

Cell Growth Simulation using SAS® Software

Karen Walker, Experis, Chandler, Arizona USA

ABSTRACT

This paper will showcase the number crunching power of SAS 9.4®. It's back to basics, with an updated twist, as we will explore the power of the CAMERA interface coupled with SAS® PROC S. Building files read into SAS 9.4®, this discussion seeks to uncover broader technological trends by integrating visuals, with the metrics power of SAS 9.4®. Our objective for this paper is to AUTOMATE the process of visualization for cell growth in a way that inspires us.

INTRODUCTION

In the not too distant past visual communication designs were limited to inform and influence only, never to impress.

Historically, font, and color were used only if available and under strict guidance. For visual communications, standardized reports were "black and white". SAS procedures like **PROC G3D** built to show space relationships between 3 or more variables had no valid justification. Dose any of this remind you of film and television in the 60s? If so, then you'll see where we're headed with this paper.

Visual communication based on observations, experiences, and research is proven to be most effective. As now special effects are no longer just for motion pictures. Special effects are used on a worldwide basis in business intelligence correspondence to help convey all aspects and properties of the data. The internet and web based publishing have made way for the "**Visual Sound bite**" often referred to as the new standard "Visual Sound Bullet".

This paper will focus on the use of SAS GRAPH® and SAS BI WEB SERVICES to develop highly productive visual communications. These illustrations will show accurate accounts of the relationships between entities even where there are infinitely many variables. As a rule, the best visual business intelligence solutions are indelibly etched in the brain, and at the same time are brief and are simple to understand.

In this paper we'll use SAS 9.4® Visual Communications by developing a SAS GRAPH® program for Cell Growth over time. We'll exploit SAS BI WEB SERVICES techniques to enhance the graph on the internet where the image will move and change across time.

We'll have the chance to look at both XML and JSON formats while uploading content to the WEB. Then accessing web services using PROC HTTP from BASE SAS® 9.4, PROC SOAP, and PROC JSON for downloading content from the WEB; where we have access permissions to do so.

Thanks for attending this session.

MATHEMATICAL BACKGROUND

Simulation is finding a solution once, and then adjusting parameters to see what happens to that solution over time.

Let's set our scene by programming two cells. We're going to chart their growth over time by noticing the frequency and size of each cell. We know from books on how to simulate biological cell growth that the rules that govern cell growth are often RANDOM. So the variables for each molecule on the cells are set equal to a variation of the RANNOR function over time.

Now let's consider our two cellular families as we vary time frequency and size not randomly, but instead by what we've discovered through empirical observation, experience, and research. After review of the most interesting genome sequence analysis, protein structure investigation, and gene expression data analysis. We've found the continuous wavelet transformation (CWT) equation meets our constraints.

This is given by the following:

$$CWT(f,a,b) = a^{*-1/2} \int_{-\infty}^{\infty} f(t) \psi (t - b / a) dt$$

Where ... ψ is the "Mother WAVE", a (the scale parameter) >0 , b (the translation parameter), t (is time), and dt is (time's derivative). The CWT maps a one-dimensional signal to a two-dimensional time-scale joint representation. It is calculated by continuously shifting a continuously scalable function over a signal and calculating the correlation between the two.

Cell Growth Simulation using SAS® Software, continued

PREPARATION

All charts, plots and files are created with SAS® 9.4 running WINDOWS 10. Before we can make a movie from the .GIF files that SAS GRAPH® creates, we must learn how to process those files.

Transporting files/ files suffix on your PC:

1. Write files using FILENAME statement in SAS, then INSERT into Microsoft PowerPoint.
2. File suffix includes: GIF, PNG, or XML to name a few.
3. Use WINDOWS utilities (e.g. PAINT, etc..) to convert files to JPEG

MOV files are captured from CAMERA files or on Social Media for example. Then can be loaded to Microsoft PowerPoint in the same manner as PNG, XML, or JPEG files.

SAS® BI WEB SERVICES to convert and UPLOAD files to XML with XMLMap utilities provided by SAS® 9.4.

- Use **PROC JSON** new to SAS® 9.4.
- Use SAS® 9.4 **XML V2 LIBNAME Engine**

SAS® XML Mapper can be best described by looking at the XMLMap files below. Specific information is requested from a WEB SITE it will be returned in XML or JSON formats.

