Focal Sum Model User Guide

Version: focalsum_04082010.mxd

Date 4 August 2010

Ву

Kayoung Lee, Danielle Vienneau, Kees de Hoogh and David Briggs

Department of Epidemiology and Biostatistics Imperial College London

Introduction

The Focal Sum Model is a GIS-based model designed to estimate air pollution exposures from source emissions. It is based on the principle that proximity to the source (e.g. roads or industry) is one of the main determinants of pollutant concentrations.

It uses the ArcGIS *focalsum* function with a weighted kernel to simulate the distance decay from sources. The weighted emissions are then calibrated with monitored air pollution data, using linear regression, to produce a concentration surface. An optional validation using a reserved set of independent monitoring is also included (Figure 1).



Figure 1. The modelling approach

Focal functions are applied to raster data, and compute an output raster where the cell value at each location (target cell) is some function of the input cells in a specified window (defined by the kernel file) around the location (Figure 2).

	Input raster					Same	e Inpu	t rast	er	Output raster				
2	4	2	1	5	2	4	2	1	5					
3	2	1	5	1	3	2	1	5	1		12.5	13.5	14.5	
1	6	2	3	4	1	6	2	3	4		17	15	13.5	
5	5	3	1	4	5	5	3	1	4		17.5	15	12	
2	2	4	1	1	2	2	4	1	1					
with mul sum writ cell	1. The values of the cells within the 3x3 window are multiplied by weights and summed. The result is written to the target (centre) cell in the output grid (blue).				move perfection calcutor of ce writt outp	es to ti orms th ulation ells. Th en to ti ut grid	he righ ne san on the ne res he targ (grey)	nt and ne e next ult is a get cel).	group gain I in the	move calcu writir grid.	across lating them	the ir he valu to the	nput gr ues, ar outpu	id nd t
	0.5	0.5	0.5											
	0.5	1	0.5											
	0.5	0.5	0.5											
3x	3 weig	hted	windo	w										

Figure 2. Focal sum example with 3x3 weighted kernel

In the Focal Sum Model, the window is moved across an input emissions raster to collect (i.e. sum) the source contributions and assign them to the target cell. With a weighted kernel one can assign greater contributions from nearby sources. Cells directly downwind of an industrial stack, for example, are expected to receive a greater source contribution than cells adjacent to but upwind of the cell in which the source is located.

In the Focal Sum Model the size, basic shape (square, circle) and weights (1/d, 1/d2 where d = distance in cells) for the kernel file can be specified by the user. Users may also wish to use existing kernel files either manually created or derived, for example, from dispersion models.

Software

The Focal Sum Model runs in ESRIs ArcGIS v9.3 ArcMap, and requires Spatial Analyst. The code is written in VBA and is embedded and saved in an ArcMap document.

To avoid overwriting the original, the user should make a copy of this document before running the model.

Data needs

Data format: ESRI's ArcGIS shape file and raster

Two main data inputs are required:

(1) Monitored air pollution at a sample of sites (point shape file).

One shape file is required for the training set. To include model validation, an optional second shape file containing an independent subset of sites is required. Both the training and validation shape files require a unique ID of 'long' integer type (i.e. in addition to the default FID). If it does not exist, the tool will prompt the user to create a new unique ID field.

(2) Emissions, or a proxy for them, to provide measures of source intensity (raster).

Test dataset has been provided by Daniela Fecht who developed urban SImulation for ENvironment health Analysis (SIENA). This includes:

- 250m grids for
 - \circ PM₁₀ and NO_x (t/km²/year)
 - o Emissions from main roads, minor roads and total road emissions
- Monitoring network shape files
 - $\circ \quad \mathsf{PM}_{10} \text{ and } \mathsf{NO}_2$
 - o Representing two-week average concentrations measured in summer
 - Routine monitoring network of 10 sites
 - Study area network of 45 sites

When saving output files (shape files, tables or rasters) please specify file names not already used. Suggested output names are provided, but the user should customise if running the model more than once.

Running the model

When opened the user will see the following toolbar:

🏟 Focalsum_26072010 - ArcMap - ArcInfo		
Ele Edit View Bookmarks Insert Selection Tools	Window Help	
D 📽 🖬 🚭 X 🖻 🎕 🗙 🗠 🗠 -	🔸 🛛 1:7,419,527,627 💽 🛃 📣 🚳 🗉	1 🔐 🎠 🕅 🖉 🔍 😋 🗱 22 전) 🕲 🗭 🔿 🖗 🔍 🔪 🛛 🗛 👶 🏄 🌮 💭 🔤 👘 🐨 🗑 🐘
Spatial Analyst - Layer		
Layers	ArcToolbox Arcyst Tools Analysis Tools Analysis Tools Canversion Tools Conversion Tools Conversion Tools Conversion Tools Deta Interceperability Secure Tools Server Tools Server Tools Sested Statistics Tools Tracking Analyst Tools	

If the 'focalsum toolbox' toolbar does not appear: go to Tools – Customize – scroll to the buttom and click 'focalsum toolbox'. The Model tool will then be displayed onto the ArcMap window.



