Institute of Actuaries of Australia

Volatilities, Correlations and Risk Capital Allocation in the US PC Insurance Industry, parts of a cross-industry study with recommendations.

Prepared by David Odell, Ben Zehnwirth and Glen Barnett

Presented to the Institute of Actuaries of Australia

Risk and Capital Management Research Conference 2008.

Sydney, Australia

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract.

We seek answers to the following questions in the context of reserve and underwriting risk for long-tail insurance portfolios. How is Economic Capital to be calculated on a company-wide level, taking into account diverse lines of business (LOBs)? What are correlations between lines and how are they manifested? What constitutes “common economic drivers” between different LOBs and how widespread are they? How much do different companies have in common in the same LOB? Implicitly we also ask what the is cost to a company if it accepts default industry figures for correlations rather than measuring them by internal modeling?

Our methodology was to analyse the portfolios of the largest US Insurance groups based on their Schedule P 10-year reports as published by A.M. Best 2006. Each line of business in each group was fitted with a model in the Probabilistic Trend Family (PTF). This provides a precise “fingerprint” of the trends determining the evolution of paid losses over the period in question, and of the variability around those trends. These models become the starting point for joint models in MPTF (Multiple PTF) which produce matrices of process and parameter correlations for the lines within each group. We argue strongly that correlations can be measured only *after* accounting for trends in this way.

Volatilities and correlations between the lines in a group enable the allocation of risk capital in a precise way which best reflects individual company experience. The Development and Calendar (or Payment) trend profiles enable us to segment this risk capital further into payment years with great advantage to asset/liability matching. In general, a wide diversity of capital allocation patterns is seen, highly specific to each of the company groups. We conclude that inter-company differences and company-industry differences are highly significant and therefore that company-specific correlations should be used in Economic Capital calculations.

Keywords: Risk, Risk Management, Economic Capital, Long-Tail Liability, Modelling, Schedule P, Correlation, Correlation Matrix, Probabilistic Trend Family, Lines of Business, Multiple Lines of Business, Predictive Capability, Forecast, Risk Capital Allocation, Trends, Drivers.

Introduction: On Risk and Modelling

The Scottish poet Don Paterson has a neat way of summing up a lot of recent thinking in evolution:

“Consciousness [is] only a tool possessed by a unit mammal which found itself in need of some half-decent predictive capability.”

In this way of thinking consciousness evolved out of the need for risk management and so in a sense you could say that risk management is as old as we are, but in spite of this, as a scientific discipline it is of very recent date. For its origin we can mention the beginnings of insurance in the coffee houses of 17th century London, and the simultaneous beginnings of probability theory as the mathematical treatment of gambling, Bernoulli’s invention of the utility function in 18th century, Bentham’s attempt to quantify utilitarian principles in early 19th century , the beginning of modern statistics and Bachelier’s study of the stock market via Brownian motion at the turn of the 20th century, mathematical economics and game theory in the 20th century and so on. In some respects the modern emphasis on risk has to do with the evolution of a technology complex enough so that we can carry out the calculations necessary to quantify the intuitive notion of risk. However, in many areas even sophisticated modern methods don’t take us much beyond “half-decent predictive capability”, although evolution (and the markets) teaches us that even a small edge in the “half-decent” can go a long way.
On the other hand where the physicist Eugene Wigner could famously note “the unreasonable effectiveness of mathematics in the natural sciences”, in financial mathematics and statistics we find mathematics to be only reasonably effective. Our goal is not to find out the truth but to make the best use of the available information in real time.

The following definitions, taken from an on-line dictionary indicate in a nutshell the move from intuitive qualitative notions of risk towards quantitative objective notions.

1. The possibility of suffering harm or loss; danger.
2. A factor, thing, element, or course involving uncertain danger; a hazard.
3.
a. The danger or probability of loss to an insurer.
b. The amount that an insurance company stands to lose.
4.
a. The variability of returns from an investment.
b. The chance of nonpayment of a debt.
5. Probability and severity of loss linked to hazards.

We start out with “possibility”, “danger”, “uncertain”, “hazard” which are qualitative or subjective terms and end up with “probability”, “stands to lose” = expectation, “variability” = variance or volatility, “probability/severity” = probability distribution, which are all quantitative and mostly objective terms, but somewhat confused together.

It is important to separate notions of risk from notions of luck. In ordinary usage these terms are fairly close to each other. Gamblers it might be said, are all in love with the idea of luck, but only the better ones also understand risk. The expressions “risk-takers” and “high-risk behaviour” often associated with young males, mean that while “risk” seems to be better thought out than “luck”, it still has a considerable glamour about it.
Roughly speaking, in this context risk = objective, quantitative and general while luck = qualitative and individual. The concept of luck is still an important one in the business world, it is linked to ideas of the charismatic individual the one who “makes his own luck” – an individual like Richard Branson or Donald Trump. We cannot entirely dismiss these concepts as unscientific just because we don’t yet

fully understand what signals our brains are processing when they come up with conclusions about them.

“Finally, we shall see that [rating an Insurance company] necessarily requires a subjective element, that not everything can be quantified, and that the capital adequacy does not come down to a single ratio.” Marc Philippe Juilliard (Fitch) in “Evaluating Insurers’ Capital Requirements” SCOR 2003

Risk had become a buzz word in the last decade in our public and corporate discourse. People are being asked to decide on matters that depend more and more on an understanding of this concept and as they do so the concept itself undergoes change, becoming more precise in some respects and more elusive in others. Examples from the present time are almost too numerous to mention, but the most obvious being climate, terrorism and the complex and unpredictable effects of globalisation. One the one hand our technology suggests to us that we can have unprecedented oversight and control over things and on the other hand the increased complexity arising from the very dissemination of technology means that a whole new level of unpredictability comes into play, the financial markets being the obvious example here.

With Solvency 2 (and Basel 2) we are seeing the evolution of a new kind of regulatory culture defined by a response to the concept of risk.

Regulators want to be able to quantify and compare the ability of corporations to meet their commitments to stakeholders in near medium and long term. Corporations wish to be able to demonstrate in a confident manner to both stakeholders and regulators that they are aware of possible future developments affecting their viability and are suitably provisioned.

Securitisation and trading in risk are another powerful source of the pressure to make risk measures more precise and more objective. These conversation have led to the promulgation of new risk measures such as Value at Risk, Expected Shortfall, and Tail Correlations, new concepts are really just new ways of looking at probability distributions. They have also set people thinking about extreme values and rare events, “black swans” and heavy tailed distributions. These can be modelled probabilistically, but parameterisation is difficult and the approriate risk measures are not so clear. There may be an analogy to the need for non-linear utility functions in behavioural economics.

[It is notable that the debates about climate change and what actions to take about it rely on concepts equivalent to probability distributions, more or less well understood. They proceed in terms of conventional statistics and the risk measures derived from them. Other perils which fall under extreme values and rare events, such as asteroid collisions, massive volcanic eruptions, solar flares, a new ice age, etc. get much less air time.]

Within corporations risk management is often divided into distinct silos such as financial risk”, “operation risk”, “legal risk” etc. and it is a problem at a higher level to sythesise the contributions from each silo into a clear picture of the overall situation. Naturally it would be best if all were expressed in the same quantitative language with consistent measures of probability but this is an ideal not yet possible to attain.

This talk focuses on reserve risk in long-tail lines of general insurance which is an area highly amenable to detailed quantitative analysis and where mathematical modelling has long been in use.
As we just saw with the dictionary definition of quantitative risk there are a number of different paths that open up in the mathematics of uncertainty which lead to different approaches with different merits, so before considering the particular modelling framework which we used in our study it is worthwhile thinking about some general issues related to mathematical modelling of risk.

What does an Economic Capital model on the enterprise level need to do? What desiderata does it need to meet? As far as the reserve risk component is concerned we mention the following although this list could be added to:

- Models need to address the actual scenarios in which the losses arise. In terms of reserve risk this is simply being called on to disburse more in losses and expenses than have been prepared for within a fixed time horizon, usually one year. For this reason it is the incrementals that should be modelled and not the cumulatives. It is cashflows and not loss ratios that represent risk. Models should be directly applicable to DFA and asset/liablity matching and so should actually model cashflows rather than drawing them off as a side-effect of modelling something else.

- Models should be mathematically rigorous but their chief features should be translatable into the ordinary language of business. A good model for a sector of business should tell a story about the performance of that sector in relation to its economic environment. This is an important point because risk models form part of a dialogue within a business, between levels and within levels of management and between the business and regulators and the business and stakeholders. For example, a dialogue with a regulator about the right level of risk capital for a line of business should be rightly understood as a discussion about the level of volatility estimated for that line (not the volatility of the estimates!). The discussion will be more fruitful if the economic factors that lead to this estimate are reflected transparently in the model.

- Modelling procedures should be consistent and smooth. Consistency means that the model outputs for two quite different lines of business are comparable since the same underlying procedures have been applied in each case. If sectors of a business are going to be competing for a share of the risk capital then the volatility estimates they each produce should be produced by the same methodology so that they can be fairly compared. Smoothness means that if two sets of data are very close to each to each other the models should also bear a close resemblance. In particular if we look at gross and net (of reinsurance) data for the same insurance portfolio we would expect to see similar salient features in the models.

- Modelling frameworks should include comprehensive diagnostics so that model fit can be assessed and compared. Where there are differences in the results from different models diagnostics provide a basis for an informed choice between alternatives.

- Berquist & Sherman[1] in an appendix to their well-known paper on testing the adequacy of reserve modelling present a list of questions that that an actuary should ask department executives in Claims, Underwriting, Ratemaking etc. A good modelling practice should prompt such questions by picking up any changes in claims behaviour. This includes modelling of counts and estimates as well as paids, and meaningful relations between the models.

- The “predictive capability” of models depends on the formulation of plausible future economic scenarios, and so models should speak directly to the kinds of factors used to formulate these.

- Correlations between lines of business in a corporation need to be understood as belonging to the volatile component of the forward estimates. This is an important and somewhat subtle point that we will go into in more detail later on. In a nutshell though, say I have two lines of

[1] *Berquist*, J.R.; and *Sherman*, R.E., ". *Loss. Reserve Adequacy Testing*: A. Comprehensive, Systematic Approach. ,". PCAS LXIV, 1977, pp. 123-184.

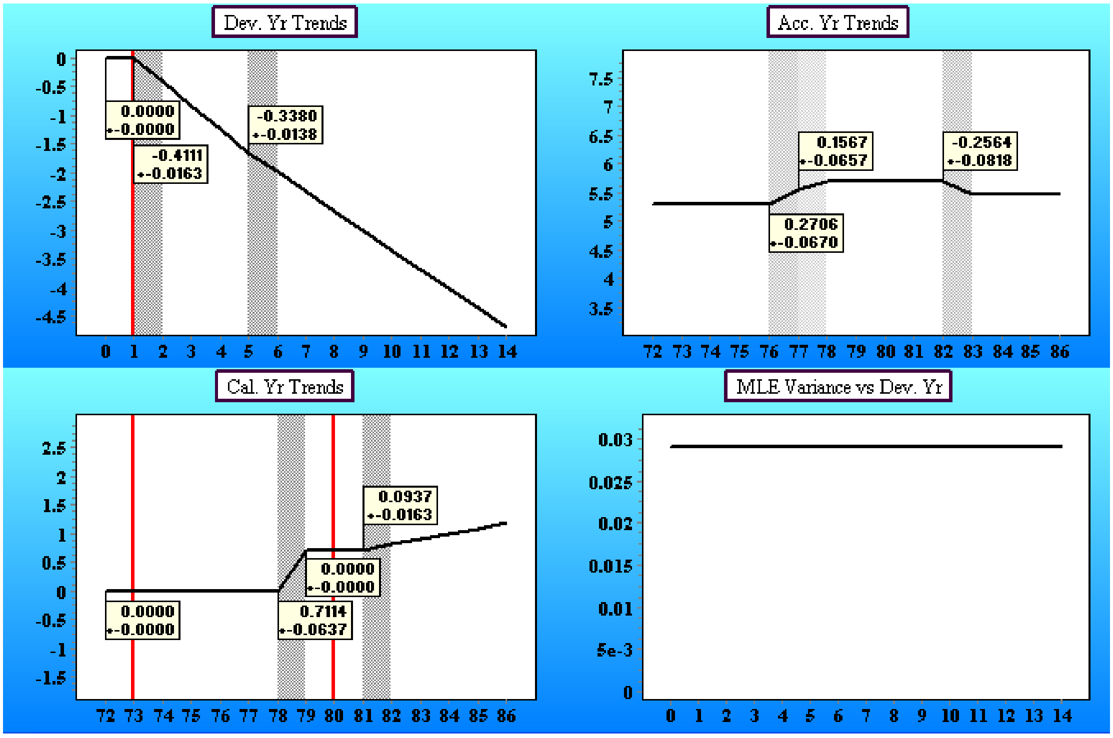
business A and B and expected losses for the next year of LA and LB. By the end of that year when I count up the paid losses the results will differ from my estimates by certain amounts, VA and VB, (so that the paids are LA+VA, and LB+VB, respectively). The amounts VA and VB may be negative or positive, but a correlation occurs when VA and VB tend to go the same way significantly more (or less) than half the time. In other words, when the volatile or unforseen components are tied together in some way. That's what needs to be measured for each individual corporation and for all lines of business.

A very brief introduction to the Probabilistic Trend Family (PTF) and Multiple PTF (MPTF) modelling framework.

The PTF modelling framework was introduced by Barnett and Zehnwirth in their paper "Best Estimates for Reserves" with which many of you may be familiar. In an appendix to this paper PTF modelling is described in some detail. At this point I only want to describe it briefly so as to make the rest of the paper understandable.

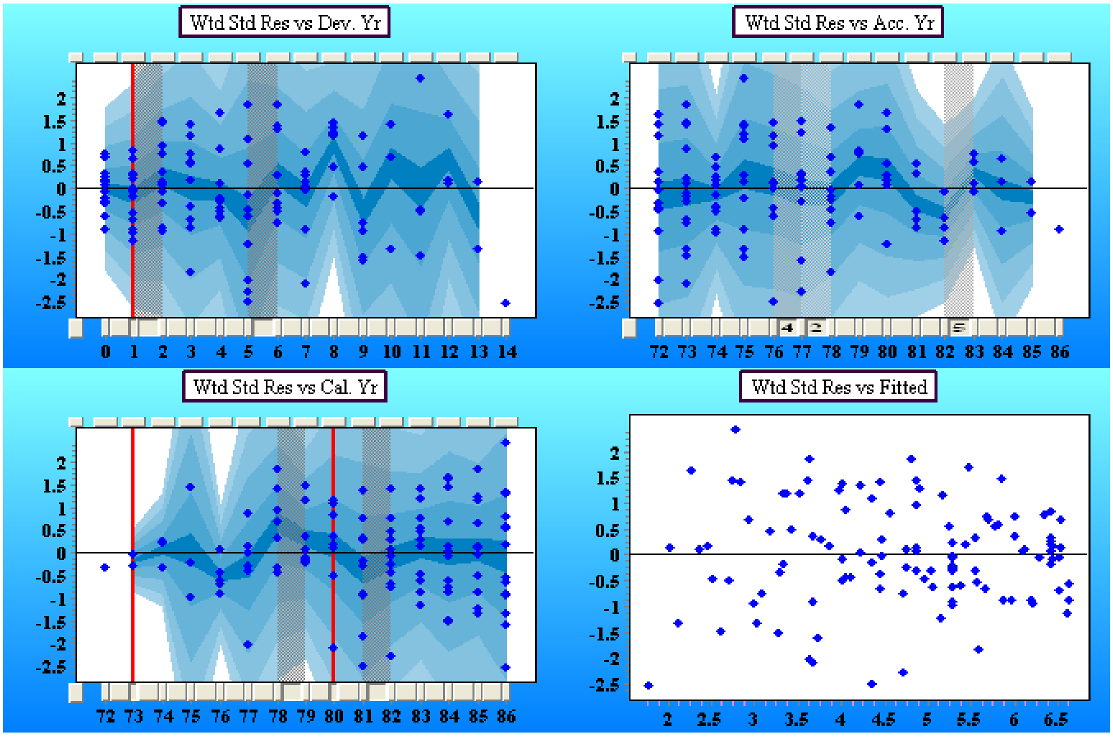
PTF works with incremental data arrays. A basic modelling unit is the "trend" understood in the usual way as a percentage trend along a time axis, and for this reason amongst others, the data is subjected to a logarithmic transformation before modelling. Percentage (i.e. multiplicative) trends on the dollar scale then appear as linear trends on the log. scale. Three time axes are considered: the development direction, the accident direction and the calendar direction. The modelling procedure seeks to fit the data by calibrating trends in the development direction and the calendar direction. In the accident direction the modelling unit is the "level" or more precisely the changes in level from year to year. These directions are not entirely independent, a trend in the calendar direction is also seen in the development direction and the accident direction, however no combination of only development trends and accident changes can capture the precise effect of a calendar year trend.

