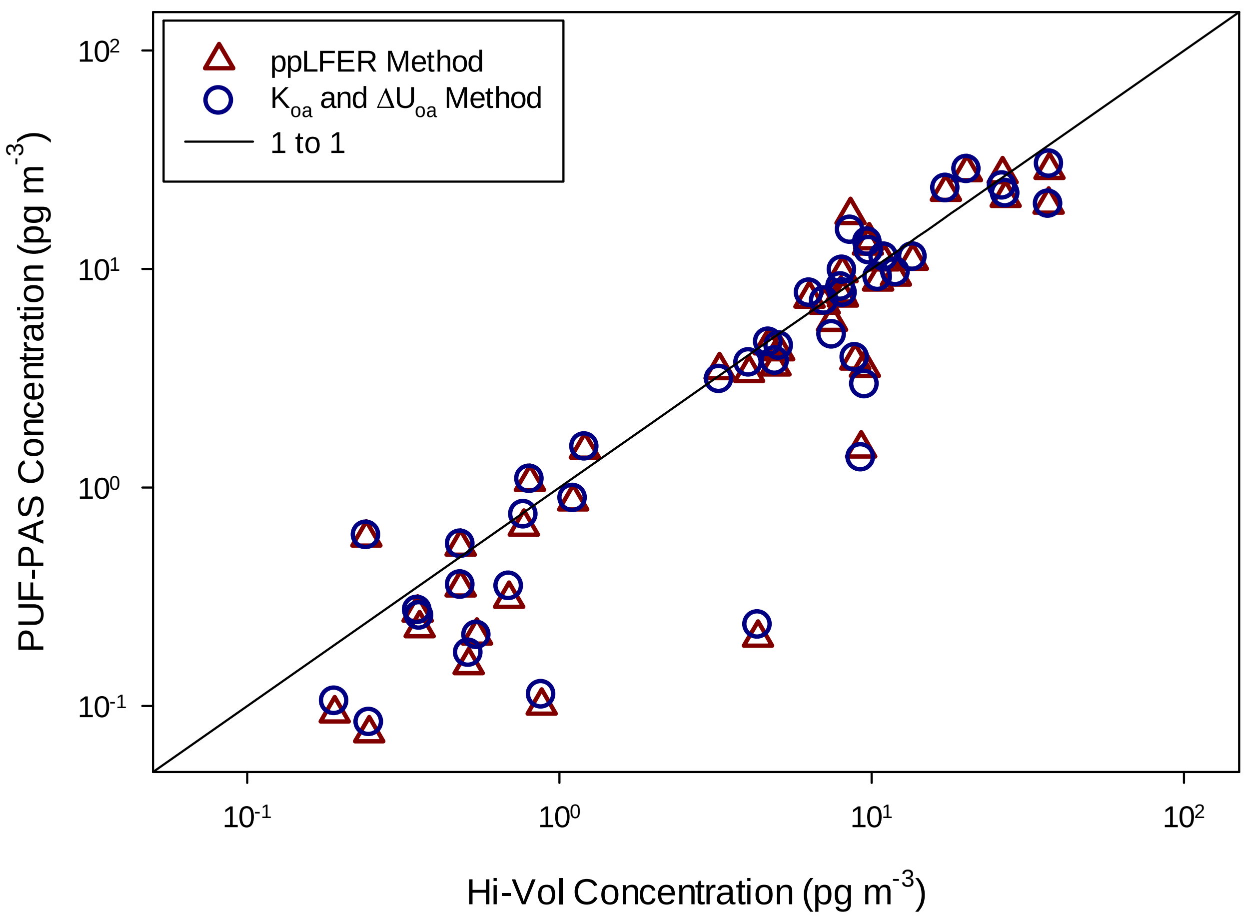


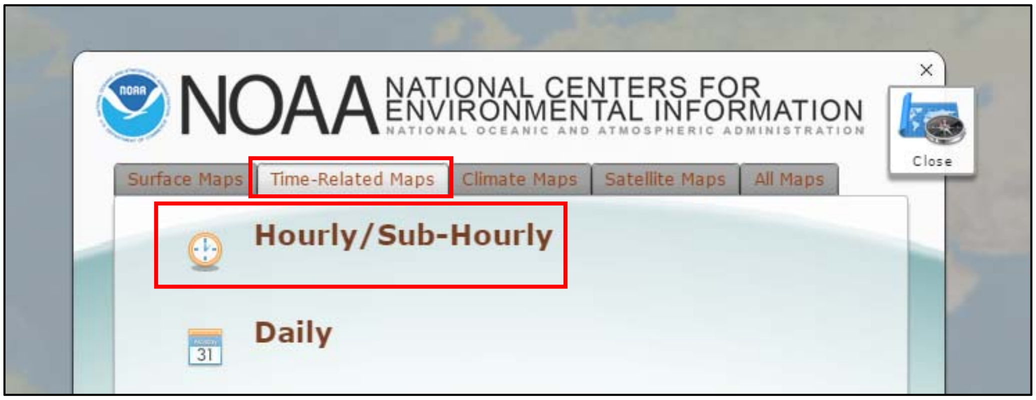
Figure S7: The model results for the Chicago comparison were recalculated using the temperature dependent LFER for polyurethane foam given be Sprunger et al.[5] modified from Kamprad and Goss.[6] It was found the average method ratio for the Chicago comparison using the LFER determined K_{PUF} was 1.63, compared to 1.59 for the comparison using the temperature corrected K_{PUF} determined by the empirical relationship proposed by Shoeib and Harner.[7]

Steps to Run Effective Volume Model:

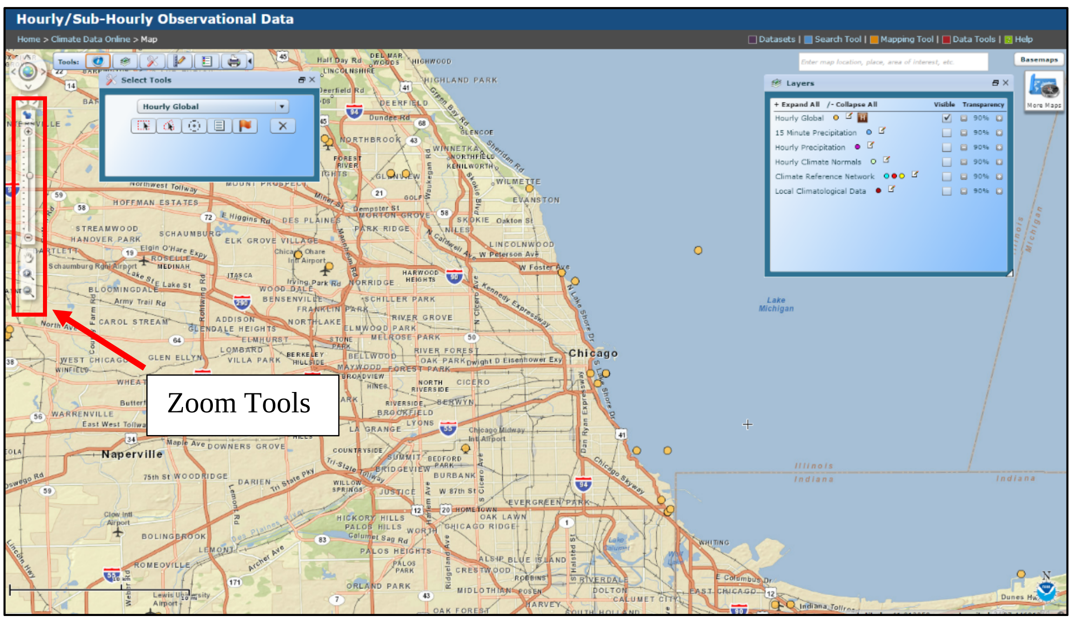
1. Identify the Site ID of the NOAA Integrated Surface Dataset that corresponds to the spatial location of your study using the Map Tool provided on the NOAA website

