A practical approach to claims reserving using state space models with growth curves

Chibisi Chima-Okereke

Active Analytics Ltd

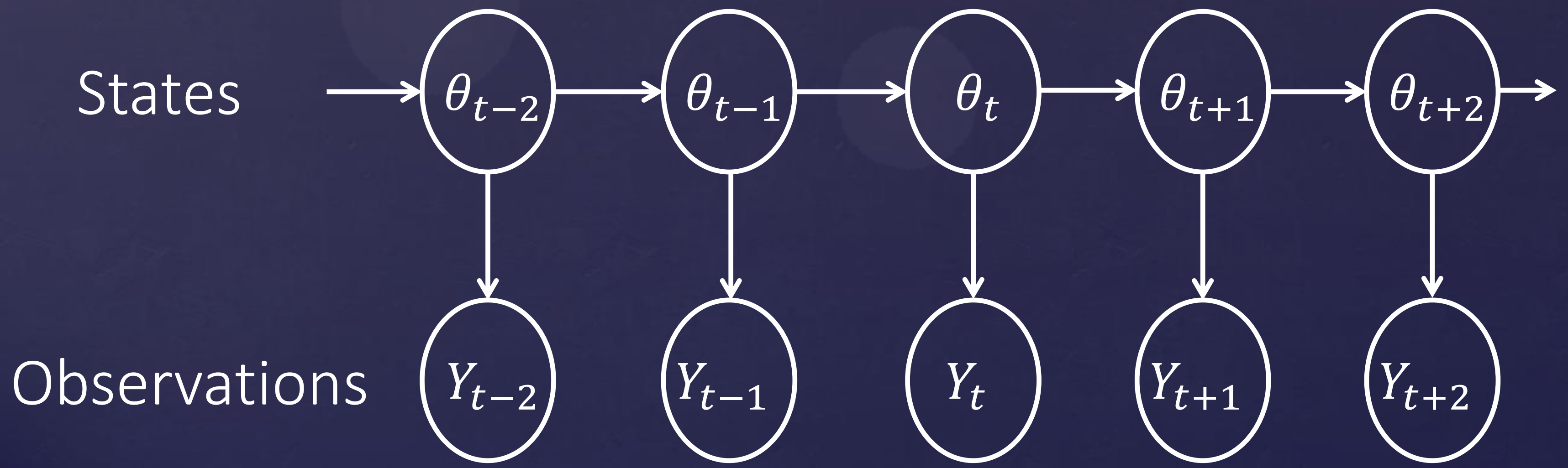
www.activeanalytics.co.uk

Program

- What are state space models
- Why we are interested in them for reserving
- Their general representation
- The chosen approach
- Univariate linear model
  - Analysis
- Multivariate linear model
  - Analysis
- Particle filters
  - Multivariate analysis
- Summary

What are state space models?

- A representation of a dynamic system
- The states (θ_t) of the system are not directly observable
- These states "drive" the observable set of values Y_t
- The state space model has conditional independence structure



Why State Space Models?

Advantages

- Can use subjective expert judgement and data from any relevant source to drive model outputs
- Allows models with meaningful dynamic parameters to be created
- States and forecasts have probabilistic representation making them useful of quantifying uncertainty for reserves
- Provides a formal framework for intervention
- Can be used as a framework for automating the reserving process

Why State Space Models?

Disadvantages

- Expert skills need to be acquired or developed to use them
- Relatively unknown in actuarial analysis so may take time to gain acceptance
- They take along time to develop and can be expensive to implement
- They can be very complex so easy to get wrong

General representation

- Observation Equation

$\{Y_t|\theta_t\}$ $\quad Y_t = F_t(\theta_t, v_t)$ $\quad v_t \sim H_v$

- System or state Equation

$\{\theta_t|\theta_{t-1}\}$ $\quad \theta_t = G_t(\theta_t, w_t)$ $\quad w_t \sim H_w$

