SOCIETY OF ACTUARIES Technology Section

CompAct

ISSUE 33 | OCTOBER 2009

1 **The Actuarial Perfect Storm** *Sean Burns, Jonathan Houk, and Paul Kosakowski*

2 **Editor's Notes** *Howard Callif*

3 **Letter from the Chair** *Tim Pauza*

6 **Be More Productive with Everyday Technology** *J. Eddie Smith, IV*

9 **European Approaches to the Modeling Process Are Changing—What Can We Learn?** *Van Beach and Bruce Keenan*

11 **Introducing Twitter** *J. Eddie Smith, IV*

15 **Cool Tech** *Matthew Wilson*

21 **R Corner—Combining Cluster Analysis and Predictive Modeling** *Steve Craighead*

The Actuarial Perfect Storm

By Sean Burns, Jonathan Houk, and Paul Kosakowski

Actuaries are being hit with a perfect storm of simultaneous and unprecedented forces as economies throughout the world have been in a state of near collapse. Risk modeling practices of the past are no longer adequate. New solvency reporting requirements have emerged, involving running and then analyzing thousands of model scenarios, whereas in the past maybe only a few were needed. Compounding this, new process and control regulations, making spreadsheets and other shadow systems no longer valid options, are making actuaries rethink the way they do business. What can be done? Companies put themselves at risk of failure if they continue with "Business as usual", and in the current economic environment, hiring more actuaries to get the work done is not feasible.

As banks and insurers began to discover in 2007, Value At Risk (VAR) calculations were lacking in effectiveness as companies routinely found themselves outside the realm of acceptable risk tolerance. On top of this need for improved financial modeling of risk, is the need for greater maintainability of auditable trails. Auditable trails first became a topic with the 2002 Sarbanes-Oxley initiative which resulted from the accounting scandals of 2001. Insurers are sure to come under even more SOX pressure as the credit crisis of 2007 continues on. SOX pressure is multileveled as insurers must not only maintain auditable trails, they must implement a stringent process and control environment. Each year-end they must then sign off on the effectiveness of that environment. That's right; corporations need to be confident enough to sign off on the effectiveness of their environments. That's a daunting requirement that can send shivers down the spines of compliance officers throughout the industry. Fortunately there is a solution. That solutions lies in the latest advances in Information Technology.

Today's information technology enables actuaries to take financial modeling farther than ever before. At the same time, Business Intelligence, a subsector of Information Technology, can greatly alleviate the pain of

CONTINUED ON **PAGE 4**

Actuaries
Risk is Opportunity.®

CompAct

ISSUE 33 OCTOBER 2009

Published quarterly by the Technology Section of the Society of Actuaries

2008-2009 SECTION LEADERSHIP

Tim A. Pauza
Chair
e: tim.pauza@ey.com

Timothy Deitz
Vice-Chair/People Coordinator
e: Tim.Deitz@omfn.com

Frank Reynolds
Spring Meetings Education Liaison
e: fgreynol@uwaterloo.ca

Holly L. Loberg
Web Coordinator/Actuarial Club Program
e: holly_loberg@allianzlife.com

Phillip Gold
Board Partner
e: Phil.Gold@ggy.com

Jeffrey Pomerantz
Annual Meeting
e: jeff.pomerantz@qrm.com

David Minches
Communications Coordinator
e: david.minches@ey.com

Carl Nauman
Scenario Manager/Webinar Coordinator
e: Carl.Nauman@ggy.com

Carl Desrochers
"Other" Meetings Coordinator /Speculative Fiction Contest
e: carl_desrochers@berkshirelife.com

Joseph Liuzzo
Secretary/Treasurer
e: jliuzzo@tiaa-cref.org

OTHER VOLUNTEERS

Howard Callif
Newsletter Editor
e: howard@callif.org

Mary Pat Campbell
Social Networking Advisor
e: marypat.campbell@gmail.com

FRIEND OF THE COUNCIL

Kevin Pledge
Council Advisor

Robert LaLonde

SOA STAFF

Sam Phillips
Staff Editor
e: sphillips@soa.org

Susan Martz
Project Support Specialist
e: smartz@soa.org

Meg Weber
Staff Partner
e: mweber@soa.org

Patrick Milani
Graphic Designer
e: pmilani@soa.org

Facts and opinions contained herein are the sole responsibility of the persons expressing them and shall not be attributed to the Society of Actuaries, its committees, the Technology Section or the employers of the authors. We will promptly correct errors brought to our attention.

Copyright © 2009 Society of Actuaries.
All rights reserved.
Printed in the United States of America

Editor's Notes

By Howard Callif

The year seems to be flying by! It's hard to believe the deadline for the October issue has arrived! We have had a lot of people step-up and provide articles on a wide range of topics. This edition has two articles on the importance of technology in the increasingly complex world of financial modeling. Both articles discuss how critical it is to embrace new technologies to successfully deal with Solvency II, financial reporting, and risk analysis. Please consider e-mailing me—or writing an article—on how your company is dealing with this strategic issue.

Some council members spoke up at the last section meeting and requested articles with an introductory slant for members who may not be familiar with some of the newer technologies available. J. Eddie Smith wrote two articles in response to this need, both very informative. The first article is a summary of shortcuts and ways to use existing technologies more efficiently. The second explains *Twitter*, which you have likely heard of, but may not be using. The article, "Introducing Twitter," will help you understand how to use this social networking tool, and why.

Thanks to the regular columnists, and to everyone who has submitted articles throughout the year—their work and contributions are appreciated! Please provide feedback on how you are using this information, and suggestions for additional articles. R Corner author Steven Craighead has also requested some feedback on the direction he should take in future articles. Please let us know whether you are finding the column useful, and what other information you would like to see. Emails I'd like to share (edited):

From: Francesco Balena [mailto:fbalena@codearchitects.com]

First of all, thanks for mentioning me and my VB Migration Partner software in your Editor's Notes column (April 2009).

You might find it interesting to have a look at the Resource section of our Web site, with tons of info useful to all developers involved in VB migration projects (regardless of whether they use our software). We have filled this section with virtually everything we learned while developing our conversion software, including the many subtle differences between the two languages (http://www.vbmigration.com/Resources/migratingfromvb6.aspx) and control libraries (http://www.vbmigration.com/Resources/migratingfromvb6controls.aspx). Most of this material isn't available anywhere else on the Internet.

EDITOR'S REPLY: Thanks for your feedback, and your contributions! ■

From: Weigel, Patrick [mailto:pweigel@Gafri.com]

Here's an addition to the list of languages from issue 31 (4/2009) for the article Cyber-Ballistic Podiatry:

Objective-C: Steve Jobs shoots you in the foot. And you like it.

EDITOR'S REPLY: Funny! Keep them coming! ■

Howard Callif, ASA, MAAA, is a product champion at Milliman IntelliScript. He can be contacted at howard@callif.org

Letter from the Chair

By Tim Pauza

Tim Pauza, ASA, is a manager at Ernst & Young, LLC. He can be contacted at tim.pauza@ey.com.

Greetings! It is hard for me to believe that the year is coming to an end. This will be my last "Letter from the Chair." It has been a pleasure serving the Technology Section as chairman. I am pleased with what we have accomplished this year, and excited about the many initiatives underway. We have a strong team with our new council and many dedicated volunteers that are entering 2010 with a lot of enthusiasm. Before I get into all of the good things going on, I want to pause to thank our retiring members: Joe Liuzzo, David Minches and Carl Nauman. It has been a pleasure working with Joe, David and Carl, and I wish them each the very best as they move on to other opportunities. Our new council members stepping in for their own terms are Lisa Lefkowitz, David Kester and Mary Pat Campbell. Along with Tim Deitz, our new chair, these folks are going to bring unique skills, ideas and energy to section leadership. Finally, I'd like to pay tribute to three members of the section who, in the true spirit of volunteerism have committed to supporting the council. To prove we are all winners when we volunteer to serve, Eddie Smith, Mark Press and Brian Reed have each agreed to take on dedicated roles as friends of the council for 2010. Please reach out to all of these folks with your appreciation for their service and with your ideas for how we can make the Technology Section a real value to you the members.

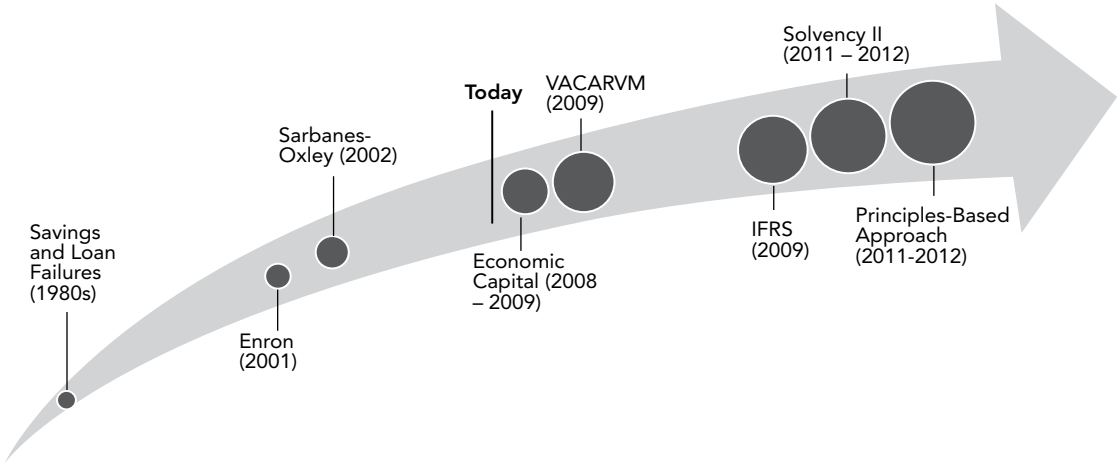
There is much going on, and I will not try to cover all the activity, but I will mention a few highlights from the past year that will continue to be major initiatives for 2010. As I mentioned in the last newsletter, we have implemented a private LinkedIn group for section members, "Society of Actuaries Technology Section." This online forum for networking and information sharing will focus on the challenges and interests of our members. As the technology demands grow for actuaries in an ever more complex and challenging economic environment, we believe LinkedIn can provide us an invaluable opportunity to keep abreast of the latest developments in arenas where technology is playing a major role in the future of our industry. Since my last report, the group has grown steadily to 84 members, and, more importantly, the activity and content have developed significantly. Nurturing this fledgling community will be a priority for the 2010 council, and I am excited about the potential here for the section. If you have not already done so, join LinkedIn, join our group, and see what your fellow section members are up to. (www.linkedin.com)

One more item I would like to highlight is a collaboration we have developed over the past two years with Gartner Group (www.gartner.com). We published a survey in 2008 to investigate the effectiveness of the collaboration between actuarial departments and information technology departments in the industry. The results of this survey were presented in a session at last year's annual meeting. The response we received from that session indicated a high level of interest in this important topic, so we repeated the survey for 2009. This year, we also partnered with the Financial Reporting section who felt their members had particular interest in this topic. In doing so, we were able to triple the number of responses. Hopefully, you will be able to attend our session in Boston this year (Session #34) where we will be discussing observations from the survey results, including trends from our finding of last year. If not, watch for an article by Steve Leigh from Gartner in an upcoming edition of CompAct.