JSON

```
proc json out="output\cellsOut.xml" pretty nosastags;
  export work.labs;
run;
```

The NOSASTAGS removes the metadata tags that are not needed for our purpose, and the PRETTY option makes the JSON MAP file easier to read.

```
[
  {
    "date": "09/21/2017"
    "summary": "Proteins",
    "maxpre": 342,
    "minpre": 0.1725
  }
]
```

XML

```
Libname cellgrowth xmlv2 "output\cellsOut.xml";
data cellsOut;
  set work.labs;
run;
```

The XML code looks like this and can be presented on the internet, and also read back into SAS® without an XMLMap file.

```
<?xml version="1.0" encoding="windows-1252" ?>
<TABLE>
  <CELLPROTEIN>
    <date>2017-09-21</date>
    <summary>Proteins</summary>
    <maxpre>342</maxpre>
    <minpre>0.1725</minpre>
  </CELLPROTEIN>
</TABLE>
```

Cell Growth Simulation using SAS® Software, continued

SAS CODE FIGURE #1

As mentioned before, we're going to chart cell growth by simulating protein molecules for hair and skin growing over time on a series of scatter plots. We know from books on how to simulate biological cell growth that the rules that govern cell growth are often RANDOM. For this reason we'll build a data set WORK.LABS that contains protein molecules; their counts as they grow over time of 6 days. We'll find the growth rate for the protein molecules using the RANNOR function. Then render the data using SAS GRAPH® SGPANEL.

LABs variable	Description
l	Iterations from 1 to 100
visitnum	1=1day,2=3days, 3=6days
labtest	Protein molecule
cell	Molecule type hair or skin
pre	$\min(4, 2.5 * (\text{abs}(\text{rannor}(123)))) / 3.0$;
result	$\min(4, 2.5 * (\text{abs}(\text{rannor}(345)))) / 3.0$;

Table 1. The Data Definitions for RANDOM WORK.LABS

Program 1 shows how SAS CODE is built for SGPANEL

```
proc sgpanel data=labs;
  panelby labtest visitnum / layout=lattice onepanel novarname;
  scatter x=pre y=result / group=cell markerattrs=(size=9);
  refline 1 1.5 2 / axis=Y lineattrs=(pattern=dash);
  refline 1 1.5 2 / axis=X lineattrs=(pattern=dash);
  rowaxis integer min=0 max=4 label=' ';
  colaxis integer min=0 max=4 label=' *';
  keylegend / title=" " noborder;
run;
```