A model is understood by means of a model display chart which consists of four graphs, one for each of the trend directions and a fourth that shows the volatility in relation to development period.



Example of a PTF model display. Black lines are the trends or levels and the vertical grey bars and red lines show changes in trend or level.

A quick way to understand this is to see the development trend graph (top left) as tracing the characteristic run-off pattern for the line of business. In this case it is Worker's Compensation and we see the typical pattern whereby the losses decay from an early peak, but at a decreasing rate. The calendar trend graph (bottom right) shows how the portfolio is impacted by economic conditions, and particularly social inflation. The sharp increase in the 78-79 period seen here after a long period of no inflation clearly requires explanation. (In this case it was due to a change in legislation affecting WC payments.) The reporting period ends with a 9% inflation trend. The accident year levels reflect the size of the business year to year. If a pertinent exposure vector is available one would expect the accident levels to remain fairly constant, although they may still reflect differences in mix of business and underwriting and claims handling policy. The volatility recorded in the fourth graph is a result of both the inherent properties of the line of business (Workers Comp. is always going to be more volatile than Commercial Auto) and the size of the portfolio. A company with a large market share is likely to have lower volatility than one with a small market share in any given line through the effects of the "pooling or risk", the basic principle of insurance, or technically the Central Limit Theorem.



Residual display for the model above. The mean of the residuals is represented by the dark shaded "river" at the centre of the three directional displays, the lighter shades are quantiles.
The most basic diagnostic in a PTF model is the residual display, this display is also an interface for adding new parameters. It is important to see that no major trends remain after modelling. The fourth graph, residuals vs. fitted should exhibit balance around the zero line consistent with draws from a normal distribution.

Once models have been identified in PTF for several lines of business they can be analysed jointly in the MPTF framework. This begins by assessing the residual correlations and then readjusts the parameter estimates on the basis of the extra information supplied by the correlation matrix.

Final Weighted Residual Correlations Between Datasets								
Datasets oName	PL0(I)	PL1(I)	PL2(I)	PL3(I)	PL4(I)	PL5(I)	PL6(I)	PL7(I)
PL0(I)	1	-0.001453	-0.663271	-0.261735	0.145582	0.096887	0.115298	0.040202
PL1(I)	-0.001453	1	-0.233650	-0.416273	0.101330	0.053068	-0.040581	-0.075512
PL2(I)	-0.663271	-0.233650	1	0.760055	-0.229879	-0.183267	-0.038093	-0.039726
PL3(I)	-0.261735	-0.416273	0.760055	1	-0.140173	-0.189140	0.093515	0.005829
PL4(I)	0.145582	0.101330	-0.229879	-0.140173	1	0.444232	-0.221013	-0.049324
PL5(I)	0.096887	0.053068	-0.183267	-0.189140	0.444232	1	-0.056441	-0.031059
PL6(I)	0.115298	-0.040581	-0.038093	0.093515	-0.221013	-0.056441	1	0.420219
PL7(I)	0.040202	-0.075512	-0.039726	0.005829	-0.049324	-0.031059	0.420219	1
11 iterations were executed								

Example of a residual correlation matrix. The blue entries are not statistically different from zero. (Display based on simulated data).

The result is a joint model for all the lines of business, incorporating the correlations in their volatility components (on the log scale). Combined forecasting for all lines can then be performed from the joint model and this gives rise to a reserve correlation matrix which has different entries from the residual correlation matrix. In particular the reserve correlations refer to the dollar scale.

Reserve Forecast Correlations Between Datasets (Totals)								
	PL0(I)	PL1(I)	PL2(I)	PL3(I)	PL4(I)	PL5(I)	PL6(I)	PL7(I)
PL0(I)	1	0.001855	-0.339170	0.080258	0.066008	0.029733	0.088957	0.014461
PL1(I)	0.001855	1	-0.087193	-0.139630	0.042163	0.016841	-0.004793	-0.016585
PL2(I)	-0.339170	-0.087193	1	-0.009125	-0.141686	-0.104765	-0.042180	-0.023351
PL3(I)	0.080258	-0.139630	-0.009125	1	-0.050955	-0.000655	0.022626	-0.001060
PL4(I)	0.066008	0.042163	-0.141686	-0.050955	1	0.242877	-0.129163	-0.034638
PL5(I)	0.029733	0.016841	-0.104765	-0.000655	0.242877	1	0.006952	-0.012772
PL6(I)	0.088957	-0.004793	-0.042180	0.022626	-0.129163	0.006952	1	0.195034
PL7(I)	0.014461	-0.016585	-0.023351	-0.001060	-0.034638	-0.012772	0.195034	1

Reserve correlation matrix for defalt forecasts based on the same model. Note for example that PL2(I) and PL3(I) have very high residual correlations but very low reserve correlations. This is because the values in PL3(I) are much higher than those in PL2(I).

Let us now turn to a more detailed consideration of the meaning of correlation.

What is correlation?

The statistical definition of correlation is a measure of the strength of a relationship between two variables. Most commonly, correlation refers to the strength of *linear association* between two variables. Unless otherwise noted, correlation is synonymous with *linear association* in this discussion.

Correlation, linearity, normality, weighted least squares, and linear regression are closely related concepts.
Y is in a linear relationship with X if Y tracks changes in X in a proportional manner, that is, if adding a certain amount *x* to X corresponds to adding a certain amount *y* to Y, then adding double the amount 2*x* to X corresponds to adding double the previous amount 2*y* to Y, and so on. Most one-to-one relationships between variables exhibit this linearity property for very small changes, but for a linear relationship to obtain there must be linearity over the entire range of possible values of X and Y. When X and Y have a linear relationship we can express Y as a function of X using an affine function: $Y = aX + b$, where the proportionality factor *a* is usually called the *slope* and *b* the *intercept*. These names obviously come from the Cartesian representation of such a function as a straight line.
In practice one or both of X and Y are derived from empirical data and so the relationship between them is not exact. In this case we often perform linear regression by least squares in order to determine the best approximation to a linear relationship between X and Y.
The method is called "least squares" for the following reason. If we take any affine function $f(X)= aX + b$ the degree to which this fits the relationship of X and Y is determined by the set of differences, $Y - f(X)$, by which the function misses the true value of Y for all the various values of X. These differences are called *residuals*, or sometimes *errors*. Generally as we move over the range of values of X the function $f(X)$ will sometimes fall short and sometimes overshoot the corresponding values of Y. In these cases the residuals will be positive or negative, respectively. Out of all possible affine functions the best one will be the one that in some sense makes the set of residuals smallest. To guage the size of this set we need to treat positives and negatives equally and not allow them to cancel each other out. One simple strategy is to equate the size of the set with the average of the squares of the residuals. If we follow this strategy to choose the best affine function, which means choosing the best pair of slope and intercept, we are doing *least squares linear regression*.

This method is best suited to situations where the underlying relationship is linear but this linearity is obscured by random fluctuations called *noise*. This noise may result from the measurement process or it may be inherent to the physical processes giving rise to the data. The noise factor by definition does not have predictable values, but it may be described in a precise manner by way of a probability distribution. If, as is commonly the case, this distribution is the Normal distribution then the least squares method fits perfectly with it, in the sense that it accurately uses all the information contained in the data about the underlying linear relationship in forming its estimate of the slope and intercept. If the relationship is linear but the noise distribution is not Normal the least squares method will usually produce close to optimal results.

The (linear) *correlation* is a number between -1 and +1 which the describes the goodness of fit between the data and the least squares linear regression function. Its sign tells us whether we have a negative or positive correlation, that is whether Y changes in the opposite or the same sense as X, and its magnitude tells us how clear the relationship is. For a given slope the correlation decreases if the intensity (technically, the variance) of the noise increases, and for a given noise level the correlation increases as the slope increases, since a steep dependence of Y on X is less masked by noise.[2] A correlation of zero, corresponds to no relationship of X and Y. It means that the value of X has no bearing on the linear approximation of Y, in this case we (linearly) estimate the same value of Y for all X, namely its average.

In situations where the underlying relationship is not known to be linear or is known not to be linear the least squares linear regression can still be applied but its results may be quite uninformative since they represent the best estimate for a linear relationship.

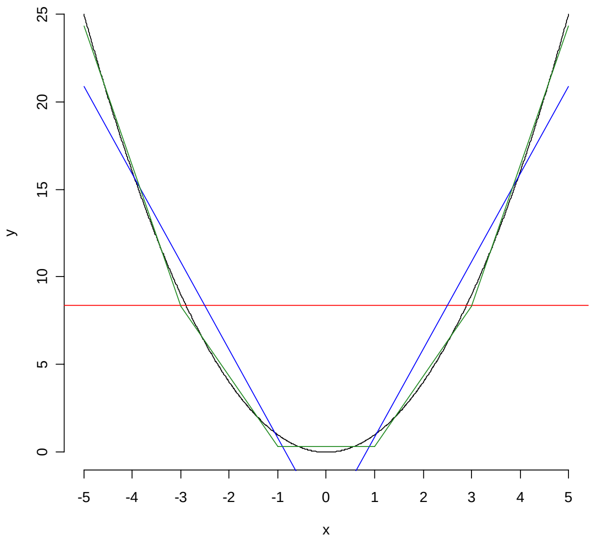
For example say Y is an exact squared function of X, ($Y = X^2$). In this case we might want to say that Y is perfectly correlated with X, since a knowledge of X yields complete knowledge of Y. If however

[2] The correlation of X with Y is the same as the correlation of Y with X, even though the regression lines have different slopes. This is because the noise intensity (variance) is measured relative to the dependent variable and so is different in each case.

the range of X is symmetric about zero then the linear correlation is zero and the regression line is horizontal.

The function $Y=X^2$ over the range (-5,5) as displayed below (the black line is the $Y=X^2$ curve). The linear correlation between Y and X is zero for X over the entire range and the red line is the least squares regression line.

The two blue lines show the approximation of the $Y=X^2$ curve over the range (-5, 0) and (0, 5) respectively. Clearly the linear correlations between the X ranges and the Y points they estimate are strong and in in fact are equal to -0.96 and +0.96 respectively. This correlation coefficient can be calculated from the X values 0,..5 and the corresponding Y values 0, …25, and equally well by comparison of the fitted values on the regression line and the Y values. The slopes of the two blues lines are -5 and +5 and each has intercept -3.3333.



If we split the X series into even smaller segments over this range, we could end up with the connected green lines (even stronger correlation, except for the interval (-1,1) which has zero correlation as it is symmetric about zero).

The above example highlights another point about correlation. **Correlation does not give us an idea of how good our approximation is to fitting our data!** The data, and predictions, must be plotted.

A mathematical transformation of the values of Y and/or X may render a non-linear relationship into a linear or almost linear relationship that can usefully be approximated by linear regression. In the example above if we take square roots and adjust the sign, so that Y_2 = +_ sqrt(Y) [where we choose + or – according to the sign of X] then Y_2 now has a perfectly linear relationship to X which can be discovered by regression and Y can be easily computed from Y_2.

A very important non-linear relationship, especially for every kind of financial data is the power relationship. In this case Y tracks changes in X not according to additive changes in each variable but by multiplicative or percentage changes. In this case if an *x*% change in X corresponds to a *y*% change in Y, then a 2*x*% change in X corresponds to a 2*y*% change in Y. The formula for the underlying relationship in this case if $f(X) = bX^a$ and in the real-life noisy instances of this relationship the noise is multiplicative as well, that is random percentage fluctuations.

This kind of relationship can be effectivly treated as a linear regression by transforming both X and Y by taking logarithms: $Y_2 = \ln(Y)$, $X_2 = \ln(X)$, then the underlying equation becomes $Y_2 = aX_2 + \ln(b)$ which is of the linear type.

Financial data is often in the form of time-series where the Y is in dollars and the X is in time units, years, quarters, days, etc. In this case we are dealing with *trends* where the basic case of a constant trend is if an additive change of *x* in X corresponds to a percentage change *y*% in Y then and additive change of 2*x* in X corresponds to change of 2*y*% in Y. The underlying equation this time is $f(X)=ba^X$. Again this can be rendered into a linear relationship by taking the logarithm of just Y: $Y_2 = \ln(Y)$, underlying equation is $Y_2 = (\ln(a))X + \ln(b)$.

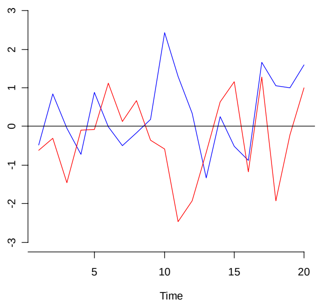
Transforming Y or X onto another scale may also significantly alter the correlations. For example, if X and Y are normally distributed with linear correlation *r*, the linear correlation between the corresponding lognormals Exp(X) and Exp(Y) can be much lower than *r*, depending on the standard deviations of *r*.

How do correlations relate to loss arrays? Process, parameter, and reserve correlation?

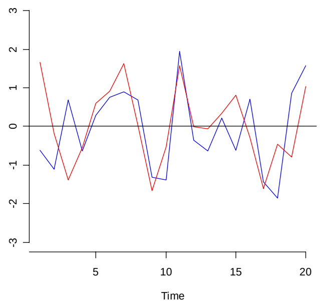
Loosely speaking correlation can be thought of a measure of the tendency that two loss arrays will fluctuate in the same direction (at the same time). A high positive correlation indicates fluctuations in the same direction; a low negative correlation indicates fluctuations in opposite directions. This is easiest explained via example.

Example 1

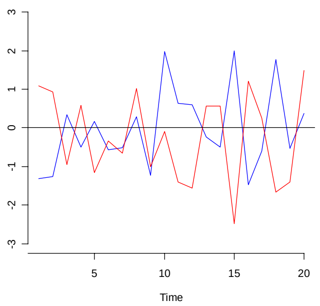
Consider the following graphs of pairs of time series on the same axis.



