

28 November – 1 December 2004

# Xth Accident Compensation Seminar

2004



Institute of Actuaries of Australia

 **ERNST & YOUNG**

 **KPMG**

**PRICEWATERHOUSECOOPERS** 

  
**IAG**  
Insurance  
Australia  
Group

28 November – 1 December 2004

# Xth Accident Compensation Seminar

2004



## Soft-Computing in Accident Compensation

**Peter Mulquiney**  
**Taylor Fry Consulting Actuaries**



Institute of Actuaries of Australia



# Outline

- Features of accident compensation data
- Overview of soft-computing methods
- Case study: CTP reserving



# Accident Compensation Data

- Features of accident compensation data that make pricing and reserving difficult
  - Changes in the rate of claim finalisation
  - Legislative changes
  - Seasonality
  - Superimposed inflation varying by
    - Experience year
    - Age of claims
- Can be dealt with using GLMs



## Example Data

- Data relates to CTP insurance from one state
- Claim file of 60,000 claims from 1994 – 2003
- Features of data illustrated by a GLM model of average size of finalised claim
- Model was developed for pricing and reserving



# GLM Model

- Model of average size of finalised claims

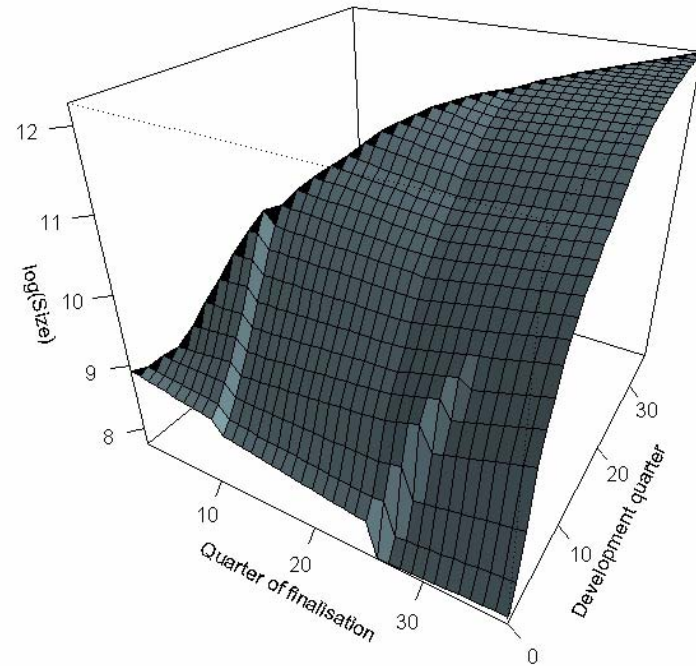
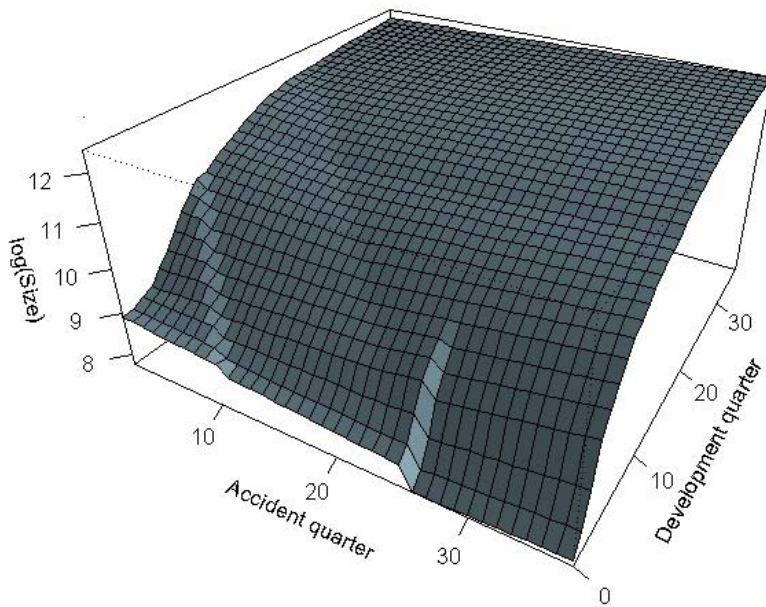
$$\begin{aligned}
 E[Y_r] = \exp \{ & \beta^d_1 t_r + \beta^d_2 \max(0, 10 - t_r) \\
 & + \beta^d_3 \max(0, t_r - 80) + \beta^d_4 I(t_r < 8) \quad [\textit{Operational time effect}] \\
 & + \beta^s I(k_r = \text{March quarter}) \quad [\textit{Seasonal effect}] \\
 & + \beta^f_1 k_r + \beta^f_2 \max(0, k_r - 2000Q3) \\
 & + \beta^f_3 I(k_r < 97Q1) \quad [\textit{Finalisation quarter effect}] \\
 & + k_r [\beta^{tf}_1 t_r + \beta^{tf}_2 \max(0, 10 - t_r)] \quad [\textit{Operational time x finalisation quarter interaction}] \\
 & + \max(0, 35 - t_r) [\beta^{ta}_1 + \beta^{ta}_2 I(i_r > 2000Q3)] \quad [\textit{Operational time x accident quarter interaction}]
 \end{aligned}$$