    a. Go to → https://gis.ncdc.noaa.gov/map/viewer/

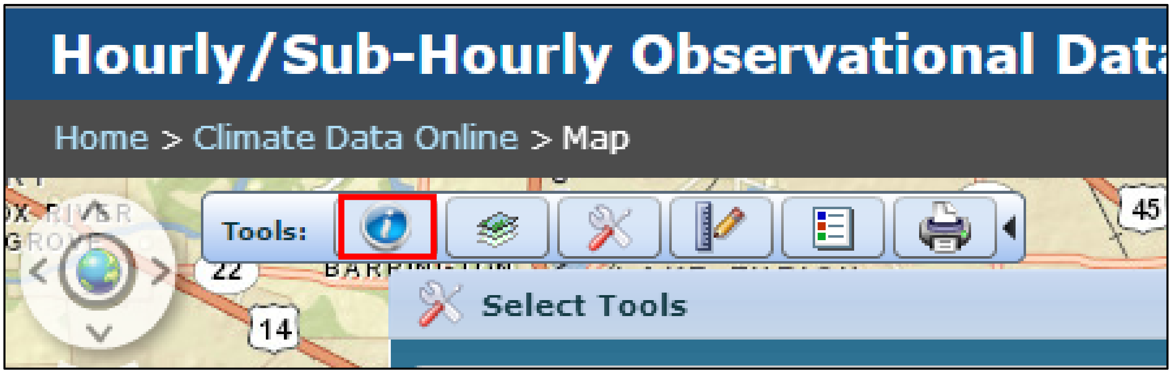
    b. Click on Hourly/Sub-Hourly under the Time-Related Maps tab on the map GUI



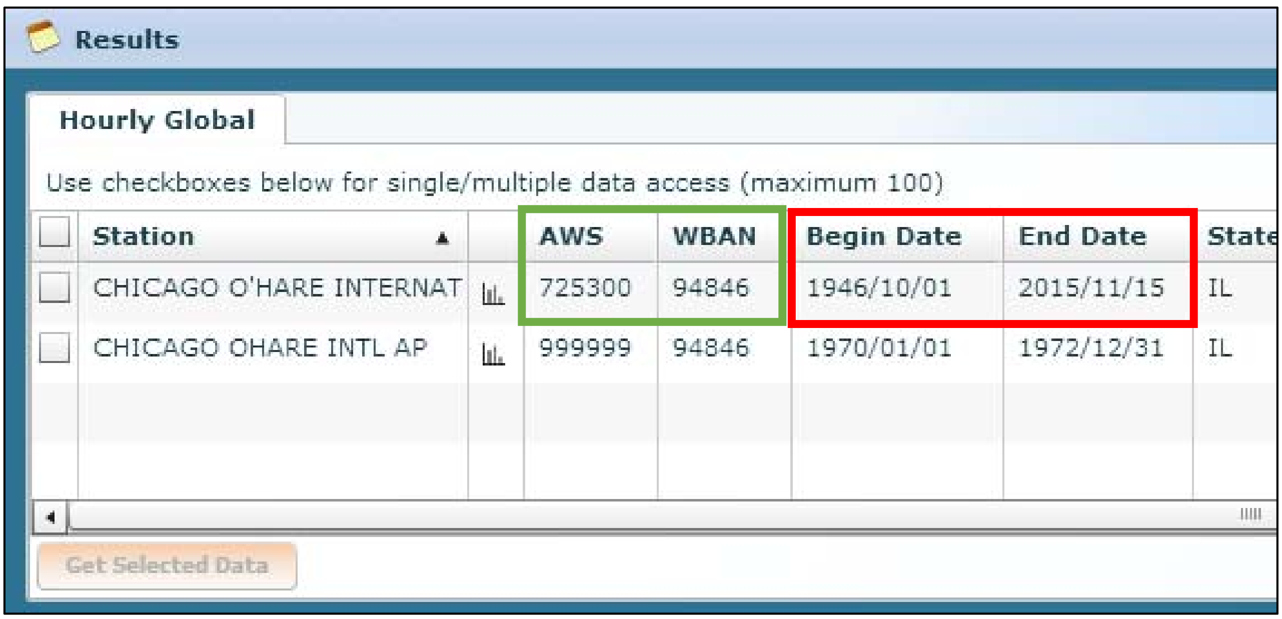
    c. Use the Map to zoom into the spatial location of your sampling location



d. Click on the information curser option in the upper left hand corner of the screen



e. Select a weather station that is spatially close to your sampling location and ensure the weather data spans the appropriate time to encapsulate your study. (From our experience the best data quality typically comes from Airport data and we recommend selecting airport data where possible.)

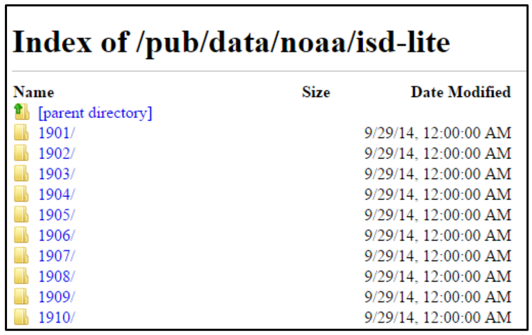


f. If the weather station spans the appropriate time span, record the AWS and WBAN number, for future reference (from above example 725300-94846)

2. Obtain the Raw NOAA ISD Lite file for the given year/years of your study and download them into your MATLAB workspace

   a. Go to → ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-lite

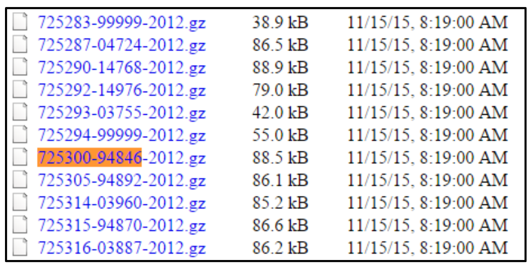
   b. Select the appropriate year for your study