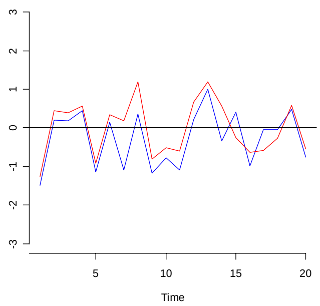
(1)



(2)



(3)



(4)

Each series consists of 20 points from a time series with mean zero and standard deviation of one (ie a random portion of a loss development array after average trends have been removed – see below). Simulations were done until the sample correlation was close to the expected correlation.

(1) Series correlation =0, sample correlation =0.02
(2) Series corr. = 0.5 sample corr = 0.5
(3) Series corr. = -0.5, sample corr. = -0.51
(4) Series corr. = 0.8, sample corr. = 0.81

A high positive correlation produces almost parallel fluctuations as in (4) above, whereas a negative correlation indicates a tendency for fluctuations to go in opposite directions, as in (3) above. Note however that even though it is possible to see evidence for the negative correlation in (3) by inspection of the graph, it requires some effort to distinguish it. Similarly, it is likely that many people would ascribe a correlation to the two unrelated series in (1) due to a tendency of the brain to infer a common pattern on the basis of slight evidence.

The correlation should concern only the random portion of the data, which means that before measuring it we need to remove all the predictable trend components. If this is not done then correlations will appear to be greatly exaggerated or appear to be present when in fact they are not. We call the correlation of the random component of two loss arrays: **process correlation**.

In the next example, we examine the relationship between two loss development arrays. One loss development array is Gross data the other array is the same data 'Net of Reinsurance. Since Net of Reinsurance data is a subset of Gross data, we expect strong process correlation.

Example 2: Common drivers - Gross vs Net

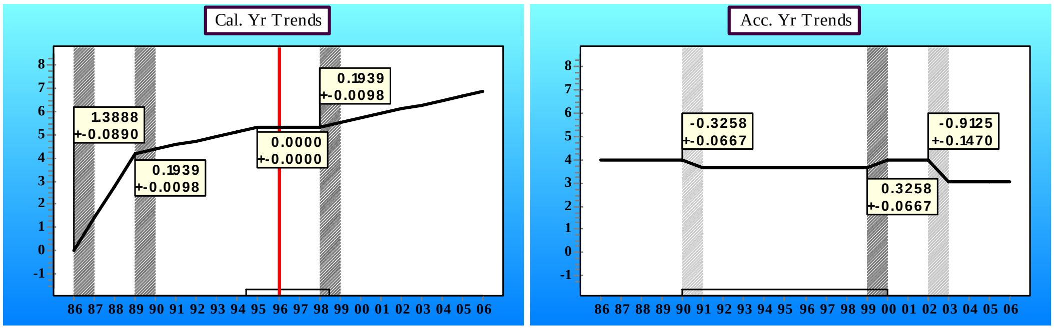
A high correlation between two lines of business is often said to be caused by "common economic drivers" acting on the liablity streams. In terms of our three-way trend analysis of loss arrays this translates largely to common calendar year trends.

There may also be observable correlations in accident year levels, however once these are taken account of by modeling or by the associated exposure vectors they exert no further influence on the forecast values used in completing the array and so they don't contribute to the final reserve correlation. The same applies to common patterns in development year trends.

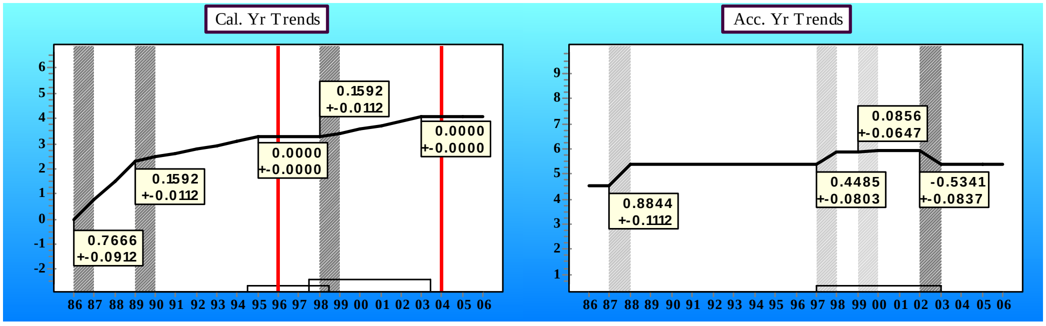
Evidence of common economic drivers where they exist should be apparent in a shared pattern of calendar year trends and should also be accompanied by high process correlation. We will give an example to illustrate this below, but we also wish to stress that process correlation is mostly found in the absence of such evidence. Two lines can have quite different calendar trend profiles but be connected in such a way that on a year to year basis the fluctuations about the respective trends go in generally the same way, or generally in opposite ways. For example Private Passenger Auto. (PPA) and Commercial Auto. Liability (CAL) are often found to be positively correlated, even though they are impacted differently by economic and social inflation and so have different calendar trends. A bad period for PPA in companies where this correlation is found is more often than not also a bad period for CAL, and likewise for good periods.

A clear example of common drivers can be found when comparing Gross with Net (of reinsurance) for the same LOB.

In the first set of graphs below we compare the calendar trends and accident year levels as found after (automatic) PTF modeling of the two parts.

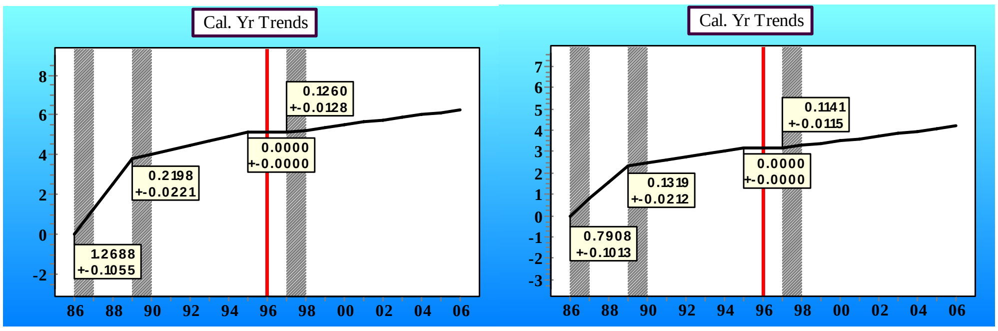


Above: Calendar year trends and accident year levels in the PTF model for the Gross data.
Below: The same trends as they appear in the PTF model for the Net data.

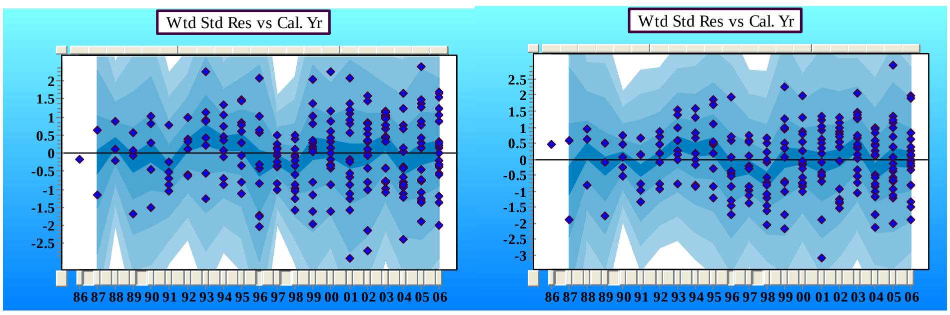


Prima facie evidence of common economic drivers is seen in the similar pattern of the calendar year trends for the Gross and Net in the graphs above. Apart from the final zero trend in the Net, model parameters occur in exactly the same places and with the same relative magnitudes in both cases. The intensity of the trends is greater however in the Gross model.

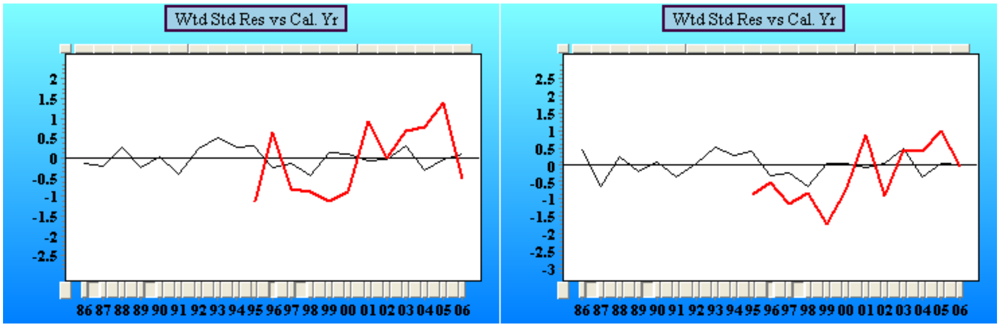
When we turn to MPTF modeling for the two datasets simultaneously the observed process correlation enables sharing of information in identifying the precise trends producing more accurate models. The calendar year trends as identified by MPTF are seen below, Gross on left and Net on right.



The MPTF joint model shows and even more marked similarity in the calendar trend structure, furthermore it enables us to measure the process correlation at the high positive value of 0.86.
If in each case we alter the model by allowing only a single average calendar trend then the measured process correlation increases to 0.88. This illustrates in a small way the fact that accurate modeling lowers the magnitude of the observed process correlation. If a statistically significant trend has been ignored by the model then it returns as a trend in the residuals, and two sets of such residuals will have artificially inflated correlations.
The common structure in the calendar year trends that we see above is not the same as significant (positive) process correlation and could in theory exist without it, although that would be quite unusual. If we wish to observe the effects of high process correlation directly we can do so via the residual plot.

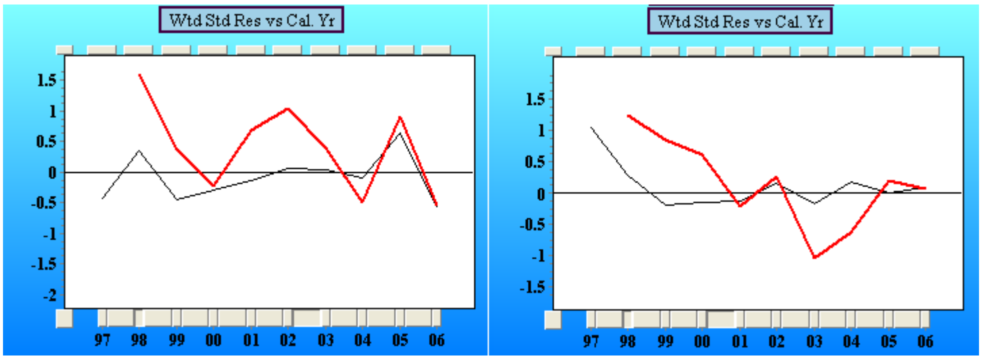


Residual plots for Gross (left) and Net (right) by calendar year. The background contours show the central tendency for each year and we can see that the pattern of these is very similar for the two datasets. This is still not quite the same as process correlation since the latter is based on a comparison of individual cell values across the two sets. In the next display we have run a line through all the residuals from the accident year 1995 and removed the other data points, so that the relationship can be seen clearly.



Residuals by calendar year: Gross (left), Net (right). The red line goes through all the residuals from values in accident year 1995. The thin black line is the annual mean of the residuals.

In the example below we see an example high process correlation with no accompanying evidence of common economic drivers.



Swiss Re, CAL (left) and PPA (right) residuals by accident year, mean with accident year 98 shown in red. We have omitted the parameter bars but their position can be seen from the buttons below the graph, one new trend starting in 02-03 for CAL and one starting in 00-01 for PPA. After modeling the two datasets have a correlation of 0.55 which is typically about as high as one encounters with real data. The reserve correlation in this case is 0.31, it is always lower than the process correlation since the total reserve is the sum of many different forecast cells.

Capital Allocation and Risk Capital Allocation

Risk Capital is money held in reserve over and above the estimated mean reserve required for servicing claims. It is a hedge against volatility in claim aggregates.
For a single line of business the amount of risk capital to be held at say the 98th percentile, so as to cover a 1 in 50 year adverse development, can be calculated from the (tail of the) reserve distribution.

There are however two other important practical considerations in the allocation of and managent of risk capital.

If there are multiple lines of business then the allocation of risk capital by line is not the same as the allocation of mean reserve by line. This can be explained with a simple example. Let's say that all our reserve distributions are such that the 98th percentile is reached at twice the standard deviation above the mean.

Line A requires a mean reserve of $100m, and has a CV of 5%. The standard deviation is $5m, and so risk capital at the 98th percentile is $10m.

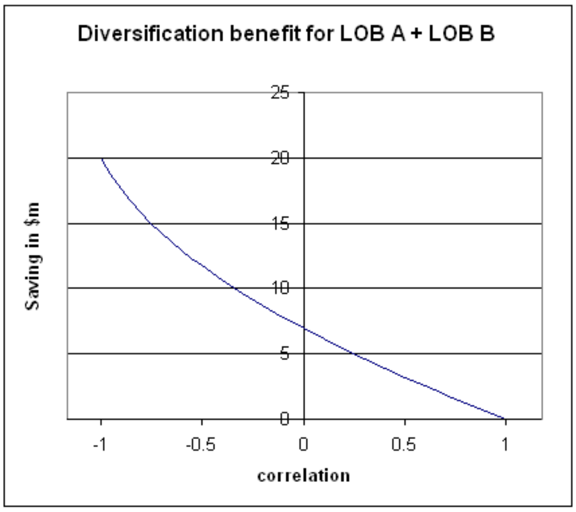
Line B has a mean reserve of $50m, but is more volatile and has a CV of 15%, the standard deviation is hence $7.5m and risk capital at 98th percentile is $15m.

Thus we see that line B requires 50% more risk capital than line A for the same level of risk, despite having only half the mean liabilities.

Say that lines A and B are the only two lines written by our company. What is the total risk capital requirement at the 98th percentile? It might seem that the answer is $25m = $10m (line A) + $15m (line B), but this is not in fact correct.

The standard deviation of a sum of two or more random variables is not the sum of the standard deviations, unless the variables are perfectly correlated, which is rare in reality.
If A and B are *independent* the standard deviation of the sum is very close to $9m, and hence the risk charge would be $18m.

The saving of $7m in risk charges in the example above is a benefit of diversification. In practice the exact value of this benefit depends on the linear correlation between reserve distributions for the two lines, as indicated in the graph below.



Diversification benefit for a company with two lines of business is highly dependent on the correlation between the reserve distributions of the lines.

