Excerpt from the Proceedings of the 2011 COMSOL Conference in Boston

Making Cartograms and Using Them for Data Acquisition

P. Kip Mercure[*1], and Robert P. Haley[2]
[1]The Dow Chemical Company, Midland, Michigan, [2]ATM Research, Midland, Michigan
*Corresponding author: 1909 Vine, Midland, Michigan 48642, pkmercure@gmail.com

Abstract: A diffusion-based method can be used in COMSOL Multiphysics to distort the boundaries in a map to represent thematic data. We use a diffusion model coupled with geometry deformation to construct a cartogram representing population differences. Thematic data, such as population, is used to establish the initial values for the diffusion problem and the "Moving Mesh Interface" is used to move the boundaries according to the velocity generated from the diffusion problem. We apply this cartogram construction technique to thermocouple locations to maximize the accuracy of measuring the spatial variation of temperature. A heat conduction model is used to generate the Laplacian of the temperature field. This is used to provide the initial values for diffusion-cartogram construction. The cartogram maps a uniform distribution of measurement points back to the original geometry. This results in a distribution of measurement points that equally distributes the approximation error across the spatial region.

Keywords: cartogram, mapping, optimization, data acquisition

1. Introduction

Measuring temperature profiles is an important part of product and process development. We often encounter the need to accurately estimate temperature or other variable over spatial regions using the fewest number of sensors. This report addresses the issue of calculating improved sensor locations. The specific problem addressed is how to accurately estimate temperature profiles in a diesel particulate filter undergoing thermal shock testing.

If a part has a great deal of symmetry, placing temperature sensors can be straightforward, but many systems have an irregular geometry making sensor location non-intuitive. Uniform cylindrical diesel particulate filters have relatively intuitive sensor placement, but when a filter is built up from segments, the resulting geometry makes sensor placement an issue.

There are suggestions in the literature for improving sampling along one dimension [1] but extending to two or three dimensions is not obvious. If you consider each sensor as representative of a "sampled region" what is needed is to change the areas of the regions such that the approximation errors become uniformly distributed. Considering that a cartogram distorts geographical regions to provide a uniform distribution of some thematic variable, this suggests a way to redistribute sensors in an optimal way.

2. Cartogram Construction

A relatively recent method for cartogram construction has been published by Gastner and Newman [2]. This starts with the idea that if there is a non-equal distribution of population, a diffusion process would cause the distribution to become more even. If the boundaries "flow" with the population, the resulting boundary areas should reflect an equal population. They start with the description of the diffusion process.

$$\frac{\partial p}{\partial t} = D\,\nabla^2 p$$

The exact value of the diffusion coefficient does not matter for the cartogram method, but keeping track of units may aid understanding. The initial condition, p[t=0], is the original population density value. If "insulating" boundary conditions are used, the total amount of population will remain constant within the geometry chosen, because it cannot "flow" outside of the figure boundary. It is useful to include non-populated areas such as lakes and oceans to aid in preserving a recognizable shape. These regions are specified with the same overall average population density as the populated region; this tends to reduce distortion.

Given the solution to the diffusion problem, p[t,x,y], the flux is expressed as:

$$J = -D\,\nabla p$$

The flux can also be expressed as a concentration value times a velocity:

$$J = p\,v$$

The velocity, which can be used to move the boundaries, is:

$$v = \frac{-D\,\nabla p}{p}$$

Note that with the units for the diffusion coefficient of length^2/time, the velocity term will have units of length/time.

We created a cartogram using a diffusion method in COMSOL Multiphysics. Making cartograms is a useful function, but not typically seen in engineering projects.

3. Cartograms in COMSOL Multiphysics

The diffusion calculation uses the COMSOL "Mathematics Interface." Thematic data, such as population, establishes the initial values for the diffusion problem. The gradient of the solution generates a velocity term for the movement of the boundaries. The COMSOL "Moving Mesh Interface" moves the boundaries according to the velocity function generated from the diffusion problem.

Simplified state borders are generated as .dxf files, and imported into the geometry. The initial values of each state are set to the population density for the state.

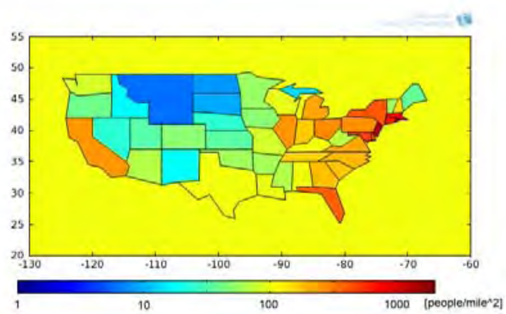

Figure 1. Initial population density. (log scale)

There is a factor of about 300 between the extremes in population density. The initial value of the surrounding region (lakes, oceans, Canada, Mexico) set to average population density of the United States. The diffusion model then allows the population to even out, using the COMSOL Mathematics General Form PDE. A zero-flux boundary condition is used on the outer rectangle. Border movement is calculated by the COMSOL Deformed Geometry form. The outer rectangle border has a prescribed mesh displacement of zero. A prescribed mesh velocity is used for the state borders. Because of sharp differences in population densities initially, the mesh velocity for border movement is damped at the beginning. The functional forms of the border velocities are:

Vx = -uy *(1-exp(-10*t))/u
Vy = -uy *(1-exp(-10*t))/u

During the diffusion process, the low population areas contract, and the high population areas expand.