With that, I will draw to a close. Thanks again to our volunteers for their willingness to get involved, and thanks to all of our council members and friends of the council who work hard to keep our section going! Got an idea for us? Let us know. Want to get involved? We can accommodate that too. Please reach out to any of the council members listed inside the front cover of this newsletter with your ideas or to volunteer.

Thanks and happy reading! ■

Tim Pauza, 2009 Section Council Chair



maintaining audit trails. Read on to find out more about the storm and how Business Intelligence combined with Information Technology can play a central role in establishing a business process that can enable actuarial organizations to thrive.

Up until 2003 a great deal of financial modeling was done in a deterministic fashion. As a result of the risk associated with variable annuities and other equity based insurance products that were exposed in the wake of the accounting crisis, it became evident that a more robust analysis of risk was required. Thus, in 2003 a push to a more stochastic assessment of risk began. Stochastic processes are extremely powerful as they allow for the testing of multiple future scenarios allowing for random elements of risk. The only limit to stochastic modeling is the power of the IT infrastructure it is running on. In 2003 it became apparent that the current environments in place were not capable of supporting the multifaceted ability of stochastic modeling. The world of scalable high performance (HPC) computing environments began.

One benefit of stochastic modeling for financial reporting is that one can forecast distributions of results for various points in the future. Adding to this is the ability to create stochastic variation within stochastic variation. Theoretically, a stochastic process can forecast an indeterminate number of scenarios. With the right high performance technology solution in place actuaries can develop and run models capable of forecasting thousands of scenarios utilizing millions of data points. In a high performance environment these models can run extremely quickly allowing the actuary more time to review the output. In fact, in a high performance environment, modeling tools can deliver distributed processing across multiple processors and PCs. This growth of infrastructure resources and software functionality has allowed actuaries to manage demand for more scenario runs and process more model points with increasing model complexity. But how can actuaries support a high performance environment consisting of multiple machines and multiple servers whereas in the past they were able to maintain software releases and configurations on single PCs? The answer here is to form an alliance with their IT organizations. With this alliance in place actuaries will no longer have to worry about the functionality of their environment and can instead focus their efforts on the daunting task of risk management.

SARBANES-OXLEY

The Sarbanes-Oxley Act is adding to the pressure of more accurate financial risk forecasting by requiring in-depth auditable trails. One section that has specifically impacted the actuarial organizations is section 404 pertaining to Management Assessment of Internal Controls.

Section 404 is affecting business as well as IT units within organizations. Specifically, section 404 has three major requirements:

- A written confirmation by which management acknowledges its responsibility for establishing and maintaining a system of internal controls and procedures for financial reporting;
- An assessment at the end of the most recent fiscal year, of the effectiveness of the firm's internal controls; and
- A written attestation by the firm's outside auditor confirming the adequacy and accuracy of those controls and procedures.

Process and control requirements resulting from SOX are dramatically making actuaries take a close look at the tools they are using to conduct business. For most, this will involve a complete assessment of the use of technology, as technology plays such a large role in the storage and manipulation of financial data from the time it leaves the business areas until it eventually reaches the offices of the CFO and CEO on various financial reports. One area in particular that will be under close scrutiny is the use of spreadsheets. In the past actuaries have relied heavily on spreadsheets due to the ease of use and high-power functionality. While this has sufficed in the past, this use of spreadsheets may put companies at risk of not meeting the requirements of SOX section 404. There is no control and accountability of manually created spreadsheets as they are also difficult to store and maintain an audit history for. New methods of Business Intelligence can have a significant positive impact on an insurer's ability to meet the requirement of section 404.

New methods must be considered that will allow flexibility yet be maintainable and controllable at the same time. Gaining a level of confidence in the audit ability of its IT systems can be a huge headache for companies.

SOLVENCY II AND U.S. SOLVENCY

The Solvency II regime is an effort aimed at reducing the probability of consumer loss or market disruption. These new requirements will require insurers across the EU to change the way in which they calculate economic capital requirements as well as their methodology for maintaining audit trails. Firms will be required to publish accurate details of their credit, capital and risk management. Solvency II will require firms to value their assets and liabilities on a market-consistent basis and that more risk-sensitive capital requirements will address asset as well as liability risks, consistent with the domestic reforms that were implemented for insurers in 2004. Under Solvency II, the capital requirement for an insurer will be calculated on two levels:

1. Minimum Capital Requirement: The level of capital below which ultimate supervisory actions (for example, license withdrawal) would be triggered; and
2. Solvency Capital Requirement: The level of capital that enables an institution to absorb large unforeseen losses.

Solvency II will impact U.S. companies conducting business in the European Union. This said, as more U.S. companies look to expand their presence overseas, it is imperative that they begin to transition to adhere to Solvency II requirements. Recently, the National Association of Insurance Commissioners (NAIC) has adopted a solvency work plan including analysis of Solvency II's impact on the U.S. market. There is also a Solvency Modernization Plan in the works within the United States. Stochastic modeling as described above will play a large role in a company's ability to meet the requirements of the Solvency II regime. An HPC computing environment capable of running high-powered stochastic modeling can greatly increase a company's ability to meet the requirements of Solvency II.

LOOKING FORWARD

Actuaries have faced adversity before. During the first half of the 20th century, actuaries dealt with the advent of group insurance, World War I, the influenza pandemic of 1918, the severe financial problems with income disability and annuity coverage, the great depression of the 1930s, social security and World War II. In each instance actuaries not only succeeded, they thrived. With a synchronistic approach of tying together financial modeling, business intelligence and information technology, there is no doubt that actuaries will once again defeat adversity. ■

Brian Harris is vice president, Strategy for InSource LLC, in Windsor, Conn. He can be reached at 860.688.9900 x1021 or b.harris@insourceconsulting.com

Be More Productive with Everyday Technology

By J. Eddie Smith, IV

J. Eddie Smith, IV, FSA, MAAA is an assistant actuary with RBC Insurance. He is also an exam instructor with The Infinite Actuary. He can be reached at j.eddie.smith.iv@gmail.com

"Do more with less." We hear and feel it constantly these days. Today's business world requires us to complete highly complex tasks in very short periods of time with an ever-diminishing availability of people resources.

Fortunately, technology resources have not been diminishing, and even though it may be difficult to invest in new technology in the current economic environment, the standard technology toolbox available in the workplace is the most advanced ever, and it can indeed make it possible to do more with less.

But only if you utilize your technology efficiently.

As a group, actuaries tend to have a proclivity for technology. However, you may be someone right out of school that has not had a chance to pick up skills on the job, you may be a senior-level executive who began working before computers became as common as staplers in the workplace, or you may be anywhere in between and have just become accustomed to doing things a certain way. Whatever group you fall into, it is very likely that there are things that you can do to make yourself more productive (and valuable) by utilizing everyday technology tools more efficiently.

In the sections that follow, I list tips for common tasks in Windows. The information presented here just barely scratches the surface of the capabilities of these applications and is intended to be a primer only. I would encourage you to practice some of these tips and then look for more on your own. Hopefully you will find ways to incorporate these tips into your workflow to make your day go a little smoother. The idea is to give you more control of your technology (rather than letting the technology control you).

GENERAL WINDOWS USE

Quick Launch. Utilize the quick launch toolbar (next to Start button) by putting shortcuts there for commonly accessed applications. Avoid going to Start > Programs to open applications that you use regularly. Additionally, move any shortcuts cluttering up your desktop into the quick launch toolbar. That way, you can access them anytime without needing to show your desktop.

Keyboard shortcuts. I encourage you to embrace your keyboard as a highly efficient interface between you and your computer. Listed below are a few general Windows keyboard shortcuts that I use all the time. If you currently use your mouse for most of these activities, I guarantee that your productivity will soar if you utilize your keyboard for these tasks instead:

General Windows Keyboard Shortcuts:

ALT+TAB	Switch from one open window to another.
ALT+F4	Close a Window. This also brings up the Windows "shut down" prompt if no Windows are in focus (i.e., the desktop is visible/active).
Win+D	Minimize all windows and show desktop.
Win+E	Open Windows Explorer.
Win+L	Lock your computer.

Keyboard Shortcuts that Work in Most Common Application

CTRL+S	Save file.
CTRL+Z	Undo last action.
CTRL+Y	Redo the last undo.
CTRL+N	Create new file (e.g., Excel workbook, Word document).
CTRL+B	Apply bold format.
CTRL+U	Apply underline format.
CTRL+C	Copy cells to clipboard.
CTRL+X	Cut cells to clipboard.
CTRL+V	Paste contents of clipboard.
F7	Spell check.

Stay organized. Keep keep your digital "desk" clean.

- Create a "temp" folder on your desktop for random files so that you avoid clutter. Keep it sorted by date modified and clean it out every so often.
- Every hour or so, look at how many windows you have open. Do you need them all right now? Close any that you don't need to: 1) de-clutter your workspace, and 2) help ensure that you are not leaving unsaved documents open or documents that others on your network may need to modify.

WINDOWS EXPLORER

Utilize Favorites. Do you spend a lot of time in Windows Explorer and often have to scroll up and down the folder tree pane? Fortunately, you can save a lot of time by using Favorites to navigate quickly to common spots rather than using the folder tree view. Yes, you can "bookmark" folders just like you do Web sites using the Favorites menu. You can also add a Favorites button to the Explorer toolbar by right-clicking the standard toolbar and choosing Customize. Bookmarking folders that you visit often allows you to get to them in one click no matter where you are on your hard drive or network—just keep the Favorites side bar displayed.

Utilize History. Did you know that Windows keeps track of the documents you've been working on just like a Web browser keeps up with your browsing history? The History feature is useful for quickly reopening, say, a spreadsheet that you just accidentally closed or a file that you were working on last Tuesday but can't remember where you saved it. You can also add a History button to your toolbar.

Keyboard shortcuts. Yes, I do love keyboard shortcuts (and you should too). Here are a few that are specific to Windows Explorer:

BACKSPACE	Move up one folder level.
CTRL+ Number Pad + -	Auto-size columns in the "Details" view.
F2	Rename the selected file or folder.
CTRL+C then CTRL+V	Quickly duplicate files.
ALT+F+N+F	Create new folder.
F5	Refresh directories (like refreshing a Web page).
CTRL+I	Show Favorites in side bar.
CTRL+H	Show History in side bar.

