A Tutorial Introduction to DCM Quantitative Characterization and Modelling of Material Microstructures Using Monochromatic Multi-Energy X-Ray CT

Y. Sam Yang a), Adrian Trinchi, Andrew Tulloh and Clement Chu

CSIRO Manufacturing, Private Bag 10, Clayton, Victoria 3169, Australia

a)Corresponding author: Sam.Yang@csiro.au

Abstract. This article is intended as a tutorial guide for new users of the DCM (data-constrained modelling) software for quantitative characterization of material 3D microstructures using multi-energy X-ray CT data. It guides users through the steps necessary for processing a small CIPS (Calcite In-situ Precipitation System) sandstone data set. It also covers some built-in and plug-in features to analyze and visualize the microstructures.

INTRODUCTION

The DCM (data-constrained modelling) software is a Microsoft Windows™-based application for 3D characterization of compositional material phase distributions and modeling of microstructure-based material properties. It takes quantitative multi-energy X-ray CT data as its input, and applies the Data Constrained Modelling (DCM) technique to generate microscopic partial volume distributions of materials and pores at the CT resolution in 3D. DCM assumes that each voxel (volume pixel) is comprised of a percentage of each defined composition, and requires that the constituent material phases are known apriori. DCM predicted voxel-level partial volumes are the effects of the fine length scale structures that lie below X-ray CT resolution. DCM can also calculate material properties using either built-in or plug-in computational modules. It has a built-in 3D display for stereo visualization. Its plug-in programming APIs let users to add other microstructure-based 3D processing capabilities, or use it as a general purpose 3D simulation development environment. Additional information is available from the DCM homepage (http://research.csiro.au/dcm) and the Wikipedia (http://en.wikipedia.org/wiki/Data_constrained_modelling).

The DCM software comes in the following releases: DcmLite (free for non-commercial use), DcmEdu (Login or USB dongle), DcmPro (Login or USB dongle). DcmLite is a 32bit application and is feature-limited. The fully featured DcmEdu and DcmPro are 64-bit applications which can handle large datasets and can use Microsoft CPU clusters to speed-up data processing tasks. The Login releases require an online account at http://ts-imaging.net, and the USB releases require a hardware dongle issued by CSIRO. The latest release of the DCM software, datasets and plug-ins are available at https://data.csiro.au. This article is based on the DCM version 2.0.1, build number 1721.

The aim of this tutorial is to run through a case study on a synthetic CIPS (Calcite In-situ Precipitation System) sandstone sample, and to show how DCM can be used to extract and visualize various information, such as 3D distribution of compositions, the connectedness of material clusters, determine various material properties and to run simulations such as fluid flow and electricity conduction on the sample. This tutorial assumes that readers are familiar with quantitative X-ray CT, although some basic considerations are presented in selecting the X-ray beam energies. DCM APIs and the plug-in framework enable it to be used for other purposes, such as writing 3D programs for running other calculations and simulations on the data, as well as being a 3D visualization platform for non-CT related processing. These aspects however will not be covered in the present tutorial.
DATA INPUT/OUTPUT AND SAMPLE SETUP

The DCM software accepts two types of input data format, these being in TIFF format and those as outputted by the CSIRO developed CT-reconstruction software called X-TRACT. In both cases, the data are a stack of 2D CT slices. Upon opening the DCM software, a default new window appears which includes a Void channel by default, which can be redefined if desired. The first step is to define the grid size for the input data, i.e. the sample space. The grid size is the dimension of the 2D slices by the number of slices, and is set by right clicking on the DCM object name as per Figure 1. For the CIPS-sandstone sample, there are 100 slices, each being 120×110 in size.

Creating Tomography Channels and Importing CT Data

CT datasets acquired at different energies are imported by first creating an X-ray CT Data channel for a specified energy. This is done by right clicking over the DCM object name and selecting “Insert X-ray CT Channel”. In the subsequent dialog box that appears, click “ok” to accept the default channel name of “ct_1”. This new “ct_1” channel will appear on the left panel. Right click over it and select “X-ray CT Properties”. From within the dialog box which appears, the X-ray beam energy can be entered (the default is 20keV). Please note that monochromatic or polychromatic beam types can be specified, as per Figure 2. The type of an X-ray dataset can be specified as either the linear attenuation/absorption coefficient (LAC) or the linear phase-shift coefficient (LPC), depending on the physical values represented in the CT dataset. For the CIPS-sandstone case study, monochromatic 35keV and 45keV LAC datasets are used. Updating the “X-ray CT Properties” dialog box so that monochromatic energy is 35keV, sees the “ct_1” channel name updated to reflect this change. The process of adding another X-ray CT Data channel can be repeated for the 45keV dataset.