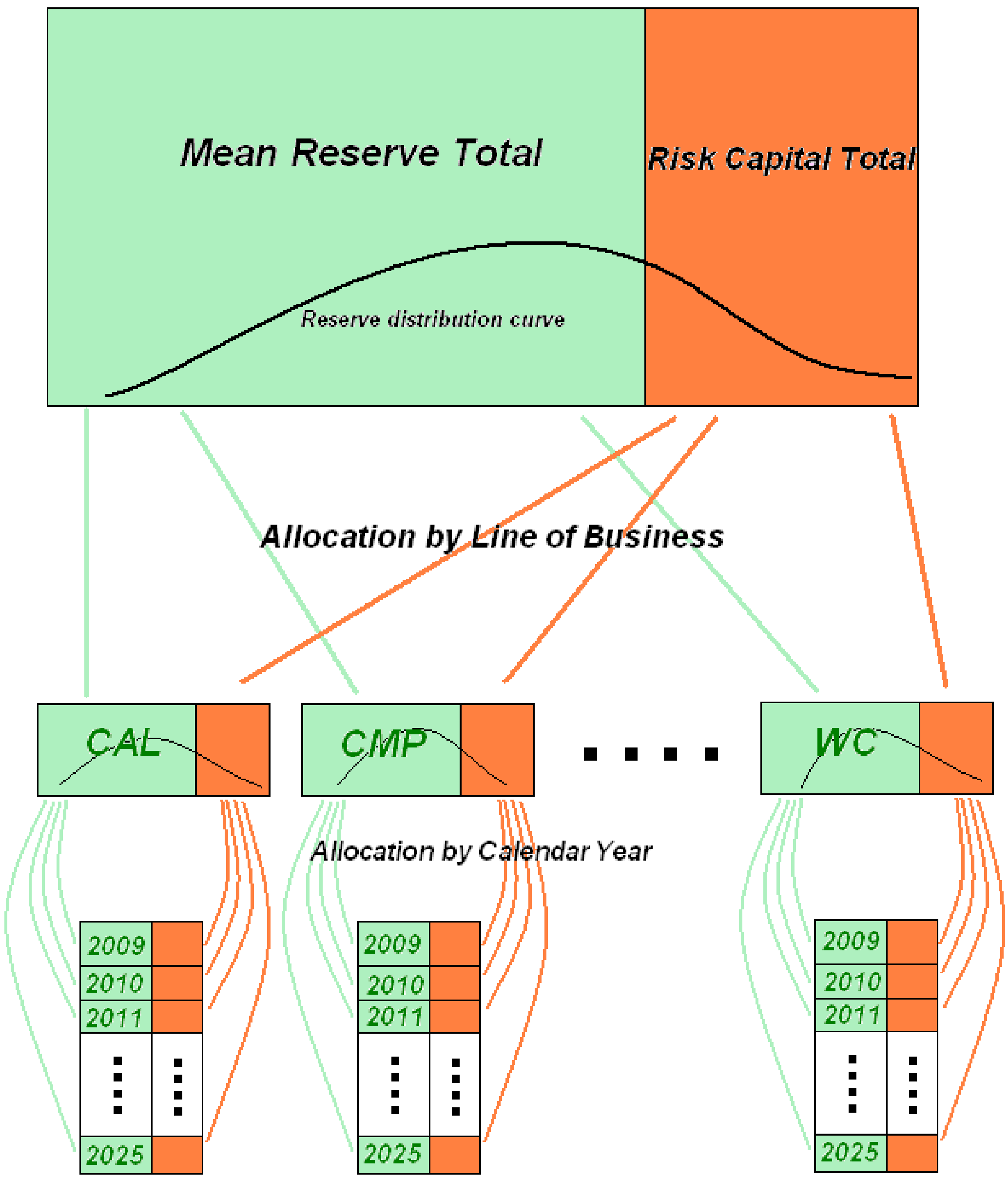
The second consideration applies even to a single line of business. The mean reserve figure estimates the total amount of outstanding and future liabilities for a set of policies and as such it is an aggregate of the means for each future year until all claims have been exhausted. If it concerns a long-tail line of business this may stretch over a considerable period of time. Whether the period is long or comparatively short prudent management requires a breakdown of the claims distribution by calendar period going forward. If the liabilities are spead fairly evenly over the mid-term horizon then greater benefits can accrue through investment of unearned premium than is the case when most claims fall due in the near-term.

The division of reserve and risk capital by calendar year is analogous to the division by line of business. The CVs of individual calendar year estimates are different and the distributions for individual calendar years are correlated with each other, due to the existence of common trends.

The breakdown of reserves and risk charges by line of business and accident year is illustrated below. The green lines connect the reserve means and these are additive in the sense that the sum of the means at one level is equal to the mean at the next higher level.

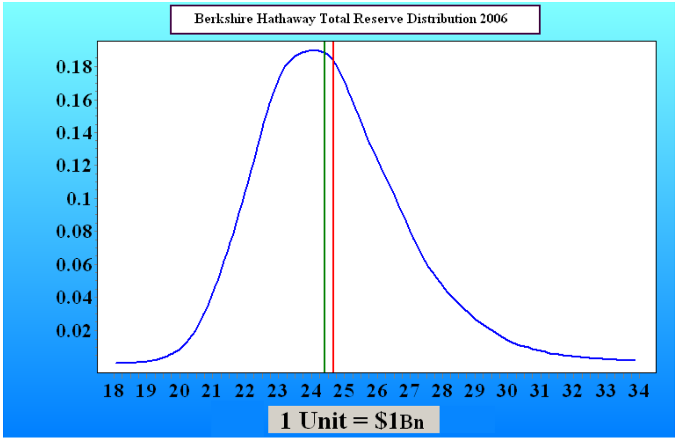
The orange lines connect risk charges. If the risk charges were all computed at the same percentile then they would be subadditive in the sense that the sum of the charges at one level would be more than the charge computed for the next higher level. It is more natural to determine the total risk capital at the highest level and then allocate it down to individual parts of the business according to a formula which takes into account the individual distributions and correlations of the constituents. If the risk charge for the Total is at the 98th percentile the effective risk for each individual lines calendar year payment may be at the 90th percentile. We can be more lax at the lower levels because diversification effects operating across years as well as across lines, with high and low years averaging out over time.

In either case the proportions represented by the orange lines are quite different from those represented by the green lines.



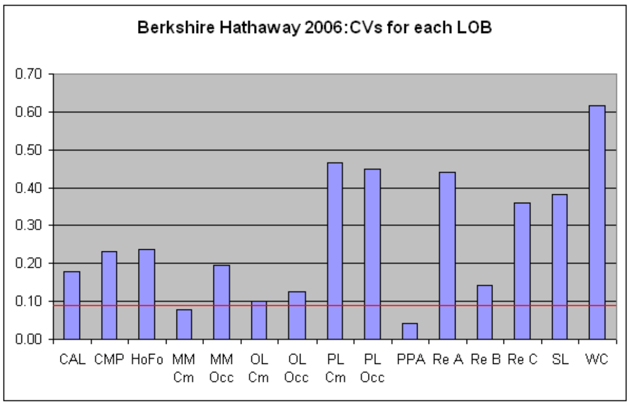
We illustrate the above with a comparative study of risk charges for Berkshire Hathaway and The Hartford based on A.M. Best 2006 Schedule P data.

Berkshire Hathaway.
Total reserves = $24.7Bn, CV = 9%

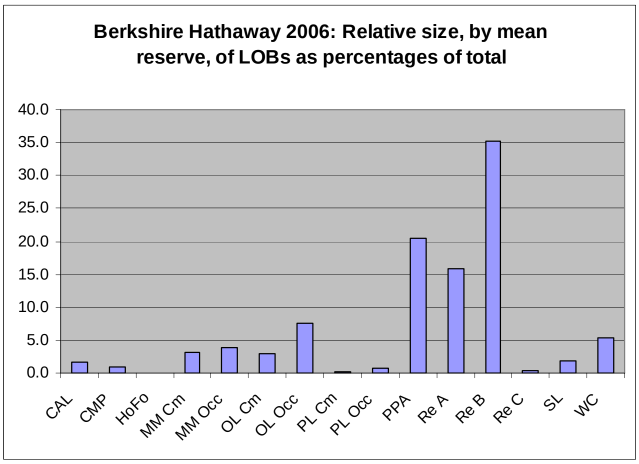


Total Reserve distribution for Berkshire Hathaway 2006, kernel smoothed curve based on 50,000 simulations from the joint distribution for all LOBs. The green line indicates the median and the red the mean of the distribution. Their distance apart is an indication of the degree of skewness.

The 98th percentile is $5.64Bn above the mean. If the distribution were Normal this figure would have been considerably lower at $4.57Bn, so we see that the effect of the heavy upper tail in the distribution is to increase the risk charge.

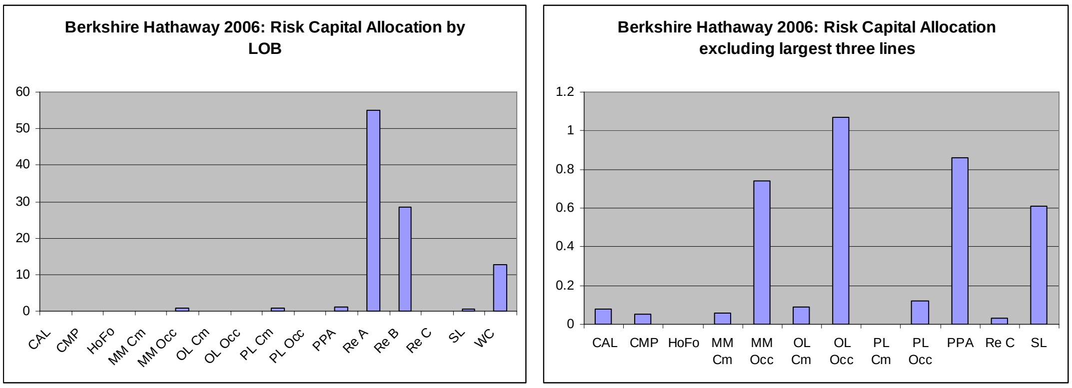


CV's for each line making up Berkshire Hathaway. The horizontal red line is at 0.09 the CV of the total. Note that this is quite a bit lower than the average of the CVs of the individual lines.



Split of the total mean reserve by Line of Business. Note that of the largest five lines, Re B, PPA, Re A, OL Occ and WC, only PPA has a CV below that of the total.

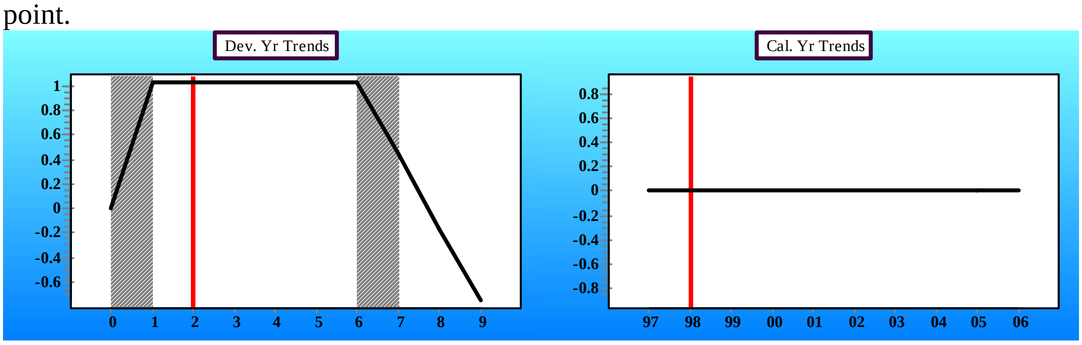
The spread of risk capital by line of business combines the information in the two previous graphs with the measured correlations between lines to produce the split amongst lines in the graphs below. Because three lines, Re A, Re B and WC soak up 96% of the risk capital needs we have added a second chart excluding these so that the relations among the other lines can be seen. Note that PPA, although the second biggest line in terms of reserves belongs in the less risky category.



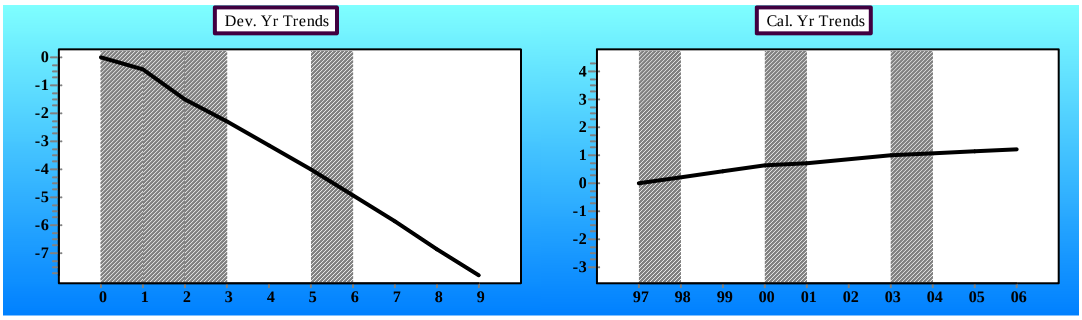
Risk capital allocation by line of business for Berkshire Hathaway 2006. The graph on the right excludes Re A, Re B and WC which together account for 96% of the need for risk capital.

The allocation of reserve by calendar year is a crucial business decision as it affects asset/liability matching and investment strategy and hence the core profitability of the company. Each long-tail line has its own cross sectional liability profile determined by the shape of its development decay and inflation trends.

A comparison of the two largest Berkshire Hathaway lines, Re B and PPA, suffices to make this point.



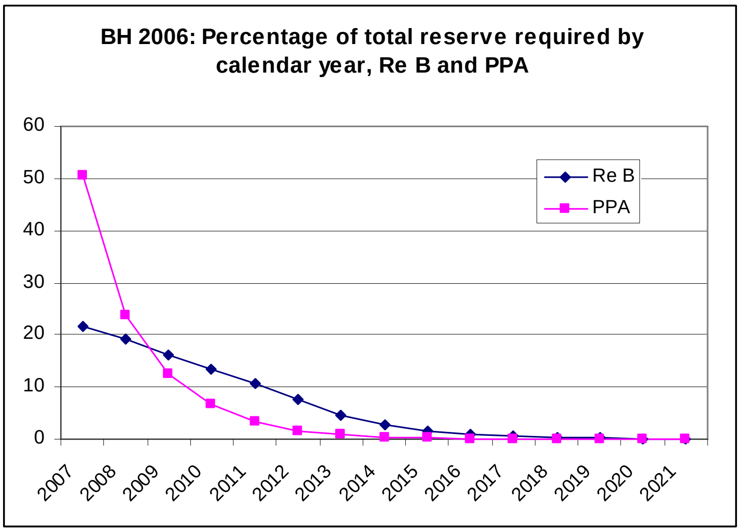
Berkshire Hathaway 2006: Re B development year and calendar year trends. The grey bars denote trend changes, the red lines the start of a period of zero trend.



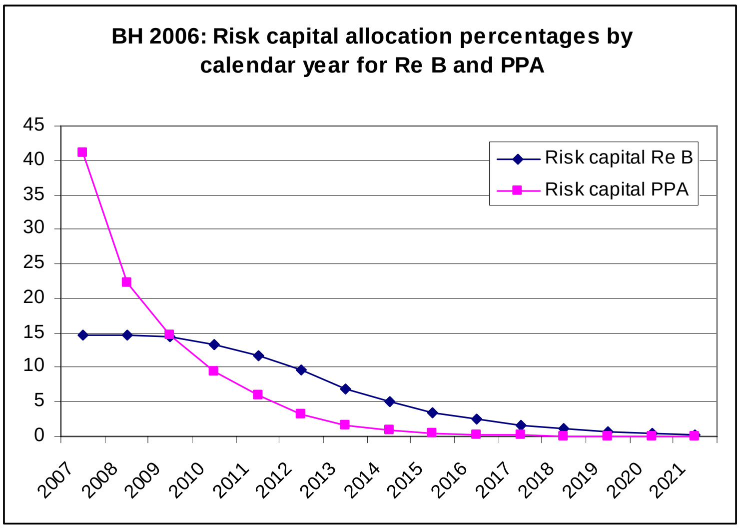
Berkshire Hathaway 2006: PPA development year and calendar year trends.