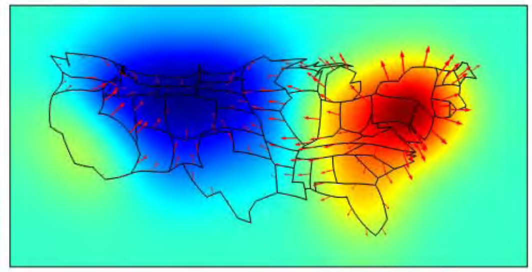

Figure 2. A stage in population diffusion.

A Damping (diffusion) Coefficient of 1 gives uniform population distributions in about 10 time units.

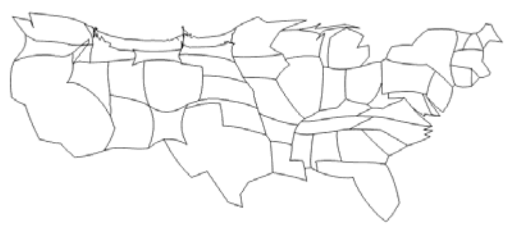

Figure 3. Final cartogram of U. S. population.

4. Application to Sensor Location

Now we apply the cartogram construction technique to distribute thermocouple locations to maximize the accuracy of the measurement of the spatial variation of temperature. An initial model of a heat generation and conduction problem is used to generate a basis solution for the system.

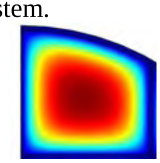

Figure 4. Typical temperature distribution.

Since the heat transfer problem is a second order partial differential equation, a reasonable approximation is by cubic functions. The error term in a cubic approximation function is on the order of the value of the second derivative of the solution value. The value of the second derivative is available in the solution to the heat conduction problem.

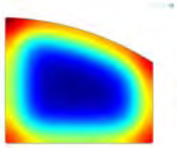

Figure 5. Estimate of approximation error.

Our procedure is:

1. Solve the heat conduction problem to find a typical temperature distribution
2. Extract the Laplacian of the temperature data as an estimate of the approximation error
3. Take the error estimate and use a diffusion process to equalize the error, producing a transform from the original coordinates to an "equal-error" coordinate system
4. Use this transform to redistribute sensor locations
    a. A uniform distribution of measurement points is applied to the cartogram.
    b. The locations are mapped back to the original geometry
    c. This results in a distribution of measurement points the equally distributes the approximation error across the spatial region.

4. Formulation of the Method

The redistribution problem is formulated as a coupled partial differential equation model. Three partial differential equations are solved on the same geometric space.

PDE #1 solves the error diffusion problem. There are no convection or source terms. Zero flux boundary conditions are used at the outer extremes of the part geometry. The initial value is the error estimate, which is the Laplacian of the temperature field from the approximate heat conduction model.

$$\frac{\partial U_1}{\partial t} = \nabla^2 U_1$$

$$U_1[t=0] = \textit{approximation error}$$

$$\frac{\partial U_1}{\partial x} = \frac{\partial U_1}{\partial y} = 0 \textit{ at edge}$$

PDEs #2 & 3 solve the boundary motion problem. In this case we move the coordinate system rather than explicit boundaries. These PDEs have convection but no diffusion or source terms. The convection velocity is calculated from the gradient of PDE #1 following Gastner-Newman method:

$$\textit{velocity} = -\nabla U_1 / U_1$$

The initial value of PDE #2 at each point is the x coordinate; the initial value for PDE #3 at each point is the y coordinate. As the error "diffuses" the coordinate values are moved, just as the boundaries in the cartogram problem. At equilibrium, when the error is equally distributed, the new coordinates can be used to transform the sensor locations.

$$\frac{\partial U_2}{\partial t} = -\frac{\nabla U_1}{U_1} \nabla U_2 \;\; \& \;\; U_2[t=0] = x$$

$$\frac{\partial U_3}{\partial t} = -\frac{\nabla U_1}{U_1} \nabla U_3 \;\; \& \;\; U_3[t=0] = y$$

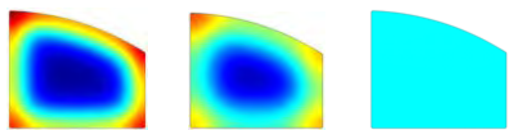

Figure 6. Progress of diffusion of the "error".