After creating the two monochromatic X-ray CT channels, they now appear as “LAC Mono_35keV” and “LAC Mono_45keV” respectively on the left window pane. The CT datasets can now be imported by right clicking over the CT channel name and selecting “Import Data” followed by “Image Files”. The “Import Data” dialog box will now appear. Using the browser, locate the directory containing the 35keV data as downloaded from the CSIRO Data Access Portal, and select all the 100 CT slices, which in this example are in floating point TIFF format (see Figure 3 - left).

In this example, the pixel data represents physical values β̃, which is the imaginary part of the complex refractive index (n = 1 – δ + iβ̃). As DCM calculation uses LAC, β̃ should be multiplied by a factor of $E^4\pi10^7/12.396$ to give ρ (cm$^{-1}$), where E is the X-ray beam energy in keV. This is illustrated in the “Rescale file data” check box under the “CT channel-specific Settings” in the Import Data dialog box.
The “Channel-common Settings” section also gives provision for rescaling the CT data with an additional constant multiplier and a zero offset. This is to compensate for any data bias arising during the experiment and CT reconstruction.

FIGURE 2. Adding X-ray CT data channels (top). The X-ray CT properties dialog (bottom).

Although not part of this tutorial, there may be situations in which the CT dataset is in fixed point format, whereby the pixel data does not represent physical properties, rather the CT image slice is simply a raster image of varying pixel intensity. In such a case the pixels take on integer values, e.g. between 0-65535 for a 16 bit resolution image. When importing fixed-point images, a known material phase in the CT slice should first be identified, and based on the value of the attenuation coefficient at the specified energy for that phase, the slices can be rescaled to a full-scale image by manually entering the upper and lower bounds in the “Manually enter Total Linear Absorption Coefficients” section of the dialog box, with “image value 0.0” and “image value 1.0” corresponding to the lower and upper bounds, respectively.

At this point, it would a good idea to save the changes by following the main menu “File”, “Save As”. The file name can be entered as “CIPS-sandstone”. The saved file extension is “.dcm”.

020029-3

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 121.219.45.4 On: Sat, 12 Mar 2016 06:37:01
Creating Compositional Channels

Following the creation and import of the CT data, it is time to create the compositional channels. DCM will use the multi-energy CT data to generate the volume fractions of each composition, and they are stored in the corresponding compositional channel. Compositional channels are created by right clicking over the DCM1 (CIPS-sandstone) root label and selecting “Insert Composition Channel” and then either “First” or “Last”. The default composition label will be “comp_1”. By right clicking over “comp_1” the “Composition Properties” can be modified (see Figure 3 - right), and the chemical formula and composition density should be entered. For the CIPS-sandstone sample, three composition channels should be created, being: void, quartz (SiO$_2$, 2.65 g/cm$^3$) and calcite (CaCO$_3$, 2.71 g/cm$^3$). Once complete, clicking “OK” will set these composition parameters and the default label will be automatically changed to “SiO$_2$” etc. The process can be repeated for the CaCO$_3$ material phase. There should now be two X-ray CT channels and three composition channels (including Void). After DCM computation, every voxel in the sample will be assigned values representing volume fractions (0 to 100%) of each composition.

Generally, the number of compositions should be no more than two or three plus void. Otherwise it would be too sensitive to CT data noise. Multiple compositions with similar X-ray attenuation properties should be combined to form a combined composition. The combined formula is a weighted sum of the constituency formulas. The weight for each composition is a product of its volume fraction in the combined composition, density and inverse atomic weight relative to the total atomic weight of the combined composition.

X-ray Attenuation Properties

The theoretical attenuation coefficients of compositions can be viewed, either individually or as a collective. The two composition channels can both be selected by holding down the SHIFT key and clicking on both the SiO$_2$ and CaCO$_3$ channels. Once highlighted, right clicking over one of the channels and selecting “X-ray Attenuation Properties” reveals the dialog box that allows the attenuation, phase-shift and X-ray penetration properties for each composition to be examined as a function of X-ray beam energy. The energy range, intervals and various parameters can be changed, and are updated by clicking “Recalculate”. The values can also be plotted by clicking “Show
Graph”. From the plots of absorption coefficient vs energy in Figure 4, the quartz phase exhibits K absorption edge at around 1.84keV, however in the energy range between 35-45keV there are no absorption edges present.