Index of /pub/data/noaa/isd-lite

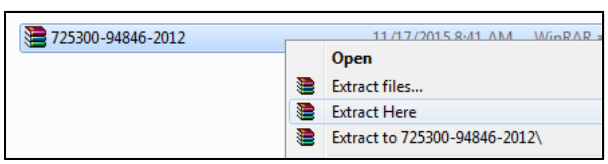
Name	Size	Date Modified
[parent directory]		
1901/		9/29/14, 12:00:00 AM
1902/		9/29/14, 12:00:00 AM
1903/		9/29/14, 12:00:00 AM
1904/		9/29/14, 12:00:00 AM
1905/		9/29/14, 12:00:00 AM
1906/		9/29/14, 12:00:00 AM
1907/		9/29/14, 12:00:00 AM
1908/		9/29/14, 12:00:00 AM
1909/		9/29/14, 12:00:00 AM
1910/		9/29/14, 12:00:00 AM

   c. Location the weather station ID previously recorded via the map tool (We recommend using the search function control-f)



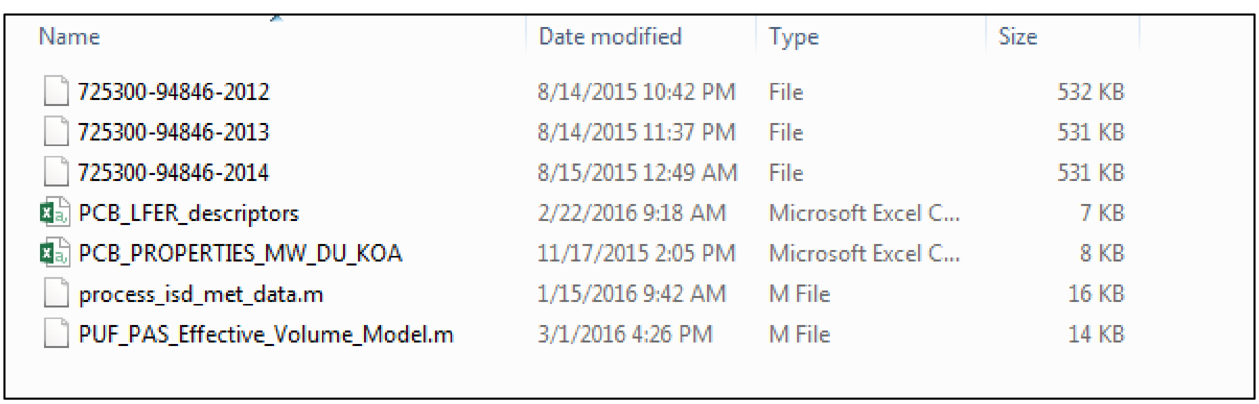
725283-99999-2012.gz	38.9 kB	11/15/15, 8:19:00 AM
725287-04724-2012.gz	86.5 kB	11/15/15, 8:19:00 AM
725290-14768-2012.gz	88.9 kB	11/15/15, 8:19:00 AM
725292-14976-2012.gz	79.0 kB	11/15/15, 8:19:00 AM
725293-03755-2012.gz	42.0 kB	11/15/15, 8:19:00 AM
725294-99999-2012.gz	55.0 kB	11/15/15, 8:19:00 AM
725300-94846-2012.gz	88.5 kB	11/15/15, 8:19:00 AM
725305-94892-2012.gz	86.1 kB	11/15/15, 8:19:00 AM
725314-03960-2012.gz	85.2 kB	11/15/15, 8:19:00 AM
725315-94870-2012.gz	86.6 kB	11/15/15, 8:19:00 AM
725316-03887-2012.gz	86.2 kB	11/15/15, 8:19:00 AM

   d. Download the appropriate file and extract the zip file.



   e. Move the extracted file to the folder containing the **process_isd_met_data.m** script

   f. Repeat Steps B through E for each required year of meteorological data (if your study spans more than one year)



Name	Date modified	Type	Size
725300-94846-2012	8/14/2015 10:42 PM	File	532 KB
725300-94846-2013	8/14/2015 11:37 PM	File	531 KB
725300-94846-2014	8/15/2015 12:49 AM	File	531 KB
PCB_LFER_descriptors	2/22/2016 9:18 AM	Microsoft Excel C...	7 KB
PCB_PROPERTIES_MW_DU_KOA	11/17/2015 2:05 PM	Microsoft Excel C...	8 KB
process_isd_met_data.m	1/15/2016 9:42 AM	M File	16 KB
PUF_PAS_Effective_Volume_Model.m	3/1/2016 4:26 PM	M File	14 KB

3. Check the data quality and convert the meteorological data to the required format of each individual year using the **process_isd_met_data.m** script.

   a. The input should be in the form of

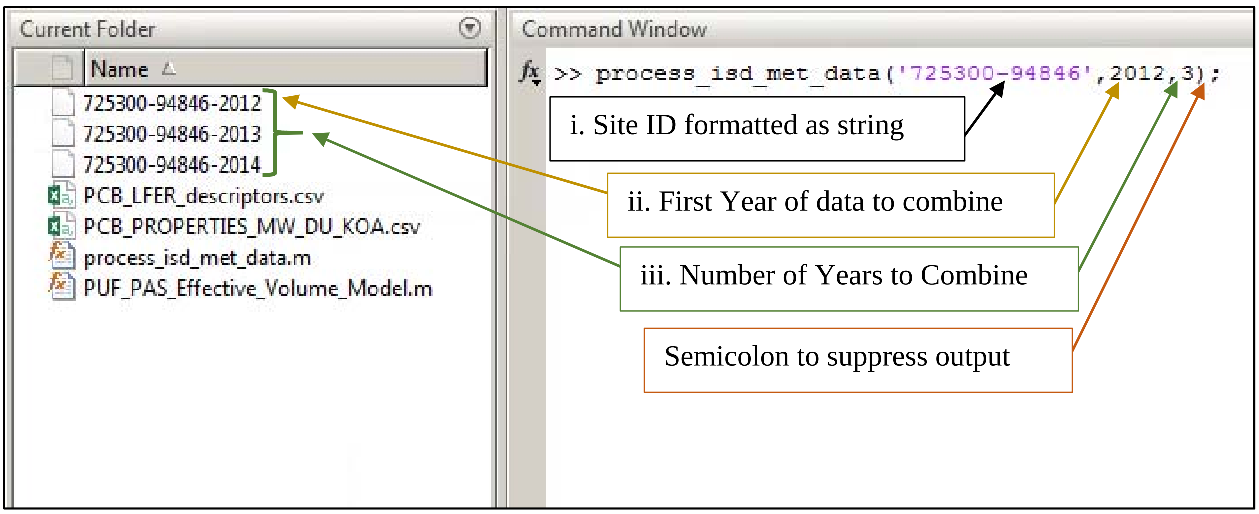
      process_isd_met_data (‘XXXXXX-XXXXX’,iYear, NumberYears);

      where,

      i. ‘XXXXXX-XXXXX’ = the site ID to be combine formatted as a string (i.e. enclosed in apostrophes ‘ ’).