Figure 7. Movement of the x-coordinate value

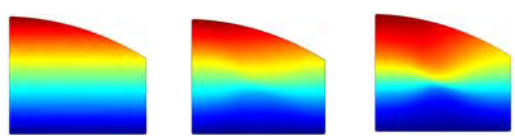

Figure 8. Movement of the y-coordinate value

Figures 6 to 8 show the progress of the diffusion calculation. The equilibrium values are the final distorted coordinate system representing an even approximation error. Sampling points can be evenly distributed across this "evenly-distributed-error" coordinate system. There are nine measurement points used, which is appropriate for the instrumentation in this experiment. If fewer or greater numbers of sensors are desired, this can be accommodated in this initial placement.

5. Redistributing sensors

The formulae for re-distributing sensors are below.

$$x_{new} = U_2[x_{original}, y_{original}, t = \infty]$$

$$y_{new} = U_3[x_{original}, y_{original}, t = \infty]$$

The values of x and y represent a transformed coordinate system.

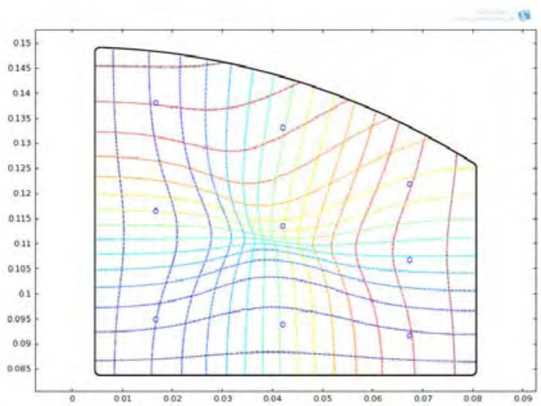

Figure 9. “Equal-error” coordinate system with evenly spaced measurement points superimposed.

Transforming the evenly distributed sensor locations back to the original coordinates gives the optimal sensor locations.

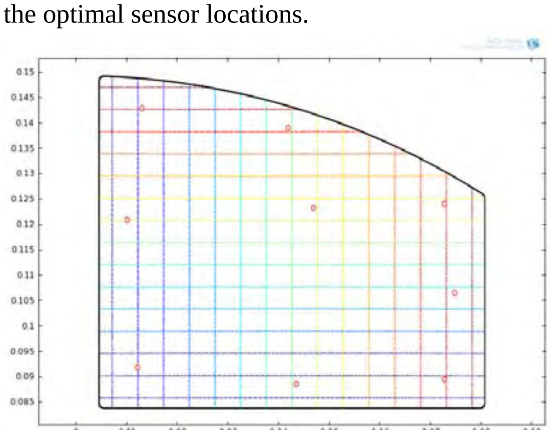

Figure 10. Optimal measurement point distribution.

The center point shows unusual behavior. The suggested location is off-center. This appears to be a result of sensitivity to the placement of the point used as a starting value. Since the temperature profile is so flat, this would not make a practical difference, and the user may choose to stay with the original point with little affect on the results.

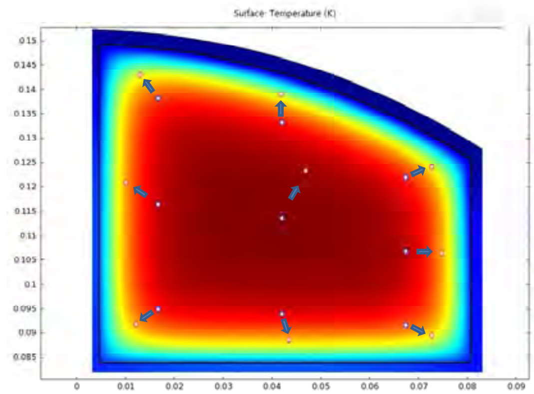

Figure 11. Movement of points on temperature field.

This method of distributing thermocouple points is a much more fundamental approach compared to statistical random sampling procedures.

The chief strength of the technique is its applicability to geometry of any complexity.

6. Conclusions

While building cartograms may not be in the project description for many engineers, the general approach to locating sensors is of great utility in many situations. While the present work is in a two dimensional space, extension to three dimensions is straight forward. The present work is optimal for a single point in time, but there is a possibility that the method could be extended to optimize sensor placement over a range of times

7. References

1. Carl de Boor, *A Practical Guide to Splines*, page numbers. Springer, place (year)
2. Michael T. Gastner and M. E. J. Newman, “Diffusion-based method for producing density-equalizing maps”, *Proceedings of the National Academy of Science*, **101(20)**, 7488-7504 (May 18, 2004)