

Cluster Modeling

Liabilities, Assets, Scenari



Presentato da:
Nazario Cannarozzi



Compressione dei dati – Tendenza nel tempo



Il Cluster Modeling lo fa meglio

- **La domanda quindi non è: comprimere o non comprimere?**

- **La domanda è: dal momento che è necessario comprimere, come farlo al meglio?**

Vivere in un mondo compresso

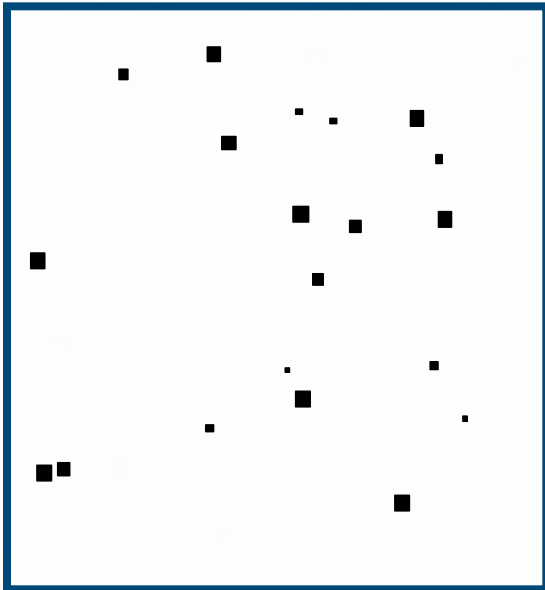
- Tecniche di compressione classiche
 - Alcune hanno delle regole (compressione in base all'età o alla data di decorrenza)
 - Altre sono soggettive (tariffe meno importanti con tariffe più importanti)
 - Sono focalizzate sui valori iniziali: si ipotizza che riprodurre i valori iniziali implichi una buona riproduzione dei flussi futuri
- Criticità
 - Rimanere aggiornati con le nuove tariffe
 - Gestire e misurare il “rumore” della compressione
 - Assecondare gli auditors

Schema bi-dimensionale del Cluster Modeling

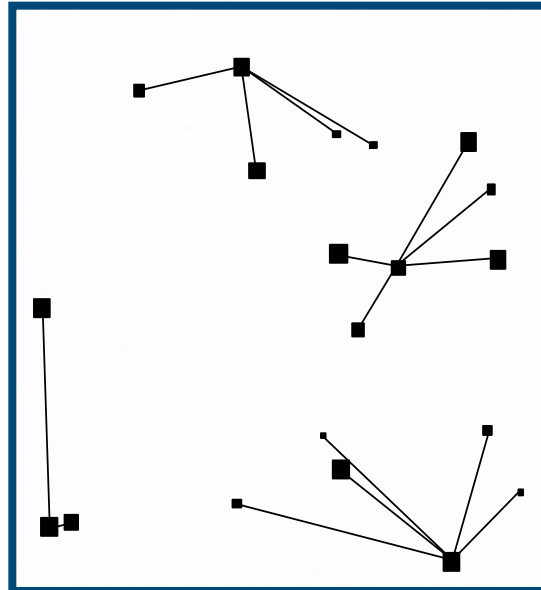
(Esempio sui passivi: Riserva iniziale e premio proiettato)

(Esempio sugli assets: Valore di carico/alla pari e rendimento a scadenza)

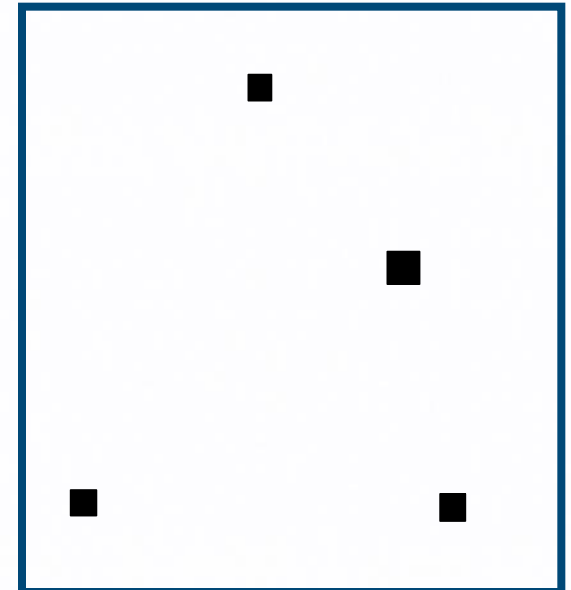
Scatter-Plot bidimensionale di polizze di diversi pesi