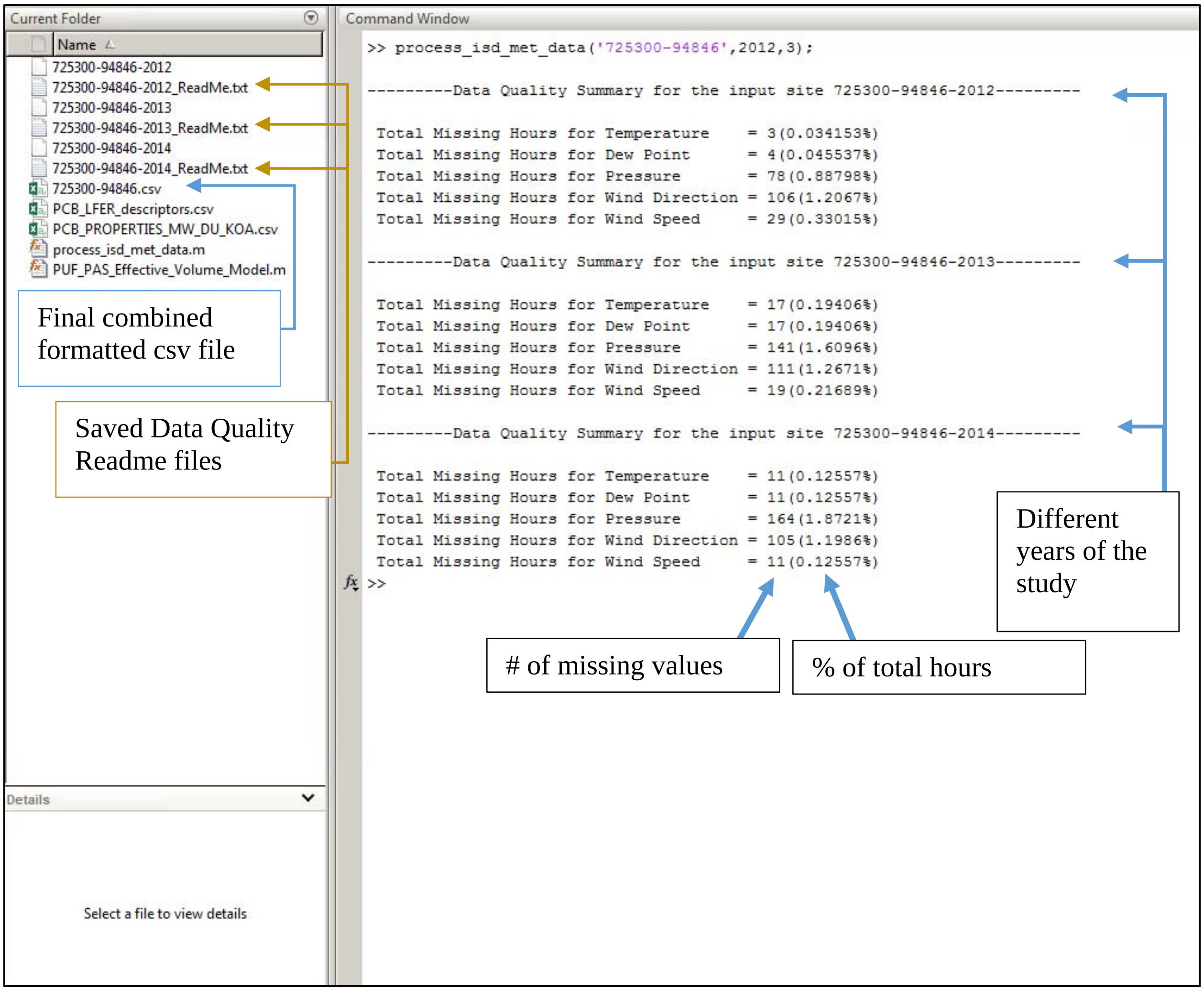
      ii. iYear = the first year of the data to be combined

      iii. NumberYears = the number of years to be combined



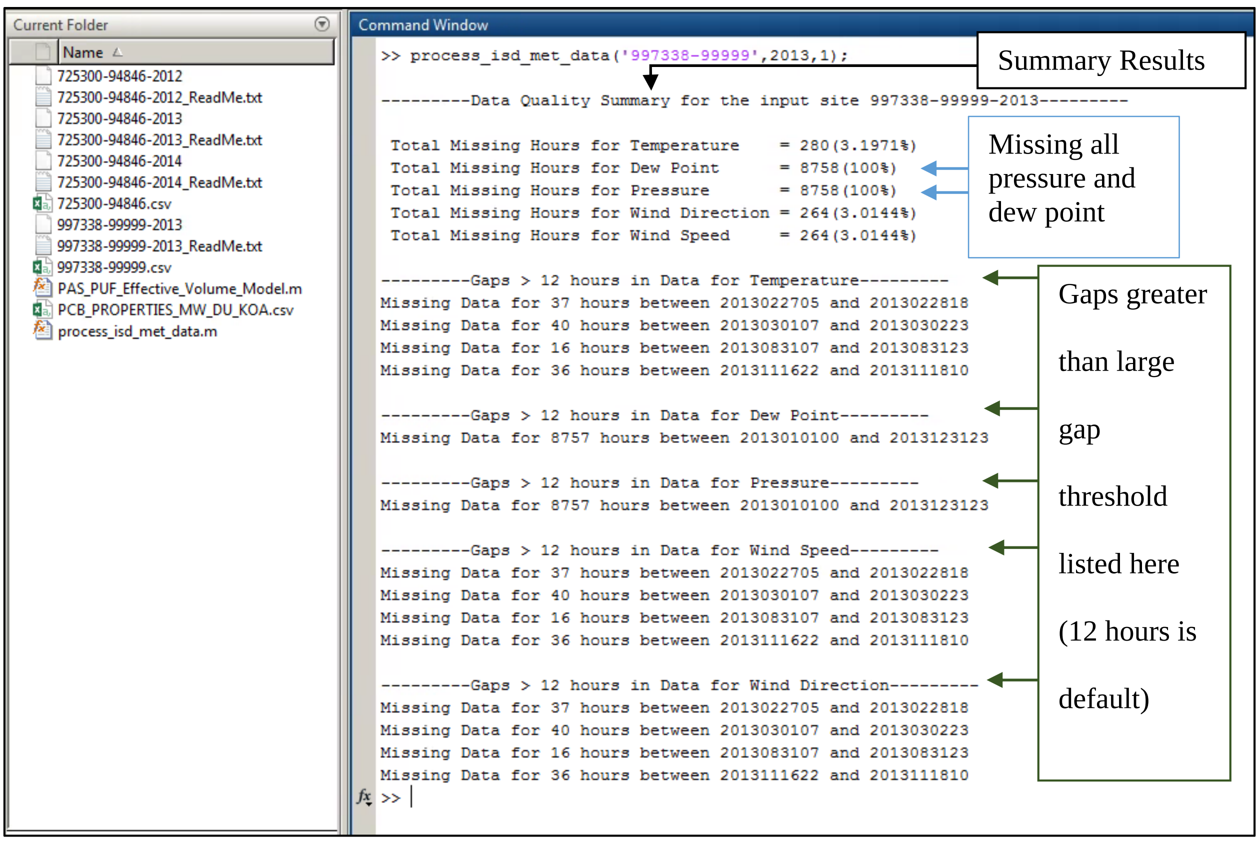
b. A final formatted csv file with the site ID will be created with a summary of the data quality for each year that will be both printed to the screen after the run and saved as a readme file to be accessed later. The readme file will overwrite any previous version of the file. The script will provide a summary of missing values for all 5 weather parameters and detect any gaps greater than 12 hours.