MICROSOFT OUTLOOK

Most of the business world relies on Outlook for e-mail, calendar and task management. I must admit that Outlook is not my preferred e-mail client, and I only use it at my day job, where it is required. Out of the box, I would argue that Outlook is designed to facilitate your rapid drowning in e-mail even if you only receive a moderate volume of messages on a daily basis. The following tips will help you make the most of Outlook and allow you to take steps toward controlling Outlook, rather than it controlling you. If you use Lotus Notes or some other e-mail client, it is likely that you can apply some of the concepts below to those applications.

Turn off new message alerts. By default, Outlook alerts you with a little pop-up message each time a new message arrives as well as playing a sound. If you are like most professionals, you receive a steady stream of e-mail throughout the day, and consequently, Outlook is constantly nagging you as you do other work. Eliminating these distractions makes it significantly easier to focus on the task at hand and makes it easier to deal with e-mail on your own terms. In Outlook 2003, the message notification settings are buried in Tools > Options > Preferences tab > Email Options > Advanced Email Options. Once you arrive there, uncheck everything under the "When new items arrive in my inbox" section, and instantly begin enjoying the benefits of a less disruptive work space.

Keep as few emails in your inbox as possible. I cannot overstate the amount of clarity and sense of control you gain by keeping e-mails out of your inbox. In fact, there is an entire movement called Inbox Zero pioneered by Merlin Mann that advocates taking control of e-mail by keeping your inbox as empty as possible. It advocates turning e-mail "noise" into actionable tasks that you methodically execute. You can read all about Inbox Zero concepts at http://www.43folders.com/izero. Another great resource is the book Getting Things Done by David Allen.

Use Rules. Outlook has some fairly sophisticated options for filtering incoming and outgoing mail. You can use Rules to automatically move messages from specific senders to another folder, automatically flag messages, and a multitude of other actions. Just go to Tools > Rules and Alerts, and begin exploring the choices. As an example, I automatically route all "newsletters" I receive to a specific folder that I can peruse once a day.

CONTINUED ON **PAGE 8**

Use e-mail folders smartly. Utilize the Favorites pane in Outlook for folders that you access often. Create search folders to scan for messages with specific criteria (e.g., flagged messages, e-mails that are more than X days old, unread items, e-mails tagged with a specific category, etc.). By carving your pile of e-mail into well-defined, actionable buckets, you can process your mail more efficiently.

Keyboard shortcuts. If you spend a significant amount of your day in Outlook (like many professionals), using keyboard shortcuts can greatly reduce the amount of time it takes to process messages because your hands can stay on your keyboard (rather than moving back and forth from the keyboard to the mouse). Here are just a few that I use incessantly:

Email shortcuts

CTRL+R	Reply to selected message or active message.
CTRL+SHIFT+R	Reply all to selected message or active message.
CTRL+F	Forward selected message or active message.
CTRL+ENTER	Send message.
CTRL+SHIFT+M	Compose new e-mail (anywhere in Outlook).
CTRL+Q	Mark selected messages read.
CTRL+SHIFT+G	Flag and/or assign reminder to selected messages.
CTRL+SHIFT+I	Go to inbox from anywhere in Outlook.
CTRL+Y	Go to any mail folder.
CTRL+SHIFT+V	move message to a folder.

Calendar shortcuts

CTRL+SHIFT+A	New calendar entry (anywhere in Outlook).
CTRL+G	Go to a specific date.

In the Day view, you can press ALT+2, ALT+3, etc., to view two days, three days, etc., at a time.

Task shortcuts

CTRL+SHIFT+K	Create a new task (anywhere in Outlook).

Type dates and times into "due" and "reminder" fields instead of using your mouse to select from the drop-down menus. Outlook recognizes many date formats. For example, you can type "today" or "tomorrow" or "fri," and it will know what you mean.

Use conversation groups. Outlook can organize the contents of mail folders in a number of ways; just check out the View menu sometime. One very useful "view" that you may not know about is the Conversation view. This view is a lifesaver if you have been away on vacation and have a bulging inbox because it groups e-mails by subject, then sorts by date received. It allows you to quickly see all messages related to a specific conversation and go straight to the most recent e-mail in that conversation. You may be able to eliminate 10+ e-mails at a time in each conversation and speed your way to an empty inbox. You can quickly organize your messages by conversation by going to Views > Arrange By and selecting Conversation. You will also notice a number of other fields that can be used as grouping keys.

Use Find Related Messages. If you right-click a message in a folder, you can select "Find Related" messages to quickly locate other e-mails in a conversation even if they are scattered across different mail folders. This is very useful for tracking down attachments related to a conversation and also finding past responses you have made.

As I said before, I hope you find some of these tips useful, and I would very much welcome any feedback or be happy to elaborate on the items discussed here. Moreover, if you are interested in other technology topics aimed at beginners, improving workflow efficiency, or general computing tips, please let me know at j.eddie.smith.iv@gmail.com. Given that actuarial work is so heavily impacted and aided by technology, it only makes sense that we educate ourselves and our staff as much as possible to get the most out of the technology resources we already have. ■

European Approaches to the Modeling Process Are Changing—What Can We Learn?

By Van Beach and Bruce Keenan

Financial projections and models have been a mainstay of actuarial departments for years. As we look ahead to an environment where models are used for financial reporting as well as financial projections, there is little question that the actuarial workload will increase. Since actuarial models historically have been used for ad-hoc processes such as pricing, business planning, strategy analysis, and cash-flow testing, there is an acknowledgement that better processes and controls will be required to meet SOX and internal audit requirements as these models will now be used for setting reserves and capital. However, it appears that few companies are anticipating significant restructuring or strategic modeling changes and are simply planning to expand existing processes. Suppose, however, that your company would need to double its actuarial staff to meet emerging regulatory requirements with current processes? Would "do nothing" continue to be an option? This is the reality that is being faced by some European insurers as they come to grips with implementation of Solvency II and look ahead to IFRS Phase 2, and market-consistent reporting.

Here in the United States, we are just starting down the path of using actuarial models for financial reporting. What insights can we gain from the European experience? Let's look at some of the actions that are being taken:

Rationalization. As the complexity of modeling requirements increases, companies are taking steps to simplify the internal modeling processes to meet these requirements. Actuarial systems are being consolidated to a single platform. Business logic and modeling approaches are being simplified and a newfound emphasis on reuse is being enforced. Data marts are being created to standardize model input—in-force records, assumptions, and product values. Output is also being standardized for easier aggregated reporting and analysis of results across a wider range of business units. By reducing the number of systems, the amount of business logic, and the variations in input and output structures, the complexity of the modeling process is greatly reduced which reduces maintenance effort.

A long-held European requirement for actuarial projection systems has been the need for flexibility to reflect company-specific product features, management actions and financial structures. While this is still a requirement, some are revisiting the need for flexibility and placing an increased emphasis on the ability to standardize core modeling approaches—i.e., provide flexibility within controlled parameters. The impact is similar to the rationalization of systems, business logic, and data described above—but is even more powerful. Not only is the maintenance effort reduced by the standardization, but the consistency increases understanding of all models across the organization and allows for better resource allocation and utilization. Further, key projection methodologies, such as asset and liability interactions, can be consistently applied and best practices leveraged across all models. This inherent consistency also positions the software vendor to provide ongoing updates and higher levels of support, which further reduces the internal burden.

Centralization. Possibly the most fundamental shift in actuarial approach is the movement to centralize actuarial modeling. This is a multi-faceted endeavor as it involves infrastructure, expertise, processes and management. For a multinational firm, added complexity is involved as the centralization spans multiple countries and regulatory jurisdictions. Nonetheless, companies have awakened to the commonalities across these boundaries and are anxious to achieve the economies of scale made possible by addressing the requirements through common resources. Beyond the efficiencies of centralized approaches, companies are also eager to achieve higher levels of quality and development of best practices made possible through collaboration and specialization.

Each aspect of centralization—computing infrastructure (typically a grid), modeling expertise and model management, modeling processes (data management, model execution, report generation, etc.), management—is in itself a sizeable challenge and is beyond the scope of this article. In many cases, formal change management protocols and expertise are being employed to navigate the transition from decentralized operations to the centralized structure.

CONTINUED ON **PAGE 10**

Van Beach, FSA, MAAA works with insurers globally to address financial modeling issues as the MG-ALFA Director of Business Development. He may be reached at van.beach@milliman.com.

Bruce Keenan, FSA, MAAA is a consulting actuary with the life insurance consulting practice of Milliman in London, UK. He may be reached at Bruce.Keenan@Milliman.com

Automation and Integration. Through the evolution of actuarial models providing one-off insights to being used for regular and external reporting, European companies have come to appreciate the totality of the actuarial modeling process—data assembly and manipulation; assumptions management; model update procedures; business logic development; efficient, scheduled execution of model runs; and aggregation and post-runtime analytics for deeper actuarial insights. With reporting windows tightening and scrutiny of results heightening, the only way to keep pace is to embrace automation, integration and pursuit of a hands-off actuarial process. While relatively new to mainstream actuarial modeling, the automated end-to-end actuarial process has already been achieved for insurance hedging applications. Not only does the automation and integration speed the execution time for the end-to-end process, a key byproduct is the additional actuarial capacity that is now free to analyze, interrogate and understand model results.

New approaches. Formulating a comprehensive view of a company's capital requirements requires assessing a complex array of risks and developing an opinion on the interactions and impact of these risks under a wide range of scenarios. New tools and approaches are currently being tested and evaluated for aggregating these risks across a diverse organization. New model compression techniques, Cluster Modeling as an example, are also being developed which enable a company to leverage existing models for higher volumes of scenarios. With advanced model compression, companies achieve the advanced stochastic analyses while losing very little accuracy and maintaining the transparency into the underlying actuarial models.

Sustainable processes. European companies have come to realize that actuarial models are no longer just producing one-off results. The results need to be produced on a regular, timely basis. Results must be reconciled across multiple products, functions and geographies. Changes made between reporting periods need to be documented, controlled, validated, and the impact assessed. All results must be fully reproducible and further, all values must be auditable and traceable both within a reporting cycle and across reporting cycles. To achieve a successful, sustainable modeling process, careful planning and consideration must be taken to structure models for change. Further, change must be embraced, managed and continually assessed—this aspect of the modeling process will never be finished.