Assegnazione delle polizze ai cluster



Estrapolazione dei Punti Centrali



Il Cluster Modeling – pro e contro

- Pro:
 - Applicabile a qualsiasi prodotto o tipo di asset
 - Compressione migliore a parità di adattamento modello-reale
 - Semplice da automatizzare con poco impiego di risorse
 - Manutenzione semplice per valutazioni successive
 - Personalizzabile
 - Semplice calibrazione del numero di model points
 - Semplici analisi al volo dell'adattamento del modello
- Contro:
 - Idee...?

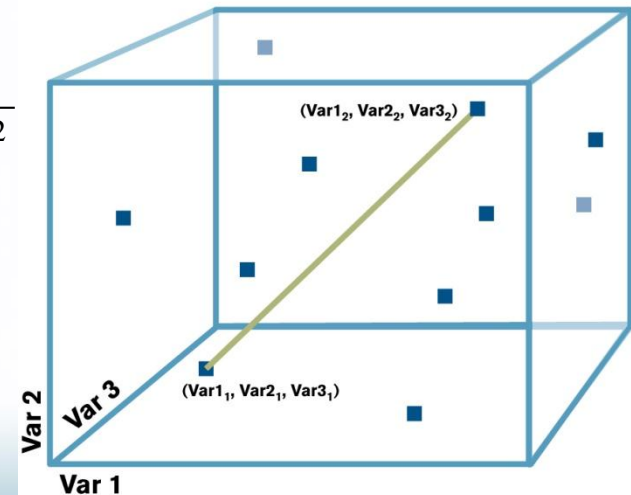
Cluster Modeling: concetti chiave

- ***Variabile di posizione:*** qualunque valore che si voglia riprodurre con cura, come:
 - Riserve iniziali o a date successive
 - Premi annualizzati, o a varie date
 - Sinistri in diversi momenti della proiezione
 - Flussi netti del passivo in ognuno dei primi cinque anni
 - Valore attuale dei profitti
- Gli utenti definiscono la lista di variabili e ne catturano i valori in un report

Cluster Modeling: concetti chiave

- **Funzione distanza:** Una misura sintetica di “quanto lontani” siano due oggetti (polizze, titoli, scenari) in uno spazio n-dimensionale
- Distanza euclidea applicata alle variabili di posizione normalizzate
- Possibilità di assegnare pesi per riflettere l’importanza di ogni singola variabile

$$\sqrt{(Var1_1 - Var1_2)^2 + (Var2_1 - Var2_2)^2 + (Var3_1 - Var3_2)^2}$$



Cluster Modeling: concetti chiave

- **Peso:** un elemento rappresentativo dell'importanza di ogni polizza/titolo, definibile dall'utente
 - Un esempio tipico è il capitale corrente, o la riserva
- **Importanza** = (Peso) * (Distanza dalla polizza più vicina)

Cluster Modeling: concetti chiave

- **Segmento:** Un gruppo di cui ogni polizza fa parte, in modo che nessuna polizza sia mappata fuori dal proprio segmento
- LOB o asset class saranno sempre un segmento
- Altri esempi: codice tariffa, durata, minimo garantito
- L'uso della segmentazione abbrevia i tempi di compressione e può migliorare la mappatura dei risultati su altri scenari
- Può ovviamente portare limitazioni