c. Below is an example of a weather data station being run for 3 consecutive years and the resulting data quality summary.



d. The script will also detect if there are any gaps in data greater than 12 hours. This threshold can be change in **line 68** of the **process_isd_met_data.m** script if desired. In the example shown below there are 4 gaps greater than 12 hours for temperature, wind speed, and wind direction, while pressure and dew point are missing all measurements for the year. This data quality would be insufficient for the flowrate model since it does not contain any pressure or dew point measurements, and a different weather station would need to be selected (steps 1-2)

```
64
65   %Subfuntion that counts the null values (-9999) and logs any large gaps.
66   %for missing values it assumes continuity with previous measurements
67
68 -      LargeGapThreshold = 12; %Threshold (in hours) for flagging consecutive missing values
69
70 -      [TA,TAcount,TALargeGaps] = correctgaps(TA,0.1,LargeGapThreshold);
71 -      [TD,TDcount,TDLargeGaps] = correctgaps(TD,0.1,LargeGapThreshold);
72 -      [Pr,Prcount,PrLargeGaps] = correctgaps(Pr,10,LargeGapThreshold);
73 -      [WD,WDcount,WDLargeGaps] = correctgaps(WD,1,LargeGapThreshold);
74 -      [WS,WScount,WSLargeGaps] = correctgaps(WS,0.1,LargeGapThreshold);
75
```



4. After the met data has been converted use the **PUF_PAS_Effective_Volume_Model.m** script to obtain congener and deployment specific effective sampling volumes and sampling rates [Note: ensure the accompanying PCB physical-chemical properties CSV file (**PCB_PROPERTIES_MW_DU_KOA.csv**) and the PCB LFER descriptors CSV file (**PCB_LFER_descriptors.csv**) are in the same location as the **PUF_PAS_Effective_Volume_Model.m** script].

    a. Prior to running the script a few modification at the beginning of the script may need to be made at the beginning of the script.

        1. **Line 12:** Output file name (Use a name descriptive to the specific run)

        2. **Line 17:** Select the method for determining the air/PUF partition coefficient
            a. Method 1 is the empirical relationship (with temperature adjusted Koa) from Shoeib and Harner (2002) [7]
            b. Method 2 is the temperature dependent linear free energy relationship for polyurethane foam from Sprunger et al. (2007)[5]

        3. **Lines 21-24:** PUF Disk Parameters specific to the study.

    Note: A screen shot of the beginning of the script where the above parameters can be adjusted is shown below.

FILE NAVIGATE EDIT BREAKPOINTS RUN

```
1   function[PAS_Array,DeploymentMetData] = PUF_PAS_Effective_Volume_Model(Deployment,Collection,PUF_ID,MET_ID)
2   %This model was originally published by Herkert et al. (2016) Environ. Sci. Technol.
3   %based upon a model first developed by Petrich et al. (2013) Environ. Sci. Technol.
4
5   %Input variable formats:
6   %     Deployment = PUF Deployment Date formatted as  YYYYMMDDHH
7   %     Collection = PUF Collection Date formatted as  YYYYMMDDHH
8   %     PUF_ID = Numerical PUF ID for identification in output data set
9   %     MET_ID = string of AWS site ID formatted as 'XXXXXX-XXXXX.csv' for the weather data to be used for the PUF deployment
10
11  %Specify output file name and location
12 -    outfilename = 'Generic_Output_File_Name.csv';
13
14  %Specify the method for determining KPUF.
15  %    1 = Empirical relationship with temperature adjusted Koa from Shoeib and Harner (2002) Environ. Sci. Technol.
16  %    2 = Temperature dependent linear free energy relationship for polyurethane foam from Sprunger et al. (2007) Anal. Chem.
17 -    KPUFMethod = 1;
18
19  %Declarations: Known variables about PUF disk parameters to be used
20  %The default parameters are for PUF from Tisch Environmental used by Herkert et al. (2016)
21 -    diameter = 0.14;                 % Length of PUF disk (m)
22 -    thickness = 0.0135;              % Thickness of PUF disk (m)
23 -    dpuf = 21173;                    % Density of PUF (g/m^3)
24 -    As = .0365;                      % Surface area of the PUF (m^2)
25
```

b. The input should be in the form

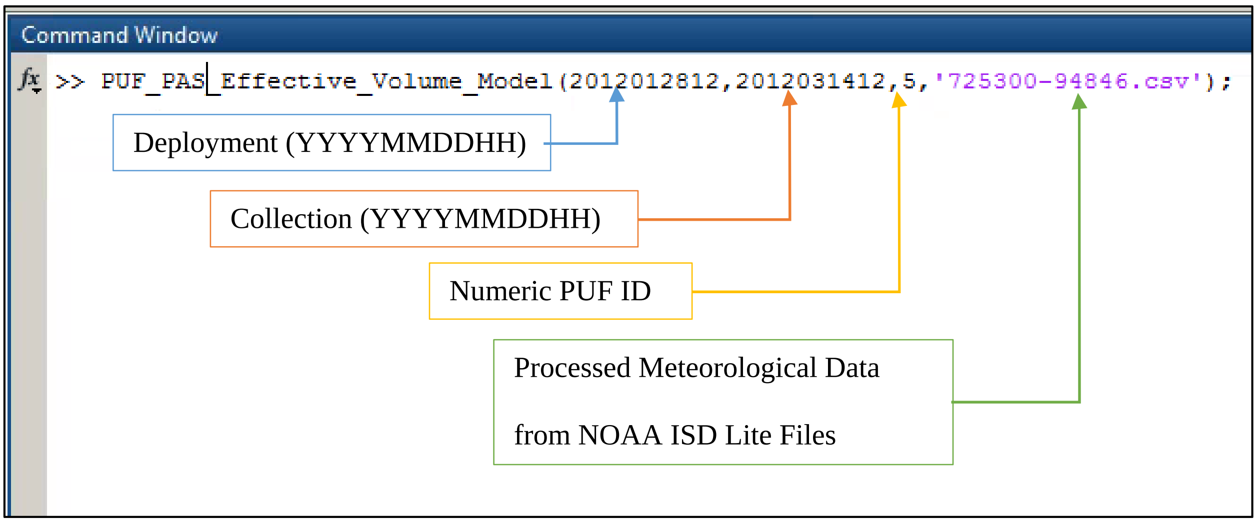
PUF_PAS_Effective_Volume_Model(Deployment,Collection,PUF_ID,MET_ID);
where,