Re B and PPA represent two extremes of trend pattern. Re B peaks after one year but remains at this level for the next six years before beginning to decay. The decay in the following three years is steady

but relatively small. Inflation however appears not to be a factor here. We can expect the calendar year distribution to be relatively flat.
PPA on the other hand decays moderately in the first year and then very sharply in the remaining years. A steady inflation figure (about 7%) may take some of the edge off this rapid decay, but we still expect the calendar year distribution of liabilities to to be mostly crowded into the first two or three years.

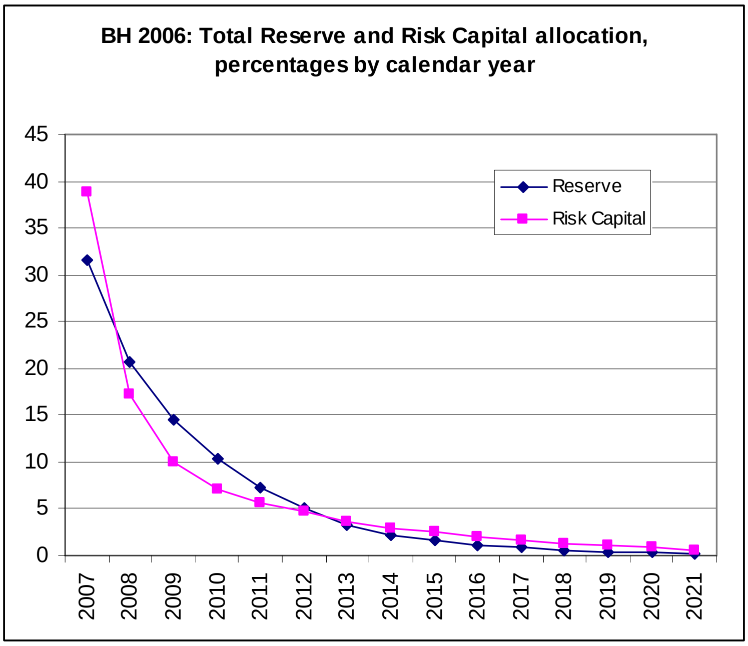


Berkshire Hathaway 2006: Percentage of the total reserve requirements by calendar year for the policies written between 1997 and 2006. As expected from the development and calendar year trend profiles the payments for PPA are far more concentrated in the early years. Average time to payment is 1.5 years for PPA and 3 years for Re B.

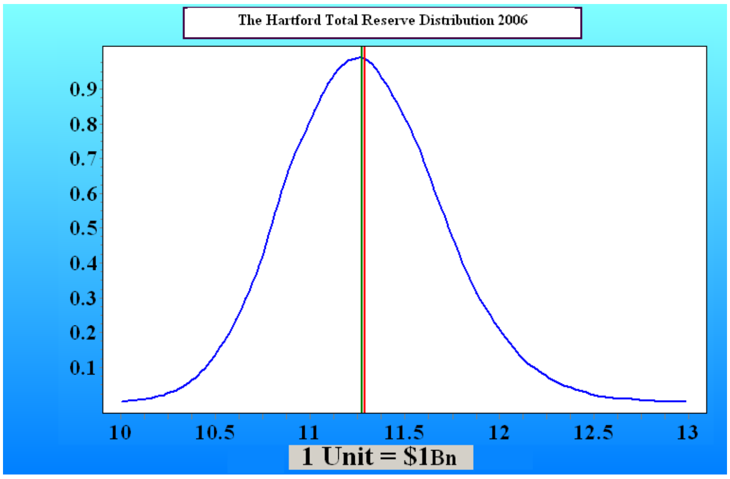


The risk capital allocation for the same two Berkshire Hathaway lines by calendar year shows the same general pattern but is somewhat flatter than the reserve allocation. This is because the CV of the later calendar year forecasts is generally higher than the earlier ones.

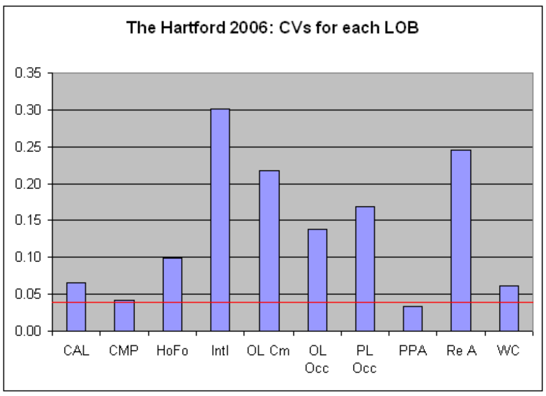
Finally let us look at the allocation of total reserve funds by calendar year and compare this to the allocation of risk capital. We can plot these, as percentages of the totals on the same graph.

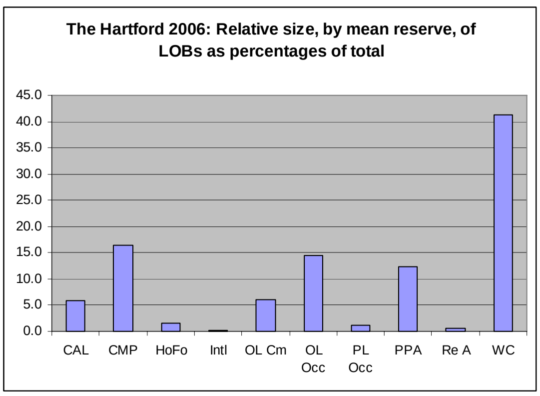


The Hartford
Total reserves = $11.3Bn, CV = 4%

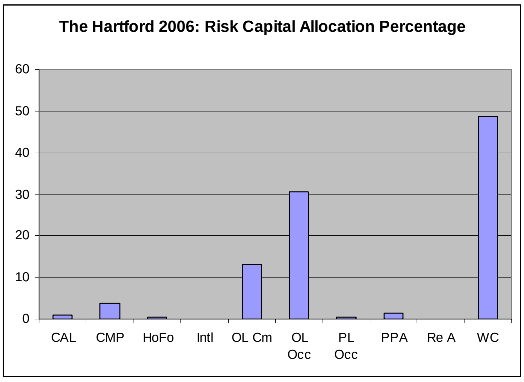


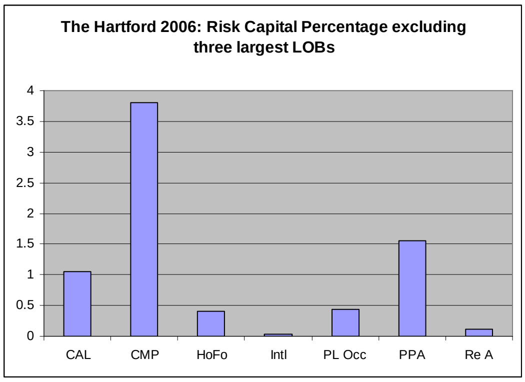
The Hartford 2006: Total reserve distribution. In comparison with BH above two differences are immediately apparent, the distribution is far less skew and far more concentrated around its mode. The 98th percentile is $903m above the mean. If the distribution were Normal this would be $844m, so the upper tail of the distribution is only marginally heavier than that of the Normal.





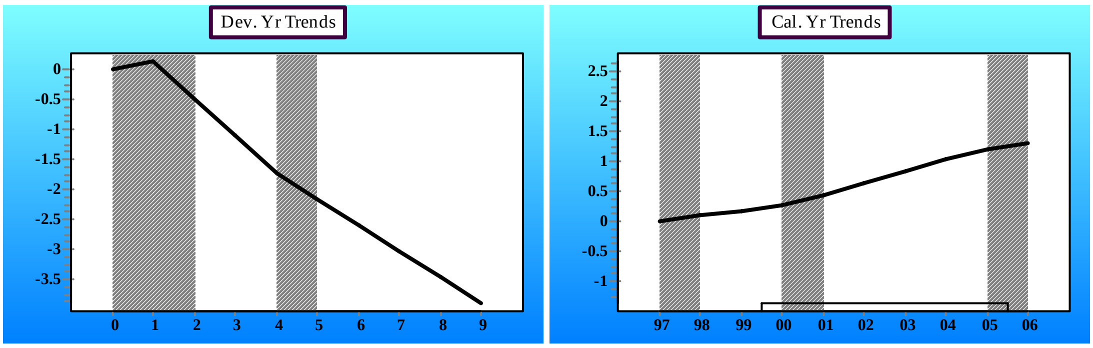
Relative sizes and CVs of the mean reserve for the LOBs making up The Hartford. There are fewer LOBs than for BH and the business is distributed somewhat more evenly over them. The two largest lines, WC and CMP are among the lowest in CV, which contributes to the very low CV of the aggregate.



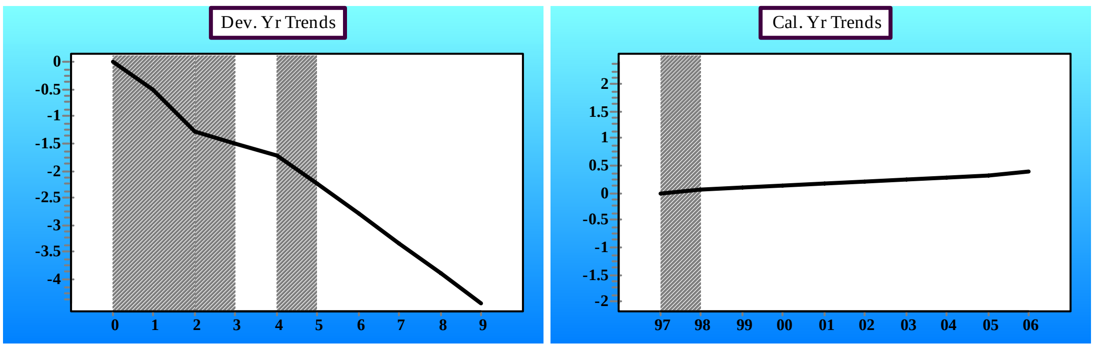


For The Hartford the three most risky portfolios soak up 93% of the risk capital (compare 96% for BH). Note that OL Cm is only the 5th largest LOB by reserve but the 3rd largest in terms of risk capital requirement.

The two largest lines are WC and CMP and these are ones we will compare in terms of calendar year liability streams.

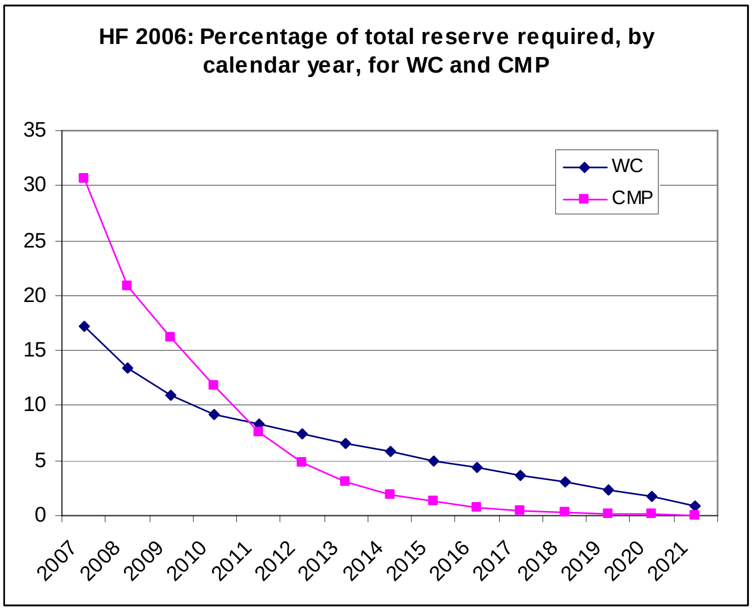


The Hartford 2006: Development year and calendar year trends for WC

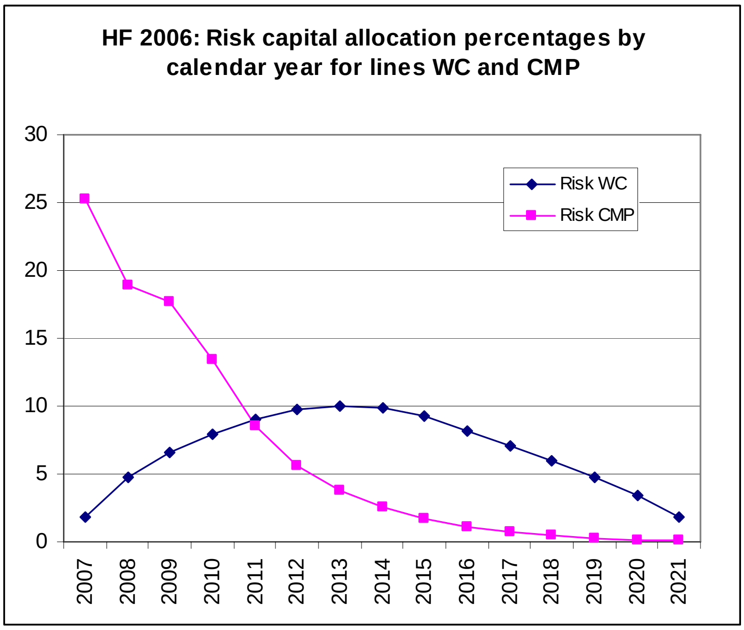


The Hartford 2006: Development year and calendar year trends for CMP

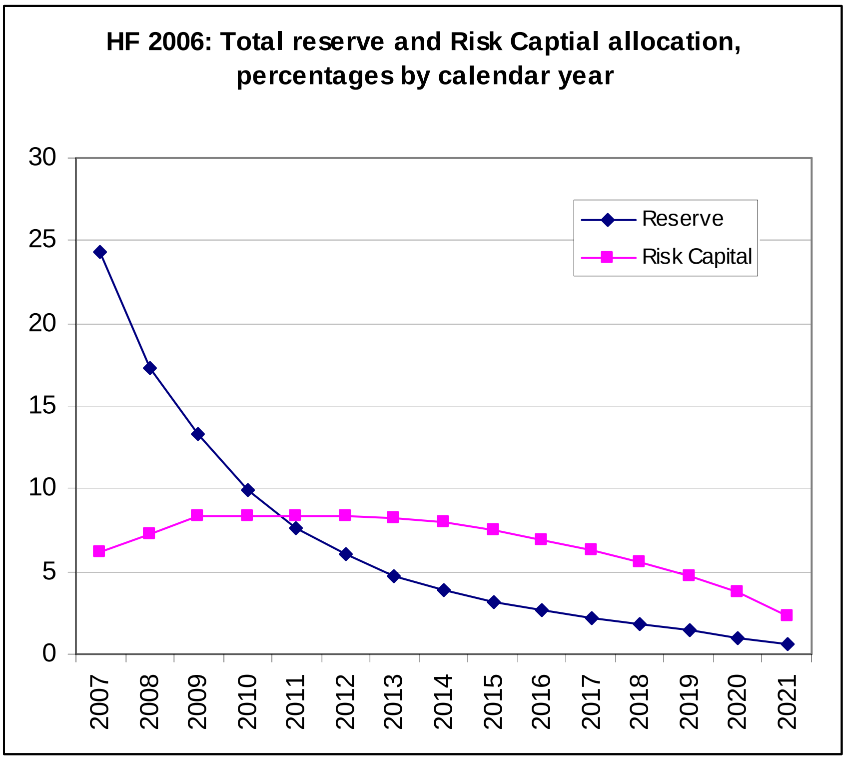
Apart from a small increase in payments after the first development year for WC, both lines show a steady decresing trend in liabilities. Comparison with BH PPA, shows that this trend is moderate, being at about half the rate seen in the BH line. In both cases there is a consistent positive inflation trend, it averages about 14% for WC and about 4% for CMP. We expect the payment streams by calendar year to be fairly spread out, but more so in the case of WC.