Ideally, it is best to conduct CT experiments at X-ray energies that are on both sides of absorption edges to maximize independence of different X-ray CT datasets. However there are practical limitations which may not make this possible. For example, although conducting CT experiments at 1.7keV and 1.9keV for a pure quartz sample would maximize datasets independence, the sample would have to be very thin for those X-rays to penetrate through it. From the “X-ray Attenuation Properties” dialog box, penetration through the sample can be calculated and it is found that such a sample should be no thicker than a few micro meters, making it unpractical. For X-ray energies of 35 and 45keV, the sample can be in the order of millimeters, making it much more suitable for practical experiments.

![Graph showing absorption coefficient vs energy](image)

**FIGURE 4.** Analyzing the X-ray attenuation properties for the material compositions.

The “normalized to channel” check box is used to calculate and plot the ratios of the X-ray attenuation or phase-shift coefficients for different compositions. A larger slope indicates more linear independence for different X-ray beam energies. This is to help users in selecting X-ray energies to maximize datasets independence. More sophisticated tools are built-in as well for optimizing the CT energies. This enables the right energies to be selected prior to carrying out a CT experiment, resulting in the optimal material phase prediction. Under certain circumstances, numerical phantoms may be used to evaluate the noise sensitivity and accuracy of composition reconstruction with different beam energy choices and other experimental conditions.

**Applying a Boundary Mask Layer**

Boundary channels can be created to obtain information from a region of interest. When applied to the DCM calculation, the statistical and volume data is defined for DCM predictions lying within the boundary region. Right click over the DCM root label and select “Insert Boundary Channel”, after which the default “bound_1” channel will
appear on the left window pane. Right clicking over “bound_1” and selecting “Build Boundary” opens up the build dialog box. Boundaries can be created based on various geometries, i.e. cubic, cylindrical and spherical, as well as a sample-based boundary where an algorithm calculates the surface of irregularly shaped samples. The boundary type is selected from the “Operation” section using the “Type” drop down menu (Figure 5). For the standard geometric types, the dimensions can be specified along with rotation and translation. For the “sample-based” boundary type, the X-ray CT channel(s) on which the calculation is based can be specified together with various other logical operands. After clicking “Apply” and “OK” the boundary is generated. The simple geometric boundaries are generated almost instantly whereas the “sample-based” based boundary is far more computationally expensive and therefore requires more time.

Right clicking over the boundary channel reveals further options, such as “Transform Boundary” where a previously defined boundary can be rotated and translated, and “Active for Computation” which as the name suggests, sets the boundary active for subsequent DCM calculations.

**FIGURE 5.** Creating a boundary mask layer dialog. Cylindrical standard geometry (left). Sample based (right).

**DISCRETE LEASE-SQUARE SEGMENTATION AND LINEAR OPTIMIZATION**

Having imported the CT data, defined the sample compositions and created a boundary, the DCM calculation can be initiated. From the “Computation” menu, select “Setup Computation Method” and a dialog box appears (Figure 6). Ensure the “Compute Composition” is checked and the others i.e. “Boundary Mask”, “Plugins”, and “Compute Property” (if it appears), are all un-checked. The “Compute Composition” section is for DCM computation of material compositional distributions. Within it are three computational modules: 1) Discrete least-square segmentation, 2) Data constrained linear optimization; and 3) Data constrained non-linear optimization (discussed in next section). The three can be used independently of each other, and their order of execution can be altered by entering different values in the edit controls (default 3) left to the names of the computational modules. Clicking on the “Set Parameters” button for each method displays the various computational options for each respective method.
Figure 7(a) shows the “Set Parameters” dialog for the “Discrete least-square segmentation” module. This is a form of binary segmentation. It works by minimizing the difference between the CT-derived and the theoretical attenuation coefficients for a discrete material composition for every voxel. From the dialog box, the “Zero Offset” and “Multiplier” options allow a linear transformation to be applied for each composition and energies, enabling fine user adjustment. The parameters can be set according to those in Figure 7(a). The remaining parameters should be left at the default values for the purpose of this tutorial. To carry out the “Discrete least squares segmentation”, click OK and then from the “Computation Menu” select “Start Computation”. Once the computation is completed, the compositional channels will be populated with the computed results, ready for visualization.

Visualization of datasets is achieved by clicking over the folder icon that lies to the left of a channel name. Each channel represents a set of 3D data. A green “tick” symbol will appear over the folder icon to indicate that it is active in the display (Figure 8). Multiple channels can be viewed simultaneously, and individual slices may also be displayed by expanding an individual channel by clicking the crossed box to the left of the channel name, then subsequently clicking over the icon next to the slice number of the relevant slice or slices. The display properties of a channel or multiple channels can also be modified by right clicking over the channel name (or selecting multiple
channels by holding the SHIFT key) and selecting “Display Properties”. A dialog box will appear in which numerous display properties can be modified, such as colours, opacity, options to display channel statistics such as average composition value etc., and also an option to limit display to a boundary mask (Figure 9(a)).