F_t: Design matrix/function
G_t: System matrix/function
v_t: Observation errors with distribution
w_t: Evolution errors with distribution

H_v and H_w are not necessarily normal
F_t and G_t are not necessarily linear
v_t and w_t are mutually independent

Approach

- Focus on filtering and forecasting
- Framework for multivariate model from de Jong & Zehnwirth's approach to system and design matrices
- A growth curves approach to reserving e.g. Dave Clark & James Guszcza
- Sequential Importance Resampling particle filter for nonlinear functions
- In general we assume that the covariance matrices V_t and W_t are constant with time

Univariate model

- Cumulative paid claims for origin year j and development period t be given by $P_{j,t}$
- The univariate model focuses on claims development for a particular origin year
- The log-transformed paid claims is our observation

$$Y_t = \log_e(P_t)$$

- The observation and system equations

$$\{Y_t | \theta_t\} \qquad Y_t = F_t \theta_t + v_t \qquad v_t \sim N(0,\ V_t)$$

$$\{\theta_t | \theta_{t-1}\} \qquad \theta_t = G_t \theta_{t-1} + w_t \qquad w_t \sim N(0,\ W_t)$$

$$F_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad G_t = \begin{bmatrix} 1 & \lambda \\ 0 & \lambda \end{bmatrix} \quad \theta_t = [\theta_{t_1}, \theta_{t_1}] \quad w_t = [w_{t_1}, w_{t_1}]$$

Univariate model

- The Gompertz, Gumbel, and Logistic curves have parameters that relate to λ
- Dave Clark & James Guszcza suggest some other curves that can be used
- For instance in the Gompertz function