Exposure and use. Pillar II of Solvency II mandates that the models that are being created to assess the solvency capital requirements are actually used by the companies to drive decisions. It is not sufficient to simply calculate the solvency capital requirement; rather companies must demonstrate that the results from risk models are being actively consulted and used to drive business decisions. This requires that actuarial models are exposed to the broader organization and the theories of risk management permeate throughout the company. While it is recognized that the core financial modeling expertise will reside in the actuarial department, the risk data must become part of the corporate cognition. For this to happen, risk data must be produced regularly and reliably, risk lexicon must be communicated, and tools for risk analysis must be delivered to stakeholders across the organization.

What do these actions tell us about the challenges ahead in the United States? Embedding a model within a mission-critical financial reporting or risk management process will require an approach to modeling that is more comprehensive and more strategic. It is a very different requirement than the ad-hoc applications which have historically been the norm. Larger volumes, complexity and scrutiny of results must be addressed—all while keeping costs and resources down. It is a unique challenge—finding a solution that will meet the paradoxical requirement for modeling increasingly complex products and regulations while simplifying the modeling environment. And this is only the start. Once a strong system foundation is in place, the perspective then broadens to the end-to-end actuarial process and developing internal management structures to ensure consistency and quality across functions and business units. Similar to our European counterparts, we'll likely also conclude that "do nothing" is not an option. ■

Introducing Twitter

By J. Eddie Smith, IV

By now, you have likely heard about the latest social media darling known as Twitter. Maybe you even have friends and colleagues that have jumped on the bandwagon. What's all the fuss about? And for that matter, why would any rational individual, already choking on the unrelenting torrent of information streaming from the Web, care about investing more energy into yet another nascent social media platform that will surely be dubbed "so 2009" once 2010 arrives?

I would say that, well … Twitter is different. Yes, its development is immature, and yes, it is unstable at times, but I think it has firmly staked its flag in the Web 2.0 landscape. Today's online experience is about accessing information not only from Web sites, but from people—real time. Twitter delivers the people element in spades.

Created in 1996 as a "side project," Twitter's initial mission was to provide a simple, free, text-based service that allowed people to tell the world what they were doing—in 140 characters or less. No pictures, no audio, no video, not even color; just text-only fragments of human consciousness flitting about cyberspace. Roughly three years later, there is little doubt that the idea is catching on. As of February 2009, the number of people on Twitter had approached 10 million, representing 700 percent-plus, year-over-year growth.

If you are familiar with other online social media sites like Facebook, MySpace or LinkedIn, you may wonder what makes Twitter so special. After all, it only has one aspect of the other social media titans: the status update. Why would you want to only be able to read that your friend just got back from their Caribbean vacation when you could also see pictures and video of it?

The best explanation of why Twitter represents a unique, stand-alone service is the quality of information contained in those 140-character messages. Also different is the fact that Twitter updates are public by default, unlike other social media sites. Twitter does give you the option of restricting who sees your updates. However, doing so can really take away from Twitter's value, which increases every time someone contributes another bit of public information to the Twittersphere. Over time, people who have interests similar to yours will begin to follow you based on what you have to say. Twitter is best viewed as a completely open forum, and the more you contribute to it, the more you tend to get out of it.

TWITTER BASICS

The essential basics of Twitter are:

1. Tweeting,
2. URL shortening,
3. Following,
4. Replying,
5. Retweeting,
6. Sending private messages, and
7. Using hashtags.

Tweeting

Messages sent through Twitter are called "tweets," which are somewhat similar to SMS text messages—only without the usage charges. Verbose a tweet is not. There is only so much you can say 140 characters at a time, and this can take some getting used to. However, since so many people use text messaging already on their mobile phones, the concept of sending succinct electronic messages is probably not all that foreign to those new to Twitter. The act of tweeting is simple: enter your message, and hit the update button on your Twitter home page. Your tweet will be visible to anyone who happens across your home page. In no time, you will have created your own micro-blog.

URL Shortening

It's not just for baking anymore. But seriously, you may have wondered by now how you could ever include a Web address in a tweet if you only have 140 characters? The clever solution has been the advent of URL shortening. A number of Web sites are dedicated to making URLs of any length into much smaller URLs by associating the original (long) URL with a shorter version. When someone clicks on the shortened URL, they are routed to the original address. A very well-known URL shortening service is offered by tinyurl.com. To make a "tiny URL," simply copy the URL of the page you want to share, paste it into a field on the tinyurl.com site, click the "Make TinyURL" button, and then copy the new, compact URL into your

CONTINUED ON PAGE 12

J. Eddie Smith, IV, FSA, MAAA is an assistant actuary with RBC Insurance. He is also an exam instructor with The Infinite Actuary. He can be reached at j.eddie.smith.iv@gmail.com

tweet. As an example, I converted the SOA Technology Section Web address into http://tinyurl.com/cxkbvy, which is only 25 characters long.

Following

The first priority for most people who have just created their first Twitter account should be to find other users, or "tweeps," to follow. When you follow someone on Twitter, you will see their tweets in your friends' time line on your Twitter home page. Following someone does not require their approval—unlike other social media sites like Facebook where approval is required to complete the connection.

Following is easy, and perhaps in this way, Twitter mirrors real life. To find people to follow, you can use the Twitter site to search by name, e-mail address, Twitter handle, location, and other criteria. There are also several directories that you can use to find people (e.g., WeFollow, Localtweeps, and Twellow). For most people new to Twitter, the natural instinct is to look for people you already know. Others may want to search for celebrities, bloggers, news organizations, even the president of the United States. However, as you begin to use Twitter more, you will probably find that the majority of tweeps that you follow are complete strangers that simply tweet about things that are interesting to you.

Replying and Retweeting

It probably will not be long before you see a tweet that contains something interesting. Maybe a friend just finished watching a movie that you liked, or maybe someone just gave a useful computer tip, or maybe you just saw a bit of breaking news. There are two constructive ways you can react: replying and retweeting.

A Twitter reply is essentially just a public tweet directed at one or more Twitter usernames. Replies to other Twitter users should begin with the @ symbol and include the username of the Twitter user(s) to whom you are replying. For example: "@TheirUserName I love that movie too - I have watched it 100 times."

While replies are useful for making direct (albeit it public) contact with another tweep, they do not add value to Twitter in the way that retweeting does. A retweet is simply a reposted tweet. Retweeting is unique to Twitter and serves to accelerate the flow of information to others. The syntax for retweeting involves putting "RT" in front of the original author's username and their tweet. For example: "RT @BreakingNews Pope makes unexpected visit to Baghdad." Obviously, the more a tweet is retweeted, the more exposure it gets in Twitter.

Retweeting is generally encouraged in the Twitter community. It is seen as both complimentary to the original tweet author, and it is an important tool for propagating information. Retweeting is also an excellent way to attract followers because they recognize that you "get it," and they also might hope that you follow back, thereby growing their follower count and giving them more voice in Twitter. In other words, they want you to retweet them too.

Direct Messages

Twitter also allows private, or direct, messaging. Direct messages can only be sent to your followers. When a direct message is sent, it shows up in the recipient's Twitter inbox, and only they see it. A direct message must start with the letter D. For example: "D @TheirUserName Meet for lunch?" This aspect of Twitter is the most similar to SMS text messaging because it is a closed one-way communication between two users. Care should be taken when sending private messages through Twitter, however. As with e-mail or any other form of electronic communication, "privacy" is a luxury afforded only in the short-term. Additionally, there have been reports that certain Twitter clients have, on occasion, incorrectly posted direct messages publicly because of program error. I would not recommend sending any truly sensitive information as a direct message.

#Hashtags

Hashtags provide a convenient way to tag or categorize tweets so that others can find information on specific topics. Hashtags are preceded by the # symbol. Examples include #politics, #BreakingNews, #actuary. Some Web sites even track hashtag use and identify trends as the general conversation on various items changes. One of the more famous hashtags is #FollowFriday, which is part of a Twitter tradition to promote the usernames of other tweeps that you follow on Friday.

TWITTER: THE NEW SEARCH ENGINE

No matter what information you choose to tweet, you can always utilize Twitter as a huge searchable pot of information. In fact, I would argue that search is where Twitter really separates itself from its Web 2.0 counterparts.

The information spewing from all those gregarious tweeters and retweeters represents a resource unlike any offered by the Web so far. While Google has come to represent the gold standard of Web searching, it cannot query the one repository that Twitter can: people's brains. But don't worry; you do not need little wires dangling from your scalp to use Twitter. Rather, the human element of Twitter coupled with its openness provides a convenient way to tap the collective knowledge of other tweeps. You cannot really "ask" a conventional search engine a question in a way that will elicit a human response, and you cannot "thank" and interact with Google in a manner that fosters an ongoing, mutually beneficial relationship. Indeed, a conventional search engine's smarts are limited only to the fetching algorithms that probe its electronic archives.

But with Twitter, you can ask questions and get human responses.

- "Where's the best place to get a cup of coffee on Michigan Ave in Chicago?"
- "Are Google Docs reliable?"
- "Does anyone have an Amazon coupon code?"
- "Anyone heard rumors that Company ABC is buying Company XYZ?"

Of course, searching Twitter goes far beyond finding good cups of coffee and shopping discounts, and you do not necessarily have to wait for a response to get your questions answered in many cases. In fact, you are often likely to find answers and information simply by searching existing tweets. There is no better source of real-time information than Twitter on breaking news and current events. If you search for keywords on a breaking news topic, you are likely to see tweets that are only seconds old. Such timeliness is not even offered by Google. If you search Google for breaking news, you are likely to see a mixture of results on the first page that will vary in age. Twitter, on the other hand, will show you what is happening now.

HOW MIGHT TWITTER FIT INTO YOUR DIGITAL LIFE?

Twitter's uses are probably as varied as the tweeps that compose the Twitterverse, and unlike other advice you might hear, you really don't need a "mission statement" before you start. Some people enjoy being able to tell the world they just ate a bologna sandwich. Others see it as a platform for promoting their business. Some organizations see it as a way to communicate to their members as well as promote the organization's interests. Local news stations are using it to send out weather alerts.

Marketers are becoming very interested in what people are saying on Twitter. Suppose someone is thinking of moving to Miami and says so in a tweet. A realtor in Miami would be very interested in finding that person. Perhaps more relevant to the actuarial profession, if you visit sellyourterm.net, you will see a running archive of tweets containing references to insurance.

It is difficult understand the Twitter experience without trying it yourself, and as such, its benefits to you will not come into focus until you actually try it. A few more examples of how people are using Twitter are given below:

- *Self-promotion and advertisement.* Tell the world who you are or about your business and why they should care.
- *Conversation.* Use Twitter like an open chat room in which you discuss any and everything that interests you.
- *Collaboration.* Create an open forum to discuss ideas, update project status, and invite ideas from outside.
- *Meet new people.* Who knows what relationships you may forge in Twitter.
- *Micro-blogging.* Don't want to spend the time writing a full blog? Just use Twitter.
- *Share information.* Maybe you are a do-it-yourselfer with a lot of information to share with others.
- *Communication at conventions and trade shows.* Use hashtags to instantly create groups and give attendees a way to find and communicate with each other while at the event.