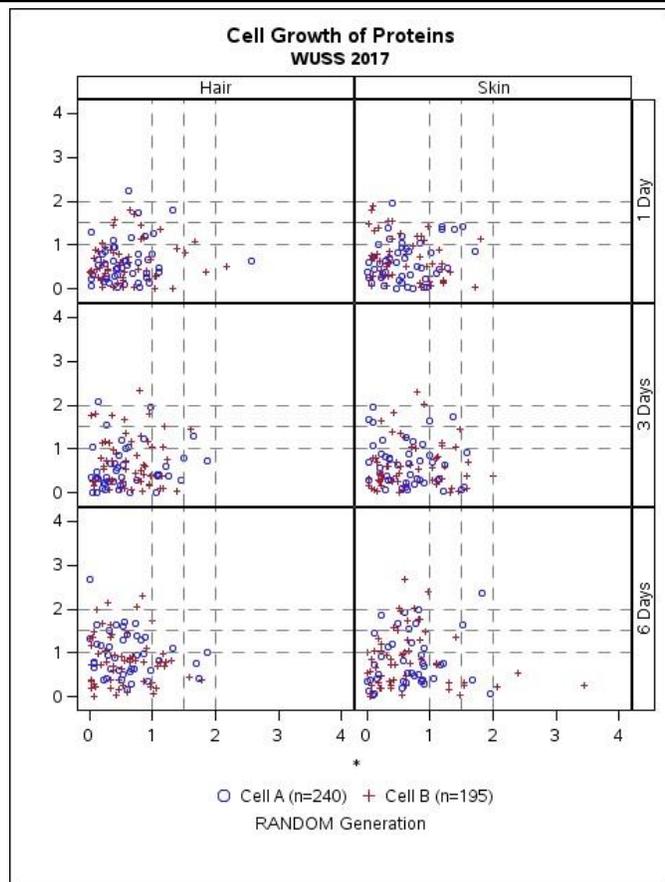


Figure 1. Scatter Plots output SGPANEL

Cell Growth Simulation using SAS® Software, continued

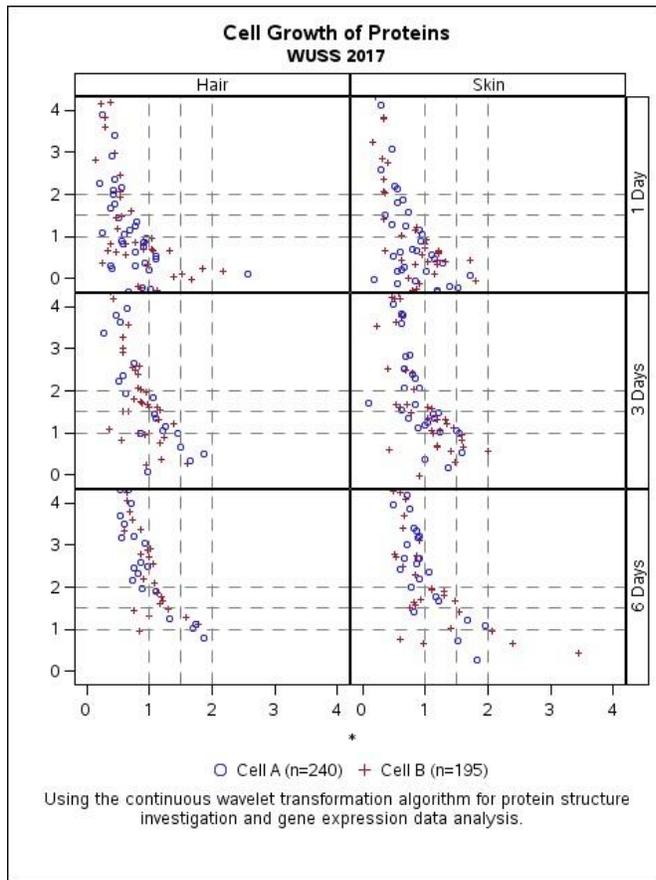
SAS CODE FIGURE #2

In contrast we'll next simulate our WORK.LABS dataset data by charting the cell growth over time using values that satisfy our given equation for modeling the data for CWT: Continuous Wavelet transformation. I must mention that since I've applied limits on the dataset it is more accurate to say DISCRETE Wavelet Transformation (DWT) instead. Since this is a "Visualization Presentation" not a statistics one we focus on the graphs for today.

LABs variable	Description
l	Iterations from 1 to 100
visitnum	1=1day,2=3days, 3=6days
labtest	Protein molecule
cell	Molecule type hair or skin
pre	Mother wave value for cell
result	DWT value

Table 2. The Data Definitions for DISCRETE WAVELET TRANSFORMATION WORK.LABS

Figure 2. Scatter Plots output SGPANEL



Imagine that we've separated our graph of Cell Growth Proteins into 6 individual plots, 3 for hair, and 3 for skin. The simulation is accomplished as we play "PROC GREPLAY" plots over time (visitnum) 1, 2, and 3. On other words, we set up a LOOP to cycle across visitnum, to display each graphic image. Our simulation using the WORK.LABS data table is complete, true cell growth based upon data can be captured as an XMLmap file for display on the internet using an XML LOOPING routine in the same manner.

Cell Growth Simulation using SAS® Software, continued

Program 2 shows how SAS CODE is built to make the GRAPHS move over time

```
libname graphcat '/sas/sas942/sasgraph'; *-- CREATE CATALOG;
symbol interpol=join;
* Create plots, put in the GRAPHCAT.PROTEIN catalog;
* Give the plots names of HAIR and SKIN each for visitnum 1,2,3;
proc gplot data=work.labs uniform gout=graphcat.protein;
  where visitnum = 1 and cell = A;
  title '';
  plot result*pre / name='Hair';
run;
<REPEAT for all PLOTS, by visitnum, cell >
...
proc greplay igout=graphcat.PROTEIN; run; quit;
```

As we can accomplish this with .GIF, and PNG files, we can also perform the same method on images from actual photos of cells. Let's explore how to make our images move using all that's available to us.