Cluster Modeling: l'algoritmo sottostante

- Calcolo della distanza di ogni polizza da ogni altra nel proprio segmento
- Calcolo dell'*importanza* di ogni polizza come $(\text{Peso}) * (\text{Distanza dalla polizza più vicina})$ per ogni polizza
- Identificazione della polizza meno importante e mappatura nella più vicina dello stesso segmento
- Iterazione fino all'ottenimento del numero di celle desiderato
- Per ogni cella risultante, selezione del punto più vicino alla posizione media di tutte le celle originarie. Utilizzo di questo punto per rappresentare la cella risultante
- Estrapolazione di tutti i dati di input associati alla cella di destinazione

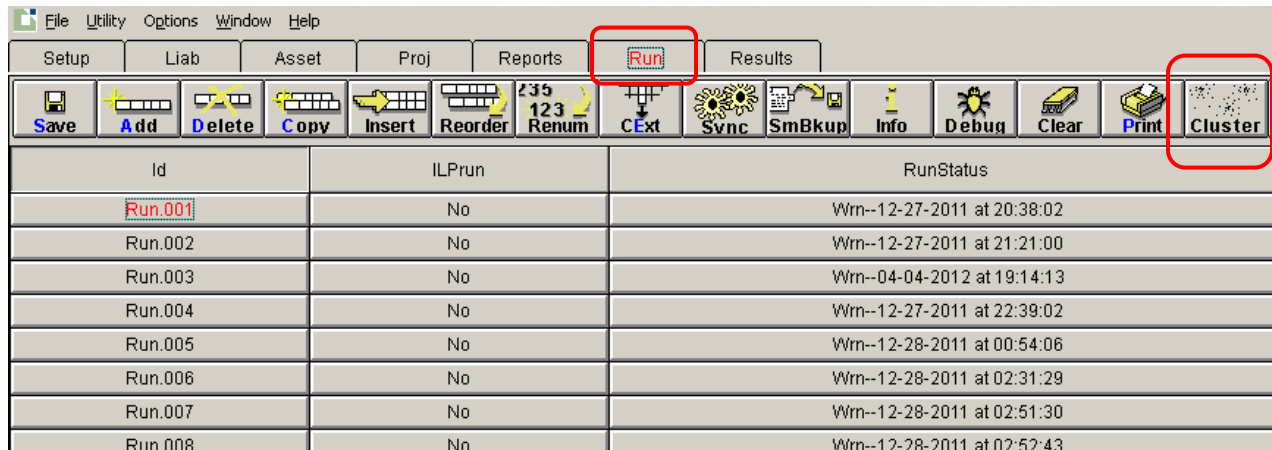
Opzioni di Cluster per un modello stocastico di grandi dimensioni

- Comprimere i passivi
- Comprimere gli attivi
- Comprimere gli scenari
- Tutte le precedenti:
 - Miglioramento della velocità secondo più ordini di grandezza

Procedura da attuare

- Definire le variabili di posizione, gli scenari di calibrazione e gli inventory reports in funzione delle finalità del modello
- Girare il modello
- Identificare il numero di celle target e assegnare i pesi alle variabili di calibrazione
- Eseguire la compressione
- Validare e verificare l'adattamento
- Ripetere raffinando se necessario (cambiare il numero di celle, i pesi, la segmentazione...)

Cluster Modeling – in pratica



The screenshot shows a software application window with a menu bar (File, Utility, Options, Window, Help) and a toolbar. The 'Run' button in the toolbar is highlighted with a red box. Below the toolbar is a table with three columns: 'Id', 'ILPrun', and 'RunStatus'. The first row of the table has 'Run.001' in the 'Id' column, which is also highlighted with a red box. The 'Cluster' button in the toolbar is also highlighted with a red box.