CONTINUED ON **PAGE 14**

- *Announcements.* Examples include new blog posts, updates to your Web site, or just general announcements about your life.
- *Stay in the know.* Follow your favorite organizations, celebrities and other prominent people; follow breaking news and identify news trends.

While I have generally extolled Twitter, some words of caution are in order.

- Twitter is still a fairly immature platform. It is not at all unusual for Twitter's main site to go down, and sometimes it can be downright quirky and unstable (e.g., tweets disappear).
- Twitter's owners have not really figured out how to make it generate revenue, so its future is uncertain.
- Some might argue that so much openness is not such a good thing. Only time will tell if what we say today will cause frustrations tomorrow.
- Spam has been on the rise, and as more people flock to Twitter, we should expect the spam war to escalate.
- You cannot (and should not) believe everything you read on Twitter, and it does provide a medium in which rumors, misinformation, and untruths can spread like wildfire. The eruption of tweets containing the term "swine flu" earlier this year is an example of how quickly something can become overblown in Twitter.
- It is difficult to convey context in 140 characters. Tweets can easily be misinterpreted.
- Is there such thing as too many tweeps? Does Twitter's value decline if too many people join?
- Processing Twitter's information can be like drinking from the firehose. Users would benefit from more tools that allow them to be more selective and filter the information.
- Worms have started invading Twitter. Clicking on links in tweets could foster the spread of worms and viruses.

The best way to learn more about Twitter is to sign up and start tweeting, following and getting followed. Online resources for Twitter abound. Not only are there a multitude of articles, there are also several blogs devoted entirely to using Twitter. A few examples are:

- The Official Twitter Blog: http://blog.twitter.com/
- TwiTip: http://www.twitip.com/
- Twittercism: http://twittercism.com/
- Mashable — The Social Media Guide: http://mashable.com/

I would like to thank Mary Pat Campbell for contributing some of her Twitter knowledge to this article. Please feel free to direct any questions or comments to me at j.eddie.smith.iv@gmail.com. I look forward to hearing from you. Oh, and do not hesitate to follow me, @eddie_smith, on Twitter. Happy tweeting.

REFERENCES AND OTHER SOURCES FOR TWITTER NEWCOMERS

- Twitter.com
- "Twitter Traffic Explodes...And Not Being Driven by the Usual Suspects!": http://www.comscore.com/blog/2009/04/twitter_traffic_explodes.html
- "Profile Twitter founders: Jack Dorsey, Biz Stone and Evan Williams" http://www.telegraph.co.uk/science-andtechnology/technology/twitter/4388880/Profile-Twitter-founders-Jack-Dorsey-Biz-Stone-and-Evan-Williams.htm
- "Swine flu: Twitter's power to misinform": http://neteffect.foreignpolicy.com/posts/2009/04/25/swine_flu_twitters_power_to_misinform
- Directories:
  - WeFollow: http://wefollow.com/
  - Localtweeps: http://localtweeps.com/
  - Twellow: http://www.twellow.com/
- URL Shorteners:
  - http://tinyurl.com
  - http://tr.im
  - http://bit.ly
- "Twitter: Why You Should Care": http://www.thisistrue.com/blog-twitter_why_you_should_care.htm
- "5 stages you will go through on Twitter": http://ianbarnett.wordpress.com/2009/04/12/6-stages-you-will-go-through-on-twitter
- "30 Essential Twitter Tutorials for Newbies and Experts": http://www.outlawdesignblog.com/2009/30-essential-twitter-tutorials-for-newbies-and-experts/ ■

Cool Tech

By Matthew Wilson

The purpose of Cool Tech is to expose you to a lot of different programs available on the Web. In the last edition of CompAct I referenced over 200 free programs available on the Web that do all kinds of cool things. In this edition I'll show you even more free programs to check out. I'm covering a cool way to hack Web pages in the Firefox browser. The general idea is that I want the world to conform to your ideas instead of the other way around. The Greasemonkey Firefox add-on plus a Javascript script helps you to move along that path.

I discuss the trials and tribulations of creating a photography Web site using Gallery 2.

I regularly make changes to my Web site (www.1913intel.com). Most of the time the changes don't work out all that well, but occasionally something does work. I effectively blew up 33 percent of my traffic by going to a membership site, but my site revenue is up anyway. I discovered a social browser that allows me to blog post about five times faster than before. So I'm spending much less time on my blog than before, but earning more.

I take a look at how you can use phpMyAdmin for the administration of your MySQL databases. Sometimes you have to snoop around a database if there is a problem, and phpMyAdmin is the way to go.

It might not be entirely clear if a particular open source package is going to work for you. That means installing several different packages to see what works. It would be nice if there were demo versions to play around with. You need to try out Click2Try.

I provide a short review of the Google browser called Chrome, and how to build a simple fulltext search engine using PHP and MySQL.

In the download section is a spreadsheet containing the Mersenne Twister code for a random number generator that started out in C++, moved to VB, and I moved it to Excel.

MY VIEW

Recently I was advising a coworker (Karl) on how to put together a photography Web site. As a hobby Karl likes to take quality portrait pictures of people. He had a free Web site, but wanted to do better.

Karl mentioned over and over that he wanted something real simple. OK, I got the hint. He wants something real simple. I was thinking about a real simple html based Web site until he mentioned that he wanted folders to be password protected. One can protect folders by setting ".htaccess" appropriately, but a much better way to go is to use a photo gallery package based on PHP and MySQL. But which one should we use?

Based on feedback from a photography forum, we chose the Gallery 2 photo package. It's a very popular open-source package that uses PHP and MySQL, and it fulfilled Karl's requirement for password protection. I thought this was going to be very easy to install given that it uses PHP and MySQL, but I made a strategic mistake in the selection of host. I suggested the Yahoo! hosting service, because it has worked for me in the past and provides quicker help than my current host—dreamhost.com. Unfortunately, Karl couldn't get the Gallery 2 package to work properly. Yahoo! was blocking some files from being uploaded and the thumbnail pictures wouldn't work properly.

Sometimes there are dependency issues that make it difficult to get a software package to work properly. You may need to install additional libraries or programs in order to get your primary program working correctly. If a host has one-click installs then they have already solved the dependency issues for you.

After two weeks of playing around it was time to bail out. I suggested that Karl use the one-click installs at dreamhost.com. The Gallery 2 package was one of the one-click installs. Within one weekend Karl had the Gallery 2 package installed and working properly. He also tracked down a new template to give his Web site a slick new look.

You can check out Karl's Web site here: http://karlkramerphotography.com

CONTINUED ON **PAGE 16**

Matthew Wilson, ASA, MAAA, can be reached at matt_wilson@farmersinsurance.com

You can get the Gallery 2 photo album package here: http://gallery.menalto.com/

In order to install Gallery 2 one should download the Gallery 2 zip file. Unzip the file on your PC, then go into the top-level directory and upload all of the individual files and folders to your Web site folder on your host's machine.

On the host's machine there exists a file folder called karlkramerphotography.com. In that folder is a file called index.php. When someone enters www.karlkramerphotography.com in their browser window the index.php program gets executed. So that program had better be there or the Web site is not going to work. Index.htm or Index.html work for Web sites not using PHP.

Next create a MySQL database (Db), and be sure to write down the key information about your database: Connection (localhost?), Db name, Db ID and password.

Gallery 2 has a quick installation guide: http://tinyurl.com/226qcc

After two weeks of failing with Gallery 2 we were getting close to the point where I was going to step in and do it for him. But changing hosts did the trick.

In the end it all worked out because Karl's Web site is now looking pretty good. Unfortunately, he had to waste a couple of weeks to get everything working correctly.

We did learn one other lesson from Karl's experience of having to switch hosts after a few weeks. When you sign up for a Web host get the monthly payment plan even though it's more expensive. After three or four months then you can switch to the yearly payment plan if you wish.

While Karl was playing around with his Web site I was trying to make some major changes to my Web site—1913intel.com. I wanted a way to accelerate the posting of news articles, and I wanted to change my Web site to a membership site.

I post news articles each day to 1913intel.com on various topics. I was thinking that I could develop my own browser that grabs an article title, link and the first X number of characters of the story. I could use the WordPress API in order to post automatically to the blog. Finally, I could hire an Indian programmer through Elance (elance.com) to put this together for me.

This type of project is actually not that hard. You can get the code for a full featured browser using VB.Net at planetsourcecode.com. This browser piggybacks on the Microsoft Internet Explorer browser. Check out Web Browser/IE Clone at http://tinyurl.com/mlf2wt

It's pretty easy to create a browser right in Excel too. I have created a simple demo browser in Excel. Here's the link: http://www.mattstechnotes.com/files/My_Web_Browser.zip

As I was putting together the specs for this project, I decided to search the Web for a browser that also does blogging. That's when I stumbled onto the Flock social browser at flock.com. Flock works with WordPress, Blogger, Twitter, Facebook, MySpace, YouTube, Flickr and a lot more. It also works with different mail systems, so you can easily fetch your e-mail. Flock is free and piggybacks on the Firefox browser.

In order to use Flock for blogging, one simply selects some text from an article. Next you right click and select "Blog This." An editor will pop up with your selected text in a block quote. You can add to the selected text if you wish. Also, you can add tags. Finally, after you publish it you can select the categories for the article.

The entire posting process is significantly faster using Flock versus doing it manually in WordPress.

I did have a problem using Flock in WordPress 2.8.1. It wouldn't work until I went into the WordPress settings area concerning remote posting. That's located at Settings, Writing, and then selecting both boxes at the remote publishing section.

The next item I worked on was converting my blog to a paid membership site. After quite a bit of research I

decided to use a strategy that separates the billing from WordPress. This means I would handle billing manually through PayPal and simply find a way to restrict article viewing to members only. After some research I selected a plug-in called Member Access.

Member Access is a simple plug-in that basically does what I need, but did require a tweak. It restricts viewing for any article that I wish. Also, I could set the default value to public or private viewing.

You can get the plug-in here: http://wordpress.org/extend/plugins/member-access/

It didn't work exactly the way I wanted, so I needed to go into the MySQL database and manually execute a query to change the member_access_visibility field in the wp_posts table. I was interested in changing the default value to public for old posts, and I wanted the default value for new posts to be private.

Here's the MySQL update query that I used in phpMyAdmin to change the table:

update wp_posts set member_access_visibility = replace(member_access_visibility,'default','public');

This query changes the default value for existing posts to public. So old posts will be public, but new posts will be default (private).