The Hartford 2006: reserve requirements as a percentage of total for the two largest lines of business. Note how much flatter the curve for WC is. The development decay is similar in the two lines but (social) inflation effects are much more pronounced for WC, which increases the mean payments in later years.



The Hartford 2006: Risk capital requirements as a percentage of total for the two largest lines of business. The stronger inflation effects experienced by WC liabilities lead to the rather dramatic result that the most risk capital is required to handle claims that come in six or seven years after the end of the underwriting period.



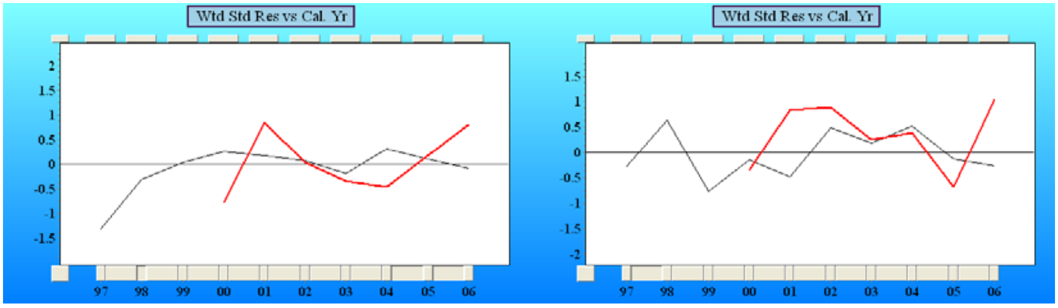
The Hartford 2006: The allocation of reserve and risk capital for The Hartford shows a much flatter contour than was observed for Berkshire Hathaway. This is a more desirable result and is due to the particular mix of risks favoured by the smaller company. The difference warrants further analysis but at this stage we can say that it is partly due to the very large contribution of PPA to the BH portfolio, PPA having an essentially short-tailed development profile.

No two companies are alike and no company is the same as the total industry.

I this section we will have a look at some instances where the correlation between two lines of business is markedly similar to or different from that seen in the industry.

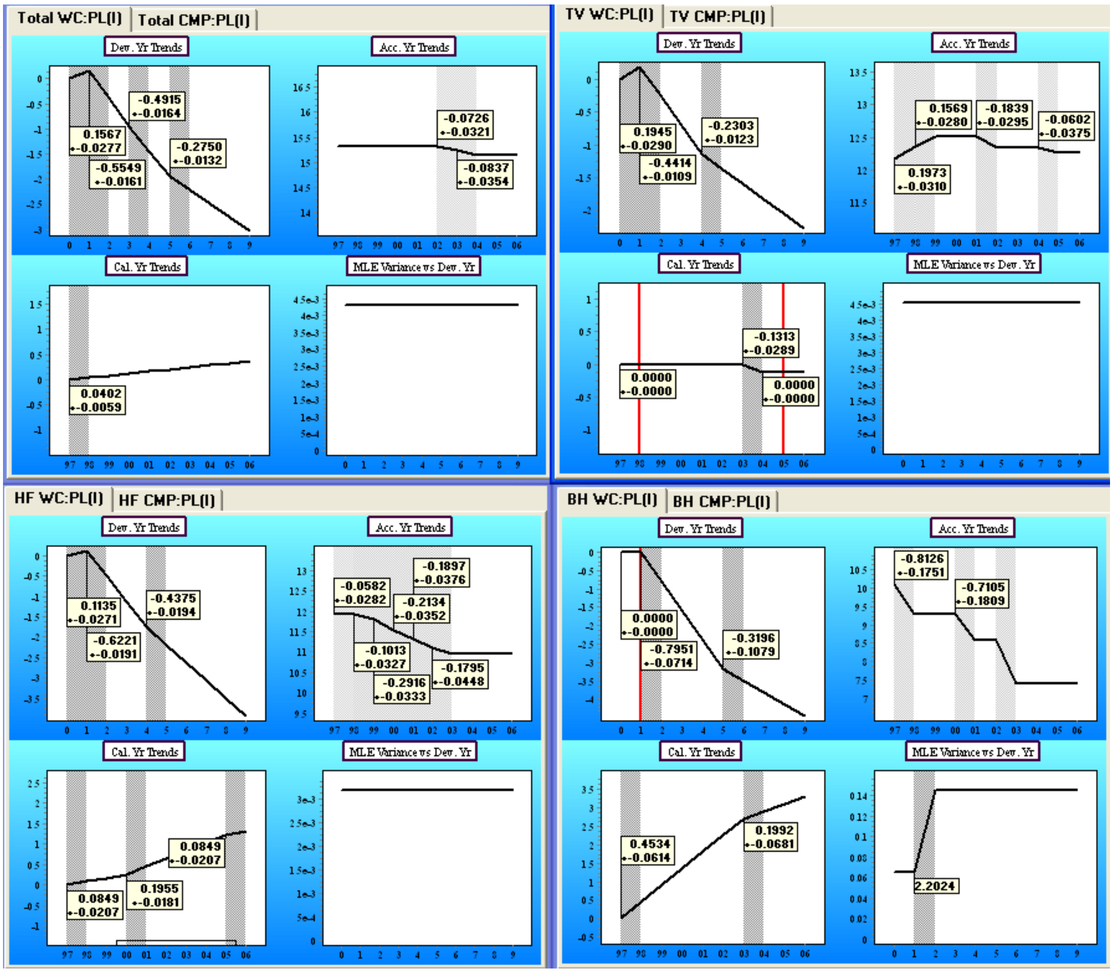
We begin with a study of Workers Compensation (WC) and Commercial Multi-Peril (CMP), lines which one might expect to bear some relation to each other.

MPTF analysis of the Total US industry figures from A.M.Best 2006, shows evidence of a high positive (process) correlation of 0.46.
We can exemplify this graphically in the display below in which we show, for each line, the calendar year trends after modeling and the trace of the accident year 2000 residuals on the same graph. The calculated correlation figure of 0.46 encompasses all the residuals from the 10 year reporting period, but we have chosen one year from around the middle of this period so that the display is easy to read.

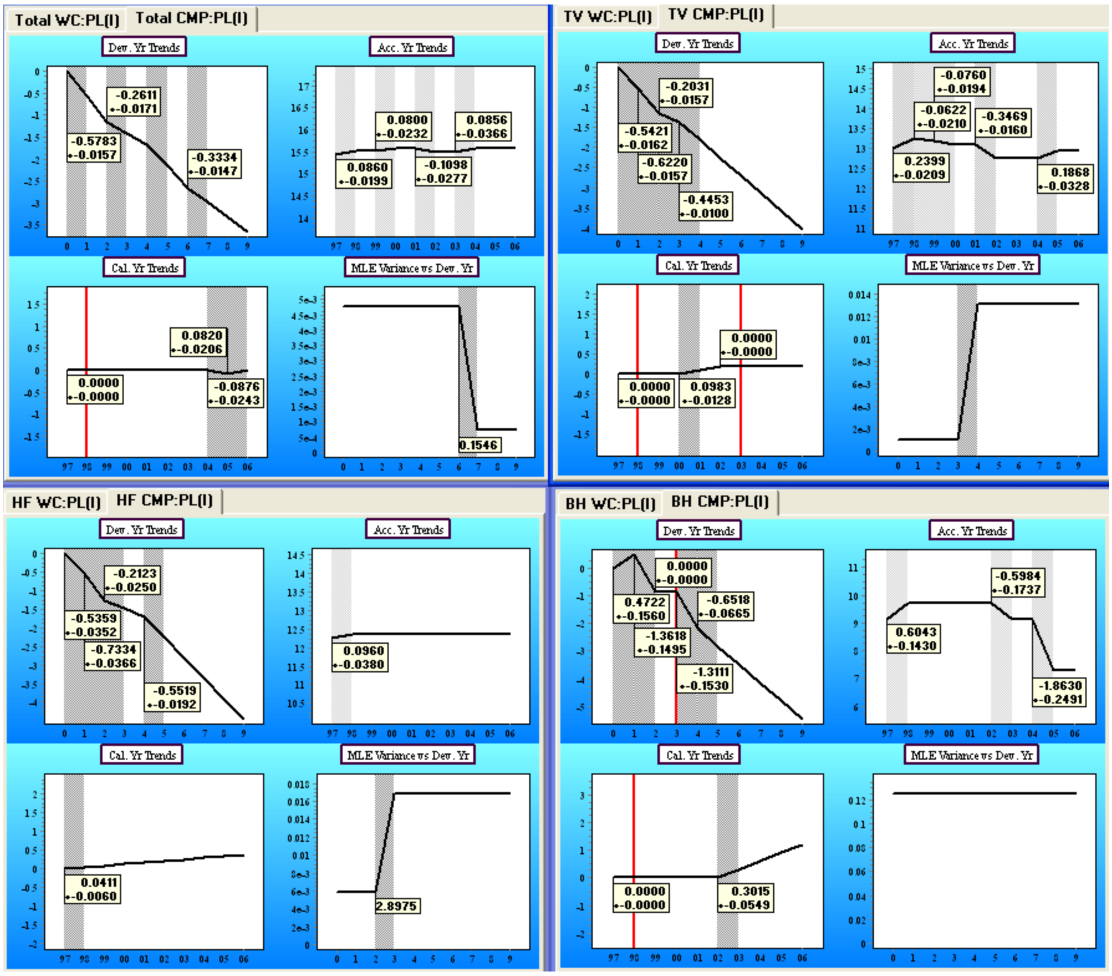


Calendar trend, after modeling (black) and residuals for accident year 2000 (red). Total US Industry, CMP (left) and WC (right).

Before going on to see if this correlation can be found in individual corporations we will pause and look at some model displays for the industry and selected companies in these two lines. These constitute an important check on the consistency of the models and alert us to differences and similarities in the individual performance of companies which are not covered under the rubric of correlation.



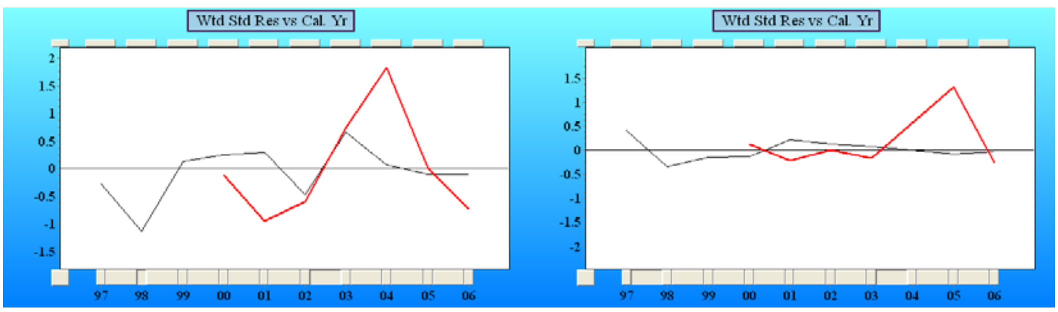
These four model displays show the WC LOB in (L to R, Top to Bottom) Total US industry, Travellers, The Hartford and Berkshire Hathaway. We note that the development pattern is very similar in all models, this is characteristic of the type of insurance. In all instances except Travellers there are positive calendar trends, although the patterns are different. Similarly the accident period patterns all show an overall drop over the period but in very different ways.



Model displays for the CMP LOB for the same four lines of business. The development patter for BH is somewhat different from the rest which indicates that it is writing a somewhat different mix of risks. The accident year patterns for TV and the Total Industry show a marked similarity in shape. TV represents about 10% of the industry so perhaps this is not surprising. TV and the Industry have a 0.28 correlation in respect of CMP.

We stress again that correlations are what is found after these individual trend patterns have been accounted for.

Turning now to Berkshire Hathaway we produce the calendar year displays for the same two lines WC and CMP, with representative accident year 2000 shown in red.

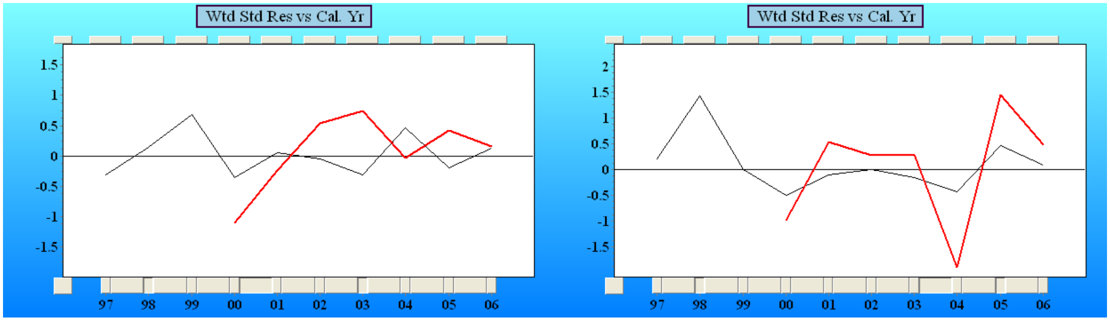


Calendar trend, after modeling (black) and residuals for accident year 2000 (red). Berkshire Hathaway, CMP (left) and WC (right).

Again the red lines indicate the likelihood of a positive correlation and this is confirmed by an MPTF computation from the entire set of residuals which yields a correlation of 0.36.

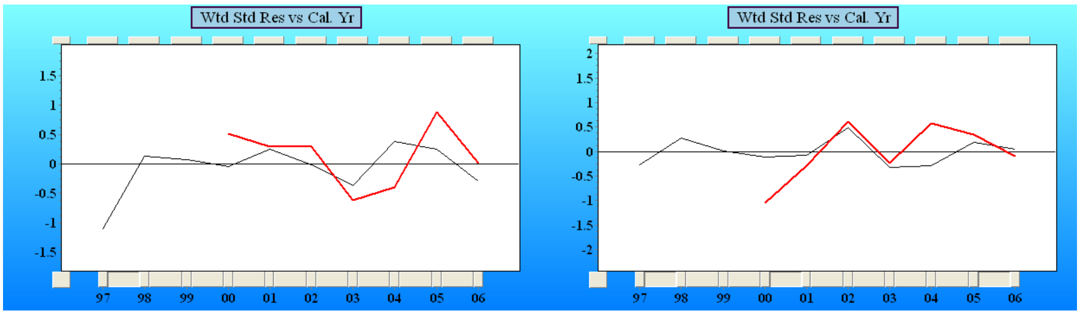
Comparing Berkshire Hathaway with its down-up-down movement in 2000-2006 and the Total Industry it appears as though there may be a negative correlation between them in one or both lines, however a full comparison yields only a statistically insignificant small negative correlation in each case.
Berkshire Hathaway's market share in these lines of business is quite low, at around 1% for WC and 0.5% for CMP (based on Ultimates in the reporting period).

Travellers and The Hartford are the two largest companies, by ultimates, in these two lines. Travellers has about 5% of the total industry in WC and 10% in CMP, and The Hartford at around 4% and 5% respectively.



Calendar trend, after modeling (black) and residuals for accident year 2000 (red). Travellers Insurance, CMP (left) and WC (right).