Id	ILPrun	RunStatus
Run.001	No	Wrm--12-27-2011 at 20:38:02
Run.002	No	Wrm--12-27-2011 at 21:21:00
Run.003	No	Wrm--04-04-2012 at 19:14:13
Run.004	No	Wrm--12-27-2011 at 22:39:02
Run.005	No	Wrm--12-28-2011 at 00:54:06
Run.006	No	Wrm--12-28-2011 at 02:31:29
Run.007	No	Wrm--12-28-2011 at 02:51:30
Run.008	No	Wrm--12-28-2011 at 02:52:43

Verifica dell'adattamento

- Un esempio di report di verifica:

Model Fit Analysis--Compress 0.3% --From 113,821 Cells to 341 Cells

Variable	Description	Actual	Model	Difference	Ratio %
CellCount	In-force file cell count	113,821	341	-113,480	0.3
FixedLives	Units in-force	113,821	112,604	-1,217	98.9
InitPolCount	Initial policy count	113,821	112,604	-1,217	98.9
InitAV	Account value in-force	1,803,127,157	1,803,127,157	0	100.0
InitPremInforce	Initial premium in-force	111,582,699	109,284,691	-2,298,008	97.9
CollPinf	Collected premium in-force	111,582,699	109,284,691	-2,298,008	97.9
TargPinf	Target premium in-force	111,582,699	109,284,691	-2,298,008	97.9
PourPinf	Pour-in premium in-force	111,582,699	109,284,691	-2,298,008	97.9
InitStatReserve	Initial statutory reserve	1,803,127,157	1,803,127,157	0	100.0
InitTaxReserve	Initial tax reserve	1,803,127,157	1,803,127,157	0	100.0
InitDB	Insurance in-force	1,803,127,157	1,803,127,157	0	100.0
InitCV	Cash value in-force	1,803,127,157	1,803,127,157	0	100.0
InitPolLoans	Initial policy loans	0	0	0	100.0
CashPrem(1)	Cash premiums	6,333,667	5,742,564	-591,102	90.7
CashPrem(12)	Cash premiums	5,098,574	4,669,526	-429,048	91.6
CashPrem(60)	Cash premiums	2,484,392	2,297,286	-187,107	92.5
CashPrem(120)	Cash premiums	900,516	848,793	-51,724	94.3
CashPrem(239)	Cash premiums	176,369	155,338	-21,030	88.1
Comm(1)	Total commissions	74,988	75,275	287	100.4
Comm(12)	Total commissions	17,615	17,496	-118	99.3
Comm(60)	Total commissions	19,209	19,327	117	100.6
Comm(120)	Total commissions	10,670	10,897	228	102.1
Comm(239)	Total commissions	160	101	-59	63.3
AcqExp(1)	Acquisition expenses	0	0	0	100.0
AcqExp(12)	Acquisition expenses	0	0	0	100.0
AcqExp(60)	Acquisition expenses	0	0	0	100.0

Variable 1 of 190

Tempistiche (dati da/per MG-ALFA®)