I installed a PayPal button on a subscription page. When a person paid for a subscription, PayPal sent me an e-mail. I then manually registered the individual and sent them an e-mail with the log-in information.

Ultimately, only 10 people were interested in signing up for the membership section at $4/month. I really needed about 50 people to express interest before it was worth my while to launch this section. So it didn't work out. While all the membership coding remains in place, I now just make all of the articles viewable to the public.

On a more positive note, my regular advertising revenue is now over $200 per month. My sources of ad revenue are Google Adsense and Text Link Ads.

I created a simple PHP ad rotator program that randomly selects different Adsense ads each time a page is selected. The ads vary by size, shape and color. Regular readers quickly become immune to one type of ad, so it's important to keep the ads changing.

You can get the ad rotator program here: http://www.mattstechnotes.com/files/admain.php

My website will probably never bring in a lot of revenue due to strategic reasons. As Felix Dennis said in How to Get Rich, "Keep your eye on the ball. The one that says—the money is here." Meaning that in order to have a decent chance of making money one must focus on areas that a lot of money is flowing into. Clearly, news on emerging risks and international conflict is not an area that much money is flowing into.

How can I jump-start my Web business if I'm only going to make between $200 and $300 per month? It turns out that the average salary in the Philippines is less than $300 per month. I can hire someone in the Philippines to start creating content for more Web sites. My next stop is Craigslist- Philippines.

There is quite a bit of competition for good English writers in Craigslist-Philippines. You can get articles written for about $3 per one-page article. A U.S. writer will typically charge $10 or more per one-page article. I have an ad in Craigslist-Manila, but I won't be able to report my results until next quarter. I'm running up against a publishing deadline.

A few months ago I used Elance to hire a Philippine company to do Web research for me. It cost me $5 per hour for the project. They did the research that I required, but there were a few issues. They were late and didn't keep me informed very well during the project. Still, I got what I wanted.

PHPMYADMIN (HTTP://TINYURL.COM/5RCRW)

Since I used phpMyAdmin to update my MySQL database, I thought I would discuss how this program works.

CONTINUED ON **PAGE 18**

phpMyAdmin should really be called MySQLAdmin. It's a great program that allows you to snoop around your MySQL databases. You can browse tables, look at structures and even change things if you want to. You can manually edit records or easily execute queries. It's a big time-saver.

If your host offers MySQL databases, then most likely it also offers phpMyAdmin too.

phpMyAdmin gives you the ability to dump your database to a text file through its export feature. This could come in handy if you want to back-up your database or port it to another Web site. It also has an import feature that will read the exported dumps.

There is a problem with the phpMyAdmin import feature. If you try to import a large database then you could get timed out by PHP. You generally have 30 seconds before a PHP script will time out. What if you need a couple of minutes to import a large database? Think BigDump!

BigDump gives one the ability to import a big dump or large database dump file. It imports the database file in discrete chunks then automatically restarts the session to ensure that you don't encounter a time out problem.

Here's the link to Bigdump: http://www.ozerov.de/bigdump.php

CLICK2TRY WEBSITE

Click2Try is a demo Web site where you can try out all kinds of open source software packages to see if they meet your needs. Here are some of the categories:

Business, Accounting, CRM/ERP, Document Mgmt, E-Commerce, Groupware, Human Resources, Project Management, Education, Course Management, Linux, Centos, Ubuntu Desktop, Office/Productivity, Office Suites, Productivity, Software Development, Developer Workstations, IDEs, Software Project, Management, System/Network Admin, Backup, Help Desk, Monitoring, Utilities, File Management, Vertical Applications, Real Estate, Web Development, Blogs, Content Management, Editors, Forums, Wikis.

Here is the link: http://www.click2try.com/catalog/

LIST OF OPEN SOURCE PACKAGES IN WIKIPEDIA

Wikipedia has a list of packages that is so large it would be difficult to list them all. This is definitely a good source for open source software.

Here is the link: http://tinyurl.com/pmst7

GREASEMONKEY – A FIREFOX ADD-0N

Greasemonkey is a seriously cool add-on for the Firefox Web browser. Actually, Greasemonkey doesn't do anything by itself, rather it gives you the ability to inject Javascript code into a Web page and execute it. Effectively, you can operate on all of the content on a Web page. In order to make Greasemonkey do something, you will need to get Javascript scripts or write your own.

Here's the link to the Greasemonkey add-on: https://addons.mozilla.org/en-US/firefox/addon/748

Check out the free online book called Dive into Greasemonkey: http://diveintogreasemonkey.org/

Here's a list of 25 useful scripts: http://tinyurl.com/3xzwaz

You can get a complete list of scripts here: http://userscripts.org/

You can use another add-on called Platypus to create your own scripts. Platypus gives you the ability to modify a Web page and save the changes as Javascript code that can be used by Greasemonkey. Here's the link: https://addons.mozilla.org/en-US/firefox/addon/737

Want to learn more about Greasemonkey? Check out the Greasemonkey Hacks book. You can get a used copy for under $5 plus shipping. Here's the link: http://tinyurl.com/l3qqcy

You can get a lot of these links at my new Web site called Matt's Tech Notes. During my Web surfing I post short article snippets on anything that I find interesting. Here's the link: http://www.mattstechnotes.com/

CHROME—THE GOOGLE BROWSER

I tried Chrome out the other day just to see what it could do. I regularly open about 50 tabs at once in Firefox, and occasionally there are stability problems. Chrome is built such that if one window (tab) crashes then none of the other windows is affected.

It turns out that Chrome is a fairly plain browser. It didn't have any trouble opening 50 tabs at one time; however, it is clearly slower when initially opening a tab than Firefox or IE. Once the tab is open, then it seems to work better. For example, moving tabs around is very easy even though there are 50 open. Also, system stability seems to be high.

Overall, there are probably not enough bells and whistles in Chrome to make switching from IE or Firefox to be worthwhile. However, if you like to have 30 or 40 tabs open at the same time then you might want to give this browser some consideration.

For most people Firefox is going to be a better browser than Chrome because of its ability to incorporate add-ons. Firefox is my favorite browser unless I'm blogging. I switch to the Flock browser for blogging purposes because of its social media abilities. Flock can accept many Firefox add-ons too.

BUILDING A FULLTEXT SEARCH ENGINE WITH PHP (HTTP://WWW.ZEZ.ORG/ARTICLE/VIEW/83)

In this section I review some PHP coding that gives you the ability to create a fulltext search engine. It's a simple search engine, but will be just fine for most Web sites.

In building this search engine we first create three tables:

1. A table holding the actual text documents.
2. A table holding all of the words in all of the documents.
3. A table providing a link between the words and the actual documents.

The first part of the process involves indexing your data. This involves counting the occurrences of each word in each document. The words and occurrences are stored in table #2. A link between the words and associated document are stored in table #3.

To determine the relevance of an article one needs to calculate the word density (percentage)—the number of times a word is used in a document as compared to total words in the document.

One could easily improve the word density calculation by determining a weight based on position. For example, the earlier a word appears in the document then the more weight it gets.

A search query involves individually searching for each word in the query to find the word density for each document. The document with the highest sum of densities wins.

I used this code to create my own search engine several years ago. I used a spider to gather news information every day (See the last issue of CompAct to get the code for these spiders). I put the information into a searchable MySQL database and used the code in this section to add search capability.

You can download all the code and read the full article here: http://www.zez.org/article/view/83

DOWNLOAD SECTION: MERSENNE TWISTER

The Mersenne Twister is a random number generator based on Mersenne prime numbers – form 2^N – 1

The code for this program was originally written in C++. Carmine Arturo Sangiovanni translated the C++ into Visual Basic. I took Carmine's code and got it working in Excel. Additionally, I created a working example showing how it works.

Here's the link: http://www.mattstechnotes.com/files/Mersenne_Twister_v2.zip ■

HELP US STAY IN TOUCH.

UPDATE YOUR PROFILE IN THE ONLINE DIRECTORY OF ACTUARIAL MEMBERSHIPS.

You can change your:

- Addresses, such as employer and work/home contact information;
- Preferences, including your name and options for including or excluding specific information from your listing in the directory; and
- Professional designations, such as non-actuarial designations, employment type and areas of interest.

Update your profile today at ***https://www.actuarialdirectory.org***.

New

Professional Development E-Courses

Take advantage of these additional opportunities to grow your knowledge on a variety of important and valuable subjects, while earning continuing professional development (CPD) credit, from the convenience of your computer.

The new e-course topics include:

- decision making and communication,
- enterprise risk management,
- financial economics,
- financial reporting and operational risk,
- fundamentals of actuarial practice,
- health systems overview,
- investment strategy,
- pricing, reserving and forecasting,
- regulation and taxation and
- social insurance.

Learn more and register today at WWW.SOA.ORG.

R Corner – Combining Cluster Analysis and Predictive Modeling

By Steven Craighead

Editor's note: R Corner[1] is a series by Steve Craighead introducing readers to the "R" language used for statistics and modeling of data. The first column was published in the October 2008 issue, and explains how to download and install the package, as well as providing a basic introduction to the language. Refer to each CompAct issue since then for additional articles in the series. The introductory article can be found on p. 24 of the October 2008 issue on the SOA Web site: *http://soa.org/library/newsletters/compact/2008/october/com-2008-iss29.pdf*

Steve Craighead, ASA, MAAA, is an actuarial consultant at TowersPerrin in Atlanta, Ga. He can be reached at steven.craighead@towersperrin.com.

In this issue, I will outline a combination approach that I used within R to estimate the 90 percent Conditional Tail Expectation (CTE90) on an old block of business that I used more than 10 years ago[1]. I have used this approach to dramatically reduce the computer processing time regarding the calculation of Principle-Based Reserves and Capital.

This is done by using representative scenarios to train a separate smaller predictive model to replicate a full business model. After the predictive model is constructed, I then use all the scenarios within this less time-expensive model and not use any probability weights at all as is done with most representative scenario techniques.

In the study I use a single clustering algorithm to select these representative scenarios.

Let's briefly discuss the method used to select representative scenarios.

REPRESENTATIVE SCENARIOS

There are several (actuarially) published, as well as commonly known, methods to determine representative scenarios from a larger collection. However, I will use a cluster analysis technique called CLARA (Clustering LARge Applications). I will not introduce any weighting within our scenario selection process and I will treat the selection process as directly formulated by the sources. However, later I will discuss the use of weighting in the control of bias.

The CLARA Cluster Algorithm[2] can either use a sum of squares or a sum of absolute values metric to measure distance. In our work below I will use a sum of absolute values to indicate distance between separate scenarios. I have found the sum of absolute differences tends to select more extreme scenarios than does the sum of squares approach. This is important because when training a predictive model, the available information obtained from extreme scenarios increases the accuracy of the predictive model.