Using PROC HTTP and PROC SOAP we can capture images from the public domain. A word of caution, please don't download and use images that are the property of someone else without their permission.

HTTP

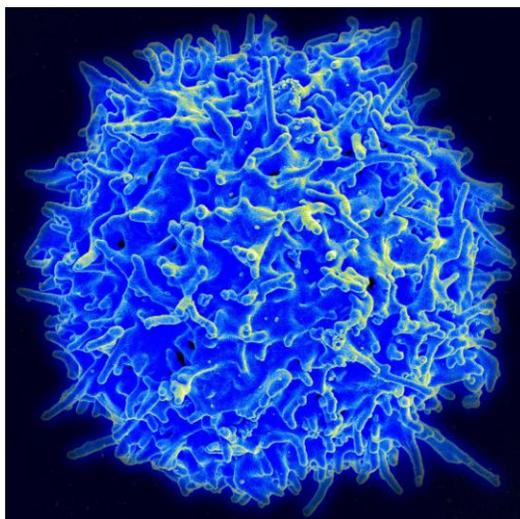
To use WEB SERVICES to capture/download content from the internet use PROC HTTP introduced on SAS® 9.2. This **HyperText Transfer Protocol** interface requires that we know the **URL** of the content we're looking for. It is often used in conjunction with the XMLV2 LIBNAME engine introduced in SAS® 9.3.

The HTTP procedure is used to obtain the T Cell image file from the National Institutes of Health(NIH) database.

```
Filename TcellImage "output\Tcell.png";

proc http out=TcellImage
  url="https://www.nih.gov/sites/default/files/styles/featured_media_breakp
oint-large-extra/public/news-events/news-releases/2017/20170727-t-
cell.jpg?itok=wg7ziCQ&timestamp=1501164766"
  method="GET"
  ct="application/x-www-form-urlencoded";

run;
```



Display 1. T-Cell from NIH

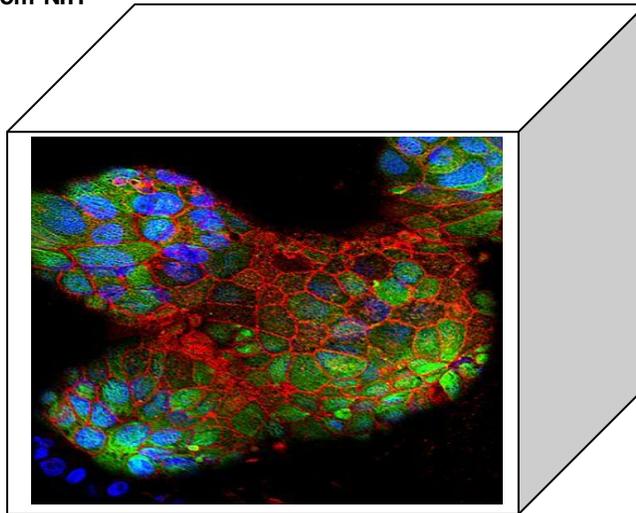
Cell Growth Simulation using SAS® Software, continued

SOAP

To use WEB SERVICES to capture content use PROC SOAP introduced on SAS® 9.2. **This Simple Object Access Protocol** web service allow us to send a request for data from a variety of formats. Each format depends on the web service targeted. SOAP uses messages built with Web Services Description Language (WSDL) in a file. It is often translated using XML. Notice that a proxy host name and port number is required to communicate with most web services. Consult your BASE SAS Procedures Guide for more information. Here's an example of a call to SOAP.

```
proc soap in=request
          out=response
          url="https://pubchem.ncbi.nlm.nih.gov/pug_soap/PUG_SOAP.html/SOAP_server.
          php"
          soapaction="http://pubchem.ncbi.nlm.nih.gov/GetDownloadUrl"
          proxyhost="your host name here"
          proxyport=usually 80
;
run;
https://www.nih.gov/sites/default/files/styles/floated_media_breakpoint-
large/public/news-events/research-matters/2016/20160315-
placenta.jpg?itok=AEvutdc2&timestamp=1482332428
```

Display 2. 3D Cells in Placenta from NIH



JSON

To use SAS® BI WEB SERVICES to EXPORT SAS datasets use PROC JSON introduced on SAS® 9.4.