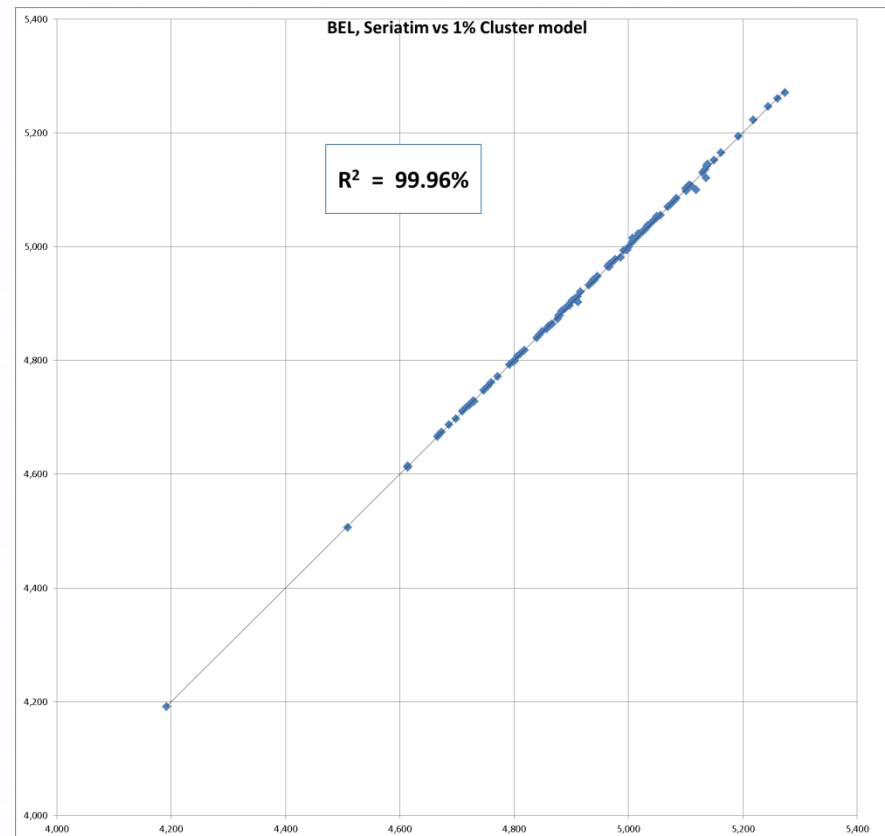
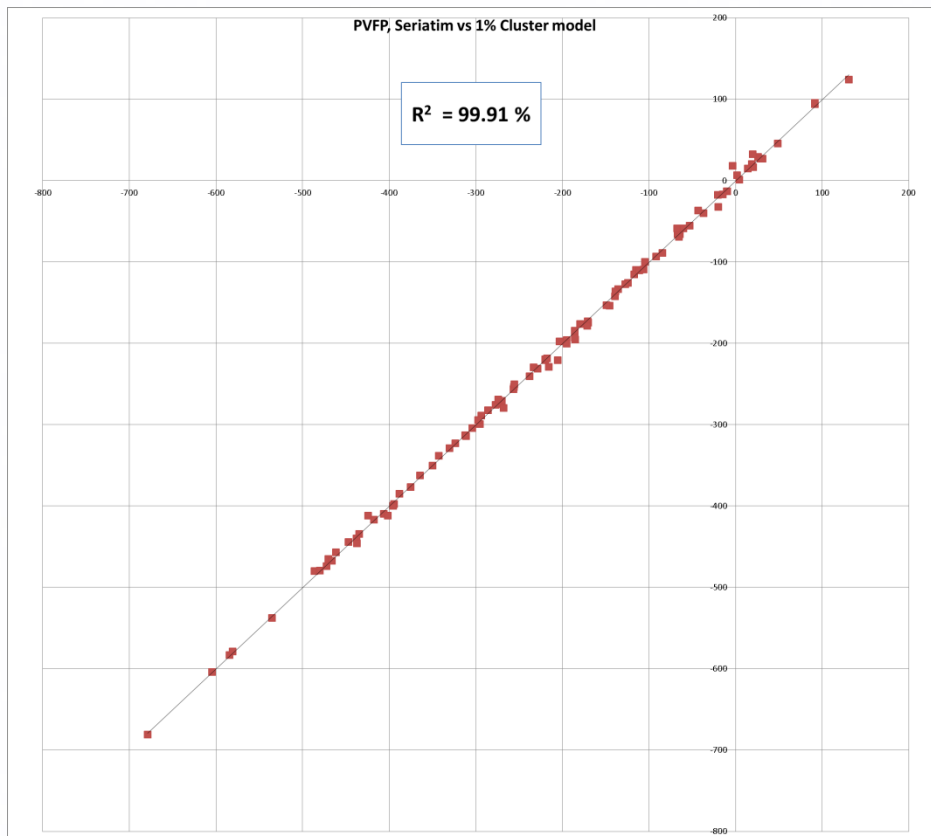
- I tempi di run dell'algoritmo dipendono soprattutto da:
 - Numero di model points di partenza
 - Numero di variabili di posizione scelte
 - Indicativamente gira in minuti, non in ore
- Re-run completo dell'algoritmo per:
 - Aggiunta/eliminazione di variabili
 - Variazione dei loro pesi
 - Variazione della segmentazione
- Re-run parziale (pochi secondi) per:
 - Variazione del numero di celle target
 - Variazione delle specifiche delle variabili di in-force (media piuttosto che somma, riproporzionamento, ...)