These are the steps of how CLARA works:

1. Choose the number of clusters (representative scenarios) desired, say M.
2. Choose an arbitrary data point out of the N sampled ones and call it Pivot #1.
3. Calculate the distances from Pivot #1 to the remaining N -1 data points.
4. Name the data point with the largest distance from Pivot #1 as Pivot #2. Randomly decide between ties.
5. Calculate the distances of the N - 2 non-pivot points to Pivot #1 and Pivot #2.
6. Assign each of the N - 2 points to the closest of Pivot #1 or Pivot #2, thus forming two disjoint clusters. Flip a coin if the distances are equal. Each of the N - 2 points now has a unique distance to its pivot point.
7. Rank these N - 2 distances in descending order. The point producing the top distance is called Pivot #3. (Break ties randomly.)
8. Following the above procedure to select the additional M - 3 pivot points.
9. If the number of points associated to a pivot #k is Nk, assign a probability of Nk/N to this pivot point.

For further insight on how CLARA is used within R, use the command:

```
help(clara)
```

Now, let's look at predictive modeling.

CONTINUED ON **PAGE 22**

FOOTNOTES

[1] Craighead, S. (2000), "Insolvency Testing: An Empirical Analysis of the Generalized Beta Type 2 Distribution, Quantile Regression, and a Resampled Extreme Value Technique," ARCH, pp. 13-149, available at http://www.soa.org/library/research/actuarial-research-clearing-house/2000-09/2000/arch-2/arch00v26.pdf.

[2] CLARA algorithm, see: http://rweb.stat.umn.edu/R/library/cluster/html/clara.object.html

PREDICTIVE MODELING

Predictive modeling is a means by which you can design or create a model that can be used to predict an outcome with approximately the same probability that is observed with the actual data. There are many different techniques, but while working independently with the Academy Valuation Basis Table subcommittee I found one type of predictive model that excelled all expectations. This modeling technique is called Projection Pursuit Regression[3] (PPR). Below, is a fairly extensive discussion around what it is and how you use it to create models.

In linear regression, you fit a response variable Y to a collection of n predictor variables X_i in the familiar form:

$$Y = \alpha + \sum_{i=1}^{n} \beta_i X_i + \varepsilon$$

Where in additive models, the $\beta_i X_i$ are replaced with various functions $f_i(X_i)$, with this form:

$$Y = \alpha + \sum_{i=1}^{n} f_i(X_i) + \varepsilon$$

Projection Pursuit Regression (PPR) was introduced by Friedman and Stuetzle[4] and it is a modification of the additive model in that there are:

- M different f_i.
- Each f_i acts on a different linear combination of all n of the X_k.
- A specific coefficient of these linear combinations is denoted by α_{ik}.
- Each f_i is multiplied by a β_i.
- The constant term is the average of the response variable.

So PPR takes on the following form:

$$Y = \bar{y} + \sum_{i=1}^{M} \beta_i f_i \left(\sum_{k=1}^{n} \alpha_k X_k \right) + \varepsilon$$

or in vector format:

$$Y = \bar{y} + \sum_{i=1}^{M} \beta_i f_i (\alpha_i \cdot X) + \varepsilon$$

where $X = (X_1, X_2, ..., X_n)$ is a vector of predictor values, and $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$ is a vector of directions.

The term Projection in PPR comes from the fact that X is projected onto the directional vector α_i for each i.

The Pursuit arises from the algorithm that is used to determine optimal direction vectors $\alpha_1, \alpha_2, \ldots, \alpha_M$.

Each f_i is called a ridge function. This is because they only have values in the α_i direction and are considered constant elsewhere. Effectively, what occurs is that the overall PPR model is a linear combination (β_i are the coefficients) of the ridge functions. These functions only take on values that arise from the projection of the predictors against the direction vectors, and the functions as assumed to take on a constant value in any other direction. So, each ridge function is like the profile of a mountain range, and you linearly combine these functions along all different ridges (as pointed out by the α_i).

On a formal basis, Y and X are assumed to satisfy the following conditional expectation:

$$E[Y \mid X_1, X_2, \ldots, X_n] = \mu_y + \sum_{i=1}^{M} \beta_i f_i (\alpha_i \cdot X)$$

with $\mu_y = E[Y]$ and the f_i having been standardized to have zero mean and a unit variance. That is: $E[f_i(a_i \lozenge X)] = 0$ and $E[f_i^2(a_i \lozenge X)] = 1$, where i takes on values from 1 to M. I assume that the realized sample values for the random variables Y and $X = (X_1, X_2, \cdots, X_n)$ are independent and identically distributed to the distributions of Y and X, respectively.

FOOTNOTES

[3] Projection Pursuit Regression, see: http://www.slac.stanford.edu/pubs/slacpubs/2250/slac-pub-2466.pdf

[4] Friedman, J. H. and Stuetzle, W. (1981) "Projection pursuit regression," Journal of the American Statistical Association, vol. 76, 817-823.

[5] R Development Core Team (2006). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

The PPR algorithm in the R stats library[5] estimates the best $\beta_i, f_i,$ and the α_i by minimizing the following target function for the mean square error

$$E\left[Y - \mu_y - \sum_{i=1}^{M} \beta_i f_i(\alpha_i \cdot X)\right]^2$$

across all the data samples for Y and X. Note: This expectation can be based on a weighted average as well.

A powerful trait of PPR models, since the predictor vector X is projected, is that interactions beween different X_j and X_k are included within the model, whereas other model algorithms cannot do this without user intervention.

Other advantanges of PPR are:

- The PPR model is a continuous function. According to Venables and Ripley,[6] they cite Diaconis and Shahshahani[7] and say that given a large enough number of ridge functions, PPR can approximate any arbitrary continuous function.

- It is the best possible fit since every component is solved for the minimization of the weighted least squares.

- Each ridge function does not extrapolate outside of its specific domain. If the specific $\alpha_i \cdot X$ is outside the domain the relevant domain endpoint is used.

- The model handles the interactions between the different predictors as I saw in the last section.

- PPR models categorical predictors as easily as continuous predictors.

- PPR models can take extremely large amounts of data and create a very good model of the underlying data. You can also adjust the model to distinguish between model fit and model smoothness.

- PPR does not suffer from the curse of dimenionality (COD). COD arises from the increased complexity of a multi-dimensional surface. Since PPR optimally is solved one ridge function at a time, the difficulty of trying to locate global optimal values for model calibration is eliminated.

But PPR is not perfect and there are a few disadvantages:

- The range of a PPR model may be outside of the range of acceptable values. For instance, if you are using PPR to model mortality, model results could fall below zero or above one. However, PPR will not extrapolate outside the existing ridge functions, so if any predictor projects on a specific α with a value outside the domain of a specific ridge function, the ridge function takes on the value either at the furthest point on the right hand side or left hand side. This no extrapolation rule can lead to biased results.

- All of the parameters are point estimates and there is no distributional consideration given to the significance of a specific parameter. Because you are not able to create a confidence interval using the R ppr function around each of the α_k or the β_i, you will not be able to determine if a specific parameter is significant to the model. In fact you are unable to actually test to see if the actual model is significant, other than the use of the goodness of fit statistic. There are complex methods that have been developed using spherical statistics to overcome this, but these require an understanding of advanced Banach Algebra in functional analysis and have not been included within the R ppr function.

- You can easily overfit or over explain the data. See Venables and Ripley for a further discussion.

R TECHNIQUES AND DIAGNOSTICS FOR PPR

The procedure when you use the R ppr algorithm is that you must first, specify that M should range between $M_{MIN} = 1$ and some positive integer M_{MAX}. The ppr algorithm then creates a PPR model for each M from M_{MAX} to M_{MIN} in a descending fashion, and

CONTINUED ON **PAGE 24**

FOOTNOTES

[6] Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S, Fourth Edition. Springer.

[7] Diaconis, P. and Shahshahani, M (1984) "On non-linear functions of linear combinations,", SIAM Journal of Scientific and Statistical Computing, vol 5, pp. 175-191.

at the same time produces a goodness of fit statistic for each value of . You then scan this list of goodness of fit values looking for a local minimum or zero. If this local minimum is at M_{MAX} you should reprocess the experiment with a larger M_{MAX}. Once you determine the local minimum, say s, reset $M_{MIN} = s$ and reprocess the ppr algorithm with the same M_{MAX} as before. The resultant model arising from the backward iteration from M_{MAX} to M_{MIN} will then be the best PPR model.

Two other components that are implemented in ppr is the concept of Bass and Optlevel. The option "bass" is to allow the calibration algorithm access to Friedman's super smoother bass tone control[8] that is used with automatic span selection. It is used in ppr to smooth the results. The range of values allowed with this component is from 0 to 10. To increase smoothing within the data, increase this value. The default is 0 and this setting gives the best fit to the underlying data. Bass is similar to the h smoothness parameter used within the Whitaker-Henderson graduation formula.

The option Optlevel is an integer from 0 to 3, which determines the optimization thoroughness. The best models usually are obtained if this is set to three. Level 0, the ridge functions, are not refitted. Level 1, the projection directions, are not refitted, but the ridge functions and the regression coefficients are. Levels 2 and 3 refit everything, but level 3 takes pains to rebalance each regressors' contribution at each step and so reduces the chance of converging to a saddle point in the sum of squares.

One diagnostic aid in PPR model building is to plot the ridge functions. If these ridge functions are very noisy or discontinuous, you should expect that the resultant PPR model will behave oddly.

Another effective diagnostic aid is to both plot the fitted $\hat{Y}$ against the actual Y and do a simple linear regression of Y against $\hat{Y}$, assuming no intercept. The scatterplot should display symmettry around the 45-degree line and the coefficient of the regression should be approximately 1. These two diagnostics will indicate how well the PPR model will perform as a predictive model.

Note: Since a PPR model does not extrapolate outside of the sample data, the resultant fitted values from PPR model will hit a maximum value and will not grow any larger no matter how you manipulate the predictors. This is not the case for linear regression models, where there are no natural limits placed on how you set any respective X_i. However, you may revise the prediction object to conduct extrapolations. However, you must first feel comfortable with the continuity of the separate ridge functions. If these functions are very noisey or appear not to be differentiable, you might want to avoid all forms of extrapolation.

COMBINATION OF METHODS

In the past, most actuarial research concentrated only on the use of representative scenarios and weighting the results based on the probabilities associated with each representative. I have found that this approach alone does not adequately represent the overall behavior that you obtain when using all of the scenarios. However, the goal of model efficiency is to reduce the entire processing time of the various reserve or capital models. This has been mostly done in the past by either reducing either the number of model points used with the liabilities or assets or by reducing the number of scenarios processed through the model.