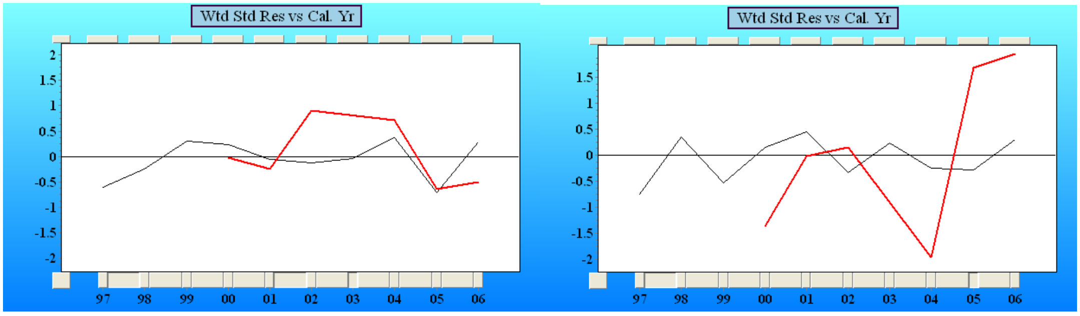
The pattern for Travellers seems closer to that of the Total Industry than Berkshire Hathaway, and MPTF modeling reports a significant correlation between CMP and WC of 0.34. Interestingly if we compare Travellers to Total Industry in these two lines we find, no correlation in WC, but a possibly significant correlation of 0.28 (p-value = 0.04) in respect of CMP.



Calendar trend, after modeling (black) and residuals for accident year 2000 (red). The Hartford, CMP (left) and WC (right).

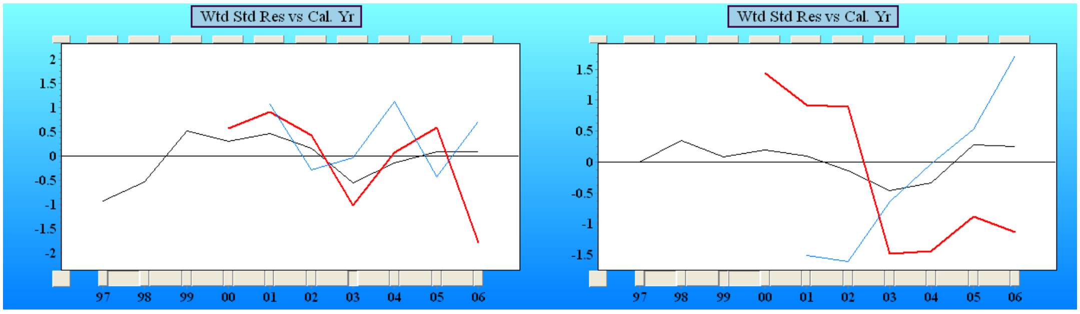
The same two lines for The Hartford show no significant correlation (in fact MPTF modeling detects a small non-significant negative correlation), a lack of relation that is reflected in the trace for the 2000 residuals.

Before leaving these LOBs let us look at two more companies, Zurich Financial Services Group, and State Farm, with 4.4Bn and 1.6Bn respectively in reserves.



Calendar trend, after modeling (black) and residuals for accident year 2000 (red). Zurich FSG, CMP (left) and WC (right).

Zurich registers a process correlation of 0.34 for these two lines. Interestingly Zurich and Travellers are significantly correlated at 0.3 and 0.35 for CMP and WC, and indications of this can be seen by comparing the accident year 2000 traces.



Calendar trend, after modeling (black) and residuals for accident year 2000 (red) and 2001 (blue). State Farm, CMP (left) and WC (right).

State Farm registers no significant process correlation for these lines. In the display above we show residual traces for two accident years, 2000 (red) and 2001 (blue).
On the basis of accident year 2000 one might be led to suspect a positive correlation, but no trace of this is visible in the 2001 residuals.

Conclusions.

The Schedule P Data runs to hundreds of megabytes giving detailed information on thousands of US insurance companies. As far as any individual company goes it represents only a small slice of the available data, limited by its 10-year reporting horizon and a number of other factors. Nonetheless it is sufficient to allow us to make our main point, which is that *it is possible* to model long-tail lines of business at an enterprise level and meet all of the desiderata that were listed at the beginning of this paper, which are briefly:

- Modelling incrementals and cashflows directly
- Producing models that tell a story in a transparent way in the language of business decision makers
- That modelling be accompanied by diagnostics that enable users to immediately assess how well they fit the data
- That modelling is consistent so that it is fair to compare results when the same general procedures are applied to different businesses or parts of businesses
- That models are synchronous with the formation and application of future economic scenarios
- That models include the capacity for measuring correlations between lines of business in a meaningful way.

Given the existence of modelling frameworks such as PTF/MPTF which meets these requirements it naturally follows that the use of *ad hoc* correlations or industry-wide development factors represent less than optimum uses of available information and are likely to lead those who follow them into a higher than necessary rate of extinction.

Appendix

Example of Modelling in the PTF Modelling framework

Outline

In the PTF modelling framework an optimal model is identified, equivalently, built or designed that captures the variability (volatility) in the incremental loss development array. The variability is described using four components of interest. Namely, trends in the three directions: **development period**, **accident period** and **calendar period**, and **the variability** of the data about the trend structure. The (process) variability is an integral part of the model.

A PTF model is succinctly described by four graphs; three graphs describe the trend structure in the three directions, the fourth graph depicts the process variability. The identified model is tested to ensure that the model assumptions are consistent with the data; including validation testing (by removal of years). The triangle is regarded as a sample path from the fitted model. Thus data simulated from the model should not be distinguishable from the original data in respect of trend structure and volatility about the trend structure.

An identified PTF model forecasts distributions in every cell in the future, conditional on a set of explicit and easily interpretable assumptions that can be directly related to the historical experience. Immediate benefits are easy calculation of Value-at-Risk and percentile (quantile) tables that can be used to calculate prediction intervals (sometimes erroneously called ranges), risk margins, and to match liabilities with assets.

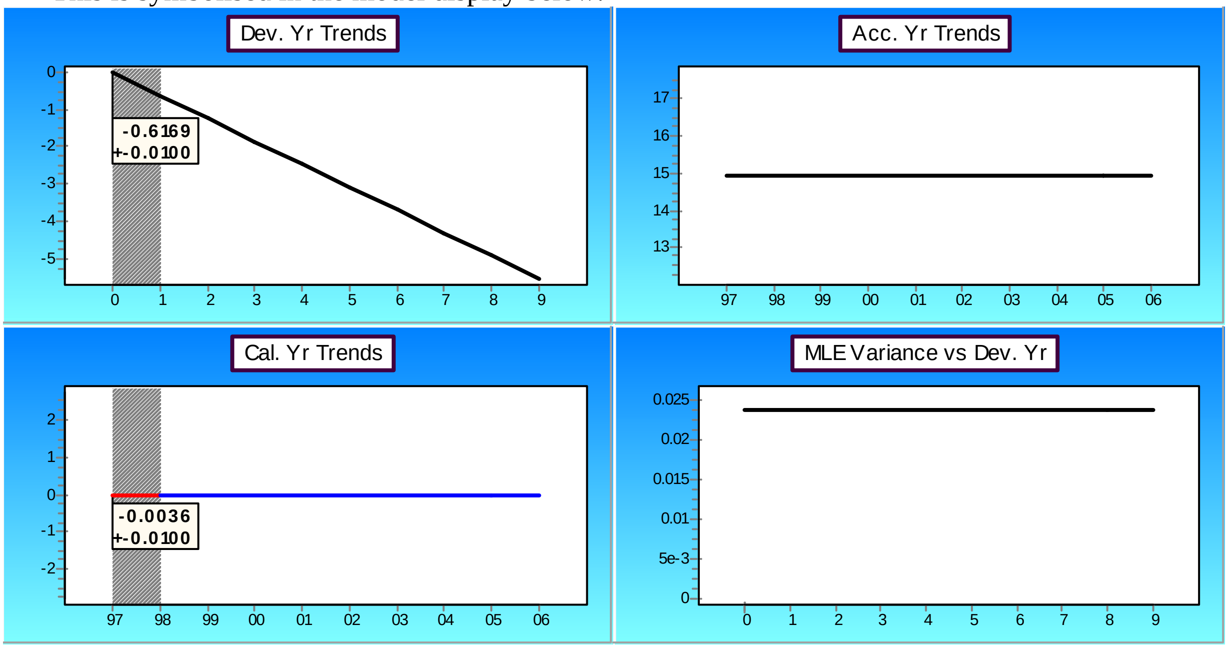
Modelling example: AI PPA

The data from AI PPA was used as an example for the modelling process.

1) Open PTF:

   The first model fitted in PTF is a single trend in the development and calendar directions and an average level in the accident direction. That is, an average "trend" in all three directions.

   This is symbolised in the model display below:



These four pictures describe the model as currently fitted to the data. We have:

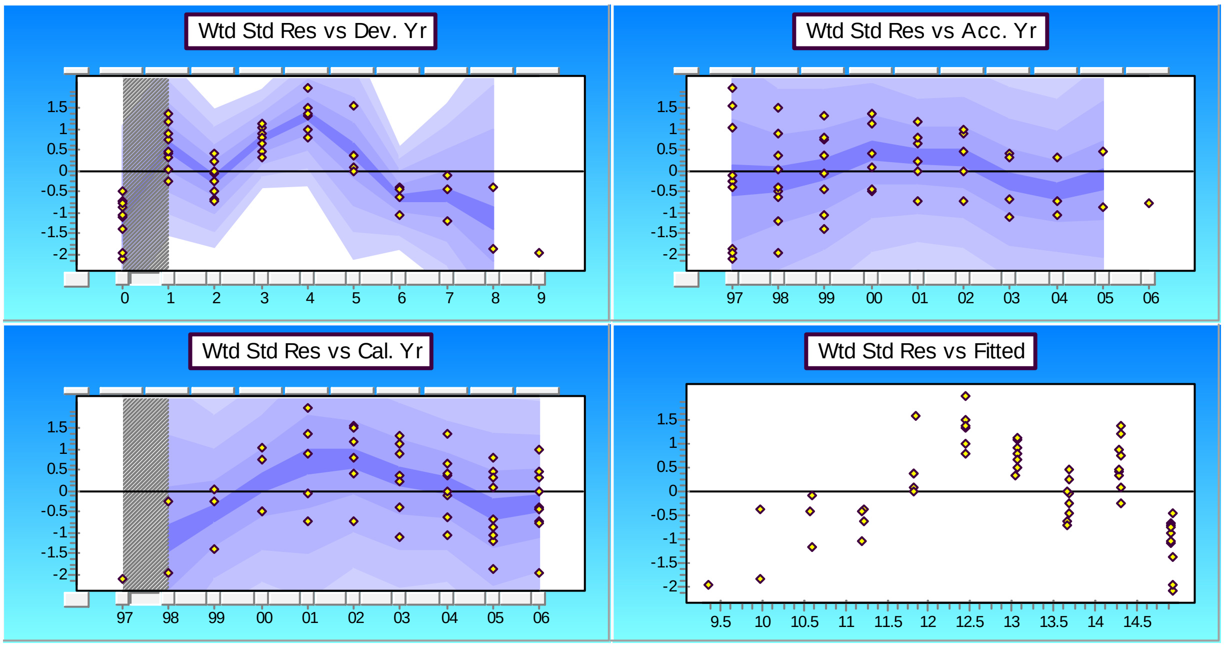
- An average decay in the development direction (upper left) of 61.7%+-1% per year.

- An average calendar trend of -0.3% +- 1% per year (lower left). Note that the average calendar trend is not statistically distinguishable from zero.

- The accident year levels (upper right) show a flat line – the average level was fitted (14.9+-0.06) on a log scale.

- The process variance is the fourth graph (lower right) shows the average variation around the development trend line.

While these four graphs describe the model fitted to the data, they do not provide any description of whether the structure in the data has been fully described in the model. In order to evaluate the model, we need to turn to another set of four pictures which display the difference between the model predictions (ie the trends predicted by the method) and what we actually see in the data. We call the differences "residuals".



The same three descriptions apply for the development, accident, and calendar year graphs. However, the process variance graph has been replaced by a graph of the differences between the data and what the model predicts – irrespective of period.

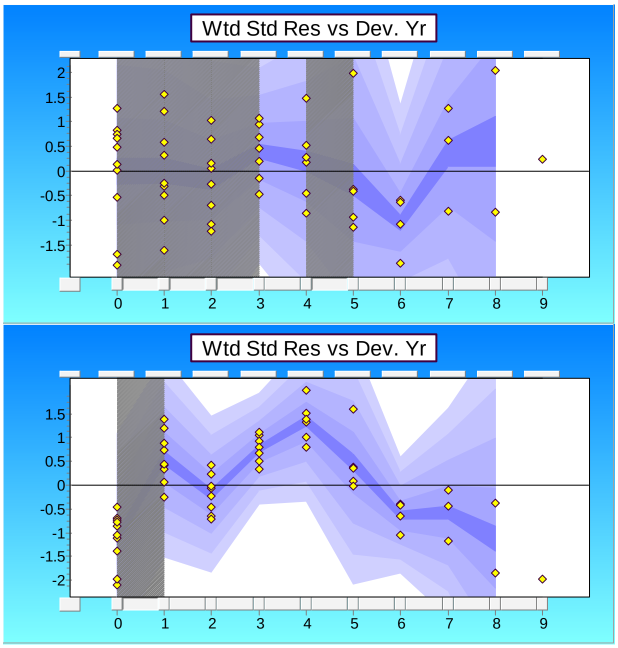
If the model fully describes the data, there should be no discernable structure (patterns) in any of these four residual graphs (bearing in mind that the human eye tends to accentuate patterns).

In this case, it is obvious that structure in the data still exists. The residuals in the development direction (for example), increase and decrease as a group around our average trend line from period to period. In the calendar direction, there is a definite increase in trend followed by a decay (relative to the trend line).

In order to capture the structure in the data, we need to add a model parameter wherever we see a ‘major’ change in direction in the trend line (central purple band). From the PTF modelling perspective, it doesn’t matter if you add ‘too many trends’ in that the optimisation routine will remove any excess parameters.