The steps to run this tool are outlined below:

1. Set environment

The set environment window prompts the user to define the folder in which all the output data from the model run will be stored.



If this step is skipped the following message will appear to prompt the user to set the workspace:



2. Create Kernel File

In this step the user specifies the parameters for creating the weighted kernel file. The tool creates a kernel with an odd number of rows and columns (e.g. 3x3 or 7x7 kernel) so that a true middle is defined. Weights are then computed using linear distance in cell units from the centroid of the 'middle kernel cell' to all other 'kernel cell' centroids.

2. Create Kernel File 🛛 🔀	
Kernel File Name:	Set the kernel shape file template name, this name will also be assigned to the kernel text file
Define Kernel Resolution	
Select emissions:	Select emission raster. Cell resolution information
Resolution: m	raster is selected
- Define Radius of Kernel	
Kernel Radius: cells	Select kernel radius by cell numbers.
Kernel shape:	
Decay Function:	Select the kernel shape – square or circle. Select decay function – $1/d$ or $1/d^2$. Decay function can
Kernel File Name:	be calculated based on cell units.
Verify Run 🐲	

If the user does not select all parameters, the following message will be shown. Once the missing parameters are entered the model can be run.



When Step 2 finishes running, a message box (shown below) tells the user that the kernel file is stored in: *'workspace which is set from step 1'\kernel_txt\'kernel file name'*. This kernel file will be used as an input in step 3.

АгсМар	
(į)	Kernel File C:\test0726\kernel_txt\wei_square.txt created!
	ОК

-	kernel_txt			
Contraction of the local division of the loc	File Edit View Favorites Tools	Help		A1
	Gask · 🕤 · 🎓 🔎 Sea	arch 😥 Folders 💷 -		
	Add ess 😂 C:(test0726)/æmel_bit			💌 🛃 Go
	Folders	× Nome ↔	Size Type	Date Modified
	GIS_workshop GIS_tworkshop GIS_tw	Kel_square	2 K8 Text Document	27/07/2010 13:33
D wei_squ	are - Notepad			
File Edit F	ormat View Help			
0.16666666 0.18274396 0.19519410 0.2 0.25 0.3 0.19519410 0.18274395 0.16666666 0.15022110	38667 0 1907 43569631 0 217 22022555 7652 0 217 2202556 0 2610387 4564 11601 0 2402507 3552 0 3690 1689437 33333353333 0 5 1 0 5 0 3355335335 1010 0 2402537 3552 0 3600 1689437 767532 0 217 12972556 0 3610387 456 1010 0 240259031 0 217 12927255 14652 0 1666666666667 0 1827 43997633 14652 0 1866666666667 0 1827 43997633 14652 0 18652 0 1827 1827 1825 14652 0 1827 1827 1825 14752 0 1855 14752 0 18752 0 18752 14752 0 18752 0 18752 0 18752 14752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 18752 0 187520 0 18752 0 18752 0 18752 0 18	6 0 240053073952 0 .25 0 .2400530733 0 390610994376 0 3333333333 0 5 0 414213562373 0 .5 0 41421356237 3 0 .25 0 .2 0 414213562373 0 .5 0 .41421356237 0 .30016994376 0 .3333333333 0 .24005307352 0 .25 0 .2402530733 0 .24005307352 0 .25 0 .2402530733 0 .196194101601 0 .2 0 .19619410160	52 0.217129272955 0.190743669831 508016994375 0.261203874964 0.212 31 0.309016994375 0.240253073352 0 31 0.309016994375 0.240253073352 0 300016994375 0.240253073352 0 300016994375 0.241203874964 0.212 52 0.217129272955 0.190743669831 10.182743997632 0.16666666666666 0	0.166666666667 129272965 0.182743997632 195194101601 195194101601 129272965 0.182743997632 0.16666666667 150221104822
in the second	II i Student II i Student			

When the user selects the 'Verify' button, the summary information for the parameters used to create the kernel file will be shown in a message box.

After clicking 'Run', if a kernel radius of less than 10 cells was defined, the calculated weights for the distance decay function will automatically be visualised in the map view. Otherwise, just the points will display, but the user can add labels to visualise.