$$P_t = \alpha e^{-\beta e^{-\gamma t}} \ (\alpha, \beta, \gamma > 0)$$

$$\lambda = e^{-\gamma}$$

Mitscherlich for claims increment

- Consider the Mitscherlich as a "log Gompertz" type function

Mitscherlich : $E(\log(P_t)|\,\theta_t) = \alpha - \beta\lambda^t$

Observation: $Y_t = \log(P_t) - \log(P_{t-1})$

Evolution: $E(\theta_t|\theta_{t-1}) = \lambda\theta_t$

Observation: $Y_t = \theta_t + v_t$

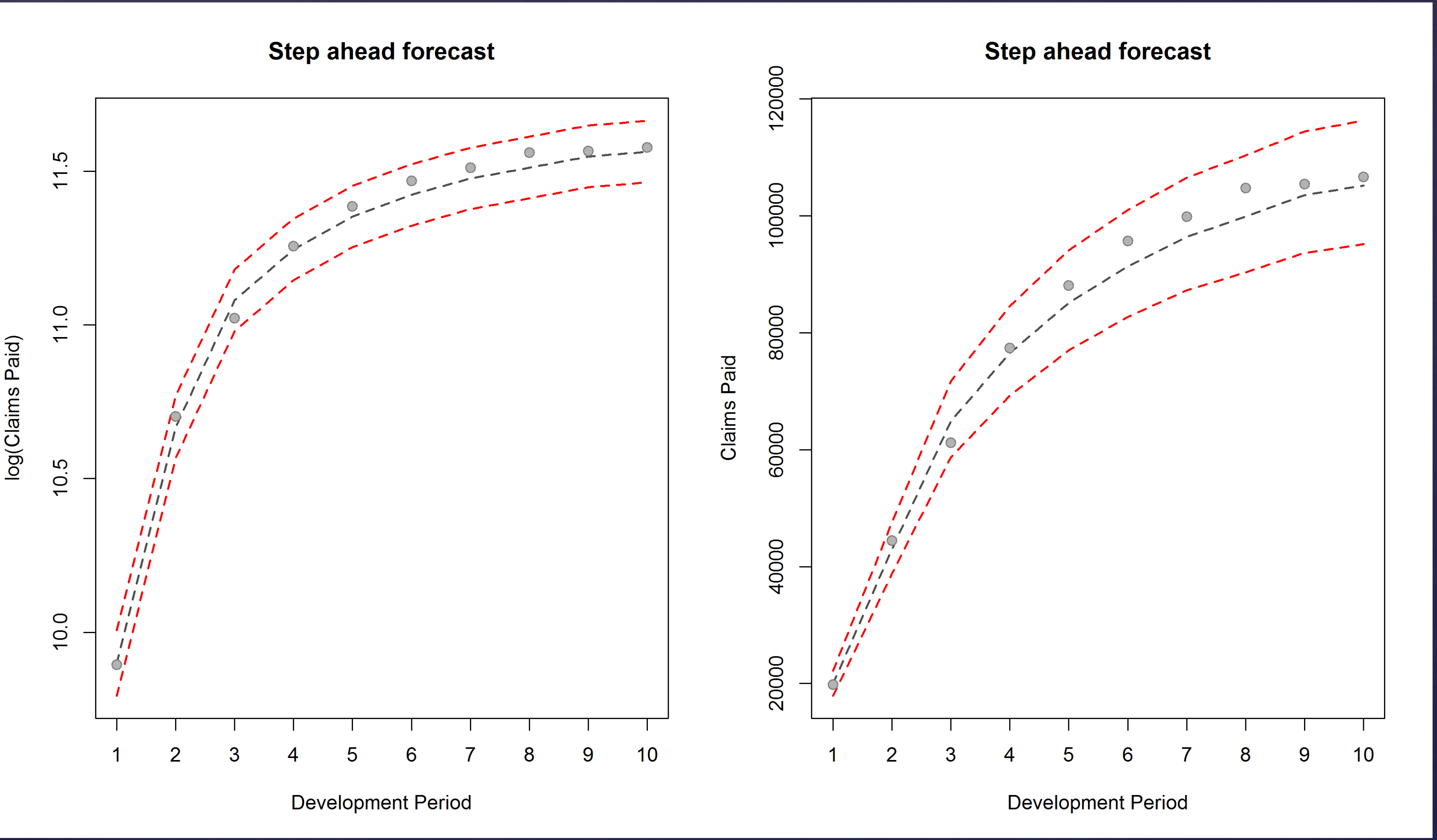
System: $\theta_t = \lambda\theta_{t-1} + w_t$

- We will stick to the original formulation

Estimation of filter parameters

- Parameter λ may be estimated from nonlinear regression and is a constant in the dynamic model
- The prior distribution of θ is $(\theta_0|D_0) \sim N(m_0, C_0)$ where D_t is the data available at time t.
- However m_0, C_0, λ, V, **W** can all be obtained by using maximum likelihood methods. This is what we do in this presentation.
- Of course they can be adjusted or created using expert judgement. V and W don't need to be constants (adaptive).

Univariate model



Multivariate model

- Now the multivariate model for the claims triangle

Development period (d) $\longrightarrow$

Origin year (j) $\downarrow$

$$
\begin{array}{cccccccc}
P_{1,1} & P_{1,2} & P_{1,3} & \cdots & \cdots & \cdots & P_{1,T-1} & P_{1,T} \\
P_{2,1} & P_{2,2} & \cdots & \cdots & \cdots & \cdots & P_{2,T-1} & \\
P_{3,1} & P_{3,1} & \cdots & \cdots & \cdots & P_{3,T-2} & & \\
\vdots & \vdots & & & \vdots & & & \\
\vdots & \vdots & \vdots & & & & & \\
P_{J-2,1} & P_{J-2,2} & P_{J-2,3} & & & & & \\
P_{J-1,1} & P_{J-1,2} & & & & & & \\
P_{J,1} & & & & & & &
\end{array}
$$