# GLM Model

- Plot of  $\log(\text{average claim size})$  without seasonality





# GLM vs Soft-Computing

- Building a GLM model takes time
- Specifying form of model can be difficult
  - Interaction terms can be troublesome
- Quality of model will depend on skill of model builder
- **Question: Are there better ways?**





# Soft-Computing

- Collection of methods
  - Designed specifically for large and/or complicated data sets
  - Largely automated
- Uses
  - Data mining
  - Modelling complex “non-linear” features
- Examples
  - Neural Networks, MART, MARS



## Case Study – CTP Data

- Aim: Compare models of the average size of finalised claims
  - GLMs
  - Neural Networks
  - MARS
  - MART



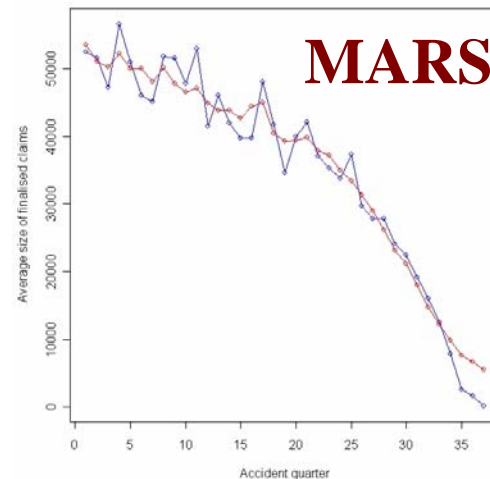
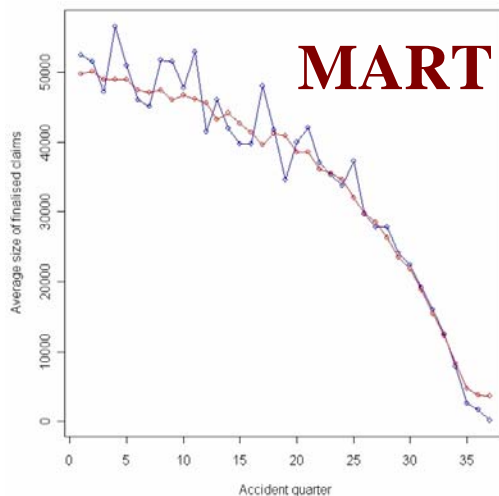
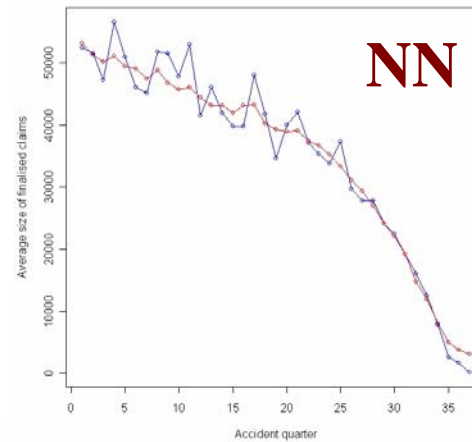
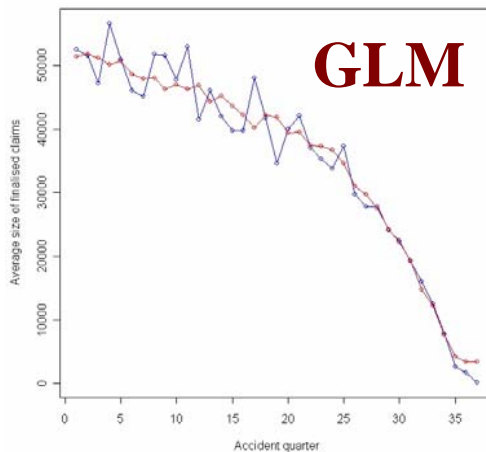
# Soft-Computing Background