i. Deployment = PUF Deployment Date formatted as YYYYMMDDHH

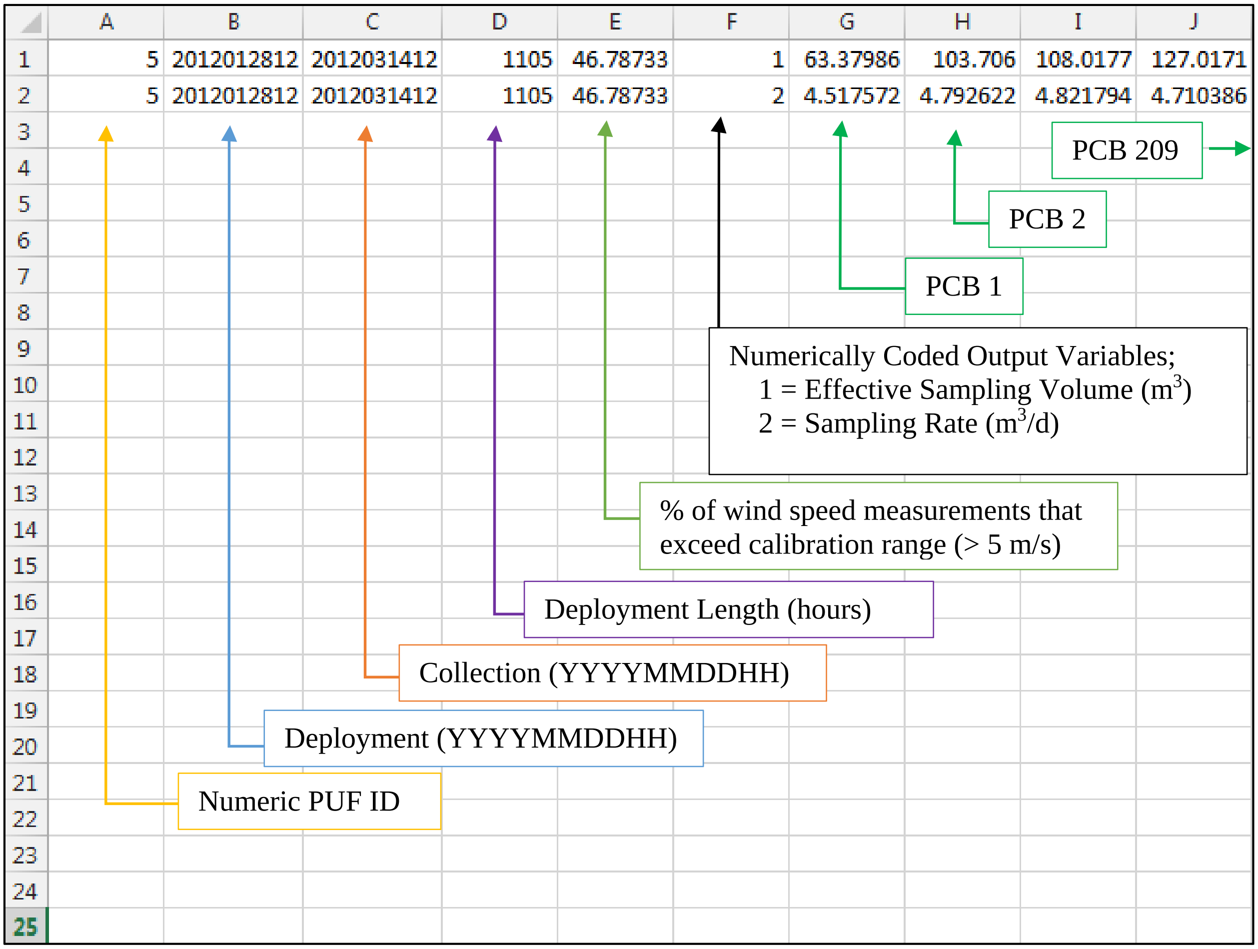
ii. Collection = PUF Collection Date formatted as YYYYMMDDHH

iii. PUF_ID = Numerical PUF ID for identification in output data

iv. MET_ID = string of AWS site ID formatted as 'XXXXXX-XXXXX' for the weather data to be used for the PUF deployment

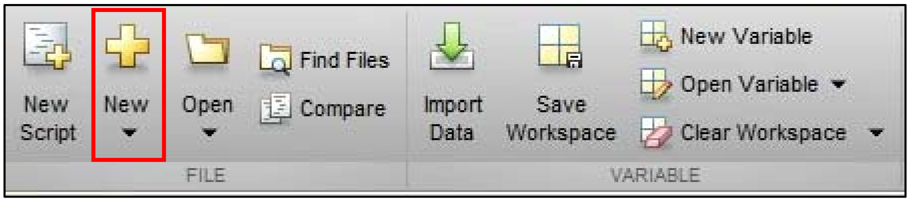


c. The final output file will produce a comma delimited file that can be opened in Microsoft Excel.