Multivariate model

- Data is a successively expanding vector of diagonals
- Y_t is the vector of log cumulative claims at time t containing $y_{j,d}$, $t = j + d$

$$Y_1 = [y_{1,1}], \quad Y_2 = \begin{bmatrix} y_{1,2} \\ y_{2,1} \end{bmatrix}, \quad Y_3 = \begin{bmatrix} y_{1,3} \\ y_{2,2} \\ y_{3,1} \end{bmatrix}, \quad \cdots, \quad Y_t = \begin{bmatrix} y_{1,t-1} \\ y_{2,t-2} \\ \vdots \\ y_{J-1,2} \\ y_{J,1} \end{bmatrix}$$

- Design and system matrices
  F_t & G_t are now block forms

(de Jong & Zehnwirth)

Alternative state matrix forms

- Off-diagonal blocks give the opportunity to take previous states into account

$$\left[\begin{array}{cc:cc} p & p\lambda & (1-p) & 0 \\ 0 & p\lambda & 0 & (1-p) \end{array}\right]$$

- Where $0 \leq p \leq 1$

- We can also alter λ to λ_t so that G_t is no longer constant with time

$$\lambda_t = \lambda_0 + \delta\left(1 - (d+1)e^{-2d}\right)$$

- The form is similar to the basis function given by de Jong & Zehnwirth

Multivariate model

- The data is adjusted for inflation having 10 development periods
- This means that data is "complete" over 5 development periods and origin years
- Fit multivariate dynamic linear model and chain ladder model to the 5 by 5 triangle
- The $\lambda_t = \lambda_0 + \delta\left(1 - (d+1)e^{-2d}\right)$ form was used
- Compare residual sums of squares

Model outputs

Actual (Inflation adjusted)

	1	2	3	4	5
1	19827.00	44449.00	61205.00	77398.00	88079.00
2	20398.16	44283.85	62835.02	84362.19	95873.43
3	18801.15	37116.70	54811.46	73788.66	85143.78
4	17627.32	39120.33	62148.34	74740.05	86238.05
5	17441.77	39836.28	58902.97	73055.92	81916.40

DLM $Log(RSS) = 19.41$

	1	2	3	4	5
1	19827.00	44449.00	61205.00	77398.00	88079.00
2	20398.16	44283.85	62835.02	84362.19	98308.83
3	18801.15	37116.70	54811.46	70202.57	81688.29
4	17627.32	39120.33	58582.04	75282.23	87904.65
5	17441.77	36235.67	54701.20	70295.07	82081.30

ChainLadder $Log(RSS) = 20.78$

	1	2	3	4	5
1	19827.00	44449.00	61205.00	77398.00	88079.00
2	20398.16	44283.85	62835.02	84362.19	96004.26
3	18801.15	37116.70	54811.46	71479.45	81343.68
4	17627.32	39120.33	55595.98	72502.55	82507.97
5	17441.77	37537.26	53346.20	69568.61	79169.14

Disadvantages

- We have static variables λ and δ that need to be suitably obtained
- Linear space state models limit us to normal error assumptions and linear system and observation equations
- Linear state space models constrain the choice of functions we can use to represent the claims development curve

Particle filters

Particle filters allow a more flexible modelling structure including

- Allows nonlinear design (F_t) and system (G_t) relationships
- Allows non-normal v_t and w_t
- Working directly curve parameters as states gives us interesting options for the state evolution matrix (G_t)
- Gives a good representation of the updated system "state" with time

- The price is that simulation is now necessary - which can take much longer depending on the number of particles
- Here some basic sequential importance sampling examples are presented

Sequential Importance Resampling

Procedure

- Sample $\theta_{t_0}^{(1)}, \dots, \theta_{t_0}^{(N)}$ from $p_0(\theta)$ prior distribution
- At time t – 1 we have particles $\theta_{t-1}^{(1)}, \dots, \theta_{t-1}^{(N)}$
- Use the evolution equation to generate a new set of particles $\tilde{\theta}_t^{(1)}, \dots, \tilde{\theta}_t^{(N)}$ by computing $G_t(\tilde{\theta}_t | \theta_{t-1}^{(i)},\ W_t^{(i)})$
- Then compute the weights from the obs. density function

$$\omega_t^{(i)} \propto \frac{p(Y_t | \tilde{\theta}_t^{(i)}, y_t)}{\sum_i p(Y_t | \tilde{\theta}_t^{(i)}, y_t)}$$