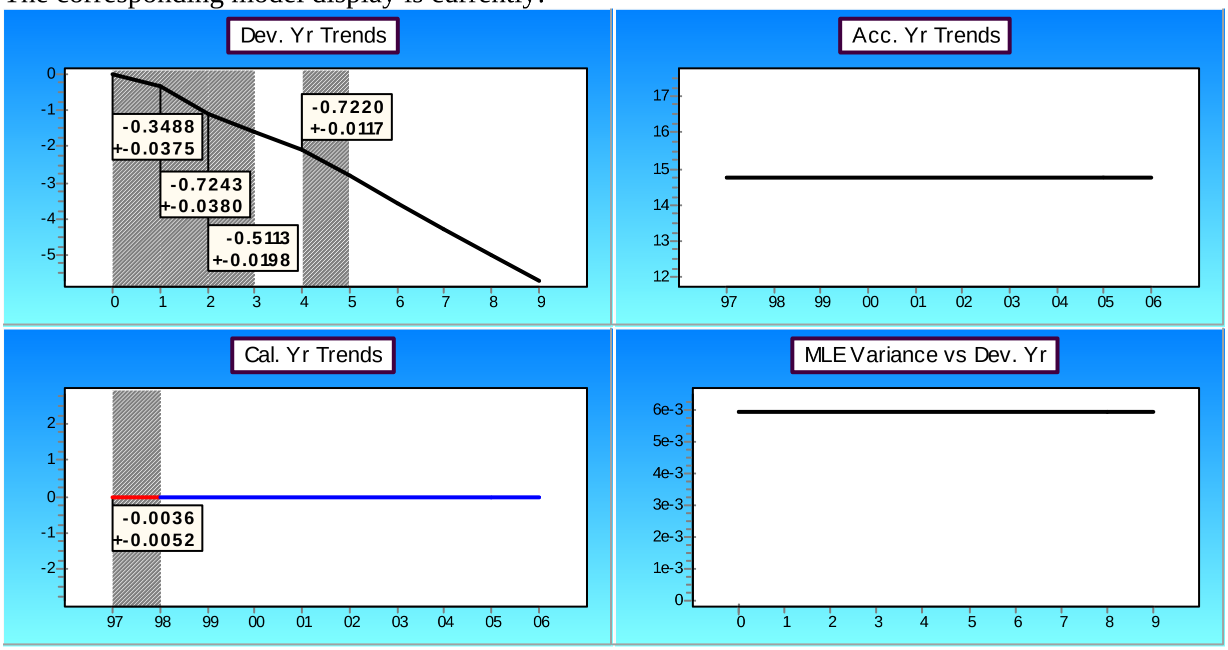
We then add parameters to the model based on where we see changes in direction of the residuals relative to the fitted (single) trend (or level for accident periods). I focus on development trend changes first, then add either accident level changes or calendar trend changes depending on which direction has the greatest change. Furthermore, as we add trend changes, it can become obvious that we need further parameters; this is a natural part of the modelling process.

For example, if we place development period parameters where we saw changes in the development trend direction (relative to the average trend), we end up with (left hand side). The original graph has been reproduced on the right hand side for comparison.



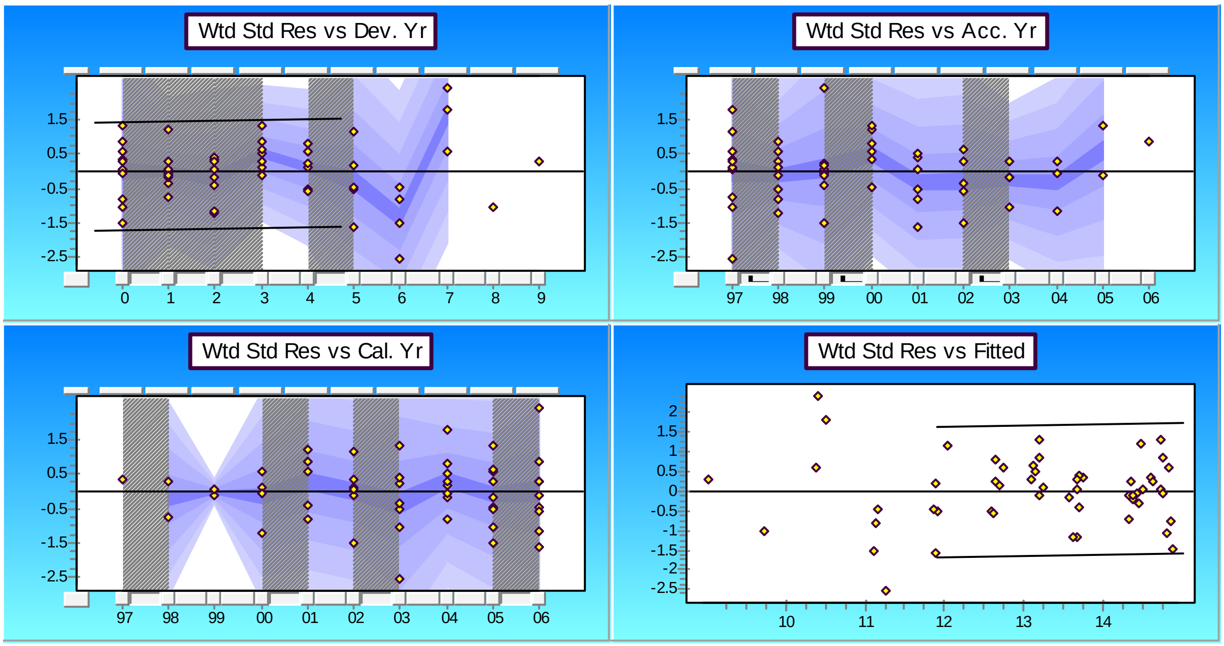
The newly added trend parameters (left hand side) show residuals that are much more randomly scattered about their respective trend lines than the display for the single trend line (right hand side).

The corresponding model display is currently:



We then proceed with this same technique for the other two directions (accident and calendar) to arrive at a model which looks like:

With a corresponding residual display of:

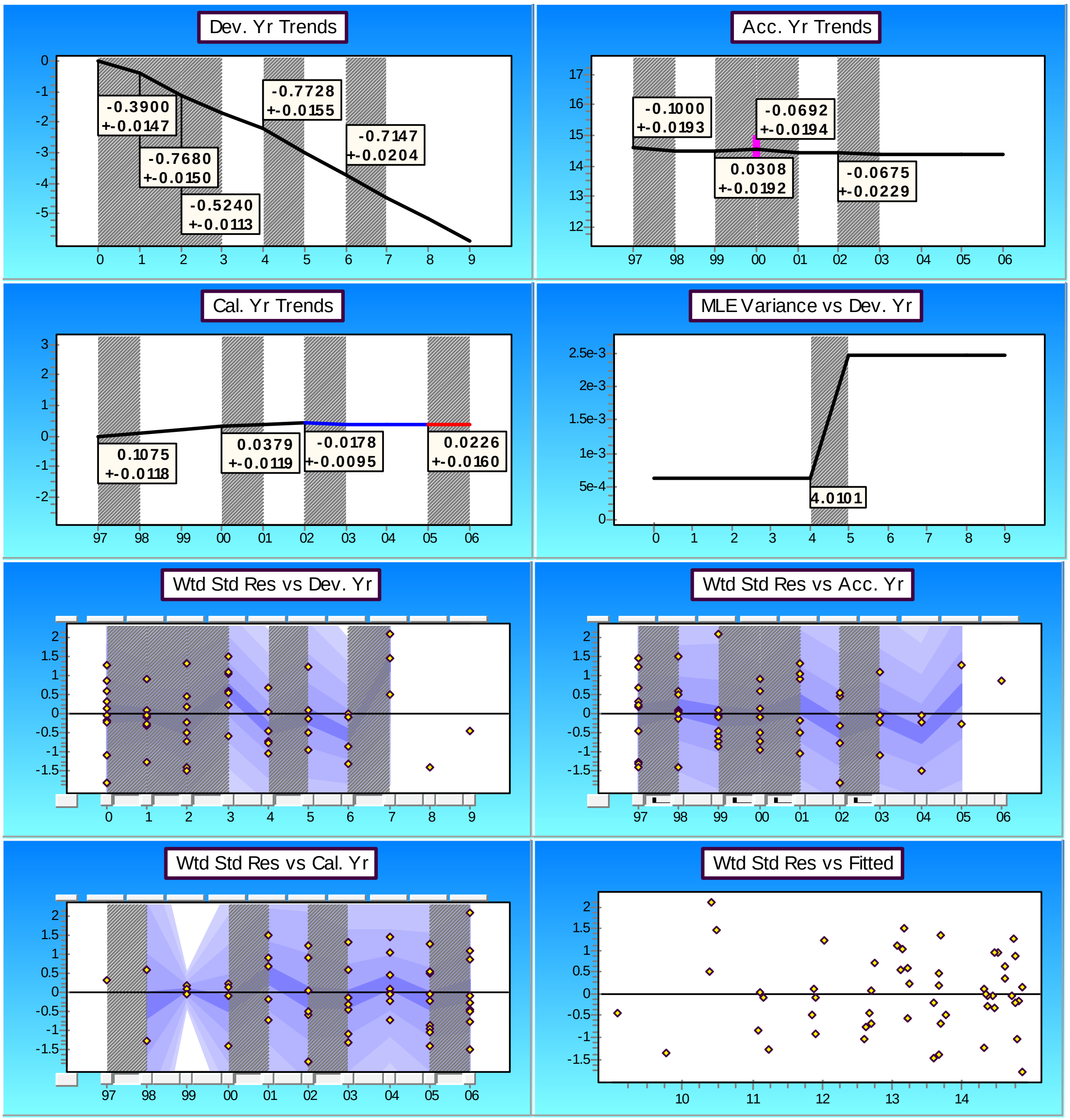


Note at this point we have not made any changes to the process variation. We have *only* concerned ourselves with the changes in the trends. From these four plots above, we would conclude there are no more trend changes (remember the last development periods are based on very few points so we expect more variation from the 'average trend').

The fourth attribute we need to model is then the process variance. We normally model this around the development period as changes in percentage variation usually occur in development periods (less claims contributing to the sum generally results in more variation of the sum). This is observed by a change in the width of the scatter in the residuals (shown by horizontal lines); in this example more clearly by the residuals versus fits where the lower payments (left hand side) show greater variability than higher payments (right hand side).

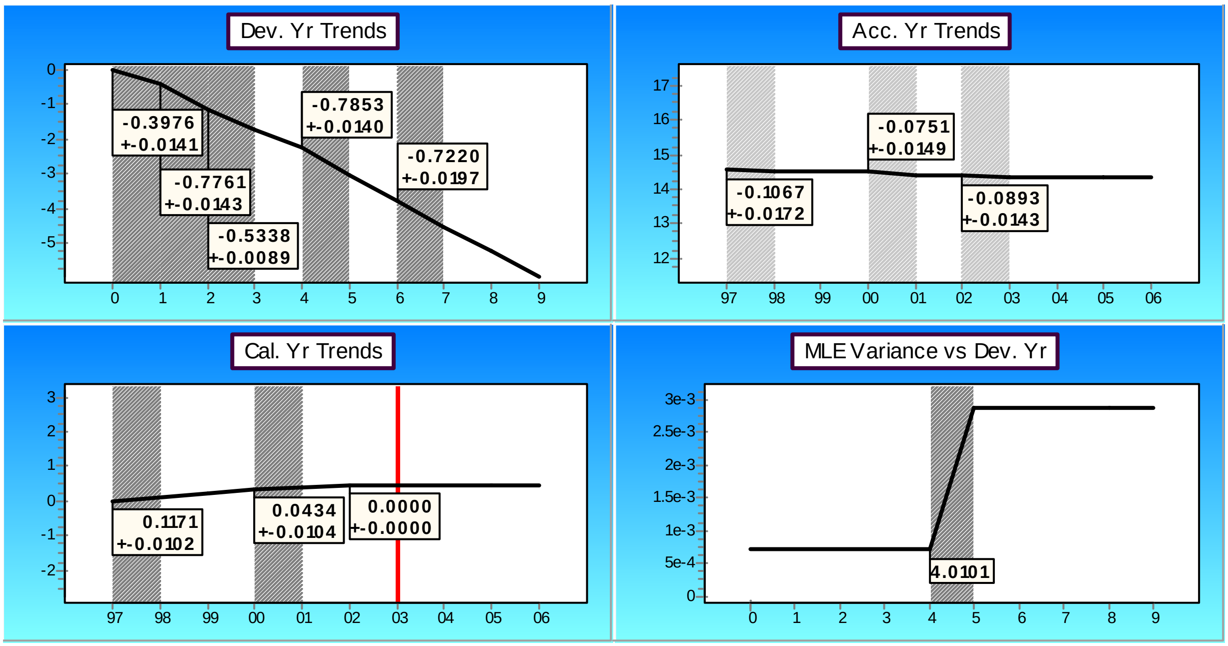
The model displays incorporating the change in process variation are shown below. The process variance analysis has determined an increase of percentage variation of 400% between development periods 4 and 5. Although this sounds a major increase, the values we are measuring are very small and even these final periods do not have large amounts of process variation.

Note that this model is not optimal – parameters were added where it was thought that they were needed, but no formal testing has been done on them.

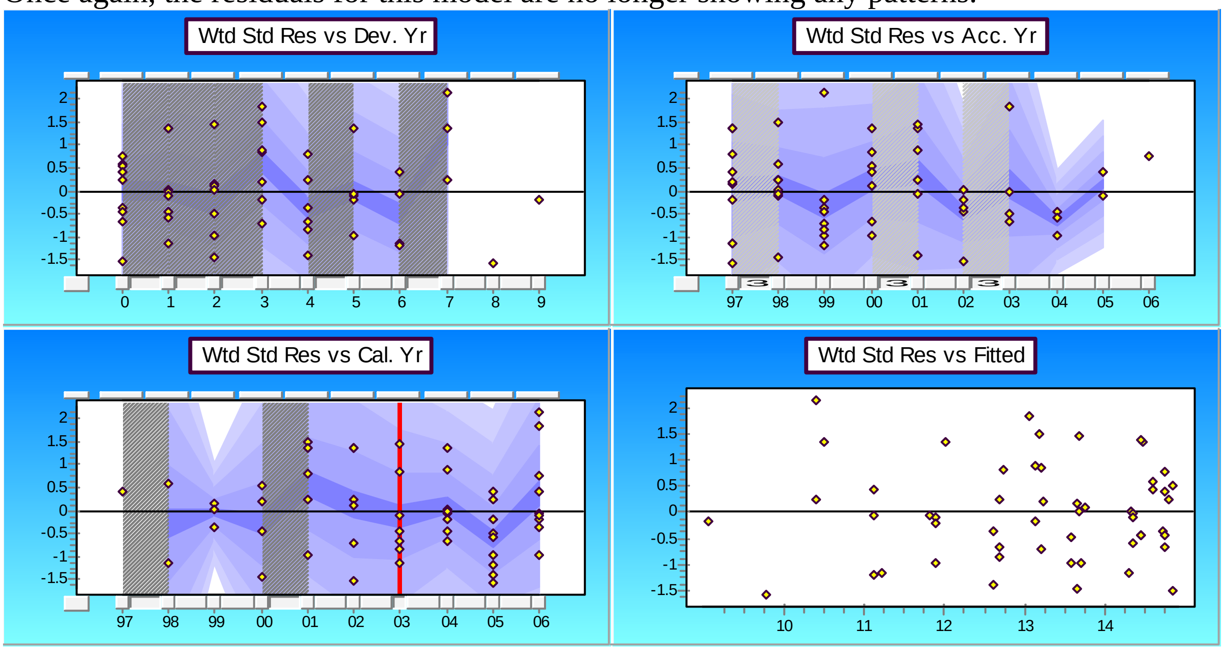


Note that the residuals versus fits are now flat. Similarly, the residuals versus development period are more evenly scattered around the trend lines. Two more parameters have been added to better capture the development and accident period structure.

We then optimise the model by conducting a series of statistical tests on the model parameters to ensure we only retain statistically significant changes. The optimisation routine conducted on this model resulted in the following “final” model.



Once again, the residuals for this model are no longer showing any patterns.



However, the final calendar year may indicate a trend change however we do not have enough data to estimate this trend sufficiently.

Forecast scenarios where a positive trend occurs will need to be conducted and the sensitivity of the reserve distribution to the return of the trend evaluated. Business knowledge would also help determine the likely cause of the calendar trends and therefore give credence to particular forecast scenarios. The importing thing to note is that at all times, modelling and forecast scenarios are in the hands of the user, are fully quantifiable, and controllable. The PTF modelling framework is fully interpretable and does not have any 'black box' properties.