- GLMs
  - Manually build model that matches the data
- Neural Networks
  - Model is so flexible that it can model almost anything
  - Various techniques used to protect against overfitting
- MART and MARS
  - Regression functions built automatically by progressively adding more terms
  - Each term constructed out of simple functions



# Results: one-way tables

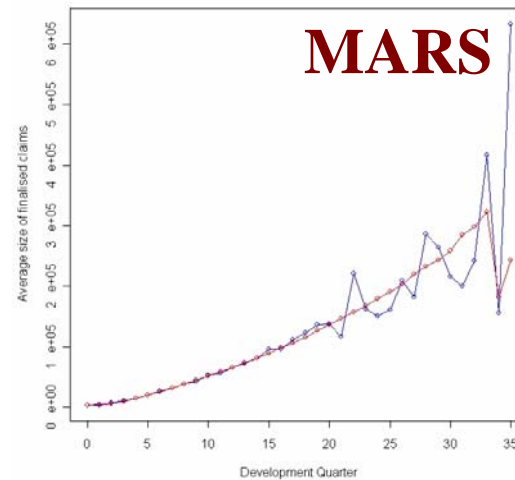
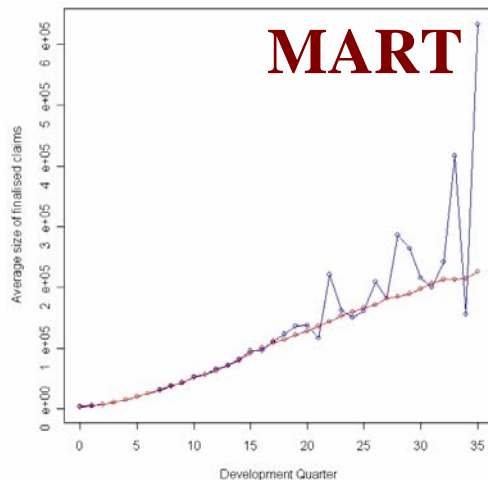
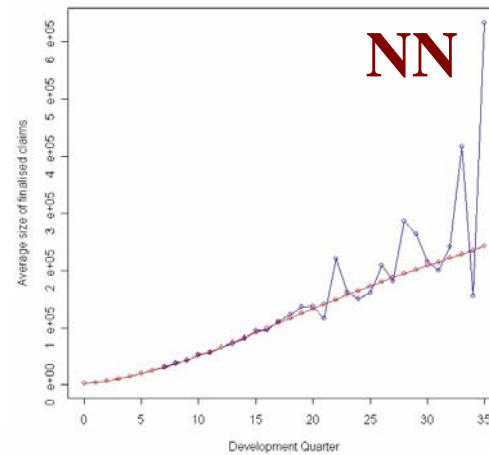
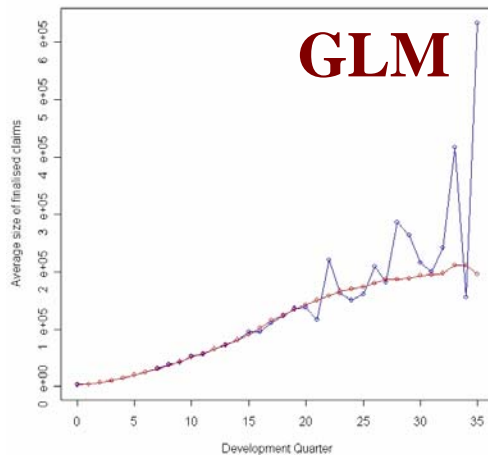
- Finalised claims against accident quarter