- Now resample $\theta_t^{(i)}$ from the pairs $\{\tilde{\theta}_t^{(i)}, \omega_t^{(i)}\} \sim p(\theta_t | D_t)$

Analysis

- Two nonlinear forms are considered
  - The Gompertz function

$$E(Y_t|\theta_t) = \theta_{t_1} e^{-\theta_{t_2}{}^{\theta_{t_3} t}}$$

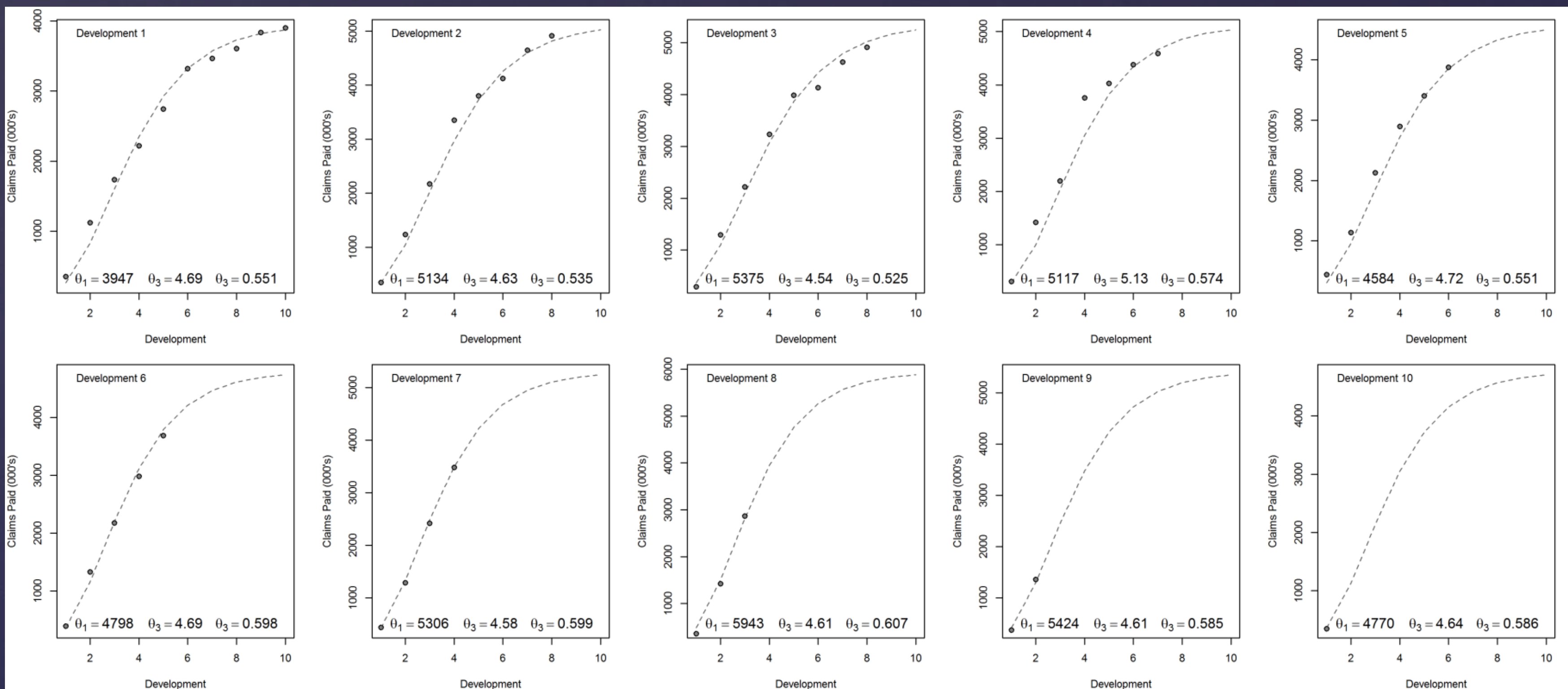
  - The Weibull function

$$E(Y_t|\theta_t) = \theta_{t_1} \left(1 - e^{-\left(\frac{t}{\theta_{t_2}}\right)^{\theta_{t_3}}}\right)$$