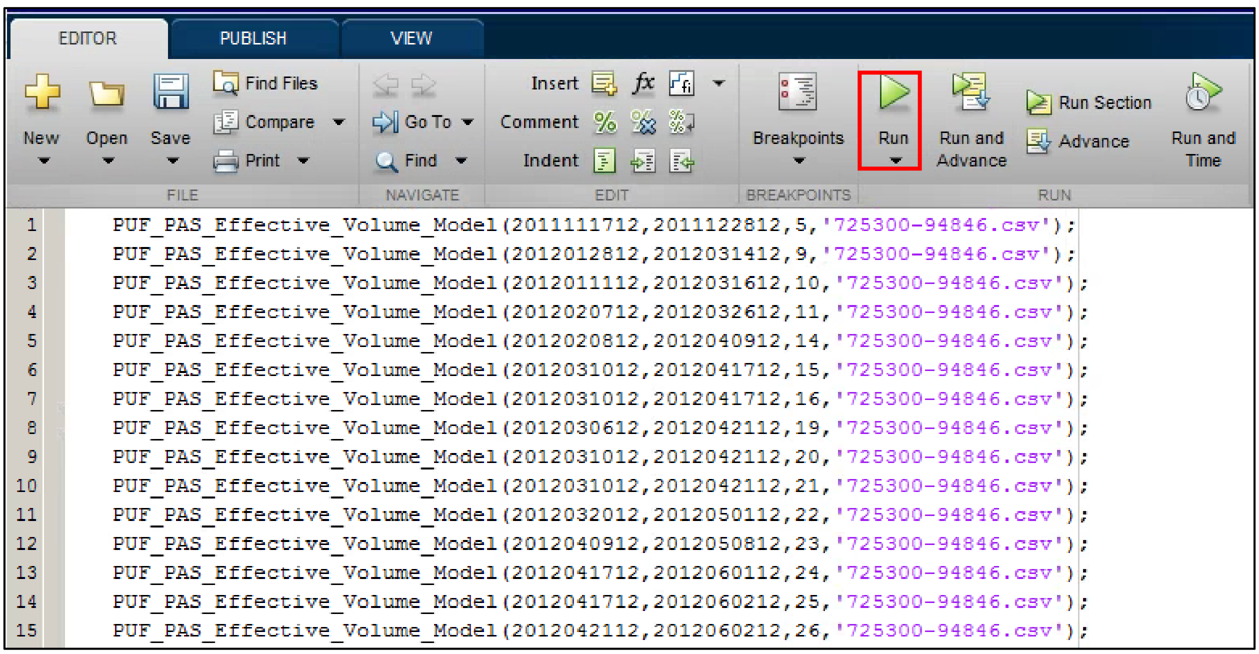


5. **(OPTIONAL)** If you want to run many samples in a row, a new script can be set up to run PUF consecutively.

   a. First open a new blank script by clicking “New Script” in the open left-hand corner of the MATLAB window.



   b. Then repeat the format from step 4B for as many PUF as needed and hit run (highlighted in red box)



c. These run commands can also be created in excel using the concatenate function and pasted into the new script to speed up this process.

SUM | =CONCATENATE("PUF_PAS_Effective_Volume_Model(",B16,",",C16,",",A16,",",D16,");")

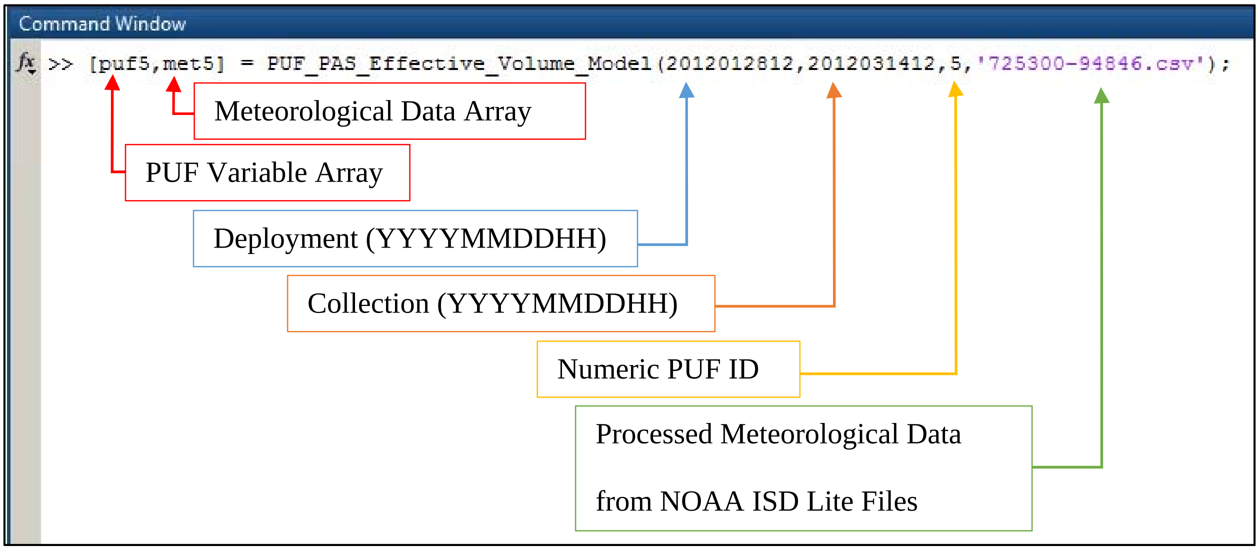
	A	B	C	D	E	F
6	PUFID -->	Numerically Assigned PUF ID for identification in output data				
7	Deployment -->	Deployment Date formatted as YYYYMMDDHH				
8	Collection -->	Collection Date formatted as YYYYMMDDHH				
9	Met_ID -->	AWS Site ID formatted as a string 'XXXXXX-XXXXX'				
10						
11	PUF_ID	Deployment	Collection	MET_ID		Run Command
12	5	2011111712	2011122812	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2011111712,2011122812,5,'725300-94846.csv');
13	9	2012012812	2012031412	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012012812,2012031412,9,'725300-94846.csv');
14	10	2012011112	2012031612	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012011112,2012031612,10,'725300-94846.csv');
15	11	2012020712	2012032612	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012020712,2012032612,11,'725300-94846.csv');
16	14	2012020812	2012040912	'725300-94846.csv'		=CONCATENATE("PUF_PAS_Effective_Volume_Model(",B16,",",C16,",",A16,",",D16,");")
17	15	2012031012	2012041712	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012031012,2012041712,15,'725300-94846.csv');
18	16	2012031012	2012041712	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012031012,2012041712,16,'725300-94846.csv');
19	19	2012030612	2012042112	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012030612,2012042112,19,'725300-94846.csv');
20	20	2012031012	2012042112	'725300-94846.csv'		PUF_PAS_Effective_Volume_Model(2012031012,2012042112,20,'725300-94846.csv');

d. If this is done the output file will have the same format as previously described preceding down the columns in the order the PUF were run