# Results: one-way tables

- Finalised claim against development quarter





# Triangles of observed to fitted ratios

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Sep-94	NA	31%	80%	154%	198%	141%	75%	80%	83%	73%	114%	100%	90%	69%	63%	101%	76%	45%	113%	231%	109%	154%	76%	81%	63%	51%	448%	5%	188%	154%	77%	NA	106%	NA	3%	324%
Dec-94	23%	106%	95%	100%	134%	129%	123%	97%	81%	124%	95%	86%	81%	139%	95%	91%	100%	92%	113%	90%	56%	99%	58%	94%	105%	202%	46%	56%	47%	81%	101%	145%	280%	298%	78%	NA
Mar-95	23%	75%	96%	105%	99%	112%	89%	84%	98%	120%	158%	94%	84%	61%	65%	84%	69%	91%	103%	187%	65%	48%	58%	40%	108%	89%	253%	52%	46%	50%	183%	133%	160%	278%	77%	
Jun-95	NA	57%	92%	112%	184%	117%	84%	98%	111%	88%	109%	112%	185%	138%	91%	102%	112%	118%	80%	86%	84%	67%	263%	222%	73%	251%	62%	96%	75%	185%	91%	172%	112%	90%		
Sep-95	6%	83%	112%	134%	106%	105%	92%	82%	120%	80%	87%	76%	95%	126%	49%	103%	147%	106%	142%	83%	50%	96%	45%	92%	115%	51%	54%	42%	668%	61%	140%	38%	20%			
Dec-95	127%	94%	90%	95%	90%	88%	93%	112%	78%	105%	83%	81%	115%	81%	131%	112%	133%	103%	110%	90%	70%	63%	94%	66%	62%	141%	58%	31%	54%	66%	63%	29%				
Mar-96	NA	101%	89%	78%	118%	80%	107%	70%	91%	76%	72%	92%	91%	109%	81%	109%	84%	155%	68%	189%	43%	56%	58%	60%	70%	74%	79%	119%	54%	462%	38%					
Jun-96	NA	77%	78%	94%	103%	91%	86%	103%	101%	79%	95%	140%	168%	101%	109%	91%	101%	89%	84%	69%	238%	66%	65%	137%	110%	222%	85%	76%	64%	169%						
Sep-96	78%	72%	107%	110%	108%	100%	96%	112%	120%	114%	104%	122%	114%	129%	112%	76%	81%	83%	104%	107%	202%	68%	103%	136%	130%	80%	173%	178%	232%							
Dec-96	NA	87%	120%	100%	83%	92%	117%	100%	82%	101%	100%	107%	75%	91%	126%	59%	93%	108%	191%	69%	80%	141%	258%	57%	50%	65%	208%	181%								
Mar-97	NA	91%	107%	100%	94%	81%	100%	119%	112%	86%	137%	91%	118%	80%	73%	91%	96%	131%	146%	103%	133%	57%	426%	110%	107%	38%	153%									
Jun-97	NA	122%	124%	96%	86%	77%	112%	86%	99%	101%	91%	81%	77%	75%	86%	113%	129%	115%	63%	81%	98%	76%	118%	34%	71%	51%										
Sep-97	2%	90%	92%	92%	98%	96%	93%	92%	91%	99%	110%	137%	88%	133%	118%	95%	75%	78%	57%	85%	103%	83%	420%	162%	116%											