Comprimere dati per sistemi diversi da MG-ALFA®

- Definire le variabili di posizione, gli scenari di calibrazione e gli inventory reports in funzione delle finalità del modello
- Girare il modello (meglio se seriatim) creando un output in formato testo (con valori delle variabili di posizione nei differenti scenari)
- Impostare la compressione in MG-ALFA® con i file in-force e i risultati da essi generati
- Procedere con gli step descritti precedentemente
- Validare i file in-force prodotti e esportarli nel sistema utilizzato

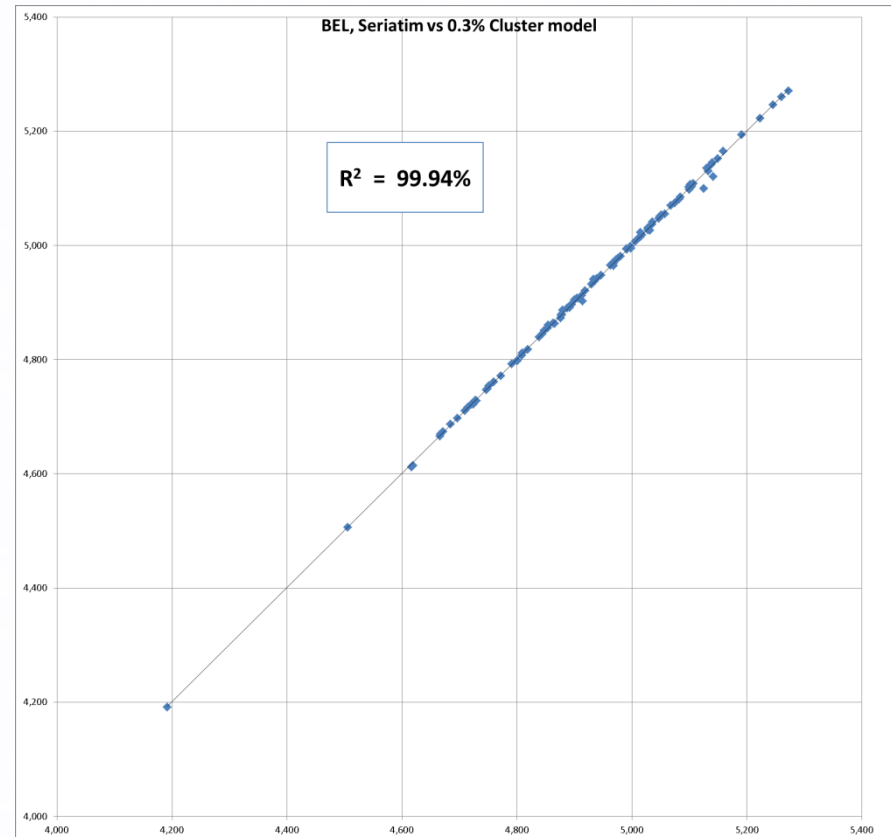
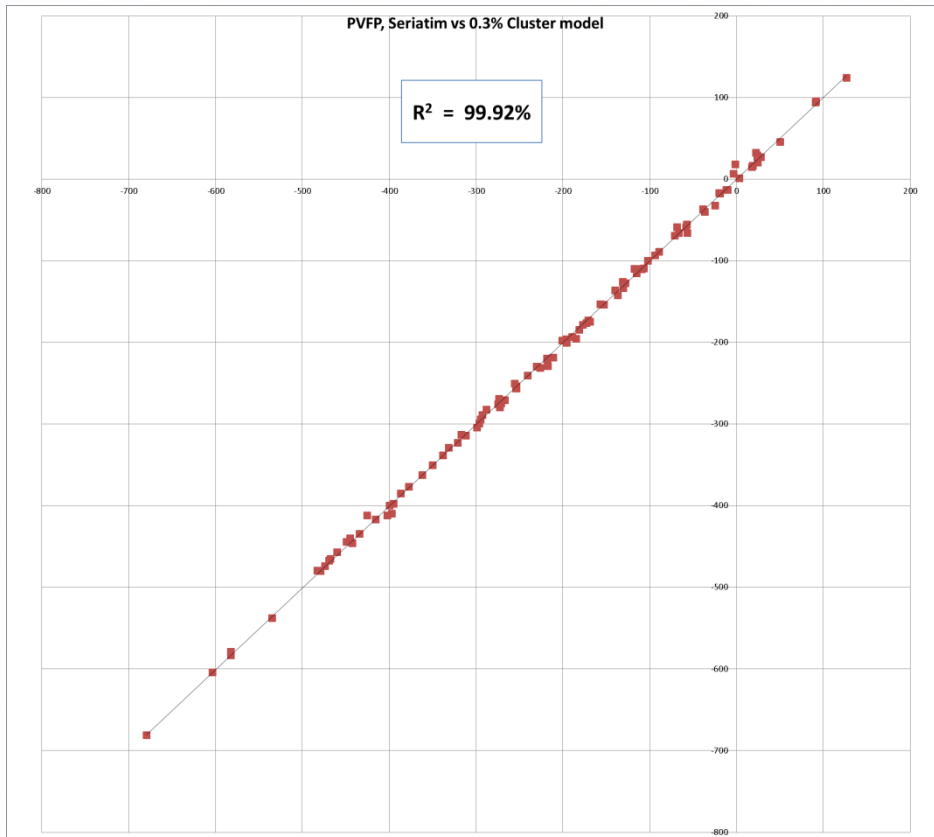
Alcuni esempi su una gestione separata