The channel display colour can be changed by clicking the “Data Colour” button, another dialogue is open as shown in Figure 9(b). For volumetric data such as the composition volume fractions, the “Single Colour” option should be used. The opacity of the display is related to the underlying volume fraction. For some physical properties such as the fluid speed or electric current, it may be more appropriate to use the “Multi Colour” option. For discrete data values, such as the calcite clusters, “One Colour per Data Bit” or “One Colour per Data Value” option should be used. Other controls should be self-explanatory.

**FIGURE 8.** Channel display control. (a) Mouse control of channel display. (b) A DCM display showing the output of the least-square segmentation.

**FIGURE 9.** Channel display colour control. (a) A display properties dialogue. (b) Data colour selection dialogue.
The output of the DCM least squares prediction can be displayed as shown in Figure 8(b). The DCM predicted compositional data can be directly accessed through the Void, SiO2 and CaCO3 composition channels. Additional display features can be accessed using the DCM “View” menu.

The “Discrete least-square segmentation” is limited in that it is a binary segmentation technique. It is not adequate for voxels with multiple compositions. The “Data constrained linear optimization” represents the first of the main DCM computational methods. It is based on assumptions that the total volume of the voxel is the sum of volumes of individual materials (including void) in the voxel; and the total attenuation of an X-ray beam by the voxel is the sum over the attenuations by the individual materials in the voxel, as described elsewhere.

From the “Setup Computation Method” dialog box (Figure 6), ensure that only the “Data constrained linear optimization” checkbox is selected. Click on its associated “Set Parameters” button and assign the parameters to the values shown in Figure 7. Once complete, click OK and then from the “Computation Menu” select “Start Computation”. Figure 10(a) is the display after the “Data constrained linear optimization” is completed. Users are encouraged to explore the effects of changing the various values in “Set Parameters” for both modules. Quite often it is useful to display the relative amounts of compositions. This can be done by simply highlighting the three compositional channels and selecting “Display Properties”. Once the dialog appears (as in Figure 9) the checkbox for “Average” can be selected in the “Data Statistics” section at the bottom right of the dialog.

![Figure 10](image_url)

**FIGURE 10.** (a) Output from the Linear Optimization module for the CIPS-sandstone sample. (b) Output from the Non-Linear Optimization module for the CIPS-sandstone sample

The discrete least-square segmentation and data-constrained linear optimization methods have been used in microstructure characterizations of various material samples, including aerospace paint primers, oil & gas reservoir rocks, and various other materials.

**NON-LINEAR OPTIMIZATION**

The data-constrained linear optimization is capable of reconstruction of microstructures with partial volumes of compositions in the same voxel. It is computationally efficient. However, it does not adequately account for the interactions among compositional phases. For instance, as shown in Figure 10(a), incorrect co-existence of calcite and pores in the quartz grains is evident. The data-constrained non-linear optimization method offers a more accurate control of such interactions. The parameters are entered through the data-constrained non-linear optimization dialogue as shown in Figure 11. The parameter-dialogue is opened by selecting the main menu “Computation”, “Setup Computation Method”, and clicking the “Set parameters” button next to the “Data-constrained non-linear optimization” label.
The lower half of the dialogue specifies the interaction matrices. As for the CIPS sandstone, it is known that the quartz crystals are dense. Their sizes are typically much larger than the voxels, i.e., they are unlikely to mix with the void or calcite in the same voxel. Therefore, large repulsive interactions in the same voxel (intra-voxel) should be specified between quartz and void, and between quartz and calcite. Under the “intra-voxel” tab, a value 1 is used for both interactions between quartz and void, and between quartz and calcite. Also under the “intra-voxel” tab, small repulsive interaction value of 0.015 is specified between calcite and void. Between the nearest neighbour voxels, a small attractive interaction value of -0.015 is specified indicating that it is more likely that the neighbouring voxels would have similar compositional distributions. These small repulsive interaction values between different compositions in the same voxel and small attractive interactions between neighbouring voxels are used to reduce the noise effect in the reconstructed composition maps. As these values are at or below the noise level of the reconstructed CT slices, there is no strict rule about their selections. It is “arbitrary” to choose interaction values at or below such noise level. The relative noise level of a typical “good” CT data set would be at the order of 1%. These small interaction values could be used as adjustable parameters to “fit” the known overall porosity value and the overall percentage of other compositional phases. This bears a degree of similarity with the threshold image segmentation technique that the arbitrary threshold values are adjusted such that the segmented image looks “right”. As it is known in statistical mechanics, a higher energy state would have a lower probability according to Boltzmann distribution.