Dec-97	94%	73%	112%	86%	89%	105%	84%	123%	113%	100%	87%	79%	123%	87%	86%	113%	68%	78%	96%	61%	131%	42%	76%	43%												
Mar-98	NA	96%	96%	104%	95%	92%	96%	94%	103%	88%	57%	100%	95%	81%	157%	91%	65%	78%	84%	137%	111%	65%	44%													
Jun-98	NA	112%	109%	103%	97%	98%	115%	114%	114%	77%	101%	91%	110%	127%	88%	136%	85%	73%	87%	52%	39%	67%														
Sep-98	14%	116%	123%	100%	111%	124%	112%	112%	115%	135%	108%	128%	89%	107%	173%	102%	128%	128%	180%	71%																
Dec-98	NA	114%	108%	182%	132%	108%	99%	93%	126%	79%	104%	184%	80%	94%	127%	82%	113%	188%	82%	90%																
Mar-99	8%	85%	109%	95%	86%	94%	70%	107%	84%	92%	87%	79%	72%	88%	79%	82%	68%	61%	80%																	
Jun-99	5%	95%	93%	96%	91%	99%	110%	111%	113%	120%	103%	82%	84%	80%	152%	104%	73%	113%																		
Sep-99	4%	90%	110%	97%	92%	111%	114%	103%	128%	99%	134%	93%	98%	104%	108%	130%	80%																			
Dec-99	12%	126%	104%	90%	111%	120%	109%	109%	97%	113%	103%	93%	97%	75%	100%	97%																				
Mar-00	112%	98%	86%	115%	86%	106%	93%	88%	93%	92%	84%	85%	143%	72%	80%																					
Jun-00	63%	92%	92%	82%	84%	91%	95%	95%	91%	85%	108%	104%	77%	90%																						
Sep-00	NA	138%	99%	81%	86%	114%	106%	126%	105%	94%	169%	98%	84%																							
Dec-00	8%	108%	103%	100%	84%	98%	91%	86%	86%	99%	101%	113%																								
Mar-01	29%	85%	108%	112%	85%	101%	104%	110%	98%	76%	80%																									
Jun-01	43%	98%	96%	165%	103%	107%	111%	110%	109%	85%																										
Sep-01	42%	108%	94%	107%	132%	114%	87%	98%	79%																											
Dec-01	45%	89%	93%	100%	118%	103%	114%	73%																												
Mar-02	54%	72%	107%	116%	98%	110%	81%																													
Jun-02	33%	102%	133%	113%	102%	110%																														
Sep-02	110%	88%	92%	107%	102%																															
Dec-02	251%	96%	110%	100%																																
Mar-03	37%	73%	55%																																	
Jun-03	70%	46%																																		
Sep-03	3%																																			