😧 Focalsum_04082010 - ArcMap - ArcInfo	1								
<u>Eile Edit View Bookmarks Insert Selection Iool</u>	s <u>W</u> indow <u>H</u> elp								
🗋 🗅 🗃 🖬 🚭 X 🖻 🛍 X 🗠 🗠	+ 1:4,551,587 -] :! 🕭 🄇) 🗆 éř 🥍 🕅	? 🛛 🍳 🤤 🕻	: 22 🖑 🥥 🛊	🖬 🔿 🖓 🖨	0 M 🕺	🚣 🐔 📃 🗟	XTools Pro 💌 🗐
Spatial Analyst - Layer: main_nox_250	- ½ 📐								
		1. Set	Environment 2. Crea	ate Kernel File 3.	Run Focal Statistics	4. Create Concer	ntration Surface		
	ArcToolbox		0.150221 ♦	0.166667 ◆	0.182744 •	0.195194 •	0.2	0.195194 •	0.182744 •
□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □	Conversion Tools Society Conversion Tools Society Coverage Too		0.166667 •	0.190744 ◆	0.217129 ◆	0.240253 •	0.25 •	0.240253 •	0.217129
	Linear Referencing Times of Multidimension Tools Multidimension Tools Some Samples Server Tools Spatial Analyst Tools		0.182744 •	0.217129 •	0.261204 •	0.309017 •	0.333333 •	0.309017 •	0.261204 •
2. Create Kernel File Kernel File Name: wei <u>-</u>	Rquare		0.195194 •	0.240253 •	0.309017 ◆	0.414214 •	• ^{0.5}	0.414214 •	0.309017 •
Define Kernel Resolution Select emissions: main_nox_3 Resolution: 250	250 _ _		0.2 •	0.25 •	0.333333 •	0.5	1 •	0.5 •	0.333333 •
Define Radius of Kernel	cells		0.195194 ♦	0.240253 ♦	0.309017 ♦	0.414214 ♦	0.5	0.414214 ♦	0.309017 ♦
Kernel shape: Square Decay Function: 1/d Kernel File Name: Well_SQ	uare.bt	ArcM You 9 × 250 wei	ap will create the followir 9 cells m resolution square shp shape file	ng kernel file:	0.261204	0.309017 •	0.333333 •	0.309017 •	0.261204 •
			square.txt weights]	0.217129 •	0.240253 •	0.25	0.240253 •	0.217129 •
			0.150221 •	0.166667 ♦	0.182744 •	0.195194 •	0.2 •	0.195194 ♦	0.182744 ◆
Display Source Selection	Favorites Index Se ()	00	Ⅱ ◀						

Once step 2 finishes, Next ⇒

button will be shown.

3. Run Focal Statistics

In this step, the kernel file is used with a *focalsum* function to weigh and sum the emissions to create a weighted emission raster. This process is illustrated in Figure 2 above, and an example raster input and output are shown below.





4. Create Concentration Surface

In these steps, the weighted emissions are calibrated with monitored air pollution data, using linear regression. In step 4.1 the weighted emission value for each monitoring site is extracted and added to the calibration shape file. The parameters used to develop the regression model are defined in step 4.2. If a suitable calibration model is obtained, the user can apply the model to the weighted emission raster in the final step to produce a concentration surface.

4. Create Concentration Surface	
Step 4.1 Extract weighted emissions to training sites	Select training sites – input shape file should have 'Long Integer' field type. If not, tool will prompt user to create a new field called 'uniqueID'
Weighted emissions: w_emission	Weighted emission raster from step 3 will be automatically added
	Define calibration shape file name. Tool gives 'calib_' as a default name

4.1. Extract weighted emissions to training sites

If the user selects an input file without a Long interger unique id field, a warning message will appear. If the user clicks 'Yes', the uniqueID field will be created automatically. If the user clicks 'No' then the tool will terminate.

8
•

The 'Extract Values to Points' tool is used in this step. If this has previously been run with the selected training sites the shape file may already contain a 'RASTERVALU' field. The tool will warn the user if a 'RASTERVALU' field already exists and ask if it should be deleted. If the user does not want to delete the existing 'RASTERVALU' field, the program will terminate.

ArcMap	АгсМар
Please delete 'RASTERVALU' field or choose different input Do you want to delete 'RASTERVALU' field?	You should delete 'RASTERVALU' or choose different input to run the model.
Yes No	 ОК

Once run by the user, the 'Run' button on 4.1 will be hidden and 4.2 tab will appear so that the user is directed to the next step. This functionality also applies to other steps with tabs.