- Run su 100 scenari, livello di compressione 1%:



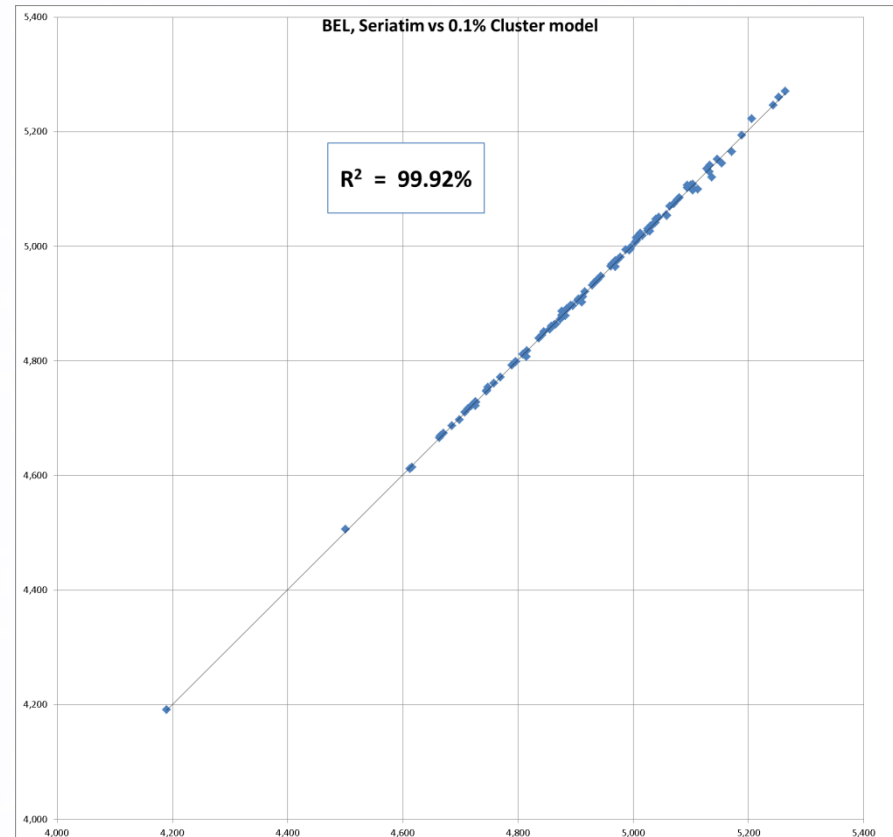
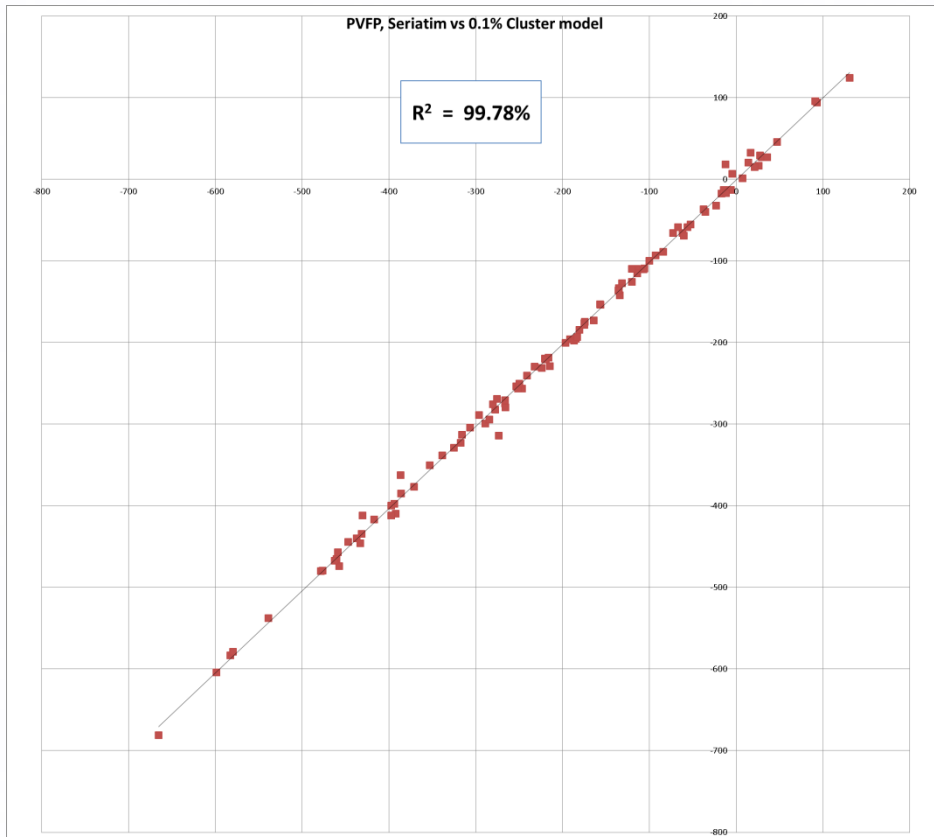
Alcuni esempi su una gestione separata

- Run su 100 scenari, livello di compressione 0.3%:



Alcuni esempi su una gestione separata

- Run su 100 scenari, livello di compressione 0.1%:



Nazario Cannarozzi

Consulting Actuary

Milliman S.r.l.

Corso Europa, 5

20122 Milano

T: +39 02. 76. 260. 524

Cell: +39 334 67 86 720

nazario.cannarozzi@milliman.com

This presentation has been prepared for illustrative purposes only. It should not be further distributed, disclosed, copied or otherwise furnished to any other party without Milliman's prior consent.

No reliance should be placed on the results or graphs presented herein and no inference made about the appropriateness of the information and different bases presented. In particular, independent verification and professional advice should be sought when establishing company bases and assumptions for the purposes of pricing, valuation and transaction purposes etc.

Actual experience may be more or less favourable than the assumptions and illustrations presented in this presentation. To the extent actual experience differs from these, so will actual results differ from those presented.