GLM

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Sep-94	NA	34%	59%	112%	155%	125%	76%	76%	84%	91%	121%	102%	91%	74%	61%	95%	72%	46%	107%	214%	102%	149%	70%	75%	58%	50%	398%	5%	162%	143%	67%	NA	91%	NA	3%	260%
Dec-94	38%	93%	81%	85%	118%	118%	117%	97%	95%	129%	95%	85%	86%	137%	93%	92%	107%	93%	117%	97%	63%	103%	58%	97%	108%	194%	45%	55%	46%	73%	89%	128%	247%	268%	70%	NA
Mar-95	28%	85%	102%	98%	94%	102%	86%	94%	100%	127%	156%	99%	83%	60%	64%	86%	69%	94%	160%	123%	74%	50%	73%	43%	104%	83%	268%	52%	43%	45%	178%	126%	135%	297%	71%	
Jun-95	NA	84%	105%	105%	108%	109%	89%	88%	108%	108%	138%	133%	89%	91%	101%	117%	87%	89%	87%	69%	248%	248%	71%	248%	60%	93%	71%	184%	86%	150%	102%	85%				
Sep-95	18%	123%	114%	122%	102%	110%	91%	80%	118%	84%	88%	76%	96%	132%	47%	100%	155%	120%	154%	86%	54%	106%	47%	93%	116%	54%	55%	41%	645%	59%	137%	37%	19%			
Dec-95	336%	112%	87%	98%	94%	85%	88%	108%	80%	107%	84%	82%	119%	77%	120%	110%	138%	102%	106%	92%	73%	64%	91%	64%	64%	134%	55%	31%	54%	61%	58%	26%				
Mar-96	NA	132%	107%	89%	89%	816%	78%	104%	71%	95%	79%	74%	96%	88%	103%	78%	109%	83%	134%	72%	198%	43%	57%	58%	64%	69%	71%	80%	128%	52%	457%	36%				
Jun-96	NA	123%	100%	94%	99%	89%	85%	108%	104%	79%	97%	133%	98%	96%	99%	137%	109%	87%	89%	66%	267%	64%	68%	194%	110%	122%	89%	75%	63%	101%						
Sep-96	251%	105%	106%	106%	106%	106%	99%	99%	116%	122%	116%	101%	114%	109%	124%	102%	72%	82%	89%	104%	199%	208%	72%	103%	139%	136%	83%	169%	178%	233%						
Dec-96	NA	90%	115%	102%	83%	96%	124%	103%	85%	102%	99%	108%	76%	86%	121%	63%	102%	115%	191%	74%	89%	148%	275%	63%	55%	68%	219%	190%								
Mar-97	NA	94%	100%	97%	99%	85%	100%	120%	113%	86%	139%	91%	111%	72%	70%	91%	91%	125%	145%	108%	135%	60%	462%	125%	110%	38%	167%									
Jun-97	NA	135%	116%	103%	93%	79%	112%	89%	104%	107%	94%	79%	95%	120%	128%	130%	73%	87%	110%	87%	141%	37%	80%	58%												
Sep-97	3%	89%	89%	97%	101%	94%	94%	95%	94%	101%	102%	132%	89%	155%	189%	94%	76%	85%	58%	92%	119%	96%	431%	112%	126%											
Dec-97	134%	86%	93%	106%	91%	88%	106%	89%	135%	120%	160%	86%	81%	130%	83%	81%	116%	75%	84%	108%	71%	100%	46%	88%	49%											
Mar-98	NA	86%	93%	106%	90%	99%	99%	92%	103%	89%	60%	103%	92%	77%	156%	96%	68%	84%	93%	161%	119%	72%	50%													
Jun-98	NA	101%	99%	110%	104%	99%	107%	110%	111%	77%	96%	81%	99%	117%	81%	118%	82%	77%	97%	55%	42%	76%														
Sep-98	149%	98%	105%	103%	113%	115%	108%	109%	110%	108%	119%	96%	117%	80%	90%	148%	93%	125%	119%	179%	73%															
Dec-98	NA	96%	89%	105%	107%	103%	101%	99%	128%	79%	105%	107%	83%	93%	132%	93%	139%	119%	94%	107%																
Mar-99	9%	77%	106%	95%	87%	95%	74%	105%	84%	94%	93%	82%	71%	93%	94%	101%	78%	74%	126%																	
Jun-99	7%	83%	92%	107%	95%	97%	100%	107%	112%	124%	102%	78%	84%	83%	156%	104%	79%	126%																		
Sep-99	5%	81%	114%	108%	93%	99%	106%	99%	123%	95%	128%	94%	105%	118%	108%	142%	94%																			
Dec-99	15%	126%	117%	100%	104%	90%	104%	110%	92%	108%	104%	100%	100%	71%	101%	105%																				
Mar-00	155%	116%	92%	124%	95%	107%	94%	82%	88%	93%	102%	88%	138%	75%	88%																					
Jun-00	142%	98%	99%	107%	94%	96%	88%	93%	94%	97%	115%	104%	83%	108%																						
Sep-00	NA	157%	125%	101%	103%	110%	104%	120%	115%	96%	164%	103%																								
Dec-00	9%	98%	77%	85%	71%	91%	88%	91%	85%	97%	106%	124%																								
Mar-01	43%	93%	84%	96%	83%	92%	109%	106%	98%	81%	90%																									
Jun-01	181%	104%	80%	82%	95%	95%	106%	111%	114%	97%																										
Sep-01	182%	100%	98%	106%	130%	107%	89%	100%	90%																											
Dec-01	71%	121%	87%	112%	106%	103%	113%	82%																												
Mar-02	126%	96%	107%	110%	102%	102%	90%																													
Jun-02	177%	126%	110%	129%	98%	112%																														
Sep-02	279%	75%	95%	107%	108%																															
Dec-02	272%	104%																																		
Mar-03	50%	66%	45%																																	
Jun-03	95%	40%																																		
Sep-03	3%																																			