![Data-constrained non-linear optimization parameter dialogue.](image)

**FIGURE 11.** Data-constrained non-linear optimization parameter dialogue.

More comprehensive discussions about the parameters and applications of the data-constrained non-linear optimization are published elsewhere\(^\text{14, 16, 17}\). It should be pointed out that the selection of the “Grid size”, “No. of
iterations” and the “Neighbour interaction range” would have a significant effect on microstructure reconstruction accuracy and computational CPU time. The data-constrained non-linear optimization module is CPU intensive. For the example quoted in this article using the parameters as in shown above, it would take an hour or more on a desktop computer. For a moderately large data set, it is advisable to start with the discrete least-square segmentation and data-constrained linear optimization to make the DCM predicted composition looks “approximately right”. Then a sub-volume of the sample is used to do the initial parameter adjustments for the data-constrained non-linear optimization. Once the correct parameters have been selected, they can be applied to the whole sample to complete the computation.

For users of DcmPro and DcmEdu, computation speed can be increased by using Windows HPC clusters. The cluster computation can be enabled by check the “Use Cluster” check box. The CPU intensive part of the computation can be executed in the cluster and the results are displayed on the desktop. The process is transparent to the user. Before using a cluster, some once-off initial setup is required. This can be accomplished by following the main menu “Edit”, “Preferences …”, and clicking the “Cluster” tab. The cluster computation feature is not available with DcmLite.

Once the parameters are set and saved, computation of the compositional microstructure can be started by following the main menu “Computation”, “Start computation”. The computation can be interrupted by following the main menu “Computation”, “Stop computation”. The progress of the computation is displayed on the bottom bar of the main DCM interface. The computed microstructure is shown in Figure 10(b). The result shows that quartz in the sample exists as pure grains. In contrast, calcite distributes either as porous coatings on the quartz grains, or resides between quartz grains with varying levels of porosity.

**PROPERTY ANALYSIS AND USER PLUG-IN DEVELOPMENT**

Some microstructure properties can be calculated using DCM built-in functions. To use these functions, at least one property channel must be added. To add a property channel, right-click the DCM object name “CIPS-sandstone”, and select the “Insert Property Channel” pop-up menu as shown in Figure 12(a). Rename the property channel to “Calcite clusters”. Once the property channel has been added, follow the main menu “Computation”, “Setup Computation” to open the dialogue as shown in Figure 12(b).

In the “Setup Computation” dialogue, drag the “Linked Cluster Analysis” to the right. A parameter dialogue will be opened as shown in Figure 13(a). Select the input channel as “CaCO3” and the output channel as “Calcite.
clusters”. The “Normalised threshold” can be set to 0.5. Run the computation by following the main menu “Computation” and “Start Computation”. Setup “Calcite clusters” display color as “One Colour per Data Value”; the channel display should resemble Figure 13(b). Details about clustering analysis and applications are published elsewhere\textsuperscript{18-20}.

![Figure 13](image1.png)

FIGURE 13. Linked cluster analysis. (a) Setup linked cluster analysis parameters. (b) Calcite linked clusters.

To add additional modules for calculation of other physical properties, various plug-in modules can be added. A DCM plug-in is a user-written computation module that can be scheduled to run with other built-in DCM computation modules or with other plug-ins. A number of plug-in modules are available at CSIRO Data-Access-Portal (https://data.csiro.au), which simulate fluid flow in the pores\textsuperscript{21-23}, and electrical properties\textsuperscript{24}.

DCM users can develop their own plug-ins to perform virtually any desired function, not necessarily associated with materials microstructure and properties. The advantage of developing a custom algorithm as a DCM plug-in over writing a standalone program is that the plug-in environment provides immediate access to all the DCM features including 3D interactive visualization and animation, high level 3D data manipulation, memory management and data input and output.

The plug-in development environment is installed on the user’s machine as a default option at DCM installation time. A prerequisite for developing DCM plug-ins is a moderate knowledge of C or C++ programming. Plug-in templates and examples are provided which can be modified or copied to produce new plug-ins. A plug-in is built from its C/C++ source code as a binary .dll file so users can share their developed plug-ins with others without needing to distribute source code.

DCM users intending to develop their own plug-ins should install Microsoft Visual Studio. The Visual Studio Express edition is freely downloadable from Microsoft. The DCM on-line help contains more information about using and developing plug-ins.

ACKNOWLEDGMENTS

The DCM software development is supported by CSIRO CSS-FSP and eResearch initiatives.
REFERENCES