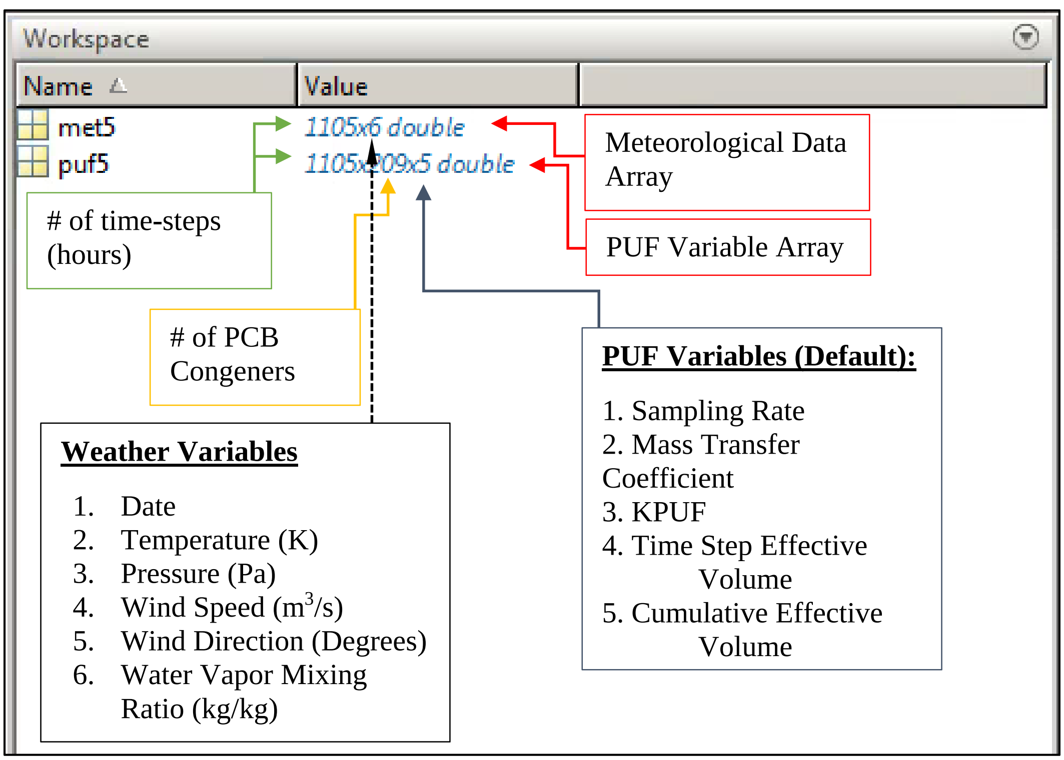
	A	B	C	D	E	F	G	H	I	J	K	L
1	5	2012012812	2012031412	1105	46.78733	1	63.37986	103.706	108.0177	127.0171	169.9775	163.9116
2	5	2012012812	2012031412	1105	46.78733	2	4.517572	4.792622	4.821794	4.710386	5.00201	4.955705
3	9	2012012812	2012031412	1105	46.78733	1	63.37986	103.706	108.0177	127.0171	169.9775	163.9116
4	9	2012012812	2012031412	1105	46.78733	2	4.517572	4.792622	4.821794	4.710386	5.00201	4.955705
5	10	2012011112	2012031612	1561	44.52274	1	58.13245	107.682	113.4634	142.424	207.1393	197.7077
6	10	2012011112	2012031612	1561	44.52274	2	4.389495	4.656748	4.685093	4.576843	4.860199	4.815207
7	11	2012020712	2012032612	1153	44.57936	1	37.76773	69.94987	74.10364	97.1599	148.5119	140.5724
8	11	2012020712	2012032612	1153	44.57936	2	4.728145	5.016016	5.046548	4.929946	5.235163	5.1867
9	14	2012020812	2012040912	1465	45.80205	1	50.97058	85.92163	90.42468	115.4995	178.3171	168.3355
10	14	2012020812	2012040912	1465	45.80205	2	4.793166	5.084996	5.115948	4.997743	5.307157	5.258027
11	15	2012031012	2012041712	913	44.24973	1	46.8062	77.45591	81.04287	98.83027	139.6872	133.5291
12	15	2012031012	2012041712	913	44.24973	2	5.087936	5.397713	5.430568	5.305094	5.633537	5.581385
13	16	2012031012	2012041712	913	44.24973	1	46.8062	77.45591	81.04287	98.83027	139.6872	133.5291
14	16	2012031012	2012041712	913	44.24973	2	5.087936	5.397713	5.430568	5.305094	5.633537	5.581385
15	19	2012030612	2012042112	1105	47.69231	1	49.67255	82.53264	86.51863	107.1889	156.6825	149.045
16	19	2012030612	2012042112	1105	47.69231	2	5.067846	5.376399	5.409125	5.284146	5.611292	5.559346

6. (OPTIONAL) The outputs (as well as meteorological data) can be saved directly in MATLAB to create figures and view trends.

a. Assign Variable for the PUF array and meteorological data



b. After running the **PUF_PAS_Flowrate_Model.m** script, the Meteorological Data Array and PUF Variable Array are available in the workspace.



c. The PUF Variables are defined on **lines 103-113** of the **PUF_PAS_Effective_Volume_Model.m** script and additional variables can be defined here by increasing the size of the **PAS_Array** to examine other trends if desired.

```
102
103        %Fill in PAS_Array for variables. This is a 3D array that gets populated during the run
104            PAS_Array(t,C,1) = Rs;
105            PAS_Array(t,C,2) = kv;
106            PAS_Array(t,C,3) = KPUF;
107            PAS_Array(t,C,4) = Veff;
108
109        %Index 5 is time integrated effective volume
110            if t == 1
111                PAS_Array(t,C,5) = Veff;
112            else
113                PAS_Array(t,C,5) = Veff  + PAS_Array((t-1),C,5);
114            end
115     end
116 end
```

References

1. National Centers For Environmental Information, Integrated Surface Database (ISD). www.ncdc.noaa.gov/isd (Accessed December 3, 2015),
2. Melymuk, L.; Robson, M.; Helm, P. A.; Diamond, M. L., Evaluation of passive air sampler calibrations: Selection of sampling rates and implications for the measurement of persistent organic pollutants in air. *Atmos. Environ.* **2011,** *45*, (10), 1867-1875.
3. Chaemfa, C.; Barber, J. L.; Gocht, T.; Harner, T.; Holoubek, I.; Klanova, J.; Jones, K. C., Field calibration of polyurethane foam (PUF) disk passive air samplers for PCBs and OC pesticides. *Environ. Pollut.* **2008,** *156*, (3), 1290-1297.
4. Harner, T.; Su, K.; Genualdi, S.; Karpowicz, J.; Ahrens, L.; Mihele, C.; Schuster, J.; Charland, J. P.; Narayan, J., Calibration and application of PUF disk passive air samplers for tracking polycyclic aromatic compounds (PACs). *Atmos. Environ.* **2013,** *75*, 123-128.
5. Sprunger, L.; Acree, W. E.; Abraham, M. H., Comment on "Systematic investigation of the sorption properties of polyurethane foams for organic vapors". *Anal. Chem.* **2007,** *79*, (17), 6891-6893.
6. Kamprad, I.; Goss, K. U., Systematic investigation of the sorption properties of polyurethane foams for organic vapors. *Anal. Chem.* **2007,** *79*, (11), 4222-4227.
7. Shoeib, M.; Harner, T., Characterization and comparison of three passive air samplers for persistent organic pollutants. *Environ. Sci. Technol.* **2002,** *36*, (19), 4142-4151.