NN



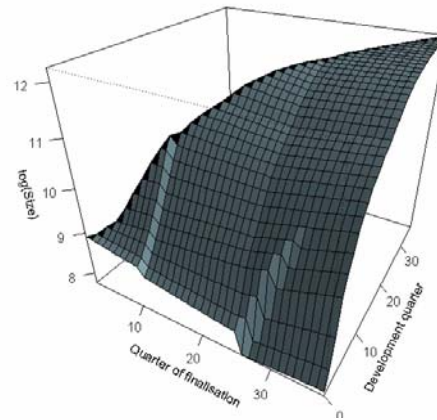
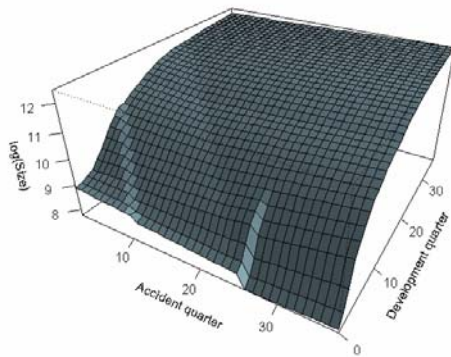
# MART

# MARS

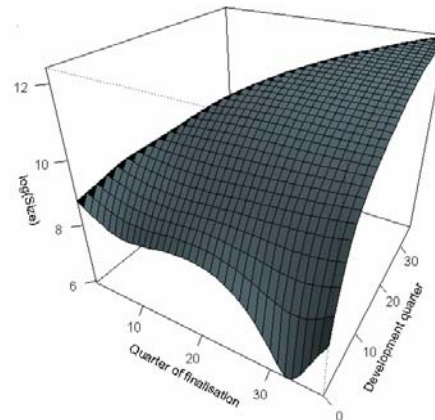
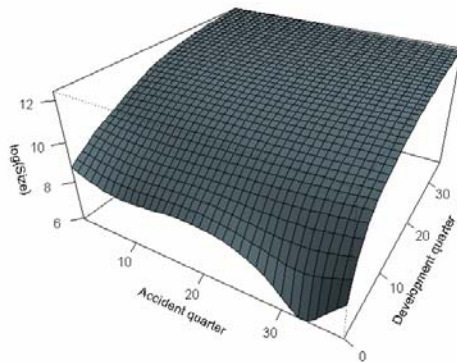


# “3D Triangles”

- Log(average size of finalised claims)



GLM

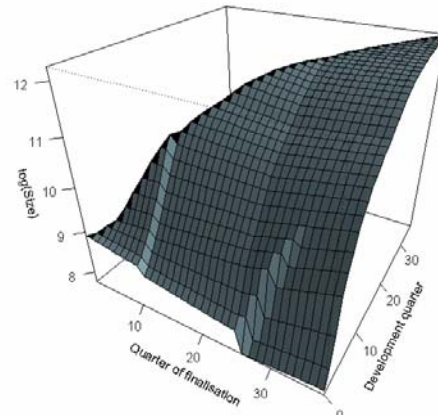
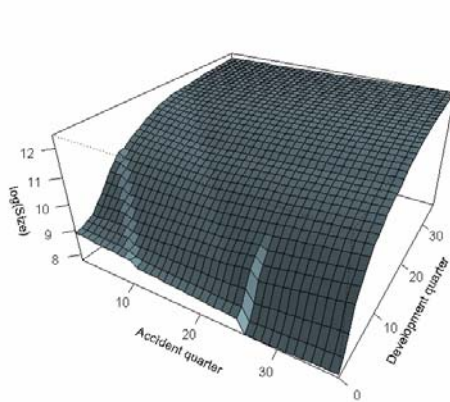


NN

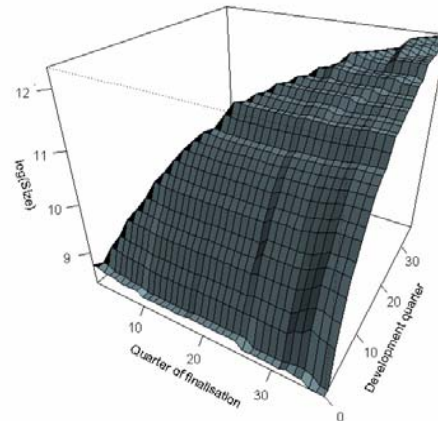
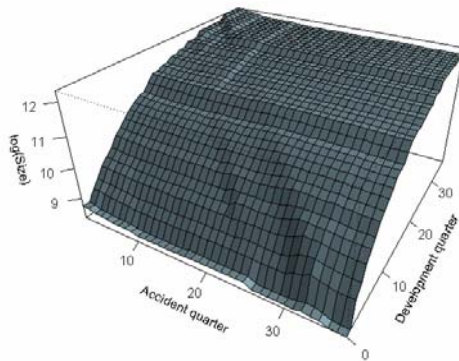


# “3D Triangles”

- Log(average size of finalised claims)



**GLM**



**MART**



## Predictive accuracy

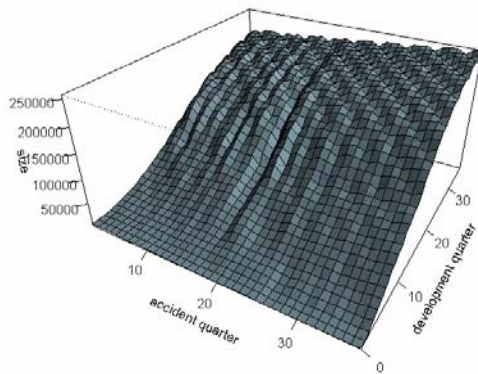
- Each model fitted to “training set” consisting of 2/3 of total data
- Remaining 1/3 of data used as “test set” to test the predictive accuracy