It's easy. Try this.

```
proc json out='C:\mywebfile.json' pretty;
write values 'Age 11';
write ioen array ;
write values 'Girls';
write open object;
export sashelp.class(where=((sex='F') and (age=11)));
write close;
write close;
;
run;
```

Cell Growth Simulation using SAS® Software, continued

MAKING THE MOVIE

Making movies from your SAS Graphics files gets complicated quickly. I recommend sticking to one format, file size, and suffix until you are comfortable with processing reports that show your simulations. Keep the special effects handy although. For it will soon be time to show off your science with flare. When it comes to porting your work you'll find it necessary to use the Cloud Foundry, or a Dropbox of some sort because simulation files are usually over 65K MB.

File Suffixes	Deploys to
WMV	Most all platforms
MP4	Most all platforms, MS PowerPoint
AVI	Smart Phones
MOV	EXCEL, WORD, MS PowerPoint
FLV	Open Source, Word Processing, Spreadsheet,
MKV	Kumbutu, YouTube
TS	YouTube
3GP	Unix Platforms
MPEG-2	Social Media sites
WEBM	Public and Private WEB Sites
GIF	SAS GRAPH all versions
MP3	Audio Players

CONCLUSION

As your need to deploy simulation content across the World Wide WEB grows, you'll need to know specific information. Use what you know about PROC JSON when porting SAS data to the internet, and PROC XML for porting internet content into SAS. Look for more about surfacing SAS data to the internet to come shortly.

Surfacing techniques can be found at SAS Business Intelligence WEB Services:
<http://support.sas.com/documentation/cdl/en/wbsvcd/62759/PDF/default/wbsvcdg.pdf>

"To laugh often and much, to win the respect of intelligent people and the affection of children, to earn the appreciation of honest critics and endure the betrayal of false friends, to appreciate beauty, to find the best in others, to leave the world a bit better, whether by a healthy child, a garden patch ... to know even one life has breathed easier because you have lived.

This is to have succeeded!"

- Emerson

Cell Growth Simulation using SAS® Software, continued

REFERENCES

- Aldroubi,A and Unser, M. (eds) (1996) Wavelets in Medicine and Biology, CRC Press, Boca Raton, FL.
- Vannucci,M and Lió,P (2001) Non-decimated wavelet analysis of biological sequences: applications to protein structure and genomics . Sankhya B, 63, 218 - 233.
- Lió , Peter. (2003). "Wavelets in bioinformatics and computational biology: state of art and perspectives" Bioinformatics Review. Vol. 19 no. 1 2003 Pages 2 - 9. Department of Zoology, University of Cambridge, Cambridge, UK and European Bioinformatics institute, Wellcome Trust Genome Campus, Hinxton, Cambs, UK: Oxford University Press
- McNeill Bill. (2013). "Beyond the Basics." *Proceedings of the SAS Global 2013 Conference*. Gary North Carolina; SAS Institute
- Musitano, Neal Jr. (1999). "User Experience SAS/GRAPH® Charts & Plots Using MXG®." *Proceedings of the SAS Users Group International 1999 Conference*>. Gary, North Carolina: SAS Institute

http://support.sas.com/sasamples/graphgallery/PROC_SGPANEL.html Support : SGPANEL Program Examples . "Sample 44025" . 08/17/2011. Available at <http://support.sas.com/kb/44/025.html>.

ACKNOWLEDGMENTS

So many have been supported me, at this time I want to thank Lyn Fenex at Experis, along with all of Experis. And Thank you Jesus for all that you are!

RECOMMENDED READING <HEADING 1>

- Base SAS® Procedures Guide
- SAS® XML LIBNAME Engine: User's Guide
- SAS® For Dummies®
- <https://www.soapui.org/>
- The Little SAS Book

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Karen Walker
Enterprise: Walker Consulting LLC
Address: 6481 West Gary Dr
City, State ZIP: Chandler, AZ. 85226
Work Phone: 480.206.7196
Fax:
E-mail: kkwalker77@yahoo.com
Web twitter: KarenErneWalker

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.