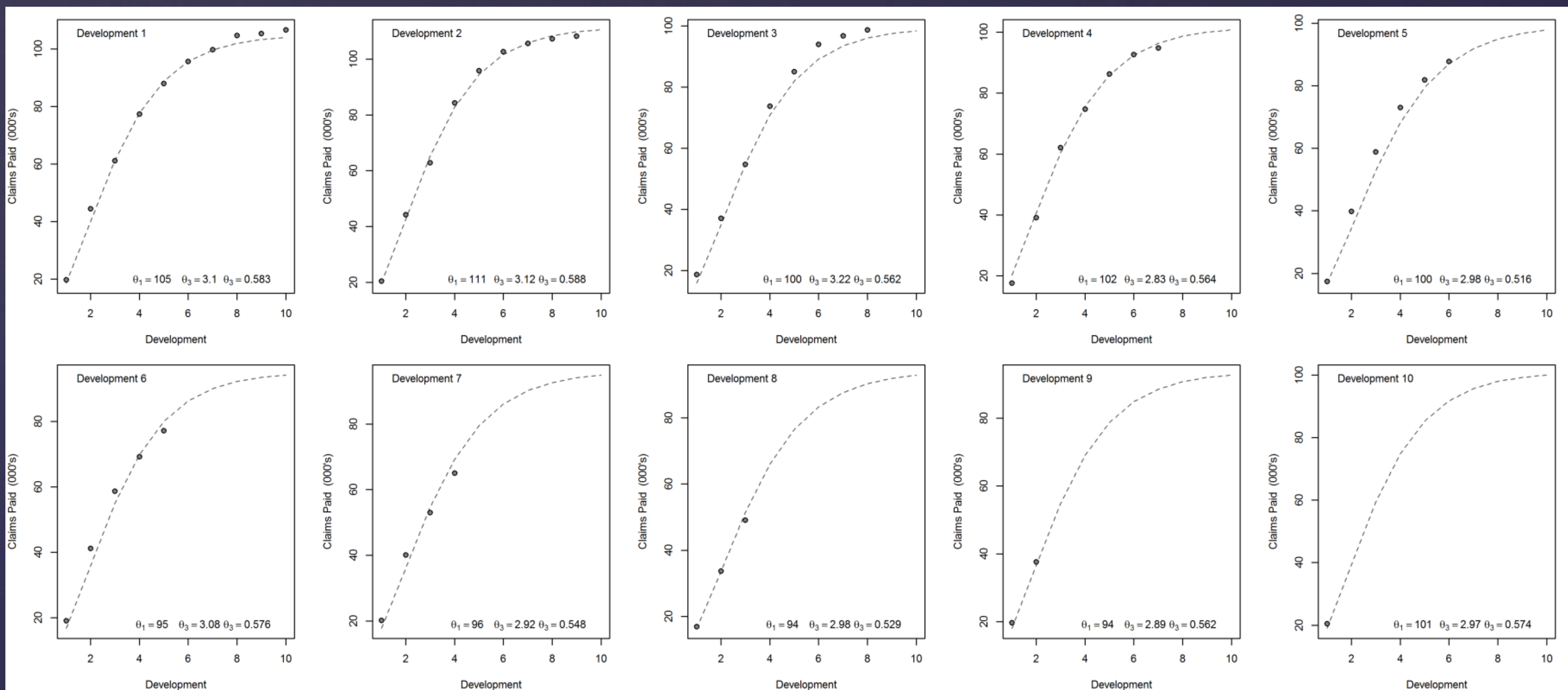
- θ_{t_1} is the ultimate loss and now exists as a state
- Claims triangles data from Dave Clark and Auto data from the ChainLadder package
- The components θ_t, v_t and w_t are normally distributed

$$v_t \sim N(0, V_t); \quad w_t \sim N(0, W_t); \quad \theta_t \sim N(m_t, C_t)$$