Model	Sum of squares	Average Absolute Error
GLM	$2.000 \times 10^{14}$	33,777
Neural Network	$1.996 \times 10^{14}$	33,476
MART	$1.999 \times 10^{14}$	33,290
MARS	$1.994 \times 10^{14}$	33,806

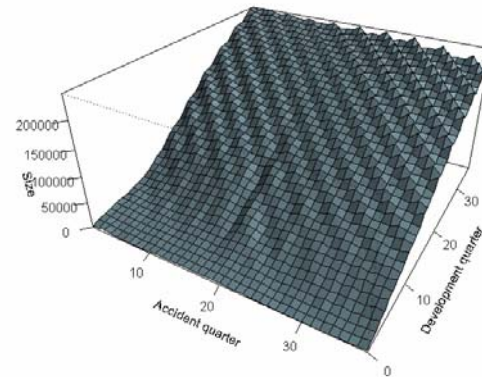




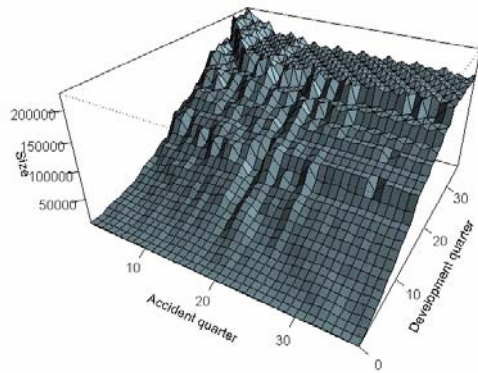
# Projections of claim size



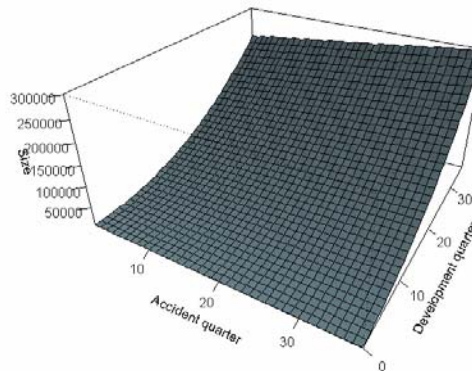
**GLM**



**NN**



**MART**



**MARS**



## Summary

- Neural Networks and MART had better predictive accuracy than GLMs
- The soft-computing methods were largely automated and quicker to use
- Soft-computing models were less suitable for projecting claim sizes into future periods





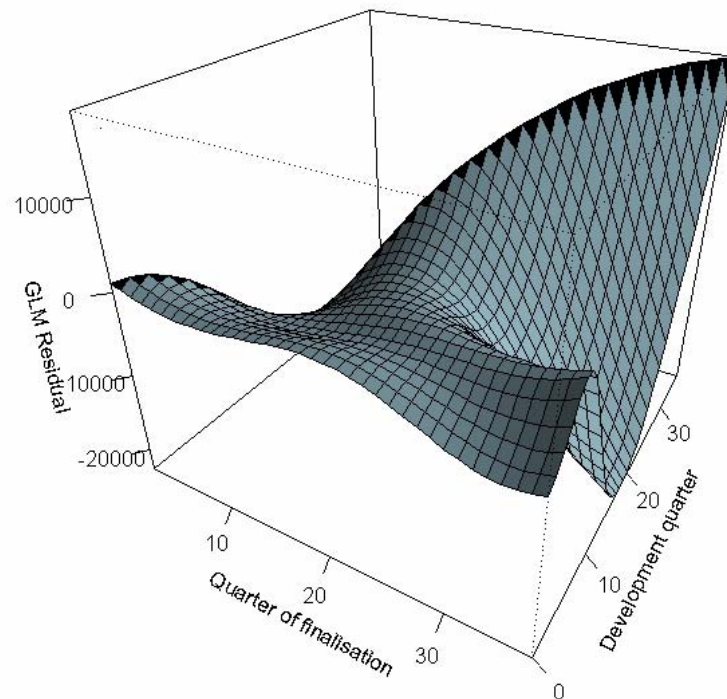
# Summary

- GLMs better for performing pricing and reserving projections
- Soft-computing techniques could play role in model checking
  - Assessing the prediction error of the GLM model
  - Help determine which interactions to include in the GLM model.



# Summary

- Features remaining in the residuals from a “main effects” GLM model of claim size





**Thank you**