Low :	: II.		
4.	Create Concentration	Surface	×
ſ	Step 4.1 Step 4.2		1
	Extract weighted emission	ns to training sites	
	Training sites:	training_sites	
	Weighted emissions:	w_emission 💌 🔄	
	Calibration shapefile:	calib_	

4.2. Run simple regression to calibrate model

4. Create Concentration Surf	ace	×	Shape file of monitoring sites containing both measurement data (dependent) and weighted emissions (explanatory) variables for analysis
Calibration shapefile:	calib_		Long integer Unique ID field
Unique Field ID: Dependent Variable:	ID •		Numeric field containing measured concentrations
Explanatory Variable:	RASTERVALU		Explanatory variable is always 'RASTERVALU' representing weighted emissions(output from step 4.1)
Residuals Shapefile: Regression Model Table:	c_residual		Name of output shape file containing estimated value and residuals
Diagnostics Table:	c_diag_table		Name of output table containing model coefficients and standard errors
	Rur		Name of output table summarising model building statistics, including coefficient of determination (R ²)

Please note that when setting parameters in Step 4.2, different inputs are required for each of Unique Field ID, Dependent variable and Explanatory variable. Otherwise the tool will show an error message and terminate.



After simple regression is run, a message appears indicating that the detailed regression output (e.g. R^2 , regression coefficients, etc) can be viewed by going to the 'source' tab as shown below.



The regression equation is also written to a text file found under: 'workspace which is set from step 1'\Calibration_stats_txt\'Regression Model Table name_calibration.txt'

Calibration_stats_txt	
File Edit View Favorites Tools Help	
🚱 Back 🔹 🕥 🕤 🏂 🔎 Search 👔	Folders
Address 🔯 C:\test0726\Calibration_stats_txt	💌 🄁 Go
Folders	Name - Size Type Date Modified
🛅 GIS_workshop	C_reg_table_calibration 1 KB Text Document 27/07/2010 15:04
🗉 🚞 Inetpub	
info	C rea table calibration Notanad
🗄 🧰 Intel	Ele Edit Except View Hele
🔲 logso –	- File Calt Format View Help
	Pred = 18.1561221143 + w_emission * 0.103252689854
🗉 🚞 MSDERelA	12 - 0.440
New Folder	
New Folder PCRaster	
New Folder PCRaster Porgram Files	
New Folder New Folder Program Files Puthon24	2

This equation is used to calculate new concentration surface in step 4.3.

4.3. Apply calibration to weighted emissions to compute concentrations

After Step 4.3 is run and the new concentration raster created, a message will appear asking whether the user will proceed the next step for validation.



If the user clicks 'Yes' the optional Model Validation window will open. Step 5.1 and 5.2 use the same process as in step 4.1 and 4.2, respectively, though should be applied to the shape file containing the monitoring sites reserved for validation. If the user clicks 'No' then the tool will exit. The new concentration raster and regression outputs are automatically saved to the set workspace.

5. Model Validation – Optional

In these steps, the new concentration raster is validated with independent monitored air pollution data, also using linear regression. The final output includes regression statistics indicating model performance.

5. Model Validation - Optional	
Step 5.1 Extract concentrations to validation sites Validation sites:	Select validation sites – input shape file should have 'Long Integer' field type. If not, tool will prompt user to create a new field called 'uniqueID'
Concentration raster: new_conc Validation shapefile: valid_ Run 🗇	New concentration raster from step 4.3 will be automatically added
	Define validation shape file name. Tool gives 'valid_' as a default name

5.1. Extract concentrations to validation sites

As previous, the 'Run' button on 5.1 will be hidden and the 5.2 tab will appear so that the user is directed to the next step.

5.2. Extract concentrations to validation sites

5. Model Validation - Optional Step 5.1 Step 5.2 Run Simple Regression to validate model	Shape file of monitoring sites containing both measurement data (dependent) and weighted emissions (explanatory) variables for analysis
Validation shapefile: valid_	Long integer Unique ID field
Unique Field ID: ID Dependent Variable NO2	Numeric field containing measured
Explanatory Variable RASTERVALU	Explanatory variable is always
Enter names for the validation outputs Residuals Shapefile: v_residual	'RASTERVALU' representing modelled concentrations (output from step 5.1)
Regression Model Table: v_reg_table	Name of output shape file containing estimated value and residuals
Run 🗇	Name of output table containing model coefficients and standard errors
	Name of output table summarising model validation statistics, including coefficient of determination (R ²)
ArcMap Validation Statistics: r2 = 0.51 Click source tab to see detailed table information Also see, C:\test0726\Validation_stats_txt\v_reg_table_validation OK	.txt

As per step 4.2, when successfully run a message appears showing the validation R^2 and indicating that the detailed regression output (e.g. R^2 , regression coefficients, etc) can be viewed by going to the 'source' tab.