Independently, I have observed that PPR models are very effective, not overly sensitive to outliers within the calibration data, and replicate the overall behavior of high dimensional models well. Another nice feature of PPR is that is very quick when asked to evaluate additional input besides that of its training data.

In past experience, I have also been observed that the CLARA algorithm is very effective in selecting representative scenarios. This is due to the fact that the process discovers a majority of the extreme scenarios, which con-

FOOTNOTES

[8] Friedman, J.H. (1984). "A variable span scatterplot smoother." Stanford University Technical Report No. 5 Laboratory for Computational Statistics.

tribute to the tail of the reserve or the capital distribution. When you use representative scenarios and then apply the associated probability weighting to the results, the final results will be very dependent upon how those weights are obtained or used. However, if you do not use these weights at all, but only use the representatives as a training dataset for a predictive model, you can then process all scenarios through the resulting predictive model. For this to work well, the number of representative scenarios needs to be rich enough to adequately span the high dimensional business model. Also, you hope that the predictive model will also adequately model the business model as well.

I will now test the hybrid approach of using representative scenarios as training data and then processing all the scenarios through the predictive model.

Next, I briefly discuss what data is being used.

DATA SOURCES

I[9] described and modeled from over 100 insurance related datasets. In the work below I will concentrate on the 1993 dataset associated with business model 4 and the associated 10,000 interest rate scenarios used in the generation of those values. I also discussed the generation process of these scenarios as well in that paper. I have restricted myself to this specific dataset because I determined that this specific data set had such complex behavior, that if you are able to adequately model the underlying data, the remaining datasets are very easily modeled.

Now let's look at our experiment.

PROCESS OUTLINE

Using the information of the 10,000 scenarios and the 10,000 associated capital values mentioned above, I conducted 100 separate experiments using random samples of 5,000 scenarios for each representative set size. On each of these scenario sets, I apply the CLARA Algorithm and choose separate representative subsets. Once a specific representative set is selected, the representative scenarios, in addition to their associated capital values, are used as the training data for a PPR model. Once the PPR model is trained (or calibrated), the entire sample of 5,000 scenarios are then projected using the resultant PPR model. Using the PPR model, I calculate the CTE90 (which I will refer to as Model). Also, based on the specific sample of 5,000 capital values, I also calculate the CTE90, (this is referred as Actual). I then calculate the relative error associated between Actual and Model, by the formula: RE = (Actual – Model)/Actual.

The CLARA algorithm uses the following distance formula:

$$\sum_{t=0}^{20} (| i_{90:t} - i^{P}_{90:t} | + | i_{10:t} - i^{P}_{10:t} |)$$

The 90:t or 10:k notation is used to indicate the 90-day or the 10-year Treasury rates in year t.

RESULTS

Regarding, the CLARA Algorithm, I use representative set sizes of 50, 75, 125, 175, 200, 250, 300, 400, and 500.

As mentioned before, for each of these representative set sizes, I repeat the random sampling of 5,000 scenarios 100 separate times. By conducting this repeated sampling, you can observe the effectiveness of the overall process and approximate the sample error associated in the tests.

Now let's examine the corresponding CLARA Algorithm results for the CTE90 capital estimation. The graph at the end of this article displays the collection of box-and-whisker plots[10] where the CLARA Algorithm is used to select the separate representative scenarios.

For the CTE90 results (shown in the table at the end of the article), notice that these results are liberally biased, but the median results as well as the minimum and maximum relative errors are very tight. I have observed

CONTINUED ON **PAGE 26**

FOOTNOTES

[9] Craighead, S. (2000), "Insolvency Testing: An Empirical Analysis of the Generalized Beta Type 2 Distribution, Quantile Regression, and a Resampled Extreme Value Technique," ARCH, pp. 13-149, available at http://www.soa.org/library/research/actuarial-research-clearing-house/2000-09/2000/arch-2/arch00v26.pdf.

from prior experience using the CLARA algorithm (set to use the sum of absolute value of the differences metric), that the algorithm chooses more tail scenarios than any other published technique that I have used.

Below is an R function that actually takes the scenarios to be clustered in the dataframe clustdata. It then uses the actual interest rate scenarios used, which for this example, is the same dataframe associated with clustdata. If you want to cluster on some transformation of the scenario data, you need to use separate dataframes. Next the actual surplus values are provided in the dataframe observed. Finally, you specify with the ssize variable how many representative scenarios you want to model.

```
randomtest<- function(clustdata, intdata, observed, ssize)
{
clus<- clara(clustdata,ssize,metric="manhattan",samples=100)
sam <- clus$i.med
sam_clusinfo<-clus$clusinfo[,1]
pprsam<- ppr(intdata[sam,],observed[sam],nterms=1, max.terms=125,optlevel=3)
best <- order(pprsam$gofn)[1]
pprsam<- ppr(intdata[sam,],observed[sam],nterms=best,max.terms=best+10,optlevel=3)
yhat<-predict(pprsam,newdata=intdata)
}
```

Below are the commands that I used to generate one example from the experiment regarding 200 representative scenarios. If you want to produce 100 experiment, you will need to insert the code below in a for loop like for(q in 1:100) { }

```
ssize <- 200
bigsampsize<-5000 #let's just choose 5000 samples from the 10000 available
bigsamp<-sample(1:10000,bigsampsize)
dep<-x[bigsamp,] #x is the scenario dataframe
obs<-y[bigsamp,4] #y is the capital values that correspond to each scenario
clu<-x[bigsamp,] #this is the scenario dataframe

yhats <- randomtest(clu, dep,obs,ssize)
bot90<-floor(.90*bigsampsize) +1 #this sets up the logical location of where the tail is
top <- bigsampsize #last scenario
list90 <- seq(bot90,top) #
cte90<- (mean(sort(obs)[list90])- mean(sort(yhats)[list90]))/mean(sort(obs)[list90])
cat( ssize, bigsampsize, i,cte90,sep=" ","\n")
```

ISSUES REGARDING BIAS

Though the results are positively biased and so understate the capital, you can see that the average error is reasonable given the speed enhancement. Of course you may increase the size of the representative set and this will also reduce the bias. Donald Krouse at AEGON has given some insight into what a practitioner may take to reduce this bias. Based on his suggestions, you could introduce weights to the scenario selection process or by experimenting with other various metrics. This may help, because the bias may arise from the fact that the training scenarios may over- or under-emphasize certain attributes within the scenarios. Also, the PPR model itself can lead to biased results just due to the fact of how it calibrates as discussed before. You may want to introduce weighting to the calibration process or manipulate other settings to see if the bias can be eliminated or reduced. Currently, I have used other predictive models such as neural networks and other types of machine learning, to eliminate the bias, but I have found that PPR is still superior because it does not suffer from the curse of dimensionality. Furthermore, it simulates the underlying structure of the complex capital model quite well, where these other techniques poorly calibrate to the representative sets.

CONCLUSIONS

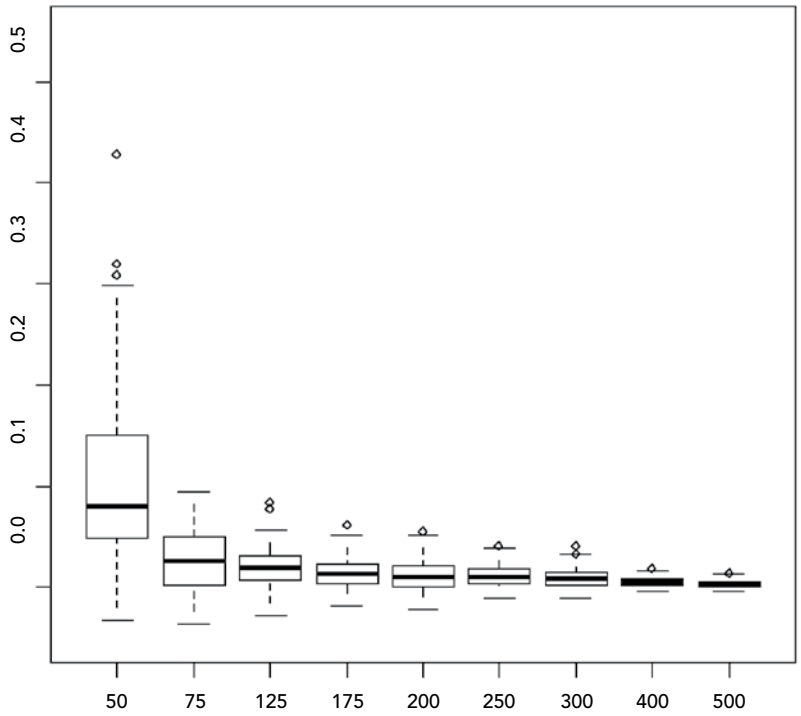
In this issue, I just discussed the results of the combination of CLARA and PPR. In the next issue, I will go into greater detail regarding the use of CLARA and the multiple applications to which I have applied that algorithm. In a future issue, I will also discuss further implementations of PPR as well.

If you want to examine this process further, the following two documents on the SOA Web site discuss this combination approach in more detail:

http://www.soa.org/files/pdf/2008-qc-craighead-dardis-51.pdf

http://www.soa.org/library/newsletters/financial-reporter/2008/june/frn-2008-iss73.pdf ■

CLARA Algorithm with PPR- CTE90



	Min	1st Qu.	Median	Mean	3rd Qu.	Max.
50	-0.03344	0.04874	0.08013	0.1027	0.1484	0.4273
75	-0.03633	0.002239	0.02523	0.02672	0.0496	0.09426
125	-0.02789	0.006763	0.02003	0.01907	0.03103	0.08259
175	-0.01918	0.003258	0.01248	0.01382	0.02327	0.06151
200	-0.02165	0.001014	0.01065	0.01294	0.0219	0.05468
250	-0.01053	0.003376	0.009998	0.01089	0.01766	0.03989
300	-0.01077	0.002906	0.007679	0.008806	0.01451	0.04095
400	-0.00422	0.001636	0.005093	0.005152	0.008281	0.01846
500	-0.00379	0.000105	0.002869	0.003092	0.005355	0.01331

FOOTNOTES

[10] Box-Whisker Plot, see http://en.wikipedia.org/wiki/Box_plot.

SOCIETY OF ACTUARIES Technology Section

CompAct

475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
p: 847.706.3500 f: 847.706.3599
w: www.soa.org

Non Profit Org
U.S. Postage
PAID
Carol Stream, IL
Permit No 475

FSC Mixed Sources
Product group from well-managed forests and other controlled sources
www.fsc.org Cert no. XXX-XXX-XXXX
© 1996 Forest Stewardship Council