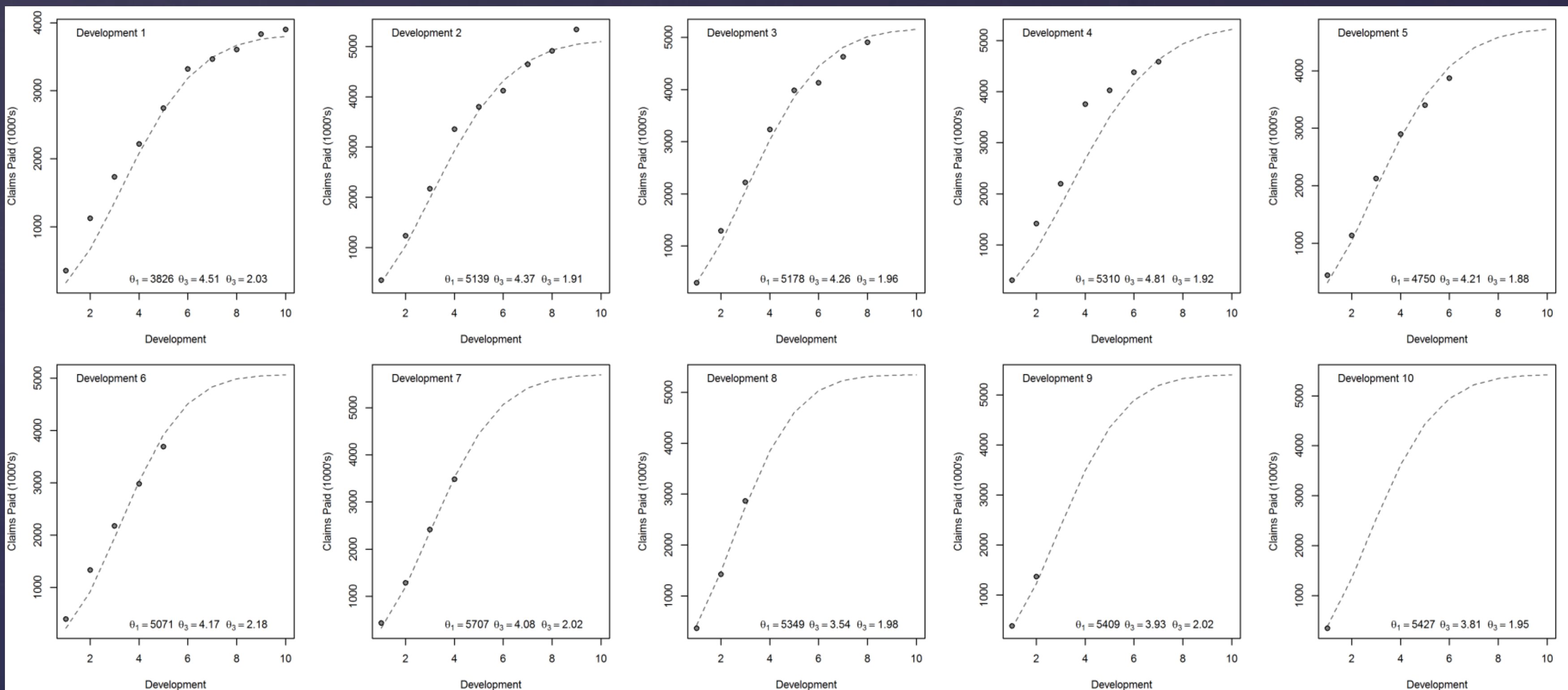
Outputs: Gompertz



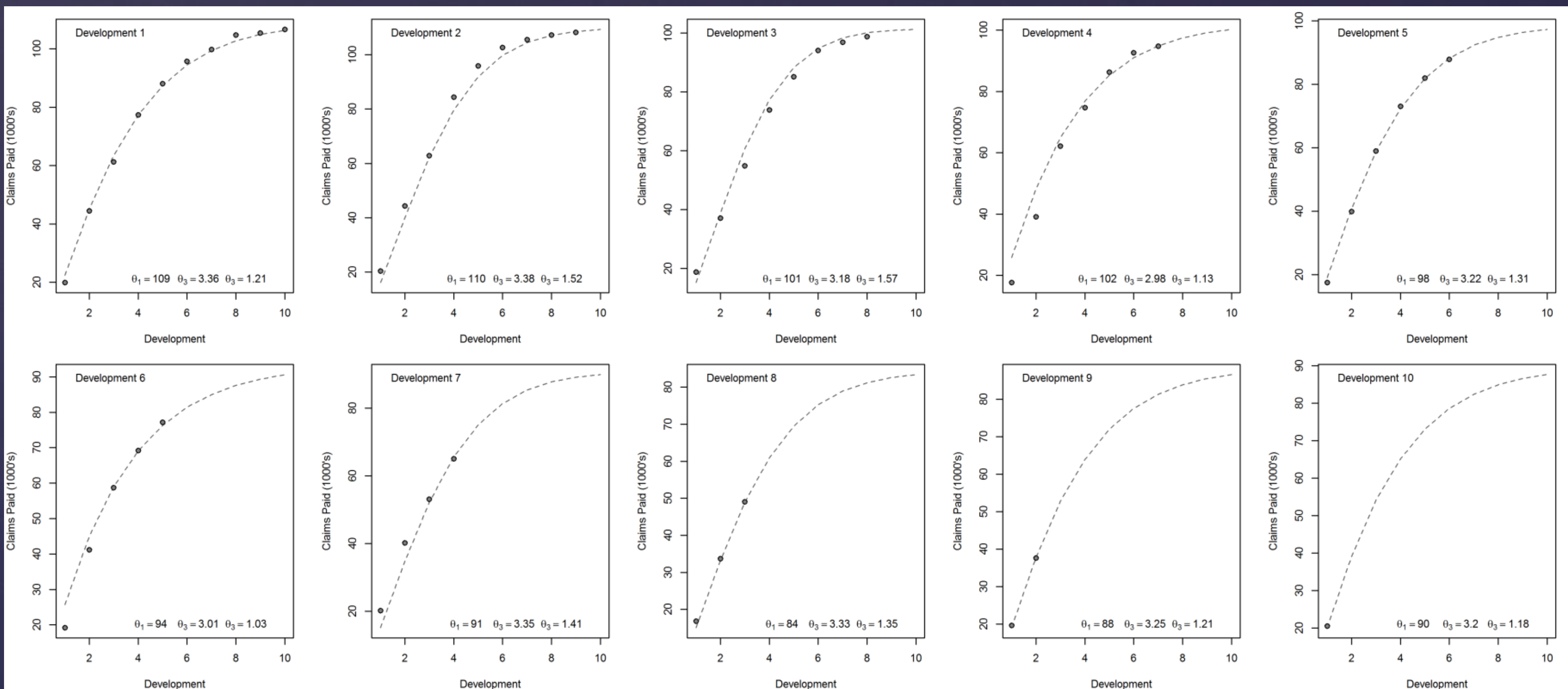
Outputs: Gompertz (Auto)



Outputs: Weibull



Outputs: Weibull (Auto)



Summary

- More work to be done to hone the model, perhaps a none parametric technique are more appropriate
- State space models offer an interesting and varied tool set
- They offer a formal framework that can be used for intervening in the forecasting process
- They can be complex, difficult to implement and take a long time to develop
- It can be a challenge to obtain an appropriate parametric curve and parameters for the state space model